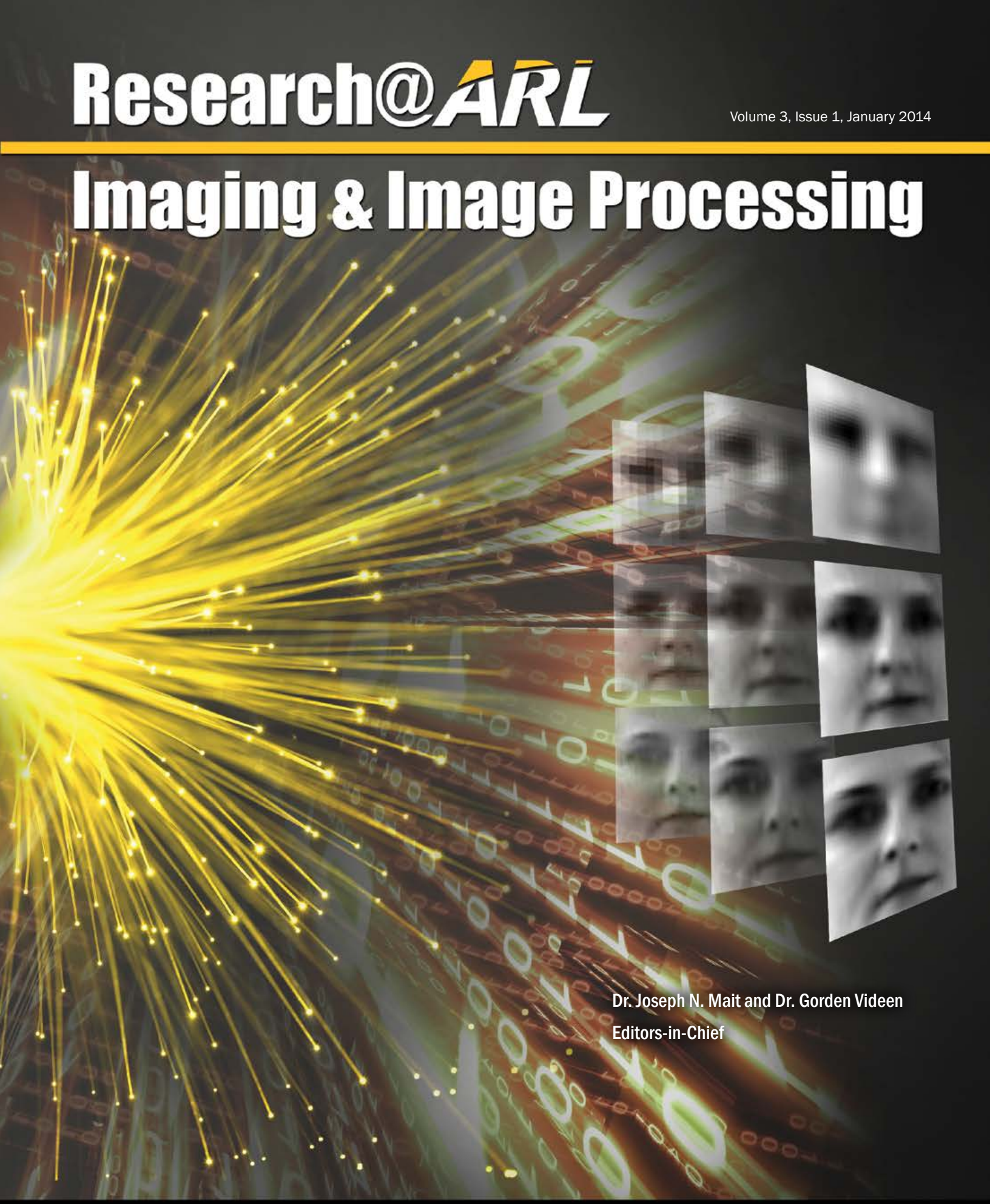
Research@ARL

Volume 3, Issue 1, January 2014

Imaging & Image Processing



Dr. Joseph N. Mait and Dr. Gorden Videen

Editors-in-Chief

U.S. Army Research Laboratory

Editorial Board

Dr. Joseph N. Mait
Editor-in-Chief

Dr. Gorden Videen
Editor-in-Chief

Dr. Nasser M. Nasrabadi

Dr. Kwong K. Choi

Advisory Board

Research@ARL can be accessed electronically at **www.arl.army.mil/ResearchARL**

Research@ARL

Imaging & Image Processing • 2014

contents

Introduction to ARL

The Army Research Laboratory of the U.S. Army Research, Development and Engineering Command (RDECOM) is the Army's corporate laboratory. ARL's research continuum focuses on basic and applied research (6.1 and 6.2) and survivability/lethality and human factors analysis (6.6). ARL also applies the extensive research and analysis tools developed in its direct mission program to support ongoing development and acquisition programs in the Army Research, Development and Engineering Centers (RDECs), Program Executive Offices (PEOs)/Program Manager (PM) Offices, and Industry. ARL has consistently provided the enabling technologies in many of the Army's most important weapons systems.

The Soldiers of today and tomorrow depend on us to deliver the scientific discoveries, technological advances, and the analyses that provide Warfighters with the capabilities to execute full-spectrum operations. ARL has Collaborative Technology Alliances in Micro Autonomous Systems Technology, Robotics, Cognition and Neuroergonomics, Network Science, an International Technology Alliance and new Collaborative Research Alliances in Multiscale Multidisciplinary Modeling of Electronic Materials and Materials in Extreme Environments. ARL's diverse assortment of unique facilities and dedicated workforce of government and private sector partners make up the largest source of world class integrated research and analysis in the Army.

ARL Mission

The mission of ARL is to provide innovative science, technology, and analyses to enable full-spectrum operations.

Our Vision

America's Laboratory for the Army: Many Minds, Many Capabilities, Single Focus on the Soldier.

ARL's Organization

- Army Research Office (ARO) - Initiates the scientific and far reaching technological discoveries in extramural organizations: educational institutions, nonprofit organizations, and private industry.
- Computational and Information Sciences Directorate - Scientific research and technology focused on information processing, network and communication sciences, information assurance, battlespace environments, and advanced computing that create, exploit, and harvest innovative technologies to enable knowledge superiority for the Warfighter.
- Human Research and Engineering Directorate - Scientific research and technology directed toward optimizing Soldier performance and Soldier-machine interactions to maximize battlefield effectiveness and to ensure that Soldier performance requirements are adequately considered in technology development and system design.
- Sensors and Electron Devices Directorate - Scientific research and technology in electro-optic smart sensors, multifunction radio frequency (RF), autonomous sensing, power and energy, and signature management for reconnaissance, intelligence, surveillance, target acquisition (RISTA), fire control, guidance, fuzing, survivability, mobility, and lethality.
- Survivability/Lethality Analysis Directorate - Integrated survivability and lethality analysis of Army systems and technologies across the full spectrum of battlefield threats and environments as well as analysis tools, techniques, and methodologies.
- Vehicle Technology Directorate - Scientific research and technology addressing propulsion, transmission, aeromechanics, structural engineering, and robotics technologies for both air and ground vehicles.
- Weapons and Materials Research Directorate - Scientific research and technology in the areas of weapons, protection, and materials to enhance the lethality and survivability of the Nation's ground forces.

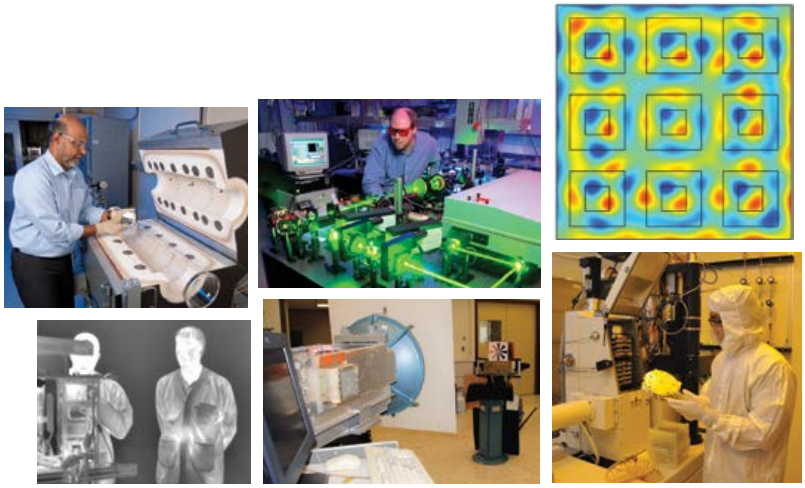
ARL Workforce in 2013

- 1,980 Civilians - 38 Military
- 1080 Contractors (1027 full-time/53 part-time)
- 1,379 Research Performing Workforce
- 552 (40%) hold PhDs
- 11 STs / 23 ARL Fellows

ARL's Primary Sites

- Aberdeen Proving Ground, MD
- Adelphi Laboratory Center, MD
- White Sands Missile Range, NM
- Raleigh-Durham, NC
- Orlando, FL

Visit ARL's web site at **www.arl.army.mil**



Unique ARL laboratory facilities and modeling capabilities provide our scientists and engineers with a world-class research environment.

ARL
U.S. ARMY RESEARCH LABORATORY

FOREWORD

Thank you for your interest in this latest edition of Research@ARL. This compendium of previously published peer-reviewed journal articles represents the best of ARL research efforts in the topic area covered. Our researchers take extreme pride in the quality of their research that will influence the way the U.S. Army operates 10, 20 and even 30 years from now. Their dedication to our mission and the mission of the U.S. Army is focused on providing enhanced capabilities for our Soldiers of the future. In this edition of Research@ARL, we take a look at the science and technology of imaging and image processing.

For over 400 years, optical instruments have been essential to military operations. From the first telescopes and binoculars, imaging has expanded to put remote cameras on satellites and on autonomous platforms. Significantly, the U.S. Army Research Laboratory's predecessor organizations developed technologies in the 1950s and 1960s that allow us to see in the dark, giving the U.S. military a strategic warfighting advantage. In its 2012 Optics and Photonics: Essential Technologies for Our Nation, the National Academies highlighted the importance of imaging and optical technologies to the military.

This volume of Research@ARL highlights recent contributions to imaging science being made by ARL researchers. These include advanced concepts in optical design, infrared technology, and image processing. Further, using the power of signal processing, ARL researchers are altering the concept of imaging itself to develop capabilities not previously possible. For example, by exploiting the quantum nature of light, ARL researchers are developing imagers that will allow cameras to see around corners.

I hope you will enjoy perusing the articles contained in this volume, and after doing so, I'm sure you'll appreciate the advances our scientists and engineers are making to ensure our Army remains technologically superior.

Dr. Thomas P. Russell
Director, U.S. Army Research Laboratory

Imaging and Image Processing Research at ARL

J. N. Mait, G. Videen, N. M. Nasrabadi, and K.-K. Choi

1. Introduction

Today cameras are ubiquitous. In addition to capturing special moments with family and friends, they monitor traffic and they monitor us in public places. They provide a full visual field as we back up our cars. We can even swallow a pill-sized camera to image our intestinal tract. Cameras are so widely used that 48 hours of video data are uploaded to YouTube every minute.[1] Through advancements in optics and photodetectors, cameras are now commodities. Further, the proximity of cameras to processing chips in smartphone platforms is driving an explosion in imaging applications.

The U.S. military has been a significant but unheralded contributor to this revolution through its development of lightweight, small-scale optics, high-pixel-count focal-plane arrays, and algorithms for pattern recognition. Further, nearly 50 years ago, predecessor organizations to the U.S. Army Research Laboratory (ARL) developed image intensifiers and infrared imaging systems to "own the night." First deployed in small numbers to soldiers in Vietnam, the prevalence of these technologies in military units during Desert Storm gave the U.S. a strategic advantage.

Today, ARL researchers are building on this heritage to alter the concept of imaging itself. For many years, one created a camera by combining optics to form an image with a detector to capture it; more recently, one uses post-detection processing to enhance it. When viewed as a whole, however, it is possible to spread the process of image formation across all three elements—optics, detectors, and signal processing. Doing so has allowed ARL researchers to develop imaging capabilities that are not possible under the old paradigm.

This volume presents some of those capabilities, as well as other contributions made by ARL researchers to imaging and image processing, and this introduction provides context for ARL's research investment in these areas.

2. Army Applications

Military applications of imaging are varied. They include, for example, intelligence gathering (collecting specific information to support a query), surveillance (watching a particular region for activity), reconnaissance (gathering task-specific military information), targeting (labeling an object unequivocally as an object for destruction), and battle-damage assessment (BDA, assessing the status of a target after an engagement). Broadly, they serve national policy, strategic policy, and tactical missions. At the national level, imaging assets are used to assess adherence to international treaties. Critical strategic applications include intelligence, reconnaissance, and surveillance (ISR), and critical tactical applications include targeting and BDA. Further, as imaging assets have become available to field commanders, situational awareness (a tactical understanding of one's surroundings) also has become a critical application.

The applications of imaging to military engagements were evident at the beginning of imaging science. Although the microscope, invented in 1590, is recognized as the first optical instrument, speculation exists that the British defeat of the Spanish Armada in 1588 was enabled in part by the secret invention of the telescope.[2] If true, the lesson of the Spanish defeat is that seeing farther than an adversary gives one the advantage of time. If untrue, the story is at least apocryphal. Two months after submitting his patent application for the telescope in 1608, the recognized inventor, Hans Lippershey, submitted another for binoculars. To this day, binoculars remain a mainstay of tactical military units.

The invention of film in 1837 removed the need for a human observer and led to the development of the camera. Consequently, the first aerial photograph was taken in 1858 by a photographer in a hot air balloon, and a camera launched on a kite in 1882 initiated the field of remote imaging. The practice of loading cameras with film canisters, placing them on aerial and high-altitude platforms, and retrieving the canisters after exposure was used to great effect in World War I and continued into the Cold War with U-2 reconnaissance aircraft and Corona, the first imaging satellite. In 1962, these platforms provided the imagery that indicated the Soviets were constructing missile launchers in Cuba.

The invention of solid-state detection by Smith and Boyle in 1969 removed the need to retrieve film canisters and, in 1976, the National Reconnaissance Office launched the KH-11, the first satellite equipped with electronic imaging. Such capabilities were used in the 1980s to "trust but verify" arms treaties with the Soviets. (The first personal digital camera was invented in 1975 by Steve Sasson of Kodak.)

In 1991, satellite imagery was used to create target lists for Desert Storm. However, targeting mobile SCUD missile launchers revealed the need to shorten the latency between target detection and engagement. Delays encountered in collecting and analyzing satellite imagery and distributing the results back to field commanders drove the need for persistent surveillance.

In a reflection of the past, cameras attached to balloons hovering over cities have returned. The aerostat-borne Persistent Threat Detection System and its aircraft-borne brethren, like Constant Hawk, have reduced collection and distribution times. But analysis remains a critical problem. These systems generate massive amounts of digital image data that require considerable time to process and interpret.

The scale reduction in imaging systems has been a boon for tactical units. Small unmanned aerial vehicles are being developed to provide overhead imaging to company commanders, and camera-equipped robots are providing squads with situational awareness of buildings, caves, and other unexplored terrain before a Soldier enters. No Soldier should ever again have to experience the fear of a "tunnel rat" in Viet Nam, armed only with a flashlight and a 45.

The imaging applications discussed up to this point require a camera that merely replicates what humans could see if their eyes were in the position of the camera. However, a camera does not function like a human eye, which is both a disadvantage and an advantage. For example, the dynamic range of the human eye is far greater than that of a camera. Humans can see in bright sun and in moonless nights; whereas, most cameras do not image well, if at all, in dim light. If they do, their images are saturated if just a little light is present. On the other hand, our eyes detect radiation only in the visible portion of the electromagnetic spectrum. Yet we can build cameras that detect radiation in different wavebands, such as infrared. Infrared cameras are useful for seeing warm objects at night, such as humans and car engines.

In the next section, we highlight contributions made by ARL researchers to address the technical challenges presented by military applications of imaging.

3. ARL Research

3.1. Optics and Optical Design

The most fundamental principle underlying all imaging systems is that the angular size of the smallest object the system can resolve is proportional to the wavelength of the illumination and inversely proportional to the diameter of the input aperture. To increase resolution, one increases the diameter of the optical system. However, this also increases system volume. A system capable of resolving objects half the size of another system nominally requires eight times the volume. Further, the field-of-view, or how much of a scene an observer can see, is a function of the extent of the image plane. An imager with a detector that is twice as large as another imager will have twice the field-of-view, but it will also have four times the volume. Thus, designers must balance optical performance and system size.

For persistent surveillance, one would like to maintain resolution across a large field-of-view. However, if one does not change the optics as the field-of-view increases, the quality of the image at its edges degrades. Light arriving at a detector edge travels a longer path than light that arrives at the detector center. For large fields-of-view, the additional path length produces a distorted, or aberrated, image.

Milojkovic and Mait investigated the trade-off between optical performance and physical size for imagers with a large field-of-view in "Space-bandwidth scaling for wide field-of-view imaging" (page 13). They considered two different types of lenses, a conventional plano-convex lens and a monocentric lens—i.e., one in which the front and back surfaces of the lens have a common center, combined with two different types of detectors, a conventional flat detector array and a curved detector array. For all cases, they quantified optical performance and physical characteristics, such as size and weight, as they varied the lens diameter from a few micrometers to a few meters. Their analysis indicates that a monocentric lens imaging onto a curved detector outperforms other systems for the same requirements on optical performance. Unfortunately, a monocentric lens requires more glass than a plano-convex lens with the same focal length and, therefore, weighs considerably more. Milojkovic and Mait quantified the trade-off between weight and optical performance, and they also determined the minimum volume an imager must have to achieve a desired optical performance. Their results allow optical designers to balance resources against performance when designing imaging systems for persistent surveillance.

The relationship between aperture diameter and wavelength holds at all wavelengths. For the same aperture diameter, a longer wavelength implies worse resolution. This is the case for infrared (IR) radiation, whose wavelength is approximately 10 times longer than visible radiation. However, IR radiation conveys different information about a scene than visible radiation, which can be advantageous.

In the visible spectrum, people and objects reveal their physical characteristics only when illuminated by a source, such as the sun or indoor lighting. Measurements made in the IR spectrum reveal the temperature of an object. That is, a vehicle imaged using an IR system will look dark if it has not been used for hours and bright if its engine has just been turned off. Similarly, a patch of soil produces a different IR image when it is in direct sunlight versus when it is in shade.

In "Remote detection of buried land-mines and IEDs using LWIR polarimetric imaging," (page 27) Gurton and Felton exploit the properties of IR imaging to distinguish between disturbed and undisturbed soil, which can potentially improve the detection of buried explosives. Gurton and Felton demonstrated experimentally that IR imaging can reveal the differences in physical structure between disturbed and undisturbed soil.

Instead of measuring just the intensity of IR radiation reflected off a patch of soil, Gurton and Felton also measured the polarization of the radiation. Polarization is a property of an optical field that indicates the orientation of its oscillations. Sunlight, for example, has no preferred polarization. But sunlight reflected off a surface does.

Through a series of field experiments, Gurton and Felton showed that IR radiation from disturbed soil is more polarized than undisturbed soil. They also identified two sources of the polarization. The IR radiation given off by warm, disturbed soil is strongly polarized. Additionally, upon reflection, the disturbed soil polarizes the sun's IR radiation. Further, Gurton and Felton showed that the difference in the degree of polarization between disturbed and undisturbed soil is sufficiently large that one can reliably distinguish between the two. When combined with other sensor modalities, such as radar, IR polarimetric imaging could enhance our ability to detect landmines, IEDs, and other buried threats.

At even longer wavelengths, beyond IR, radiation phenomenology changes again. The region between optical and radio frequencies is the terahertz (THz) and gigahertz (GHz) regime, where wavelengths range between 300 μm and 3 mm. In this frequency band, it is possible to image through obscurants such as dust, smoke, fog, and even textiles. Thus, millimeter wave (mmW) imaging, as it is sometimes referred, offers potential solutions for helicopter pilots operating in degraded visual environments and for security personnel monitoring people for body-borne explosives at checkpoints or urban sites.

For the latter application, as in any persistence surveillance application, one desires a large field-of-view. However, with wavelengths four orders of magnitude longer than visible, this is physically not possible. Thus, a large field-of-view is achieved by scanning the imager across a scene. Presently, scanning is achieved mechanically. However, a more efficient method is to alter the radiation phase within the aperture of the imager so that its limited field-of-view is scanned across a scene electronically. In "Design of 220 GHz electronically scanned reflectarrays for confocal imaging systems," (page 45) Hedden, Dietlein, and Wikner considered the performance of an integrated reflectarray to provide electronic scanning. They examined the tradeoffs between reflectarray size, system size, and the number of resolvable image pixels. Consequently, they designed an imager that operates at 220 GHz (λ = 1.36 mm) with 8.3 cm resolution at 50 m. This resolution allows one to discern the barrel of a pistol or blade of a short knife at 50 m. The system aperture is 1 m and system length is 0.23 m, which provides the requisite resolution, a 30-degree full field-of-view, and still allows the system to be portable. The reflectarray is 5.4 cm × 5.4 cm, with 78 × 78 phase shifting elements spaced a half-wavelength apart.

Hedden, Dietlein, and Wikner also characterized the impact on the quality of an imaged point if one uses a 1-bit reflectarray—i.e., one that can realize only two phase shifts, 0 and π, in comparison to a 2-bit reflectarray, one that can realize four phase shifts, 0, $\pi/2$, π, and $3\pi/2$. In simulations, the 1-bit reflectarray generated significantly more noise and errors than the 2-bit reflectarray and did not meet the requirements for imaging. Their design and analysis provide circuit designers with requirements for reflectarray performance and systems designers with an architecture for future systems to detect body-borne devices in cluttered urban environments.

3.2. Detection

Given the large commercial market, development of detectors for visible imaging is primarily the domain of industry. However, the military continues to dominate the development of IR detectors. By observing the natural radiation given off by an object, the military can detect and track targets at a great distance without relying on a light source. For the same reason, it is also a more reliable way to achieve situational awareness. The many uses of IR imaging include, for example, night vision, large area surveillance and reconnaissance, helicopter piloting in degraded visual environments, detecting and countering unmanned aerial systems, gun-sights for armored vehicles and dismounted soldiers, covert search and rescue, and decoy countermeasures.

To ensure military dominance on the battlefield, the Army seeks to maintain its superiority in infrared detection. To achieve this goal, the focal plane arrays (FPAs) the Army deploys must excel in all areas of performance including thermal sensitivity, image resolution, speed of detection, pixel uniformity and operability, system reliability and robustness, operation readiness, and simplicity. To enable large-scale deployment, the technology must also be manufacturable and affordable.

As the corporate research laboratory of the Army, ARL is exploring new frontier science and technology to revolutionize the Army's IR capabilities. To ensure no opportunities are overlooked, the Army engages in a wide range of research, from the most conventional to the most exotic infrared materials. Currently, mercury cadmium telluride (HgCdTe) is the most sensitive IR material among competing materials. However, HgCdTe detector arrays are difficult to produce because the substrates needed to grow the material are not suitably large. HgCdTe is traditionally grown on bulk cadmium zinc telluride (CZT) substrates, which are lattice-matched to HgCdTe. However, CZT substrates are available only in relatively small sizes. Further, the difference in thermal expansion coefficients between a CZT substrate and its silicon (Si) read-out integrated circuitry reduces the reliability of large format FPAs due to repeated thermal cycling.

Some in the community believed this problem could be overcome by growing HgCdTe on composite substrates consisting of cadmium selenide and telluride, and silicon (Cd(Se)Te/Si). They also believed this approach could potentially provide a route to affordable, robust third-generation FPAs. However, due to the lattice mismatch between Cd(Se)Te and Si, this approach leads to high dislocation densities (greater than mid x 10^6 cm^{-2}), which degrades performance.

Alternatively, one can change the IR material, and researchers at ARL have considered mercury cadmium selenide (HgCdSe)

as an alternative to HgCdTe. With a lower variance in lattice constant, HgCdSe yields better lattice-matching to alternate substrates and offers better multi-color performance.

Additionally, HgCdSe offers flexibility in the choice of substrate. With a lattice constant near 0.61 nm, HgCdSe is well-suited to be grown on either gallium antimonide (GaSb) or zinc telluride (ZnTe) substrates. GaSb is available as a bulk substrate with dislocation densities ~10^4 cm^{-2}, and ZnTe can be alloyed with zinc selenide (ZnSe) to form lattice-matched $ZnTe_{1-x}Se_x$. There have been attempts to produce good quality material in the past, but ARL has produced the highest quality material reported.

The paper "Mercury cadmium selenide for infrared detection" by Doyle et al. (page 57) represents a significant step in reducing the background electron concentration of HgCdSe to a level suitable for fabricating devices. As described in the paper, by switching to high purity 6N Se source material, Doyle and his co-authors reduced the background electron concentration of HgCdSe samples by over an order of magnitude, from concentrations greater than 10^{17} cm^{-3} to the mid-10^{16} cm^{-3}. They reduced this further by annealing under Se. However, before HgCdSe devices are developed, this background concentration must be reduced by at least another order of magnitude. Current work is moving towards reducing the carrier concentration to develop IR detectors.

For less conventional IR technologies, quantum well infrared photodetector (QWIP) technology is a promising candidate except for one major weakness. Made from one of the highest quality semiconductor materials besides silicon, gallium arsenide (GaAs) QWIPs could easily fulfill all the Army's IR requirements. Unfortunately, this material requires an unusual detection scheme. To produce an electrical signal, incident light must travel sideways—i.e., parallel to the material layers. When IR light shines on the detector surface, as in most other detector technologies, no light is detected. The standard solution to this problem is to place a diffraction grating on top of individual detectors to disperse incoming light into different angles, thereby altering the propagation direction of the light. A portion of the light travels at a large angle and is detected. However, this approach has a quantum efficiency (QE) of only 5%. With 95% of the incident light undetected, QWIPs cannot provide the military with the IR sensitivity and imager speed it needs. Presently, 25 years after its invention, QWIP technology has not improved significantly, and it has long been deemed to be a low QE technology.

Nonetheless, ARL is determined to increase the QE of a QWIP and recent work shows great promise to increase it significantly. ARL's approach includes developing a highly accurate electromagnetic (EM) model to calculate the EM field inside a complex detector geometry, a capability that the infrared community heretofore did not recognize or pursue. After considerable effort, as reported in "Electromagnetic Modeling and Design of Quantum Well Infrared Photodetectors," Choi et al. (page 63) succeeded in establishing a finite-element EM model that one can apply to any arbitrary three-dimensional detector geometry. This model enabled the researchers to explain, for the first time, all the previously unexplained open literature experimental data, and for the first time to predict a definitive QE from any QWIP design. Aided by this EM model, the authors advanced a new detector concept, referred to as the resonator-QWIP or R-QWIP. The R-QWIP utilizes the detector volume as a resonant cavity to diffract, capture, and store the otherwise unabsorbed light until it is eventually absorbed by the detector. Choi predicted a QE for the R-QWIP as large as 75%. Subsequently, the authors tested the R-QWIP concept on five different detector materials and observed QEs ranging from 15 to 71%. One of the more modest QE materials was fabricated into imaging arrays and yielded a QE of 30%, all in accordance with the model predictions. Even with this modest QE, thermal sensitivity—i.e., the detector's lowest measureable temperature change is already 15 mK when it is operated at a 2.4 ms integration time. These metrics are many times better than the 20 mK operated at 20 ms in standard QWIP cameras, proving the potential of the new detector concept. Imaging arrays with higher QEs are under production.

Additionally, the developed EM model allows designers to have much greater control of the detector's optical properties. In the near future, different R-QWIPs will be produced to suit a wide range of applications—for example, narrow band imaging through dust clouds, narrow band detection of chemical gases, simultaneous two-color detection for infrared search and track, sequential two-color detection for target detection and identification, broadband detection for hyperspectral imaging, and circular polarization detection for biological imaging.

In addition to advancing QWIP technology, using the resonant storage of light to enhance absorption is also applicable to other infrared materials and other optical devices, such as solar cells. The paper by Choi et al. shows how ARL's basic and applied research can make an ineffective infrared technology useful to the Army and how this research can have an even broader impact on other scientific and technological areas.

To further strengthen the Army's IR capability, ARL also works on IR detection by altering and manipulating a material's optical properties in this wavelength regime. The paper "Passive infrared sensing using plasmonic resonant dust particles" by Mirotznik, et al. (page 75) reported a new way to control the IR emission spectrum from surfaces and particles. Most objects, either manmade or found in nature, reflect and emit IR radiation in a relatively smooth and broad spectrum; however, by applying structures with resonant absorption to the surface of those materials, the reflection and emission spectra can be enhanced or reduced at particular wavelengths. Moreover, by mixing small resonant particles (<100 μm) designed for several different wavelengths, one can form an IR dust that reflects or emits with a characteristic spectral signature. Such material-by-design

particles would be useful for a variety of practical applications. For example, when applied to a base surface, the resonant particles could be used to tune IR reflectance to mimic other natural or manmade surfaces.

The resonant particles and surface treatments are of particular interest for the Army. Potential applications include atmospheric sensing of chemical agents, calibration and training aids for hyperspectral imaging systems, and creation of custom infrared spectral signatures for passive friend-or-foe identification. For example, in the training-aid application, Soldiers could learn to use hyperspectral equipment to identify the spectral signature of dangerous chemical agents by observing an assortment of safe plasmonic samples that are designed to match the spectral signatures of the dangerous agents.

The paper presents computational and experimental results for particles that can be tuned to preferentially reflect or emit IR radiation within the 8–14 μm infrared band. The particles consist of thin metallic subwavelength gratings patterned on the surface of a simple quarter wavelength cavity. This design creates distinct IR absorption resonances by combining the plasmonic resonance of the grating with the natural resonance of the cavity. The resonance peaks are easily tuned by varying either the geometry of the grating or the thickness of the cavity. Measurements of reflection and emission from fabricated particles agreed with predicted performance. The tested particles use a one-dimensional grating that works for one polarization of incident light, but the paper also shows that a two-dimensional "fish net" grating should yield high-contrast spectral features for both incident polarizations. The author's next step is to design and fabricate surfaces or particles for practical Army applications so that the new structures can be demonstrated and tested under field conditions.

3.3. Post-Detection Processing

ARL researchers have been at the forefront of research in image processing and scene analysis. ARL actively conducts research on a large number of topics, such as automatic target recognition, multimodal sensor fusion, personnel detection, super-resolution, face recognition, object tracking from video sequences, and the use of biometrics for human identification. The following discussion addresses novel image processing techniques for four different applications: superresolution for video face recognition, multimodal sensor fusion, automatic target recognition in FLIR imagery, and target detection in hyperspectral imagery are presented.

Video imagery has become the most common and versatile form of media for capturing, analyzing, and disseminating a variety of information. Video surveillance systems have led to their widespread use on commercial properties and for residential monitoring. One major application of cheap, low-resolution video surveillance cameras is face recognition, which is crucial to aiding the law enforcement community and homeland security in identifying suspects and suspicious individuals on watch lists. However, face recognition performance is severely affected by the low resolution of individuals in typical surveillance footage, often due to the long distance between individuals and cameras, as well as the small pixel count of low-cost surveillance systems. Fortunately, super-resolution algorithms have the potential to improve face recognition performance by using a sequence of low-resolution images of an individual's face in the same pose to reconstruct a more detailed high-resolution facial image.

In "Face Recognition Performance with Super-resolution," (page 85) Hu, Maschal, and Young from ARL, in collaboration with Hong and Phillips from NIST, developed a super-resolution algorithm for face recognition, and conducted an extensive performance evaluation using a methodology and experimental setup consistent with real world settings, including multiple subject-to-camera distances.

Using the same low-resolution camera, facial images were obtained at far (~13 m), mid (~9 m), and close (~5 m) range. At the ranges, the face resolutions in terms of eye-to-eye distances were 5–10, 15–20, and 25–30 pixels, respectively. Hu et al. doubled the effective resolution of the system using digital processing using a sequence of eight low resolution images. They then submitted the super-resolved images to a state-of-the-art face recognition algorithm.

For recognition of faces at 9 m and, assuming a fixed 5% false alarm rate—i.e., one out of 20 times the system makes a false identification—the use of super-resolved images improved the rate of correct identification from 31% using original images to 45%. Their results show that super-resolution image reconstruction can improve face recognition performance considerably at the examined midrange and close range.

Forward-looking IR (FLIR) cameras provide the U.S. Army with the capability to see through darkness to detect and track objects of interest, such as humans and vehicles. One of the major military applications of FLIR imaging sensors is automatic target recognition (ATR), which seeks to detect and recognize objects of interest (targets) in an environment full of clutter and imaged by an imperfect sensor, which introduces noise into the resulting signal.

An ATR system consists of several stages. In the first, the system scans an entire image to identify regions of interest within which a potential target is detected. In the second stage, background clutter is removed. (Roughly, anything that is not considered part of the target is considered clutter.) In the third stage, the system computes a set of features that are used in the fourth stage to classify the target—e.g., bus, sedan, or tank, or even more specifically, an M1 tank versus a T-72 tank.

In “Sparsity-motivated automatic target recognition,” (page 97) Nasrabadi of ARL, in collaboration with Patel and Chellappa from the University of Maryland, developed a new classifier for long-wave IR imagery. The new classifier is based on the concept of using a dictionary of target templates and the theories of compressive sensing (CS) and sparse representation to reduce the amount of data required for high confidence classification. Their idea is to create a dictionary matrix of training samples from all classes of targets and represent the test sample as a few selected (sparse) linear combinations of the dictionary column vectors. This sparse representation is then used to infer the target type (class) of the input sample. They demonstrated that the performance of the proposed classifier is significantly better than classical classifiers, as well as previously developed classifiers from the ARL.

They also investigated the use of compressive sensing technique to reduce the dimensions of both the test samples and the training samples in the dictionary. They demonstrated that by reducing the dimensionality from the original target templates 40 × 65 pixels to only 256 incoherent measurements (features), the decrease in the classifier performance is insignificant. When only 64 features are used the classifier performance is still 78%; however, when a mere 16 features are used; the classifier performance degrades drastically to 43%.

As mentioned, IR imaging systems provide capabilities that visible cameras cannot, such as seeing through darkness, shadows, fog, clouds, rain, snow, and smoke. They are, however, subject to a number of inherent limitations, such as low resolution (in comparison to visible), the loss of non-thermal but important visual features (such as color and texture), and, under certain combinations of ambient and target temperatures, yield low thermal contrast between targets and background. Given that visible cameras are relatively low-cost, easy to use, and capable of producing high-quality imagery under favorable conditions, researchers have considered exploiting the advantages of cameras in each spectral band to improve target detection.

In “Fusing concurrent visible and infrared videos for improved tracking performance,” (page 107) ARL researchers Chen and Schnelle studied the usefulness of fusing visible color and long wave IR imageries to improve the detection and tracking of moving targets. Although a given sensor may be easily fooled sometimes, it is much harder to trick a number of sensors simultaneously at any given time. Consequently, Chen and Schnelle investigated several pixel-based image fusion algorithms using image pyramids generated by the Laplacian, contrast, gradient, morphological, and several variations of the Discrete Wavelet Transform (DWT) methods. Pixel-based methods perform fusion at the lowest level of image representation, the pixel. They do not require high-level abstract information and are, therefore, the simplest to implement. Chen and Schnelle performed digital detection and tracking on the fused images, and compared the performance across the fusion algorithms against the performance of just using visible imagery and just using IR imagery. Their results indicate that, in comparison to using just IR imagery, detecting and tracking performance degraded for fusion algorithms based on combining visible pixels with IR pixels. Performance was mixed for several pyramid-based algorithms, which use physical scale as a basis for representation, but were deemed inferior due to their high computational cost. Fusion algorithms based on the DWT provided improved performance with the lowest computational costs of all pyramidal methods. By exploiting the complementary strengths of visible and IR imagery, Chen and Schnelle demonstrated that fusion algorithms based on DWT image pyramids are capable of improving target detection and tracking.

Improved target detection based on the fusion of visible and IR imagery leads naturally for one to consider the advantages of using multiple wavebands. This is the motivation behind hyperspectral imaging. Hyperspectral cameras collect and process information across the electromagnetic spectrum by dividing a region of the spectrum into many narrow spectral bands, ranging typically from 50 to possibly 400 spectral bands. Hyperspectral sensors look at objects using a vast portion of the electromagnetic spectrum well beyond the visible range. Since different materials have unique spectral signatures, hyperspectral sensors are useful in agriculture and mineralogy to distinguish between crops and soil types. However, one also can use spectral signatures to locate materials that pertain to a particular target. Thus, hyperspectral imagery, which combines spectroscopy and imaging, is useful for object detection and classification.

A byproduct of hyperspectral imaging is the large amount of data that results. How one processes this data efficiently and effectively is considered by Nasrabadi from ARL, in collaboration with Chen and Tran from Johns Hopkins University, in “Sparse Representation for Target Detection in Hyperspectral Imagery” (page 121). Nasrabadi et al. developed a novel sparsity-based target detection algorithm to locate military targets in hyperspectral imagery in desert and forest environments. Their proposed algorithm uses known target and background signatures from a spectral library to construct a composite dictionary consisting of target and background sub-dictionaries. Then a hyperspectral test pixel is reconstructed approximately using very few training samples (i.e., sparse representation) from both target and background sub-dictionaries after imposing a sparsity constraint on the reconstruction. The recovered sparse representation is used directly to detect the presence or absence of a target in the hyperspectral test pixel. For targets that consist of multiple pixels, a smoothing constraint called the joint sparsity model is enforced in the reconstruction process to incorporate the assumption that neighboring spatial pixels consist of similar materials. By incorporating this contextual information directly into the classifier through the joint sparsity model, it is possible to enforce the classifier to make a joint decision on all the neighboring spatial pixels and improve the target detection performance. This also avoids the usual post-processing fusion of the classifier outputs on the neighboring pixels.

3.4. Computational Imaging

Image formation discussed up to this point uses optics to form an image, uses a detector to convert photons to electrons, and uses electronic processing to enhance information in the image. However, if one redesigns the front-end optics in an unconventional way and, concurrently, redesigns the post-detection processing, it may be possible to generate the desired information in a simpler manner, or it may be possible to generate information that would otherwise be costly to produce. Linking optical design and post-detection processing in this way is referred to as computational imaging, a field in which ARL is a leader.

Although the field of computational imaging grew out of advances in electronic detection and processing, holography, when viewed in retrospect, is one of the first computational techniques developed to improve resolution. As originally proposed, the post-detection processing was performed optically.

Developed in 1948 and made popular after the invention of the laser in 1960, holography interferes a reference beam coherently with a beam reflected from an object. The resulting interference pattern is the hologram. When the hologram is illuminated by the reference beam, the hologram produces an image of the object. With the advent of electronic detection and processing, the interference pattern is now recorded on a FPA and the image is reconstructed digitally. This is referred to as digital holography.

In "Digital holographic imaging of aerosol particles in flight," (page 135) Berg and Videen apply digital holography to characterize aerosol particles. Aerosol particle characterization has been a research priority in monitoring pollutants for many decades. In the atmospheric-sciences community, such characterization has overlapped strongly with atmospheric dynamics and chemistry. More recently, biological aerosols have been recognized as a health threat within the medical and security communities. While counting and sizing strategies have been around for many years, rapidly attaining other properties of aerosols has proven elusive. In addition to accuracy, the basic requirements are speed, low cost, and automation. Traditional imaging techniques are hampered by the depth-of-field required to produce a sharp image, as the uncertainties of the aerosol position within a flow generally are well beyond the depth of focus. Elastic light scattering also has been pursued, but retrieving information from the scattered field has proven difficult.

Berg and Videen overcome these problems by interfering the scattered light from the aerosol with a reference beam to form a digital hologram. By applying the Fresnel-Kirchhoff approximation in different focal planes, Berg and Videen overcame the focusing problem and reconstructed images from different depths. Morphological information pertaining to the aerosols can be retrieved directly from the reconstructed images. By applying holographic techniques to image aerosols in a flow, Berg and Videen provided a new tool to the aerosol community.

The depth-of-field problem—i.e., the depth of a region over which an imaging system is considered in focus—is inherent to all optical systems. As discussed in Sec. 3.1, millimeter-wave technology allows one to scan individuals for body-borne explosives. In controlled situations, such as an airport, authorities can scan individuals in a portal. In a dynamic urban setting, such control may not be possible, yet one would still like to scan individuals as they pass through a volume. However, the 1-m aperture of the system designed by Hedden et al. has an extremely narrow depth-of-field. In "94-GHz Imager with Extended Depth of Field," (page 145) Mait et al. used computational imaging techniques to extend the region over which a millimeter wave image remains in focus. Their approach is to aberrate the system in a known, controlled fashion (in their case, by using an optical element that has a cubic phase) and to perform simple post-detection processing. It is not possible to extend the depth-of-field by five times, as Mait and Wikner did, using conventional means without incurring considerable cost. The computational approach requires only one additional optical element and unique post-detection processing. However, the computational approach requires that one alter their notion of the function of optics in an imaging system given the availability and capability of post-detection processing.

Whereas the cubic phase approach is a fixed solution to a problem, the ability to sense an environment and adapt the optics based on those measurements to improve system performance is another aspect of computational imaging. In "Experimental demonstration of coherent beam combining over a 7 km propagation path," (page 153) Weyrauch et al. apply adaptive techniques for imaging horizontally. Unlike remote sensing, where cameras peer down through the atmosphere, tactical imaging is horizontal, i.e., parallel to the surface of the earth. Under these conditions, atmospheric turbulence limits system resolution and fidelity.

Weyrauch et al. demonstrated that they can control phase variations in the optical path with sufficient precision to combine coherently seven laser beams emerging from an adaptive fiber-collimator array over a 7 km atmospheric propagation path. This is a significant achievement, as this is not only more than an order-of-magnitude greater distance than previous experiments, but also extends the range into that of practical working distances of energy transmission through the atmosphere. To perform this feat, the wavefront phase at each fiber-collimator subaperture was controlled by its internal micro-fiber positioner. The output beams were combined coherently and focused onto their target. The phase locking and control of the wavefront were achieved by maximizing the target-return optical power using stochastic parallel gradient descent (SPGD) techniques.

This technique offers a lighter and more efficient system in resolving the complex atmospheric turbulence problem and improves the resolution of the system. Direct sensing from the target and beam control through the SPGD mechanism eliminated an external bulky and costly wavefront-detection system. The coherent beam-combining technique delivers kHz-rated phase-locking compensation and the maximum power to the target. The technology directly supports the Army's needs in developing tactical and long-distance sensing, imaging, communication, and directed energy systems.

High-resolution imaging of distant objects has many military applications. While lasers are used for long-distance illumination through the atmosphere, they create image speckle due to coherent phase aberrations. Atmospheric turbulence can cause these speckle spots to wander on the target, which can limit resolution.

In "Turbulence-free ghost imaging," (page 157) Meyers, Deacon, and Shih present a fundamentally new approach to meet this challenge. Turbulence-free ghost imaging is a computational imaging technique that can reconstruct an object that is not in a conventional sense imaged by the system. Through imaging experiments performed at ARL, Meyers et al. suggest ghost-imaging can be performed free of the adverse effects of turbulence.

Meyers, Deacon, and Shih used two-photon interference and the superposition of the quantum properties of light to image through turbulence. While a single-pixel light sensor sensed the total light reflected from an object, a second sensor camera imaged the photons coming just from the laser light source. The coincident measurements were combined computationally, creating the ghost image of the target.

Ghost imaging shares some attributes with conventional holography. In conventional holography, signal and reference beams interfere as coherent waves to form a pattern that generates an image. In ghost imaging, signal and reference beams combine to form an image based on the correlation between their quantum properties. To function properly, though, ghost imaging requires fast, single-photon-sensitive cameras.

The work by Meyers, Deacon, and Shih provides a unique, fundamental contribution to advanced imaging techniques. Ghost imaging is different from conventional imaging in that the illuminating light is imaged, as opposed to the object. Quantum two-photon interference provides improved resolution to the images as the aberrations caused by atmospheric turbulence are cancelled out in the quantum process.

4. The Future of Imaging

As the previous discussion indicates, advances in detectors and electronic processing have expanded the capabilities of optical sensing beyond replicating the appearance of an object and the computational imaging examples are only just the beginning. Faster detectors, such as the ones used for ghost imaging, and smaller cameras will allow optical designers to exploit both time and space to create even more capabilities. For example, wafer-scale cameras enable the development of small, multi-aperture cameras, which has allowed designers to create very thin cameras,[3] cameras with high dynamic range,[4] and cameras sensitive to polarization.[4]

Extremely fast detectors enable fast cameras and have enabled femtography.[5] In a single femtosecond (10^{-15} seconds), light propagates a few millimeters. Using a detector array capable of femtosecond exposure times, it is possible to track the propagation of optical beams and, consequently, to label rays. Using these labels, it is possible to keep track of rays as they go out of sight and return after being reflected by an object out of the camera's line of sight. In this way it is possible to perform non-line-of-sight imaging, to see around corners, without relying upon the quantum nature of light. Such capabilities will no doubt be of extreme value to the military. The future of imaging remains very bright.

5. References

[1] http://royal.pingdom.com/2012/01/17/internet-2011-in-numbers/

[2] C. Ronan, "Leonard and Thomas Digges," Journal of the British Astronomical Association Vol. 101 No. 6 (1991).

[3] A. Portnoy, N. Pitsianis, X. Sun, D. Brady, R. Gibbons, A. Silver, R. Te Kolste, C. Chen, T. Dillon, and D. Prather, "Design and characterization of thin multiple aperture infrared cameras," Appl. Opt. 48, 2115-2126 (2009). (doi: 10.1364/AO.48.002115)

[4] M. Mirotznik, S. Mathews, R. Plemmons, P. Pauca, T. Torgersen, R. Barnard, B. Gray, Q. Zhang, J. van der Gracht, P. Curt, M. Bodnar, and S. Prasad, "A practical enhanced-resolution integrated optical-digital imaging camera (PERIODIC)," Proc. SPIE 7348, Modeling and Simulation for Military Operations IV, D. A. Trevisani, ed., 734806 (2009). (doi: 10.1117/12.819484)

[5] A. Velten, D. Wu, A. Jarabo, B. Masia, C. Barsi, C. Joshi, E. Lawson, M. Bawendi, D. Gutierrez, and R. Raskar, "Femto-Photography: Capturing and Visualizing the Propagation of Light," ACM Transactions on Graphics (PROC. SIGGRAPH), 2013.

Space-Bandwidth Scaling for Wide Field-of-View Imaging

Predrag Milojkovic and Joseph N. Mait

Applied Optics 2012, 51 (4).

Space-bandwidth scaling for wide field-of-view imaging

Predrag Milojkovic and Joseph N. Mait*

U.S. Army Research Laboratory, 2800 Powder Mill Road, Adelphi, Maryland 20783, USA

*Corresponding author: joseph.mait@us.army.mil

Received 3 October 2011; revised 6 January 2011; accepted 10 January 2012;
posted 10 January 2012 (Doc. ID 154244); published 27 January 2012

We examine the space-bandwidth product of wide field-of-view imaging systems as the systems scale in size. Our analysis is based on one conducted to examine the behavior of a plano-convex lens imaging onto a flat focal geometry. We extend this to consider systems with monocentric lenses and curved focal geometries. As a means to understand system cost, and not just performance, we also assess the volume and mass associated with these systems. Our analysis indicates monocentric lenses imaging onto a curved detector outperform other systems for the same design constraints but do so at a cost in lens weight.

OCIS codes: 110.0110, 220.4830, 220.3620.

1. Introduction

The proliferation of imaging assets for security and defense has generated a demand for high resolution imaging across an ever increasing field-of-view (FOV). The off-the-shelf engineering solution to this problem is to tile the desired FOV with numerous high resolution cameras. This approach, though simple to implement, uses limited resources inefficiently, such as volume and weight. A more inspired approach is to redesign the optics. However, designers very quickly realize how difficult it is to maintain high resolution as the FOV increases. Using simple arguments reflected in Fig. 1, Lohmann showed that performance is limited primarily by optical aberrations [1].

Lohmann and his colleagues used space-bandwidth as a metric for imaging performance in a second, more analytic, study and examined the limits on space-bandwidth product (SBWP) as a function of FOV [2]. Their analysis of a simple plano-convex lens imaging onto a flat detector provided a rough indication of the physical scales when geometric aberrations, as opposed to diffraction, dominate imaging performance.

To provide high resolution over a wide FOV, researchers have proposed alternate lenses and alternate detector geometries to overcome aberrations [3,4,5,6,7,8,9,10]. For example, a curved detector removes geometric image distortions introduced at the edge of the FOV when imaging onto a flat detector [4] and fabrication of such detectors is an ongoing research topic [5,6,7,8]. Further, by its nature, a monocentric system, i.e., one in which the front and back surfaces of the lens have a common center, reduces the aberrations imposed on large off-axis rays. Monocentric triplets are typically used to correct chromatic aberrations in eyepieces [11]. Although interest in monocentric systems has recently increased [9,10], the advantages of such systems have a long history, as evidenced by the Sutton panoramic water lens patented in 1859 and the Baker ball lens from 1942 [3]. The aberrations can be reduced even further if the refractive index of the lens is graded [12].

In particular, [4] presents a quantitative analysis of the performance of three different imaging systems at a single scale: a plano-convex lens imaging onto a flat focal geometry, a Cooke triplet imaging onto a flat focal geometry, and a ball lens imaging onto a curved focal geometry. The linear scale of each system is approximately 10 mm, and, in addition to comparing on- and off-axis point spread functions, the analysis examines chromatic behavior for three visible wavelengths. The results indicate that a ball lens imaging onto a curved focal plane provides the best overall performance.

In our work, we explore how changes in optical design affect imaging performance as a function of system size. To do so, we base our work on that of Lohmann's and extend it to account for changes in lenses and in detector geometries. Using our analysis, we examine the impact on system size and weight

as performance demands increase and we highlight general trends. We do so, however, for only a single wavelength. We do not consider chromatic behavior. Further, our approach can be used by others as a framework for generating quantitative data when necessary.

In Section 2 we present the metrics Lohmann used to characterize optical performance and, in Section 3, we describe our method for imaging analysis. We present our data and provide a discussion of it in Section 4 and conclude in Section 5 with additional discussion and summary remarks.

2. Space-Bandwidth Analysis

In his analysis, Lohmann used the space-bandwidth product S as a measure of image quality. The space-bandwidth S is the number of resolvable points in an image plane,

$$S = \frac{A}{a_{\text{res}}}, \tag{1}$$

where A is the image plane area and a_{res} is the area of a single resolvable spot. One can approximate the resolution spot size a_{res} as the sum of contributions from diffraction and aberration [13],

$$a_{\text{res}} = (\delta x)^2 + (\delta \xi)^2 = (\lambda f/D)^2 + \bar{\xi}^2, \tag{2}$$

where λ is the wavelength of illumination, f and D are the lens focal length and diameter, respectively, and ξ is a measure of lateral geometric aberrations. The term $\bar{\xi}^2$ is its variance. To increase space-bandwidth, one can increase the size of the image plane, i.e., increase the FOV, reduce the size of the resolvable spot, or both. However, resolution area and FOV are linked to the image system and cannot be controlled independent of one another.

In [1], Lohmann assumed a constant FOV and considered only the impact on space-bandwidth as a function of the resolution spot size. In this case, if one considers only diffraction, an increase in lens diameter reduces the size of a resolvable spot, which, in turn, increases the space-bandwidth. However, because increasing the diameter of a lens also increases the impact of aberrations, the relationship between the size of a lens and its space-bandwidth is more complex than that specified simply by diffraction.

Lohmann used Eqs. (1) and (2) to generate the heuristic curves represented in Fig. 1, which indicate the relationship between the scale of an imaging system and its space-bandwidth. Figure 2 indicates how the imaging system is scaled by a factor M, while the f-number,

$$f_{\#} = f/D, \tag{3}$$

is held constant.

In the absence of aberrations, S will increase without bound as the imaging system increases in size. In contrast, because aberrations scale with the size of the imaging system, in the absence of diffraction, there would be no improvement in S as the system increases. Thus, the actual performance of an imaging system is a combination of these behaviors: increasing with scale in regions where diffraction dominates and constant in regions where aberrations dominate.

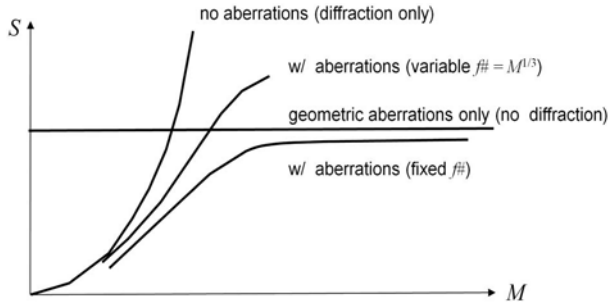


Fig. 1. Space-bandwidth of an optical system as a function of scale. Reproduced from [1].

The last curve generated by Lohmann indicates how designers compensate for the limitations imposed by aberrations. To increase S as system scale increases, the f-number of the system must necessarily also increase. In a second, coauthored publication, Lohmann explored this link between f-number and space-bandwidth, and considered also the FOV of the system [2]. It is this publication that provides the starting point for our analysis.

With reference to Fig. 3, Lohmann and his colleagues considered the space-bandwidth properties of the simple plano-convex lens represented in the upper left. In comparison to the analysis in [1], which considered fixed f-number lenses and a single scaling factor M, the analysis in [2] considered variable f-number lenses. In addition, it considered the FOV of the imaging system. As represented in the figure, we extend Lohmann's analysis to include monocentric lenses in addition to plano-convex as well as two different focal plane geometries, flat, and curved. We also consider the special case of a Luneburg lens with a curved detector on its surface. A Luneburg lens is monocentric with a variable refractive index,

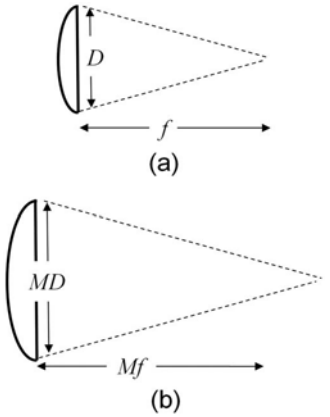


Fig. 2. Imaging system scaling considered by Lohmann. (a) Base system. (b) Scaled system.

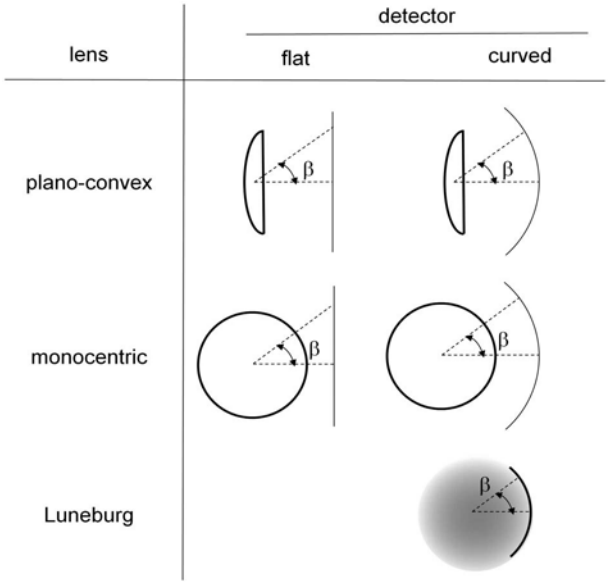


Fig. 3. Representative imaging systems considered for analysis.

$n(r) = [2 - (r/R)^2]^{1/2}$, that images collimated light perfectly on its surface [12].

A. Space-Bandwidth as a Function of Focal Plane Geometry

It is important to consider the geometry of the focal plane since it affects the numerator of Eq. (1) and the size of the detector sets the FOV. With reference to Fig. 4, the image area A as a function of the half-angle β for a flat focal plane represented is

$$A_{\text{flat}} = (2f_{\#}D \tan \beta)^2, \tag{4}$$

and for a curved focal plane,

$$A_{\text{curv}} = 2\pi(f_{\#}D)^2(1 - \cos \beta). \tag{5}$$

In the absence of aberrations, the size of a resolvable spot for a flat detector is

$$a_{\text{res,flat}} = \left(\frac{\lambda f}{D \cos^3 \beta}\right)^2, \tag{6}$$

and for a curved detector

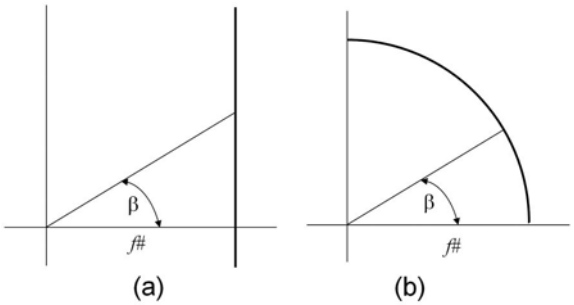


Fig. 4. Geometry for rays assuming a (a) flat and (b) curved focal plane.

$$a_{\text{res,curv}} = \left(\frac{\lambda f}{D \cos \beta}\right)^2. \tag{7}$$

The cosine-cube scaling of the conventional diffraction limited spot on a flat detector results from the oblique incidence of the illumination on the aperture, the increased distance rays have to travel (relative to on-axis), and the oblique incidence of the illumination on the detector. With a curved detector, the latter two factors are removed and only the cosine scaling due to the oblique incidence of the illumination on the aperture remains. As a point of reference, if this cosine scaling is absent, the resolvable spot follows from the conventional definition,

$$a_{\text{ideal}} = \left(\frac{\lambda f}{D}\right)^2. \tag{8}$$

This assumes a uniform beam illuminates the lens aperture. Also, in essence, we approximate the diameter of the spot by its full-width at half-maximum.

Thus, Eq. (1) in combination with Eqs. (4)–(7), yields closed-form expressions for the space-bandwidth given the two detector geometries,

$$S_{\text{flat}} = 4(D/\lambda)^2(1 - \cos^2 \beta)\cos^4 \beta, \tag{9}$$

$$S_{\text{curv}} = 2\pi(D/\lambda)^2(1 - \cos \beta)\cos^2 \beta. \tag{10}$$

These functions are represented in Fig. 5. The existence of a maximum is due to the fact that, for small angles, the rate at which the spot size grows is slower than the rate at which the detector area increases. For large angles, this rate behavior reverses. The angle at which this behavior switches is the one that yields maximum space-bandwidth. For a flat detector, this angle is $\beta = 35.2°$ and for a curved detector, $\beta = 48.2°$. Not only is the angle that maximizes space-bandwidth larger for a curved detector than a flat one, the value of the space-bandwidth is 1.5 times larger. This provides some indication of the advantages of a curved detector over a flat one.

The ideal behavior noted in Fig. 5 is derived assuming no aberrations and no geometrical scaling.

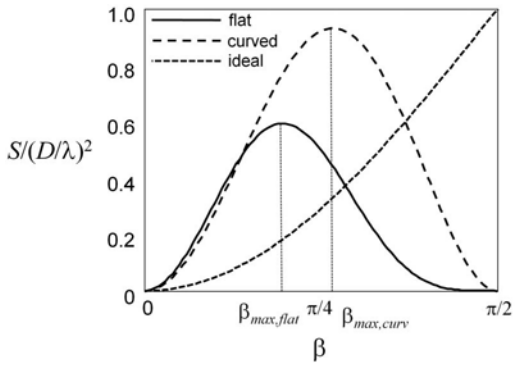


Fig. 5. Space-bandwidth product S as a function of half-field angle β for flat and curved detectors.

Thus, as the area of a curved detector increases, the number of resolvable spots increases without limits. The normalized curve shown follows $1 - \cos \beta$.

However, this analysis is valid only in the absence of aberrations. How far the space-bandwidth deviates from the values in Eqs. (9) and (10) is a measure of the severity of aberrations. To assess this, we need to analyze the optical performance of the lenses in combination with these detectors. In the next section, we describe the procedure we used to do so.

3. Lens Analysis

In this section we present the results of our optical simulations to assess the impact of both diffraction and geometrical aberrations on space-bandwidth. The lenses we analyzed (plano-convex, monocentric, and Luneburg) are represented in Fig. 6.

Our analysis followed as closely as possible the approach in [2]. We analyzed performance at a single wavelength $\lambda = 500$ nm, i.e., we did not consider chromatic behavior, and, except for the Luneburg lens, assumed $n = 1.5$. (We selected real glass materials that met this criteria.) We varied the lens diameter, D, from 50 μm to 5 and m the f-number, $f_{\#}$, from 1 to 1000.

For each lens and detector geometry, we varied the angle of the incident beam over the FOV and determined its corresponding spot size on the detector. To do so, we launched a large number of rays from a particular field point, including the chief ray, into the entrance pupil of the lens, and traced them through the optics to the image surface. We calculated the spot size as the root-mean-square of the differential distance between the location of the chief ray and the locations of the other rays. We also calculated the diffraction spot size from Airy disk.

We inserted these values into Eq. (2) to determine S and, in accordance with [2], assumed the value of resolution spot size was valid across the entire detector plane. We designated the maximum value of space-bandwidth $S_{\max}$ and the angle at which this maximum was achieved, $\beta_{\max}$.

The details pertaining to each lens and detector geometry are described in Appendix A. It is important to note that, again in accordance with the approach used in [2], we did not consider all potential means to improve lens performance. For example, we did not use any aspherical surfaces.

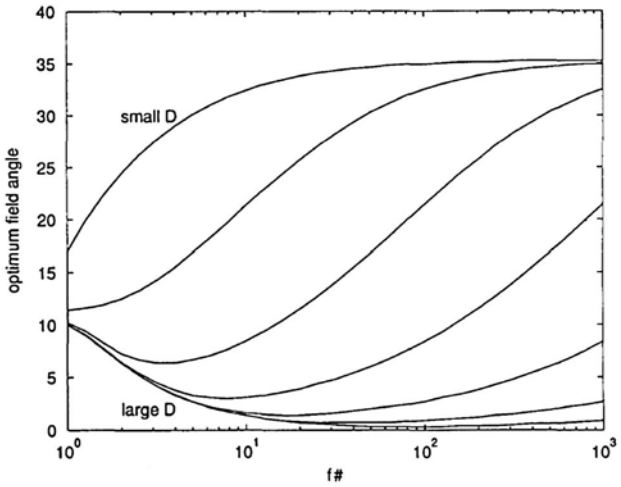


Fig. 7. $\beta_{\max}$ as a function of $f_{\#}$ for variable lens diameters for a plano-convex lens imaging onto a flat detector. Reproduced from [2].

One example of the results presented in [2] is reproduced in Fig. 7. Our results are presented in Figs. 8–12. We feel the consistency between Figs. 7 and 8(a) validates our approach and gives us confidence in drawing conclusions from the data for different systems, which we present in Section 4.

In addition to Fig. 7, we compared other data from [2] to ours. Our data were consistent with [2] in all cases, e.g., spot size as a function of angle. However, we felt that including all comparisons would have overwhelmed the reader and added little to the discussion. We present the space-bandwidth data for its relevance to this discussion. To aid the comparison between our work and [2], Fig. 8 contains data for an additional lens with $D = 5$ μm that the other lens systems do not. ZEMAX was unable to converge consistently to a design for all of systems with such a small diameter.

We note the special case of the Luneburg lens in Fig. 12. The Luneburg lens, in effect, provides ideal imaging because the area of the smallest resolvable spot is independent of the angle of incidence. The space-bandwidth is maximum at 80° because that was the upper limit of angles we considered. (Beyond 80° ZEMAX was unable to generate consistently stable data). Based on our data, the space-bandwidth of the Luneburg lens as a function of angle β is

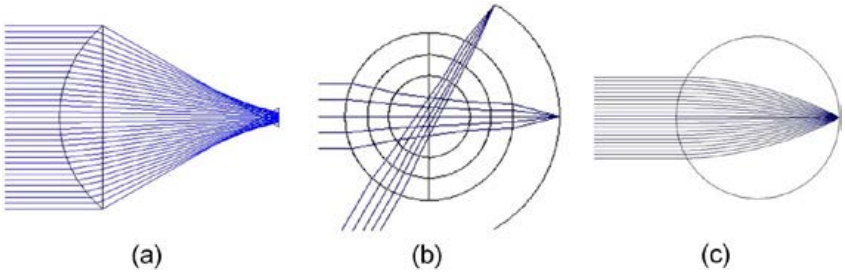


Fig. 6. (Color online) Lenses analyzed. (a) Plano-convex. (b) Monocentric. (c) Luneburg.

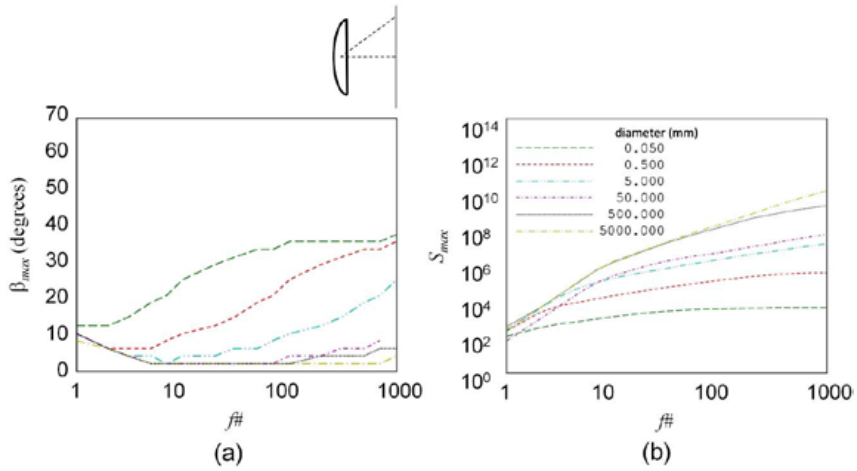


Fig. 8. (Color online) Analysis of a plano-convex lens imaging onto a flat detector. (a) $\beta_{\max}$ and (b) $S_{\max}$ as a function of $f_{\#}$ for variable lens diameters.

$$S_{\text{Luneburg}} = 1.34(D/\lambda)^2(1-\cos\beta). \tag{11}$$

The compact presentation of data presented in Figs. 8–12 belies the effort required to generate it. Data collection for a single combination of lens type and detector requires the analysis of elements with six different diameters and, for each diameter, 16 f-numbers, for a total of 96 lenses. Since we analyzed 31 field points for plano-convex lenses and 41 field points for monocentric lenses, we therefore determined 2976 spot size data values for a single detector geometry with a plano-convex lens and 3731 values for a single detector geometry with a monocentric lens. Not including the analysis of the Luneburg lens, this means we generated and analyzed over 13,000 optical spots to generate the graphs presented in this section.

Needless to say, our data collection was not performed manually. Instead, we implemented an automated method of data collection using a ZEMAX extension file written in C++ to drive ZEMAX externally. Because of the unstable nature of the optimization algorithm in the programming environment required for this task, development of the automation method took a considerable amount of time to test and validate.

4. Data Interpretation

In this section, we attempt to draw conclusions from our results. We need to be cautious when making comparisons between systems because each was optimized using slightly different criteria. Nonetheless, we feel the trends displayed are consistent and allow us to provide some explanation for the behavior exhibited. For example, the improvement in maximum space-bandwidth in Fig. 9(b) over Fig. 8(b) is due most likely to the close match between the curved detector and the Petzval surface shape.

Figures 8 and 10 reaffirm the conclusions from [2]. Only for physically small lenses and lenses with large f-numbers are aberrations sufficiently small that diffraction dominates their performance. Small f-number lenses and large diameter lenses are most affected by aberrations.

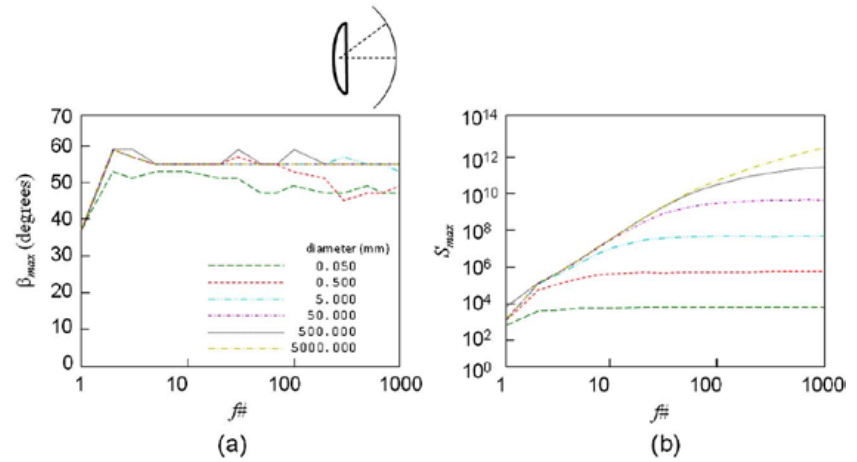


Fig. 9. (Color online) Analysis of a plano-convex lens imaging onto a curved detector. (a) $\beta_{\max}$ and (b) $S_{\max}$ as a function of $f_{\#}$ for variable lens diameters.

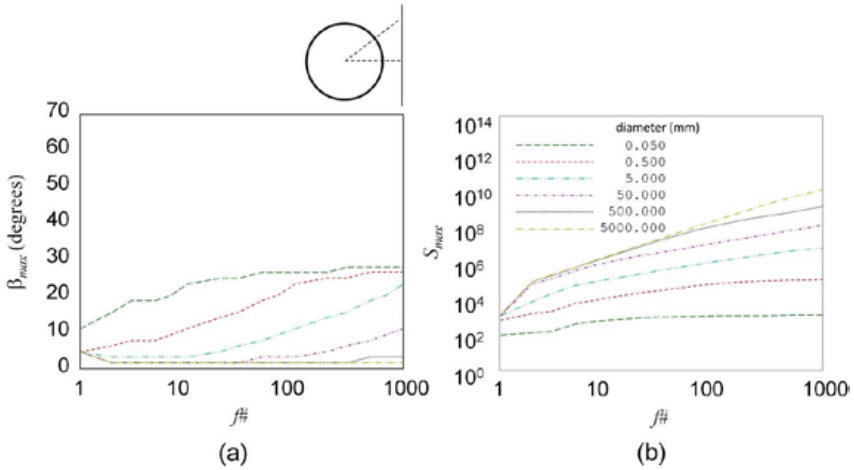


Fig. 10. (Color online) Analysis of a monocentric lens imaging onto a flat detector. (a) $\beta_{\max}$ and (b) $S_{\max}$ as a function of $f_{\#}$ for variable lens diameters.

We note also in these figures that the FOV defined by diffraction is an upper bound. This indicates that, for small f-number lenses, the size of the resolution spot from aberrations is larger than the diffraction spot size. But, as f-number increases, the impact of aberrations lessens and behavior is dominated by diffraction. This behavior is true for both plano-convex and monocentric lenses used in conjunction with a flat detector.

As is evident in Figs. 9 and 11, this changes with a curved detector. In fact, the behavior of plano-convex and monocentric lenses differs considerably. For a plano-convex lens, the angle that achieves the maximum space-bandwidth starts at 38° for an $f/1$ lens, increases rapidly to a peak angle, and decreases slowly to the angle defined by diffraction. Lenses with small diameters approach the diffraction-defined angle more rapidly than large lenses. Large f-number lenses are again dominated primarily by diffraction.

The shape of the curves in Fig. 9(a) reflect the interplay between aberrations and diffraction for off-axis angles. For large off-axis angles, the aperture is effectively stopped down. This increases the diffractive spot to a size that is, apparently, comparable to the one generated by off-axis geometric aberrations. We note that, although a curved detector geometry increases the value of $\beta_{\max}$ for small $f_{\#}$ lens, the increase in $S_{\max}$ is small. This is due to the definition of space-bandwidth, which assumes a constant spot size across the entire image. The spot sizes produced for large off-axis angles are larger than those for on-axis spots. An alternative definition of space-bandwidth, one that allows for variation in spot size across the image, might show a more dramatic increase in space-bandwidth with angle.

For a monocentric lens with a curved detector, all rays provide effectively on-axis performance. However, geometric aberrations are reduced due to the stopped-down aperture at large off-axis angles. Figure 11 therefore reflects this reduction in off-axis geometric aberrations.

The behavior exhibited in Fig. 12 is easily explained by Eq. (11). Since the value of $\beta_{\max}$ is the

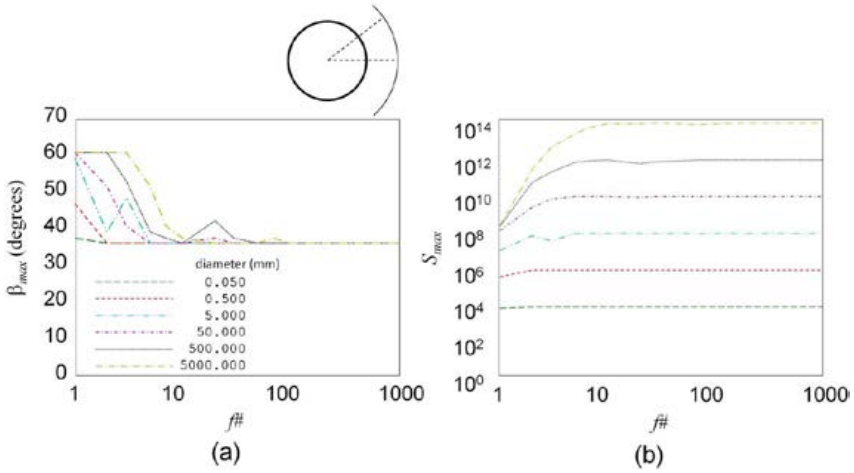


Fig. 11. (Color online) Analysis of a monocentric lens imaging onto a curved detector. (a) $\beta_{\max}$ and (b) $S_{\max}$ as a function of $f_{\#}$ for variable lens diameters.

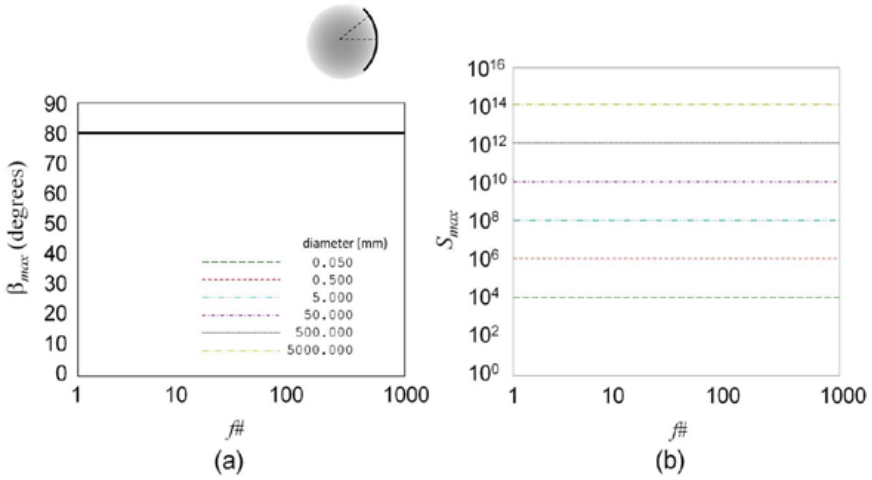


Fig. 12. (Color online) Analysis of a Luneburg lens imaging onto a curved detector. (a) $\beta_{\max}$ and (b) $S_{\max}$ as a function of $f_{\#}$ for variable lens diameters. Since the value of $\beta_{\max}$ is independent of $f_{\#}$ and D, all curves in (a) lie on top of one another.

same in each case, the space-bandwidth is constant for a fixed diameter.

Figure 13 compares the performance of each system we analyzed for two different scales, $D = 50$ mm and $D = 500$ mm. These scales represent the ones a designer might consider for imaging with 100 megapixel to 10 gigapixel detectors. As noted in the beginning of this section, we exercise caution in making these comparisons and concentrate primarily on trends and relative magnitudes. For f-numbers greater than 10, detector geometry dominates and a curved detector provides about 2 orders of magnitude improvement in space-bandwidth over a flat one. Given that large f-number lenses have long focal lengths, most likely detector shape dominates because the distinction between lens characteristics is marginal.

For f-numbers less than 10, lens properties dominate. For a flat detector, a monocentric lens provides approximately 2 orders of magnitude improvement in space-bandwidth over a plano-convex lens and approximately 4 orders of magnitude improvement for a curved detector.

For both scales, the performance of a plano-convex lens with a curved detector is comparable to a monocentric lens with a flat detector. That is, equivalent performance can be achieved using either a poor quality lens with a curved detector or high quality lens with a flat detector. However, replacing a flat detector in a monocentric system with a curved detector improves performance for low f-number lenses by several orders of magnitude. The curved detector, no doubt, takes full advantage of the increased FOV provided by a monocentric lens. The Luneburg lens, which essentially represents an upper bound on performance, provides additional improvement over extremely fast ($\approx f/1$) monocentric lenses.

To underscore the link between our work and [1], we reformat our data in Fig. 14. For each imaging system, we present space-bandwidth as a function of lens diameter D for variable f-number lenses. Results for the Luneburg lens are plotted as a dashed

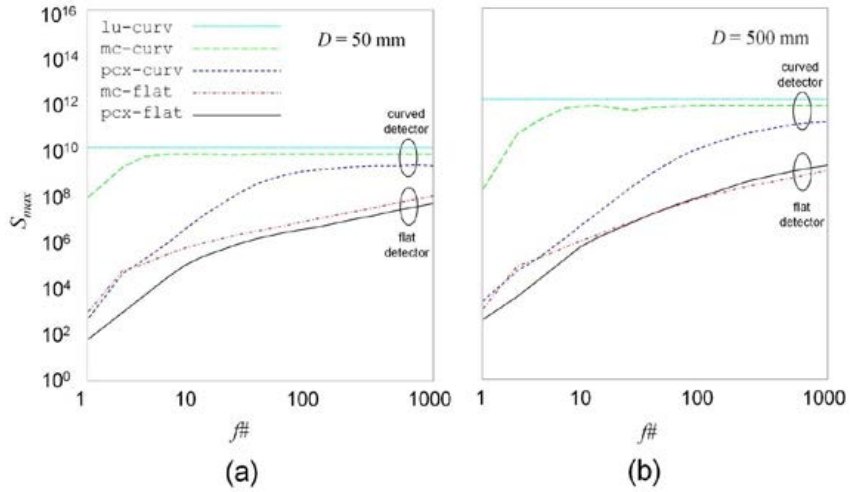


Fig. 13. (Color online) $S_{\max}$ as a function of $f_{\#}$ for various imaging systems with (a) $D = 50$ mm and (b) $D = 500$ mm. Labels for the graphs indicate lens type (pcx—plano-convex, mc—monocentric, lu—Luneburg) and detector geometry (flat versus curved).

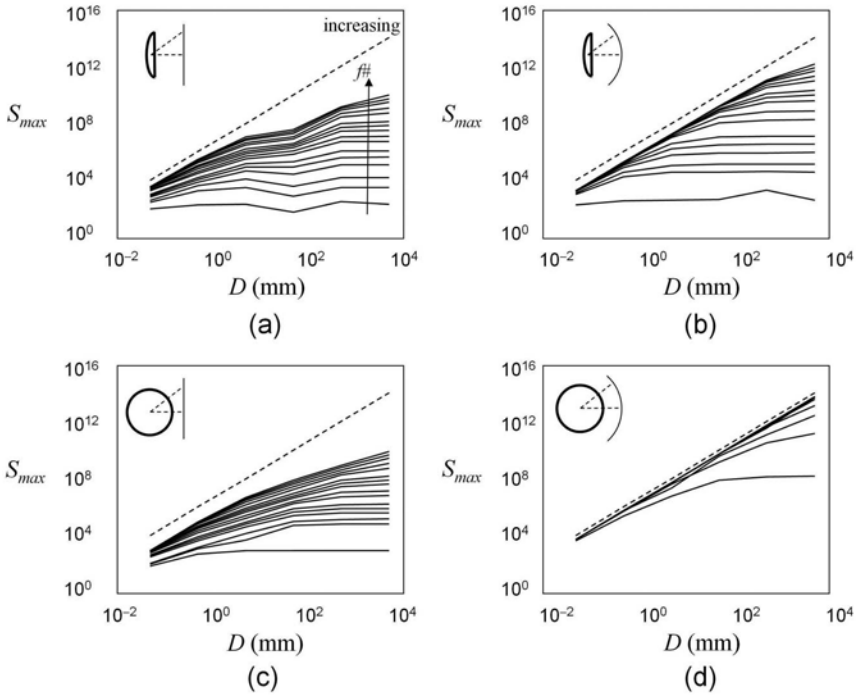


Fig. 14. $S_{\max}$ as a function of D for various imaging systems. (a) Plano-convex lens and flat detector. (b) Plano-convex lens and curved detector. (c) Monocentric lens and flat detector. (d) Monocentric lens and curved detector.

line in each figure. Note that, with space-bandwidth plotted on a logarithmic scale, a quadratic function is a straight line with slope 2 (i.e., an order magnitude change in linear scale yields 2 orders of magnitude change in space-bandwidth).

As predicted by Lohmann in Fig. 1, for a fixed f-number lens, as the size of the lens increases, the space-bandwidth saturates and the only way a designer can increase space-bandwidth is to change the imaging system. Given the discussion in [1], this is possible only by increasing the f-number of the lens. This is in contrast to our approach, which considers alternate lenses and detectors.

Figure 14 highlights again the advantages of a monocentric lens and a curved detector. Most systems exhibit quadratic behavior. However, the space-bandwidth of $f/1$ and $f/2$ lenses are already saturated when $D = 100$ mm.

In addition to supporting Lohmann's explanation for lens behavior, our analysis can also be used as an aid to design. Figure 15(a) indicates the system volume V_s required to achieve a desired space-bandwidth for each scale and lens system we analyzed, where

$$V_s = V_\ell + V_a, \tag{12}$$

V_ℓ is the lens volume, and V_a is the volume required to image onto the detector. Although we used all lens systems to generate Fig. 15(a), not all data points are shown. Instead, the values shown for each system are those that produce maximum space-bandwidth for a minimum amount of volume.

The formulae we used for V_ℓ and V_a are listed in Tables 1 and 2. For plano-convex lenses, we used the volume determined by ZEMAX. For all systems, h is the axial image distance from the lens center. We assumed the volume of the Luneburg lens is the same as the monocentric lens. There is no volume of air for the Luneburg lens because the image surface is coincident with the lens surface.

The trends indicate that systems with flat detector planes are least efficient in terms of volume for a given space-bandwidth. An order of magnitude increase in space-bandwidth requires a three-order magnitude increase in volume or, more simply, an order magnitude increase in each dimension of the box that defines the imaging system. Note that a monocentric lens, with its spherical shape, requires more volume than a plano-convex lens. This is reflected as a constant volume offset.

Systems with curved detector geometries use volume more efficiently. With a detector plane attached to the surface of the lens, the Luneburg lens uses volume most efficiently. The volume increases by only 1.5 orders of magnitude for an order of magnitude change in space-bandwidth. The monocentric lens has a slope of 1.6 and the plano-convex lens, 2.2. If one discounts the Luneburg lens as an impossible ideal, one can conclude from Fig. 15(a) that, given a fixed amount of volume, the monocentric lens with a curved detector yields the highest space-bandwidth.

However, in addition to volume, scaling of lens mass m_ℓ (which is easily converted to weight) is also a critical parameter of interest. This is represented in Fig. 15(b). Again, the values shown are those that produce maximum space-bandwidth for a minimum amount of mass. Not surprisingly, the spherical shape of a monocentric lens is detrimental in terms

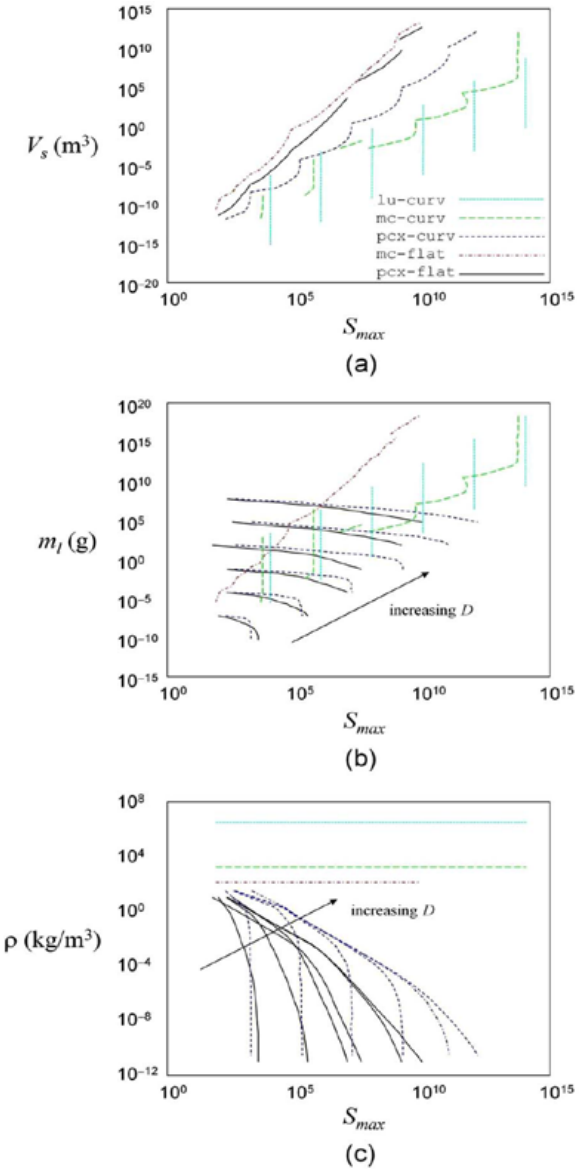


Fig. 15. (Color online) Physical characteristics as a function of space-bandwidth for imaging systems analyzed. (a) System volume. (b) Lens mass. (c) System density.

of mass. A monocentric lens imaging onto a flat detector geometry is least efficient in terms of mass because a larger lens is required to achieve the same space-bandwidth when compared to a curved detector geometry. The mass of a plano-convex lens is reduced with increasing space-bandwidth because lens f-number increases and, therefore, its thickness is reduced. [See Fig. 16 and the discussion contained in Appendix A.] Thus, an order of magnitude increase in space-bandwidth reduces the mass by one-third an order of magnitude.

Table 1. V_ℓ for Imaging Systems Analyzed

lens	V_ℓ
plano-convex	ZEMAX
monocentric	$\frac{4\pi}{3}R_\ell^3$

System density ρ as a function of space-bandwidth is represented in Fig. 15(c), where

$$\rho = \frac{m_\ell}{V_s}.$$

We use this definition based on our assumption that the lens is the primary contributor to system mass. Note that systems with monocentric lenses exhibit a constant density, whereas the density of plano-convex systems decreases with increasing space-bandwidth. This reduction in density is due to the reduction in mass as mentioned previously. The constant density for monocentric lens systems indicates lens size is the dominant characteristic. An increase in space-bandwidth demands an increase in lens radius, which dictates system volume and mass.

If one compares Figs. 15(a), 15(b), and 15(c), it is apparent that the weight advantage offered by plano-convex lenses is negated by its large volumetric costs. Conversely, a heavy system is the price one pays for the optical performance provided by a monocentric lens.

In our final comment, we note that space-bandwidth is a measure of optical resolution elements, or resels, which differs from the number of pixels in a detector array. Depending upon the size of a single detector pixel, the number of pixels per resel can vary. There should be at least one pixel per resel but the number can be between four and eight for a well sampled system. Given this caveat, Fig. 15 can provide some indication about the system scale and weight required to achieve a given imaging capacity. For example, Fig. 15 indicates that a 10-gigapixel (Gpx) imager should require on the order of 1 m^3 of volume.

5. Summary and Conclusion

We examined the space-bandwidth of wide FOV imaging systems as the systems scale in size. We extended Lohmann's analysis for plano-convex lenses imaging onto flat focal geometries to systems with monocentric lenses and curved focal geometries, which have been proposed as technologies to provide high resolution wide FOV imaging. To understand

Table 2. V_a for Imaging Systems Analyzed

	detector geometry	
lens	flat	curved
plano-convex	$\frac{\pi}{3}h^3\tan^2\beta$	$\frac{2\pi}{3}h^3(1-\cos\beta)$
monocentric	$\frac{\pi}{3}h^3\tan^2\beta - \frac{2\pi}{3}R_\ell^3(1-\cos\beta)$	$\frac{2\pi}{3}h^3(1-\cos\beta) - \frac{2\pi}{3}R_\ell^3(1-\cos\beta)$

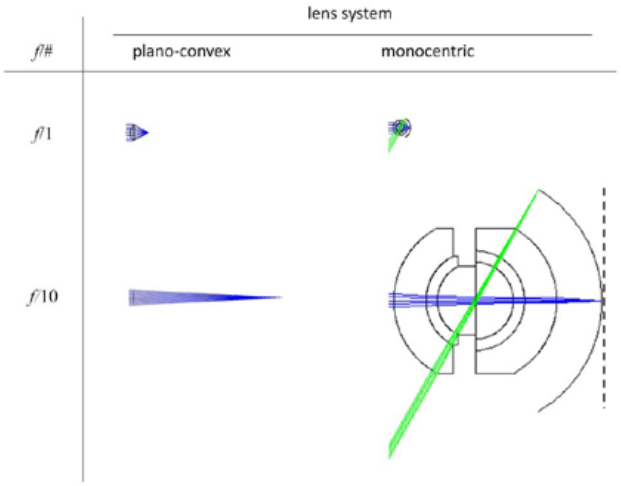


Fig. 16. (Color online) Optical system scaling.

system cost, and not just performance, we also assessed the volume and mass associated with these systems.

The most straightforward conclusion we draw, that a curved detector improves the performance of an imaging system, provides justification for the DARPA Hemispherical Array Detector for Imaging program [8]. Our analysis, however, also supports development of monocentric lens systems. Our analysis indicates that monocentric lenses imaging onto a curved detector outperform other systems for the same design constraints but do so at a cost in lens weight.

More important than simply providing these simple conclusions, our work also provides a framework for analysis. Our particular instantiation of the framework suggests the above conclusions but we recognize that further investigation is required to validate their universality. For example, one can use our analysis technique to consider the performance of multiscale optical systems, such as those for chip-to-chip optical interconnects [14] and those to reduce aberrations in wide FOV imaging [15].

We note that Fig. 15 considers only scaling the optics. An analysis useful for a designer must consider scaling of the postdetection electronics as well. In fact, a more exact analysis would consider the curved detector approximated by a collection of two-dimensional flat detector arrays. Is there an optimal array size under such conditions?

Finally, in this work we have considered only optical means to reduce the impact of aberrations. We have not considered the performance of optics combined with electronics. A recent analysis based on Lohmann's heuristic link between f-number and scale indicates some improvement is possible [10]. If the scale factor $M = f_{\#}^{1/3}$ as Lohmann prescribes, space-bandwidth scales roughly as $M^{4/3}$. Using computation after detection, it is possible to change this to $M^{2.8/1.9} \approx M^{3/2}$. This remains to be verified. Even if true, it is imperative that a designer know the cost of implementation for a given level of performance. That is, which implementation is more costly, a monocentric lens with a curved detector or a plano-convex lens with a curved detector followed by computation? We hope to create an analysis framework that will allow us to address this point.

Appendix A: Lens Analysis

We describe in this appendix the procedures we used to optimize and determine lens performance. The procedures were embodied in several C++ programs which we used to drive the optical system design and analysis software package, ZEMAX, externally. We used its raytracing and optimization capabilities to collect a huge set of data. Figure 16 indicates the relative scale of the systems we considered. The width of the illuminating beam is constant in each case. Because of the wide range of lens parameters in which we were interested for this study, ZEMAX had trouble converging to physical solutions in all cases. Each lens type and size, therefore, required some manual adjustment to generate usable and repeatable data. Nevertheless, the capability to automate the data collection process proved invaluable in collecting over 13,000 spot size data for further analysis. Figure 17 presents representative spot diagrams at different angular positions for each of the imaging systems at a single scale and f-number. The spot size predicted by diffraction in each case is considerably less than the scales shown.

Plano-Convex Lenses

To achieve a desired f-number for a plano-convex lens with a fixed diameter, we adjusted the curvature of the front surface, the thickness of the lens, and the distance between the lens back surface and the on-axis focal point. We located the flat detector plane at the paraxial focus.

With the detector plane fixed, we varied the angle of the incident beam between 0 and 60° in 2° increments and, for each angle, adjusted the lens' curvature and thickness to maintain a constant image plane and constant $f_{\#}$. We determined the spot size assuming a flat detector plane by launching a large number of rays from a particular field point, including the chief ray, into the entrance pupil of the lens, and tracing them through the optics to the image surface. We calculated the spot size as the root-mean-square of the differential distance between the location of the chief ray and the locations of the other rays. We also calculated the diffraction spot size from Airy disk.

To determine S, we inserted these values into Eq. (2) and, in accordance with Lohmann, assumed the value of resolution spot size was valid across the entire detector plane. We selected the maximum value of S over the angles calculated.

Without changing the lens, we modified the curvature of the image surface to measure the spot size on a curved detector plane. We set the curvature of a spherical surface such that the lens $f_{\#}$ was maintained yet produced the smallest spot size for that field angle (i.e., we determined the circle of least confusion). Although the curvature is slightly different for each field point, all surfaces intersect the optic

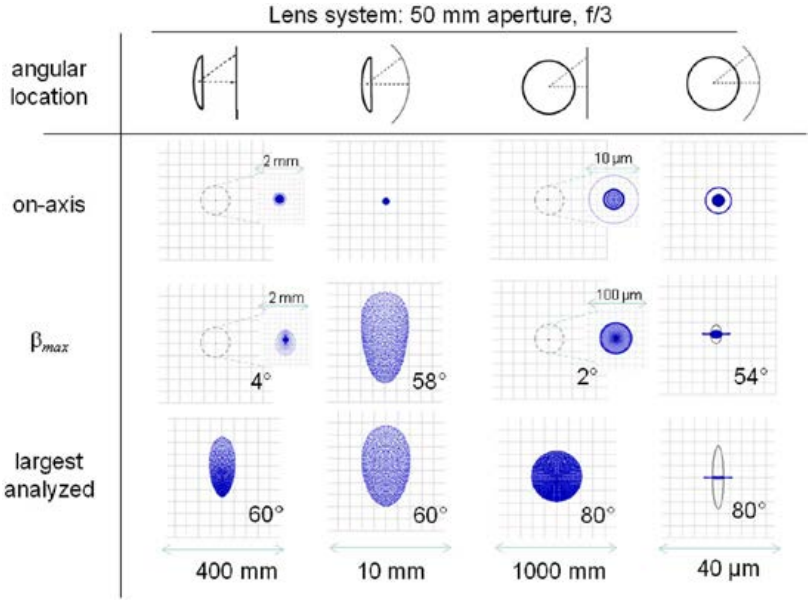


Fig. 17. (Color online) Representative spot shapes from analysis.

axis at the paraxial focus. The curved surface, in effect, followed that of the Petzval surface.

Monocentric Lenses

The monocentric lens consisted of a central spherical element and two concentric outer spherical shells, all concentric around the same point. We specified that all elements were made of commercial glasses and cemented to each other. Analysis of the lens was similar to that for plano-convex lens.

To determine the optimum position of the curved image surface, we assumed it was spherical and concentric with the lens, and we optimized the lens performance (primarily $f_{\#}$) only at two field points, 0° and 60°.

We determined the geometrical spot size for incident angles ranging from 0 to 80° in 2° increments. It was necessary to increase the range of angles for the monocentric lenses because they reached their maximum above 60°. Small lenses reached their maximum space-bandwidth below 80°, but ZEMAX did not produce stable results when we analyzed large lenses above 80°. We therefore limited our maximum angle for analysis to 80°.

Using raytracing, we calculated first the geometrical spot size on a curved image surface. We then calculated the geometrical spot size on a flat image plane by placing the image plane at the point where the curved plane intersected the optic axis. To achieve the required focal length for the spherical elements, we changed the lens f-number by adjusting the lens size while keeping its aperture constant.

Luneburg Lenses

We used the Luneburg lens as the limiting case on performance for a similarly sized monocentric lens. We scaled the Luneburg lens to match the f-number and aperture of a monocentric lens yet simultaneously provide aberration-free imaging. Because a fully illuminated Luneburg lens is always $f/0.5$, we had to stop down the lens to achieve $f/1$. To comply with Lohmann's approach, we changed the lens f-number by changing its lens size while keeping its aperture size constant.

Because ZEMAX was unable to trace accurately the off-axis rays through this type of lens, we generated most of our data for the Luneburg imaging system by calculating the diffraction spot size numerically and zeroing out the geometrical aberrations generated by a monocentric lens. We used ZEMAX to validate some of our results to insure the performance for rays at any field angle was equivalent to that for an on-axis ray.

We note that we did not use the Luneburg lens in a way that provided the best possible performance in the smallest possible volume. To do so, we could have modified the index of the Luneburg lens to reduce volume, but we would have been forced to use a different index profile for each f-number. This violated our basic postulate not to expend considerable effort to optimize lens performance.

References

1. A. W. Lohmann, "Scaling laws for lens systems," Appl. Opt. **28**, 4996–4998 (1989).
2. H. M. Ozaktas, H. Urey, and A. W. Lohmann, "Scaling of diffractive and refractive lenses for optical computing and interconnections," Appl. Opt. **33**, 3782–3789 (1994).
3. R. Kingslake, *A History of the Photographic Lens* (Academic Press, 1989), pp. 50–52.
4. S.-B. Rim, P. B. Catrysse, R. Dinyari, K. Huang, and P. Peumans, "The optical advantages of curved focal plane arrays," Opt. Express **16**, 4965–4971 (2008).
5. H.-C. Jin, J. R. Abelson, M. K. Erhardt, and R. G. Nuzzo, "Soft lithographic fabrication of an image sensor array on a curved substrate," J. Vac. Sci. Technol. B **22**, 2548–2551 (2004).
6. H. C. Ko, M. P. Stoykovich, J. Song, V. Malyarchuk, W. M. Choi, C.-J. Yu, J. B. Geddes III, J. Xiao, S. Wang, Y. Huang, and J. A. Rogers, "A hemispherical electronic eye camera based on

compressible silicon optoelectronics," Nature **454**, 748–753 (2008).
7. O. Iwert and B. Delabre, "The challenge of highly curved monolithic imaging detectors," Proc. SPIE **7742**, 774227 (2010).
8. http://www.darpa.mil/Our_Work/MTO/Programs/Hemispherical_Array_Detector_for_Imaging_(HARDI).aspx.
9. D. L. Marks and D. J. Brady, "Gigagon: A monocentric lens design imaging 40 gigapixels," in *Imaging Systems*, OSA Technical Digest (CD) (Optical Society of America, 2010), paper ITuC2.
10. O. Cossairt, D. Miau, and S. K. Nayar, "Gigapixel computational imaging," in *IEEE International Conference on Computational Photography* (2011).
11. *Handbook of Optical Systems Vol. 4: Survey of Optical Instruments*, H. Gross, F. Blechinger, and B. Achtner, eds. (Wiley-VCH Verlag, 2008), pp. 109–110.
12. R. K. Luneburg, *Mathematical Theory of Optics* (Brown University, 1944), pp. 189–213.
13. N. Bareket, "Second moment of the diffraction point spread function as an image quality criterion," J. Opt. Soc. Am. **69**, 1311–1312 (1979).
14. M. P. Christensen, P. Milojkovic, M. J. McFadden, and M. W. Haney, "Multiscale optical design for global chip-to-chip optical interconnections and misalignment tolerant packaging," J. Sel. Topics Quantum Electron. **9**, 548–556 (2003).
15. D. J. Brady and N. Hagen, "Multiscale lens design," Opt. Express **17**, 10659–10674 (2009).

Remote Detection of Buried Land-Mines and IEDs Using LWIR Polarimetric Imaging

Kristan P. Gurton and Melvin Felton

Optics Express 2012, 20 (20) 22344-22359

Remote detection of buried land-mines and IEDs using LWIR polarimetric imaging

Kristan P. Gurton* and Melvin Felton

U.S. Army Research Laboratory, 2800 Powder Mill Rd., Adelphi, Maryland 20783, USA
**kristan.p.gurton.civ@mail.mil*

Abstract: We report results of an ongoing study designed to assess the ability for enhanced detection of recently buried land-mines and/or improvised explosive devices (IED) devices using passive long-wave infrared (LWIR) polarimetric imaging. Polarimetric results are presented for a series of field tests conducted at various locations and soil types. Well-calibrated Stokes images, S0, S1, S2, and the degree-of-linear-polarization (DoLP) are recorded for different line-of-sight (LOS) slant paths at varying distances. Results span a three-year time period in which three different LWIR polarimetric camera systems are used. All three polarimetric imaging platforms used a spinning-achromatic-retarder (SAR) design capable of achieving high polarimetric frame rates and good radiometric throughput without the loss of spatial resolution inherent in other optical designs. Receiver-operating-characteristic (ROC) analysis and a standardized contrast parameter are used to compare detectability between conventional LWIR thermal and polarimetric imagery. Results suggest improved detectability, regardless of geographic location or soil type.

OCIS codes: (280.4991) Passive remote sensing; (040.2480) FLIR, forward-looking infrared; (110.5405) Polarimetric imaging.

References and links

1. S. Reed, Y. Petillot, and J. Bell, "Model-based approach to the detection and classification of mines in sidescan sonar," Appl. Opt. **43**(2), 237–246 (2004).
2. C. Bohling, K. Hohmann, D. Scheel, D. Nodop, C. Bauer, J. Burgmeier, W. Schade, and G. Holl, "Real-time detection of mines and explosives by laser-induced breakdown spectroscopy," in Conference on Lasers and Electro-Optics, 2006 and 2006 Quantum Electronics and Laser Science (CLEO/QELS. 2006).
3. C. Bohling, D. Scheel, K. Hohmann, W. Schade, M. Reuter, and G. Holl, "Fiber-optic laser sensor for mine detection and verification," Appl. Opt. **45**(16), 3817–3825 (2006).
4. J. A. Shaw, N. L. Seldomridge, D. L. Dunkle, P. W. Nugent, L. H. Spangler, J. J. Bromenshenk, C. B. Henderson, J. H. Churnside, and J. J. Wilson, "Polarization lidar measurements of honey bees in flight for locating land mines," Opt. Express **13**(15), 5853–5863 (2005).
5. T. H. Chua and C. L. Chen, "Fiber polarimetric stress sensors," Appl. Opt. **28**(15), 3158–3165 (1989).
6. J. S. Tyo, B. M. Ratliff, J. K. Boger, W. T. Black, D. L. Bowers, and M. P. Fetrow, "The effects of thermal equilibrium and contrast in LWIR polarimetric images," Opt. Express **15**(23), 15161–15167 (2007).
7. R. Harr and M. Polcha, "Preliminary investigation of the reststrahlen phenomenology at low-grazing angles," Proc. SPIE **5794**, 978–987 (2005).
8. Y. Wang, L. Li, and Y. Sun, "Adaptive imaging for forward-looking ground penetrating radar," IEEE Trans. Aerosp. Electron. Syst. **41**(3), 922–936 (2005).
9. J. Kositsky, R. Cosgrove, and C. Amazeen, "Results from a forward-looking GPR mine detection system," Proc. SPIE **4742**, 206–217 (2002).
10. T. W. Du Bosq, J. M. Lopez-Alonso, and G. D. Boreman, "Millimeter wave imaging system for land mine detection," Appl. Opt. **45**(22), 5686–5692 (2006).
11. K. Stone, J. Keller, K. Ho, M. Busch, and P. D. Gader, "On the registration of FLGPR and IR data for a forward-looking landmine detection system and its use in eliminating FLGPR false alarms," Proc. SPIE **6953**, 695314, 695314-12 (2008).
12. E. Winter, M. Miller, C. Simi, and A. Hill, "Mine detection experiments using hyper-spectral sensors," Proc. SPIE **5415**, 1035–1041 (2004).
13. E. M. Winter and M. S. Silvious, "Spectral method to detect surface mines," Proc. SPIE **6953**, 69530R, 69530R-9 (2008).

14. A. C. Goldberg, T. Fischer, and Z. Derzko, "Application of dual-band infrared focal plane arrays to tactical and strategic military problems," Proc. SPIE **4820**, 500–514 (2003).
15. G. Koh, E. Winter, and M. Schatten, "Rainfall degradation of LWIR disturbed soil signature," Proc. SPIE **6217**, 62170G, 62170G-8 (2006).
16. W. Wolfe and G. Zissis, *The Infrared Handbook*, Environmental Research Institute of Michigan, Office of Naval Research, Dept. of Navy, Washington, DC (1978).
17. G. Zissis, ed., *The Infrared & Electro-Optical System Handbook, Sources of Radiation*, (SPIE Optical Press, 1993), Vol. 1.
18. J. Sergio, Z. Wang, J. Tyo, and B. Hoover, "Target Detection with Partial Mueller Polarimeters," in *Frontiers in Optics*(FIOS), OSA Technical Digest (Optical Society of America, 2008), paper FThO7.
19. J. S. Tyo, B. M. Ratliff, J. K. Boger, W. T. Black, D. L. Bowers, and M. P. Fetrow, "The effects of thermal equilibrium and contrast in LWIR polarimetric images," Opt. Express **15**(23), 15161–15167 (2007).
20. K. P. Gurton and R. Dahmani, "Effect of surface roughness and complex indices of refraction on polarized thermal emission," Appl. Opt. **44**(26), 5361–5367 (2005).
21. E. Hecht and A. Zajac, *Optics* (Addison-Wesley, 1979), Vol. 22, pp. 4223–4227.
22. M. Kudenov, L. Pezzaniti, and G. Gerhart, "Microbolometer-infrared imaging Stokes polarimeter," Opt. Eng. **48**(6), 063201 (2009).
23. C. S. L. Chun, D. L. Fleming, and E. J. Torok, "Polarization sensitive, thermal imaging," in *Automatic Object Recognition IV,* F. A. Sadjadi, ed., Proc. SPIE **2234**, 275–286 (1994).
24. J. L. Pezzaniti and D. B. Chenault, "A division of aperture MWIR imaging polarimeter," Proc. SPIE **5888**, 58880V, 58880V-12 (2005).
25. J. L. Pezzaniti and R. A. Chipman, "Imaging polarimeters for optical metrology," in Polarimetry: Radar, Infrared, Visible, Ultraviolet, and X-Ray, R. A. Chipman and J. W. Morris, eds., Proc. SPIE **1317**, 280–294 (1990).
26. J. L. Pezzaniti and R. A. Chipman, "Mueller matrix imaging polarimetry," Opt. Eng. **34**(6), 1558–1568 (1995).
27. J. S. Tyo, D. L. Goldstein, D. B. Chenault, and J. A. Shaw, "Review of passive imaging polarimetry for remote sensing applications," Appl. Opt. **45**(22), 5453–5469 (2006).
28. T. Fawcett, "An introduction to ROC analysis," Pattern Recognit. Lett. **27**(8), 861–874 (2006).
29. A. P. Bradley, "The use of the area under the ROC curve in the evaluation of machine learning algorithms," Pattern Recognit. **30**(7), 1145–1159 (1997).
30. J. K. Gohagan, E. L. Spitznagel, M. M. McCrate, and T. B. Frank, "ROC analysis of mammography and palpation for breast screening," Invest. Radiol. **19**(6), 587–592 (1984).
31. W. Dillon and M. Goldstein, *Multivariate Analysis Methods and Applications* (John Wiley & Sons, 1984).
32. M. Felton, K. P. Gurton, J. L. Pezzaniti, D. B. Chenault, and L. E. Roth, "Measured comparison of the crossover periods for mid- and long-wave IR (MWIR and LWIR) polarimetric and conventional thermal imagery," Opt. Express **18**(15), 15704–15713 (2010).

1. Introduction

Both military and civilian personnel are facing an ever-evolving threat from buried/concealed landmines and improvised explosive devices (IEDs). There has been significant research dedicated to the detection of buried explosive devices [1–6]. One particular technology that has shown promise is forward-looking ground penetrating radar (FLGPR) [7–9]. However, recent studies have identified several inherent problems associated with FLGPR, e.g., buried explosives that are formed from dielectric- or polymer-based materials (plastics) are difficult to detect due to the small electromagnetic (EM) radar cross-sections for non-conducting materials [10,11]. In addition, FLGPR systems are plagued by unacceptable false-alarm rates due to the detection of commonly buried debris. A consensus has emerged that two or more complimentary technologies will most likely be required to improve detectability while reducing false-alarm rates. One such complimentary approach may involve a combination of a FLGPR system with an optically based imaging platform capable of detecting surface anomalies, i.e., disturbed earth (DE), that result when explosive devices are buried/concealed near the surface of a given terrain.

One suggested imaging technique for the remote detection of DE involves various forms of spectroscopic imaging in the thermal infrared (IR), sometimes termed multi- or hyper-spectral imaging [12–14]. These techniques attempt to exploit the so-called "reststrahlen" effect, in which the bulk emissivity for a particular soil changes within 8–10μm spectral range due to absorption at the reststrahlen frequencies, which are approximately equal to the natural frequencies of certain crystalline structure associated with small semi-transparent silica-based particles [15–17] Although well-documented, the reststrahlen effect has been shown to be

quite variable depending on geographic location and soil composition. Most research shows a 3–5% variance in the IR emissivity associated to the reststrahlen phenomena under optimal conditions.

We consider a new imaging approach based on changes in polarization state associated with radiation that is emitted and/or reflected from a surface that has recently been altered [18–20]. The premise for considering polarimetric imaging is based on the fact that both manmade and naturally occurring terrain establishes an "average" polarization profile or pattern that results from vehicle traffic, weathering, or just the passage of time. Since the polarization state of the image forming radiation is extremely sensitive to subtle changes in the geometry of reflecting/emitting surface, resultant differences in polarization signatures arise for localized surface regions that have recently been disturbed.

For our application, we chose to use a Stokes parameter approach to describe the polarization state of the radiation that is emitted and/or reflected from a target area [21]. We apply the Stokes methodology to an imaging application where we define the Stokes "images" S1, S2, and S0, by the usual convention shown in Eqs. (1)–(3),

$$\mathrm{S1} = \mathrm{I}(0) - \mathrm{I}(90)\ \left(\mathrm{w/sr-m^2}\right), \tag{1}$$

$$\mathrm{S2} = \mathrm{I}(+45) - \mathrm{I}(-45)\ \left(\mathrm{w/sr-m^2}\right). \tag{2}$$

For total linear polarization, the total radiance image, S0, is defined as,

$$\mathrm{S0} = \mathrm{I}(0) + \mathrm{I}(90)\ \text{total radiance}\ \left(\mathrm{w/sr-m^2}\right), \tag{3}$$

and the degree-of-linear polarization (DoLP) image is expressed as,

$$\mathrm{DoLP} = \frac{\sqrt{S1^2 + S2^2}}{S0}, \tag{4}$$

where I(0), I(90), I(+ 45), and I(–45) represent well-registered (spatially) images produced with polarimetrically filtered radiance (w/sr-m^2) at orientation angles 0°, 90°, + 45°, and –45°, respectively, where 0° is defined as the vertical with respect to the image plane. As one can see from Eqs. (1)–(3), the S1 image represents a relative measure of the vertical compared to the horizontal component, the S2 image represents a relative measure of the difference between the two ± 45° diagonal states, and the S0 image is merely a conventional "intensity only" image.

This study represents a compilation of results spanning a three-year period. The three field-tests presented here were conducted in 2008, 2009, and 2011, to assess the viability for using passive LWIR (8-12μm) polarimetric imaging to identify regions of recently DE. The primary goal was threefold—1) objectively measure the ability to detect regions of disturbed soil associated with the placement of buried land-mines and/or IEDs; 2) assess how detectability is effected varying soil type and composition; and 3) determine how LWIR polarimetric DE signatures are effected by atmospheric conditions, e.g., clear sky, cloud cover, rain, wind, etc.

During the time period that this report encompasses, three different LWIR polarimetric imaging platforms were used, and all were produced by Polaris Sensor Technologies, Inc., located in Huntsville, AL. It should be noted that the use of three different LWIR polarimetric imagers was dictated not by choice, but rather by necessity. After an initial proof-of-concept study in 2008, a first-generation LWIR polarimetric imager experienced technical issues that delayed further study. Early in 2009, a LWIR microbolometer-based polarimetric sensor became available, which allowed for our work to continue. Finally, in 2010, a state-of-the-art LWIR polarimetric imager became available and was used in the final phase of the study.

Although there are a variety of optical configurations appropriate for polarimetric imaging, (e.g., division-of-amplitude (DoA), division-of-focal-plane (DoFP), and division-of-

aperture (DoAP)), we chose a division-of-time (DoT) approach based on a spinning achromatic retarder (SAR) design for recording calibrated LWIR Stokes imagery [22–26]. Because the DoT method relies on the capture and differencing of sequentially recorded images, it is only appropriate for imaging objects that are slowly moving or static within the scene. Although somewhat limited by the sequential nature of the recorded imagery, it is by far the best choice for basic research applications due to maximum radiometric throughput, spatial resolution, and polarimetric sensitivity.

2. Spinning achromatic retarder (SAR) polarimetric sensors

A spinning achromatic retarder (SAR) imaging polarimeter operates by capturing a sequence of images in time. Each image in the sequence is recorded at a different orientation position of a spinning achromatic retarder. In its principle mode of operation, the system acquires a set of 16 images per rotation of the retarder—i.e., images are captured at 0, 22.5, 45… to 337.5°.

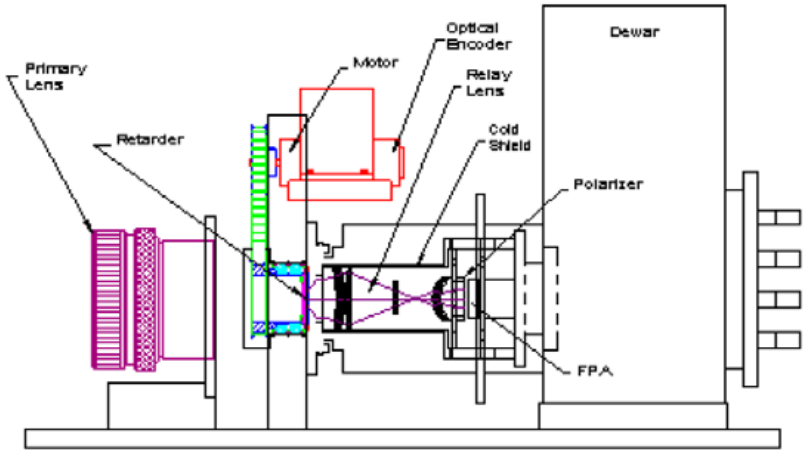


Fig. 1. Basic design of a spinning achromatic retarder (SAR) LWIR polarimetric imager.

Figure 1 shows the basic design of a LWIR SAR-based imaging polarimeter in which either a room-temperature microbolometer, or a cryogenically cooled Mercury Cadmium Telluride (MCT) focal-plane-array (FPA) detector is positioned at the image-plane of the sensor. In general, we have found the cooled MCT-based FPAs to exhibit a noise-equivalent DoLP or NEDoLP (similar to NEΔT for conventional thermal) on the order of ± 0.1%, whereas the microbolometer-based systems typically exhibit NEDoLP values in the range of ± 0.3-0.5%.

For MCT-based systems, much of the optical train is held under vacuum within a Dewar and cooled to an approximate temperature of 88K. The achromatic retarder is mounted just outside the Dewar window and is mounted to a precision set of frictionless bearings. A series of relay optics are used to reduce any beam wander generated by the rotating optic, and by using this configuration pixel, registration error between sequential frames is typically less than 1/20th of a pixel. The retarder is rotated continuously by a stepper motor at variable rates, depending on the specified integration period and application. All three SAR-based systems used in this study are designed to record/display a user-specified set of Stokes and/or polarimetric image products at processing rates approaching real-time. An excellent review of current polarimetric imaging technologies can be found in Tyo, Goldstein, Chenault, and Shaw [27].

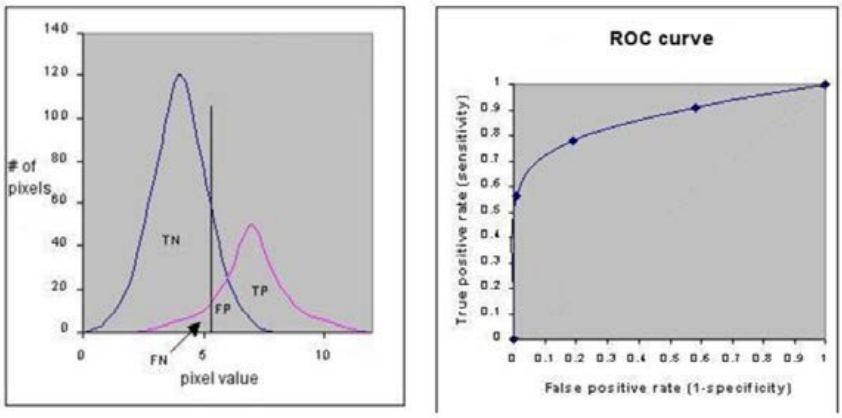
3. Detection analysis

Perhaps one of the more difficult tasks involved with image detection analysis is developing an objective evaluation metric that consistently and correctly identifies the "best" image type for maximum detectability. Much of the difficulty arises from the fact that optimum detectability is so heavily dependent on the type of end-user one considers, e.g., human or computer algorithm. To address this, we chose to use two established image evaluation metrics—i.e., detection calculations based on a receiver-operational-characteristic (ROC) curve approach, and the more intuitive, standardized contrast parameter method. Since both techniques have inherent strengths and weaknesses, we present to the reader actual LWIR thermal and polarimetric image sets for subjective, yet sometimes more informative, evaluation.

The receiver-operational-characteristic (ROC) analysis is often the tool of choice among researchers within the artificial intelligence (Ai) and automated-target-recognition (ATR) community. The ROC method was originally developed for signal detection analysis but is now widely applied in many different disciplines [28–30]. ROC curve analysis is used to compare target detectability between different image sets recorded or processed by different means. In order to do this, some *a priori* knowledge about the location of the actual target is necessary so that a "truth" image can be generated. Figure 2(a) shows an image histogram of an example truth image where the large Gaussian like curve on the left represents the pixel values associated with the background, and the smaller distribution on the right represents the pixel values associated with the target. The vertical line in the figure represents an arbitrary threshold point. Also shown are regions defined by the intersection of the two histograms, as well as regions to the right and left of the threshold line identified as true-negative (TN), true-positive (TP), false-positive (FP), and false-negative (FN) regions.

A ROC curve is generated by comparing the overlapping regions TN, TP, FP, and FN, as the threshold point is swept right to left across the histogram. Figure 2(b) shows a resultant ROC curve for the histograms shown in Fig. 2(a). The area under the ROC curve is defined as the normalized probability for detection of the target identified in the truth image.

One inherent weakness associated with the ROC curve approach stems from the fact that it is a purely statistical method and fails to take into account addition spatial information associated with localized target pixel location and/or clustering—an important aspect of visual cognitive detection. The human eye can often decipherer target regions within a scene based on very subtle variations among clusters of pixels that form a particular distinguishable shape. Nevertheless, the ROC method is a readily accepted metric among Ai/ATR community and does offer an objective measure of target detectability.



(a) (b)

Fig. 2. 2(a) A image histogram where the target (right Gaussian) and background (left Gaussian) regions are defined and, 2(b) the corresponding ROC curve, where the area under the curve is related to the probability for detection the target within the scene.

A second evaluation method for grading imagery for maximum detectability involves calculating a "standardized" contrast parameter [31]. At the most fundamental level, the ability to detect a given object within an image is heavily dependent on the magnitude of the difference between pixel values associated with the object and its associated background, i.e., contrast. However, in order to compare pixel values that result from different image types—e.g., thermal, Stokes, DoLP, etc.—a standardization process must be applied to the entire image set. This is a common procedure used to normalize multivariate image sets before applying a particular evaluation metric. The standardization process effectively translates each image histogram (derived from different physical quantities) onto the same basis set of coordinate axes. The standardization procedure involves subtracting the mean pixel value derived by the entire image, and dividing the resultant histogram by the standard deviation. This multivariate normalization process by no means affects overall integrity and information content of the image. After image standardization, separate ROIs are defined for the target and background regions for a given image set, and the average pixel value for each region is computed. Finally, a standardized contrast parameter is calculated for each image and is defined as,

$$C^{o}\left|\bar{u}_T - \bar{u}_B\right|, \tag{5}$$

where $\bar{u}_T$ and $\bar{u}_B$ represent the average pixel values for the target and background ROIs, respectively.

4. Experiment (test sites) and results

The first test was conducted on May 22, 2008, and was located at the U.S. Army Research Laboratory (ARL), Adelphi, MD, on a test surface best described as a well-traveled dirt road consisting of a gravel-clay-soil mixture that was well-compacted. The test was conducted over a 6-h period during mid-afternoon under clear skies, with relative humidity approximately 50% and temperatures varying from 77 to 81 °F. Holes were dug approximately 12 in into the hardened road surface and surrogate IED targets were buried at different locations, see Fig. 3.

For this particular test, we used our lowest resolution 256 x 256 MCT FPA-based SAR polarimetric camera system. The polarimetric imager was mounted on a tripod and positioned 2.75 m above the ground and was focused on the DE region at an approximate distance of 10 m away as shown in Fig. 4. A 50 mm LWIR objective lens was fitted to the polarimetric

sensor, which produced an effective field-of-view (FOV) of 15°. The camera LOS was angled to the DE region, resulting in a range of grazing angles from 15 to 20° defined by the LOS and the road surface. It should be noted that for this first proof-of-concept test, great care was taken to camouflage the disturbed region as best as possible, i.e., not readily noticeable to a casual observer, see Fig. 3(b). After the disturbed regions reached thermal equilibrium after approximately 1 h, a series of four image sets were recorded at 15-min intervals.

Fig. 3. Photograph of DE region May 22, 2008 test conducted at U.S. Army Research Laboratory, Adelphi, MD site. 3(a) DE test-bed (red oval area represents buried target and 3(b) a close up of the DE region (clay-gravel-soil mixture).

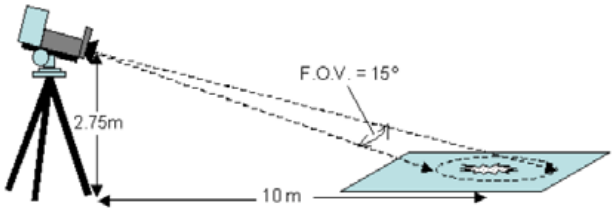


Fig. 4. Schematic for the initial test conducted on May 22, 2008 showing the positioning of the LWIR polarimetric camera with respect to the DE test region.

Figure 5 shows the resultant imagery consisting of a conventional LWIR thermal image, S0, the two Stokes images, S1 and S2, and DoLP image, where a false color has been applied to all the original grey-scale images. Note that all Stokes image values presented here are normalized with respect S0, and range from –1 to 1. Table 1 shows the average absolute radiance and normalized Stokes values for ROIs that are defined as either the DE or background regions.

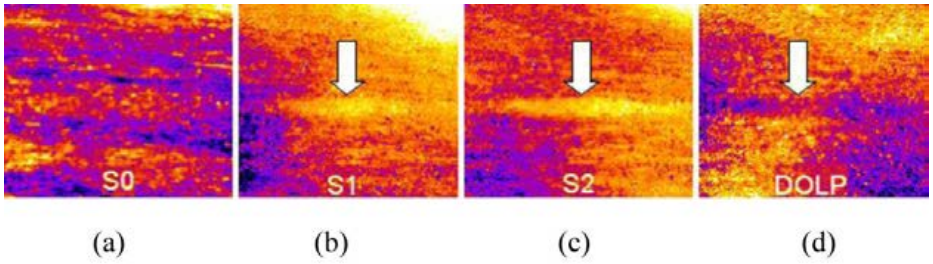


Fig. 5. (a) conventional LWIR thermal image, S0, for the DE region highlighted in Fig. 3 recorded on May 22, 2008. Figures 5(b)-(d) show the resultant Stokes images S1, S2, and degree-of-linear-polarization (DoLP) image where the DE region is shown by an identifying arrow.

As one can see in Fig. 5(a), the ability to distinguish disturbed from undisturbed soil regions is quite poor for the conventional LWIR thermal image, S0, and reflected by the lowest calculated ROC curve value, 0.256 (Fig. 6), and the lowest contrast parameter value, 0.056, shown in Table 2. The DE region becomes visible in the Stokes image S1 (Fig. 5b), where contrast arises from the fact that the DE region emits thermal radiation that is slightly less polarized, when compared to the surrounding undisturbed area (Table 1). Note that since all normalized S1 values are negative, the majority of the polarization lies in the horizontal plane, based on the definitions shown in Eqs. (1-2), which is associated with "emission" dominant polarization. Conversely, a positive S1 value implies that the vertical component is dominant, and the majority of the received radiance is due to "reflection" of the ambient optical background.

Table 1. Average radiant and polarimetric values for DE and background ROI regions for images shown in Figs. 5(a)-5(d), recorded on May 22, 2008.

May 22, 2008 Test	S0 (watt/sr-m^2)	S1/S0	S2/S0	DoLP(%)
DE region (ROI average)	27.36	0.041	–0.081	2.843
background (ROI average)	27.19	0.048	–0.058	3.044

Similar evaluation of the S2 image shows further improvement in target detectability and is reflected by the highest calculated ROC curve and contrast parameter values of 0.958 and 1.745, respectively. Again, since the values for the normalized S2 image shown in Table 1 are negative, the dominant polarization state is oriented at –45°, with respect to the vertical. In a scene in which the LOS to the lay of the surface is perfectly symmetric, we would expect the values of the S2 image to be nearly zero—i.e., ground is surface flat and level, and the region of interest is centered. However, due to the slope of the ground surface and the fact that the camera mount was off-center with respect to the DE region, a larger than normal difference arose between the + 45° and –45° states. Figure 5(d) shows the DoLP image, which is merely the normalized superposition of the Stokes images S1 and S2. A lower contrast parameter value of 0.993 is not unexpected since the DoLP product contains noise components from S1, S2, and S0, which in this case, results in a slight reduction in the overall contrast for the DoLP product image.

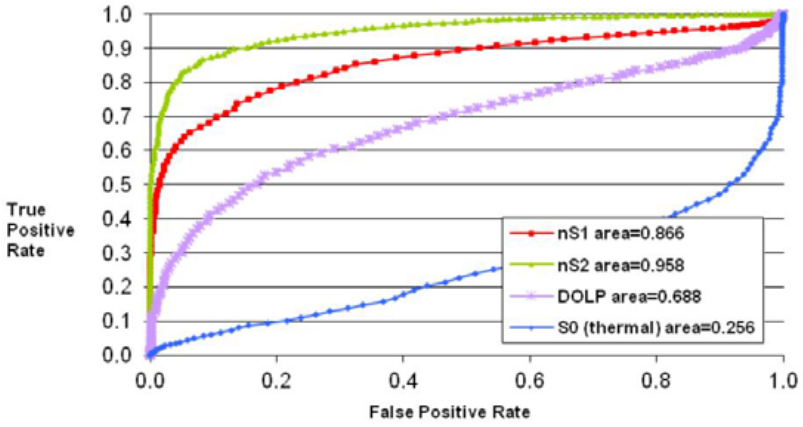


Fig. 6. Corresponding ROC curves calculated for images shown in Figs. 5(a)–5(d), in which the probability of detection is defined as the integrated area under each respective curve.

Table 2. Comparison of the contrast parameter and ROC curve results for the image set shown in Figs. 5(a)-5(d), recorded on May 22, 2008.

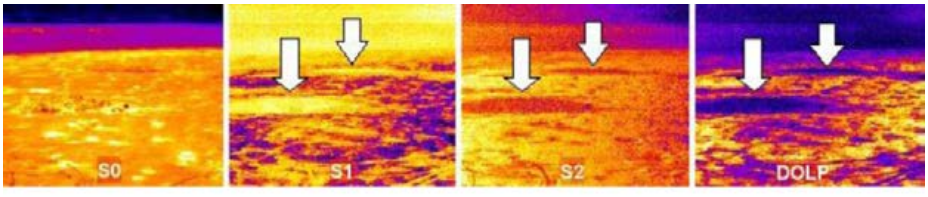
May 22, 2008 Test	S0	S1	S2	DoLP
contrast parameter	0.056	0.790	1.745	0.993
probability of detection	0.256	0.866	0.958	0.688

A follow-on series of DE tests were conducted over a multi-day period from August 27–September 4, 2009. The location was again the Adelphi, MD, area, which included the original May 22, 2008, dirt road site, as well as two new locations in which the soil compositions for each location is characterized as red-clay-silt mixture, and a topsoil type material, rich in organic material and small stone. As previously mentioned, the original liquid nitrogen (LN2)-cooled 256 x 256 MCT FPA SAR polarimetric sensor was unavailable during this period and a new 324 x 256 FPA microbolometer-based SAR polarimetric imager was substituted in its place.

The first test was conducted on August 27, 2009, at a local baseball field in Adelphi, MD. The site was chosen due to its unique soil type similar to what is found in various regions of Southeast Asia, see Fig. 7(a). Once again, holes were dug and surrogate objects were buried at various locations at depths < 1 m. Similar to the first test conducted in 2008, each hole was carefully raked and brushed over to camouflage the fact that digging had occurred, see Fig. 7(b). The microbolometer SAR polarimetric imager was set up in a similar manner as in prior tests, with the exception that the camera was mounted at a slightly lower to a height of 2 m above the ground. This resulted in a viewing angle (as defined by the LOS and the soil surface) that ranged from 10 to 16°, depending on the location of interest within the scene. Once the disturbed regions reached thermal equilibrium with the surrounding background, capture of polarimetric imagery began and was recorded every 15 min for approximately 4 h.

(a) (b)

Fig. 7. Test area recorded on August 27, 2009. 7(a) The packed red-clay-silt soil field used to generate the DE regions, and 7(d) the subsequent camouflage and raking of the area.



(a) (b) (c) (d)

Fig. 8. The polarimetric image set for August 27, 2009 test where the DE regions are shown by the identifying arrows. (a) Conventional LWIR thermal image S0; (b) Stokes image S1; (c) Stokes image S2; and (d) DoLP image, recorded with the microbolometer based SAR polarimetric sensor.

Figure 8 shows the resultant thermal and polarimetric image set recorded at the red-clay-silt site on August 27, 2009. Two large DE regions are visible only in the Stokes and DoLP images shown in Figs. 8(b)-8(d) and are identifiable as either light regions in image, 8(b), or dark regions in images 8(c) and 8(d). Average thermal and polarimetric values for the disturbed and background ROIs are shown in Table 3.

Review of the values shown in Table 3 again shows negative values for the normalized S1 image, which implies that the polarization state results primarily from surface emission, rather than reflection of the ambient background radiation. Because the camera was well-centered with respect to the disturbed regions, greater symmetry was created within the scene, resulting in normalized S2 values that were effectively zero. As with the earlier test, the disturbed region was less polarized than the undisturbed background, resulting in DoLP values of 1.00% and 1.43%, respectively.

The results of the ROC curve and contrast parameter analysis are shown Table 4 and indicate that the DoLP image is ranked highest in detectability, followed closely by the S1 image. However, issues associated with the ROC curve approach become apparent when considering the S2 image, which registered the lowest probability of detection with a value of 0.312, although the DE regions are clearly visible in the S2 image, shown Fig. 8(c).

Table 3. Average radiant and polarimetric values for the DE and background regions shown in Figs. 8(a)-8(b) for the red-clay-silt test site recorded on August 27, 2009.

Aug. 27, 2009 Test	S0 (watt/sr-m^2)	S1/S0	S2/S0	DoLP(%)
DE region (ROI average)	16.41	–0.001	–0.002	1.0
background (ROI average)	15.46	–0.014	0.004	1.4

Table 4. Comparison of the contrast parameter and ROC curve results for the red-clay-soil surfaces shown in Fig. 7(a)-7(d), recorded on August 27, 2009.

Aug. 27, 2009 Test	S0	S1	S2	DoLP
contrast parameter	0.023	1.438	1.022	1.549
probability of detection	0.444	0.478	0.312	0.795

On September 3, 2009, we returned to the first test-site location of May 22, 2008, and repeated the measurement with the microbolometer-based SAR polarimetric sensor. This time, two surrogate targets were buried, and after thermal equilibrium between the DE and backgrounds surfaces was reached, polarimetric images sets were recorded, see Fig. 9. Note that the bright spots seen in the S0 image, Fig. 9(a), was a result of localized heating due to direct sunlight that was filtered through a series of trees located on the right side of the test site.

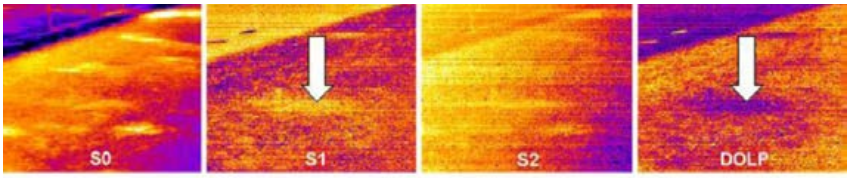


Fig. 9. Radiant and polarimetric imagery recorded on Sept. 3, 2009 at the original May 22, 2008 site. 9(a) Conventional thermal LWIR image, S0, 9(b) Stokes image, S1, 9(c) Stokes image, S2, and 9(d) DoLP image, recorded using a microbolometer based SAR imaging polarimeter.

On September 4, 2009, we chose a test location containing a topsoil mixture consisting of fine dirt and gravel that was "shadowed" from the afternoon sun by a large building, see Fig. 10(a). This location was specifically chosen to assess how shadowing, as well as radiant loading resulting from the building, would affect the ability to polarimetrically resolve regions of DE. Only one surrogate target was buried for the test, and after a period of about 2 h (needed for thermal equilibrium to occur), recording of polarimetric imagery was started, see Fig. 10.

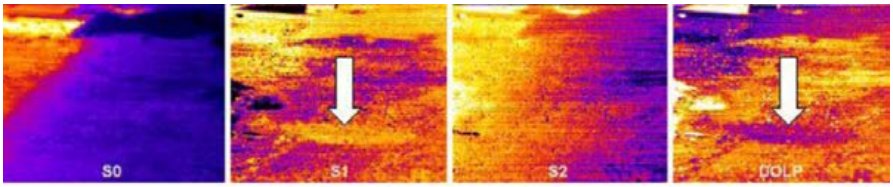


Fig. 10. Radiant and polarimetric imagery recorded on Sept. 4, 2009 showing the effect of shadowing and ambient radiant loading on a DE region. Target site consisted of a topsoil mixture of fine dirt and gravel. (a) Conventional thermal LWIR image S0; (b) Stokes image S1; (c) Stokes image S2; and (d) DoLP image, recorded using the microbolometer based SAR imaging polarimeter.

Table 5. Average radiant and polarimetric values for DE and background ROIs recorded on September 3 - 4, 2009 for the image sets shown in Figs. 9 and 10 using the LWIR microbolometer SAR polarimetric sensor.

Sept. 3, 2009 Test	S0 (watt/sr-m^2)	S1/S0	S2/S0	DoLP(%)
DE region (ROI average)	13.13	–0.010	–0.003	1.1
background (ROI average)	13.13	–0.016	–0.006	1.7
Sept. 4, 2009 Test	**S0 (watt/sr-m^2)**	**S1/S0**	**S2/S0**	**DoLP(%)**
DE region (ROI average)	13.61	–0.005	0.001	0.6
background (ROI average)	13.26	–0.009	–0.006	1.1

Until now, all of the studies have been short in duration—i.e., a single day—and the obvious question is, "How long and under what conditions will such disturbances continue to be detectable using a polarimetric imaging method?"

The resultant radiance and polarimetric values for the DE and background ROIs, as well as the corresponding detection metrics for the September 3–4, 2009, tests are shown in Tables

5 and 6. Review of the contrast parameter and ROC curve results shows that the degree of detectability for the polarimetric derived imagery is consistently greater than the conventional thermal image, S0. However, there is disagreement on whether the S1 or DoLP image is actually superior, since the September 3 contrast parameter implies the DoLP image to be best, while the ROC curve results for both September 3 and 4 show the S1 to be the superior detection image. Again, the dilemma of what is the "best" image is left to the observer, and after comparing the image sets shown in Figs. 9 and 10, one could make a good argument for either the S1 or DoLP for being the superior detection image.

To address this issue partially, we conducted a multi-day field test that occurred over a five-day period spanning October 1–5, 2011. The term "partially" is used because even after a week in the field, in which we experienced a variety of weathering conditions including a series of modest rain events and monsoon type winds, many of the DE regions were still visible to the polarimetric sensor.

Table 6. Comparison of the contrast parameter and ROC curve results for the DE tests recorded on September 3 - 4, 2009 for the image sets shown in Figs. 9 and 10 using the LWIR microbolometer SAR polarimetric sensor.

Sept. 3, 2009 Test	S0	S1	S2	DoLP
contrast parameter	0.082	0.775	0.744	0.913
probability of detection	0.394	0.754	0.522	0.248
Sept. 4, 2009 Test	**S0**	**S1**	**S2**	**DoLP**
contrast parameter	0.0210	1.569	0.594	1.473
probability of detection	0.399	0.613	0.678	0.397

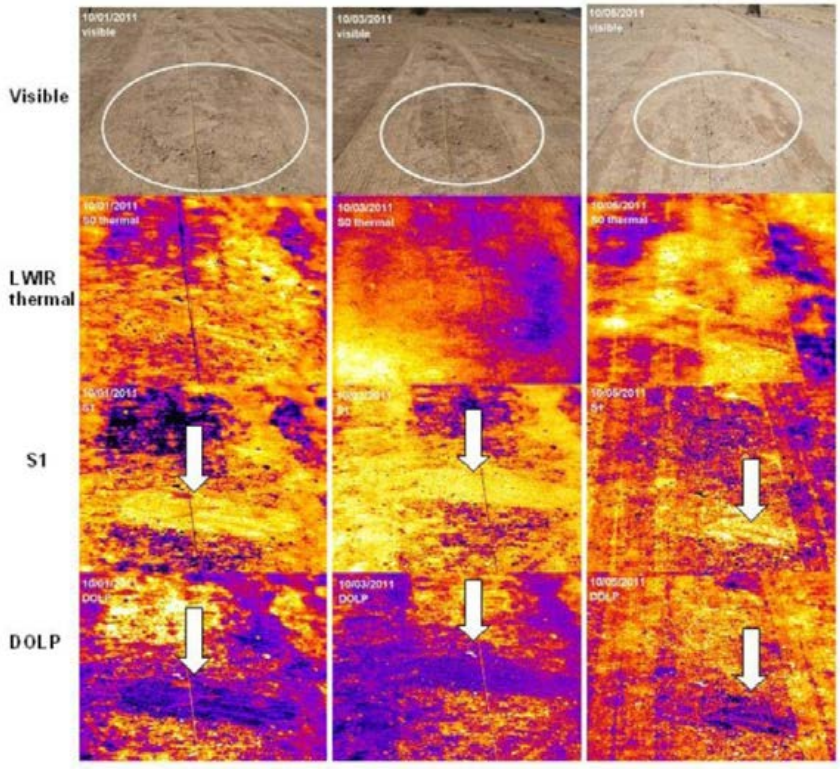
This final test was conducted in an arid southwest region of the United States located at the Energetic Materials Research and Testing Center (EMRTC) in Socorro, NM, where the soil type is characterized as loam—i.e., a soil composed of sand, silt, and clay at about 40-40-20% concentrations respectively. For this particular study, our most sensitive LWIR MCT-based polarimetric sensor became available. The 640 LWIR SAR Polarimetric imager, produced by Polaris Sensor Technologies, housed a Stirling-cooled 640x480 MCT FPA detector, which was provided by DRS Technologies, Inc. The system was designed for maximum radiometric throughput and sensitivity that results from efficient sensor design, and a usually wide spectral response of 7.5 to 11.1μm, for the FPA.

The final test was held in conjunction with another experiment in which a vehicle-mounted, forward-looking-ground-penetrating-radar (FLGPR) was being evaluated. In order to get quasi-registered LWIR polarimetric imagery with the FLGPR system, the polarimetric sensor was mounted on an elevated platform located above the FLGPR transmitter/receiver. This resulted in the sensor being 4.5 m above the ground, as shown in Fig. 11(a). The polarimetric sensor-FLGPR platform was tilted downward at an angle of 24° with respect to the LOS and test surface. A variety of surrogate objects were buried and camouflaged by hand in one of three test lanes, see Figs. 11(b) and 11(c). A series of five DE regions were imaged every other day during the period of October 1–5, 2011, in which meteorological conditions varied greatly. Weather conditions for each of the three days were as follows: October 1—clear sky, low relative humidity with a temperatures range of 69-78 °F; October 3—overcast with, low relative humidity, temperatures slightly cooler in the range of 65-73°F; and October 5—ground surface damp as a result of a rain storm that occurred during the prior day and much cooler, with a temperature range of 50-60°F.

The effects of weathering are displayed in a series of images shown in Figs. 12(a)-12(c). Each column displays a visible image, a conventional LWIR thermal image, S0, the Stokes image, S1, and a DoLP image for the DE target 2 (TG 2). Tabulated radiant and polarimetric values for all DE targets (TG) 1-5, recorded from October 1–5, 2011, as well as their corresponding background (BG) regions, are shown in Table 7. A similar list of tabulated ROC curve and contrast parameter values for the same period is shown in Table 8.

(a) (b) (c)

Fig. 11. (a) The FLGPR vehicle platform in which the 640x480 SAR polarimetric sensor was mounted and positioned at an angle of 24 degrees with respect to the LOS and the target surface. (b) Shows a typical surrogate IED being buried in the arid desert soil. (c) A typical DE region after burial.



(a) Oct. 1, 2011 (b) Oct. 3, 2011 (c) Oct. 5, 2011

Fig. 12. Typical evolution of visible, thermal, and polarimetric signatures for a DE region over the five day period from Oct. 1-5, 2011, where the DE region is identified by the identifying arrow. (a) Target 2 recorded on Oct. 1, 2011 under a clear sky condition. (b) Target 2 recorded on Oct. 3, 2011, during dense overcast conditions. (c) Target 2 recorded on Oct. 5, 2011, under light cloud cover and after a moderate rain event that occurred on Oct. 4, 20011.

The overall trend seen in the 2009 and 2010 studies continues, and is reflected in the values shown in Table 7. Specifically, 1) the polarization state for the S1 images continue to result from emission dominant radiance; 2) the DoLP for DE regions is always less than the undisturbed background surfaces; and 3) the highest ranked image type for detectability continued to be the S1 image, followed closely by the DoLP images. The values for the normalized S1 images record on October 3 (dense cloud cover) were, on average, slightly lower than S1values recorded on either October 1 or October 5.

Table 7. Radiance and polarimetric values for DE targets (TG)1-5 and corresponding background (BG) regions.

10/1/11	TG 1	BG 1	TG 2	BG 2	TG 3	BG 3	TG 4	BG 4	TG 5	BG 5
S0 (W/cm^2-sr)	35.84	35.29	36.64	35.75	35.64	34.63	37.66	37.19	36.86	36.34
S1/S0	–0.007	–0.013	–0.008	–0.012	–0.007	–0.017	–0.008	–0.012	–0.007	–0.011
S2/S0	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000
DoLP(%)	0.67	1.35	0.81	1.15	0.71	1.77	0.84	1.25	0.71	1.11
10/3/11	**TG 1**	**BG 1**	**TG 2**	**BG 2**	**TG 3**	**BG 3**	**TG 4**	**BG 4**	**TG 5**	**BG 5**
S0 (W/cm^2-sr)	22.62	22.16	23.73	23.56	22.78	22.22	23.53	23.01	22.76	22.28
S1/S0	–0.005	–0.008	–0.005	–0.009	–0.006	–0.010	–0.006	–0.009	–0.005	–0.008
S2/S0	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000
DoLP(%)	0.56	0.83	0.55	0.95	0.57	1.07	0.57	0.86	0.52	0.83
10/5/11	**TG 1**	**BG 1**	**TG 2**	**BG 2**	**TG 3**	**BG 3**	**TG 4**	**BG 4**	**TG 5**	**BG 5**
S0 (W/cm^2-sr)	23.42	23.34	22.69	23.72	21.56	21.96	21.86	22.34	22.27	22.94
S1/S0	–0.014	–0.018	–0.013	–0.018	–0.014	–0.023	–0.015	–0.018	–0.014	–0.021
S2/S0	0.001	0.001	0.001	0.002	0.002	0.003	0.002	0.002	0.001	0.002
DoLP(%)	1.45	1.86	1.38	1.87	1.47	2.38	1.52	1.87	1.44	2.16

However, based on our prior experience, we would have expected even lower S1 values for the October 3 test data, considering the overcast conditions, which is often associated with large ambient radiant loading. This is because the linear polarization that occurs when ambient radiance is reflected from a surface is always orthogonal to the linear polarization that results purely from emission, and the net effect results in an overall reduction in the total linear polarization exhibited by the surface. This effect is most often seen in MidIR

polarimetry, where it is not uncommon to see the sign of various regions within an S1 image flip with quickly changing meteorological conditions [32].

Table 8. ROC curve and contrast parameter (CP) values for DE regions 1-5 recorded on Oct. 1-5, 2011.

10/1/2011	1 ROC	1 CP	2 ROC	2 CP	3 ROC	3 CP	4 ROC	4 CP	5 ROC	5 CP
S0	0.633	0.843	0.53	0.179	0.587	1.301	0.655	0.831	0.646	0.211
S1	0.828	0.889	0.746	0.641	0.638	1.067	0.865	1.250	0.874	0.881
S2	0.494	0.045	0.48	0.119	0.476	0.248	0.399	0.307	0.431	0.232
DoLP	0.201	0.867	0.296	0.621	0.418	1.080	0.167	1.227	0.114	0.876
10/3/2011	**1 ROC**	**1 CP**	**2 ROC**	**2 CP**	**3 ROC**	**3 CP**	**4 ROC**	**4 CP**	**5 ROC**	**5 CP**
S0	0.718	0.904	0.801	1.302	0.742	0.799	0.718	0.707	0.493	0.195
S1	0.808	0.952	0.853	1.682	0.73	1.543	0.808	1.610	0.791	1.065
S2	0.495	0.013	0.439	0.472	0.551	0.050	0.495	0.209	0.543	0.102
DoLP	0.278	0.786	0.198	1.655	0.419	1.501	0.278	1.559	0.582	1.045
10/5/2011	**1 ROC**	**1 CP**	**2 ROC**	**2 CP**	**3 ROC**	**3 CP**	**4 ROC**	**4 CP**	**5 ROC**	**5 CP**
S0	0.541	0.459	0.452	0.304	0.504	0.556	0.541	0.713	0.662	0.435
S1	0.896	1.255	0.754	0.583	0.693	0.880	0.896	2.518	0.851	1.162
S2	0.489	0.311	0.53	0.066	0.519	0.045	0.489	0.427	0.432	0.331
DoLP	0.162	1.236	0.303	0.576	0.408	0.844	0.162	2.450	0.515	1.166

Perhaps the most interesting aspect of the study resulted after the rain event that occurred on October 4. Although not readily apparent in the image set shown in 12c, the modest rainfall of October 4 appeared to actually improve the contrast between the DE and surrounding undisturbed areas. We have always expected that after a sufficient amount of weathering and/or traffic has occurred, the ability to polarimetrically detect regions of recently disturbed soil would not be possible. However, based on these preliminary results, modest weathering events like blowing wind and rain may actually serve to enhance the effect, at least initially. In this particular case, the rain events appear to produce a net "smoothing" of the undisturbed regions, while the DE regions were far less affected. This is also reflected by the relatively large DoLP values recorded on October 5 for both the target (TG) and background (BG) surfaces shown in Table 7.

5. Conclusion

We have shown that by using passive LWIR polarimetric imaging, one can improve the ability to remotely detect localized regions of recently DE that is often associated with the

bearing of landmines and IEDs. The results stem from a series of tests conducted at a variety of different geographic locations with varying soil types in which three different LWIR polarimetric imaging platforms were used. In light of the multitude of changing parameters and sensors, the final results were surprisingly consistent.

First, based on objective detection metrics used in this study—i.e., ROC curve analysis and standardized contrast parameter calculations—the ability to detect localized regions of DE was greatest for the polarimetric images S1 and DoLP. We believe the DE contrast seen in the Stokes images S1 is a direct result of the symmetric slant-path imaging arrangement used in the study. We expect for an elevated nadir type detection arrangement that the DoLP image would be the image type of choice for detecting regions of DE.

Second, all measured S1 signatures were due to emission induced polarization—i.e., linearly polarized in the horizontal plane. As mentioned earlier, situations in which the ambient optical background is changing, as is the case when stratus or nimbostratus cloud-cover is present, the sign of specific regions in the Stokes images S1 or S2 are often observed changing, signifying that the mechanism for generating linear polarization has switched from being emission to reflection dominant, or vice versa. However, observations also show that even during these events, the polarimetric contrast between a given target and the corresponding background is preserved [32]

Finally, the polarimetric contrast necessary to distinguish DE regions from the undisturbed surrounding areas results from the fact that undisturbed surfaces tend to exhibit higher degrees of linear polarization compared to DE areas. Put more generally, polarimetric contrast between disturbed and undisturbed surface regions arises when symmetry of the surface is altered. Such symmetry may result from naturally occurring events, e.g., prolong wind and rain storms, or by a manmade process associated with vehicular and pedestrian travel. Once a soil surface is altered, very subtle, yet quite measurable, differences in the polarization state of the reflected or emitted radiation occurs at the boundary that defines disturbed from undisturbed surface areas.

We have shown for the cases considered here a net reduction in the linear polarization for the DE regions (relative to the surround undisturbed regions) on the order of 20-100% or more. Although all of the imagery recorded involved a slant-path LOS, the authors believe the ability to detect regions of DE using passive LWIR polarimetric imaging would greatly benefit if conducted from an aerial platform. This would allow for imaging of much larger surface areas in which a “change-detection” method could be applied. We are currently planning a series of such studies using a nadir LOS where DE regions are polarimetrically imaged in both the MidIR and LWIR to assess the benefit for using a dual-band approach.

Design of 220 GHz Electronically Scanned Reflectarrays for Confocal Imaging Systems

Abigail S. Hedden, Charles R. Dietlein and David A. Wikner

Optical Engineering 2012, 51 (9), 10. [DOI: 10.1117/1.OE.51.9.091611]

Design of 220 GHz electronically scanned reflectarrays for confocal imaging systems

Abigail S. Hedden
Charles R. Dietlein
David A. Wikner
U.S. Army Research Laboratory
2800 Powder Mill Road
Adelphi, Maryland 20783
E-mail: abigail.s.hedden.civ@mail.mil

Abstract. The authors analyze properties of a 220 GHz imaging system that uses a scanned reflectarray to perform electronic beam scanning of a confocal imager for applications including imaging meter-sized fields of view at 50 m standoff. Designs incorporating reflectarrays with confocal imagers have not been examined previously at these frequencies. We examine tradeoffs between array size, overall system size, and number of achievable image pixels resulting in a realistic architecture capable of meeting the needs of our application. Impacts to imaging performance are assessed through encircled energy calculations, beam pointing accuracy, and examining the number and intensity of quantization lobes that appear over the scan ranges of interest. Over the desired scan range, arrays with 1 and 2-bit phase quantization showed similar array main beam energy efficiencies. Two-bit phase quantization is advantageous in terms of pointing angle error, resulting in errors of at most 15% of the diffraction-limited beam size. However, both phase quantization cases considered resulted in spurious returns over the scan range of interest and other array layouts should be examined to eliminate potential imaging artifacts. [DOI: 10.1117/1.OE.51.9.091611]

Subject terms: millimeter wave technology; millimeter wave imaging; scanning radar architecture; scanning reflectarray; phased arrays.

Paper 120206SS received Feb. 14, 2012; revised manuscript received Apr. 16, 2012; accepted for publication Apr. 19, 2012; published online May 18, 2012.

1 Introduction

Detection of concealed explosives and other devices easily masked by clothing is a significant application of passive and active millimeter-wave/terahertz imaging systems.[1] Developing methods for detection at long ranges is crucial for effectively dealing with this threat. For applications involving mobile targets, imaging large fields of view at near video frame rates is needed. A common approach implemented in state-of-the-art millimeter-wave and terahertz systems uses mechanical scanners for rapid beam steering. A drawback of mechanical scanning is the associated hardware,[2] which can be physically large, heavy, and cause vibrations. Many radar applications, unlike passive millimeter-wave imaging, do not need the large bandwidth provided by scanning mirrors. In this context, rapidly rotating and oscillating elements are often a hindrance. Developing electronic scanning capabilities in the millimeter-wave and terahertz regimes avoids these downsides, which yields robust, portable, and lightweight systems.

Phased array technology and electronic beam scanning techniques are well developed at frequencies below 100 GHz. Extending this technology to the upper millimeter-wave/terahertz regime is an active area of research.[3–5] Phase shifters have already been demonstrated at frequencies above 200 GHz[6] and low-loss transmission line structures and terahertz integrated circuits have been demonstrated at frequencies beyond 600 GHz.[7] Although this technology is in the early stages of development, it is paving the way for future high frequency phased arrays.

0091-3286/2012/$25.00

To the authors' knowledge, designs incorporating reflectarrays with confocal reflector systems have not been explored at frequencies as high as 220 GHz. In this work, we analyze an architecture that uses a reflection-type phased array, or scanning reflectarray, to perform electronic beam scanning of a confocal imager to image humans at standoff distances of 50 m. This architecture is particularly useful for applications that simultaneously require a high degree of lateral resolution and modest bandwidth (several percent), such as radar imaging. We concentrate, in particular, on characteristics of the reflectarray that impact design, for example, its linear dimensions, the number of elements, element spacing, and the number of phase quantization levels. Section 2 describes the system architecture and the constraints on the imager that result from our desired application. Results of a geometrical optics analysis are presented, and the implications for electronic scanning are considered. In Sec. 3, uniform lattice arrays compatible with the imager architecture are used to analyze properties including beam symmetry and encircled energy as a function of scan angle and array element spacing. Phase quantization is considered for two-dimensional (2-D) rectangular lattice arrays, and impacts to image quality including beam pointing error and the relative number and intensity of quantization lobes are analyzed for arrays with 1- and 2-bit phase shifters. A brief study of reflectarray element design and implementation is also presented.

2 Confocal Imager

2.1 *Basic Constraints*

For imaging humans in meter-sized fields of view at standoff distances of approximately 50 m, centimeter-scale lateral

resolution is desirable. However, apertures larger than 1 m are required to achieve diffraction-limited spot sizes of 5 cm or less at frequencies below about 400 GHz. Since many standoff imaging applications are not well suited to bulky, heavy optics, this problem can be solved by operating at higher frequencies (>400 GHz) with smaller apertures. Although component technology is not well developed at these frequencies, recent and ongoing programs[5] make the development of electronic scanning capabilities above 400 GHz an attractive consideration for future efforts.

In this work, we investigate the impact of electronic scanning on system design by considering a system that operates in the 220 GHz atmospheric window. We selected this frequency as a compromise between aperture diameter, diffraction-limited spot size, component availability, signal attenuation, and the transmission and reflectivity characteristics of targets. Compared with higher frequency operation, 220 GHz offers increased transmission through clothing[1,8] and more mature component technology.

For our design, we chose a confocal design for its ability to scan the output beam of a large aperture using a relatively small scanning element. Confocal reflector systems have been demonstrated with mechanically scanned terahertz imagers[9,10] for concealed weapons detection and were proposed for use with phased arrays for satellite-based applications.[11] Additionally, they can be implemented without obstructing the primary aperture, accommodate a fully-illuminated aperture for high resolution, and be scanned up to several degrees off axis without significant vignetting of the scanned beam by the primary. Subsequently, we refer to this as spillover.

2.2 *System Architecture*

Figure 1 is a schematic representation of our design along with an equivalent simple lens model of the reflector system. A paraboloidal feed reflector illuminates the scanning reflectarray with a plane wave. The diameter of the active area of the array is d_A. The reflectarray scans the beam $\pm\theta_2$ over an under-illuminated secondary reflector (element 2). The primary and secondary reflectors, with apertures and focal lengths of f_1, D_1, and f_2, D_2, respectively, share a common focus and are designed so that the primary D_1 is fully filled

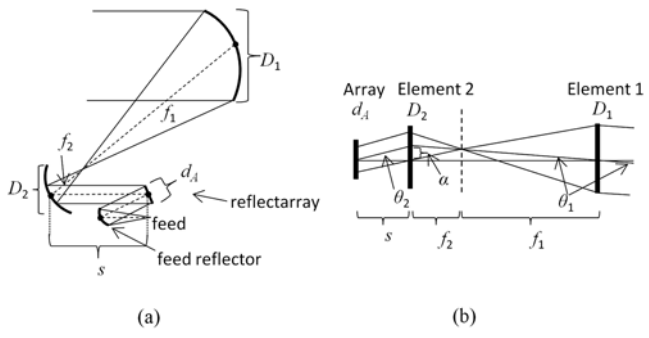


Fig. 1 (a) Diagram of the layout of a scanned reflect-array incorporated into a confocal reflector system. The array scans the beam over an under-filled secondary reflector, thereby steering the beam at the primary. The primary element is fully filled over a scan range of $\pm\theta_1$. (b) An equivalent lens model of the system showing the array and two focusing elements.

by the array beam with negligible spillover as the beam is scanned off axis by $\pm\theta_1$.

We use geometrical optics to obtain rough system dimensions and properties of the elements shown in Fig. 1. As we show, the design is sensitive to d_A. Although increasing d_A narrows the width of the scanned beam in object space, it also increases the length of the optical system. Our goal is to design a system that meets our scanning criteria but is moderately sized.

To balance the physical size of the system and its spatial resolution, we set $D_1 = 1$ m. At 220 GHz, this yields a diffraction-limited spot size of 1.66 mrad, which in turn yields a spot size of 8.3 cm at 50 m. Because we fix the aperture diameter of the primary to 1 m, we can control system size by controlling its total length,

$$f_{\text{tot}} = s + f_2 + f_1. \tag{1}$$

Although the optical design could be made more compact, the total focal length provides a sense of its overall size.

In order to fill the primary aperture without significant spillover as the array scans the beam off axis, element 2 maps the array aperture distribution onto the element 1 aperture. From the thin lens formula, the distance s is related to the focal lengths of the elements according to

$$\frac{1}{f_2} = \frac{1}{s} + \frac{1}{f_1 + f_2}. \tag{2}$$

Furthermore system magnification is given by

$$M = \frac{-(f_1 + f_2)}{s}. \tag{3}$$

Note that the image of the reflectarray is inverted in the aperture of the primary, which indicates M is negative in our system. Equation (3) shows that if f_2 and s are fixed, then f_1 grows with magnification. Therefore, from Eq. (1) it is obvious that the entire system grows as magnification increases.

From an alternative definition of system magnification,

$$M = -\frac{D_1}{d_A}, \quad = \frac{\tan\theta_2}{\tan\theta_1}. \tag{4}$$

We note that, for fixed D_1, $|M|$ decreases as the reflectarray increases in size. If we assume the spacing between array elements is fixed at $\lambda/2$, increasing d_A increases the number of elements that contribute to a scanned beam, which increases the beam efficiency.[4] This tradeoff is reflected in Fig. 2, which represents the relationship between reflectarray size and both system magnification and reflectarray elements. Magnification is representative of system length, whereas number of elements is an indication of beam quality. In the remainder of this work, we consider reflectarray diameters compatible with common wafer sizes from 2 inch to 6 inch (~5 cm to 16 cm).

In addition to illuminating the primary aperture fully, we must ensure that element 2 is large enough to accept the beam from the reflectarray as it scans over a range $\pm\theta_2$. Thus,

$$D_2 \geq d_A + 2\alpha. \tag{5}$$

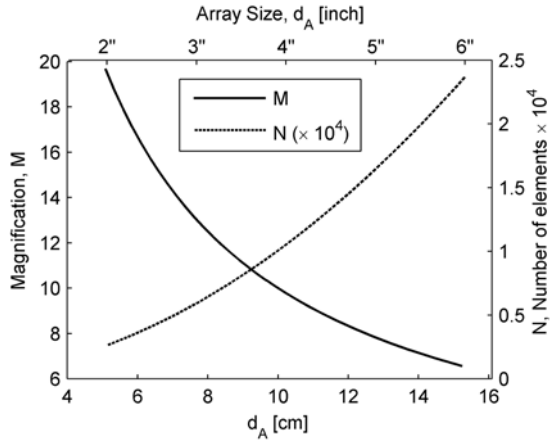


Fig. 2 System magnification and reflectarray elements as a function of reflectarray size expressed in inches and in cm. Magnification is representative of system length, whereas number of elements is an indication of beam quality. This highlights the tradeoff between overall system size and array complexity.

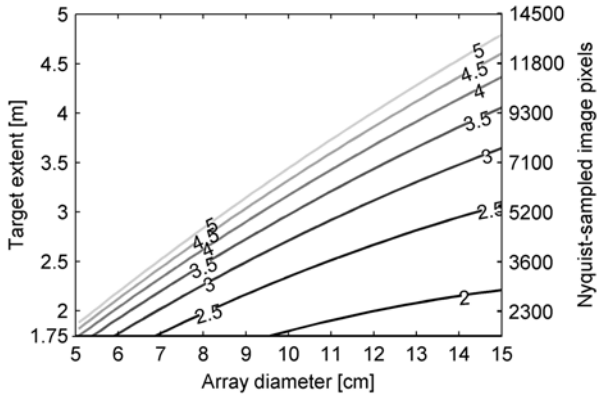


Fig. 3 Plot of field of regard (expressed as both target size and number of half-beamwidth image pixels) as a function of reflectarray size for various system *f*-numbers, $(f_1 + f_2 + s)/D_1$ calculated for a range of 50 m and $D_1 = 1$ m.

In order to ensure small system size, we assume element 2 is an $f/1$ optic, $D_2 = f_2$, which constrains d_A,

$$d_A \leq f_2[1 - 2(1 - M)\tan\theta_1]. \tag{6}$$

One can rewrite Eq. (6) using Eqs. (2) through (4) to indicate explicitly the limitations the design imposes on the scanner's field of view (FOV), defined by scan angle θ_1,

$$\left(\frac{1}{2}\right)\left[\frac{1 - \left(\frac{D_1}{f_1}\right)}{1 + \frac{D_1}{d_A}}\right] \geq \tan\theta_1. \tag{7}$$

Figure 3 illustrates the impact of this limitation on system operation assuming $D_1 = 1$ m and a 50 m object distance. If total FOV is approximately $2\theta_1$, the largest target the system can scan at 50 m is $100 \times \theta_1$ m for θ_1 expressed in radians. We plot target extent in Fig. 3 as a function of array diameter with system f-number, f_{sys}, as a parameter, where

$$f_{\text{sys}} = f_{\text{tot}}/D_1. \tag{8}$$

The upper left portion of the plot corresponds to an unphysical system space where the desired target region exceeds the maximum achievable scan angle. The right, vertical axis refers to the number of pixels in an image for a given target extent, if a pixel is defined at every half beamwidth. Note that, at small array diameters, system size (indicated by f_{sys}) increases rapidly with field of view. For a fixed system size, mapping large fields of view requires large array diameters. It is worth noting that tripling the array diameter yields only a modest increase in FOV, roughly double, whereas, from Fig. 2, the complexity of the array increases exponentially with the number of elements.

Because each sample within the FOV requires a minimum integration time to ensure low noise detection, our desire to generate images at 30 Hz frame rates also imposes limits on the FOV. Figure 3, therefore, also indicates the number of samples the system generates within the target region assuming an 8.3 cm beamwidth and half beamwidth sampling. If we assume a 30 Hz frame rate, a 1 m diffraction-limited primary aperture, and a 2 m target region, the maximum integration time per sample is 15 μs. The position update rates of phase shifters used to scan the array limit the field of view and, given that 10 μs switch times have been demonstrated at microwave frequencies with variable microelectromechanical system (MEMS) capacitors,[12] an FOV corresponding to a 3 m target extent is the limit at which images can be captured at 30 Hz frame rates.

Based on our analysis, we adopt a 3-inch-diameter reflectarray design. Table 1 provides an overview of system details assuming $d_A = 7.5$ cm (3 inches), $D_1 = 1$ m, and a maximum desired steering angle of $\theta_1 = \pm 1$ deg. This yields $M = 13$, $f_{\text{sys}} = 2.3$, and an array size of approximately $N = 6000$ elements (assuming half wavelength spacing). From Table 1, the overall size of the 220 GHz system is at the edge of what may be considered portable. However, a benefit of the design is its potential to be scaled for higher frequency operation. For example, a system operating at 410 GHz with $D_1 = 0.53$ m achieves a similar diffraction-limited spot size to the one shown in Table 1. If the array diameter and desired total scan angle remain the same, ($d_A = 7.5$ cm and $\theta_1 = \pm 1$ deg) then $M = 6.6$, $f_{\text{tot}} = 90$ cm, and $f_{\text{tot}}/D_1 = 1.8$ for the 410 GHz system.

3 2-D Array Architecture

3.1 *Uniform Arrays: Size and Element Spacing Effects*

In this section we consider the properties of the far-field pattern generated by uniform rectangular lattice arrays whose size and element spacing are compatible with the confocal reflector system described in Sec. 2.2. To do so, we generate the complex wave-amplitude of the far-field pattern $\text{AF}(\theta, \phi)$, referred to as the array factor, for different scan angles and array layouts. For an $N = A \times B$ element array with d_x and d_y element spacing along orthogonal array axes, the array factor is given by:[13]

Table 1 Details of a confocal reflector system incorporating a ~3-inch diameter reflection-type phased array including approximate sizes and focal lengths of optical elements, system magnification, and desired scan angles.

d_A (cm)	s (cm)	θ_2 (deg)	D_2 (cm)	f_2 (cm)	f_1 (cm)	D_1 (cm)	M (unitless)	θ_1 (deg)	Target (m)
7.5	16	13	15	15	200	100	13	1	1.7

$$\mathrm{AF}(\theta,\phi) = \sum_{a-1}^{A} \exp[j(a-1)kd_x \sin\theta\cos\phi - \alpha_{x,a}] \cdot \sum_{b-1}^{A} \exp[j(b-1)kd_y \sin\theta\sin\phi - \alpha_{y,b}],$$
$$\alpha_{x,a} = (a-1)kd_x \sin\theta_0\cos\phi_0 \quad \text{and}$$
$$\alpha_{y,b} = (b-1)kd_x \sin\theta_0\sin\phi_0, \tag{9}$$

where $a_{x,a}$ and $a_{y,b}$ are the phase values for each element given a desired steering angle (θ_0, ϕ_0) in the plane of the reflectarray.

In our analysis, we consider arrays with uniform spacing $\lambda/2$ or larger. Antenna structures and the physical size of high frequency phase shifters[3,6,14] make half-wavelength spacing difficult to achieve. We assume the arrays consist of identical elements with equal amplitudes and, initially, continuous phase. We consider the effects of phase quantization in Sec. 3.2.

Table 2 shows properties of arrays with different element spacings, scan angles, and sizes compatible with fabrication on 2-inch, 3-inch, and 6-inch wafers. All arrays shown can be used with the confocal reflector system in Sec. 2.2. The designs in rows 5 to 7 of Table 2 are consistent with the system described in Table 1. Pattern properties, including full-width half-max (FWHM) values of array main beams, were determined from two-dimensional Gaussian functions fit to the far-field power pattern $|\mathrm{AF}(\theta,\phi)|^2$. We found the best fit by minimizing the sum of the squares of the errors between the generated pattern and the Gaussian functions iteratively. The pattern properties extracted from Gaussian fits to the main beam yield far-field characteristics that match those predicted by analytical formulae. For example, consider the data presented in the first two rows of Table 2. As expected, for a square array of equally spaced elements, the fitted orthogonal beamwidths for the broadside array (row 1) are the same. Also, as the beam is steered off broadside to 15 deg, the fitted width in the dimension parallel to $\phi = \phi_0 = 0$ deg grows as sec θ_0, while the width in the perpendicular dimension is unaffected. Comparing the far-field beamwidths from the numerical calculation with the results from analytic expressions that are valid for large uniform rectangular arrays whose maxima are not steered far from broadside[13] reveals <3% difference between the FWHM values.

For uniform arrays with element spacings greater than $\lambda/2$, the main beam FWHM characteristics are largely unaffected by element spacing for a given scan angle since each array configuration occupies the same physical area. Encircled energy calculations are sensitive to array element spacing and provide a quantitative way of assessing the impact of grating lobes and scan angle error on beam characteristics.

The encircled energy was calculated by dividing the total energy contained within the FWHM due to the array by the total energy expected for a solid reflector, E_a/E_s, of the same size. The FWHM of the beam pattern due to the solid reflector is calculated as a normalized, Gaussian function fit to the Airy pattern:

Table 2 The properties of different uniform arrays are shown for rectangular lattice arrays with various layouts. Properties were calculated by fitting Gaussian functions to the array patterns. E is the percentage of energy contained within an angular range equal to the FWHM of F_{theory}.

d_A (cm)	size ($M \times N$)	spacing (λ_0)	θ_0 (deg)	FWHM (deg)	E (%)
3.6	51 × 51	0.5	0	[1.96, 1.96]	97
3.6	51 × 51	0.5	15	[2.00, 1.94]	90
3.6	36 × 36	0.7	15	[2.03, 1.96]	88
3.6	25 × 25	1.0	15	[2.05, 1.98]	46
5.4	78 × 78	0.5	15	[1.31, 1.27]	90
5.4	55 × 55	0.7	15	[1.33, 1.28]	89
5.4	39 × 39	1.0	15	[1.31, 1.27]	43
10.8	155 × 155	0.5	15	[0.66, 0.64]	91
10.8	110 × 110	0.7	15	[0.66, 0.64]	91
10.8	77 × 77	1.0	15	[0.66, 0.64]	52

$$F_{\text{theory}}(x,y) = \exp\left[-(x-x_0)^2/2\sigma_x^2 + (y-y_0)^2/2\sigma_y^2\right]$$
$$\sigma_x = 1.028\lambda_0/(d_A \cos\theta_0) \quad \text{and} \quad \sigma_y = 1.028\lambda_0/d_A, \tag{10}$$

where the pattern widths are given by $1.028\lambda_0/D$ and $D = d_A \cos\theta_0$ is the projected array diameter. The results are shown in Table 2 and Fig. 4 where encircled energy is plotted as a function of angular radius for array layouts shown in Table 2 rows 5 to 7 for 15 deg scan angles. From Fig. 4 it is clear that arrays with smaller element spacing rapidly approach 90% encircled energy as angle increases. However, for the array with element spacing, only about 40% of the total energy is contained within the FWHM of F_{theory} due to grating lobes. This is also reflected in the results shown in Table 2 for all layouts with λ_0 element spacing; about 50% or less energy is contained within a small angular radius of the main beam.

Since grating lobes can contain a large fraction of the total power, this limits the spacing of array elements to $<\lambda_0$. The results show that $0.7\lambda_0$ spacing is quite comparable to $0.5\lambda_0$ and a good option. Having a significant fraction of power located in grating lobes is detrimental to the system efficiency since the gain of the main beam and the energy coupled into the optical system are reduced. Even more problematic is the increased level of image clutter that can arise if unwanted lobes propagate through the optical system. The reflectivity characteristics of terrestrial scenes at millimeter-wave frequencies can vary by greater than 20 dB, and false returns may be confused with those of the main beam. This problem, including the number and intensity of such lobes is explored further for our optical system in the next section.

3.2 *Element Phase Quantization*

3.2.1 *Continuous phase shift*

The previous section assumes a continuous phase shift is applied across the array to achieve beam scanning. The amount of phase shift between elements needed to steer the main beam of a uniform array over a given angular range is determined by the array geometry and the operating wavelength. The progressive phase shift required to achieve a particular scan angle is:

$$\delta_{\text{shift}} = \frac{2\pi}{\lambda_0} d \cos(90 - \theta_2), \tag{11}$$

where d is the element separation and λ_0 is the operating wavelength. For large arrays with uniform spacing of order $\lambda_0/2$ or more, each element must be capable of 360 deg of phase shift. The maximum required progressive phase shift between elements to cover the full field of view is about 40 deg. Scanning the beam over this field of view in half-beamwidth increments requires a minimal progressive phase shift of about 2 deg. This implies that an eight-bit phase shifter is required.

Phase shifters with a 360 deg tuning range and an accuracy and reproducibility of 2 deg have not been demonstrated at 220 GHz. Also, developing digital phase shifters with seven to eightbits is prohibitive at these frequencies due to overall unit cell size and the complexity of bias routing, among other considerations. At frequencies around 200 GHz, several phase shifter designs have been successfully demonstrated. For example, phase shifters using GaAs Schottky varactor diodes integrated with 90 deg hybrid microstrip circuits exhibited 180 deg of phase shift with errors of ±15 deg over a 195 to 250 GHz frequency range.[15] Also, distributed MEMS transmission lines have been demonstrated at G-band (140 to 220 GHz), including 4- and 15-element designs based on switched MEMS capacitors,[6] and single-bit switched-line phase shifters have been proposed for 600 GHz operation.[3] Achieving beam scanning via continuous progressive phase shift is not practical given the stringent requirements on total tunable range and accuracy mentioned above. For this reason, the rest of this section explores properties of arrays with quantized phase shift capabilities.

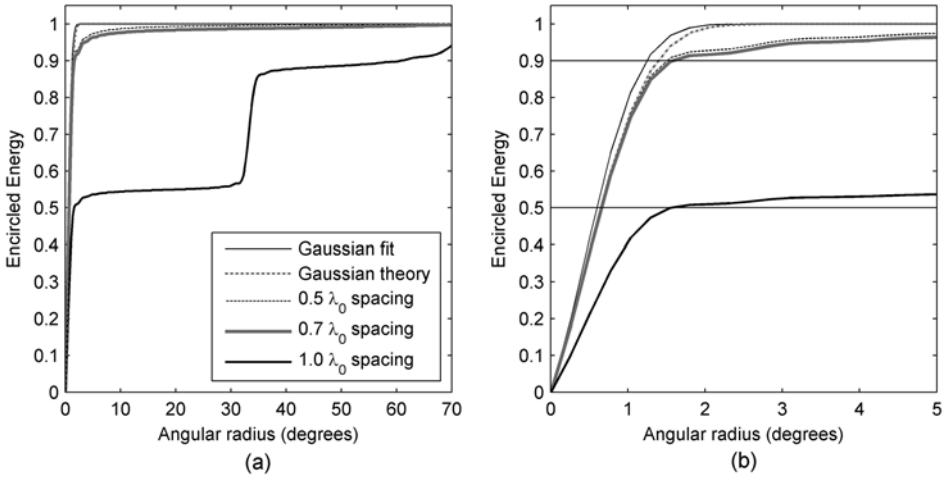


Fig. 4 (a) Encircled energy calculations for similar array layouts containing different element spacings are shown for a desired scan angle of 15 deg. Two best case curves are presented as Gaussian fits to both the Airy pattern of a solid reflector (Gaussian Theory) and array factor power pattern (Gaussian Fit). (b) A zoomed version of (a) showing the same encircled energy curves along with the results of fitted and theoretical Gaussian curves. For all layouts except the one with λ_0 element spacing, the encircled energy rapidly approaches >90% levels within small angular radii.

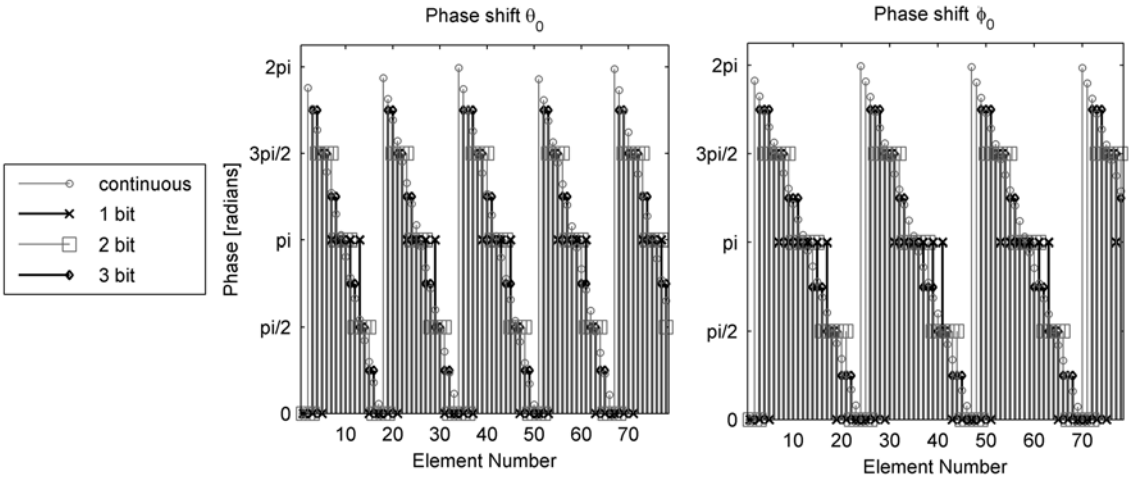


Fig. 5 Phase needed to achieve a desired scan angle of $(\theta_0, \phi_0) = (7\ \text{deg}, 5\ \text{deg})$ is shown as a function of element number across a 78×78 element half-wavelength spaced array for different phase quantization cases up to three bits.

3.2.2 *2-D rectangular lattice array with quantized phase shift*

Using Eq. (9) phase quantization is applied by replacing the ideal phases with quantized phase states, $\alpha_{q,a}$ and $\alpha_{q,a}$, that are multiples of $2\pi/2^l \bmod 2\pi$, where l is the number of bits. As an example, Fig. 5 shows the phase shifts needed for the 78×78 element half-wavelength spaced array in Table 2 (see row 5) to achieve an arbitrary scan angle of $(\theta_0, \phi_0) = (7\ \text{deg}, 5\ \text{deg})$. Phase shift is shown as a function of array element number for continuous ideal phase shift and for one to three bit phase quantization. The resulting array factors are shown in Fig. 6 for arrays with one to three bit phase shifters. The array factors are normalized to the peak intensity values and are plotted in direction cosines. The grayscale range has been restricted to bring out lower level structure in the images, highlighting features at the -10 to -20 dB level.

The most important feature in Fig. 6 is the appearance of quantization lobes, which occur in antenna arrays or diffractive optical elements with finite phase quantization.[16–18] The quantization lobes are most significant, in terms of quantity and magnitude, for the array with one-bit phase quantization. For the one-bit case, three quantization lobes with equal amplitude to the main beam appear, symmetric about broadside. Quantization lobes also appear in array factors with two-bit and three-bit phase shifters, but their significance decreases as the number of quantization levels increases. These lobes are weaker for a reflectarray illuminated by a spherical wave front compared with a focusing feed reflector.

Quantization lobes affect system performance in several ways, resulting in scan angle errors and decreased main beam directivity. Additionally, quantization lobes are clearly harmful if they are in close angular proximity to the main beam and can therefore propagate through the optical system. Their effects on the imaging performance of the system presented in Sec. 2.2 will be discussed in terms of beam pointing accuracy, the number and intensity of lobes appearing over the entire desired scan region, and their impact on the fraction of total energy contained within the main beam as a function of scan angle.

The remainder of this section compares the relative imaging performance of the system described in Sec. 2.2 with a half-wavelength spaced, 3-inch diameter array with one- and two-bit phase quantization. The decision to focus on one- and two-bit phase quantization is based on a desire to minimize fabrication complexity of the reflectarray. This is particularly important for large half-wavelength spaced arrays where minimization of unit cell area is highly beneficial. For example, the 5.4 cm 78×78 element array shown in Table 2 requires more than 6000 elements for a one-bit phase shifter design. Since even the simplest unit cell designs, such as switched line phase shifters, can require many switches to realize 2^b phase states where b is the

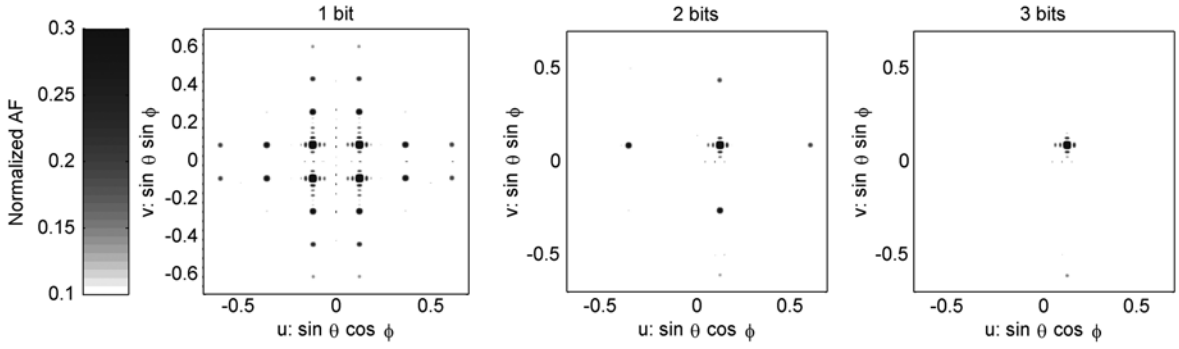


Fig. 6 Plot of the array factors for arrays with one- to three-bit phase quantization shown in Fig. 5. The array factors have been normalized to the peak intensity of the main lobes and are shown plotted in direction cosines.

number of bits, the feasibility of bias routing quickly becomes a significant concern along with overall array complexity. The physical size of phase shifting elements is an additional hurdle that makes it difficult to implement multi-bit phase shifters while maintaining half-wavelength spacing. For this reason, we choose to examine tradeoffs between systems with one- and two-bit phase quantization.

3.2.3 *Beam pointing accuracy*

The introduction of phase quantization leads to beam pointing errors resulting from array phase error. This error is most severe for scan angles near broadside where the change between phase states is large and for angles with an integer number of elements per phase step. Studies of one-dimensional (1-D) arrays reveal scan errors ranging from a few tenths to a few thousandths of a degree depending on array size, number of elements, and number of bits.[19] We analyze the effects of beam pointing error for the system described in Sec. 2.2 with a half-wavelength spaced 78×78 element array (Table 2, row 5) and examine the magnitude of this error with one- and two-bit phase quantization over the desired scan range of the optical system.

To evaluate beam pointing accuracy, array factors were calculated for arrays with one- and two-bit phase quantization using Eq. (9) over all scan angles (θ_0, ϕ_0) covering azimuth and elevation ranges compatible with the optical design shown in Table 1. Az/El scan angles of ± 13 deg (for a total of 26 deg) are needed at the secondary to map the desired field of view. We assume a minimum scan angle increment of 0.65 deg on the secondary. The 26 deg scan region in Az/El was divided into individual desired pointing angles $(\mathrm{Az}_0, \mathrm{El}_0)$ based on the minimum scan angle increment, and array factors were generated for each location. For all scan angles, array factors were generated over a finely meshed grid with angular spacing much smaller than the minimum scan angle in regions surrounding the expected location of the main lobe. The beam pointing error (Δ_p) was determined by the magnitude of the difference between the desired scan angle $(\mathrm{Az}_0, \mathrm{El}_0)$ and the location of the array factor maximum peak $(\mathrm{Az}_{\mathrm{peak}}, \mathrm{El}_{\mathrm{peak}})$: $\Delta_p = \sqrt{(\mathrm{Az}_0 - \mathrm{Az}_{\mathrm{peak}})^2 + (\mathrm{El}_0 - \mathrm{El}_{\mathrm{peak}})^2}$.

The results are shown in Fig. 7 with the magnitude of beam pointing error plotted as a grayscale image over the total Az/El region for arrays that differ only in their phase quantization. The range of beam pointing error shown in the figure has been clipped at 0.3 deg to highlight lower-level detail; for comparison purposes both panels of Fig. 7 have been plotted with this intensity scale. Several features are apparent in the plots. The relatively large errors observed around the origin and in stripes near the plot axes and at ± 15 deg elevation are due to large changes in phase states. The differences between achieved and desired scan angles are particularly large for angles that are nearly on-axis for the one-bit phase shifter case, where the phase states abruptly change by π. This error also appears to a lesser degree in the two-bit phase shifter plot. The average and median scan angle errors over the entire regions are shown in Fig. 7, (0.07, 0.05 deg) and (0.04, 0.03 deg) for both cases, respectively. However, as highlighted in the plot, there are regions where the scan angle error is much greater, and the maximum error levels for the one- and two-bit phase quantization cases are 0.6 and 0.2 deg, respectively. This is a significant fraction of the required minimum scan angle increment of 0.65 deg.

The result is that for the one-bit phase shifter case, a restricted scan range is needed in order to avoid these regions. For example, scanning in the $+\mathrm{Az}/+\mathrm{El}$ quadrant away from the origin and axes leads to maximum pixel misregistration levels of less than 0.3 deg. Beam pointing accuracy with an array that has two-bit phase quantization alleviates pointing error issues even further. Regions with the largest scan angle errors result in pointing errors of about 15% of the diffraction-limited resolution. In this case, no restriction of scan range is necessary, enabling more flexibility in the optical system design if needed.

3.2.4 *Effects of quantization lobes*

Another problem associated with phase quantization that results in degraded imaging performance is the appearance

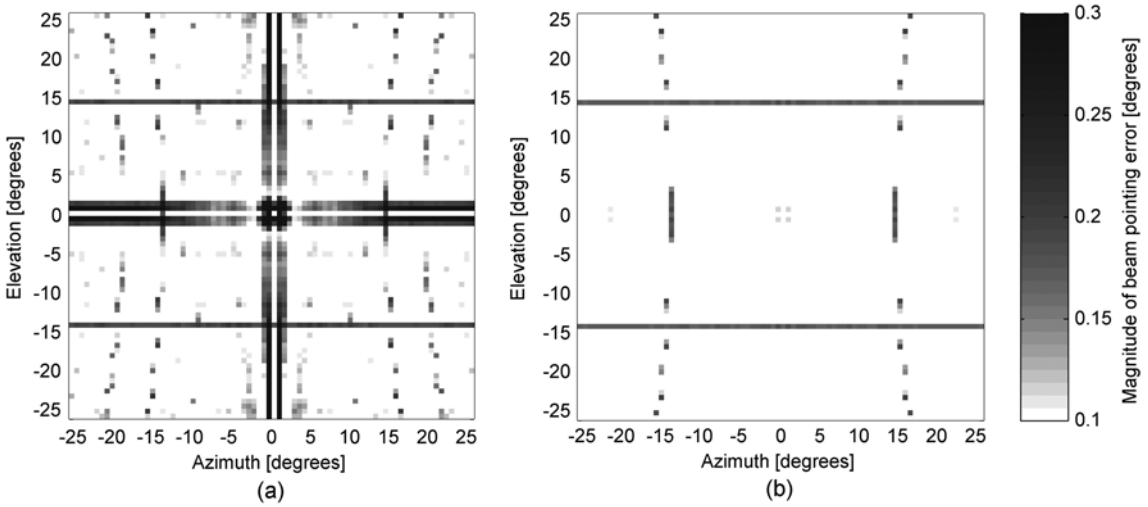


Fig. 7 (a) The magnitude of steering angle error in degrees for an array with one-bit phase quantization is shown as a function of scan ranges of interest for our optical system. (b) Shows steering angle error as a function of scan angle for the two-bit phase quantization case. For comparison purposes, the plot range matches the results of (a).

of quantization lobes in close angular proximity to the main beam. If these lobes are not spatially blocked and propagate through the optical system, they can cause false returns and introduce artifacts. This problem is compounded for applications involving imaging of moving targets at high frame rates where the scene may be rapidly changing from frame to frame. If one assumes that energy outside of the desired angular scan region is terminated in an absorptive load, the optical system presented in Sec. 2.2 will perform some spatial filtering. This is particularly important for scanned arrays with one-bit phase quantization, where quantization lobes including ones of equal intensity to the main beam appear symmetrically about the origin in Az/El (see Fig. 6). However, this does not address the appearance of quantization lobes within the desired Az/El scan region. For this reason, we compare characteristics of arrays with one- and two-bit phase quantization in terms of their effects on the imaging performance of the optical system shown in Table 1. This work examines the angular location, number, and intensity of quantization lobes calculated over the scan range of interest for these particular setups.

Figure 6 displays the angular locations of quantization lobes for a 78×78 element half-wavelength spaced square lattice array for one particular scan angle. To assess the impact of quantization lobes over the total desired scan region, array factors were generated in a manner similar to the previous section for scan angle increments of 0.65 deg over a total scan region including 0 to 26 deg in $+\mathrm{Az}/+\mathrm{El}$. The fraction of energy contained inside the main beam was calculated and used to determine the impact of quantization lobes. Individual summed energy results were normalized to the total energy calculated over the whole mapped region for each array factor. The angular radius of the main beam was determined by the diffraction limited $1/e^2$ radius of a theoretical Gaussian due to the array Eq. (10) calculated for each desired scan angle.

Using this method, the fraction of energy in the main beam was computed over all scan angles in the $+\mathrm{Az}/+\mathrm{El}$ quadrant for the one- and two-bit phase quantization cases. The beam efficiency results are shown in Fig. 8 as a function of pointing angle. The grayscale range has been restricted to highlight energy fractions between 10% and 50%. There is comparatively little structure in the plots near the origin where the energy fraction is the greatest (>70%). The average fraction of energy in the main beam for arrays with both one- and two-bit phase shifters is small, about 20% and 23%, respectively, ignoring energy lost to other quadrants. It is evident that the impact of quantization lobes and side lobes on the main beam energy fraction is significant. There is an additional 25% efficiency inherent to one-bit phase quantization since this plot neglects the radiated energy appearing symmetrically about the Az/El origin (see Fig. 6). Although this decreases the energy efficiency, it is not an insurmountable problem as long as enough source power is available to the radar imaging system and the other quadrants are spatially filtered.

The array main beam energy efficiency results do not provide direct information about the number and intensity of the lobes. To this end, we used the same array factors generated over the $+\mathrm{Az}/+\mathrm{El}$ scan region to calculate the number of quantization lobes appearing over the scan range of interest as a function of lobe intensity. To perform this calculation, we set a peak threshold relative to the main beam of -20 dB. All lobes at and above this threshold were tabulated in bins of 1 dB for each array factor. The results were summed by bin and normalized by the total number of array factors calculated. Figure 9 shows a histogram of the results for arrays with one- and two-bit phase shifters.

As expected, the total number of lobes/AF summed across all bins is greater for one-bit phase quantization. Generally, for the two-bit phase quantization case, there are fewer overall lobes, including those of comparable intensity to the main beam (this is seen down to the -6 dB level). For the one- and two-bit phase quantization cases, the total number of lobes/AF appearing over the scanned region down to the -20 dB level is about 60 and 50, respectively. Most of these lobes

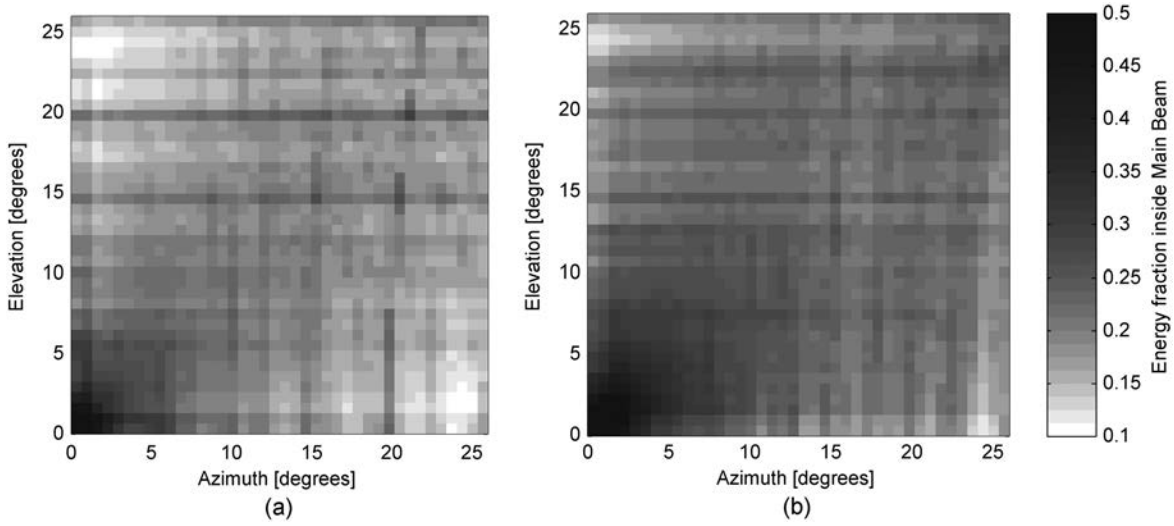


Fig. 8 (a) The fraction of energy contained within the main beam is shown over the desired Az/El scan range for an array with one-bit phase quantization. The average energy fraction is <20%. (b) Shows the energy fraction within the main beam for the two-bit phase quantization case calculated over the same scan region. The average energy fraction is about 25% and for easy comparison; the plot range matches the results of (a).

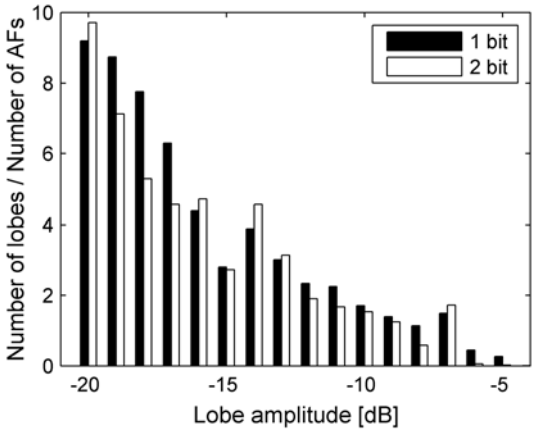


Fig. 9 Histograms of the number of quantization lobes normalized to the total number of array factors calculated over the scan range of interest are shown binned by lobe amplitude relative to the main beam. Results are shown for one- and two-bit phase quantization cases.

appear at lower intensity levels. For example, considering lobes ranging between 0 dB and −15 dB reduces both of these numbers by more than half to about 20 lobes total.

As shown in Fig. 9, a significant number of lobes appear above the −20 dB threshold over the total desired scan region. These lobes are potentially harmful since they propagate through the optical system and can result in spurious returns. As discussed above, they are less severe in terms of number and intensity for the array with two-bit phase shifters, but they are prevalent. In order to mitigate or eliminate the possibility of false returns and irreparable imaging artifacts, it may be advantageous to consider alternatives. Since larger arrays and arrays with more phase quantization states add considerably to overall complexity, contributing to future fabrication difficulties, examining the performance of other array geometries, such as aperiodic arrays, may provide better results.

3.3 *Reflectarray Elements*

3.3.1 *Antennas*

Antenna elements compatible with large scanning reflectarrays must meet several criteria. They need to be compatible with wafer scale assembly and phase shifter integration. They must also be compatible with unit cell separations less than λ_0 and incorporation into large-format 2-D arrays with low mutual coupling. At lower frequencies up to 60 GHz, reconfigurable reflectarrays have been proposed and demonstrated with microstrip patches,[20–22] although a variety of planar antenna elements including dipoles, coplanar waveguide (CPW), and slot antennas may also suffice. In the high frequency millimeter-wave region, suitable antenna element choices are more limited. At frequencies higher than 120 GHz, the small widths required for slot antennas and CPW make physical implementation of these structures challenging.[3] Planar dipoles are a possibility but may not be compatible with common phase shifter geometries, including MEMS switches. On the other hand, microstrip patches meet all of the criteria and are a favorable choice for implementing a scanned 220 GHz reflectarray.

3.3.2 *Element integration*

Adopting a simple element structure is important for large array construction, particularly since integration of antennas with phase shifters has not yet been demonstrated at frequencies as high as 220 GHz. A possible unit cell configuration that has been implemented at 60 GHz[21] and proposed for use at frequencies as high as 600 GHz[3] uses a microstrip patch antenna connected to a short-circuited stub that is loaded with either a diode or MEMS switch. The switch or diode acts as a one-bit phase shifter, providing either an open or short impedance, and results in a 180 deg phase shift between the two states. This simple geometry is promising for realizing 220 GHz unit cells, although the size of phase shifters and MEMS switches is a concern[4] at these frequencies. Their dimensions can be comparable to patch antennas and viable unit cell designs must minimize the possibility of radio frequency (RF) signal coupling to phase shifters. Digital control is a key capability for a large array. Unit cell designs incorporating electronically controlled components are therefore advantageous. G-band MEMS-based varactors and switched capacitors have been demonstrated with digital control,[6] providing a path to realizing array unit cells. An added benefit of these components for large arrays is the use of fabrication techniques that are compatible with CMOS processes and large-volume production methods.

4 Conclusions

In this work we analyze the performance of a system that uses a scanning reflectarray, or reflection-type phased array, to perform electronic scanning of a confocal reflector system for applications including video-rate imaging of meter-sized targets at standoff distances of 50 m. This work analyzes a system design for operation in the 220 GHz atmospheric window. Designs incorporating reflectarrays with confocal imagers have not previously been explored at this frequency range. The 220 GHz window was chosen as a favorable compromise between aperture diameter, resolution, component availability, signal attenuation, and transmissivity/reflectivity properties of targets. Details of the design are presented in this work. A confocal reflector system was chosen mainly for its ability to scan a large aperture with a much smaller scanning element. This is particularly attractive since large-format electronically scanned arrays have not been demonstrated at these frequencies. A 1 m primary aperture was selected as a tradeoff between system size and resolution. A 3-inch wafer-scale array was chosen based on compromises between system size, total array elements required, and achievable number of image pixels. Properties of uniform rectangular lattice arrays of different sizes and element spacings compatible with the confocal reflector system were examined. Encircled energy calculations revealed that grating lobes can contain large fractions of the total energy (>50%) over the angular scan ranges of interest for arrays with element spacings near λ_0. Other array geometries that would result in lower grating lobe levels should be considered if element spacing near $\lambda_0/2$ is unachievable. Effects of array element phase quantization were considered since achieving continuous phase shift capabilities compatible with the demands of our application (360 deg total phase shift with 2 deg precision) is impractical. Properties of 2-D rectangular arrays with one- and two-bit phase shifters and their impacts on the imaging

performance of our system were examined. The choice to focus on one- and two-bit phase quantization was a result of our desire to minimize the complexity of the array for fabrication. Beam pointing error was examined over scan ranges of interest for our optical system. For the one-bit phase quantization case, large scan angle errors were encountered, exceeding 25% of the diffraction-limited resolution at the worst locations. Two-bit phase quantization was better, resulting in pointing angle errors of at most 15% of the spot size. As a method of assessing the impact of quantization lobes as a function of scan angle, array main beam efficiency was calculated over the scan range of interest. The average energy fraction in the main beam was comparable for the one- and two-bit phase quantization cases when normalized to the total energy over the $+\mathrm{Az}/+\mathrm{El}$ scanned quadrant. The number and intensity of quantization lobes appearing over the desired scan range was examined. For both the one and two-bit phase quantization cases, a significant number of lobes appeared at intensity levels above the -20 dB threshold. These lobes are potentially harmful to our imaging system since they propagate through the optical system and may compete with the main beam, producing artifacts and false returns. In order to mitigate these problems, other array geometries will be considered in the future. Design and integration of reflectarray elements was briefly considered. Based on currently demonstrated component technology and a desire to simplify element layout, unit cells incorporating patch antennas and one-bit digitally controlled phase shifters are favorable for G-band array implementation.

Acknowledgments

Sincere thanks are due to our colleague Dr. Joseph N. Mait who spent many hours reviewing this article and working with the authors to reorganize and improve clarity throughout.

References

1. R. Appleby and B. H. Wallace, "Standoff detection of weapons and contraband in the 100 GHz to 1 THz region," *IEEE Trans. Antennas Propag.* **55**(11), 2944–2956 (2007).
2. K. B. Cooper et al., "An approach for sub-second imaging of concealed objects using terahertz (THz) radar," *J. Infrared Milli. Terahz Waves* **30**(12), 1297–1307 (2009).
3. A. Luukanen et al., "Rapid beamsteering reflectarrays for mm-wave and submm-wave imaging systems," *Proc. SPIE* **8022**, 80220M (2011).
4. A. S. Hedden, C. R. Dietlein, and D. A. Wikner, "A 220 GHz reflection-type phased array concept study," *Proc. SPIE* **8022**, 80220N (2011).
5. J. D. Albrecht et al., "THz electronics projects at DARPA: transistors, TMICs, and amplifiers," *IEEE MTT-S Int. Microw. Symp. Dig.*, 1118–1121 (2010).
6. T. Vähä-Heikkilä and M. Ylönen, "G-band distributed microelectromechanical components based on CMOS compatible fabrication," *IEEE Trans. Microw. Theory Tech.* **56**(3), 720–728 (2008).
7. W. Deal et al., "THz monolithic integrated circuits using InP high electron mobility transistors," *IEEE Trans. THz Science and Tech.* **1**(1), 25–32 (2011).
8. J. E. Bjarnason et al., "Millimeter-wave, terahertz, and mid-infrared transmission through common clothing," *Appl. Phys. Lett.* **85**(4), 519–521 (2004).
9. K. B. Cooper et al., "THz Imaging Radar for Standoff Personnel Screening," *IEEE Trans. THz Science and Tech.* **1**(1), 169–182 (2011).
10. N. Llombart et al., "Confocal ellipsoidal reflector system for a mechanically scanned active terahertz imager," *IEEE Trans. Antennas Propag.* **58**(6), 1834–1840 (2010).
11. Y. Imaizumi et al., "A study on an onboard Ka-band phased-array-fed imaging reflector antenna," *Antennas and Propagation Society International Symposium, 2002. IEEE*, **Vol. 4**, pp. 144–147 (2002).
12. C. L. Goldsmith et al., "RF MEMs variable capacitors for tunable filters," *Int. J. RF and Microwave Computer-Aided Eng.* **9**(4), 362–374 (1999).
13. C. A. Balanis, "Planar array," Chap. 6.10 in *Antenna Theory Analysis and Design*, 2nd ed., John Wiley & Sons, Inc., New York, pp. 309–321 (1997).
14. N. S. Barker and G. M. Rebeiz, "Optimization of distributed MEMS transmission-line phase shifters—U-band and W-band designs," *IEEE Trans. Microw. Theory Tech.* **48**(11), 1957–1966 (2000).
15. Z. Liu et al., "Broad-band 180° phase shifters using integrated submillimeter-wave Schottky diodes," *IEEE Trans. Microw. Theory Tech.* **53**(9), 2949–2955 (2005).
16. W. J. Dallas, "Phase Quantization—A Compact Derivation," *Appl. Opt.*, **10**(3), 673–674 (1971).
17. R. J. Mailloux, "Array Error Effects," Chap. 7 in *Phased Array Antenna Handbook*, 2nd ed., pp. 353–378, Artech House, Norwood, MA (2005).
18. R. C. Hansen, "Beam steering and quantization lobes," Chap. 2.3 in *Phased Array Antennas*, John Wiley & Sons, New York, pp. 25–33 (1998).
19. C. R. Dietlein, A. S. Hedden, and D. A. Wikner, "Digital reflectarray considerations for terrestrial millimeter-wave imaging," submitted to *IEEE Antennas Wireless Propag. Lett. AWPL*, **11**, 272–275 (2012).
20. H. Kamoda et al., "60-GHz electronically reconfigurable large reflectarray using single-bit phase shifters," *IEEE Trans. Antennas Propag.* **59**(7), 2524–2531 (2011).
21. L. Boccia, G. Amendola, and G. Di Massa, "Performance improvement for a varactor-loaded reflectarray element," *IEEE Trans. Antennas Propag.* **58**(2), 585–589 (2010).
22. B. Mencagli et al., "Design of large mm-wave beam-scanning reflectarrays," *Microwave Conference, 2005 European*, **Vol. 4**, pp. 4 (4–6 Oct. 2005).

Abigail S. Hedden received a BA degree in physics from Carleton College, Northfield, Minnesota, in 2001, and a PhD degree in astronomy from the University of Arizona, Tucson, Arizona in 2007. She was a submillimeter array postdoctoral fellow with the SMA Receiver Lab at the Center for Astrophysics, Cambridge, Massachusetts. In 2010, she joined the Millimeter-wave Sensor Technology team in the Sensors and Electron Devices Directorate at the U.S. Army Research Laboratory in Adelphi, Maryland. Current research interests include developing millimeter-wave/terahertz instrumentation, phenomenology, and experimentation.

Charles R. Dietlein received the BS degree from Seattle Pacific University, Seattle, Washington, in 2003, and the MS and PhD degrees from the University of Colorado, Boulder, Colorado, in 2005 and 2008, respectively, all in electrical engineering. From 2006 to 2010, he was a member of the Quantum Information and Terahertz Technology Project at the National Institute of Standards and Technology in Boulder. Since 2010, he has been with the Sensors and Electron Devices Directorate at the U.S. Army Research Laboratory, Adelphi, Maryland. His research interests include millimeter-wave/terahertz radar, communications, and imaging systems, their associated phenomenology and metrology, and their components including detectors, antennas, and quasi-optical devices.

David A. Wikner is the Millimeter-wave Sensor Technology Team Leader in the Sensors and Electronic Devices Directorate of Army Research Laboratory. He has worked in the area of mmW radar and imaging for 23 years. Most recently his work has involved designing 220 GHz imaging systems for the detection of concealed explosives using electronically adaptable reflectarrays. Other work has involved the exploration of 94 GHz polarimetric phenomenology, system design for passive mmW imaging, and the demonstration of novel device technology suitable for the mmW and THz bands. He has co-chaired the Passive Millimeter Wave Imaging Technology Conference at SPIE's Defense and Security Symposium since 2002. He and his co-authors were awarded the 2011 IEEE Antennas and Propagation Society Harold A. Wheeler Applications Prize Paper Award. He received the MS degree in physics from Stanford University in 1988 and the BA degree in Physics and Mathematics from Ohio Wesleyan University in 1986.

Mercury Cadmium Selenide for Infrared Detection

Kevin Doyle, Craig H. Swartz, John H. Dinan, Thomas H. Myers, Gregory Brill, Yuanping Chen, Brenda L. VanMil, and Priyalal Wijewarnasuriya

Journal of Vacuum Science & Technology B 2013, 31 (3), 5.

Mercury cadmium selenide for infrared detection

Kevin Doyle,[a] Craig H. Swartz, John H. Dinan, and Thomas H. Myers
Texas State University-San Marcos, Materials Science, Engineering and Commercialization Program, 749 N Commanche St, San Marcos, Texas 78666

Gregory Brill, Yuanping Chen, Brenda L. VanMil, and Priyalal Wijewarnasuriya
U.S. Army Research Laboratory—Sensors and Electronic Devices Directorate, 2800 Powder Mill Road, Adelphi, MD 20786

(Received 21 November 2012; accepted 18 March 2013; published 8 April 2013)

Samples of HgCdSe alloys were grown via molecular beam epitaxy on thick ZnTe buffer layers on Si substrates. Two Se sources were used: an effusion cell loaded with 5N source material that produced a predominantly Se_6 beam and a cracker loaded with 6N material that could produce a predominantly Se_2 beam. The background electron concentration in as-grown samples was significantly reduced by switching to the Se cracker source, going from 10^{17}–$10^{18}\,cm^{-3}$ to $3–5 \times 10^{16}\,cm^{-3}$ at 12 K. The concentration remained low even when the cracking zone temperature was lowered to produce a predominantly Se_6 beam, which strongly suggests that a major source of donor defects is impurities from the Se source material rather than Se species. Secondary ion mass spectroscopy was performed. Likely donors such as F, Br, and Cl were detected at the ZnTe interface while C, O, and Si were found at the interface and in the top 1.5 μm from the surface in all samples measured. The electron concentration for all samples increased when annealed in a Cd or Hg overpressure and decreased when annealed under Se. This suggests the presence of native defects such as vacancies and interstitials in addition to impurities. Overall, by switching to higher purity Se material and then annealing under Se overpressures, the background electron concentration was reduced by an order of magnitude, with the lowest value achieved being $9.4 \times 10^{15}\,cm^{-3}$ at 12 K. [http://dx.doi.org/10.1116/1.4798651]

I. INTRODUCTION

Currently, the infrared material of choice is mercury cadmium telluride (MCT). MCT is a ternary alloy with a bandgap that can be tuned from the short wave infrared (SWIR) to the very long wave infrared (VLWIR). High quality MCT can be grown via molecular beam epitaxy (MBE) on bulk lattice-matched cadmium zinc telluride (CZT), with dislocation densities $\sim 10^5\,cm^{-2}$. However, bulk CZT has a maximum area of roughly 50 cm^2, making it unsuitable for the manufacture of a large area focal plane array (FPA). MCT can also be grown by MBE on silicon (Si) with a cadmium telluride (CdTe) buffer layer. Si wafers are available in diameters at least as large as 10 in., but the 19% lattice mismatch between MCT and Si results in large dislocation densities that limit device performance, particularly for long wave infrared (LWIR) MCT.[1]

An alternative material is mercury cadmium selenide (MCS). Like MCT, MCS is a ternary alloy with a bandgap tunable from the SWIR to the VLWIR. MCS belongs to a family of materials with lattice parameters near 6.1 Å. GaSb, another member of this family, is now available in wafers with a diameter of 4 in., with 6 in. diameter GaSb wafers currently under development. Additionally, this 6.1 Å family also includes materials with band gaps suitable for detection applications in the visible and ultraviolet spectral ranges. Therefore, one could conceivably create a device made from lattice-matched materials capable of sensing from the ultraviolet to the VLWIR on a single chip.[2]

One obstacle to the use of MCS for devices has been the large background electron concentration that has been reported for this material. Despite not being intentionally doped, MCS samples typically had electron concentrations greater than $10^{17}\,cm^{-3}$ at 77 K whether in the form of bulk samples[3] or of epitaxial layers deposited by MBE.[4] The electron concentration remained high with little variation even at temperatures as low as 4 K, suggesting the presence of a shallow donor level located near or within the conduction band. The background concentration could either be reduced or increased by annealing under various conditions, suggesting the presence of native defects such as vacancies and interstitials.[5] Sources of these donor defects need to be identified so that a process to eliminate them either during growth or through postgrowth annealing can be developed.

II. EXPERIMENT

MCS samples were grown via MBE on Si substrates with zinc telluride (ZnTe) buffer layers.[6] The samples were grown in an ultrahigh vacuum MBE chamber made by DCA Instruments. The substrates were mounted on molybdenum blocks with colloidal graphite. Immediately prior to loading, the ZnTe/Si substrates were etched in a 0.2% Br:Methanol solution for 30 s followed by a brief methanol rinse, a 10 s etch in 10% HCl, a 60 s rinse in deionized water, and then blown dry with N_2. Once loaded, the substrate was heated under a Te overpressure while monitored *in situ* by reflection high energy electron diffraction to remove any remaining oxides prior to growth. Clips held the edges of each substrate, and the thickness (and therefore the growth rate) of

[a]Electronic mail: Kevin.doyle.30.ctr@mail.mil

each sample could be determined by measuring the "step" created by the clip with a profilometer.

MCS samples were grown using elemental mercury (Hg), cadmium (Cd), and selenium (Se) sources. The beam equivalent pressure (BEP) emanating from all sources was measured with a beam flux monitor (BFM) consisting of a nude ion gauge placed directly in the path of flux. Quadruple distilled Hg was supplied by a 600cc valved effusion cell and Cd with 99.999% (5N) purity was supplied by a 400g SUMO Cell, both made by Applied EPI. Initially, a Model VSb110 effusion cell made by ADDON was used to supply 5N Se. However, Se vapor consists of many polyatomic species from Se_2 to Se_8, and at this effusion cell's typical operating temperature of 325 °C (598 K), the predominant species of Se flux was Se_6 (uncracked Se).[7]

The Se effusion cell was replaced with a Mark V Selenium Valved Cracker made by Veeco, which directed the Se vapor through a cracking zone, which could be heated up to 800 °C (1073 K) to produce a predominantly Se_2 beam (cracked Se).[7,8] Differences in the ionization efficiencies of the Se atomic species resulted in different sensitivities of the BFM depending on which species were dominant. For a fixed reservoir temperature of 250 °C and a fixed valve position (to maintain a constant flux), the BEP measured for the cracker source was found to vary with the cracking zone temperature, tracking with the data found in Ref. 7. This suggests that the Se flux transitions from predominantly Se_6 to predominantly Se_5 at around 650 K and then to predominantly Se_2 near 900 K (Fig. 1). The Se BEP measured for the typical cracking zone temperature of 800 °C was found to be close to a factor of two lower than at the typical effusion cell temperature of 325 °C for the same amount of exiting Se reflecting a difference in ionization energy for the various species. This correction factor was applied to Se BEP from the cracker source when comparing the two sources. While the effusion cell was loaded with 5N purity Se, the cracker was loaded with Se with 99.9999% (6N) purity.

Samples were grown at different temperatures using various Cd to Se and Hg to Se BEP ratios. Substrate temperature was measured with a pyrometer as well as a thermocouple on the sample manipulator. However, samples grown using the cracker source presented some difficulty in measuring the substrate temperature. Heat from the high temperature cracking zone was reflected off the substrate into the pyrometer, making it more difficult to obtain an accurate measurement. An estimate of substrate temperature was determined comparing thermocouple and pyrometer temperature readings.

Cut off wavelength was determined via transmittance measurements using a Fourier transform infrared spectroscopy and the molar fraction of CdSe in the MCS alloy, or x-value, was determined from this measurement using the relationship between band gap and x-value developed by Summers and Broerman.[9] Hall measurements were performed over a range of temperatures from 4 to 300 K, on samples subjected to various postgrowth anneals. Finally, secondary ion mass spectroscopy (SIMS) was performed by the Charles Evans Analytical Group.

III. RESULTS AND DISCUSSION

A. Growth parameters

Due to the very low sticking coefficient of Hg, samples were grown with large Hg BEPs ($\sim 10^{-4}$ Torr). For a fixed substrate temperature and Hg overpressure, the growth rate varies linearly with Se BEP (Fig. 2), and for a fixed Se BEP, the x-value can be controlled by the Cd/Se BEP ratio (Fig. 3). It was found that samples grown with cracked Se had a higher x-value than samples grown with uncracked Se with the same Cd/Se ratio, suggesting greater incorporation of Cd with Se_2 than Se_6. Growth rates began to decrease at approximately 130 °C for the valved source and 150 °C for the cracker source (Fig. 4). The optimal MBE substrate temperature for MCS grown with an Hg BEP of 2.5×10^{-4} Torr was ~100 °C. This is lower than the optimal temperature for MCT with a similar Hg BEP (~185 °C), most likely due to the higher vapor pressure of Se compared to Te.

B. Hall measurements

The electron concentration versus temperature was measured for samples grown with both Se sources using Hall

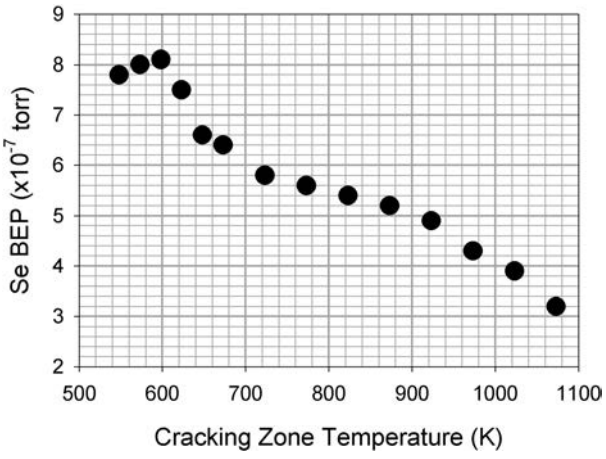


FIG. 1. Se BEP vs cracking zone temperature for a fixed Se reservoir temperature of 250 °C and valve position of 150 mils.

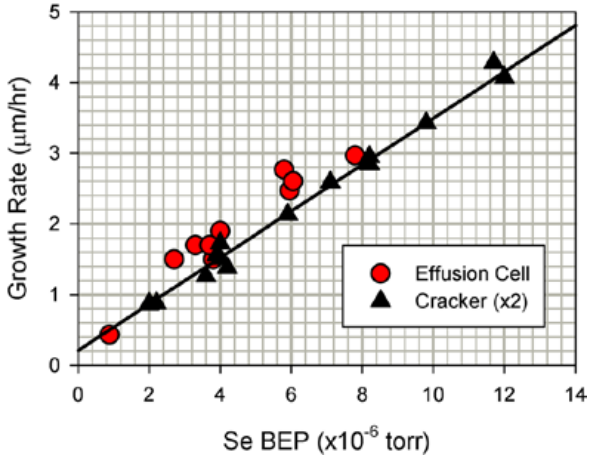


FIG. 2. (Color online) Growth rate vs Se BEP for both the effusion cell (Se_6) and the cracker source (Se_2).

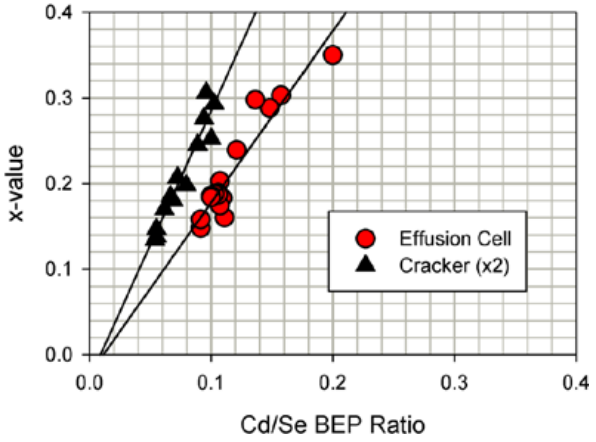


FIG. 3. (Color online) Cd composition vs Se/Cd BEP ratio for both the effusion cell (Se_6) and the cracker source (Se_2).

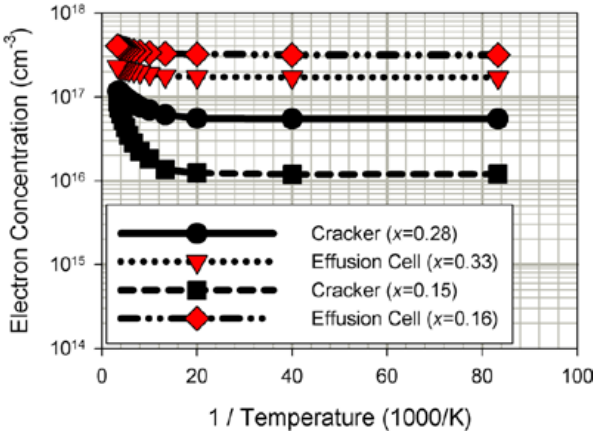


FIG. 5. (Color online) As-grown electron concentration vs temperature for MCS samples grown with both Se sources.

Effect measurements with a standard magnetic field of 0.1 T. A previous study of MCS grown via MBE using uncracked Se reported little variation in electron concentration with temperature below 100 K, with the electron concentration remaining in the 10^{17}–$10^{18}\,cm^{-3}$ range at temperatures as low as 30 K.[4] This was consistent with samples grown by the effusion cell, but MCS samples grown with the cracker source exhibited a temperature dependency with as-grown 12 K electron concentrations in the 10^{16}–$10^{17}\,cm^{-3}$ range (Fig. 5). The relative lack of carrier freeze-out and the lower electron concentrations with the cracker source suggest the presence of donors with energy levels located near or within the conduction band that were significantly reduced by switching to the Se cracker source.

Two differences between the Se effusion cell vs the Se cracker source that could explain the lower electron concentrations are the different atomic species of the Se beam ($\sim Se_6$ vs $\sim Se_2$) and the higher purity source material in the cracker source (5 N vs 6 N). MCS samples were grown using the 6 N Se in the cracker source, with the cracking zone temperature lowered to 325 °C to produce an uncracked predominantly Se_6 beam. The electron concentration remained low even when the cracking zone temperature was reduced to 325 °C (Fig. 6), strongly suggesting that the reduction in concentration was due to the higher purity source material and not the predominantly Se_2 flux. Electron mobility for the MCS samples increased as the x-value decreased (Fig. 7).

A prior study of HgSe annealed under Hg and Se suggested that Hg interstitials (n-type), Se vacancies (n-type), and Hg vacancies (p-type) were possible native defects in MCS.[5] These possibilities were investigated by subjecting MCS samples to separate 24 hour, 250 °C anneals under vacuum, a Hg overpressure, a Cd overpressure, or a Se overpressure in sealed quartz ampoules. The electron concentration always increased when annealed under Hg or Cd, and the electron concentration was reduced for samples grown with the cracker source and then annealed under Se. No significant changes were observed for samples annealed under vacuum (Fig. 8). RBS measurements indicated an increase in x-value when annealed under Cd, but no significant change in composition with the other anneals.

The lowest 12 K electron concentration achieved was for a Se-annealed sample with an x-value of 0.15. The as-grown 12 K concentration of $1.2 \times 10^{16}\,cm^{-3}$ was reduced to $9.4 \times 10^{15}\,cm^{-3}$ after annealing under Se. Overall, switching

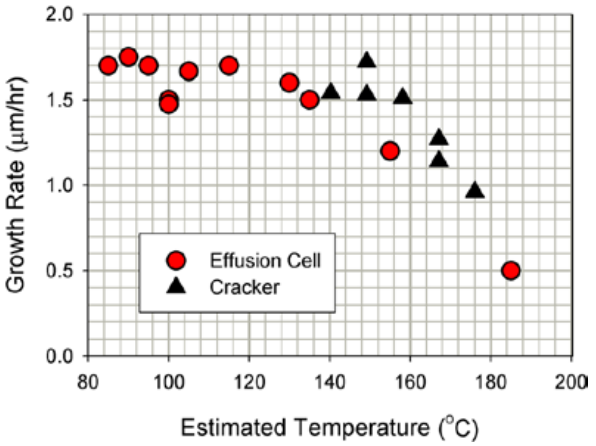


FIG. 4. (Color online) Growth rate vs estimated substrate temperature with a fixed Se BEP for both the effusion cell (Se_6) and the cracker source (Se_2).

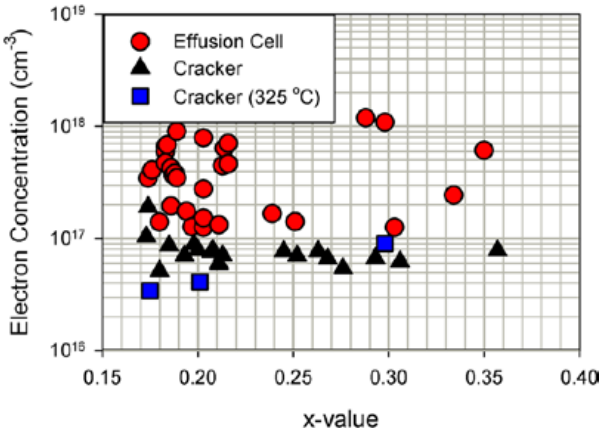


FIG. 6. (Color online) As-grown electron concentration at 77 K vs x-value for the effusion cell (Se_6), the cracker source at typical operating temperatures (Se_2), and cracker source with the cracking zone temperature at 325 °C (Se_6).

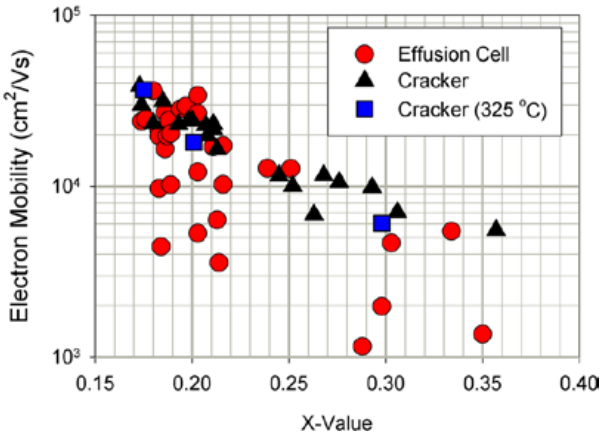


FIG. 7. (Color online) As-grown electron mobility at 77 K vs x-value for the effusion cell (Se_6), the cracker source at typical operating temperatures (Se_2), and cracker source with the cracking zone temperature at 325 °C (Se_6).

to the higher purity Se and then annealing under a Se overpressure typically reduced the background electron concentration by an order of magnitude at 77 K. Direct studies of vacancies and interstitials are being performed through positron annihilation spectroscopy and Rutherford backscattering channeling spectroscopy, respectively. These results will be presented at a later date.

C. SIMS measurements

SIMS measurements were conducted on an MCS sample grown with the effusion cell (thickness = 7.3 μm), an MCS sample grown with the cracker source (thickness = 3.5 μm), and a HgSe sample grown with the effusion cell (thickness = 3.4 μm) in order to identify unintentional impurities (Fig. 9). For all three samples, group VII elements such as Br, Cl, and F were detected at the interface with the ZnTe buffer layer. Br and Cl could be introduced during the substrate preparation process and could serve as n-type dopants if they substituted group VI Se lattice sites. C and O were detected in a region approximately 1.5 μm thick at the surface of the MCS and at the interface between MCS and ZnTe. The source of these impurities and whether they are electrically active in MCS needs to be established. Two other contaminants listed in the Cd source material certificate of analysis were group VI S and group IV Si, both of which were detected in all samples but significantly reduced in the HgSe sample where the Cd source was not

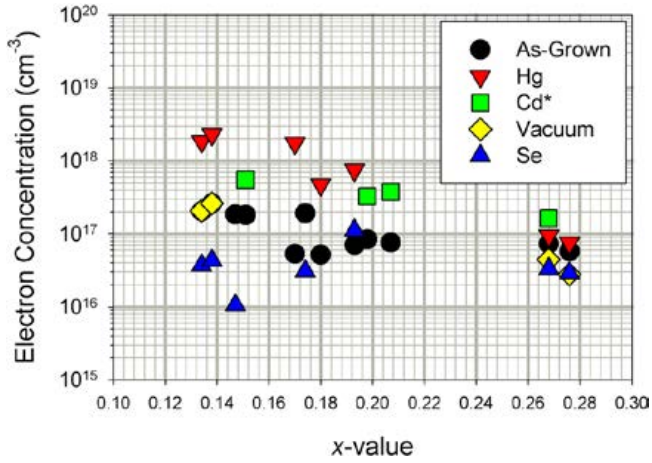


FIG. 8. (Color online) 77 K electron concentration both as-grown and after annealing under various overpressures of Hg, Se, Cd, and under vacuum. All anneals were performed in quartz ampoules pumped down to $\sim 10^{-5}$ Torr, then sealed and kept in a furnace at 250 °C for 24 h, followed by a 3 h cool-down. *x-value listed is prior to annealing.

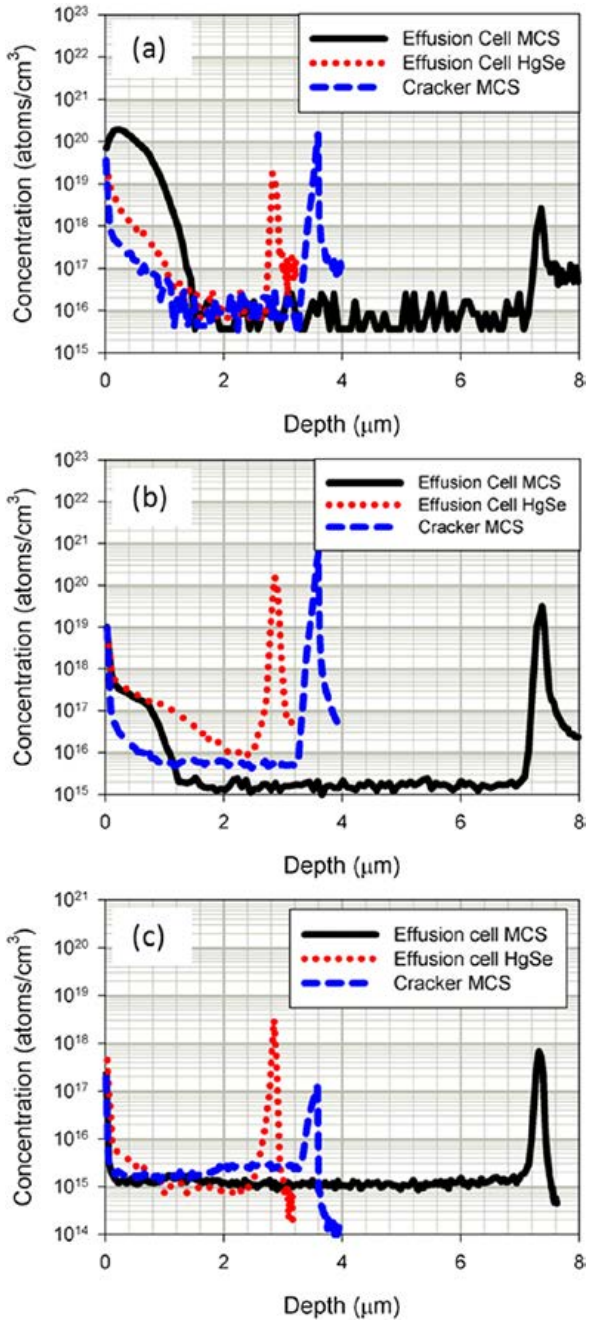


FIG. 9. (Color online) SIMS results of a MCS sample grown with the effusion cell (thickness = 7.3 μm), MCS sample grown with the cracker source (thickness = 3.5 μm), and an HgSe sample grown with the effusion cell (thickness = 3.4 μm) for (a) carbon (b) oxygen and (c) bromine. SIMS measurements were performed by the Charles Evans Analytical Group.

used—strongly suggesting they are contaminants in the Cd source material.

Unfortunately none of the SIMS measurements to date have differed significantly between the MCS samples grown with the 5N and 6N Se source material, and so the impurities that were reduced by switching to higher purity Se source material have yet to be identified.

IV. SUMMARY AND CONCLUSIONS

MCS samples were grown via MBE on ZnTe/Si substrates using two different Se sources: an effusion cell loaded with 5N source material that produced a predominantly Se_6 beam and a cracker source loaded with 6N source material that could be varied to study other Se polyatomic species. Samples grown with the Se_2 had greater x-values with lower Cd/Se BEP ratios, suggesting greater Cd incorporation with Se_2. The growth rate began to decrease when the substrate temperature was raised above $\sim$130 °C under an Se_6 flux and $\sim$150 °C under an Se_2 flux. The optimal substrate temperature for MCS grown with the effusion cell was found to be $\sim$100 °C for an Hg BEP of 2.5×10^{-4} Torr—lower than the optimal temperature for MCT growth with a similar Hg overpressure ($\sim$185 °C).

Electron concentrations remained high even at low temperatures, with as-grown 12 K concentrations ranging from 10^{17} to 10^{18} cm^{-3} for samples grown with 5N Se source material and 10^{16}–10^{17} cm^{-3} for samples grown with 6N Se. Impurities can produce energy levels located in the conduction band of narrow-gap materials, such as In dopants in MCT. As a result, these impurities do not freeze-out at lower temperatures and the concentration remains high even at temperatures as low as 4 K.[10] The fact that the electron concentration remains high in MCS even at low temperatures indicates the presence of energy levels in the conduction band similar to MCT, and the fact that the 12 K concentration was lower for 6 N Se strongly suggests that impurities are introduced from contaminants in the Se source material.

SIMS measurements detected impurities which could be acting as donors, the most prevalent of which was C. Br and Cl were detected at the MCS/ZnTe interface, suggesting they could be introduced by the substrate preparation process. Significant levels of C and O were detected at the MCS/ZnTe interface and in the top 1.5 μm of the MCS layer from the surface. Further measurements are required to determine how these impurities are introduced, whether they are electrically active, and how they can be eliminated.

The MCS electron concentration could also be changed by postgrowth annealing. Anneals under Hg and Cd overpressures raised the electron concentration, while anneals under Se or vacuum lowered the electron concentration. This would suggest the presence of native defects such as interstitials and vacancies in addition to the background impurities. The identity of these native defects and an annealing process to eliminate them is currently under investigation.

If MCS is to be used for LWIR applications, the background electron concentration needs to be reduced to at most $\sim 10^{15}$ cm^{-3}(assuming a similar lifetime to MCT). Switching from 5 N Se to 6 N Se reduced the electron concentration from 10^{17}–10^{18} cm^{-3} to 3–5 $\times$ 10^{16} cm^{-3}, suggesting that the background concentration could be further reduced by using 7 N or higher purity Se source material. Further study of native defects present in MCS is required so that a process for removing them through postgrowth annealing can be optimized. Once the background electron concentration has been fully minimized, p-type doping of MCS can be developed so that MCS device layers can be produced.

ACKNOWLEDGMENTS

Portions of this work were supported by the US Army Research Laboratory and the US Army Research Office under Contract/Grant Nos. W911NF-10-2-0103 and W911NF-10-1-0335. Additional funding was provided by the Air Force Office of Scientific Research. SIMS measurements were performed by the Charles Evans Analytical Group under contract number C0CGM100.

[1]G. Brill, Y. Chen, P. Wijewarnasuriya, and N. Dhar, Proc. SPIE **7419**, 74190L-1 (2009).
[2]G. Brill, Y. Chen, and P. Wijewarnasuriya, J. Electron. Mater. **40**, 1679 (2011).
[3]C. R. Whitsett, J. G. Broerman, and C. J. Summers, Semicond. Semimetals **16**, 53 (1981).
[4]Y. Lansari, J. W. Cook, and J. F. Schetzina, J. Electron. Mater. **22**, 809 (1993).
[5]K. Kumazaki, E. Matsushima, and A. Odajima, Phys. Status Solidi A **37**, 579 (1976).
[6]Y. Chen, S. Simingalam, G. Brill, P. Wijewarnasuriya, N. Dhar, J. J. Kim, and D. Smith, J. Electron. Mater. **41**, 2917 (2012).
[7]A. Olin, B. Nolang, E. G. Osadchii, L.-O. Ohman, and E. Rosen, *Chemical Thermodynamics of Selenium* (Elsevier Press, Amsterdam, the Netherlands, 2005), p. 89.
[8]Selenium Valved Cracker Mark V-500 CC User Guide, Veeco (2011).
[9]C. J. Summers and J. G. Broerman, Phys. Rev. B **21**, 559 (1980).
[10]A. Rogalski, Rep. Prog. Phys. **68**, 2267 (2005).

Electromagnetic Modeling and Design of Quantum Well Infrared Photodetectors

Kwong-Kit Choi, Murzy D. Jhabvala, David P. Forrai, Augustyn Waczynski, Jason Sun and Robert Jones

IEEE Journal of Selected Topics in Quantum Electronics 2013, 19 (5), 3800310 (10 pp.)

Electromagnetic Modeling and Design of Quantum Well Infrared Photodetectors

Kwong-Kit Choi, *Fellow, IEEE*, Murzy D. Jhabvala, David P. Forrai, *Senior Member, IEEE*, Augustyn Waczynski, Jason Sun, and Robert Jones

***Abstract*—The quantum efficiency (QE) of a quantum well infrared photodetector (QWIP) is historically difficult to predict and optimize. This difficulty is due to the lack of a quantitative model to calculate QE for a given detector structure. In this paper, we found that by expressing QE in terms of a volumetric integral of the vertical electric field, the QE can be readily evaluated using a finite element electromagnetic solver. We applied this model to all known QWIP structures in the literature and found good agreement with experiment in all cases. Furthermore, the model agrees with other theoretical solutions, such as the classical solution and the modal transmission-line solution when they are available. Therefore, we have established the validity of this model, and it can now be used to design new detector structures with the potential to greatly improve the detector QE.**

***Index Terms*—Electromagnetic field modeling, infrared detector, quantum efficiency (QE).**

I. Introduction

A WELL-known property of quantum well infrared photodetector (QWIP) materials is their lack of optical absorption under normal incident condition. Consequently, each detector in an array is outfitted with a light-coupling structure for detection. Although the coupling designs are usually guided by certain physical principles, their exact quantum efficiencies (QEs) are not always predictable. In the cases where analytical models do exist, they inevitably contain assumptions or simplifications, such as infinite detector sizes or infinite metal conductivities, that render the predictions inaccurate. The lack of a quantitative model has thus far prevented the QE improvement, and as a result, the QWIP technology has generally been regarded as a low QE technology.

One attempt in the past to yield a quantitative prediction is through rigorous electromagnetic (EM) modeling [1]–[4], but the success was rather limited. Recently, we showed that by expressing QE as an integral of the vertical electric field E_z, its value can be readily and reliably computed by a commercial finite element EM solver [5]. We applied this model to several detector structures and obtained quantitative agreements with experiments. In this paper, we expand this study to include all known geometries in the literature and report its finding. For completeness, we also report those described in [5]. The structures surveyed in this paper are: edge-coupled QWIPs, linear- and cross-grating QWIPs, random-grating QWIPs, corrugated-QWIPs of prism and pyramidal geometries, enhanced-QWIPs, quantum grid infrared photodetectors, plasmonic-enhanced QWIPs, and photonic-crystal-slab QWIPs. In addition, we also compare the numerical solutions with the analytical classical solutions in the cases of edge-coupled detectors and corrugated QWIPs and the modal transmission-line solutions in the case of quantum grid infrared photodetectors. The agreements turn out to be satisfactory in all examples. With a verified model, we use it to design and optimize new detectors. The result shows that the theoretical QE can reach 70–80% in some cases without an antireflection (AR) coating. Therefore, QWIPs can have a potential for high QE.

Manuscript received July 17, 2012; accepted August 28, 2012.

K.-K. Choi and J. Sun are with the U.S. Army Research Laboratory, Adelphi, MD 20833 USA (e-mail: kwong.k.choi.civ@mail.mil; guifu.n.sun.civ@mail.mil).

M. D. Jhabvala and A. Waczynski are with the NASA Goddard Space Flight Center, Greenbelt, MD 20771 USA (e-mail: Murzy.D.Jhabvala@nasa.gov; augustyn.waczynski-1@nasa.gov).

D. P. Forrai and R. Jones are with L-3 Cincinnati Electronics, Mason, OH 45040 USA (e-mail: Dave.Forrai@L-3Com.com; Robert.Jones@L-3Com.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTQE.2012.2216861

II. EM Model

Previously, we have established that by performing finite-element EM computation to the following expression, the absorption QE, labeled as η, of any detector geometry can be predicted [5]:

$$\eta = \frac{n\alpha}{\mathrm{AE}_0^2} \int_V |E_z(\vec{r})|^2 \, d^3r \tag{1}$$

where n is the material refractive index of the detector material, α is the absorption coefficient for vertically polarized light, A is the detector area, E_0 is the incident electric field from the air, V is the detector active volume, E_z is the self-consistent vertical electric field. Equation (1) states that QE can be calculated from the volume integral of $|E_z|^2$ in the presence of a finite α.

Since E_0 and E_z are linearly proportional to each other, E_0 can be set arbitrarily, and the only input parameter in (1) is the wavelength-dependent $\alpha(\lambda)$, which can be calculated based on the material layer structure [6]. For a known $\alpha(\lambda)$, there will be no more free parameters, and the value of $\eta(\lambda)$ is uniquely and unambiguously determined. To solve E_z numerically, we use a commercial finite element solver. In addition to η, we also define another quantity, the external QE or η_{ext}, which is QE $\times$ pixel area fill factor ($\equiv A/A_{\mathrm{pitch}}$).

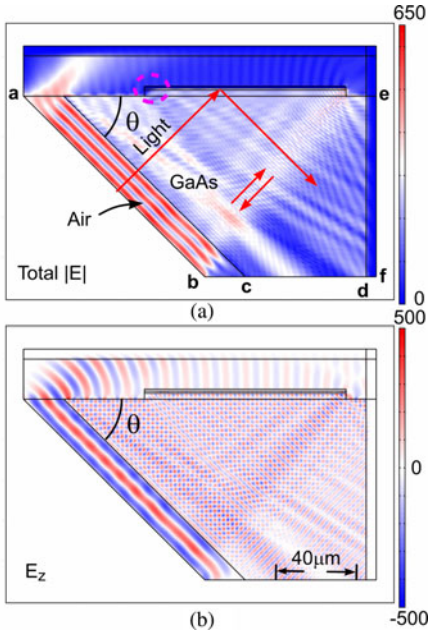


Fig. 1. The figure shows a $\theta = 45°$ edge-coupled QWIP. The figure also shows $|E|$ and E_z distributions obtained from EM modeling at $\lambda = 10\ \mu$m with $E_0 = 377$ V/m.

III. EM Solutions

To assess the reliability of our model, we first apply (1) to a light-coupling scheme that has a classical solution. It is the edge coupling via a 45° polished facet [7]. Fig. 1 shows the detector geometry.

The classical solution η_c for this geometry is

$$\eta_c(45°) = T_{\rm sub}\frac{\cos 45°}{2}\left[1 - \exp\left(-\alpha(45°)L(45°)\right)\right] \qquad (2)$$

where $T_{\rm sub} = 4n/(1+n)^2$ is the transmission coefficient of the GaAs substrate; $\alpha(45°) = \alpha \sin^2 45°$ is the material absorption coefficient at 45°; $L(45°) = 2t/\cos(45°)$ is the optical pathlength inside the detector for two passes of light, and t is the active material thickness. In (2), the first cos(45°) accounts for the smaller projected detector area in the direction of light and the factor $\frac{1}{2}$ accounts for the fact that only half of the light, the transverse magnetic (TM) mode, is coupled. In this classical model, if α is constant with respect to wavelength, η_c will also be wavelength independent. For a typical α of 0.15 μm^{-1} and t of 3 μm, $\eta_c(45°)$ can be conveniently calculated to be 12.3%.

Since under the present detector geometry, the direction perpendicular to the plane of incidence is invariant, we reduce (1) to 2-D, in which

$$\eta(\lambda) = \frac{n\alpha}{2dE_0^2}\int_X |E_z(r,\lambda)|^2\, d^2r \qquad (3)$$

where the factor $\frac{1}{2}$ accounts for one coupled polarization (the TM mode), $d = 100\ \mu$m is the assumed detector linear dimension in the horizontal direction, $E_0 = 377$ V/m, and $X = d \times t$ is the detector cross-sectional area. The present example consists of a 3-μm active layer on top of a GaAs substrate, a 1-μm top GaAs contact layer, and a 1-μm gold contact layer. The

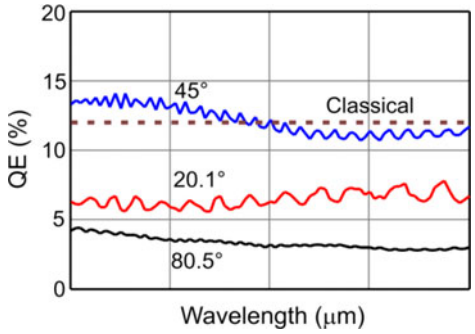


Fig. 2. The figure shows the calculated QE as a function of incident λ at three different edge angles. The dashed line shows the classical value at 45°.

properties of gold are represented by a wavelength-dependent complex refractive index. The value of α is again taken to be 0.15 μm^{-1}, independent of λ. Although α is a constant in this model, the computed E_z distribution and, thus, η nevertheless are λ-dependent because of the wave nature of light. The detector is placed near the edge of a polished GaAs substrate, and the length scale is indicated in Fig. 1(b).

The color plot in Fig. 1(a) shows the absolute magnitude of the total E field. Most of the detector active region is found to be uniformly illuminated at this angle except those near the corners. On the left, the circled region is shadowed by the substrate in the front, which prevents the light incident directly into the detector. On the right, the reflection surfaces at the top, at the side, and at the substrate, form a Fabry-Perot etalon and produce the rapidly varying intensity. The E_z component, which is responsible for absorption, is plotted in Fig. 1(b). Due to the interference between the incident light and the reflected light from the top surface, a standing wave of local maxima and minima is established.

By integrating $|E_z|^2$ within the active cross section according to (3), η can be evaluated. The result is shown in Fig. 2, along with two other edge angles. The QEs only weakly depend on λ, partially validating the classical assumption. At 45°, η varies with two distinct frequencies. The slow variation is due to the gradual shift of the standing wave along the vertical axis as λ changes. It is centered around ∼12.3%, in agreement with the classical model. The higher frequency oscillations are the Fabry–Perot oscillations produced at the right-hand corner. This example shows that the EM model, being a numerical solution to Maxwell equations, is not only consistent with the classical model but also accounts for all other optical effects neglected in the classical model [7].

For 3-D modeling, we first examine the external QE of a linear grating that was used in a polarization detection experiment [8]. The detector area A is 18.6 × 18.6 μm^2 and the pixel pitch area $A_{\rm pitch}$ is 20 × 20 μm^2. The detector structure is shown in Fig. 3(a). The calculated E_z distribution at the center cross section is shown in Fig. 3(b) for $E_0 = 377$ V/m perpendicular to the grating lines. The corresponding (unpolarized) $\eta_{\rm ext}$ is shown in Fig. 3(c) based on the α spectrum calculated from the material structure. The theoretical peak $\eta_{\rm ext}$ is 14.1% for this optical polarization. Experimentally, the peak responsivity

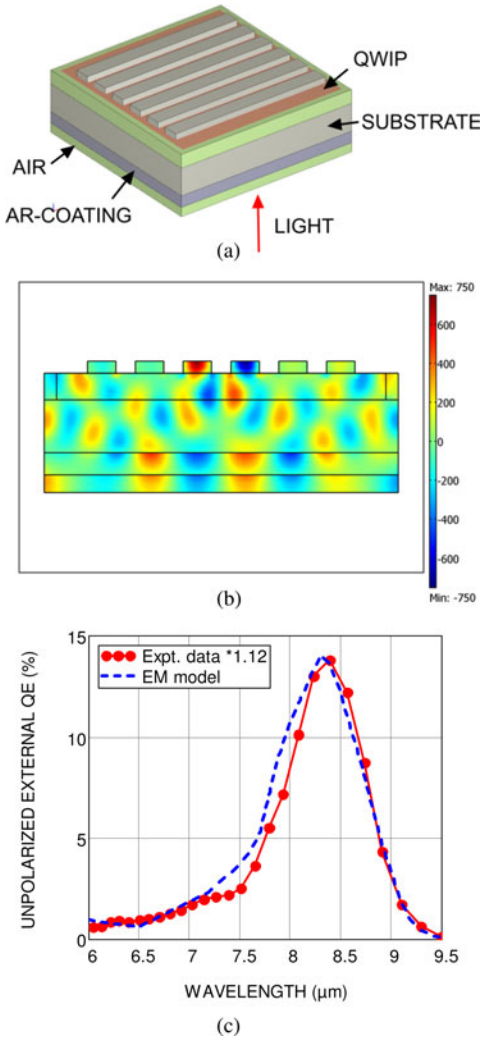


Fig. 3. (a) Grating having 2.7-μm grating period and 0.68-μm depth. (b) E_z distribution at the center cross section at $\lambda = 8.3$ μm. (c) Calculated (dashed curve) and measured QE (solid curve).

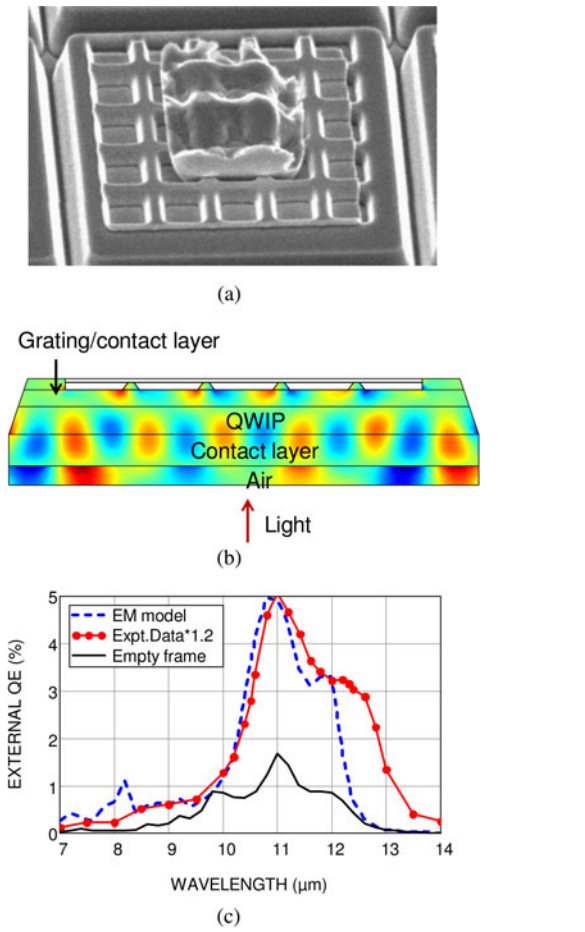


Fig. 4. (a) Experimental grating structure. (b) Modeled structure and the E_z distribution in the center cross section at $\lambda = 11.0$ μm. (c) Experimental data scaled by a factor of 1.2 (solid curve with circles), the calculated QE of the grating (dashed curve), and the calculated QE without the grid lines (solid curve).

R for this polarization was measured to be 0.48 A/W. With the reported photoconductive gain g of 0.57, the deduced $\eta_{\rm ext}$ is 12.5%, which is only slightly lower than the prediction by a factor of 1.12. Besides, the agreement in the QE magnitude, the calculated lineshape also matches well with the measured spectrum as shown in Fig. 3(c). Therefore, the EM model successfully explains the light-coupling characteristics of a linear grating.

Fig. 4(a) shows the next example of a cross grating [9]. This grating-QWIP consists of a 1.5-μm active QWIP material, a 1.5-μm top contact layer, a 1.5-μm bottom common contact layer, and a 0.1-μm etch stop layer. The pixel pitch is 25 μm. To model the experimentally realized structure, we set the pixel dimension to be 25 μm at the base and 23 μm at the mesa top. The square frame around the grating grid has a height of 0.6 μm and a width of 2 μm at the top. Inside the frame, the grating height is 0.4 μm and the grating period is 4.0 μm. The widths of the base and tip of the grid lines are 0.9 and 0.3 μm, respectively. Fig. 4(b) shows the center cross section of the modeled structure without the top metal cover layer. This top metal cover is replaced by the perfect electric conductor (PEC) boundary condition in the model. The calculated peak α of this material is 0.10 μm^{-1} at 11.3 μm, and the 50% cutoff is at 12.1 μm.

Experimentally, the peak conversion efficiency CE ($\equiv \eta \times g$) is measured to be 2.41% at 0.78 V. Together with an estimated $g = 0.56$, which is scaled from a 60-period structure, the external QE is 4.31%. This experimental QE is 20% lower than the 5.0% predicted from the EM model shown in Fig. 4(c). By multiplying the experimental value with a factor of 1.2, a good match of the spectral lineshape is obtained below 12 μm. The theoretical QE beyond 12 μm is limited by the assumed material absorption, which cuts off at 12.1 μm. Overall, the present EM model is adequate in explaining the grating efficiency after taking the detailed detector structure into account. As seen in Fig. 4(a), the present mesa has substantially inclined sidewalls. By repeating the calculation without the grid lines, we found the sidewall reflection contributes to about 33% of the QE near the peak wavelengths.

Fig. 5(a) shows the geometry of a random grating invented by Brill and Sarusi [10], [11]. The pixel size is 28 μm $\times$ 28 μm and the grating height is 0.67 μm. The thickness of the active

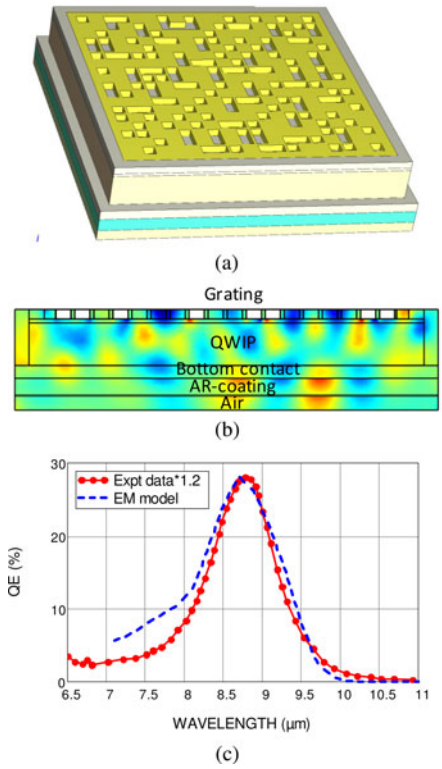


Fig. 5. (a) Random grating QWIP. (b) E_z distribution at the center cross section at $\lambda = 8.7$ μm. (c) Calculated (dashed curve) and measured QE (solid curve).

material is 3.28 μm and the thickness of the bottom contact is 0.95 μm. The substrate is completely removed and an AR-coating is applied to the detector. Fig. 5(b) shows the layer structure. The QWIP material contains 50 periods of 55-nm-thick AlGaAs barriers and 4.9-nm-thick quantum wells doped to 5×10^{17} cm^{-3}, with which the peak α is calculated to be 0.104 μm^{-1}. Experimentally, the peak detector responsivity is measured to be 0.6 A/W, at which bias, the gain is 0.32. The measured QE is thus 23.4%. On the other hand, the calculated QE is 27.9%, which is 20% larger than the experimental value. Fig. 5(c) shows the calculated and the measured lineshapes, which are in satisfactory agreement.

Fig. 6(a) shows another grid structure, which is known as the enhanced-QWIP [12]. But different from a cross grating, the active detector material in this case is etched to form the grid structure and the radiation is incident directly onto the grid. Fig. 6(a), (b), and (c) show the detector structure, the E_z field distribution at λ of 9.2 μm, and the calculated QE for a constant α of 0.21 μm^{-1}, respectively. This value of α is calculated from the material structure at the absorption peak. Fig. 6(c) also plots the measured QE of two different detectors at their peaks. The theory and experiment are in agreement with each other.

Similar to the enhanced-QWIP, the structure of a quantum grid infrared photodetector (QGIP) is shown in Fig. 7. It consists of a linear array of grid lines of active materials. In the modeled structure, the top gold layer thickness t_m is 0.2 μm, the top contact layer thickness t_c is 0.1 μm, the active layer thickness t_a is 1.18 μm, the bottom contact layer thickness t_b is 1.83 μm,

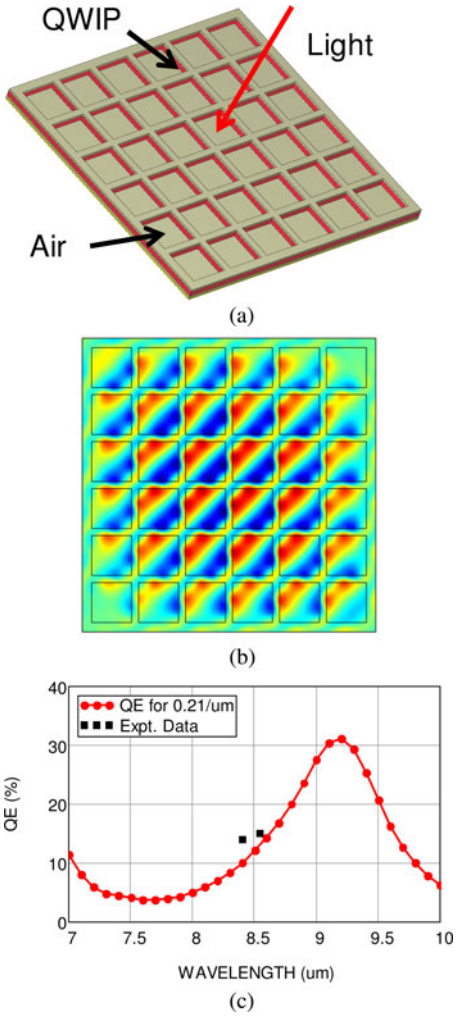


Fig. 6. (a) E-QWIP geometry. (b) Color plot of E_z distributions at 9.2-μm incident wavelength. The displayed plane is located at the center the grid layer. The grid has a period of 8 μm, a strip width of 1.2 μm, and a strip height of 1.43 μm. The back of the grid is coated with a 0.4-μm layer of gold. (c) Calculated (curve) and the observed QEs (squares) are shown.

and the separation among the grid lines s is 4.65 μm. The width of the grid line w, which determines the detection wavelength λ_p of the detector, varies among different detectors. The substrate is assumed to be thick such that the light enters into the detector from the substrate side classically with a transmission coefficient $T_{\rm sub}$.

Previously, λ_p of the detector had been designed using the 2-D modal transmission-line (MTL) method [13]. Fig. 8 shows that by choosing w appropriately, the detectors can be made to detect at each integral wavelengths from 8 to 15 μm. In that modeling, a constant relative dielectric constant of $9.722 + i$ had been adopted for the active layer, and the refractive index of GaAs was set to be 3.118. Based on the present finite element method (FEM), we repeat the same calculation using the identical detector parameters. Fig. 8 shows that the present FEM solution agrees closely with the previous MTL solution, confirming both methods. The small differences in the shorter wavelengths could be due to the truncation of certain infinite Fourier series in the MTL approach.

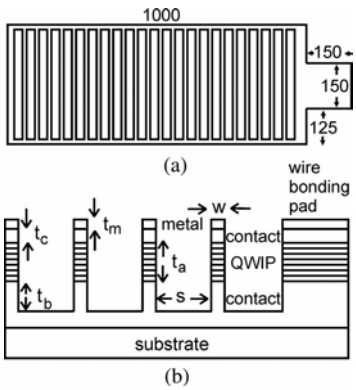


Fig. 7. (a) Top view of a QGIP. The numbers are dimensions in microns. (b) Side view.

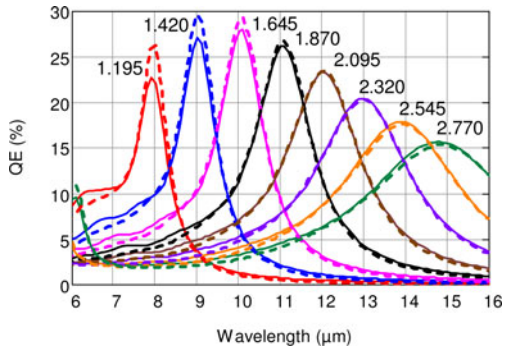


Fig. 8. QE spectra calculated based on (a) FEM (solid curves) and (b) MTL method (dashed curves) having the same detector parameters. The numbers are w in microns.

In Fig. 9, we plot the experimental coupling efficiency of two QGIP detector elements [13], which is defined as the ratio of the responsivities of the grid and the edge-coupled detector. This ratio cancels out the material α dependence and it is directly proportional to the detector QE with a constant α. In Fig. 9, we also show the theoretical QE with a constant α of 0.20 μm^{-1} and $n = 3.239$ using the FEM model. Since the substrate of these detectors is about 200-μm thick, the classical substrate transmission is applicable in this calculation. The calculated spectrum in Fig. 9 explains the overall lineshape in the experiment, although one cannot compare the absolute QE in this plot. Note that the experimental spectrum is averaged over 180 narrow and long (400-μm) grid lines. The width fluctuations along the grid lines are expected to broaden the QE peak predicted by the theory. Therefore, the present EM modeling is also applicable to the QGIP structure.

In the previous MTL analysis, it was determined with a large $s = 4.65$ μm, the diffraction effect among the grid lines is small and the metal contact on top of each grid line serves as a half-wave antenna. An absorption peak will occur whenever the incident $\lambda \approx 2nw$. To verify this conclusion, we plot the E_z distribution at $\lambda_p = 11.1$ μm of the $w = 1.8$ μm detector in Fig. 10. The dipole scattering field distribution inside the grid lines is evident in this plot, which validates the previous conclusion.

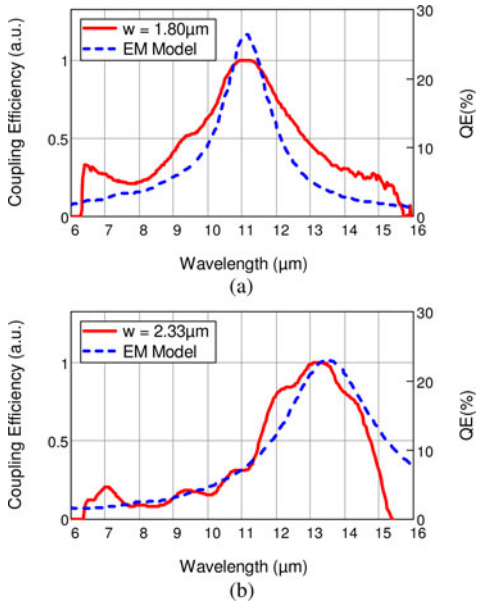


Fig. 9. Measured (solid curve) coupling efficiency and the calculated QE (dashed curve) for (a) $w = 1.80$ μm and (b) $w = 2.33$ μm. The value of w is measured using scanning electron microscope.

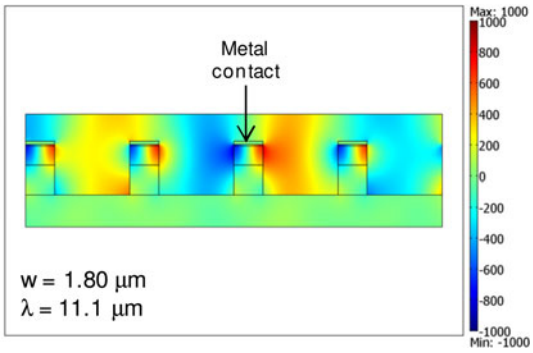


Fig. 10. E_z distribution based on FEM with $E_0 = 377$ V/m.

Fig. 11 shows the 3-D geometry and the E_z distributions of a prism-shaped corrugated QWIP. This detector geometry uses optical reflection at the angled sidewalls to create the needed E_z. This detector geometry also accepts a classical solution for QE based on ray optics [15]. Fig. 12 shows the classical solution, the rigorous EM solution, and the average experimental QE spectra for two focal plane arrays (FPAs) having different cutoff wavelengths. Fig. 12(a) is for a QWIP containing 60 periods of 700-Å $Al_{0.166}Ga_{0.834}As$ and 60-Å GaAs. This active material is placed in the middle of the corrugation and the rest of the volume is filled with contact materials. The array is not antireflection (AR) coated. The calculated α spectrum has a peak at $\lambda = 11.9$ μm with a value of 0.105 μm^{-1} and a 50% cutoff at $\lambda = 12.7$ μm. Fig. 12(b) is from another QWIP that is made of 60 periods of 700-Å $Al_{0.23}Ga_{0.73}As$ and 48-Å GaAs. The calculated α is peaked at $\lambda = 8.7$ μm with a value of 0.145 μm^{-1}. Fig. 12 again shows the agreement among the two theoretical models and the experimental data in terms of the spectral lineshape and the absolute magnitude. Accounting for

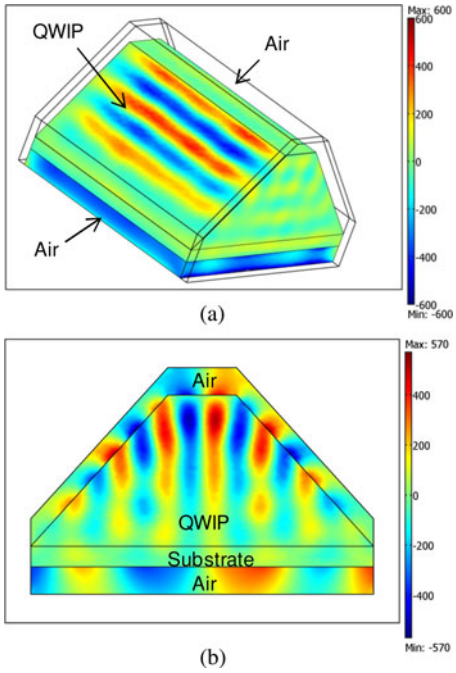


Fig. 11. Detector geometry, which is without an antireflection (AR) coating. The E_z distribution is shown at $\lambda = 11.2$ μm with $E_0 = 377\sqrt{2}$ V/m.

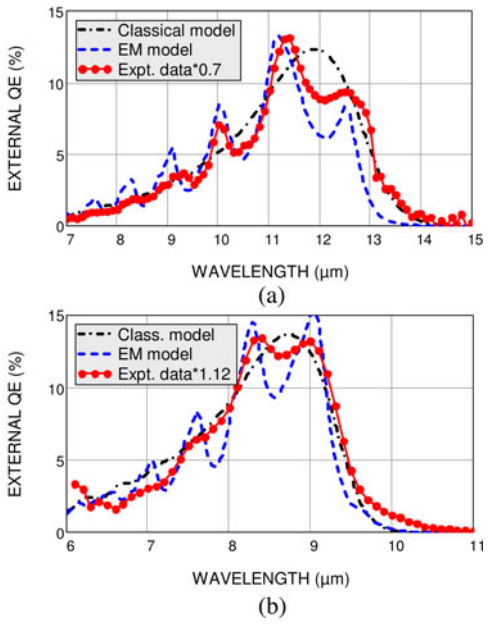


Fig. 12. Figure shows the calculated and measured external QE of two detector materials without an AR-coating.

the optical interference, the EM model is better equipped than the classical model in describing the QE oscillations.

Fig. 13 shows the geometry of a pyramidal C-QWIP, which is AR coated (ARC). This detector geometry has four angled sidewalls to reflect light. As shown in Fig. 14(a), both the classical [15] and the EM models predict the peak QE correctly. The large discrepancy in the spectrum around 8 μm is due to the known epoxy glue absorption used in the FPA integration.

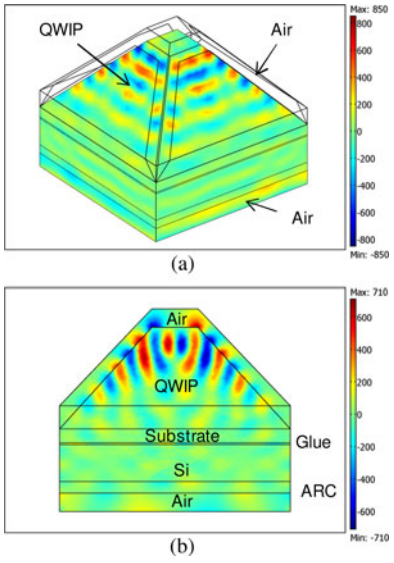


Fig. 13. Detector geometry and the E_z distribution at $\lambda = 8.8$ μm where $E_0 = 377\sqrt{2}$ V/m. This detector is glued to a piece of silicon using epoxy.

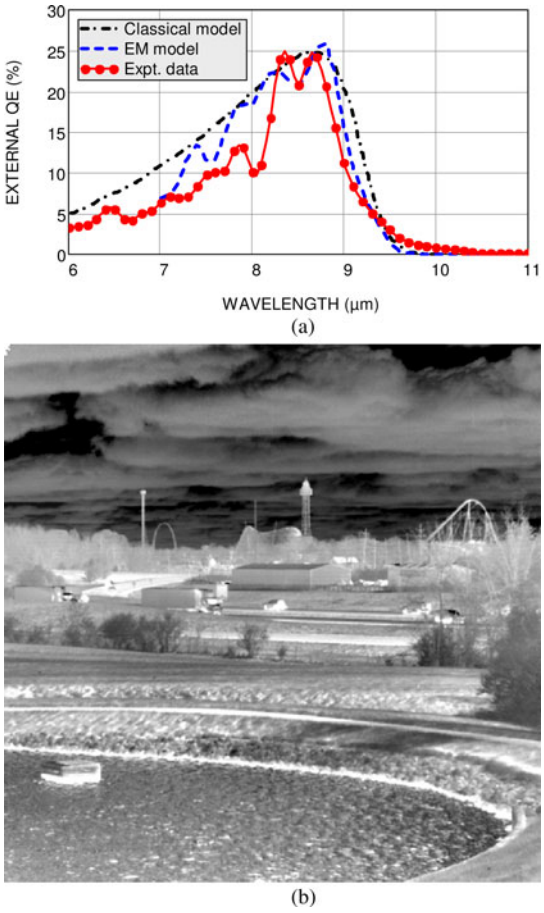


Fig. 14. (a) Calculated and measured external QE spectra of a pyramid-shaped C-QWIP FPA. (b) Infrared image taken by the corresponding 1-MP FPA.

The infrared image in Fig. 14(b) was taken by the corresponding 1-megapixel FPA camera.

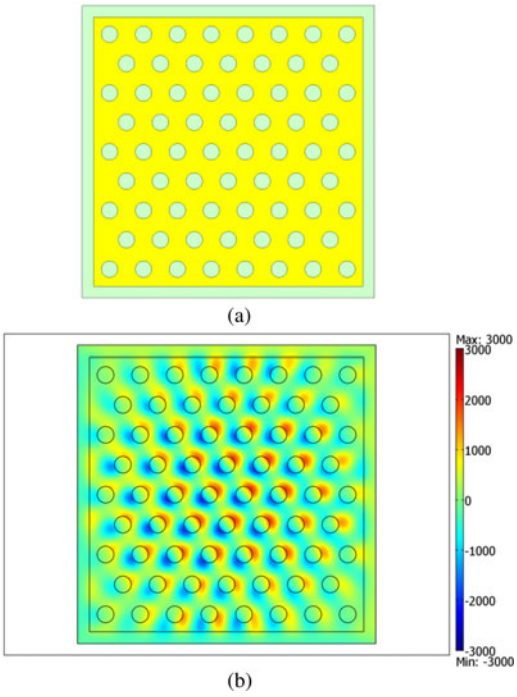


Fig. 15. (a) Plasmonic-enhanced QWIP structure. (b) Calculated E_z at $\lambda = 8.3$ μm with $E_0 = 377\sqrt{2}$ V/m.

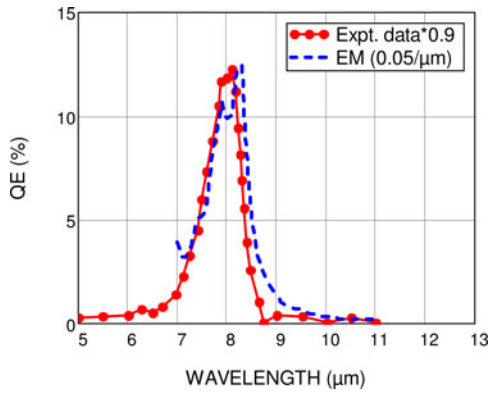


Fig. 16. Measured (solid curve) and the calculated QE with a constant $\alpha = 0.05$ μm^{-1} (dashed curve) of a plasmonic enhanced QWIP.

Fig. 15(a) shows the top view of a plasmonic enhanced QWIP studied by Wu *et al*. [16]. In this structure, a 400-Å-thick gold film perforated with circular holes is deposited on a 0.528 μm-thick InGaAs/InP active material. The spacing between two holes is 2.9 μm and their diameter is 1.4 μm. The active material has a low doping such that the peak α is calculated to be 0.05 μm^{-1}. The modeled E_z distribution at 0.25 μm below the gold film is shown in Fig. 15(b). From the E_z distribution, the calculated QE for this constant α is shown in Fig. 16. It is peaked at $\lambda = 8.3$ μm. Meanwhile, from the measured R spectrum [16] and the estimated gain of 8.5 from a similar detector structure [17], the experimental QE is deduced to be 12.6%, which agrees with the theory to within 10%, and the two spectra have similar lineshapes.

Fig. 17 shows the cross sections of a photonic-crystal-slab-QWIP (PCS-QWIP) studied by Kalchmair *et al*. [18]. The nominal structure consists of an array of air holes with hole spacing $a = 3.1$ μm and hole diameter $d = 1.24$ μm etched through the active and contact materials. The active material thickness t_a is 1.5 μm, and the bottom contact thickness t_c is 0.5 μm. The PCS is suspended in the air at a nominal height $t_{\rm air} = 2.0$ μm above the GaAs substrate. The measured photocurrent spectrum is shown in Fig. 18. Based on the material α spectrum shown in Fig. 18, which is deduced from the measured photoresponse of the edge-coupled detector [18], the QE spectrum is calculated and it indeed contains the characteristic sharp peaks. However, the calculated sharp peaks do not align exactly with the measurement. To obtain a better alignment as that shown in Fig. 18, the theoretical parameter a is reduced slightly from 3.1 to 2.9 μm. This discrepancy could be due to experimental uncertainties or theoretical assumption of an average n of 3.239 over a wide range of wavelengths. In reality, the value of n varies from 3.34 at 4 μm to 3.04 at 11 μm. In addition to the adjusted a, the magnitudes of these peaks depend weakly on $t_{\rm air}$. To obtain a larger peak at position A in Fig. 18, $t_{\rm air}$ is adjusted from 2.0 to

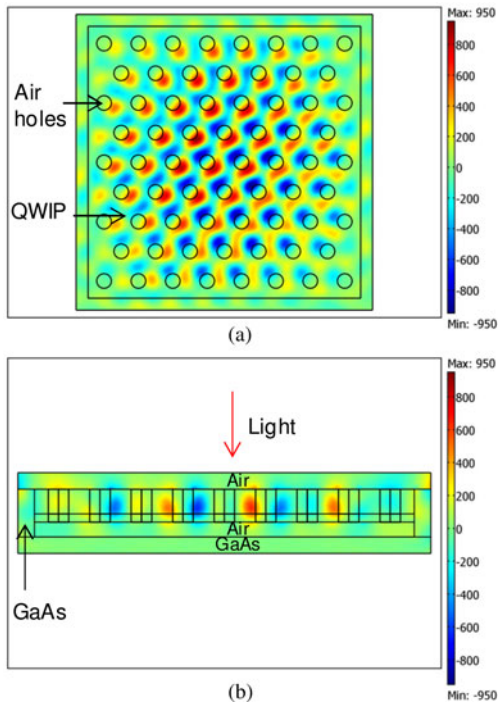


Fig. 17. Cross sections of the PCS-QWIP and the calculated E_z at one of the sharp peaks with $\lambda = 6.44$ μm or $\nu = 1552$ cm^{-1}.

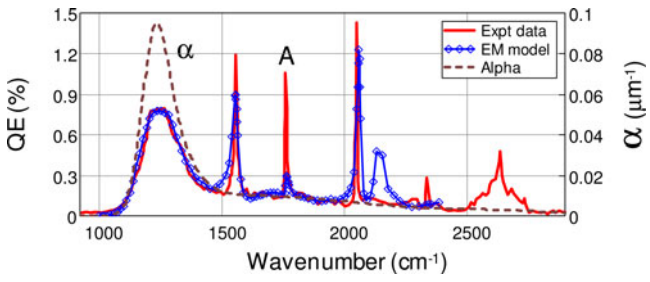


Fig. 18. Measured photocurrent spectrum in arbitrary unit (solid curve), and the calculated QE spectrum based on the displayed α spectrum.

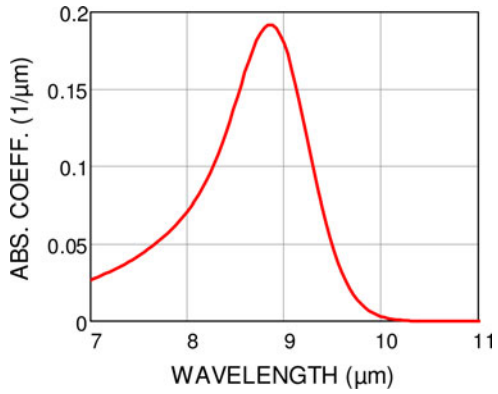


Fig. 19. Absorption coefficient α assumed in the EM modeling.

0.9 μm. A different t_{air} could be due to the sagging of the PCS in the air at the operating temperature. From the modeling, the sharpness of these peaks is caused by two factors. One is the nearly symmetrical detector structure, both in vertical and horizontal directions, that induces strong resonances. The second is the weak material absorption at the peak wavelengths, which introduces only small damping effects on the resonances. The close match of the main peak and some of the side peaks lends support for the present modeling approach.

IV. EM Design

After the EM model is verified by the existing experiments, we can use it for detector design. In the past, the design of a light-coupling structure has been focused mainly on the diffractive element (DE) placed on top of the detector. The size, thickness, and shape of the detector were not part of the consideration. The present modeling instead allows the design of the DE and the detector volume as one integral light-coupling entity. We found that the detector volume actually plays a crucial role in determining QE, in which it acts as a resonant cavity to the light diffracted from the DE. With the versatility of the finite element method, one is also able to consider a much wider variety of DEs whose patterns can be far more complex than that of a regular grating. In general, a DE can be in the form of a photonic Bravais lattice with a basis or in the form of irregularly distributed scatterers. The basis and scatterers can be of any 3-D geometrical objects. The opening up of these arbitrary patterns offers tremendous choices of QE characteristics both in spectral lineshape and in absolute magnitude. This versatility in the detector geometrical design adds to the well-known versatility in the QWIP material design. The combination of the two will yield a tremendous flexibility in designing the specific detector optical properties. The integration of a DE and a resonant cavity is referred to as the resonator QWIP [19], or the R-QWIP. With different DE designs to suit different applications, there will be different types of R-QWIPs. We have since studied a large number of these detector designs and obtained a wide range of coupling characteristics. Here, we describe two of the simplest designs for illustration purposes: one is the grating-resonator-QWIP or GR-QWIP and another is the ring-resonator-QWIP or RR-QWIP.

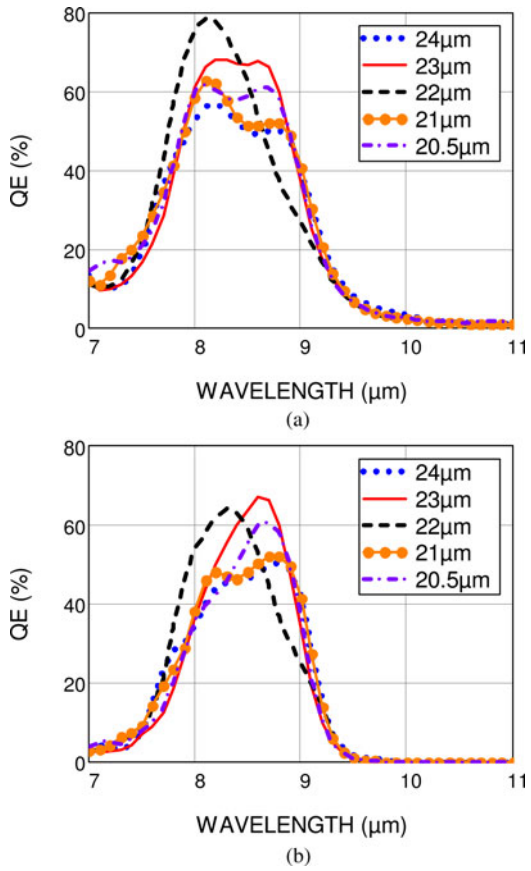


Fig. 20. Calculated QE for different detector size p for (a) a constant α of 0.20 μm^{-1} and (b) a varying α according to Fig. 19.

First, we optimize a 25-μm pitch GR-QWIP without an AR-coating for 8–9-μm detection. The value of α is either assumed to be constant at 0.20 μm^{-1} or a narrowband spectrum shown in Fig. 19. The period of the grating is first set at 2.7 μm. All the rest of the parameters, such as the active layer thickness, the grating height, the bottom contact thickness, and the pixel linear size p, are adjusted to give the maximum QE in the 8–9-μm band. Fig. 20 shows one of the optimizing procedures by varying p alone, while all other parameters have been optimized. The result shows that the pixel size has a modest effect in the QE in this case. For a constant α, the maximum QE is 78.6% achieved by a $p = 22$-μm GR-QWIP at $\lambda = 8.1$ μm.

The aforementioned modeling shows that the GR-QWIP is a promising detector structure for narrowband detection. Using an array of square rings as the DE instead, the coupling bandwidth can be widened as shown in Fig. 21(a). In this structure, the outer dimension of each ring is 4 μm and the inner dimension is 1.4 μm. The wider bandwidth is beneficial even for a narrow band material as seen in Fig. 21(b). It reduces the spectral variations with different p and preserves the absorption lineshape of the material. The largest QE in Fig. 19(a) is 73.1% achieved at $p = 21.5$ μm and $\lambda = 9.9$ μm.

Since these GR-QWIPs and RR-QWIPs will be built on a very thin active material layer, the photoconductive gain can be as large as 0.6 at full bias. Therefore, the estimated conversion

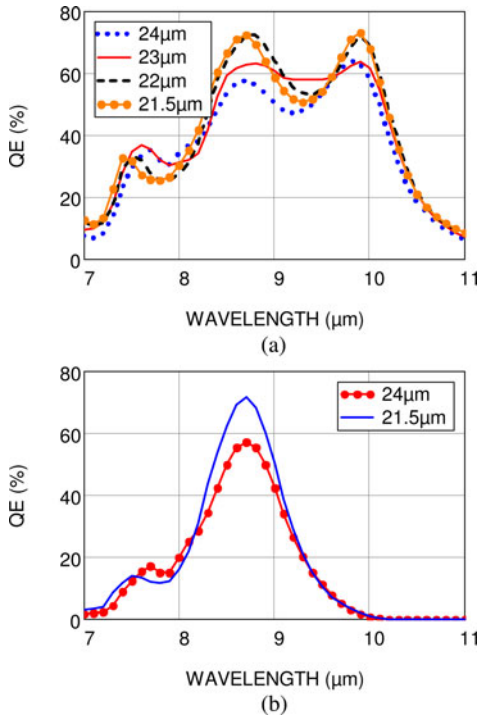


Fig. 21. Calculated QE for different detector size p for (a) a constant α of 0.20 μm^{-1} and (b) a varying α.

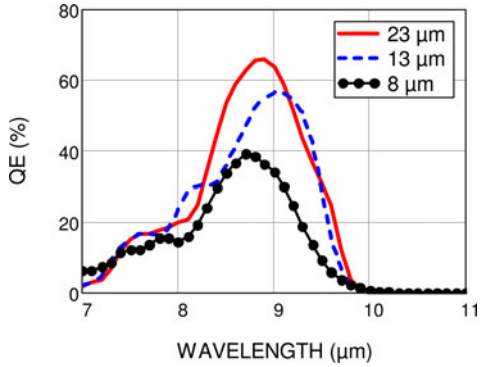


Fig. 22. Calculated QE for different RR-QWIP detector sizes for a varying α according to Fig. 19.

efficiency is about 40%, which is adequate for many high-speed applications.

V. Modeling of Imperfections and Crosstalks

The above calculations assumed that the designed structures can be produced faithfully in experiment, which may not always be feasible technologically or economically. The present model is adept in determining the impacts of processing imperfections such as size nonuniformity and deformed DEs by modeling the actual fabricated structures. The subsequent optimization can be performed on the realizable patterns. Fig. 22 shows one of such examples in modeling the optimized RR-QWIPs with rounded ring corners. The result for the 23-μm detector shows that by adjusting other detector parameters, the rounded rings can have very similar QE as the square rings. Based on the rounded ring structure, one can also design efficient detectors for smaller pixel sizes. As shown in Fig. 22, the detector can have 57% QE for 13-μm pixels and 40% QE for 8-μm pixels.

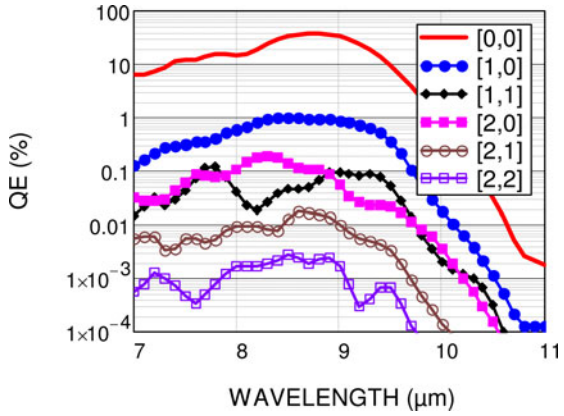


Fig. 23. Calculated QE of the neighboring RR-QWIP pixels when a plane wave is incident onto the center pixel. The legend specifies the pixel coordinates.

EM modeling can also be used to determine other FPA properties such as pixel crosstalk. For small pitch arrays, crosstalk due to pixel optical diffraction is a concern. In order to determine the amount of crosstalk, one can evaluate the values of QE of the surrounding pixels while only the center pixel pitch area is illuminated. Fig. 23 shows an example for the 10-μm pitch arrays, in which the result for five nearest neighbors is plotted. From this calculation, the crosstalk is estimated to be less than 2.5% for the RR-QWIP structure.

VI. Conclusion

The QE of a detector uniquely determines its sensitivity under background-limited infrared performance condition and, hence, it is a critical figure of merit to consider for a detector technology. QWIPs possess many unique advantages, but historically suffer from a low QE. This low QE is due to the lack of a quantitative model to perform detector design and optimization. In this paper, we have established an EM model of using (1) to calculate QE explicitly. This approach is shown to be able to provide a quantitative answer to any detector geometry in any degree of desired detail. We verified its accuracy and reliability with experiments and with analytical classical and modal transmission-line solutions. With this approach, one can now ascertain the optical coupling properties according to its physical construct. It is also well known that the absorption properties of the detector material can be calculated accurately from its layer structure. With the advent of a rigorous model for the light-coupling structure, the QWIP technology can now enter into a new era, in which all of its optical properties can be engineered in precision. This feature will be invaluable to FPA production and application. In this paper, we also optimized a grating resonator to achieve a high QE and designed a ring resonator to broaden its coupling bandwidth. Other coupling lineshapes can also be similarly designed to suit any applications.

REFERENCES

[1] A. De Rossi, E. Costard, N. Guerineau, and S. Rommeluere, "Effect of finite pixel size on optical coupling in QWIPs," *Inf. Phys. Technol*., vol. 44, pp. 325–330, 2003.

[2] K. K. Choi, K. M. Leung, T. Tamir, and C. Monroy, "Light coupling characteristics of corrugated quantum well infrared phorodetectors," *IEEE J. Quantum Electron*., vol. 40, no. 2, pp. 130–142, Feb. 2004.

[3] D. W. Wilson, "Electromagnetic modeling of multi-wavelength QWIP optical coupling structures," *Inf. Phys. Technol*., vol. 52, pp. 224–228, 2009.

[4] J. Wang, X. Chen, Z. Li, and W. Lu, "Study of grating performance for quantum well photodetectors," *J. Opt. Soc. Amer. B*., vol. 27, pp. 2428–2432, 2010.

[5] K. K. Choi, M. D. Jhabvala, D. P. Forrai, A. Waczynski, J. Sun, and R. Jones, "Electromagnetic modeling of quantum well infrared photodetectors," *IEEE J. Quantum Electron*., vol. 48, no. 3, pp. 384–393, Mar. 2012.

[6] K. K. Choi, *The Physics of Quantum Well Infrared Photodetectors*. Singapore: World Scientific, 1997, pp. 122–125.

[7] K. K. Choi, "Electromagnetic modeling of edge coupled quantum well infrared photodetectors," *J. Appl. Phys*., vol. 111, pp. 124507-1–124507-4, 2012.

[8] A. Nedelcu, H. Facoetti, E. Costard, and P. Bois, "Small pitch, large format long-wave infrared QWIP focal plane arrays for polarimetric imagery," *Proc. SPIE*, vol. 6542, pp. 65420U-1–9, 2007.

[9] M. Jhabvala, K. Choi, A. Waczynski, A. La, M. Sundaram, E. Costard, C. Jhabvala, E. Kan, D. Kahle, R. Foltz, N. Boehm, M. Hickey, J. Sun, T. Adachi, N. Costen, L. Hess, H. Facoetti, and M. Montanaro, "Performance of the QWIP focal plane arrays for NASA's landsat data continuity mission," *Proc. SPIE*, vol. 8012, pp. 80120Q-1–14, 2011.

[10] B. Brill and G. Sarusi, "QWIP research and development of 320×256 QWIP arrays in EL-OP," *Proc. SPIE*, vol. 3061, pp. 781–788, 1997.

[11] B. Brill and G. Sarusi, "System considerations in the design of QWIP-based thermal imagers," *Proc. SPIE*, vol. 3436, pp. 270–277, 1998.

[12] T. R. Schimert, S. L. Barnes, A. J. Brouns, F. C. Case, P. Mitra, and L. T. Claiborne, "Enhanced quantum well infrared photodetector with novel multiple quantum well grating structure," *Appl. Phys. Lett*., vol. 68, pp. 2846–2848, 1996.

[13] K. K. Choi, G. Dang, J. W. Little, K. M. Leung, and T. Tamir, "Quantum grid infrared spectrometer," *Appl. Phys. Lett*., vol. 84, pp. 4439–4441, 2004.

[14] L. Yan, M. Jiang, T. Tamir, and K. K. Choi, "Electromagnetic modeling of quantum-well photodetectors containing diffractive elements," *IEEE J. Quantum Electron*., vol. 35, no. 12, pp. 1870–1877, Dec. 1999.

[15] K. K. Choi, C. J. Chen, and D. C. Tsui, "Corrugated quantum well infrared photodetectors for material characterization," *J. Appl. Phys*., vol. 88, pp. 1612–1623, 2000.

[16] W. Wu, A. Bonakdar, and H. Mosheni, "Plasmonic enhanced quantum well infrared photodetector with high detectivity," *Appl. Phys. Lett*., vol. 96, pp. 161107-1–161107-3, 2010.

[17] S. U. Eker, Y. Arslan, A. E. Onuk, and C. Besikci, "High conversion efficiency InP/InGaAs strained quantum well infrared photodetector focal plane array with 9.7 μm cut-off for high-speed thermal imaging," *IEEE J. Quantum Electron*., vol. 46, no. 2, pp. 164–168, Feb. 2010.

[18] S. Kalchmair, H. Detz, G. D. Cole, A. M. Andrews, P. Klang, M. Nobile, R. Gansch, C. Ostermaier, W. Schrenk, and G. Strasser, "Photonic crystal slab quantum well infrared photodetector," *Appl. Phys. Lett*., vol. 98, pp. 011105-1–011105-3, 2011.

[19] K. K. Choi, "Photodetectors using resonance and method of making," U.S. Patent. 2012/0012816 A1, 2012.

Kwong-Kit Choi (SM'99–F'07) received the B.S. degree from the University of Hong Kong, Kowloon, Hong Kong, in 1979, and the Ph.D. degree in physics from Yale University, New Haven, CT, in 1984.

He is currently a Senior Research Scientist for Physical Sciences at the U.S. Army Research Laboratory, Adelphi, MD. His publications include *The Physics of Quantum Well Infrared Photodetectors* (World Scientific, 1979). He has been awarded 13 patents in the QWIP area. His current interests are focal plane array demonstration, new detector design, and infrared optoelectronics.

Dr. Choi is a Fellow of the American Physical Society and the Army Research Laboratory, and a member of SPIE. He is an inductee of the NASA Space Technology Hall of Fame and the recipient of the Distinguished Presidential Rank Award.

Murzy D. Jhabvala received the Ph.D. degree from the University of Maryland, College Park, MD.

He is a Chief Engineer of the Instrument Systems and Technology Division, NASA Goddard Space Flight Center, Greenbelt, MD. He has designed, fabricated, and flight qualified PMOS/CMOS integrated circuits, and has developed the internal reference source for the Cosmic Background Explorer. He was the co-inventor of the polysilicon microfilament IR sources used on the Spitzer Telescope, and led/codeveloped (with ATT/Bell Labs and Rockwell Science Center) the first long wave GaAs QWIP 128 × 128 array. He was the Principal Investigator for the development of the 640 × 512 GaAs QWIP Hyperspectral (4–15 μm) Imaging System and the LWIR 1K × 1K QWIP array. He led the development of the microshutter arrays for the James Webb Space Telescope project and led the QWIP-based focal plane development team for the thermal IR instrument on NASA's next Landsat mission.

David P. Forrai (SM'07) received the B.S. degree in electrical engineering from the University of Wisconsin-Madison, Madison, and the M.S. degree in electrical engineering from The Ohio State University, Columbus.

He is a Senior Electrooptics Engineer at L-3 Communications Cincinnati Electronics, Mason, OH. His current research focuses on advanced focal plane arrays including dual-band/multicolor FPAs, high-quantum efficiency corrugated QWIP FPAs, and polarimetric FPAs.

Augustyn Waczynski received the M.S.E.E. degree from Warsaw Polytechnic, Warsaw, Poland.

He is a Senior Electronics Engineer in the NASA Goddard Space Flight Center, Greenbelt, MD. He works in Detector Characterization Laboratory specializing in electro optical measurements and calibration of large format focal planes for space applications.

Jason Sun received the B.S. degree in physics from Nanjing University and the M.S. degree in physics from University of Kansas.

He is a Physicist at the U.S. Army Research Lab, Adelphi, MD. He has experiences in a wide range of optoelectronic and RF device physics and fabrication. He conducted research and development on C-QWIP FPA fabrication, RF tunable capacitor design, and characterization and silicon microbolometer uncooled FPAs. He has more than ten years' experience in research on conventional and high-Tc superconductors and expertise in high-Tc superconductor epitaxial thin-film growth and characteristic test and analyze. He has three patents and over 30 publications.

Robert Jones received the B.S. degree in physics from the University of South Florida, Tampa, and the M.S. degree in physics and the Ph.D. degrees in electrical engineering from the University of Cincinnati, Cincinnati.

He is a Senior Engineer at L-3 Communications Cincinnati Electronics, Mason, OH. His current research focuses on advanced focal plane array photo detector structures and materials including QWIP, barrier, and SLS structures with emphasis on developing processes with high manufacturability.

Passive Infrared Sensing Using Plasmonic Resonant Dust Particles

Mark Mirotznik, William Beck, Kimberly Olver, John Little and Peter Pa

International Journal of Optics 2012, 651563 (8 pp.)-651563 (8 pp.)651563 (8 pp.).

Research Article

Passive Infrared Sensing Using Plasmonic Resonant Dust Particles

Mark Mirotznik,[1] William Beck,[2] Kimberly Olver,[2] John Little,[2] and Peter Pa[1]

[1] *Department of Electrical and Computer Engineering, University of Delaware, 106 Evans Hall, Newark, DE 19716, USA*
[2] *Sensors and Electron Devices Directorate, U.S. Army Research Laboratory, Adelphi, MD 20783, USA*

Correspondence should be addressed to Mark Mirotznik, mirotzni@ece.udel.edu

Received 5 May 2012; Accepted 24 June 2012

Academic Editor: Georgios Veronis

We present computational and experimental results of dust particles that can be tuned to preferentially reflect or emit IR radiation within the 8–14 μm band. The particles consist of thin metallic subwavelength gratings patterned on the surface of a simple quarter wavelength cavity. This design creates distinct IR absorption resonances by combining the plasmonic resonance of the grating with the natural resonance of the cavity. We show that the resonance peaks are easily tuned by varying either the geometry of the grating or the thickness of the cavity. Here, we present a computational design algorithm along with experimental results that validate the design methodology.

1. Introduction

Most objects, either manmade or found in nature, reflect and emit infrared (IR) radiation in a relatively smooth spectrum; however, by applying structures with resonant absorption to the surface of those materials, the reflection and emission spectra can be enhanced or reduced at particular wavelengths (as illustrated in Figure 1). Moreover, by mixing small resonant particles (<100 μm) designed for several different wavelengths, we can create IR dust that reflects or emits with a characteristic spectral signature. Such material-by-design particles would be useful for a variety of practical applications. For example, when applied to a base surface, the resonant particles could be used to tune an IR reflectance to mimic other natural or manmade surfaces. This could be useful as a calibration standard for hyperspectral imaging systems. Additionally, if the particles are chemically functionalized, there are a number of remote atmospheric sensing applications that could be explored.

2. Infrared Absorbers Using Plasmonic Gratings

It is well known that metallic surfaces patterned on a subwavelength scale exhibit unusual electromagnetic properties at optical wavelengths. In particular, the presence of localized surface plasmon resonances creates well-defined absorption bands. This phenomenon has been studied and exploited by a number of investigators to realize new types of sensors, optical filters, and absorbers [1–5]. The goal of this work was to numerically and experimentally study plasmonic-based resonant absorbers in the long-wavelength IR (LWIR) band (8–14 μm) that could be fashioned into small (~100 μm × 100 μm × 25 μm) dust particles. The dust particles could then be used to tailor the reflectivity/emissivity of a surface or dispersed in air and used for atmospheric sensing applications.

There are a number of small resonant absorbing "dust like" structures that could be used to preferentially absorb, and thus thermally emit, IR radiation at specific wavelengths including dielectric ring resonators, resonant patch antennas, and plasmonic-based resonator. These various structures were compared based on (1) their ability to efficiently absorb IR energy at selected wavelengths within the 8–14 μm band, (2) the ability to easily tune the resonant absorption, (3) ease of fabrication, and (4) manufacturing cost. Based on these criteria, we chose to investigate, in detail, the relatively simple surface plasmon-based designs shown in Figure 2. The building blocks for this design are two thin resonant

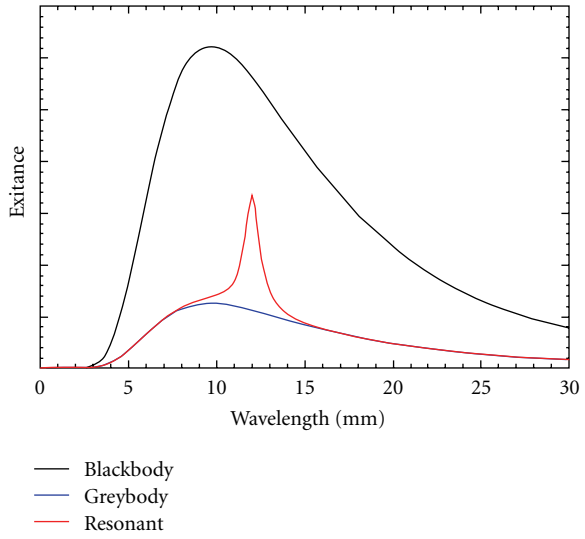


Figure 1: Notional diagram that illustrates the normally smooth thermal exitance curves from blackbody and graybody objects compared to the resonant behavior of our "engineered" IR resonant dust.

cavities, one on the top of Figure 2 and the other on the bottom. Each cavity is composed of a thin gold ground plane, a thin dielectric substrate layer (formed from zinc selenide (ZnSe) in our design), and a subwavelength metallic grating made from gold. In the middle of the structure is a relatively thick silicon layer needed for mechanical rigidity. The symmetry of the top and bottom layers was needed since the particles, when dispersed, would orient themselves randomly.

The strong resonant behavior of this design is due to a combination of two different resonant phenomena. The first is a surface plasmon resonance that is excited within the subwavelength gold grating. The second is a cavity resonance excited in the ZnSe substrate region that is between the grating layer and the metallic ground plane layer. By adjusting the thickness of the ZnSe substrate for a given grating period and duty cycle, a strong absorption resonance can be excited at any wavelength within the 8–14 micron band. To create small dust particles, a large sample is diced into small ($\sim 100\,\mu\text{m} \times 100\,\mu\text{m} \times 25\,\mu\text{m}$) particles.

3. Computational Modeling and Design

Two different computational models were employed to rigorously design and validate the resonant structure shown in Figure 2. The first method is a fully periodic planar method called the rigorous coupled wave method. The second method, finite element method (FEM), was used to investigate finite-sized particle effects. A brief description of these two methods along with simulation results are presented in the next two sections.

3.1. Modeling of Infinitely Periodic Structures Using Rigorous Coupled Wave Analysis. Two approaches are used extensively for simulating the electromagnetic properties of infinitely periodic subwavelength gratings. The first uses effective media theory to provide closed-form approximations for the effective dielectric constants as a function of the grating structure [6]. Although attractive from a computational perspective, the approximate expressions are accurate only for gratings whose period is much smaller than the wavelength of illumination. As the grating period approaches the wavelength, which is referred to as the resonance regime, the assumptions on which these closed-form expressions are based are no longer valid. For our designs, we assumed grating periods only slightly smaller than the material wavelength and thus could not accurately utilize effective media theory.

We instead employed a second approach using a rigorous electromagnetic model. Although computationally more difficult, this approach is capable of generating accurate results for gratings of any period size and shape. Several rigorous electromagnetic models can be used for this calculation. We chose the rigorous coupled wave (RCW) algorithm originally presented by Moharam and Gaylord [7]. Our specific implementation is based on the enhanced transmittance matrix approach introduced by Moharam and Gaylord [7] and later refined by Lalanne [8] and Noponen and Turunen [9]. For the sake of brevity, we refer the reader to the references above for details on the RCW method. While being accurate, the RCW method does assume the grating structure, shown in Figure 2, is infinite in the transverse directions. The effect of finite-sized samples is investigated in Section 3.3.

3.1.1. RCW Simulation Results for Infinitely Periodic Surfaces. Figure 3 presents typical simulation, results calculated using the RCW method. In the figure, the reflectivity of the sample is calculated as a function of wavelength and polarization for a normally Incident Planewave.

For this simulation the ZnSe substrate thickness was assumed to be $1.8\,\mu$m, the gold grating period was $3.0\,\mu$m with a 50% duty cycle. The gold gratings were assumed to be 100 nm thick, and the gold ground planes were 300 nm thick. The electromagnetic material properties of the gold were determined using the model given in [10]. For the ZnSe layer, a lossless index of refraction of $n = 2.41$ was used in all simulations. The incident field was assumed to be normally incident from the top. For this design, a very strong resonance absorption, near-perfect absorption, is predicted near $9.5\,\mu$m for the case of parallel polarization (E-field polarized along the axis of the grating) and only weak resonances occur for the case of perpendicular polarization (E-field polarized perpendicular to the axis of the grating).

3.1.2. Reflectance Sensitivity to Geometrical Parameters. Given a specific substrate and metallization material, such as ZnSe and gold, the dust particle's reflectance can be tuned by proper selection of the geometrical parameters shown in Figure 2: specifically, (1) thickness of the ZnSe layer, denoted by h in Figure 2; (2) grating period, denoted by Λ in Figure 2; (3) the grating's duty cycle given by (w/Λ in

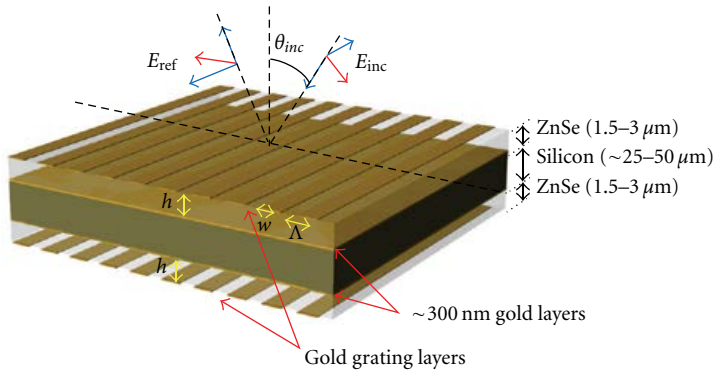


FIGURE 2: Illustration of our surface plasmon-based IR resonant particles. The gold subwavelength gratings along with cavity resonances produce distinct resonant absorption phenomenon within the LWIR band.

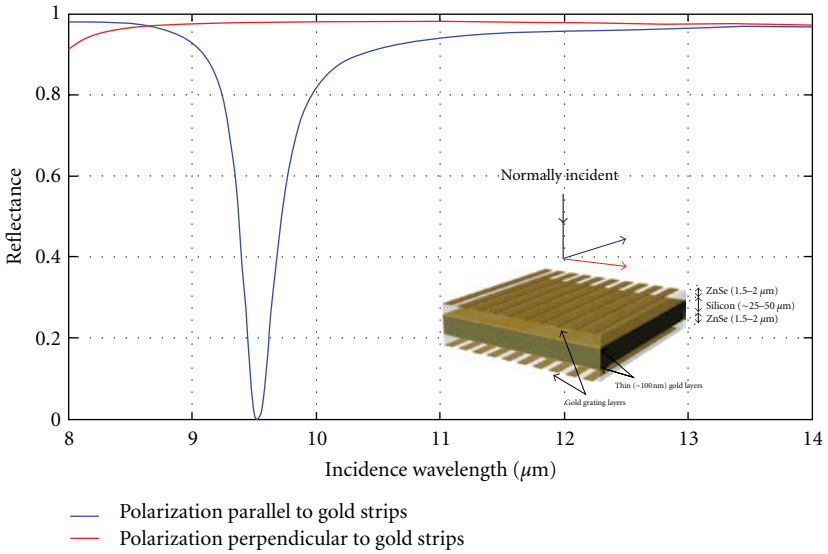


FIGURE 3: Simulation results, using the RCW method, that present the reflectivity at normal incidence within the LWIR band. The reflectivity as expected is polarization dependent due to the anisotropic nature of the gratings.

Figure 2); (4) thickness of the gold grating layer and gold ground plane. Assuming the gold layers are thick enough to prevent transmission (i.e., much thicker than the penetration depth), the variables given by 1–3 above will have the most effect on the LWIR reflectance.

In Figure 4, we present the effect of the ZnSe substrate thickness on the resonant behavior. As the thickness is increased from 1.5 to 2.5 μm, the resonant dip shifts from 8.3 to 13.2 μm, respectively. Thus the resonant behavior can be tuned by simply varying the thickness of the ZnSe substrate.

Alternatively, for a given substrate thickness, the resonant absorption characteristics can be tuned by varying the grating period and duty cycle. Shown in Figure 5 is the simulated reflectance of a sample in which the substrate thickness was fixed at 2.0 μm and the grating period was varied from 1.0 to 3.0 μm. For this simulation, the duty cycle was fixed at 50%. While the resonant wavelength clearly varied with grating period, the change was less sensitive than varying substrate thickness. Moreover, by just changing the grating period, with all other parameters fixed, the amplitude of the resonance would vary considerably. Lastly, we varied the grating's duty cycle while holding the substrate thickness and grating period fixed at 2.0 and 3.0 μm, respectively. As shown in Figure 6, the grating duty cycle has a large effect on not only the resonant wavelength but also on the amplitude and bandwidth of the resonance.

The sensitivity to incident angle was also evaluated using the RCW code. A typical result for the case of both parallel and perpendicular polarization is shown in Figure 7. Here, the simulation results predict that the resonant frequency for parallel polarization (Figure 7(a)) should slowly increase as the incident angle increases from normal incidence (0 degrees in the figure) to near grazing angles (80 degrees). It is interesting to note that for the case of parallel polarization (Figure 7(a)) the variation in resonant wavelength with incident angle is relatively small (<1 μm) even with near-grazing incident angles. For the given application of resonant dust particles, this is an attractive feature since the orientation of

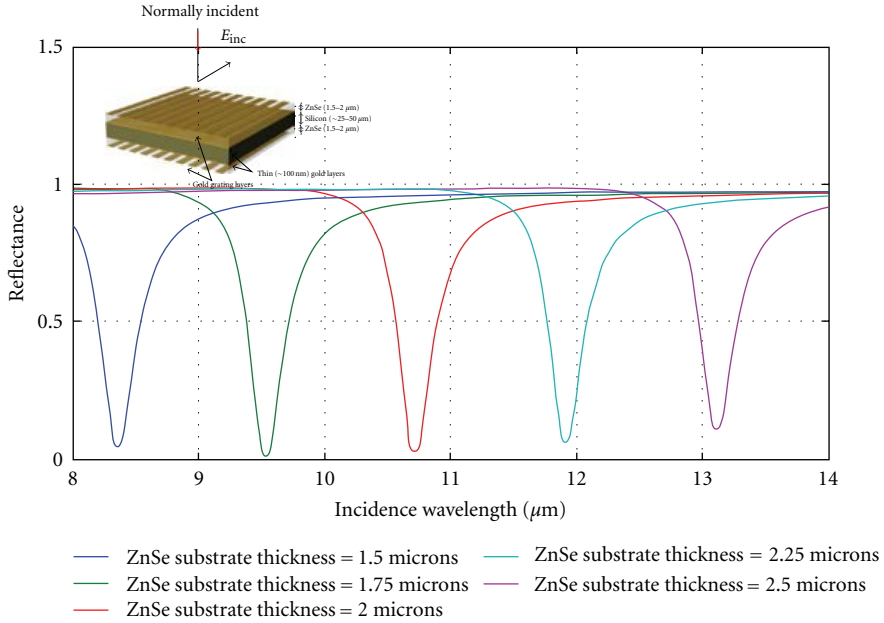


FIGURE 4: Simulation results, using the RCW method, that present the reflectance at normal incidence within the LWIR band as the ZnSe substrate thickness is varied from 1.5 to 2.5 μm. For this simulation, the grating period is fixed at 3.0 μm with a 50% duty cycle. The incident wave was normally incident with parallel polarization. As the substrate thickness is increased, the resonant absorption peak shifts to longer wavelengths but still remains strong. The bandwidth of the resonance also remains relatively fixed as the substrate thicknesses is varied.

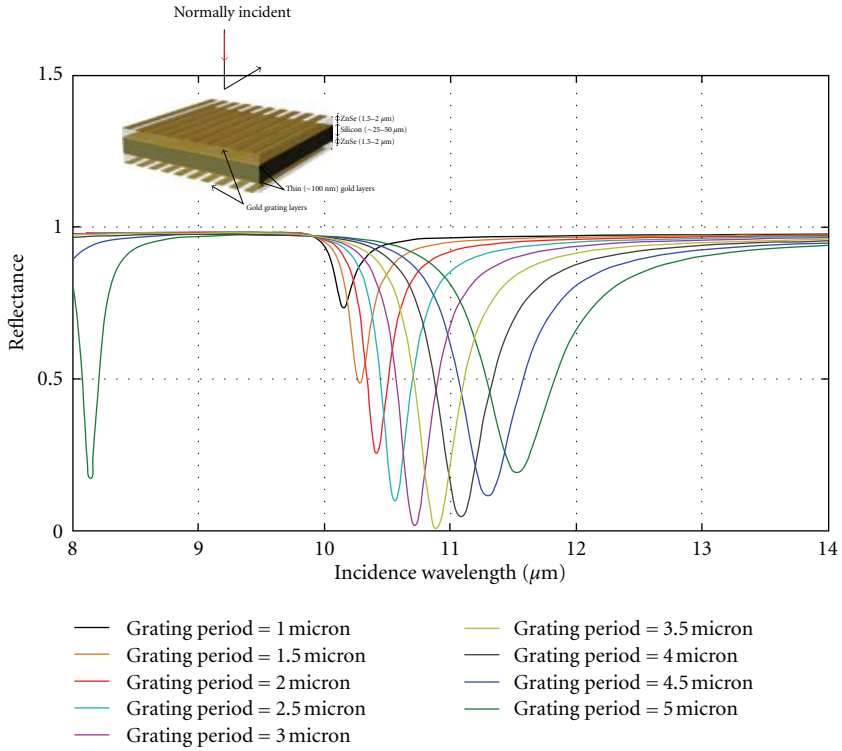


FIGURE 5: Simulation results, using the RCW method, that present the reflectance at normal incidence within the LWIR band as the gold grating period is varied from 1.0 to 5.0 μm. For this simulation, the substrate thickness is fixed at 2.0 μm with a grating duty cycle of 50%. The incident wave was normally incident with parallel polarization.

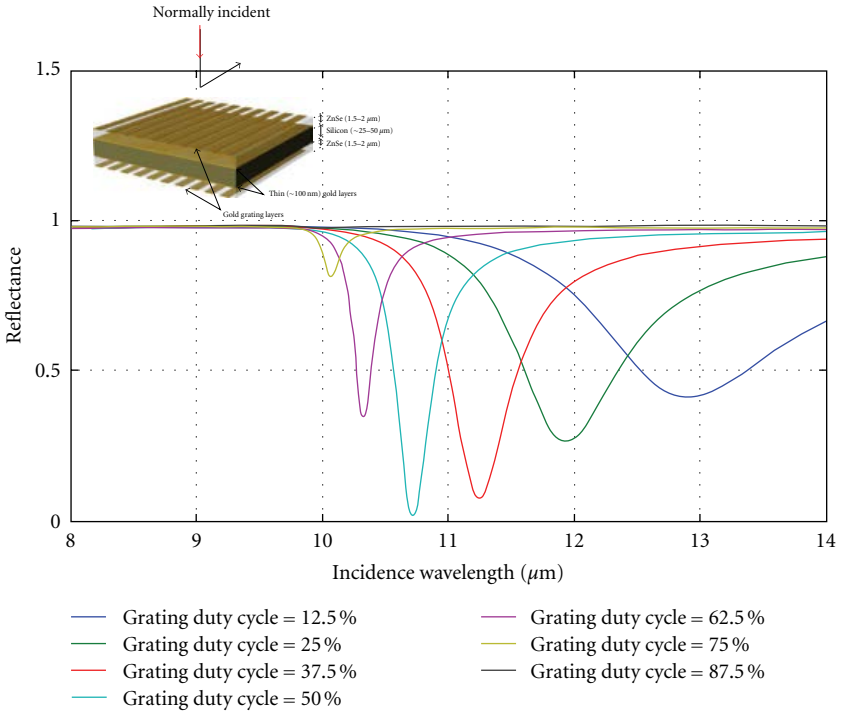


FIGURE 6: Simulation results using the RCW method that present the reflectance at normal incidence within the LWIR band as the duty cycle of the gold grating period is varied from 12.5% to 87.5%. For this simulation, the substrate thickness is fixed at 2.0 μm with a grating period of 3.0 μm. The incident wave was normally incident with parallel polarization.

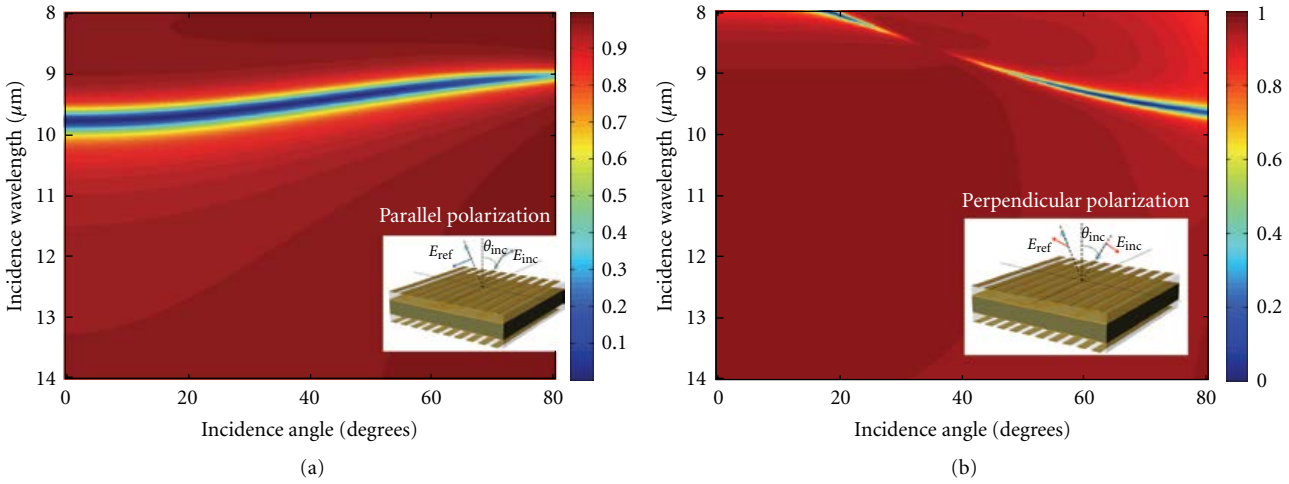


FIGURE 7: RCW predictions illustrating the sensitivity of our resonant structure with incident angle. The plot on the left (a) is for parallel polarization, while the plot on the right (b) is for perpendicular polarization.

the particles with respect to the incident field cannot be well controlled.

3.2. Iterative Design. As Figures 4 through 7 demonstrate, the resonant absorption properties of the structure shown in Figure 2 have a complicated dependence on a number of geometrical parameters. As a result, it is unlikely that any simple analytical design equation could be derived and used to determine an optimal structure for a given desired response. Consequently, we implemented a numerical iterative design algorithm. Here the RCW method is used to calculate the full wave solution for the reflectance as a function of wavelength, polarization, and angle of incidence for a geometry of a given substrate thickness, grating period, and duty cycle. An optimization algorithm is then used to refine the geometry until an objective function is minimized. The objective function may vary depending on the application, but in most cases we chose to minimize the total reflectance over some desired

wavelength band. A number of iterative optimization algorithms could be employed including traditional derivative-based algorithms, genetic algorithms, or direct pattern search algorithms. An advantage of both genetic and pattern search algorithms is that they do not require derivatives, and as a consequence work well on nondifferentiable, stochastic, and discontinuous objective functions. Both simple genetic algorithms and direct pattern search algorithms were implemented and tested for the application of interest here. While both methods produced comparable results, the pattern search algorithm was often computationally less expensive.

3.3. Modeling of Finite Grating Effects Using the Finite Element Method. The RCW method, while accurate and computationally efficient, assumes the gratings to be infinitely periodic. For our application, the samples are actually diced into small ($\sim 100\,\mu\text{m} \times 100\,\mu\text{m} \times 25\,\mu\text{m}$) particles. Consequently, it is important to understand the effects of relatively small (<10 wavelengths) finite-sized particles on the overall effectiveness of the design. To conduct these simulations we used the commercial EM solver, HFSS from Ansys. Simulations were conducted using HFSS's FEM solver with grating structures that varied from 25 to 100 μm on a side.

Figure 8 plots the simulated current density on the surface of a $50\,\mu\text{m} \times 50\,\mu\text{m} \times 5\,\mu\text{m}$ plasmonic particle at a fixed incident wavelength of 10 μm. The spatial distribution of current is a direct consequence of its finite lateral size and will affect the total absorbed energy. In Figure 9, we plot the average reflectance of the same particle as a function of wavelength. While the total absorption is slightly less and the resonance wavelength is slightly shifted towards longer wavelength, the finite-sized particles still behave with the same general absorption characteristics as the infinitely periodic predictions described previously.

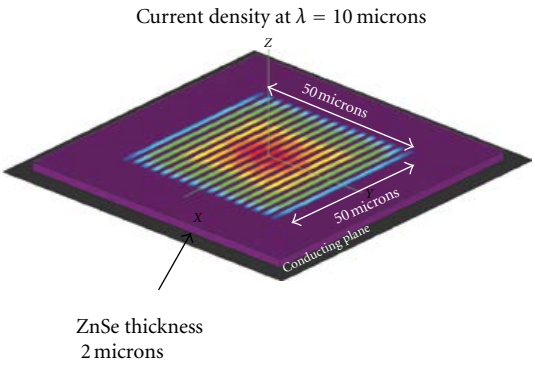


Figure 8: Current density distribution for a finite-sized resonant particle. Simulations were conducted using HFSS FEM solver.

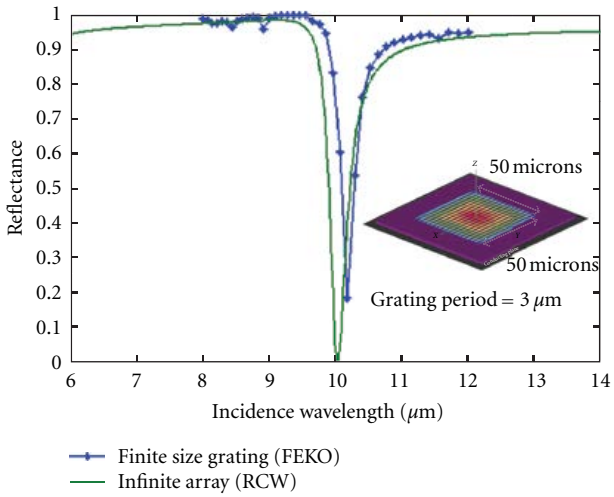


Figure 9: Predicted reflectance curves for a finite-sized particle compared to the infinitely periodic calculations. The edge effects of the finite-sized sample are evident but do not significantly alter the resonant peak.

4. Experimental Fabrication

To fabricate the samples, a thin (80 microns) 2-inch silicon wafer was first mounted onto a 3-inch (350–500 micron) silicon carrier wafer using Aquabond 55 Adhesive Products wax. The carrier wafer was placed on a hot plate at a temperature of 80°C. A small amount of wax was smeared on the surface starting at the center and working outward. The thin silicon wafer was carefully placed on top of the wax. A flat glass plate was placed on top of the thin wafer, followed by a brass weight. This was to keep the silicon wafer as flat as possible during the mounting procedure. The hot plate was turned off, and the wax was allowed to cool to room temperature. Excess wax on and around the mounted silicon wafer was removed by gently swabbing it away with a 1% solution of Aqua Clean. The wafer assembly was placed in a vacuum electron beam evaporator. A blanket metallization of 300 Å of chromium followed by 2000 Å of gold was evaporated on to the wafer. The assembled structure was then moved to another vacuum e-beam evaporator, and a 1.8-micron-thick layer of ZnSe was evaporated onto the surface. Depositions were performed at 145°C, with a base pressure of 1×10^{-6}. A 120-Å layer of yttrium oxide (Y_2O_3) was deposited first to promote adhesion between the substrate and the ZnSe.

Photolithography on the ZnSe was achieved by first spin coating the wafer assembly with AZ # 5214 image reversal photoresist at a speed of 4000 rpm for 40 seconds. This photoresist was hot plate baked at 110°C for 2 minutes, exposed on a JBA vacuum contact aligner for 20 seconds with a bulb intensity of 4 mW/cm^2, hotplate baked (reversal bake) at 124°C for 40 seconds, and flood exposed for 25 seconds. The resist was then developed in AZ 300 MIF photoresist developer for 60 seconds and rinsed in deionized (DI) water for 1 minute. The wafer was then dried with nitrogen gas. The resulting photolithography was inspected under a microscope for clearing. Prior to loading the wafer assembly into the e-beam evaporator for the grating structure, a photoresist cleaning in a barrel plasma asher was performed. The patterned wafer assembly was placed into the vacuum e-beam evaporator and a metallization of 300 Å titanium (Ti) followed by 1000 Å of gold was completed. A metal liftoff

using acetone, isopropyl, and DI water removed the excess metal. This fabrication process is graphically illustrated in Figure 10.

5. Experimental Characterization

Experimental characterization results for samples that were fabricated using the method described earlier are shown in Figures 11 and 12. For these samples, the ZnSe substrate thickness was fixed at 1.8 μm and the linear gold gratings were spaced 3.0 μm with a 50% duty cycle.

The IR reflectance and emission measurements were made using a Nicolet 560 Fourier transform infrared (FTIR) spectrometer with a near-normal incidence reflectivity module and an input port for collecting IR emission or photoluminescence. The reflectivity was taken at room temperature as a function of incident polarization. Although the polarized emission could be easily detected at room temperature, the signal-to-noise ratio was improved by taking the data at an elevated temperature. The experimental results, which closely match the modeled results, demonstrate a strong resonant absorption and thermal emission near the designed wavelength.

6. Alternative Polarization Insensitive Designs

One disadvantage of using the resonant particles described in Figure 2 is their sensitivity to polarization. This reduces the total absorbed energy by one half. To address this issue, we explored a number of designs that were less sensitive to incident field polarization. These structures, shown in Figure 13, consist of 2D arrays of gold strips (known commonly as a fishnet structure), metallic patches, and circular holes. Each of the structures shown in Figure 11 was analyzed using the RCW method. Of those structures analyzed, the inductive grid array (Figure 13(c)) showed the most promise. Figure 12 presents numerical simulations of normal incident reflectance as a function of wavelength. A strong, nearly perfect, absorption is predicted for both parallel and perpendicular polarization. Moreover, as in the previous designs, the resonant wavelength was easily tuned by simply varying the thickness of the dielectric substrate layer. It should be noted that the results shown in Figure 14 have not been experimentally validated yet.

7. Conclusions

In this paper, we presented a design methodology to create small particles characterized by a strong resonant absorption within the LWIR (8–14 μm) band. Our method combined a surface plasmon resonance, created using a subwavelength metallic grating with a dielectric cavity resonance. We showed that by varying the thickness of the cavity substrate the resonances could be tuned anywhere within the LWIR band. Experimental samples were fabricated using photolithography and experimentally characterized. The experimental results compared favorably with the calculated results. We believe that material-by-design particles, such

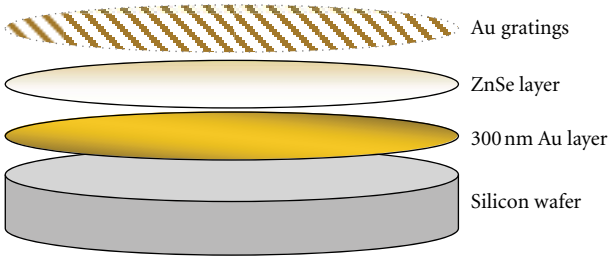


Figure 10: Illustration of the fabrication steps.

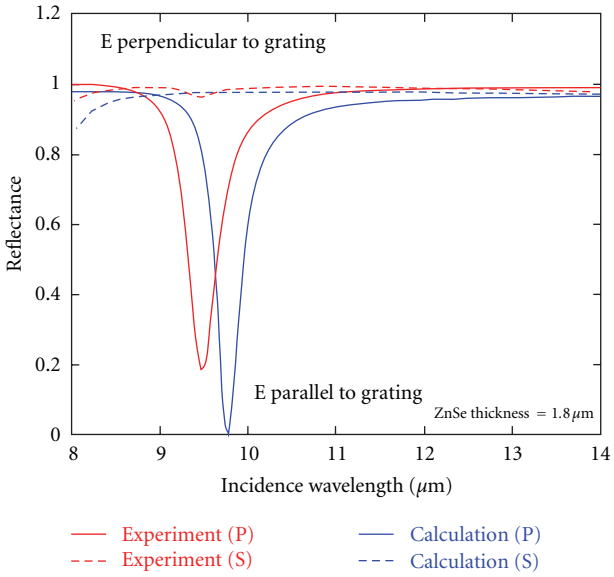


Figure 11: A comparison of predicted (using RCW code) and measured reflectance for both parallel (P type) and perpendicular (S type) polarizations.

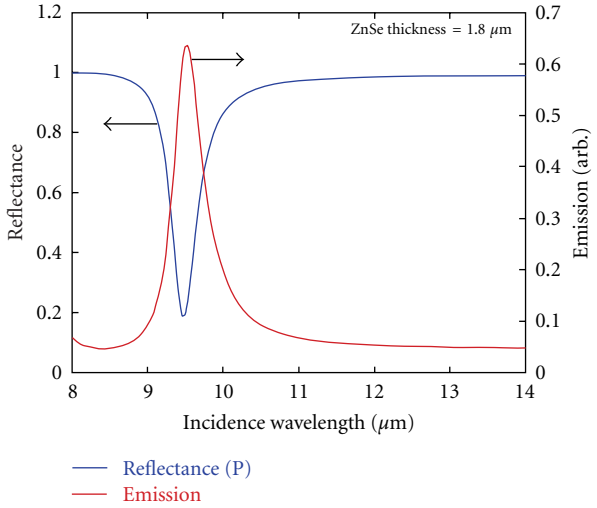
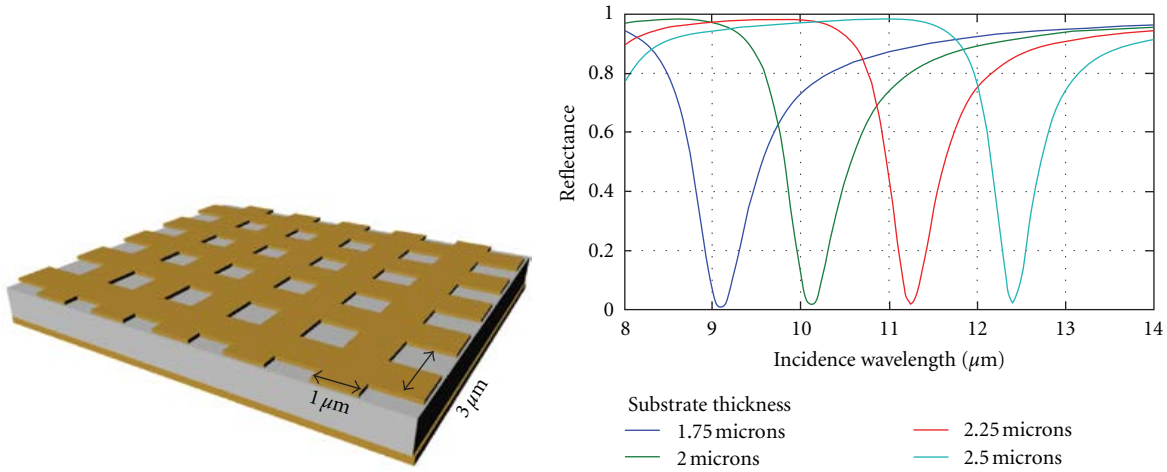


Figure 12: Experimentally measured reflectance curve and emission curve clearly demonstrating the resonant nature of our design.

(a) Hole array (b) Capacitive patch array (c) Inductive grid array (fishnet)

Figure 13: Designs less sensitive to polarization effects.



(a) (b)

Figure 14: RCW simulations for inductive grid array shown in Figure 11(c). Here as the substrate, assumed to be ZnSe, is varied from 1.75 to 2.5 μm. The resonant absorption wavelength shifts to longer wavelengths; however, the magnitude of the absorption remains near perfect.

as the ones described here, would be useful for a variety of remote atmospheric sensing applications. In those applications, which require relatively small particles, a custom spectral signature with multiple wavelengths would be achieved by mixing batches of single-wavelength particles designed for the component wavelengths. But in other applications, such as calibrated surfaces for hyperspectral imager testing and training, the surfaces could be larger and the multiple wavelengths could be designed into a single surface by implementing a checkerboard subcells with different grating periods across the surface. By properly selecting the frequencies and relative areas of the emitting subcells, the emission spectrum could be designed to mimic the spectral emission from specific natural surfaces.

References

[1] K. Masuno, S. Kumagai, and M. Sasaki, "Reflection-type wavelength-selective infrared emitter using surface plasmon polaritons," *Optics Letters*, vol. 36, no. 3, pp. 376–378, 2011.

[2] J. Hendrickson, J. Guo, B. Zhang, W. Buchwald, and R. Soref, "Wideband perfect light absorber at midwave infrared using muliplexed metal structures," *Optics Letters*, vol. 37, no. 3, pp. 371–373, 2012.

[3] K. Boratay Alici, A. Burak Turhan, C. M. Soukoulis, and E. Ozbay, "Optically thin composite resonant absorber at the near-infrared band: a polarization independent and spectrally broadband configuration," *Optics Express*, vol. 19, no. 15, pp. 14260–14267, 2011.

[4] N. Liu, M. Mesch, T. Weiss, M. Hentschel, and H. Giessen, "Infrared perfect absorber and its application as plasmonic sensor," *Nano Letters*, vol. 10, no. 7, pp. 2342–2348, 2010.

[5] C. H. Lin, R. L. Chern, and H. Y. Lin, "Polarization-independent broad-band nearly perfect absorbers in the visible regime," *Optics Express*, vol. 19, no. 2, pp. 415–424, 2011.

[6] P. Lalanne and J. P. Hugonin, "High-order effective-medium theory of subwavelength gratings in classical mounting: application to volume holograms," *Journal of the Optical Society of America A*, vol. 15, no. 7, pp. 1843–1851, 1998.

[7] M. G. Moharam and T. K. Gaylord, "Rigorous coupled-wave analysis of planar-grating diffraction," *Journal of the Optical Society of America*, vol. 71, no. 7, pp. 811–818, 1981.

[8] P. Lalanne, "Improved formulation of the coupled-wave method for two-dimensional gratings," *Journal of the Optical Society of America A*, vol. 14, no. 7, pp. 1592–1598, 1997.

[9] E. Noponen and J. Turunen, "Eigenmode method for electromagnetic synthesis of diffractive elements with three-dimensional profiles," *Journal of the Optical Society of America A*, vol. 11, no. 9, pp. 2494–2502, 1994.

[10] P. G. Etchegoin, E. C. Le Ru, and M. Meyer, "An analytic model for the optical properties of gold," *Journal of Chemical Physics*, vol. 125, no. 16, Article ID 164705, 3 pages, 2006.

Face Recognition Performance with Superresolution

Shuowen Hu, Robert Maschal, Susan Young, Tsai Hong Hong and P. Jonathon Phillips

Applied Optics 2012, 51 (18).

Face recognition performance with superresolution

Shuowen Hu,[1,*] Robert Maschal,[1] S. Susan Young,[1] Tsai Hong Hong,[2] and P. Jonathon Phillips[2]

[1]United States Army Research Laboratory, 2800 Powder Mill Road, Adelphi, Maryland 20783, USA

[2]NIST, 100 Bureau Drive, Gaithersburg, Maryland 20899, USA

*Corresponding author: shuowen.hu.civ@mail.mil

Received 29 September 2011; revised 19 April 2012; accepted 24 April 2012;
posted 30 April 2012 (Doc. ID 155384); published 20 June 2012

With the prevalence of surveillance systems, face recognition is crucial to aiding the law enforcement community and homeland security in identifying suspects and suspicious individuals on watch lists. However, face recognition performance is severely affected by the low face resolution of individuals in typical surveillance footage, oftentimes due to the distance of individuals from the cameras as well as the small pixel count of low-cost surveillance systems. Superresolution image reconstruction has the potential to improve face recognition performance by using a sequence of low-resolution images of an individual's face in the same pose to reconstruct a more detailed high-resolution facial image. This work conducts an extensive performance evaluation of superresolution for a face recognition algorithm using a methodology and experimental setup consistent with real world settings at multiple subject-to-camera distances. Results show that superresolution image reconstruction improves face recognition performance considerably at the examined midrange and close range.

OCIS codes: 100.0100, 100.6640, 100.4995, 100.2980.

1. Introduction

The affordability of surveillance systems has led to their widespread usage on commercial properties and for residential monitoring. Consequently, video footage of criminal activity is often available to law enforcement to help identify suspects. Therefore, face recognition software is a crucial tool that the law enforcement community may use to search watch lists and criminal databases to identify the suspect acquired on video. However, typical low-cost surveillance systems have small pixel counts. Furthermore, the suspect could be far away from the camera, resulting in images with very limited number of pixels on the face (i.e., low face resolution).

Research studies have shown that while face recognition algorithm performance is dependent on face resolution, this dependence is highly nonlinear. Boom *et al.* [1] examined the effect of resolution on face recognition and observed that performance became severely degraded for face images with sizes less than 32×32 pixels. However, performance was observed to be fairly similar for face images with sizes ranging from 32×32 pixels to 128×128 pixels [1], substantiating the highly nonlinear nature of face recognition performance with respect to face resolution. The Facial Recognition Vendor Test 2000 [2] also observed that the evaluated face recognition systems yielded similar performance for face images with face resolutions of 30 to 60 pixels measured in terms of eye-to-eye distance, but that performance severely degraded for some algorithms at an eye-to-eye distance of 15 pixels. In the authors' experience of working with law enforcement agencies, it is not uncommon for faces in typical surveillance footage from residential and commercial properties to have resolutions less than 30 pixels in terms of eye-to-eye distance, especially when the suspect is far away from the camera. Therefore, the limited face resolution within surveillance footage is a major obstacle for face recognition software. Pennsylvania Justice Network (JNET) states that low resolution and distance are two of the main factors that limited face recognition effectiveness of its statewide implementation of a facial recognition search system for investigators [3].

Face recognition continues to be an active area of research focused on improving performance through the development of new feature transforms, classification techniques, and mathematical frameworks to handle the large variability of face imagery found in real life. Many factors contribute to this variability: illumination, pose, and scale/resolution are several of the main factors. While research has been predominantly focused on solving the pose and illumination challenges for face recognition, some efforts have been devoted to solving the face resolution problem through the use of superresolution image reconstruction, which utilizes a sequence of low-resolution (LR) images containing the face in the same pose to reconstruct a high-resolution face image with more details. Boult *et al.* [4] proposed a superresolution method via image warping for face recognition, and Baker and Kanade [5] proposed hallucinating faces through a Gaussian pyramid-based method; however, these works did not conduct a performance evaluation to assess the benefit of superresolution for face recognition. More recently, Wheeler *et al.* [6] developed a multiframe face superresolution method with an active appearance model for registration and evaluated the face recognition improvement using the Identix FaceIt software. However, only 138 images (split between six ranges in terms of eye-to-eye distance) from three test subjects were used in [6]. Due to the small sample size, the observed improvement with superresolution is unlikely to be statistically meaningful in [6]. Whereas [4–6] perform superresolution image enhancement in the pixel domain prior to the face recognition algorithm, Gunturk *et al.* [7] developed an eigenface-domain superresolution technique for face recognition. The algorithm of [7] performs superresolution reconstruction in a low-dimensional face space through principal component analysis (PCA)-based dimensionality reduction and showed an improvement in face recognition performance with a minimum distance classifier in the eigenspace. In contrast to [4–7], Hennings-Yeomans *et al.* [8] proposed an algorithm incorporating face features into superresolution as prior information, and Huang *et al.* [9] developed a superresolution approach based on correlated features and nonlinear mappings between low-resolution and high-resolution features. Fookes *et al.* [10] conducted the most recent work on superresolution for face recognition, evaluating the performance of two face recognition algorithms with three superresolution techniques. However, as in [7–9], Fookes *et al.* [10] also utilized synthetically generated LR face images by downsampling the original high-resolution imagery. Downsampled face imagery does not accurately depict real-world compressed surveillance face images at varying subject-to-camera distances, especially since compression is highly nonlinear with more pronounced effects on facial details for far subject-to-camera ranges. Although [10] also used a white Gaussian noise corrupted version of the downsampled sets, the added white Gaussian noise does not resemble compression artifacts. The goal of this work is to conduct a comprehensive performance assessment of a state of the art baseline face recognition algorithm [11,12] with the pixel-level superresolution method of Young *et al.* [13] using a large database containing videos similar to real-world surveillance footage.

Specifically, the objectives of this work are to (a) assess the benefit of superresolution for face recognition with respect to subject-to-camera range, (b) assess face recognition performance using superresolved imagery reconstructed using varying numbers of LR frames, and (c) evaluate face recognition performance of individual frames within the LR sequence as well as the performance of a decision level fusion of the sequence to compare with superresolution face recognition results. The database of moving faces and people acquired by O'Toole *et al.* [14] was used for this study, specifically the parallel gait video datasets and close-up mug shots. Face recognition performance with the LR and superresolved imagery was assessed with the local region principal component analysis (LRPCA) face recognition algorithm [11,12] developed at Colorado State University. Correct verification rates are calculated and compared at three face resolutions/scales in terms of eye-to-eye distance corresponding to different subject-to-camera distances within the video footage. Results show that superresolution image reconstruction significantly improves face recognition verification rates at the examined mid- and close ranges, with some improvement at the far range.

2. Methodology

A. Database

Parallel gait videos and static mug shot images from the video database of moving face and people [14] are used for this work. The parallel gait video shows the subject moving towards the camera from 13.6 m away to approximately 1.5 m away, providing a large sequence of face imagery at different face resolutions from which query sets can be formed. A sample frame containing the subject at the far range is shown in Fig. 1. Since faces in the parallel gait videos are acquired from the frontal perspective, the corresponding frontal mug shots are used to form the gallery set. The resolution of the videos as well as of the frontal mug shot is 720×480 pixels (note that the corresponding pixel count is 345,600 pixels, substantially less than even one megapixel). The videos were acquired with compression using a Canon Optura Pi digital video camera. Figure 2 shows (a) close range face image of a subject, (b) close range face image downsampled by a factor of 3 to simulate far range using procedure of [10], and (c) far range face image of the subject taken from the same video. The downsampling procedure of [10] involved convolving the close range face image with a Gaussian filter of $d/4$ and then downsampling by d, where d is the

Fig. 1. (Color online) Sample frame extracted from a subject's parallel gait video in the database of moving faces and people [14]. Subject is at the far range (resulting eye-to-eye distance of 5–10 pixels).

downsampling factor. Note that while the compression artifacts are not obvious in the close range face image (simply because facial features consist of many pixels), the compression distortions are highly pronounced in the actual far range face image taken from the same video. Simulating the far range face image by downsampling, as past studies have done in examining superresolution for face recognition, does not closely resemble actual far range face imagery due to the effects of compression. The parallel gait videos used in this work emulate real-world compressed surveillance footage and enable a realistic assessment of superresolution benefit for face recognition.

B. Superresolution

This study used the reconstruction-based superresolution algorithm of Young and Driggers [13], which utilizes a series of undersampled/aliased LR images to reconstruct an alias-free high-resolution image. This reconstruction-based superresolution algorithm consists of a registration stage and a reconstruction stage. The registration stage computes the gross shift and subpixel shift of each frame in the sequence with respect to the reference frame using the correlation method in the frequency domain. The reconstruction stage uses the error-energy reduction method with constraints in both spatial and frequency domains, generating a superresolved image that improves the high-frequency content that was lost or corrupted due to the undersampling/aliasing of the sensor. The resolution improvement factor of the superresolved image is the square root of the number of frames used to reconstruct the superresolved image. A necessary condition for superresolution benefit is the presence of different subpixels shifts between frames to provide distinct information from which to reconstruct a high-resolution image. The natural movement of the subject in the parallel gait video provided this necessary subpixel shift.

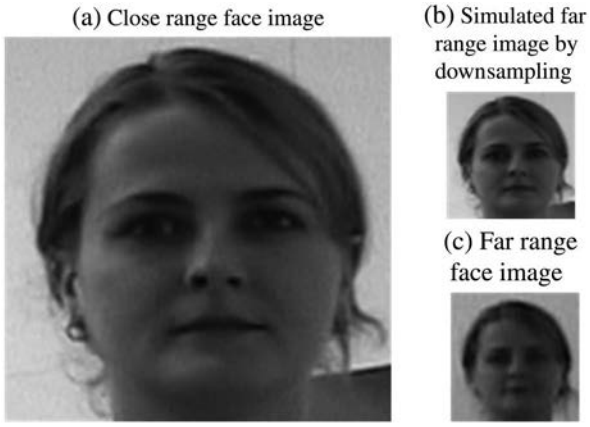


Fig. 2. (a) Close range face image of a subject, (b) close range face image downsampled by a factor of 3 to simulate far range using procedure of Fookes *et al.* [10], and (c) far range face image of the subject taken from the same video.

C. Query Sets

Frame sequences at three different subject-to-camera distances are extracted from each subject's parallel gait video: far range (~13 m), midrange (~9 m), and close range (~5 m). The face resolutions (in terms of eye-to-eye distances) corresponding to the far, mid-, and close ranges are 5–10, 15–20, and 25–30 pixels, respectively. Three query sets are constructed for each range: (a) original LR imagery (taken as the first frame within the sequence), (b) superresolved imagery using four consecutive LR frames (SR4), and (c) superresolved imagery using eight consecutive LR frames (SR8). SR4 and SR8 enable an assessment of the impact of the number of frames used for superresolution on face recognition performance. The resolution improvement factor in the x and y directions is 2 and 2.8 for SR4 and SR8, respectively. Consequently, the size of the SR4 face image is a factor of 2 larger in the x and y dimensions than the corresponding LR face image; the size of the SR8 face image is a factor of 2.8 larger in the x and y dimensions than the LR face image. A total of nine different query sets (Table 1) are generated to evaluate the improvement in face recognition with superresolution; each query set contains 200 subjects with one image per subject.

D. Face Recognition

This study used the state-of-the-art baseline LRPCA face recognition algorithm developed by Bolme *et al.*

Table 1. Query Set Nomenclature[a]

	5–10 Pixels	15–20 Pixels	25–30 Pixels
Low-resolution	$LR_{5–10}$	$LR_{15–20}$	$LR_{25–30}$
Superresolved 4 frames	$SR4_{5–10}$	$SR4_{15–20}$	$SR4_{25–30}$
Superresolved 8 frames	$SR8_{5–10}$	$SR8_{15–20}$	$SR8_{25–30}$

[a]Top row represents subject-to-camera range in terms of eye-to-eye distance.

[11,12]. The high-frequency content recovered in the superresolved imagery is expected to aid principal component analysis (PCA)-based methods, since current PCA-based algorithms often employ a large number of basis vectors (on the order of thousands for this study). As a preprocessing step, all query and gallery images are cropped and normalized to 256×256 pixels through bilinear interpolation using manually defined eye coordinates. The LRPCA algorithm was trained using the "The Good, The Bad, and The Ugly" (GBU) subset of the Multiple Biometric Grand Challenge, containing a total of 522 subjects. Training on a separate dataset distinct from the query and gallery sets avoids biasing the performance of the algorithm. The gallery set corresponding to each query set consists of one frontal mug shot for each subject.

E. Performance Measurement

For each query set and gallery, the output of the LRPCA face recognition algorithm is a similarity matrix S containing the similarity measure between every probe in the query set and every gallery image. Note that for this work, both the query and gallery sets contain a single image of each subject ($N = 200$ subjects total); therefore, the similarity matrix is a $N \times N$ square matrix with the diagonal elements containing the N match scores and $N(N-1)$ off-diagonal elements containing the nonmatch scores.

1. *Receiver Operating Characteristic Curves*

The similarity matrix is used to compute the correct verification rates as well as the corresponding false accept rates (FARs). In the verification model, the face recognition system is tasked with deciding whether the person in the probe image p_i is the same as the person in the gallery imagery g_j [15]. The decision is made based on the Neyman–Pearson theorem, testing whether the similarity score between p_i and g_j exceeds a given threshold t_0. The correct verification rate is computed by tallying the number of diagonal elements (match scores) that exceed t_0, and the FAR is computed by tallying the number of off-diagonal elements (nonmatch scores) that exceed t_0 [15]. Receiver operating characteristic (ROC) curves were generated by thresholding the similarity matrix S at various thresholds across the range from $S_{\min}$ to $S_{\max}$. For each of the nine query sets listed in Table 1, a ROC curve was constructed in this manner.

2. *Performance with Respect to Range*

To visualize face recognition performance with respect to subject-to-camera range, the correct verification rates are plotted with respect to range at commonly used FARs of 0.01 and 0.05 for LR, SR4, and SR8. Confidence intervals are calculated and overlaid onto the plots to assess the statistical reliability of the performance improvement achieved with superresolution image reconstruction.

3. *Confidence Intervals*

To indicate the reliability of the calculated correct verification rates, 95% confidence intervals are determined using the bootstrap method, specifically following the procedure for biometrics detailed in [16]. The bootstrap is a nonparametric approach that makes no assumptions about the error distribution and is preferable to parametric techniques when the underlying distribution is unknown, as is the case for biometrics. Bootstrap involves resampling the available data (match scores for this study) many times with replacement to generate confidence intervals. For this work, the probe set contained one image per subject and the gallery contained one image per subject for LR, SR4, and SR8 at each range, satisfying the independent and identically distributed (*i.i.d.*) requirement of the bootstrap.

Recall that the output of the LRPCA algorithm is a similarity matrix S containing scores of the similarity between a probe and all gallery images. Also recall that S contains $M = N$ match scores along the diagonal and $N(N-1)$ mismatch scores, where $N = 200$ is the number of subjects. For a given t_0, let the verification rate estimate be defined by the equation

$$\hat{F}(t_0) = \frac{1}{M}\sum_{i=1}^{M} 1(X_i \geq t_0), \tag{1}$$

where X denotes the set of M match scores and 1 is the indicator function [16]. The bootstrap generates $X^* = \{X_1^*, \ldots, X_M^*\}$ by resampling with replacement, and then calculates $\hat{F}^*(t_0)$. This resampling procedure is repeated B times ($B = 10,000$ for this work), generating bootstrap estimates $\hat{F}^* = (\hat{F}_1^*, \hat{F}_2^*, \ldots, \hat{F}_B^*)$. The lower and upper bounds of the 95% confidence interval is determined as values corresponding to the 2.5th and 97.5th percentile of the histogram of the B bootstrap estimates $\hat{F}^*$.

4. *Face Recognition Performance of Individual Frames and Decision Level Fusion*

To address the question of how face recognition performance with superresolution compares to face recognition performance of individual frames within the LR sequence as well as to the performance of a decision level fusion scheme, further analysis was conducted. Superresolution exploits the additional spatial information contained in the temporal dimension (i.e., multiple frames) to reconstruct a more detailed face image for recognition, and therefore is expected to exceed the face recognition performance of any single frame within the LR sequence. To validate this expectation, face recognition performance was computed for each of the eight LR frames used to reconstruct SR8 and compared to face recognition performance of SR8. Furthermore, a simple fusion scheme for the LR frame sequence was implemented by averaging the similarity matrices from the eight LR frames. Fusion by averaging of similarity

matrices exploits the spatial information in the temporal domain at the decision level and is expected to be an upper bound on the face recognition performance of any individual frame. Face recognition performance with superresolved imagery is then compared to the performance of this decision level fusion scheme.

A total of 24 query sets (eight query sets per range corresponding to each of the eight frames in the LR sequence used to reconstruct SR8; Table 2) was generated to assess the variation in face recognition performance with respect to individual frames. Note that the LR sequence is an eight frame clip of the subject walking towards the camera. Due to the fast frame rate (30 Hz) and relatively slow speed of the subjects (walking speed), the change in pose is insignificant across the eight frames in the sequence. At the far range, since the change in face size across the eight frames does not exceed a single pixel, the same eye coordinates in terms of (x, y) pixel locations are used for all eight frames. At the mid- and close ranges, face size does enlarge by a few pixels across the frames; therefore, eye coordinates are manually picked for all eight frames instead of for just the first frame as in the far range. Once the similarity matrix for each query set is computed with the LRPCA algorithm, ROC curves of face recognition performance with respect to individual frames can be generated. The decision level fusion method (denoted LR^{ave}) averages the similarity matrices across the eight frames at each range and generates the ROC curve using the averaged similarity matrix for comparison with face recognition using superresolved imagery.

3. Results and Discussion

A. Superresolved Imagery

Superresolved face imagery and original low resolution face imagery are shown in Fig. 3 at different ranges. At the far range, the LR image is heavily pixilated and distorted by compression, yielding a coarse facial outline and few facial features. Superresolution with four and eight frames enhances the facial outline and some facial details, but compression artifacts have almost completely eliminated facial details in the low resolution frames, preventing significant facial feature enhancement.

As range decreases, the camera captures finer details and the detrimental impact of compression on facial features lessens because the size of these features is now larger. At the midrange, SR4 and SR8 produce considerable enhancement of the subject's facial details. As range continues to decrease to the close range, superresolution benefit decreases as facial features become more and more defined in the low resolution imagery. Although the close range SR images may not appear significantly enhanced visually, facial recognition algorithms may still benefit from superresolution as these algorithms operate on different principles than the human visual system.

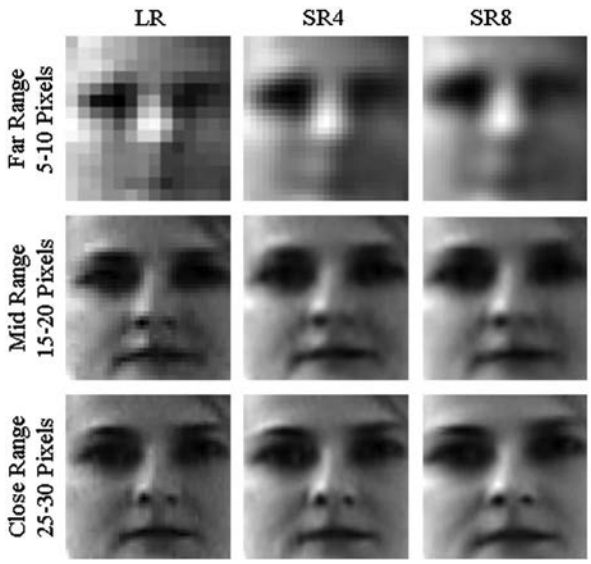


Fig. 3. Low-resolution (LR) imagery and superresolved imagery (4 frames—SR4, 8 frames—SR8) at eye-to-eye distances of 5–10, 15–20, and 25–30 pixels. All images at all ranges have been resized to a fixed size for comparison.

To provide a more objective assessment of the increase in high-frequency content with superresolution, spectral analysis was conducted using LR and superresolved face imagery. Let the following equations denote the cumulative power spectrum in wavenumber k_x and k_y, respectively, where $F(k_x, k_y)$ is the Fourier transform of the considered image:

$$S_1(k_x) = \sum_{k_y} |F(k_x, k_y)|^2, \tag{2}$$

$$S_2(k_y) = \sum_{k_x} |F(k_x, k_y)|^2. \tag{3}$$

For this study, the cumulative power spectrum is computed over the spatial region consisting of the eyes, which is a critical area for face recognition algorithms. Figure 4 shows the computed k_x-domain spectrums for LR and SR8 eye region at the midrange. The circled part of the plot in Fig. 4 represents the

Table 2. Query Set Nomenclature for Evaluation of Face Recognition with Respect to Low-Resolution Frame[a]

	Frame 1	Frame 2	Frame 3	Frame 4	Frame 5	Frame 6	Frame 7	Frame 8
Far range	$LR^1_{5–10}$	$LR^2_{5–10}$	$LR^3_{5–10}$	$LR^4_{5–10}$	$LR^5_{5–10}$	$LR^6_{5–10}$	$LR^7_{5–10}$	$LR^8_{5–10}$
Midrange	$LR^1_{15–20}$	$LR^2_{15–20}$	$LR^3_{15–20}$	$LR^4_{15–20}$	$LR^5_{15–20}$	$LR^6_{15–20}$	$LR^7_{15–20}$	$LR^8_{15–20}$
Close range	$LR^1_{25–30}$	$LR^2_{25–30}$	$LR^3_{25–30}$	$LR^4_{25–30}$	$LR^5_{25–30}$	$LR^6_{25–30}$	$LR^7_{25–30}$	$LR^8_{25–30}$

[a]Superscript denotes frame number, while subscript denotes range in terms of eye-to-eye distance in pixels.

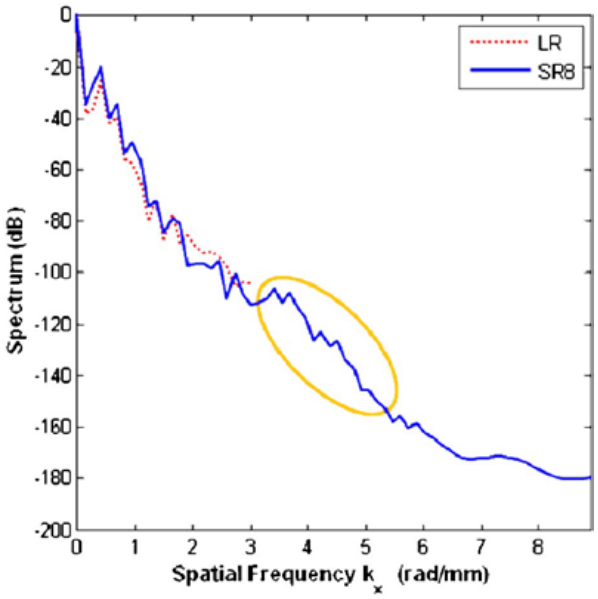


Fig. 4. (Color online) Computed k_x-domain cumulative power spectrums of LR eye region and SR8 eye region at the midrange. The circled part of the plot represents high-frequency band recovered from using a sequence of eight aliased low-resolution frames.

high-frequency band recovered with superresolution using a sequence of eight aliased LR images.

Furthermore, the enhancement in edge contrast is demonstrated in Fig. 5, which plots the intensity values of LR and SR8 at the midrange along a horizontal profile across the eyes. The improvement in edge contrast is especially noticeable across the pupils (horizontal axis = ±10) in Fig. 5.

B. Receiver Operating Characteristic Curves

ROC curves at the 5–10, 15–20, and 25–30 pixel eye-to-eye distance are shown in Figs. 6, 7, and 8, respectively. Each figure contains three ROC curves corresponding to the LR (red dotted line),

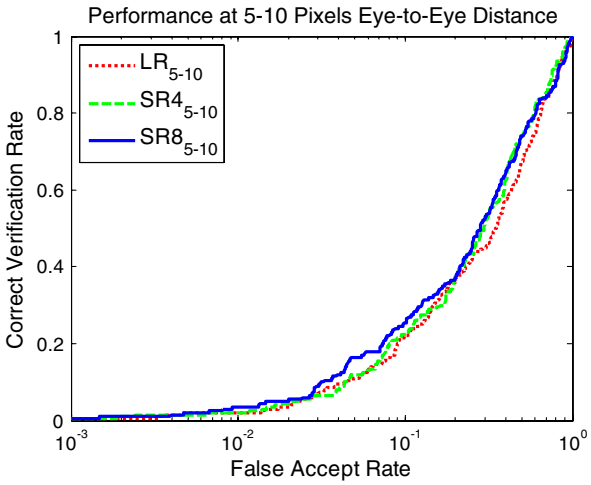


Fig. 6. (Color online) ROC curves at the far range for original low resolution ($LR_{5–10}$) query set and the corresponding superresolved query sets using four ($SR4_{5–10}$) and eight ($SR8_{5–10}$) frames.

superresolved using four frames (SR4; dashed green line), and superresolved using eight frames (SR8; solid blue line) imagery. At the far range, the ROC curves for $SR4_{5–10}$ and $SR8_{5–10}$ lay slightly but consistently above the ROC curve for $LR_{5–10}$, suggesting that face imagery at the far range possessed too few details for superresolution to provide any substantial enhancement to aid the LRPCA face recognition algorithm. At the midrange, $SR8_{15–20}$ outperformed $SR4_{15–20}$, which in turn outperformed $LR_{15–20}$ across the FARs from FAR = 0.001 to 0.6; superresolution effectively enhances facial details at the midrange to yield a large improvement in face recognition performance over the baseline LR imagery. At the close range, while both $SR4_{25–30}$ and $SR8_{25–30}$ produced higher face recognition performance than $LR_{25–30}$ at all FARs, the improvement is not as large as

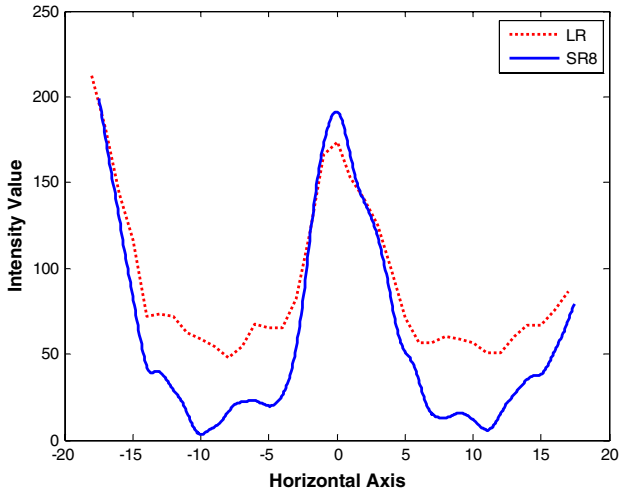


Fig. 5. (Color online) Pixel intensity value plots of LR and SR8 along a profile across the eye region at the midrange, showing the improved edge contrast with SR8 in the spatial domain.

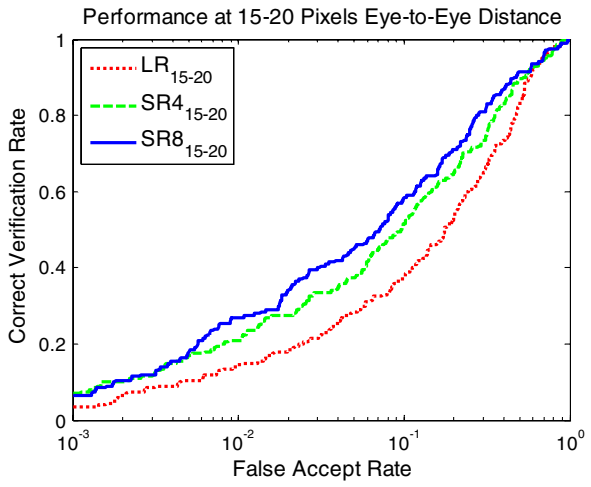


Fig. 7. (Color online) ROC curves at the midrange for original low resolution ($LR_{15–20}$) query set and the corresponding superresolved query sets using four ($SR4_{15–20}$) and eight ($SR8_{15–20}$) frames.

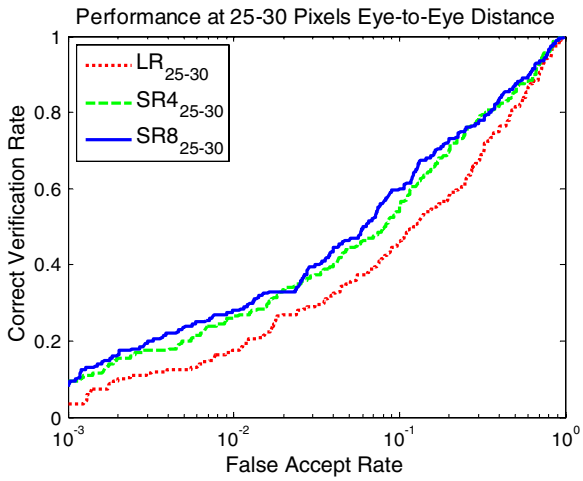


Fig. 8. (Color online) ROC curves at the close-range for original low resolution ($LR_{25–30}$) query set and the corresponding superresolved query sets using four ($SR4_{25–30}$) and eight ($SR8_{25–30}$) frames.

achieved at the midrange since the original imagery already contains detailed facial features.

C. Performance with Respect to Range

For practical applications, performance at low FARs is of particular interest; therefore, in verification rate as a function of range is examined at FAR = 0.01 and 0.05 in Fig. 9. At all FARs, the LR curve exhibits a slight knee at the midrange; the knee is more pronounced for the SR4 and SR8 curves, signifying that the change in performance with respect to range is more nonlinear for SR imagery.

At the far range, the already limited facial details are distorted by compression, preventing substantial enhancement by superresolution image reconstruction. At the midrange where the greatest benefit from superresolution is observed, the correct verification rate is 21.0% for $SR4_{15–20}$ and 27.0% for $SR8_{15–20}$ compared to 14.5% for $LR_{15–20}$, resulting in an improvement of 44.8% and 86.2% at FAR = 0.01, respectively. At FAR = 0.05, the midrange correct verification rate is 37.5% for $SR8_{15–20}$ and 45.0% for $SR8_{15–20}$ compared to 28.5% for $LR_{15–20}$, resulting in an improvement of 31.6% and 57.9%, respectively.

A large improvement of the verification rate occurs from the far range to the midrange, but the improvement is visibly smaller from the midrange to the close range. For SR8, which produced effective eye-to-eye distances ~2.8 times the original size; the verification rate exhibited only a small improvement from the midrange to the close range. These results are consistent with the findings of [1,2] that showed the improvement in face recognition performance slowed considerably once the eye-to-eye distance surpassed approximately 30 pixels.

To generate the confidence intervals shown in Fig. 9, the procedure described in Subsection 2.E.3 was performed using the similarity matrix S for LR, SR4, and SR8 at each of the three ranges. For the far range, although face recognition improves with superresolution, the confidence intervals overlap for LR, SR4, and SR8, suggesting that no significant benefit is achieved with superresolution. At the close range, the confidence interval for SR4 exhibits a partial overlap with that of LR, and the confidence interval for SR8 exhibits only a slight overlap with that LR. The small overlaps suggest that superresolution improves recognition rates for face recognition with high reliability, especially when using eight frames. At the midrange, the confidence interval for SR4 partially overlaps with that of LR, and the confidence interval for SR8 exhibits no overlap at all with that of LR. The lack of any overlap demonstrates that the face recognition performance improvement achieved with superresolution using eight frames is not only highly reliable, but is also significant for the midrange where eye-to-eye distance is between 15–20 pixels.

D. Face Recognition Performance of Individual Frames and Decision Level Fusion

To examine the LRPCA face recognition performance of individual frames within the LR sequence, the ROC curves of each LR frame are computed and shown in Figs. 10–12. The ROC curve of the simple decision level fusion method derived from averaging the similarity matrices across the eight

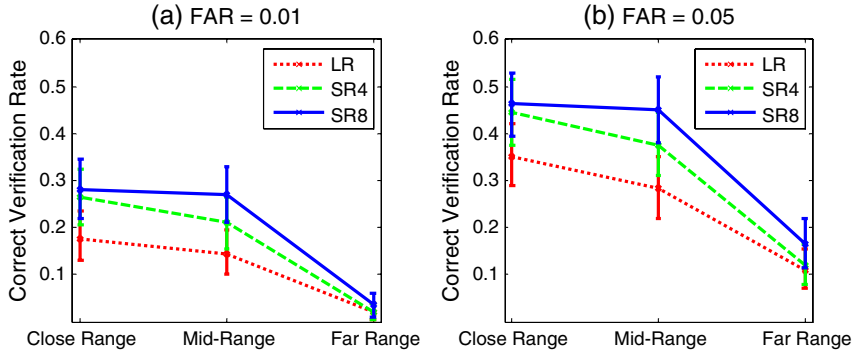


Fig. 9. (Color online) Performance as a function of range at FARs of (a) 0.01 and (b) 0.05. Error bars show the 95% confidence interval for each correct verification rate.

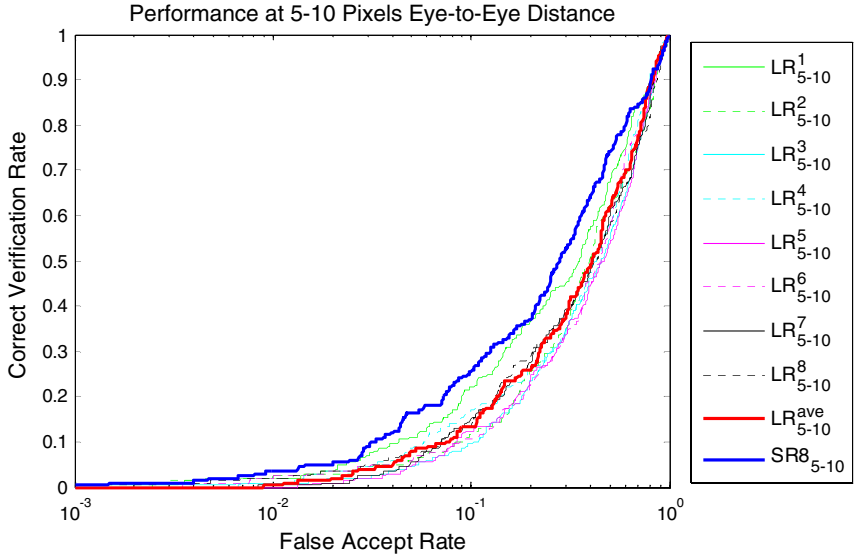


Fig. 10. (Color online) ROC curves at the far-range for each low resolution frame (superscript 1–8). The ROC curve for LR^{ave} is generated by averaging similarity matrices of the eight individual frames and generating the ROC curve.

frames (LR^{ave}) at each range is overlaid onto the plots. Averaging the similarity matrices exploits the spatial information across the eight frames at the decision level for face recognition and is compared against face recognition with superresolution (SR8).

Figures 10–12 show the ROC curves for each of the 8 LR query sets corresponding to different frames at the far, mid-, and close ranges, respectively. The ROC curve for the average similarity scores (LR^{ave}) across frames is shown in bold red along with the SR8 ROC curve shown in bold blue. The ROC curves for the LR frames exhibit some variation, but tend to be clustered together and lay within the confidence intervals as computed in Subsection 3.C. Note that there is no observable ordering of the ROC curves for LR^1 to LR^8 from lowest to highest. This verifies that the scale change across the eight frames as subject walks towards the camera is minor and does not produce any patterns in the ordering of the ROC curves. At the far range, the ROC curve for the first frame (LR^1_{5-10}) interestingly lay above the seven other LR frames which closely overlap with each other. This may be due to the definition of the eye coordinates based on the first frame, which were then reused for the remaining seven frames at the far range. This eye coordinate definition procedure may have produced a slightly more accurate eye coordinate selection for the first frame than the other frames, even though the change in eye coordinates did not exceed an integer pixel across the frames in the sequence at the far range.

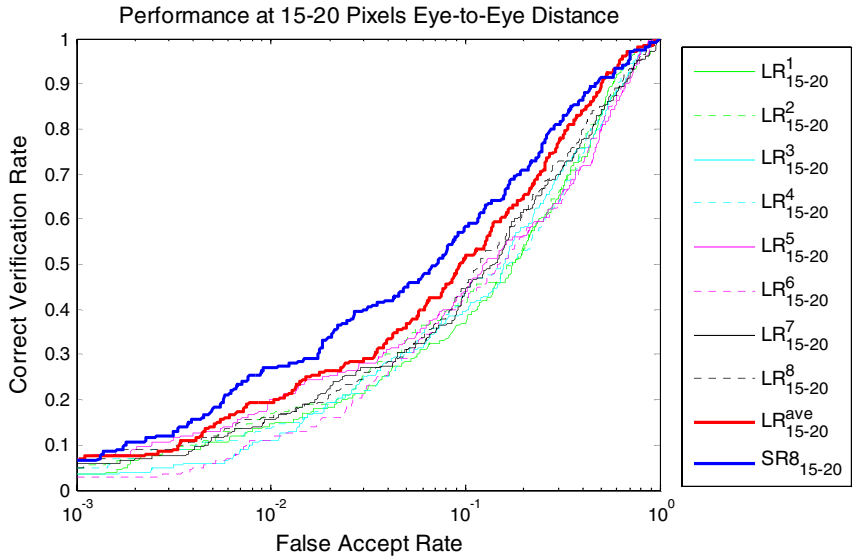


Fig. 11. (Color online) ROC curves at the midrange for each low resolution frame (superscript 1–8).

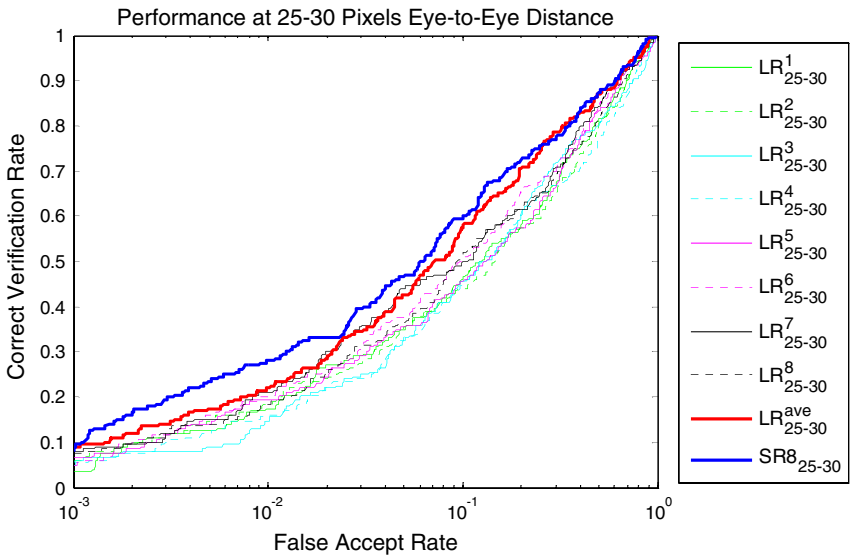


Fig. 12. (Color online) ROC curves at the close-range for each low resolution frame (superscript 1–8).

The ROC curves for LR^{ave} in general lay above the ROC curve of any individual LR frame. Since LR^{ave} is a decision level fusion in exploiting the spatial information across the eight frames, it is not unexpected that the ROC for LR^{ave} tends to be an upper bound for the ROC curve of any individual frame. However, the ROC curve for SR8 lay above the ROC curve of LR^{ave} at all three ranges, showing that superresolution image reconstruction is a more effective method in exploiting the spatial information across the temporal dimension to improve face recognition performance.

To provide a quantitative measure of overall face recognition performance of the LRPCA face recognition algorithm using superresolved and LR imagery, the area under the curve (AUC) is computed for each ROC curve in Figs. 10–12 across FAR $\in [0, 1]$ and tabulated in Table 3. Note that the maximum possible value for the AUC is 1. The AUCs for LR are generally consistently close to each other across frames 1–8 at the three ranges. The "best frame" in terms of AUC is the 1st frame for the far range, 8th frame for the midrange, and 6th frame for the close range as underlined in Table 3. The AUC for SR8 is 5.1% larger than the best frame at the far range, 7.65% larger than the best frame at the midrange, and 5.04% larger than the best frame at the close range. Furthermore, the AUC for SR8 is 14.36% larger than LR^{ave} at the far range, 2.88% larger than LR^{ave} at the midrange, and 1.38% larger than LR^{ave} at the close range. Note that although the AUC for SR8 is only a few percent better than LR^{ave} at the mid- and close ranges, the increase is still substantial as the AUC was computed over the whole range of FARs (FAR $\in [0, 1]$); typically, ROC curves for detection/classification algorithms in a given experiment overlap at higher FARs (ex. FAR > 0.1). Therefore, the results of Table 3 show that superresolution provides improvement in LRPCA face recognition performance compared to any frame as well as to the decision level fusion across the eight frames.

Using the best frames in terms of AUC from Table 3 at each range (denoted LR^*), verification rates with respect to range are plotted in Fig. 13 at FAR of 0.01 and 0.05. SR8 outperforms LR^* as well as LR^{ave}, with small to no overlap of confidence intervals at the mid- and close ranges. At the midrange, the verification rate is 0.45 for SR8, 0.37 for LR^{ave}, and 0.31 for LR^*, representing a 21.6% improvement and a 45.2% improvement over LR^{ave} and LR^* at FAR $= 0.05$, respectively. At the close range, although the performance improvement achieved with SR8 is not as significant as the midrange, the benefit of superresolution is still substantial.

For surveillance systems on residential and commercial properties where low cost cameras are prevalent, faces of individuals captured on camera are commonly between 15 and 30 pixels across in terms of eye-to-eye distance which correspond to the examined mid- and close ranges. Superresolution

Table 3. Area under the Curves for LR^i, Where $i \in [1.8]$ Denotes the Frame Number, AUC for LR^{ave} (Computed from the ROC of the Average Similarity Scores across the Eight Frames), and AUC for SR8 at Far, Mid-, and Close Ranges[a]

	LR^1	LR^2	LR^3	LR^4	LR^5	LR^6	LR^7	LR^8	LR^{ave}	SR8
Far range	0.6200	0.5553	0.5339	0.5655	0.5311	0.5474	0.5583	0.5566	0.5697	0.6516
Midrange	0.7529	0.7514	0.7557	0.7542	0.7467	0.7411	0.7625	0.7756	0.8115	0.8349
Close range	0.7559	0.7428	0.7503	0.7352	0.7584	0.7823	0.7805	0.7679	0.8105	0.8217

[a]Underlined numbers denote the best frame in terms of AUC at each range.

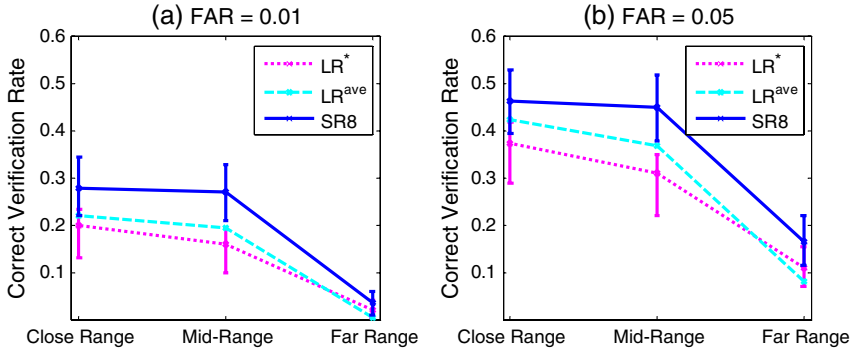


Fig. 13. (Color online) Performance as a function of range at FARs of (a) 0.01 and (b) 0.05. Error bars show the 95% confidence interval for each correct verification rate (not shown for LR^{ave} because LR^{ave} represents averaged similarity scores, and not actual similarity measurements). LR^{*} denotes the best of the eight LR frame (in terms of AUC) at each range.

is expected to provide significant benefits in enhancing the LR face images and improving facial recognition performance.

4. Conclusion

Using a video database similar to real world surveillance footage, this study shows that superresolution provides considerable benefits for the state of the art baseline LRPCA face recognition algorithm at the examined mid- and close ranges. In surveillance applications, low-cost cameras and oftentimes the far distance of individuals result in a very limited number of face pixels, severely affecting face recognition performance. Superresolution image reconstruction can be used to enhance the high-frequency content of low resolution surveillance imagery, improving face recognition performance and potentially aiding the nation in law enforcement and homeland security applications.

The authors thank Professor Ross Beveridge, David Bolme, Stephen Won, and Martha Givan for their help, as well as the reviewers for their valuable comments and suggestions.

References

1. B. J. Boom, G. M. Beumer, L. J. Spreeuwers, and N. J. Veldhuis, "The effect of image resolution on the performance of face recognition system," in *Proceedings of the 7th International Conference on Control, Automation, Robotics, and Vision* (IEEE, 2006), pp. 1–6.
2. D. M. Blackburn, M. Bone, and P. J. Phillips, "Facial Recognition Vendor Test 2000," http://www.frvt.org/FRVT2000/.
3. Pennsylvania Justice Network, "JNET facial recognition investigative search tool and watchlist," http://www.pajnet.state.pa.us/.
4. T. E. Boult, M.-C. Chiang, and R. J. Micheals, "Super-resolution via image warping," in *Super-Resolution Imaging*, S. Chaudhuri, ed. (Springer, 2001), pp. 131–169.
5. S. Baker and T. Kanade, "Hallucinating faces," in *Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition* (IEEE, 2000), pp. 83–88.
6. F. W. Wheeler, X. Liu, and P. H. Tu, "Multi-frame super-resolution for face recognition," in *Proceedings of IEEE 1st International Conference on Biometrics: Theory, Applications and Systems* (IEEE, 2007), pp. 1–6.
7. B. K. Gunturk, A. U. Batur, Y. Altunbasak, M. H. Hayes, III, and R. M. Mersereau, "Eigenface-domain super-resolution for face recognition," IEEE Trans. Image Process. **12**, 597–606 (2003).
8. P. H. Hennings-Yeomans, S. Baker, and B. V. K. V. Kumar, "Simultaneous super-resolution and feature extraction for recognition of low-resolution faces," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2008), pp. 1–8.
9. H. Huang and H. He, "Super-resolution method for face recognition using nonlinear mappings on coherent features," IEEE Trans. Neural Netw. **22**, 121–130 (2011).
10. C. Fookes, F. Lin, V. Chandran, and S. Sridharan, "Evaluation of image resolution and super-resolution on face recognition performance," J. Vis. Commun. Image Represent. **23**, 75–93 (2012).
11. D. S. Bolme and J. R. Beveridge, "CSU LRPCA baseline algorithm," www.cs.colostate.edu/facerec/algorithms/lrpca2010.php.
12. D. S. Bolme, J. R. Beveridge, M. Teixeria, and B. A. Draper, "The CSU face identification evaluation system: its purpose, features, and structure," Lect. Notes Comput. Sci. **2626**, 304–313 (2003).
13. S. S. Young and R. G. Driggers, "Super-resolution image reconstruction from a sequence of aliased imagery," Appl. Opt. **45**, 5073–5085 (2006).
14. A. J. O'Toole, J. Harms, S. L. Snow, D. R. Hurst, M. R. Pappas, J. H. Ayyad, and H. Abdi, "A video database of moving faces and people," IEEE Trans. Pattern Anal. Machine Intell. **27**, 812–816 (2005).
15. S. A. Rizvi, J. P. Phillips, and H. Moon, "The FERET verification testing protocol for face recognition algorithms," NIST IR 6281 (National Institute of Standards and Technology, 1998).
16. R. M. Bolle, N. K. Ratha, and S. Pankanti, "Error analysis of pattern recognition systems—the subsets bootstrap," Comput. Vis. Image Underst. **93**, 1–33 (2004).

Sparsity-Motivated Automatic Target Recognition

Vishal M. Patel, Nasser M. Nasrabadi and Rama Chellappa

Applied Optics 2011, 50 (10), 1425-1433.

Sparsity-motivated automatic target recognition

Vishal M. Patel,[1,*] Nasser M. Nasrabadi,[2] and Rama Chellappa[1]

[1]Department of Electrical and Computer Engineering, Center for Automation Research, University of Maryland, College Park, Maryland 20742, USA

[2]U.S. Army Research Laboratory, 2800 Powder Mill Road, Adelphi, Maryland 20783, USA

*Corresponding author: pvishalm@umiacs.umd.edu

Received 20 October 2010; revised 21 January 2011; accepted 29 January 2011; posted 1 February 2011 (Doc. ID 136904); published 29 March 2011

We present an automatic target recognition algorithm using the recently developed theory of sparse representations and compressive sensing. We show how sparsity can be helpful for efficient utilization of data for target recognition. We verify the efficacy of the proposed algorithm in terms of the recognition rate and confusion matrices on the well known Comanche (Boeing–Sikorsky, USA) forward-looking IR data set consisting of ten different military targets at different orientations.

OCIS codes: 100.0100, 100.5010, 100.3008, 330.5000.

1. Introduction

The objective of an automatic target recognition (ATR) algorithm is to detect [1] and classify each target image into one of a number of classes [2]. The recognition algorithm may consist of several stages. For example, in the first stage a target is detected on the entire image; in the second stage, background clutter is removed; in the third stage, a set of features is computed and finally, in the fourth stage, classification is done by means of a classifier. In this paper, we mainly focus on the last two stages.

Target recognition using forward-looking IR (FLIR) imagery of different targets in natural scenes is difficult due to large variations in the thermal signatures of targets. Many ATR algorithms have been proposed for FLIR imagery. Wang *et al.* proposed a modular neural-network-based ATR algorithm in [2]. In their algorithm, several neural networks are trained, each optimized for a local region in the image, whose classification decisions are combined to determine the final classification. Wavelet-based vector quantization was used for FLIR ATR in [3] by Chan and Nasrabadi, where a discriminative dictionary was created in the wavelet domain using learning vector quantization. A recognition method based on hidden Markov tree that uses a Karhunen–Loeve representation was proposed by Bharadwaj and Carin in [4]. See [5] for an excellent survey of papers and experimental evaluation of FLIR ATR. The algorithms evaluated in [5] include convolutional neural network (CNN), principal component analysis (PCA), linear discriminant analysis (LDA), learning vector quantization (LVQ), modular neural networks (MNN), and two model-based algorithms, using Hausdorff metric-based matching (H-M) and geometric hashing (G-H).

FLIR images often contain unwanted thermal signatures of the background clutter whose characteristics change with environment such as changes in fog, rain, and heat, which can make target detection and recognition difficult for automated as well as human observers. Recently, Wright *et al.* [6] introduced a sparse-representation-based classification (SRC) algorithm for face recognition, which is claimed to be robust to varying expressions, illumination, occlusion, and disguise, and has been shown to outperform many state-of-the-art algorithms. This approach is based on the theories of compressive sensing (CS) and sparse representation (SR). The idea is to create a dictionary matrix of the training samples as column vectors. The test sample is also represented as a column vector. Different dimensionality-reduction

0003-6935/11/101425-09$15.00/0

methods are used to reduce the dimensions of both the test vector and the vectors in the dictionary. One such approach for dimensionality reduction is using random projections [6]. Random projections, using a generated sensing matrix, are taken of both the dictionary matrix and the test sample. It is then simply a matter of solving an ℓ_1 minimization problem in order to obtain the sparse solution. Once the sparse solution is obtained, it can provide information as to which training samples the test vector most closely relates to. Furthermore, it was shown that if the sparsity of the solution is properly harnessed, the choice of features (e.g., dimensionality-reduction method) is no longer critical. The number of features for a given class and the sparse solution become critical.

Motivated by the SRC algorithm, in this paper, we investigate the effectiveness of SR and CS for the recognition of FLIR target images. In particular, we exploit the inherent block structure of the sparse solution induced by ℓ_1 minimization. Furthermore, our method utilizes a redundant dictionary that includes training data at various azimuth angles, hence achieving orientation invariance. As a result, our algorithm has the ability to identify targets at different orientations.

This paper is organized as follows: the theory of sparse representation along with its use for ATR is presented in Section 2. Its extensions based on block sparsity (BS) are presented in Section 3. In Section 4, we present some experimental results on a FLIR data set consisting of ten different targets. Section 5 concludes the paper with a brief summary and discussion.

2. Recognition Based on Sparse Representation

Following [6,7], in this section we briefly describe the use of SR and CS for FLIR ATR. Figure 1 shows the overview of our method.

A. Sparse Representation

Suppose that we are given L distinct target classes and a set of n training images per class. We identify an $l \times p$ gray-scale image as an N-dimensional vector that can be obtained by lexicographically stacking its columns, where $N = lp$. Let $\mathbf{A}_k = [\mathbf{x}_{k1}, \ldots, \mathbf{x}_{kn}]$ be an $N \times n$ matrix of training images from the kth class. That is, $\mathbf{A}_k$ represents the dictionary for class k. Define a new matrix or dictionary, $\mathbf{A}$, as the concatenation of subdictionaries from all the classes as

$$\begin{aligned}\mathbf{A} &= [\mathbf{A}_1, \ldots, \mathbf{A}_L] \in \mathbb{R}^{N \times (n.L)} \\ &= [\mathbf{x}_{11}, \ldots, \mathbf{x}_{1n} | \mathbf{x}_{21}, \ldots, \mathbf{x}_{2n} | \ldots\ldots | \mathbf{x}_{L1}, \ldots, \mathbf{x}_{Ln}].\end{aligned}$$

We consider an observation vector, $\mathbf{y} \in \mathbb{R}^N$, of unknown class as a linear combination of the training vectors as

$$\mathbf{y} = \sum_{i=1}^{L} \sum_{j=1}^{n} \alpha_{ij} \mathbf{x}_{ij} \tag{1}$$

with coefficients $\alpha_{ij} \in \mathbb{R}$. The above equation can be more compactly written as

$$\mathbf{y} = \mathbf{A}\alpha, \tag{2}$$

where

$$\alpha = [\alpha_{11}, \ldots, \alpha_{1n} | \alpha_{21}, \ldots, \alpha_{2n} | \ldots\ldots | \alpha_{L1}, \ldots, \alpha_{Ln}]^T \tag{3}$$

and $.^T$ denotes the transposition operation. Now we make an assumption that given sufficient training samples of the kth class, $\mathbf{A}_k$, any new test image $\mathbf{y} \in \mathbb{R}^N$ that belongs to the same class will approximately lie in the linear span of the training samples from the class k. This implies that most of the coefficients not associated with class k in Eq. (3) will be close to zero. Hence, α is a sparse vector.

In order to represent an observed vector, $\mathbf{y} \in \mathbb{R}^N$, as a sparse vector, α, one needs to solve the system of linear equations in Eq. (2). Typically, $L.n \gg N$,

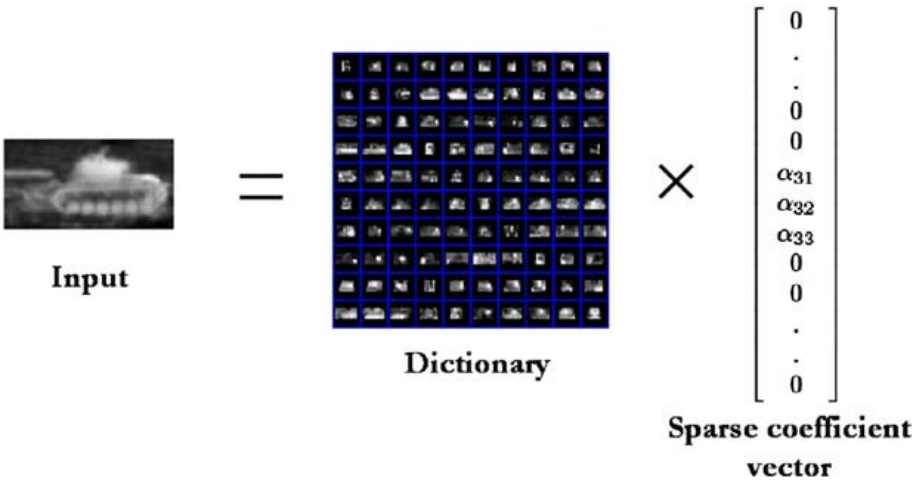


Fig. 1. (Color online) Overview of our approach using SR. Test target chip is represented as a linear combination of image chips from a dictionary containing all training images.

which makes the system of linear equations in Eq. (2) underdetermined and has no unique solution. It has been shown that if α is sparse enough and $\mathbf{A}$ satisfies certain properties, then the sparsest α can be recovered by solving the following optimization problem [8–12],

$$\hat{\alpha} = \arg\min_{\alpha'} \|\alpha'\|_1 \quad \text{subject to } \mathbf{y} = \mathbf{A}\alpha', \tag{4}$$

where $\|\mathbf{x}\|_1 = \sum_i |(x_i)|$. This problem is often known as basis pursuit and can be solved in polynomial time [13]. Note that the ℓ_1 norm is an approximation of the ℓ_0 norm [14]. The approximation is necessary because the optimization problem in Eq. (4) with the ℓ_0 norm (which seeks the sparsest α) is NP-hard and is computationally difficult to solve. In the case in which noisy observations are given, basis pursuit denoising (BPDN) can be used to approximate α,

$$\hat{\alpha} = \arg\min_{\alpha'} \|\alpha'\|_1 \quad \text{subject to } \|\mathbf{y} - \mathbf{A}\alpha'\|_2 \le \varepsilon, \tag{5}$$

where we have assumed that the observations are of the following form:

$$\mathbf{y} = \mathbf{A}\alpha + \eta \tag{6}$$

with $\|\eta\|_2 \le \varepsilon$. One condition that is required for both the ℓ_0 norm based method and the ℓ_1 norm based method to have the same solution and for Eq. (5) to stably approximate the sparsest near solution of Eq. (6) is known as the restricted isometry property (RIP) [10–12]. A matrix, $\mathbf{A}$, is said to satisfy the RIP of order K with constants $\delta_K \in (0, 1]$ if

$$(1 - \delta_K)\|v\|_2^2 \le \|\mathbf{A}v\|_2^2 \le (1 + \delta_K)\|v\|_2^2 \tag{7}$$

for any v such that $\|v\|_0 \le K$. Also, in certain cases, greedy algorithms, such as orthogonal matching pursuit [15], can also be used to recover sparse representations of images.

B. Feature Extraction

For a recognition method to work well, one needs to find good features that can separate the classes in lower dimensional spaces, often known as the feature space. The method presented in the previous section can be easily extended to the case in which different features are used. Let $\mathbf{G} \in \mathbb{R}^{\tilde{M} \times N}$ be a matrix with $\tilde{M} \le N$ that maps the vectors from the original space to the feature space. Here we have assumed that the transformation is approximately a linear operation. Some examples of $\mathbf{G}$ include random projection (RP) matrix, downsampling matrix, PCA dimensionality reduction matrix or some other orthogonal basis such as wavelet transform matrix. Equation (6) can be rewritten as

$$\tilde{\mathbf{y}} \doteq \mathbf{G}\mathbf{y} = \mathbf{G}\mathbf{A}\alpha + \tilde{\eta} \in \mathbb{R}^{\tilde{M}}, \tag{8}$$

where $\tilde{\eta} = \mathbf{G}\eta$. In general, $\tilde{M}$ is chosen to be much smaller than N. This in turn implies that the system of Eqs. (8) will be underdetermined. So long as α is sparse enough and $\mathbf{GA}$ satisfies certain conditions, one can approximate α by solving the following problem:

$$\hat{\alpha} = \arg\min_{\alpha'} \|\alpha'\|_1 \quad \text{subject to } \|\tilde{\mathbf{y}} - \mathbf{G}\mathbf{A}\alpha'\|_2 \le \tilde{\varepsilon}, \tag{9}$$

with $\|\tilde{\eta}\|_2 \le \tilde{\varepsilon}$.

C. Sparse Recognition

Given an observation vector $\mathbf{y}$ from one of the L classes in the training set, we compute its coefficients $\hat{\alpha}$ by solving either Eq. (4) or Eq. (5). We perform classification based on the fact that high values of the coefficients $\hat{\alpha}$ will be mainly associated with the columns of $\mathbf{A}$ from a single class. We do this by comparing how well the different parts of the estimated coefficients, $\hat{\alpha}$, represent $\mathbf{y}$. The minimum of the representation error or the residual error is then used to identify the correct class. The residual error of class k is calculated by keeping the coefficients associated with that class and setting the remaining coefficients not associated with class k to zero. This can be done by introducing a characteristic function, $\chi_k : \mathbb{R}^n \to \mathbb{R}^n$, which selects the coefficients associated with the kth class as follows:

$$r_k(\mathbf{y}) = \|\mathbf{y} - \mathbf{A}\chi_k(\hat{\alpha})\|_2. \tag{10}$$

Here, the vector χ_k has value one at locations corresponding to class k and zero for other entries. The class, d, that is associated with an observed vector is then declared as the one that produces the smallest approximation error:

$$d = \arg\min_k r_k(\mathbf{y}). \tag{11}$$

D. Image Quality Measure

From the previous discussion, one would expect the solution α to be sparse and that it should belong to only one class. For instance, the test image from Target 1 should only belong to the corresponding Target 1 class rather than a combination of different classes. To measure the quality of the coefficient vector, α, the notion of the sparsity concentration index (SCI) [6] has been introduced. The SCI of a coefficient vector, $\alpha \in \mathbb{R}^{(L.n)}$, is defined as

$$\mathrm{SCI}(\alpha) = \frac{\frac{L \cdot \max \|\chi_i(\alpha)\|_1}{\|\alpha\|_1} - 1}{L - 1}. \tag{12}$$

SCI takes values between 0 and 1. SCI values close to 1 correspond to the case in which the test image can be approximately represented by using only images from a single class. In this case the test vector has enough discriminating features of its class and hence has high quality. If SCI = 0, then the coefficients are spread evenly over all classes. So the test

vector is not similar to any of the classes and hence is of poor quality.

3. Block-Sparsity-Based Recognition

The following problem is related to Eq. (5) and is often known as the least absolute shrinkage and selection operator (LASSO) [16]:

$$\min \|\mathbf{y} - \mathbf{A}\alpha'\|_2^2 \quad \text{subject to } \|\alpha'\|_1 \le \tau, \tag{13}$$

where $\tau > 0$. The constrained optimization problems [Eq. (5) and (13)] are closely related to the following unconstrained optimization problem:

$$\min \frac{1}{2}\|\mathbf{y} - \mathbf{A}\alpha'\|_2^2 + \lambda\|\alpha'\|_1, \tag{14}$$

where λ is a nonnegative parameter. The Lagrange multiplier, λ, is related to the LASSO parameter, τ, of the constraint in Eq. (13) and to the reciprocal of the parameter of the constraint in Eq. (5). Hence, for appropriate selections of ε, λ, and τ, the solutions of Eqs. (5), (13), and (14) coincide [17]. That is, these formulations can all be used to identify sparse approximate solutions to the underdetermined system (2).

It has been observed that the LASSO method tends to select a single sample from a group of correlated training samples [18]. However, in our application, we have L target classes. Hence, our dictionary, $\mathbf{A}$, consists of blocks of training samples corresponding to L different targets. Thus, the resulting sparse coefficients, $\hat{\alpha}$, occur in a block. This means that we can use a regularization method that selects an entire block of correlated training samples belonging to the same class. This can be achieved by adapting the following relaxation:

$$\begin{aligned} \hat{\alpha} = \min \|\mathbf{a}_1\|_2 + \|\mathbf{a}_2\|_2 + \ldots + \|\mathbf{a}_L\|_2 \\ \text{subject to } \|\mathbf{y} - \mathbf{A}\alpha'\| \le \varepsilon, \end{aligned} \tag{15}$$

where $\mathbf{a}_i = (\alpha'_{(i-1)n+1}, \alpha'_{(i-1)n+2}, \ldots, \alpha'_{(in)})$ for $i = 1, 2, \ldots, L$. Note that this method requires the labels of each group in $\mathbf{A}$. This presents no obstacles, because the label of each training sample is known *a priori*. It is shown in [19] that the solution of Eq. (15) satisfies

$$\|\alpha - \hat{\alpha}\|_2 \le C_1 K^{-\frac{1}{2}} \|\alpha - \alpha^K\|_{2,\mathcal{J}} + C_2\varepsilon,$$

provided that $\|\mathbf{y} - \mathbf{A}\alpha\|_2 \le \varepsilon$ and $\delta_{2K|\mathcal{J}} < \sqrt{2} - 1$, where C_1 and C_2 are some constants, α^K denotes the best block K sparse approximation to α, and $\|\alpha\|_{2,\mathcal{J}} = \sum_{i=1}^{L} \|\mathbf{a}_i\|_2$, and $\mathcal{J} = \{\mathcal{J}_i\}_{i=1}^{L}$ is the partition of the set $\{1, 2, \ldots, L\}$, that is, $\bigcup_{i=1}^{L} \mathcal{J}_i = \{1, 2, \ldots, L\}$ and $\sum_{i=1}^{L} |\mathcal{J}_i| = n.L$. Here the block-restricted isometry constant $\delta_{K|\mathcal{J}}$ is defined as the smallest $\delta_{K|\mathcal{J}}$ such that $\mathbf{A}$ satisfies

$$(1 - \delta_{K|\mathcal{J}})\|v\|_2^2 \le \|\mathbf{A}v\|_2^2 \le (1 + \delta_{K|\mathcal{J}})\|v\|_2^2 \tag{16}$$

for any v that is block K sparse over $\mathcal{J}$. It follows that $\delta_{K|\mathcal{J}} \le \delta_K$, where δ_K is the conventional restricted isometry constant [e.g., Eq. (7)], corresponding to nonBS vectors [19]. This, in turn, implies the existence of improved performance guarantees compared to BPDN [19,20]. For this reason, in our ATR algorithm, we use Eq. (15) to harness the underlying BS structure of α. In recent years, an enormous amount of research has been done regarding related regularization methods. It is also possible to use grouped LASSO [21,22] and elastic net [18] for regularization in our recognition method. Note that the reconstruction of BS signals from the compressive measurements has been studied extensively in [20,23]. Our ATR algorithm based on SR is summarized in Fig. 2.

4. Experimental Results

In this section, we present some simulation results of different ATR methods promoting sparsity on the Comanche (Boeing–Sikorsky, USA) FLIR data set consisting of different military targets at different orientations. The images are of size 40×75 pixels. In all of our experiments, the dimension of each target image (chip) was reduced from 40×75 to 16×16 unless otherwise stated. A number of approaches have been suggested for solving BS-promoting optimization problems (15). In our approach, we employed a highly efficient algorithm that is suitable for large scale applications, known as the spectral projected gradient (SPGL1) algorithm [17]. The threshold value for SCI was set equal to 0.15. The performance of our algorithm is compared with that of several different methods reported in [2,3,5]. Our algorithm is also tested using several features, namely PCA features, RP features, two-dimensional Haar wavelet features, and downsampled images.

A. Data Set

In our data set, there are 10 different vehicle targets. We will denote these targets as TG1, TG2, ..., TG10. For each target, there are 72 orientations, corresponding to the aspect angles of $0°, 5°, \ldots, 355°$ in azimuth. The range to all the targets is given so that all the target chips are analyzed at 2 km. The data consist of a training set and a test set. We will refer to the training set as the SIG set and the test set as the ROI set. The SIG data set has about 13,816 target chips, while there are 3353 images in the ROI data

Given a matrix of training samples $\mathbf{A} \in \mathbb{R}^{N \times (n.L)}$ for L classes and a test sample $\mathbf{y} \in \mathbb{R}^N$:

1. Solve the optimization problem (15) to obtain α.
2. Compute the reconstruction error using (10).
3. Identify $\mathbf{y}$ using (11).

Fig. 2. SR-based ATR algorithm.

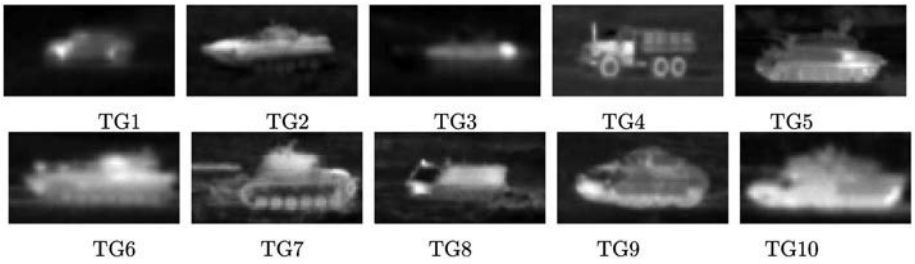


Fig. 3. Side view of all 10 targets present in the SIG data set.

set. The SIG data set consists of images that were collected under very favorable conditions. The SIG data set contains 874 to 1468 images per target class spanned over 72 different aspects. In Fig. 3, we display the side view of all the 10 targets present in the SIG set.

The ROI set consists of only five targets, namely, TG1, TG2, TG3, TG4, and TG7. The target images for the ROI set were taken under less favorable conditions, such as targets with different weather conditions, in different backgrounds, in and around clutter; hence these data are very challenging. There are 577 to 798 images for each of these five target classes. Some of the target chips from the ROI data set are shown in Fig. 4. All the images in the SIG and ROI sets were normalized to a fixed range with the target put approximately in the center. The orientation in the ROI set was given very coarsely; every 45°.

B. Results on SIG Data Set

In the first set of experiments, the training and test images were chosen from the SIG data set. For training, we randomly chose 11 target chips for each target per aspect angle, called TRAIN-SIG. Because we have a total of 72 aspects (i.e., $0°, 5°, \ldots, 355°$) for each target, we used a total of $11 \times 72 = 792$ targets per class. Hence, the resulting dictionary, $\mathbf{A}$, is of size 256×7920. Another set of 1000 targets, disjoint from the training data, called TEST-SIG, was used for testing. We solve the following problem favoring BS based on target class per aspect:

$$\hat{\alpha} = \min \|\mathbf{a}_1\|_2 + \|\mathbf{a}_2\|_2 + \ldots + \|\mathbf{a}_{720}\|_2 \quad \text{subject to } \|\mathbf{y} - \mathbf{A}\alpha'\| \le \varepsilon, \tag{17}$$

where $\mathbf{a}_i = (\alpha'_{(i-1)n+1}, \alpha'_{(i-1)n+2}, \ldots, \alpha'_{(in)})$ for $i = 1, 2, \ldots, 11$. Once the BS vector, α, is found, we compute the reconstruction error using Eq. (10) and identify the novel target chip, $\mathbf{y}$, using Eq. (11). We applied this BS-based algorithm to various features on the TRAIN-SIG data set. Examples of different features extracted for this experiment are shown in Fig. 5.

The probabilities of correct classification for these experiments are 98.48%, 99.18%, 99.96%, and 99.95% for the downsampled, RP, PCA, and Haar wavelet features, respectively. All the features performed approximately the same for these experiments. The confusion matrices [24] corresponding to these experiments are shown in Figs. 6(a)–6(d).

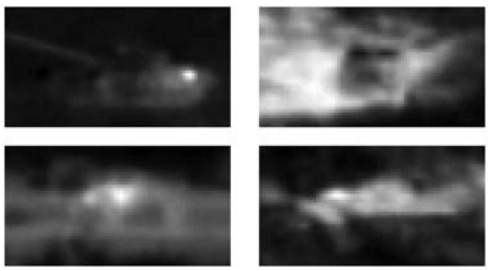


Fig. 4. Some sample target chips from the ROI data set.

C. Results on ROI Data Set

In the second set of experiments, we randomly selected 11 targets per aspect angle from the SIG data set for training. The resulting dictionary, $\mathbf{A}$, is of size 256×7290. We randomly selected 1000 images from the ROI set for testing, called the TEST-ROI set. Again, we extracted various features and applied our BS-based algorithm to these features as was done for the TRAIN-SIG data set. The probabilities of correct classification for these experiments are 75.10, 76.30, 78.89, and 76.45% for the downsampled, RP, PCA, and Haar wavelet features, respectively. The PCA features gave the best performance. The confusion matrices corresponding to these experiments are shown in Figs. 7(a) and 7(d). In these experiments, the TEST-ROI set contained only five targets, but all of the outputs were active. Note that we have included five rows with zeros for clarity due to the fact that five other targets are not present in this data set.

The best recognition results on the TEST-SIG and TEST-ROI data sets were obtained by using the PCA features. Performance of our algorithm using various features on TEST-SIG and TEST-ROI is compared in Fig. 8. Also, we report the performance of different techniques [2,3,5] on these data sets in Table 1. As can be seen from the table, our method achieves recognition rates of 99.96% and 78.89% on TEST-SIG

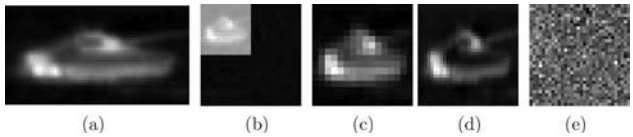


Fig. 5. Examples of different features used in this paper: (a) original target chip, (b) Haar wavelet features, (c) downsampled image, (d) PCA features, (e) random projection.

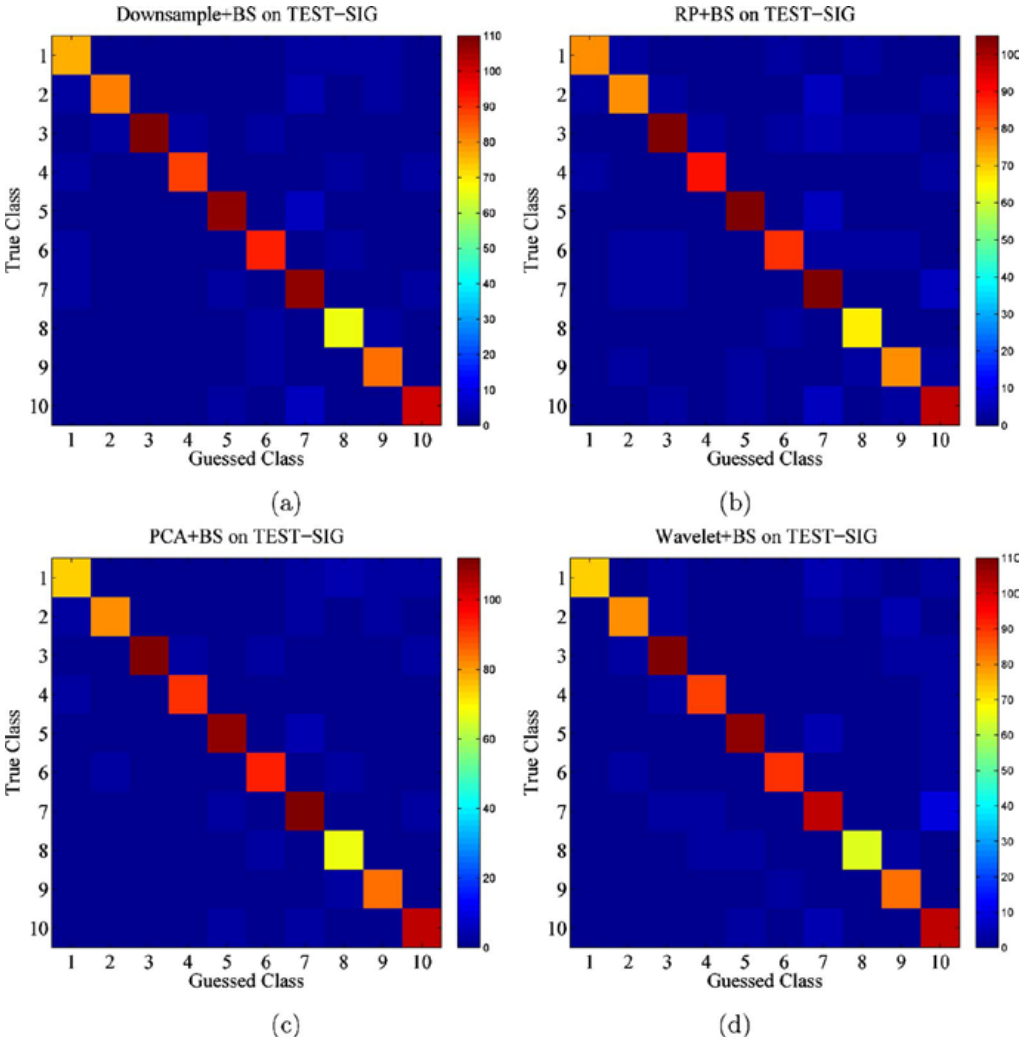


Fig. 6. (Color online) Confusion matrices corresponding to the SIG data set using different features: (a) downsampled, (b) random projection, (c) PCA, (d) Haar wavelet.

and TEST-ROI, respectively, and it outperforms other methods such as CNN, MNN, PCA, LVQ, LDA, H-M and G-H [2,3,5]. Also, note that our method is more general than the competing methods presented in [2,3]. In their methods, to deal with the background artifacts, they use several rectangular windows of different sizes based on the ground truth silhouette computer-aided design models. As a result, their performance significantly depends on the choice of windows. In contrast, the method presented here does not require any windowing or prior knowledge about the size of the targets.

D. Target Pose Estimation

Because the dictionary, **A**, contains target chips at different known orientations, we can reliably estimate the pose of a target from the test set. To illustrate this, consider a test target chip shown in Fig. 9. This chip belongs to TG1, and its orientation is 5°. The values of the sparse coefficients obtained by solving a BS-promoting problem are shown at the bottom of Fig. 9. Using the correct labels of columns in **A** and the sparse coefficients, one can identify the corresponding aspect angle of this target.

Note that such pose information can be utilized to exclude the background in each target chip as was done in [5]. Furthermore, if we know the possible orientation of a target, we can validate the target type by using some features unique to the target in that orientation (see [3,5] for details).

E. Recognition Rate versus Feature Dimension

In this section, we show how the performance of our algorithm changes as we change the feature dimension. For this experiment we again randomly selected 11 targets per aspect angle from the SIG data set for training. Another set of 1000 targets, disjoint from the training data that were used for testing. PCA features, was used for this experiment. Figure 10 shows the recognition rates for this experiment corresponding to various feature dimensions. As can be seen from this figure, the recognition rate increases as we increase the feature dimension. Above the feature dimension of 256, the recognition rate stays

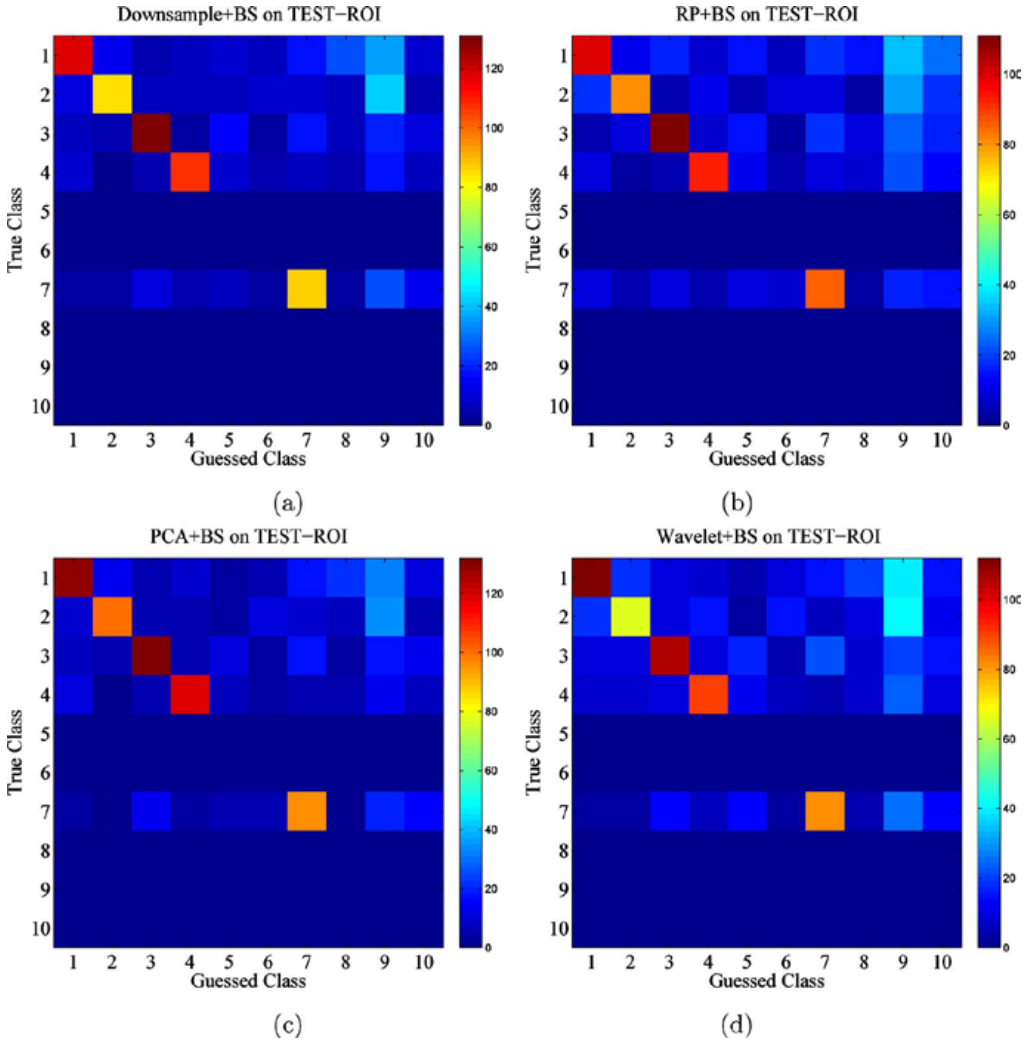


Fig. 7. (Color online) Confusion matrices corresponding to the ROI data set using different features: (a) downsampled, (b) random projection, (c) PCA, (d) Haar wavelet.

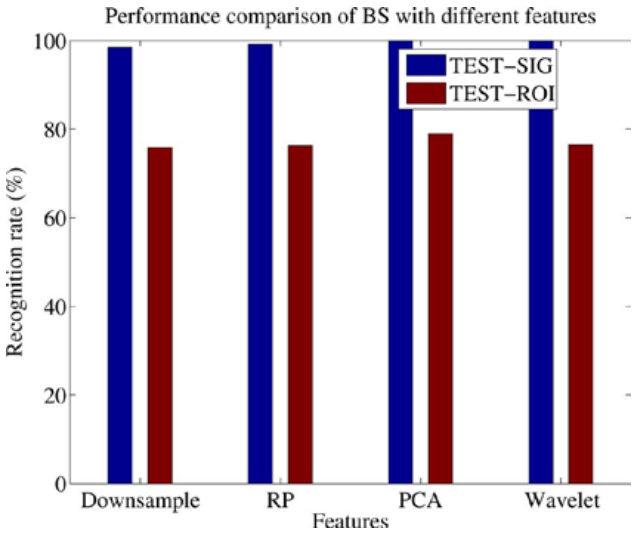


Fig. 8. (Color online) Recognition results on the TEST-SIG and TEST-ROI sets using different features.

approximately the same. This is no surprise because it has been observed by many researchers that, in practice, features (e.g., number of compressive measurements) of the order of 3 to 5 times the number of sparse coefficients suffice for a good recovery [25,26]. From our assumption that any test image that belongs to class k will approximately lie in the linear span of the training samples from the same class, and because each class per aspect angle contains 11 training images, our method requires the feature dimension to be more than $5 \times 11 = 55$ for good recovery of sparse coefficients. Hence, increasing the feature dimension more than is required by the ℓ_1 minimization will not necessarily improve the quality of recovered sparse coefficients. This in turn implies that after a certain feature dimension, the recognition rate will approximately stay the same [27]. Also, from the previous experiments, we see that the BS-based recognition algorithm gives approximately the same performance when different features are used, provided that the dimension of

Table 1. Recognition Rates (in %) for Different Methods

Methods	BS	CNN4	LVQ	MNN	PCA	LDA	H-M	G-H
TEST-SIG	99.96	95.16	99.72	95.49	95.44	86.92	93.73	80.24
TEST-ROI	78.89	59.25	75.12	75.58	52.17	50.32	62.86	50.09

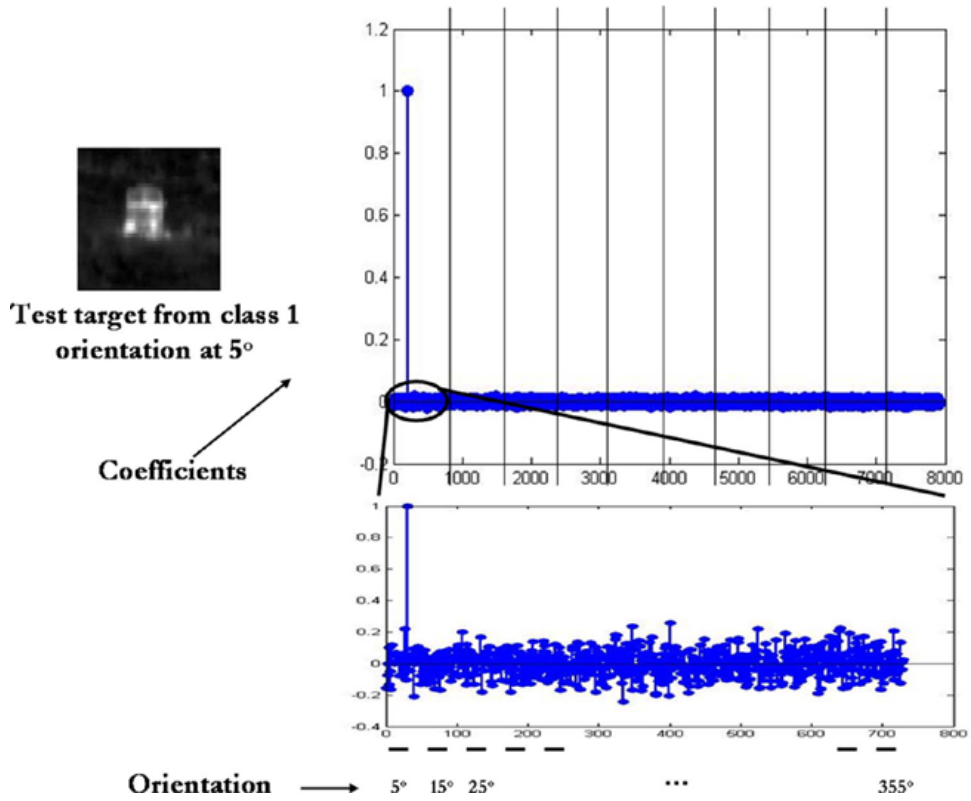


Fig. 9. (Color online) Target orientation detection. Dictionary matrix **A** contains training images with known orientation. This can be used to identify the aspect angle of a test target.

the features is kept high enough. This shows that the choice of features is not critical but the dimension of features is.

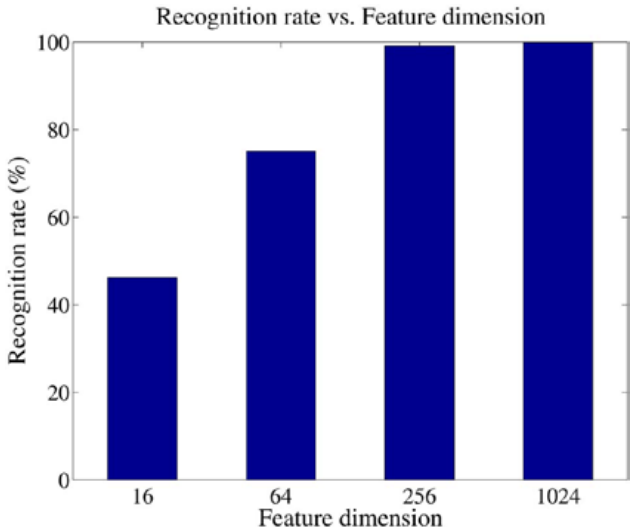


Fig. 10. (Color online) Recognition rate versus feature dimension.

5. Discussion and Conclusion

We have developed a framework for ATR using the theory of SR and CS. This entails solving a BS-promoting optimization problem on various features. Various experiments on the Comanche (Boeing–Sikorsky, USA) FLIR data set have shown promising results.

Several future directions of inquiry are possible considering our new approach to ATR. For instance, instead of using the ℓ_1 minimization, one can consider greedy pursuits such as orthogonal matching pursuit and compressive sampling matching pursuit [15,28,29]. Greedy pursuits are known to converge much faster than optimization-based methods and have the same theoretical guarantees as some of the optimization-based methods. Even though, in this paper, we took a reconstructive approach to dictionary learning for ATR, it is possible to learn discriminative dictionaries for the task of target recognition [30,31]. Note that the sparsity-motivated methods for ATR presented here for FLIR images can be easily extended to the other ATR problems based on ladar, underwater optical imagery [32], or synthetic aperture radar imagery.

References and Notes

1. F. Sadjadi, "Infrared target detection with probability density functions of wavelet transform subbands," Appl. Opt. **43**, 315–323 (2004).
2. L. Wang, S. Z. Der, and N. M. Nasrabadi, "Automatic target recognition using a feature-decomposition and data-decomposition modular neural network," IEEE Trans. Image Process. **7**, 1113–1121 (1998).
3. L. A. Chan and N. M. Nasrabadi, "Application of wavelet-based vector quantization in target recognition," Int. J. Art. Intel. Tools **6**, 165–178 (1997).
4. P. Bharadwaj and L. Carin, "Infrared image classification using hidden Markov trees," IEEE Trans. Pattern Anal. Mach. Intel. **24**, 1394–1397 (2002).
5. B. Li, R. Chellappa, Q. Zheng, S. Der, N. M. Nasrabadi, L. Chan, and L. Wang, "Experimental evaluation of forward-looking IR data set automatic target recognition approaches —a comparative study," Comput. Vision Image Underst. **84**, 5–24 (2001).
6. J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," IEEE Trans. Pattern Anal. Mach. Intel. **31**, 210–227 (2009).
7. V. M. Patel, N. M. Nasrabadi, and R. Chellappa, "Sparsity-Inspired automatic target recognition," Proc. SPIE **7696**, 76960Q (2010).
8. D. L. Donoho and X. Huo, "Uncertainty principle and ideal atomic decomposition," IEEE Trans. Inf. Theory **47**, 2845–2862 (2001).
9. D. L. Donoho and M. Elad, "On the stability of the basis pursuit in the presence of noise," EURASIP J. Adv. Signal Processing **86**, 511–532 (2006).
10. E. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," Comm. Pure Appl. Math. **59**, 1207–1223 (2006).
11. H. Rauhut, K. Schnass, and P. Vandergheynst, "Compressed sensing and redundant dictionaries," IEEE Trans. Inf. Theory **54**, 2210–2219 (2008).
12. R. Baraniuk, "Compressive sensing," IEEE Signal Process. Mag. **24**, 118–121 (2007).
13. S. Chen, D. Donoho, and M. Saunders, "Atomic decomposition by basis pursuit," SIAM J. Sci. Comput. **20**, 33–61 (1998).
14. $\|\mathbf{x}\|_0$ returns the number of nonzero elements of $\mathbf{x}$.
15. J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," IEEE Trans. Inf. Theory **53**, 4655–4666 (2007).
16. R. Tibshirani, "Regression shrinkage and selection via the lasso," J. R. Stat. Soc. Ser. B. Methodol. **58**, 267–288 (1996).
17. E. van der Berg and M. P. Friedlander, "Probing the Pareto frontier for basis pursuit solutions," SIAM J. Sci. Comput. **31**, 890–912 (2008).
18. H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," J. R. Stat. Soc. B **67**, 301–320 (2005).
19. Y. C. Eldar and M. Mishali, "Robust recovery of signals from a structured union of subspaces," IEEE Trans. Inf. Theory **55**, 5302–5316 (2009).
20. Y. C. Eldar, P. Kuppinger, and H. B. Bolcskei, "Block-sparse signals: uncertainty relations and efficient recovery," IEEE Trans. Signal Process. **58**, 3042–3054 (2010).
21. M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," J. R. Stat. Soc. B **68**, 49–67 (2006).
22. L. Meier, S. van der Geer, and P. Buhlman, "The group lasso for logistic regression," J. R. Stat. Soc. B **70**, 53–71 (2008).
23. M. Stojnic, F. Parvaresh, and B. Hassibi, "On the reconstruction of block-sparse signals with an optimal number of measurements," IEEE Trans. Signal Process. **57**, 3075–3085 (2009).
24. Given C classes, the confusion matrix is of size $C \times C$, whose diagonal entries are the number of targets that are correctly classified, while the off-diagonal entries are the number of misclassifications.
25. Y. Tsaig and D. L. Donoho, "Extensions of compressed sensing," Signal Process. **86**, 549–571 (2006).
26. E. Candès and J. K. Romberg, "Signal recovery from random projections," Proc. SPIE **5674**, 76–86 (2005).
27. Exact recovery of sparse coefficients also depends on the matrix $\mathbf{GA}$. That is, $\mathbf{GA}$ must satisfy certain properties such as the mutual coherence property and the restricted isometry property. See [11,10,20] for more details.
28. P. J. Phillips, "Matching pursuit filters applied to face identification," IEEE Trans. Image Process. **7**, 1150–1164 (1998).
29. D. Needell and J. A. Tropp, "CoSaMP: iterative signal recovery from incomplete and inaccurate samples," Appl. Comput. Harmon. Anal. **26**, 301–321 (2009).
30. F. Rodriguez and G. Sapiro, "Sparse representations for image classification: learning discriminative and reconstructive nonparametric dictionaries," Tech. Report (University of Minnesota, Dec. 2007).
31. P. Sprechmann and G. Sapiro, "Dictionary learning and sparse coding for unsupervised clustering," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing* (IEEE, 2010), pp. 2042–2045.
32. F. M. Caimi, D. M. Kocak, F. Dalgleish, and J. Watson, "Underwater imaging and optics: recent advances," Oceans 2008 **2008**, 1–9 (2008).

Fusing Concurrent Visible and Infrared Videos for Improved Tracking Performance

Alex Lipchen Chan and Stephen R. Schnelle

Optical Engineering 2013, 52 (1), 12. [DOI: 10.1117/1.OE.52.1.017004]

Fusing concurrent visible and infrared videos for improved tracking performance

Alex Lipchen Chan
U.S. Army Research Laboratory
Attention: RDRL-SES-E
2800 Powder Mill Road
Adelphi, Maryland 20783-1197
E-mail: Alex.L.Chan.Civ@mail.mil

Stephen R. Schnelle
Rice University
Electrical and Computer Engineering
MS-366, P.O. Box 1892
Houston, Texas 77251-1892

Abstract. Visible and infrared video cameras are the most common imaging sensors used for video surveillance systems. Fusing concurrent visible and infrared imageries may further improve the overall detection and tracking performance of a video surveillance system. We performed image fusion using 13 pixel-based image fusion algorithms and examined their effects on the detection and tracking performance of a given target tracker. We identified five fusion methods that produced significantly better performance, three of which also managed to achieve that with a relatively high efficiency. [DOI: 10.1117/1.OE.52.1.017004]

Subject terms: infrared imagery; image fusion; target detection; target tracking.

Paper 121099P received Jul. 30, 2012; revised manuscript received Dec. 4, 2012; accepted for publication Dec. 12, 2012; published online Jan. 7, 2013.

1 Introduction

As sensor technology, network communication, computing power, and digital storage capacity have all dramatically improved, still and video imageries have become the most common and versatile forms of media for capturing, analyzing, and disseminating a variety of information. Visible cameras are the prevailing imaging sensors because they are relatively cheap, easy to use, and capable of producing high-quality imagery under favorable conditions. However, visible cameras can be severely affected by common environmental factors such as darkness, shadows, fog, clouds, rain, snow, and smoke. Infrared (IR) imaging systems may overcome or alleviate some of these problems, but they are subject to a number of limitations of their own. IR-specific difficulties include a much lower sensor resolution; total loss of nonthermal but important visual features (such as color and text); blockage by visually transparent thermal signal shields (such as car windshields and glass doors); and very low thermal contrast between targets and background under certain combinations of ambient and target temperatures. Due to these highly complementary strengths and limitations of visible and IR cameras, more advanced target detection and tracking systems may want to acquire and process both visible and IR imageries concurrently and jointly for critical surveillance and force protection applications.

To study the usefulness of fusing visible and IR imageries for detecting and tracking moving targets, we have relied on a large collection of concurrent color visible and long-wave IR (LWIR) video sequences that are officially referred to as the Second Dataset of the Force Protection Surveillance System (FPSS).[1] These FPSS video sequences were collected using the Sentry Personnel Observation Device (SPOD) that includes a LWIR microbolometer and a color visible camera. The LWIR images were acquired with a focal plane array (FPA) of 320×240 pixels in resolution, while the color visible images were captured at the resolution of 460 TV lines. Both the original color visible and LWIR images were cropped and scaled to a common image size of 640×480 pixels, in order to attain a coarse level of coregistration between the corresponding color-LWIR images captured at any given time.

Image fusion can be performed at several different levels.[2] At the lowest levels, the raw image data can be fused. This can either be performed on the original signal or, more likely, after the image has been preprocessed and the resulting pixel values are used. Pixel-level fusion is very common due to its simplicity and universality, and it is the focus of this work as well. At higher levels, feature-based detection uses structural image characteristics, such as edges and corners, to enhance the image. For example, one could extract the edge information from a pair of images using Sobel filter and fuse the images based on the edge information. However, this approach is much more application-specific, often requiring an understanding of the image, itself, either through direct human intervention or automatic object classification algorithms. Therefore, this approach requires much more complex computation, complicated training methods, and nonreal-time intervention. One example of a higher level fusion system uses Bayesian analysis to sum the probabilities of detected human silhouettes falling within each pair of visible and infrared images. Oftentimes, detections are based on whether the probability exceeds a predefined threshold.[3] For a stationary camera installed in a specific setting, training such a system may be feasible because its background does not vary significantly. At the highest level of image fusion, symbolic fusion methods are often heavily rule-based and rely on a lot of prior or external knowledge to perform the image fusion. Nonetheless, symbolic image fusion methods can carry similar tradeoffs as the fusion methods at the feature-level.

There are many ways to measure performance of image fusion algorithms, including subjective analysis, complex similarity metrics, signal-to-noise ratio (SNR), and tracking performance. Motwani et al. suggested parameters for subjective analysis, but they concluded that subjective measures were not particularly helpful for tracking systems, except in the case of incorporating human feedback into the detection loop.[4] Cvejic et al. discussed a number of objective similarity

0091-3286/2013/$25.00

metrics, including the Piella metric, Petrovic metric, and Bristol metric.[5] The Piella metric measures structured similarity (which is based on luminance, contrast, and structure information) over local window regions and then averages these similarity measures over all windows. Weighting is given to the relative importance of each input image toward the fused image, window by window. The Petrovic metric specifically evaluates edge structure (using a Sobel edge operator) by determining the strength of edge information retained from each of the original images in the fused image. The Bristol metric, in contrast to the Piella metric, uses a slightly different weighting scheme based on the ratio of covariances between the original and fused images. Cvejic et al. compared the tracking performance of a particle filter based on these objective metrics and found that the tracking performance was actually worsened by the fusion of images. Mihaylova et al., of the same research group, later adopted a performance metric of normalized overlapping ground truth and tracking system bounding boxes in their work.[6] Their results showed that IR images alone performed just as well or better than most fusion algorithms (including contrast pyramid, dual-tree complex wavelet transform, and discrete wavelet transform) in tracking, while visible spectrum images lagged behind under harsher conditions like occlusions.

There are many possible methods of tracking a moving target, including background subtraction, optical flow, moving energy, and temporal differencing. Because the FPSS dataset was collected with a stationary SPOD with minimal background interference, we decided to use an existing FPSS tracker,[7] which is based on background subtraction method, to examine the tracking performance of various image fusion methods. Instead of the FPSS tracker, one of many other moving target tracking algorithms may be used for a similar study, as well. For instance, Trucco and Plakas described a wide range of alternative tracking algorithms in their paper.[8]

In the next section, we provide brief discussions on the 13 image fusion methods of interest. These fusion methods fall into two broad categories—simple combination and pyramid structure. A brief description of the FPSS tracker is provided in Sec. 3, while the experimental results on the tracking performance of various image fusion methods are presented in Sec. 4. Finally, some concluding thoughts are given in Sec. 5.

2 Fusion Methods

In this paper, we focus on 13 pixel-level image fusion methods, ranging from the simplest pixels averaging method to the very complicated dual-tree complex wavelet transform method. There are other interesting but less popular image fusion algorithms, including one that relies on factorizing an image $\mathbf{V}$ into two nonnegative matrix components, $\mathbf{W}$ and $\mathbf{H}$, with $\mathbf{W}$ representing a basis optimized for representing $\mathbf{V}$.[9] Another approach to image fusion is to use training sets and supervised classifiers, as explored by Chan.[10] In this work, however, we assume no prior training data are available.

To evaluate the image fusion algorithms examined here, we used all FPSS coarsely registered color visible and LWIR images as input data, a pair of which is shown in Fig. 1. To allow fusion with LWIR images, the color visible (RGB) images were converted to grayscale images using a simple weighting of $0.2989R + 0.5870G + 0.1140B$, which yielded the intensity value but removed the hue and saturation information.[11] For many automatic target detection and tracking algorithms, it is, indeed, more efficient to process grayscale images internally, while providing color outputs for human consumption only.

2.1 *Simple Combinations*

The most intuitive pixel-level fusion methods examined here are simple averaging, intelligent weighting, and selecting maximum or minimum pixel values between the visible and LWIR images. All these methods involve only simple pixel operations, which require traversing the two input images to be fused pixel-by-pixel, leading to a simple $O(m \times n)$ operation for an image of size $m \times n$. Pixels $(\mathbf{I}_1)_{ij}$ and $(\mathbf{I}_2)_{ij}$ in images $\mathbf{I}_1$ and $\mathbf{I}_2$ need only be compared against each other once.

In the first fusion method, a fused image $\mathbf{I}_f$ was generated through simple averaging by calculating $(\mathbf{I}_f)_{ij} = [(\mathbf{I}_1)_{ij} + (\mathbf{I}_2)_{ij}]/2$, and the resulting $\mathbf{I}_f$ is shown in Fig. 2(a).

Fig. 1 Example of a color visible (a) and an LWIR (b) image in the FPSS dataset.

(a) (b)

Fig. 2 Fused image through simple average (a) and PCA-weighted average (b).

Because the visible and LWIR images have differing resolutions and salient features, this method tends to muddle the details.

We boosted the influence of the better image using the principal component analysis (PCA) derived from the covariance matrix between the two input images. A simple way to do this is to consider each image as a single vector $\mathbf{I}_1$ and $\mathbf{I}_2$, creating a 2×2 covariance matrix when we compute the covariance of $[\mathbf{I}_1 \quad \mathbf{I}_2]$. A resulting eigenvector provides the weights for fusing the pixels: $(\mathbf{I}_f)_{ij} = (\mathbf{v}_k)_1(\mathbf{I}_1)_{ij} + (\mathbf{v}_k)_2(\mathbf{I}_2)_{ij}$, where $\mathbf{v}_k$ represents the normalized eigenvector associated with λ_k, the larger one of the two eigenvalues. Generally, the PCA-weighted averaging method strongly favors the image with the highest variance, which may or may not contain more informative and useful details. In fact, this selection criterion can be a disadvantageous one when dealing with noisy images. As shown in Fig. 2(b), the fused image produced by this method closely matches the original visible spectrum image because the visible image has more details and a higher variance.

Choosing the maximum pixel value, $(\mathbf{I}_f)_{ij} = \max[(\mathbf{I}_1)_{ij}, (\mathbf{I}_2)_{ij}]$, from a pair of LWIR and visible images, as shown in Fig. 3(a), may be appropriate to find some hidden targets. A man may be occluded in the visible spectrum, for example, but he can still be located in the LWIR image. For a background subtraction method, it may be desirable to boost the relative intensity of targets through this fusion method, if these targets tend to be brighter than their immediate background.

Choosing the minimum pixel value, $(\mathbf{I}_f)_{ij} = \min[(\mathbf{I}_1)_{ij}, (\mathbf{I}_2)_{ij}]$, may not be very useful in general because it tends to deemphasize the strong foreground objects, as evident from Fig. 3(b). In some rare occasions, this method may be helpful in extracting weak targets (with both weak but detectable visible and LWIR signatures) from busy backgrounds by deemphasizing stronger and brighter neighboring background pixels.

2.2 *Pyramid Structures*

Pyramid decompositions were introduced by Burt and Adelson in 1983 as a compact encoding scheme.[12] The original idea is that a Gaussian kernel (low-pass filter) is applied to the top-level image of a pyramid, $\mathbf{I}_1 * \mathbf{G}_1$, representing the convolution of the image $\mathbf{I}_1$ with a Gaussian blurring matrix

(a) (b)

Fig. 3 Fusion by selecting maximum (a) and minimum (b) pixel intensities.

$\mathbf{G}_1$. This image is then down-sampled to form the next level of these pyramids. The difference between the low-pass version and its previous-level image represents the high frequency or detail information of the previous-level image. At each step down the pyramid, we continue to filter and down-sample in the same manner. A Laplacian pyramid is formed by computing the difference between each level of the pyramid, iteratively separating an image into low and high frequency components, except that the lowest level contains the remaining low-frequency information.

Since each level is a down-sampled version of the previous level, we need to up-sample and interpolate the decimated version in order to compute the difference between the two adjacent levels. For example, the Laplacian image at level k of $\mathbf{I}_m$, denoted as $(\mathbf{L}_m)_k$, can be computed as $(\mathbf{L}_m)_k = (\mathbf{I}_m)_k - f_{k+1}[(\mathbf{I}_m)_{k+1}]$, where $f_{k+1}[\,]$ denotes the function consisting of up-sampling and an interpolation filter with similar blurring response as $\mathbf{G}_k$, while k denotes the level of decomposition. As we proceed down the pyramid, $(\mathbf{I}_m)_k$ denotes the blurred and decimated version of $(\mathbf{I}_m)_{k-1}$. By decomposing each set of the original LWIR and visible images, we form compact representations separated into detail and approximation information. Hence, we can then weight the coefficients in each pyramid. To reconstruct the fused image, we then reverse the decomposition process, starting with a synthesis image at level $k+1$, denoted as $(\mathbf{S}_m)_{k+1}$, expanding it, and adding it to $(\mathbf{L}_m)_k$ to get $(\mathbf{S}_m)_k$. The initial synthesis image is the background coefficients found at the bottom of the Laplacian pyramid. If we select the maximum coefficients between the two pyramids by taking max $\{[(\mathbf{L}_1)_k]_{ij}, [(\mathbf{L}_2)_k]_{ij}\}$ for each level k, and all ij coefficients during this reconstruction process, then a Laplacian fused image is generated [see Fig. 4(a)].

A filter-subtract-decimate (FSD) pyramid is similar to the Laplacian pyramid, but the levels are subtracted prior to decimations. This makes the method simpler and reduces delay, therefore, allowing easier real-time implementation. Slight frequency distortions are introduced, thus a correction factor is required for perfect reconstruction. This term can be dropped in practice, though variations can make minor adjustments in the synthesis phase to account for this. Figure 4(b) shows the result of image fusion based on the original FSD technique proposed by Anderson.[13] Both images in Fig. 4 may look similar, except for a slight shading difference, but their differences in tracking performance could be larger than that.

The ratio-of-low-pass (ROLP) pyramid and the contrast pyramid use the ratio of levels of the Gaussian pyramid to compute the coefficients at the next level, instead of their differences.[14,15] Otherwise, the decomposition process resembles that of the Laplacian pyramid. Since the stored coefficients are not used to compute levels of the Gaussian pyramids, the underlying Gaussian pyramid decomposition of the image does not change. The primary difference between ROLP and contrast pyramids is the use of a local background to normalize the ratio. The contrast pyramid computes $(\mathbf{L}_m)_k = \{(\mathbf{I}_m)_k / f_{k+1}[(\mathbf{I}_m)_{k+1}]\} - 1$, and the offset of 1 is reversed during reconstruction. On the other hand, the ROLP pyramid computes $(\mathbf{L}_m)_k = (\mathbf{I}_m)_k / f_{k+1}[(\mathbf{I}_m)_{k+1}]$. Instead of summing coefficients during synthesis (as in the case of Laplacian pyramid), we now reverse-decomposition by expanding $(\mathbf{S}_m)_{k+1}$ and multiplying it with $(\mathbf{L}_m)_k$ to get $(\mathbf{S}_m)_k$. A small epsilon factor is often added to the denominator to prevent division-by-zero issues. Figure 5 shows the resulting fused images from the ROLP and contrast pyramid methods. These decomposition methods are designed to emphasize the contrast in an image.

The gradient pyramid chooses the largest directional derivative in each of four directions: horizontal, vertical, and the two diagonal directions.[16] These derivatives can be computed using simple matrix operators. For example, at each level of the pyramid, the four operators $[\,1 \quad -2 \quad 1\,]$,

$$\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix},$$

$$\begin{bmatrix} 0 & 0 & 1/2 \\ 0 & -1 & 0 \\ 1/2 & 0 & 0 \end{bmatrix},$$

and

(a) (b)

Fig. 4 Fusion by selecting the maximum coefficient of Laplacian pyramids (a) and FSD pyramids (b).

Fig. 5 Fusion by selecting the maximum coefficient of the ROLP pyramids (a) and contrast pyramids (b).

$$\begin{bmatrix} 1/2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1/2 \end{bmatrix}$$

can be convolved with a filtered image

$$(\mathbf{I}_m)_k + (\mathbf{I}_m)_k * \begin{bmatrix} 1/16 & 1/8 & 1/16 \\ 1/8 & 1/4 & 1/8 \\ 1/16 & 1/8 & 1/16 \end{bmatrix},$$

where * represents the convolution operator. Coefficients are selected for each of the four directions independently during the fusion process and then added together to represent the combined gradient strength at a given pixel location. A synthesized image is reconstructed using the same procedure as in the Laplacian pyramid case. An example of the fused image produced by the gradient pyramid method is shown in Fig. 6(a). These methods are designed to preserve orientation information, which can be useful in some applications.

Morphological operations, such as opening and closing, can be applied to the Gaussian pyramid without harmful effects under certain circumstances and result in a morphological pyramid.[17] For example, we can apply the following operations to compute the next set of coefficients from $(\mathbf{I}_m)_k$: morphologically open $(\mathbf{I}_m)_k$ by first replacing the value of a given pixel with the smallest pixel value found within a predefined neighborhood of that pixel (erosion), and then on the resulting image, replacing the value of a given pixel with the largest pixel value found in the same neighborhood (dilation). The resulting image can then be closed by reversing the process—namely, first performing a dilation and then an erosion operation. The opening operation will remove small objects, while the closing operation will remove noise and smooth transitions. We decimate the resulting image to obtain our image for the next level of the pyramid, $(\mathbf{I}_m)_{k+1}$. We obtain the pyramid coefficients of level $k+1$ as the difference between $(\mathbf{I}_m)_k$ and an up-sampled and dilated version of $(\mathbf{I}_m)_{k+1}$ While these morphological operations may produce good-looking results, as shown in Fig. 6(b), they are quite computationally intensive in nature, and their usefulness in enhancing tracking performance is not necessarily great.

Finally, many specialized pyramid decompositions, such as contourlets and wavelets, separate an image into approximations and detail. We examined a simple discrete wavelet transform (DWT) using the Daubechies Symmetric Spline

Fig. 6 Fusion by selecting the maximum coefficient of the gradient pyramids (a) and morphological pyramids (b).

wavelet, as well as a shift-invariant discrete wavelet transform (SIDWT), using the Harr wavelet. The DWT is applied to an input image using two filters, $g_1 = [-2 \quad 4 \quad -2]$ and $h_1 = [-1 \quad 2 \quad 6 \quad 2 \quad -1]$. In this case, g_1 is a high-pass filter and h_1 a low-pass filter. These filters are applied to the columns and rows of an image consecutively in one of these four combinations: $g_1' * g_1$, $g_1' * h_1$, $h_1' * g_1$, and $h_1' * h_1$. The output of $g_1' * g_1$ is the high frequency content of the image, while the output of $h_1' * h_1$ contains only the low-pass one. All four combinations of the outputs are then decimated by two to form four subband images. The resulting low-pass image is used for the next iteration of decomposition, while the maximum coefficients from the other three sets are stored in the wavelet tree. An example of the fused images produced by DWT pyramids is shown in Fig. 7(a).

For SIDWT, the filters g_1 and h_1 are defined as $g_1 = [0 \ \ldots \ 0 \ 0.5 \ 0 \ \ldots \ 0 \ -0.5 \ 0 \ \ldots \ 0]$ and $h_1 = [0 \ \ldots \ 0 \ 0.5 \ 0 \ \ldots \ 0 \ 0.5 \ 0 \ \ldots \ 0]$, with $2^{(k-2)}$ zeroes in the first and last set of zeroes, and $2^{(k-1)}$ zeroes in the middle set of zeroes for level k of the pyramid. While the SIDWT is very redundant (because it up-samples the filter response instead of decimating the image at each level of the pyramid), the dual-tree complex wavelet transform (DT-CWT) can achieve approximate shift invariance and only slight oversampling by filtering the image with a pair of complementary filters. DT-CWT produces real and complex coefficients at each level of the decomposition for a total of 2^d oversampling, where d is the number of levels of decomposition. Figure 7 shows an example of the fused images produced by SIDWT pyramids [Fig. 7(b)] and DT-CWT [Fig. 7(c)], respectively.

The simple DWT can be prone to artifacts as a function of position in the image, which could be particularly problematic when using the FPSS background subtraction tracker to detect motion information. As an object moves slightly, artifacts could shift in the image, resulting in many unnecessary false alarms. Hence, a SIDWT or DT-CWT is expected to perform better in a tracking task. Similar to other pyramid methods, we use the maximum coefficient from either wavelet tree at each level during the image fusion phase.

3 FPSS Tracker

The effects of different image fusion methods were examined and compared using the existing FPSS moving target tracking algorithm that was developed and tested with the original FPSS datasets. In this study, the FPSS tracker was run on the original color and LWIR images, as well

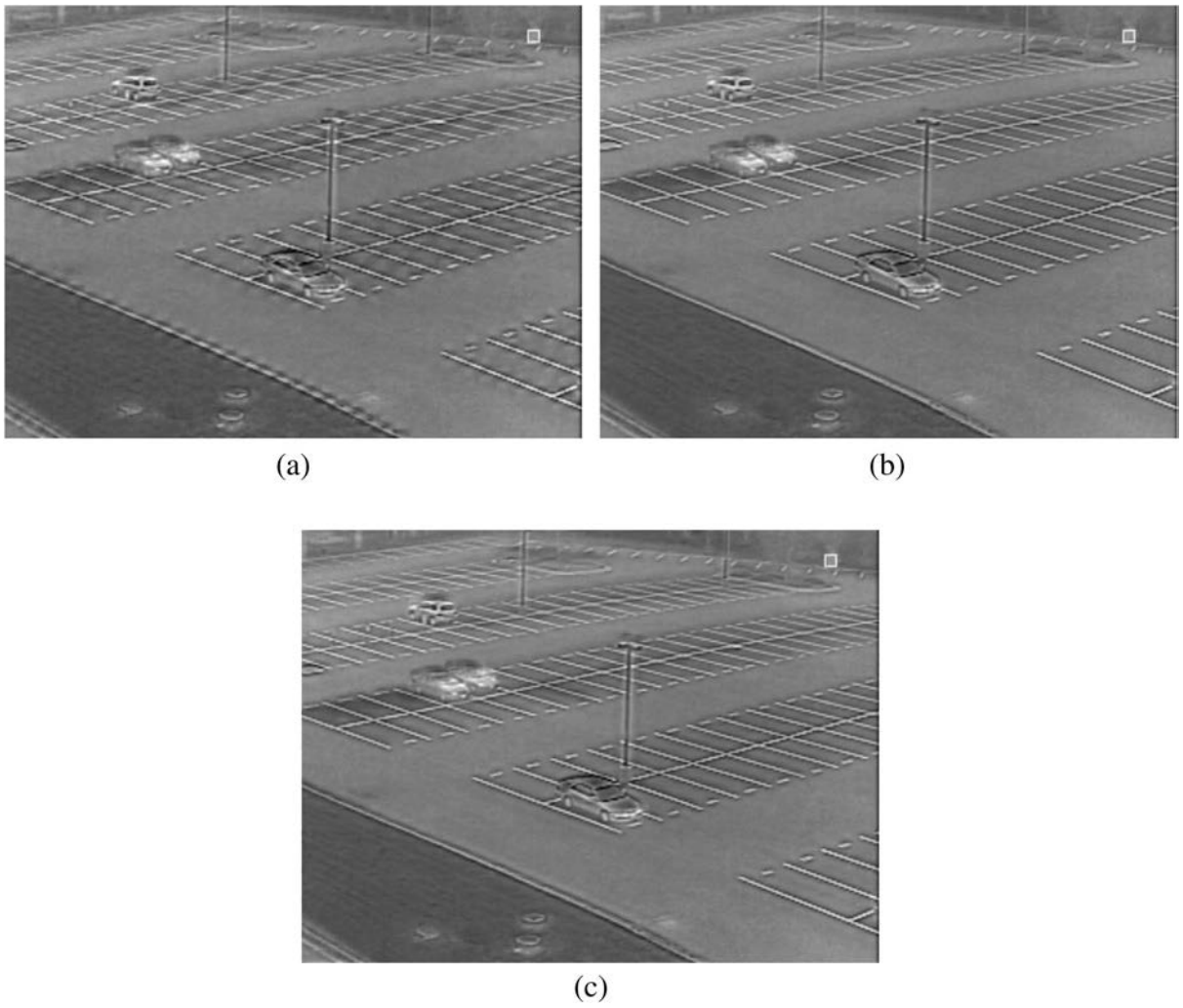


Fig. 7 Fusion by selecting the maximum coefficient of DWT pyramids (a), SIDWT pyramids (b), and DT-CWT pyramids (c).

as the fused images generated by all fusion methods described in the previous section.

3.1 *Background Modeling*

The key component of the FPSS tracker is its background modeling and subtraction process, which is depicted in Fig. 8. Each input image is first filtered by a stability mask and then channeled through four image buffers of equal size and depth. The images in buffers 2 and 4 are used to generate background models 1 and 2, respectively. Instead of being created originally from buffer 4, background model 2 can also be obtained from a buffer of models that is continuously replenished by the outgoing representations of background model 1. By subtracting the next input frame from these background models, we obtain two difference images. A difference-product image (DPI) is obtained by multiplying these two difference images pixel by pixel.

To begin the background modeling process, the first successfully preprocessed input image frame is used to fill up all image buffers and to become the initial background models. For each of the subsequent input image frames, a simple frame registration procedure is used to reduce any potential jitter effects incurred by shaking cameras. Typically, a jitter-free image contains a mostly stable background with a number of small but volatile areas caused by moving objects and other transient events. In order to prevent rapidly changing foreground pixels from ruining the background models, a stability mask is used to filter out all unstable pixels from the input image frame. Updated by the information from DPI, this stability mask looks for significant intensity changes based on a predefined threshold of variability and maintains a record of the stability index at each pixel location. Only those stable pixels on a jitter-free image are fed to buffer 1, while the once-stable but now actively changing pixels are blocked and substituted by the corresponding stable pixels available from buffer 1. Without the stable background models, it will be much harder to detect and extract legitimate moving objects in the scene, while additional false alarms will likely be generated.

Each incoming set of pixel values from the stability mask replaces the corresponding pixel values in the oldest frame in buffer 1 to form the newest frame in buffer 1, while the oldest frame of buffer 1 becomes the newest frame in buffer 2. The same mechanism of first-in first-out (FIFO) frame-shift and update is applied to all image buffers continuously. The role of buffer 1 is merely a time-delay buffer to induce a noticeable gap in time—and potentially in content—between the current input image and the image frames in buffer 2. Background model 1 is derived from the images in buffer 2, which can be as simple as taking the average of all images in buffer 2. Similar to buffer 1, buffer 3 is just another buffer to separate buffer 2 and buffer 4 in time. Background model 2 can be obtained by either processing (e.g., averaging) the images in buffer 4 or drawing from the buffer of models supplied by background model 1. The same background modeling structure depicted in Fig. 8 can be extended to include four or any larger even number of background models for more stable background representations and higher target enhancement capabilities at the expense of additional computational resources and a longer initialization period before the actual tracking begins.

One of the advantages of using multiple disjoint background models to generate a DPI is that the problematic "trailing effect," which is often associated with background subtraction method, can be suppressed effectively in this process. Since those gradually fading trails carved out by the moving objects are showing up in different parts of the disjoint difference images, as shown by the two difference images on the left side of Fig. 9, they are likely to diminish or disappear when the corresponding DPI is formed. For the same reason, time-dependent noises on the difference images are also suppressed during the formation of DPI. Another advantage of this method is that the target trails are now clearly detached from the moving objects, which allows the subsequent target detection module to estimate the size and location of those movers more accurately. With improved estimation in target size and location, the target tracking module may also perform better motion estimation and track maintenance.

An even number of background models is needed in the formation of DPI to address the problem of target polarity, which is a common target detection problem. Due to clothing and ambient temperature change, the same type of moving targets may assume different polarity of pixel intensity with respect to their immediate background. Figure 10 shows a pair of LWIR images that exhibit polarity change in human signatures during different seasons of the year. Using a single difference image or a DPI computed with any odd number of difference images to detect the moving targets will have to pick the locations with both positive and negative values simultaneously and appropriately, which is not always easy or straightforward. This problem is alleviated, however, simply as a by-product of forming the DPI using an even number of difference images.

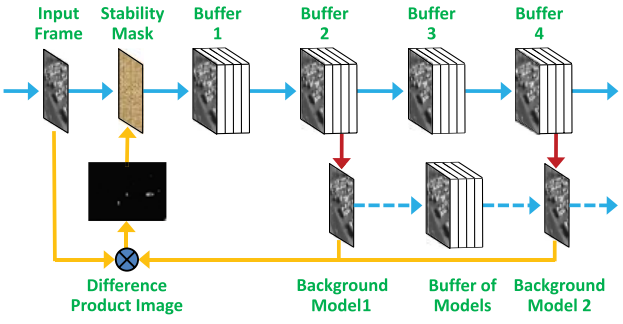


Fig. 8 The background modeling and subtraction process in the FPSS tracker.

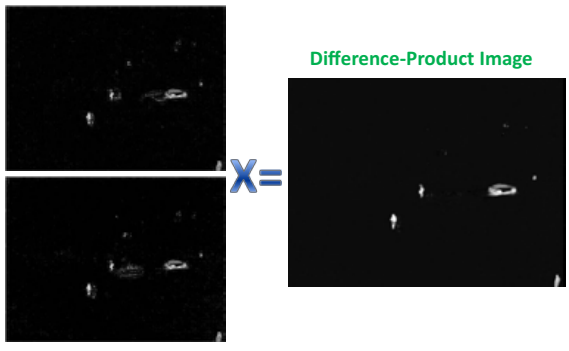


Fig. 9 Enhancement of target signatures and suppression of trailing effects and noises via a DPI.

Fig. 10 Human LWIR signatures reverse polarity in winter (a) and summer (b).

3.2 *Target Detection and Tracking*

After a DPI is generated, a morphological operation is used to remove small spikes and to fill up small gaps in the DPI. Furthermore, a pyramid-means method is used to enhance the centroid and overall silhouette of the moving targets. The moving target detection process begins with finding the brightest pixel on the post-processed DPI, which is usually associated with the most probable moving target in the given input frame. The size of this target is estimated by finding all the surrounding pixels that are deemed connected to the brightest pixel. After the first moving target is detected, all the pixels within a rectangular target-sized area of that target are suppressed to exclude them from subsequent detections. The detection process is repeated by finding the next brightest one among the remaining pixels until all the pixels are suppressed, a predefined number of detections are obtained, or other user-defined stopping criteria are reached.

Using the detection results on consecutive input images, tracks of all moving targets are built and maintained. In order to build a meaningful track, a noticeable moving target must appear in multiple contiguous frames in a video sequence. This requirement may not be met when the target is moving across the field of view of the camera at a very short range and/or a very high speed; when the camera is operated at a very low frame rate; when the target is occluded for an extended period of time and/or behind a very large obstacle; or when a combination of these and other detrimental factors occur. The FPSS tracker also uses previous locations, velocity, and target size of a moving target to predict the destination of its next movement.

4 Experimental Results

The Second FPSS dataset consists of 53 pairs of concurrent color-LWIR video sequences for a total of 71,236 frames, which depict various staged suspicious activities around a big parking lot. Each video sequence was obtained at a frame rate of 10 frames per second. No frames were dropped from any sequence in our experiments, therefore, the same frame rate was maintained across the board. Ground-truth information (target type and target location) associated with each observable moving target on each image frame was semi-manually generated using a ground-truthing GUI, storing the location of all people, vehicles, animals, and other objects for each frame. The ground truth files associated with a given pair of color-LWIR sequences may vary slightly in their content, as some moving targets may sometimes be observable in one but not both of the imageries. Because we used the LWIR approximation coefficients during the pyramid decompositions, and because LWIR ground truth files usually contain more information on the targets, we chose the LWIR ground-truths files for the purpose of verifying the detections on fused images. Based on the ground-truth information and the target size estimated by the FPSS tracker, we may compute the tracking performance achievable by the original color and LWIR sequences, as well as the performances pertaining to the fused image sequences generated by different image fusion methods.

To qualify as a correct detection or a hit, the ground-truth location must be included in the bounding box (target size) estimated by the FPSS tracker for the given detection. Multiple detections on the same target were counted as only one hit, but multiple detections on a nontarget were treated as multiple false alarms (FAs). When multiple targets in proximity were covered by a single detection, it would be treated as multiple hits. Ground-truth targets that were not included by the bounding box of any detection were regarded as misses. An adjustable acceptance threshold was used to vary the tradeoff between hits and FAs. While a range of acceptance thresholds from 0.1 to 25,000 was initially considered, we actually used the acceptance thresholds from 30 to 25,000 because very few FPSS responses had an activation level of under 30 in our experiments. Instead of normalizing the FPSS responses from different fusion methods and comparing their performance at different acceptance thresholds, we just compared their hit rates at certain fixed FA rates. By plotting the FA rate (FAR) (average number of incorrect detections per frame) against the hit rate (percentage of true targets that were correctly detected) at different acceptance thresholds, a receiver operating characteristic (ROC) curve results. To emphasize the critical differences between the ROC curves, we focused on the two end zones of these curves and examined the performance at FAR of less than 0.1 FA/frame and at hit rates exceeding 80%. The ROC curves for the original LWIR and color sequences were first generated, as shown in Fig. 11, serving as the benchmark

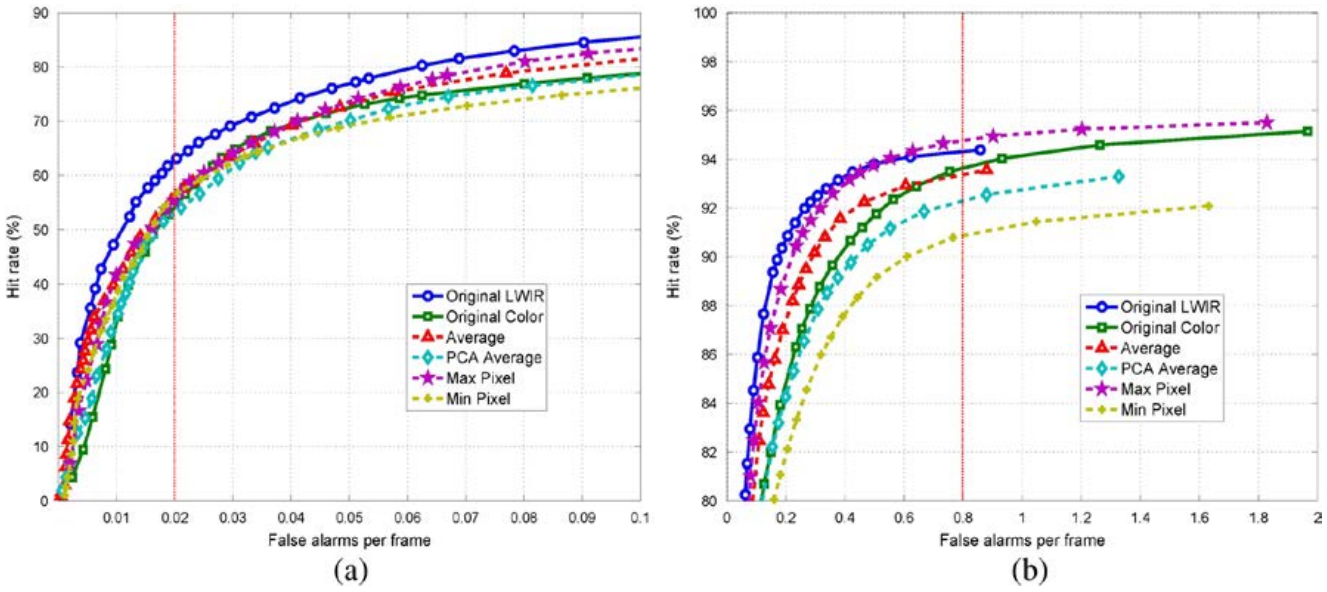


Fig. 11 The performance of four simple-combination methods at low FAR region (a) and high hit rate region (b).

performance curves that are included in all performance-related figures for comparison purposes.

Figure 11 also shows the ROC curves associated with the fused images generated by the four simple-combination methods: simple averaging, PCA-weighted averaging, maximum pixel selection, and minimum pixel selection. Their performances at the low FAR region are shown on the left graph, while the right graph shows their performance as more FAs are allowed. From the left graph, it is clear that the original LWIR images performed the best at low FAR among this group of six candidates. On the other hand, the original color images were lagging behind their LWIR counterparts consistently due to a significant increase in the number of FAs caused by headlight glares and windshield reflections in the evening hours, and protracted shadows under the slanted sun. The right graph shows that the advantage of LWIR sequences over color sequences continues to hold at the high FAR region.

Given the nature of simply averaging or selecting the pixels of the original color and LWIR images by the four simple-combination methods, we expected that their resulting fused images would perform somewhere between the original color and LWIR images. As evident from Fig. 11(a), this was, indeed, the case for the FAR region of 0.02 or less FAs per frame. As the allowable number of FAs was increased, the fused images produced by simple averaging and maximum pixel selection methods continued to yield hit rates that were between those produced by the original color and LWIR images, as demonstrated in Fig. 11(b). The performance associated with the fused images generated by the PCA-weighted averaging and minimum pixel selection methods, however, gradually fell below the performance of the original color images. In other words, there was no performance gain in tracking at any FAR by using the images fused with simple combination methods over the original LWIR images. At FARs higher than 0.02 FA per frame, even the original color images outperformed the fused images produced by the PCA-weighted averaging and minimum pixel methods.

The fusion methods based on pyramid structures were performed using an identical set of configuration parameters, which is using five levels of decomposition and a 7×7 neighborhood size when running a saliency/match measure. Based on their resulting ROC curves, these pyramid-based fusion methods were categorized into two groups for subsequent discussions: four inferior methods and five superior methods. As shown in Table 1, all nine pyramid-based methods are much more computationally intensive than the four

Table 1 CPU time (sec) needed to fuse 30 images using Matlab code on a Dell T7400 workstation.

Simple combinations	CPU time	Inferior pyramids	CPU time	Superior pyramids	CPU time
Simple average	1.280	FSD	21.670	Laplacian	24.040
PCA average	2.030	Gradient	78.970	ROLP	23.050
Maximum pixel	1.560	DWT	22.740	Contrast	23.240
Minimum pixel	1.840	Morphological	62.530	SIDWT	209.600
				DT-CWT	49.940

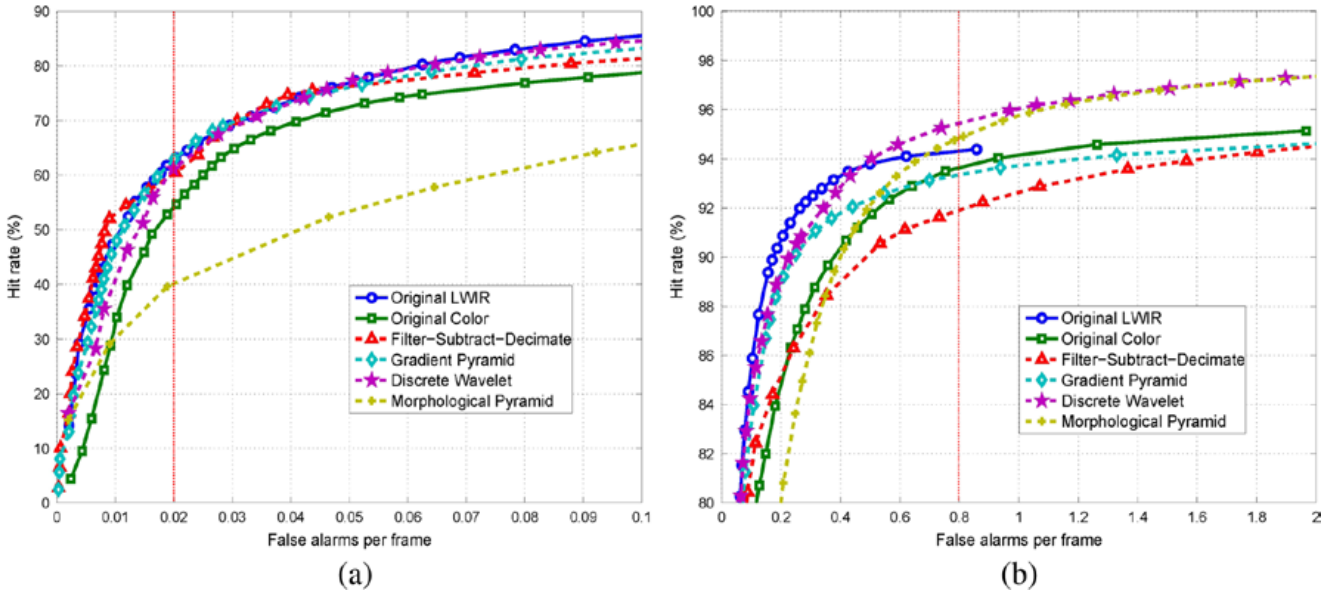


Fig. 12 The performance of four inferior pyramid-based fusion methods at low FAR region (a) and high hit rate region (b).

simple combination methods, especially the SIDWT, gradient, and morphological pyramids. Although the DT-CWT is more than four times more efficient than its more redundant variant, SIDWT, it is still considerably slower than the five simpler pyramid-based methods, three of which are ranked together in the superior pyramid column. More computations do not always generate better results, and as we can see, among the pyramid-based methods there are faster and slower candidates in both the inferior and superior columns of Table 1.

As shown by Fig. 12(a), the FSD, gradient, and DWT achieved slightly worse performance than the original LWIR images at low FARs, whereas the morphological pyramid method clearly lagged behind others under the same conditions. The picture is somewhat different at the other end of these ROC curves, as shown by Fig. 12(b), where the DWT and morphological pyramid methods were able to surpass the LWIR curve at the FAR region of 0.7 FA per frame or higher. Since alternative pyramid-based methods offer more consistent gains over the complete range of FAR, we deem these four pyramid-based methods inferior.

Finally, there are five pyramid-based fusion methods that have achieved good results on both ends of the ROC curves: the Laplacian, ROLP, contrast, SIDWT, and DT-CWT pyramid methods. As shown by Fig. 13(a), these five fusion methods clearly outperformed the original color and LWIR images from the beginning and attained the largest advantage at the FAR of around 0.02 FA per frame. At this FAR, the hit rates for the original color and LWIR images are 54.29% and 62.99%, respectively. As shown in Table 2, the corresponding hit rates of the images fused by contrast pyramid and ROLP pyramid methods are 76.94% and 75.11%,

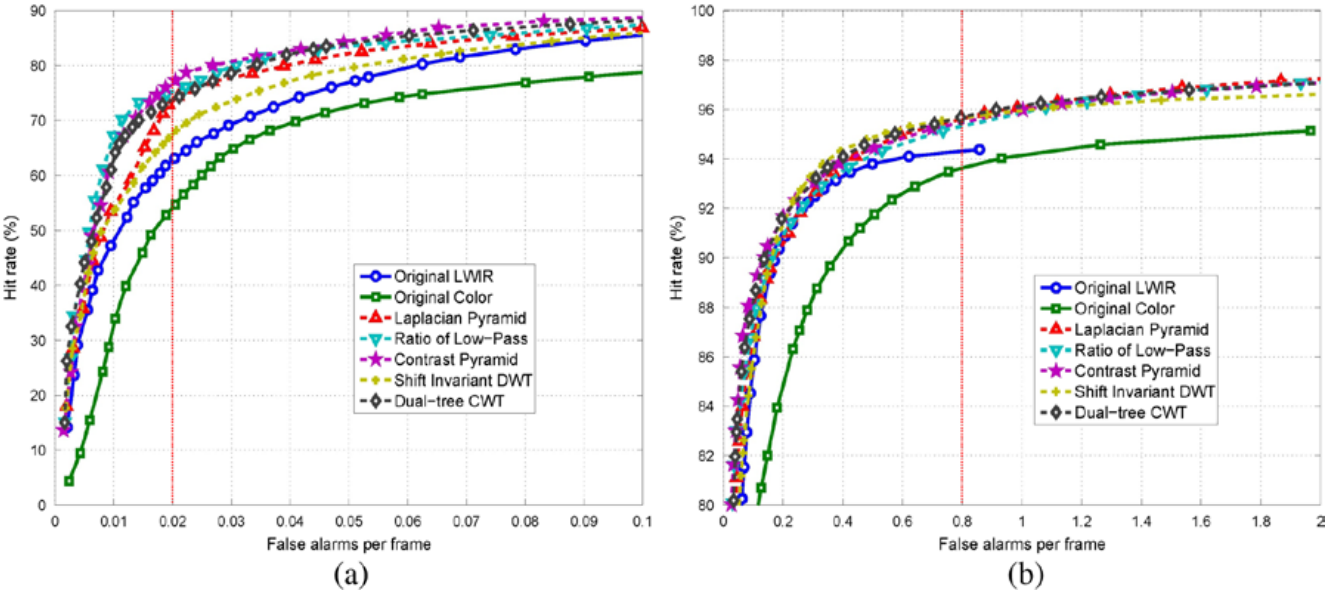


Fig. 13 The performance of five superior pyramid-based fusion methods at low FAR region (a) and high hit rate region (b).

Table 2 Performance (hit rate in %/FA per frame) of the 13 fusion methods at low FAR region.

Simple combinations	HR/FAR	Inferior pyramids	HR/FAR	Superior pyramids	HR/FAR
Simple average	56.14/0.02005	FSD	60.34/0.02008	Laplacian	73.51/0.02008
PCA average	53.43/0.02008	Gradient	62.90/0.02005	ROLP	75.11/0.02005
Maximum pixel	55.71/0.02008	DWT	61.47/0.02008	Contrast	76.94/0.02005
Minimum pixel	56.17/0.02008	Morphological	40.27/0.02008	SIDWT	67.53/0.02005
				DT-CWT	73.80/0.02002

respectively. With improvements of 12 to 14% over the LWIR images, the performance gains achieved by these two fusion methods are quite remarkable at this FAR.

Among the five superior pyramid-based methods, SIDWT is clearly lagging behind other methods in performance. Furthermore, the computational complexity of SIDWT is about nine times of that of the contrast pyramid and ROLP pyramid methods. Therefore, SIDWT is the least desirable method among this group. Although the performance of DT-CWT is competitive to those of the contrast, ROLP, and Laplacian methods, it requires more than twice as much CPU time to complete the same image fusion task, which makes it less attractive among the group of superior pyramids.

Based on the performance of the superior fusion methods at the high FAR region shown in Fig. 13(b), it is obvious that the advantage of these methods over the original color and LWIR images is still maintained at every point in the high FAR region, even though the performance gain is less significant than that in low FAR region. For example, the hit rates of the images fused by contrast pyramid and ROLP pyramid methods at a FAR of 0.80 FA per frame are 95.52% and 95.37%, respectively, exceeding those of color (93.66%) and LWIR (94.31%) images by slightly more than 1%.

As an object moves, it appears as a constant shift of some pixels in the image. For a method such as the regular DWT, large coefficient variations can be incurred by this shift. If it occurs in either the LWIR or color image, these large image coefficients can significantly outweigh the expected ones during the image fusion process. Instead of appearing as slight movements, some random variations may appear instead, which are more likely to be interpreted as noise. Due to the shift invariance property of DT-CWT, this transform is less affected by this problem and hence performs among the best. SIDWT should have offered similar advantages as the DT-CWT does, but the oversampling during the decomposition process actually produces some conflicts when the color and LWIR images are merged, thus hampers the SIDWT performance somewhat. The ROLP, contrast, and Laplacian methods work exceedingly well, because they all emphasize the contrast (brightness variation in color or LWIR images) information in the input images. On the other hand, gradient and morphological methods that emphasize the edge information in the input images do not perform very well because many faint targets may not have strong edges.

5 Conclusions

Although a given sensor may be easily fooled sometimes, it is much harder to trick a number of sensors simultaneously at any given time. For this reason, we explored and exploited the rather complementary natures of two common imaging sensors: LWIR and color visible sensors. Instead of harnessing prior background knowledge and external information sources (such as metadata on weather conditions, time of the day, season of the year, site characteristics, number of targets, target ranges, depression angle, speed of movement, and other related information) to perform symbolic-level image fusion, we focused solely on pixel-level image fusion in this work. Therefore, the techniques examined and the results obtained in this work are more readily transferrable to other applications and scenarios that process color and LWIR imageries.

Based on the results generated by the four simple-combination methods examined in this work, we conclude that these simple methods are not useful, because their performances were worse than using the original LWIR images alone. Among the nine pyramid-based image fusion methods, the gradient and FSD methods are the worst candidates because they required 10 to 60 times more CPU time than those required by the simple combination methods, but performed even worse at the high FAR region. The morphological and DWT methods are slightly better than the gradient and FSD methods, primarily because they managed to outperform LWIR in the high FAR region. Given their performances and computational requirements, these four pyramid-based methods are deemed as inferior methods in general.

The Laplacian, ROLP, contrast, SIDWT, and DT-CWT are found to be superior image fusion methods, because they consistently outperformed LWIR in every FAR region. The contrast and ROLP methods are considered the best image fusion methods to pair with the FPSS tracker because their ROC curves are consistently on top of all other ROC curves produced in this work. Furthermore, the computational requirements of these two methods are almost the lowest among the pyramid-based methods. On the other hand, SIDWT is ranked at the bottom in this group, as it performed the worst and consumed four to nine times more CPU time than its counterparts in this group did.

For future work, a potential way of improving image fusion performance is to treat each color image as three separate images (R, G, and B images) and fuse these three images with the LWIR image together. The fusion algorithms

examined in this work do not limit the number of images that can be fused together. Therefore, short-wave infrared, mid-wave infrared, and hyperspectral imageries could also be considered, if they are properly coregistered. Performance may also be improved by linking the image fusion process with the tracking algorithm, through which the information that is critical to the tracker may be better preserved or enhanced. For instance, a region-based segmentation algorithm may be incorporated into the DT-CWT image fusion process.[18,19] The segmentation algorithm could exploit the limited redundancy in DT-CWT and tie the feature level and pixel level fusion algorithms together.

Acknowledgments

This work was partially supported by the U.S. Army Research Office MURI W311NF-07-1-0185 grant, National Science Foundation Graduate Fellowship Program, National Defense Science and Engineering Graduate Fellowship Program, and the Texas Instruments Leadership University Program.

References

1. A. L. Chan, "A description on the second dataset of the U.S. army research laboratory force protection surveillance system," *ARL-MR-0670*, U. S. Army Research Laboratory, Adelphi, Maryland (2007).
2. M. Smith and J. Heather, "Review of image fusion technology in 2005," *Proc. SPIE* **5782**, 29–45 (2005).
3. J. Han and B. Bhanu, "Fusion of color and infrared video for moving human detection," *Pattern Recogn.* **40**(6), 1771–1784 (2007).
4. M. Motwani et al., "Towards benchmarking of video motion tracking algorithms," in *Proc. Int. Conf. Signal Acquisition and Processing*, pp. 215–219, IEEE (2010).
5. N. Cvejic et al., "The effect of pixel-level fusion on object tracking in multi-sensor surveillence video," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.* Vol. **372**, pp. 1–7, IEEE (2007).
6. L. Mihaylova et al., "The influence of multi-sensor video fusion on object tracking using a particle filter,"in *Proc. 2nd Workshop on Multiple Sensor Data Fusion*, pp. 354–358 (2006).
7. A. L. Chan, "A robust target tracking algorithm for FLIR imagery," *Proc. SPIE* **7696**, 769603 (2010).
8. E. Trucco and K. Plakas, "Video tracking: A concise survey," *IEEE J. Ocean. Eng.* **31**(2), 520–529 (2006).
9. V. Tsagaris and V. Anastassopoulos, "Fusion of visible and infrared imagery for night color vision," *Displays* **26**(4–5), 191–196 (2005).
10. A. L. Chan, S. Z. Der, and N. M. Nasrabadi, "Dualband FLIR fusion for automatic target recognition," *Inform. Fusion* **4**(1), 35–45 (2003).
11. W. H. Buchsbaum, *Color TV Servicing*, 3rd ed., Prentice Hall, Englewood Cliffs, New Jersey (1975).
12. P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.* **31**(4), 532–540 (1983).
13. H. Anderson, "A filter-subtract-decimate hierarchical pyramid signal analyzing and synthesizing technique," U. S. Patent No. 718104 (1987).
14. A. Toet, "Image fusion by a ratio of low-pass pyramid," *Pattern Recogn. Lett.* **9**(4), 245–253 (1989).
15. A. Toet, J. J. van Ruyven, and J. M. Valeton, "Merging thermal and visual images by a contrast pyramid," *Opt. Eng.* **28**(7), 789–792 (1989).
16. P. Burt, "A gradient pyramid basis for pattern selective image fusion," in *SID Int. Symp. Dig. Tec.*, Vol. 23, pp. 467–470, Society for Information Display (1992).
17. L. C. Ramac, M. K. Uner, and P. K. Varshney, "Morphological filters and wavelet based image fusion for concealed weapon detection," *Proc. SPIE* **3376**, 110–119 (1998).
18. J. J. Lewis et al., "Region-based image fusion using complex wavelets," in *Proc. Int. Conf. Information Fusion*, International Society for Information Fusion, Stockholm (2004).
19. J. J. Lewis et al., "Pixel- and region-based image fusion with complex wavelets," *Inform. Fusion* **8**(2), 119–130 (2007).

Alex Lipchen Chan received a BS in electrical engineering from the Chung Yuan Christian University, Chung-Li, Taiwan. He also received a MS in industrial engineering and a MS and PhD in electrical and computer engineering from the State University of New York at Buffalo. Since 1995, he has been a research scientist at the U.S. Army Research Laboratory, Adelphi, Maryland. His research interests include automatic target recognition, computer vision, and artificial neural networks.

Stephen R. Schnelle received a BS in electrical and computer engineering and a BS in mathematics from Rose-Hulman Institute of Technology (Terre Haute, Indiana) in 2007. He also received a MS degree in electrical and computer engineering from Rice University (Houston, Texas) in 2011. He is currently a PhD student in the Department of Electrical and Computer Engineering at Rice University (Houston, Texas). His research interests include compressive sensing, matrix completion, statistical signal and image processing, data analysis, and engineering education.

Sparse Representation for Target Detection in Hyperspectral Imagery

Yi Chen, Nasser M. Nasrabadi and Trac D. Tran

IEEE Journal of Selected Topics in Signal Processing 2011, 5 (3), 629-640.

Sparse Representation for Target Detection in Hyperspectral Imagery

Yi Chen, Nasser M. Nasrabadi, *Fellow, IEEE*, and Trac D. Tran, *Senior Member, IEEE*

***Abstract*—In this paper, we propose a new sparsity-based algorithm for automatic target detection in hyperspectral imagery (HSI). This algorithm is based on the concept that a pixel in HSI lies in a low-dimensional subspace and thus can be represented as a sparse linear combination of the training samples. The sparse representation (a sparse vector corresponding to the linear combination of a few selected training samples) of a test sample can be recovered by solving an ℓ_0-norm minimization problem. With the recent development of the compressed sensing theory, such minimization problem can be recast as a standard linear programming problem or efficiently approximated by greedy pursuit algorithms. Once the sparse vector is obtained, the class of the test sample can be determined by the characteristics of the sparse vector on reconstruction. In addition to the constraints on sparsity and reconstruction accuracy, we also exploit the fact that in HSI the neighboring pixels have a similar spectral characteristic (smoothness). In our proposed algorithm, a smoothness constraint is also imposed by forcing the *vector Laplacian* at each reconstructed pixel to be minimum all the time within the minimization process. The proposed sparsity-based algorithm is applied to several hyperspectral imagery to detect targets of interest. Simulation results show that our algorithm outperforms the classical hyperspectral target detection algorithms, such as the popular spectral matched filters, matched subspace detectors, adaptive subspace detectors, as well as binary classifiers such as support vector machines.**

***Index Terms*—Hyperspectral imagery, sparse recovery, sparse representation, spatial correlation, target detection.**

I. Introduction

HYPERSPECTRAL remote sensors capture digital images in hundreds of narrow spectral bands (about 10 nm wide), which span the visible to infrared spectrum [1]. Pixels in HSI are represented by B-dimensional vectors where B is the number of spectral bands. Different materials are usually assumed to be spectrally separable as they reflect electromagnetic energy differently at specific wavelengths. This property enables discrimination of materials based on the radiance spectrum obtained by hyperspectral imagery. HSI has found many applications in various fields such as military [2]–[4], agriculture [5], [6], and mineralogy [7]. One of the important applications of HSI is target detection, which can be viewed as a two-class classification problem where pixels are labeled as target (target present) or background (target absent) based on their spectral characteristics. Support vector machines [8], [9] have been a powerful tool to solve supervised classification problems and have shown a good classification performance for hyperspectral classification [10], [11]. A number of algorithms also have been proposed for target detection in HSI based on statistical hypothesis testing techniques [2]. Among these approaches, spectral matched filters [12], [13], matched subspace detectors [14], and adaptive subspace detectors [15] have been widely used to detect targets of interests. The details of these classical algorithms will be described in the next section.

Recently, a novel signal classification technique via sparse representation have been proposed for face recognition [16]. It is observed that aligned faces of the same object with varying lighting conditions approximately lie in a low-dimensional subspace [17]. Thus, a test face image can be sparsely represented by training samples from all classes. The most compact representation can be obtained by solving a sparsity-constrained optimization problem. This algorithm exploits the discriminative nature of sparse representation and the reconstruction of the test sample provides directly its classification label. This idea naturally extends to other signal classification problems such as iris recognition [18], tumor classification [19], and HSI unmixing [20].

In this paper, we propose a target detection algorithm based on sparse representation for HSI data. We use the same sparsity model in [16] where a test sample is approximately represented by very few training samples from both target and background dictionaries, and the recovered sparse representation is used directly for detection. In addition to the constraints on sparsity and reconstruction accuracy, we show that it is necessary to exploit the fact that neighboring HSI pixels usually have a similar spectral characteristics as well. To achieve this, we impose a smoothing constraint on the reconstructed image by forcing the vector Laplacian, as defined in Section III-D, of the reconstructed pixels to be zero. By incorporating this spatial correlation, the detection performance is significantly improved for images in which targets consist of multiple pixels.

One of the advantages of our proposed approach is that there is no explicit assumption on the statistical distribution characteristics of the observed data as in the previous target detection algorithms [12]–[15]. Furthermore, in the spectral matched filter, the target spectral signature is a single vector, usually obtained

Manuscript received April 20, 2010; revised September 24, 2010 and December 20, 2010; accepted January 31, 2011. Date of publication February 10, 2011; date of current version May 18, 2011. This paper was supported in part by the Army Research Office (ARO) under Grant 58110-MA-II and in part by the National Science Foundation (NSF) under Grant CCF-0728893. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Gustavo Camps-Valls.

Y. Chen and T. D. Tran are with the Department of Electrical and Computer Engineering, The Johns Hopkins University, Baltimore, MD 21218 USA (e-mail: ychen98@jhu.edu; trac@jhu.edu).

N. M. Nasrabadi is with the U.S. Army Research Laboratory, Adelphi, MD 20783 USA (e-mail: nnasraba@arl.army.mil).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTSP.2011.2113170

by averaging the training target samples or from a spectral library. However, using a single target spectrum is usually insufficient to represent the target spectral characteristics since the target spectrum changes with the environmental situation. This problem can be avoided by using a target subspace model represented by training samples that account for the target spectrum under various conditions of illumination and atmospheric conditions, making the dictionary invariant to the environmental variations [21], [22]. This environmental invariant approach can easily be incorporated into our algorithm by augmenting the target and background dictionaries with synthetically generated spectral signatures in order to construct better target and background subspaces. Moreover, unlike the other detectors based on statistical hypothesis testing, the sparsity model in our approach has the flexibility of imposing additional restrictions corresponding to the characteristics of HSI such as smoothness across neighboring hyperspectral pixels.

The paper is structured as follows. Section II briefly describes several previously proposed approaches commonly used in automatic target detection in HSI. Our sparsity-driven target detection algorithm is presented in Section III. The effectiveness of the proposed method is demonstrated by simulation results presented in Section IV. Conclusions are drawn in Section V. Throughout this paper, matrices and vectors are denoted by upper and lower case boldface letters, respectively.

II. Previous Approaches

In this section, we briefly introduce previously developed approaches for target detection in HSI. Specifically, we describe problem formulation of support vector machines (SVMs), followed by the signal models and detector expressions of the classical detectors including spectral matched filter (SMF), matched subspace detectors (MSDs), and adaptive subspace detectors (ASDs). Implementation details of the three statistical detectors and their nonlinear (kernel) versions can be found in [23], whereas details of SVM can be found in [24].

A. Support Vector Machines

The SVM approach [8] solves the supervised binary classification problem by seeking the optimal hyperplane that separates two classes with the largest margin. A nonlinear SVM (called kernel SVM) is often implemented to further improve the separation between classes by projecting the samples onto a higher dimensional feature space. In kernel SVM, the dot products in the original SVM formulation are replaced by a nonlinear kernel function using the kernel trick [8].

It has also been shown that the integration of the contextual information via composite kernels in SVM (i.e., contextual SVM) leads to an improvement in HSI classification over the traditional spectral-only SVM [24], [25]. In contextual SVM, a pixel $\boldsymbol{x}_i$ is redefined as a combination of the spectral pixel $\boldsymbol{x}_i^w$ and its spatial feature $\boldsymbol{x}_i^s$ (e.g., the mean and standard deviation per spectral band) extracted in a small neighborhood. In this paper, we implemented contextual SVM with a composite kernel that fuses the spectral and spatial information via a weighted summation

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \mu K_s\left(\boldsymbol{x}_i^s, \boldsymbol{x}_j^s\right) + (1-\mu) K_w\left(\boldsymbol{x}_i^w, \boldsymbol{x}_j^w\right) \tag{1}$$

where $\mu \in [0,1]$ is the tradeoff between spatial kernel K_s and spectral kernel K_w. Examples of possible kernels can be found in [26].

B. Spectral Matched Filter

Let $\boldsymbol{x} = [x_1 \quad x_2 \quad \cdots \quad x_B]^T$ be a spectral observation consisting of B spectral bands. The model for SMF can be expressed by

$$\begin{aligned} H_0 &: \boldsymbol{x} = \boldsymbol{n}, \quad \text{target absent} \\ H_1 &: \boldsymbol{x} = a\boldsymbol{s} + \boldsymbol{n}, \quad \text{target present} \end{aligned} \tag{2}$$

where a is the target abundance measure ($a = 0$ when no target is present and $a > 0$ when a target is present), $\boldsymbol{s} = [s_1 \quad s_2 \quad \cdots \quad s_B]^T$ is the spectral signature of the target, and $\boldsymbol{n}$ is the additive background noise.

Assume $\boldsymbol{n}$ is zero-mean Gaussian random noise. Using the generalized likelihood ratio test (GLRT), the output of SMF for a test input $\boldsymbol{x}$ is given by [12]

$$D_{\text{SMF}}(\boldsymbol{x}) = \frac{\boldsymbol{s}^T \hat{\boldsymbol{C}}^{-1} \boldsymbol{x}}{\boldsymbol{s}^T \hat{\boldsymbol{C}}^{-1} \boldsymbol{s}} \tag{3}$$

where $\hat{\boldsymbol{C}}$ represents the estimated covariance matrix for the centered observation data. If the output $D_{\text{SMF}}(\boldsymbol{x})$ is greater than a prescribed threshold δ, then the test sample will be determined as a target; otherwise, it will be labeled as background.

Variations of SMF include the adaptive SMF (ASMF) where the background clutter covariance matrix is estimated from a small number of samples in the neighborhood of the test sample and the regularized SMF [27] where a regularization term is added to force the filter coefficients to shrink and become smooth. The regularized SMF is implemented in Section IV for detector performance comparison.

C. Matched Subspace Detectors

In the previous SMF approach, only a single target spectral signature is used. However, in MSD, a pixel is modeled in terms of target subspace and background subspace which are obtained using target and background training data, respectively. The target detection set-up for MSD is

$$\begin{aligned} H_0 &: \boldsymbol{x} = \boldsymbol{B}\boldsymbol{\zeta} + \boldsymbol{n}, \quad \text{target absent} \\ H_1 &: \boldsymbol{x} = \boldsymbol{T}\boldsymbol{\theta} + \boldsymbol{B}\boldsymbol{\zeta} + \boldsymbol{n}, \quad \text{target present} \end{aligned} \tag{4}$$

where $\boldsymbol{B}$ and $\boldsymbol{T}$ represent matrices whose columns are linearly independent and span the background and target subspaces, respectively; $\boldsymbol{\zeta}$ and $\boldsymbol{\theta}$ are unknown vectors whose entries are coefficients accounting for the abundances of the corresponding column vectors of $\boldsymbol{B}$ and $\boldsymbol{T}$, respectively; and $\boldsymbol{n}$ is additive Gaussian noise.

The GLRT for the above model is [14]

$$D_{\text{MSD}}(\boldsymbol{x}) = \frac{\boldsymbol{x}^T(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{B}})\boldsymbol{x}}{\boldsymbol{x}^T(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{T}B})\boldsymbol{x}} \tag{5}$$

where $\boldsymbol{P}_{\boldsymbol{B}}$ is the projection matrix associated with the background subspace $\langle \boldsymbol{B} \rangle$, and $\boldsymbol{P}_{\boldsymbol{T}B}$ is the projection matrix associated with the target-and-background subspace $\langle \mathbf{TB} \rangle$. Usually, the eigenvectors corresponding to the significant eigenvalues of

the target and background covariance matrices are used to generate the columns of $\boldsymbol{T}$ and $\boldsymbol{B}$, respectively. For a prescribed threshold δ, if the output $D_{\mathrm{MSD}}(\boldsymbol{x}) > \delta$, then $\boldsymbol{x}$ will be labeled as target; otherwise, it will be labeled as background.

D. Adaptive Subspace Detectors

A scaled background noise under H_1 is used in ASD because in the case of subpixel targets, the amount of background covered area may be different from that of a pure background pixel. For ASD, the detection model for a measurement $\boldsymbol{x}$ is

$$\begin{aligned} H_0 &: \boldsymbol{x} = \boldsymbol{n}, \quad \text{target absent} \\ H_1 &: \boldsymbol{x} = \boldsymbol{U\theta} + \sigma\boldsymbol{n}, \quad \text{target present} \end{aligned} \tag{6}$$

where $\boldsymbol{U}$ is a matrix whose columns are linearly independent vectors that span the target subspace $\langle \boldsymbol{U} \rangle$, $\boldsymbol{\theta}$ is an unknown vector of the abundances of the corresponding columns of $\boldsymbol{U}$, $\boldsymbol{n}$ is Gaussian random noise, and σ is a scalar. The measurement $\boldsymbol{x}$ is assumed to be background noise under hypothesis H_0 and a linear combination of a target subspace signal and scaled background noise under hypothesis H_1.

The GLRT for the above problem is given by [15]

$$D_{\mathrm{ASD}}(\boldsymbol{x}) = \frac{\boldsymbol{x}^T\hat{\boldsymbol{C}}^{-1}\boldsymbol{U}(\boldsymbol{U}^T\hat{\boldsymbol{C}}^{-1}\boldsymbol{U})^{-1}\boldsymbol{U}^T\hat{\boldsymbol{C}}^{-1}\boldsymbol{x}}{\boldsymbol{x}^T\hat{\boldsymbol{C}}^{-1}\boldsymbol{x}} \tag{7}$$

where $\hat{\boldsymbol{C}}$ is the estimated background covariance. Similar to the cases of SMF and MSD, if $D_{\mathrm{ASD}}(\boldsymbol{x}) > \delta$, then $\boldsymbol{x}$ will be declared as target; otherwise, it will be labeled as background.

III. Sparsity-Based Target Detection

In this section, we introduce the first sparsity-based HSI target detection algorithm by sparsely representing the test sample using a structured dictionary consisting of target and background training samples. We first describe the details of the sparse subspace model employed in the proposed algorithm, and then demonstrate its ability as a classifier.

A. Sparsity Model

Let $\boldsymbol{x}$ be a hyperspectral pixel observation, which is a B-dimensional vector whose entries correspond to responses to various spectral bands. If $\boldsymbol{x}$ is a background pixel, its spectrum approximately lies in a low-dimensional subspace spanned by the background training samples $\{\boldsymbol{a}_i^b\}_{i=1,2,\ldots,N_b}$. The pixel $\boldsymbol{x}$ can then be approximately represented as a linear combination of the training samples as follows:

$$\begin{aligned} \boldsymbol{x} &\approx \alpha_1\boldsymbol{a}_1^b + \alpha_2\boldsymbol{a}_2^b + \cdots + \alpha_{N_b}\boldsymbol{a}_{N_b}^b \\ &= \underbrace{[\boldsymbol{a}_1^b \quad \boldsymbol{a}_2^b \quad \cdots \quad \boldsymbol{a}_{N_b}^b]}_{\boldsymbol{A}_b}\underbrace{[\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_{N_b}]^T}_{\boldsymbol{\alpha}} \\ &= \boldsymbol{A}_b\boldsymbol{\alpha} \end{aligned} \tag{8}$$

where N_b is the number of background training samples, $\boldsymbol{A}_b$ is the $B \times N_b$ background dictionary whose columns are the background training samples (also called atoms), and $\boldsymbol{\alpha}$ is an unknown vector whose entries are the abundances of the corresponding atoms in $\boldsymbol{A}_b$. In our model, $\boldsymbol{\alpha}$ turns out to be a sparse vector (i.e., a vector with only few nonzero entries). To better illustrate this model, an example is shown in Fig. 1. A background sample $\boldsymbol{x}$ consisting of $B = 150$ bands (blue solid) and its approximation $\boldsymbol{A}_b\boldsymbol{\alpha}$ (red dashed) are shown in Fig. 1(a). The background dictionary contains $N_b = 1300$ training samples which are randomly picked from the entire image including spectral signature for multiple background materials (e.g., vegetation, dirt road, and soil). The sparse representation $\boldsymbol{\alpha}$ is shown in Fig. 1(b). We see that only 4 out of the 1300 entries of $\boldsymbol{\alpha}$ are nonzero. The four atoms (background training samples) of $\boldsymbol{A}_b$ corresponding to the nonzero entries are shown in Fig. 1(c). The test sample $\boldsymbol{x}$ is approximated by a linear combination of only four training atoms with a small reconstruction error of mean squared error $(\mathrm{MSE}) = 7.25 \times 10^{-6}$.

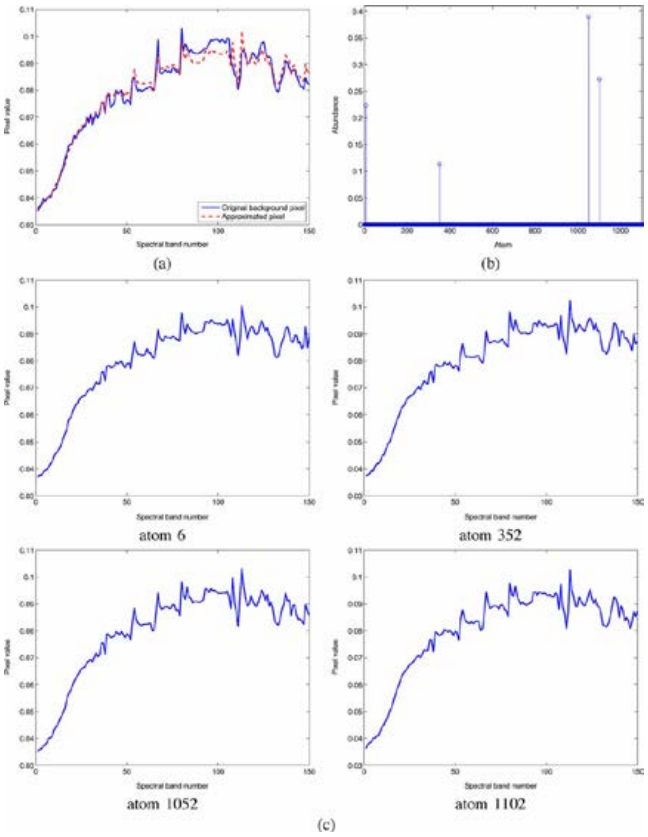


Fig. 1. Example of sparse representation of a background pixel. (a) The original pixel $\boldsymbol{x}$ (blue solid) and its approximation $\boldsymbol{A}_b\boldsymbol{\alpha}$ represented by four training samples in $\boldsymbol{A}_b$ (red dashed). The MSE between $\boldsymbol{x}$ and $\boldsymbol{A}_b\boldsymbol{\alpha}$ is 7.25×10^{-6}. (b) The sparse representation $\boldsymbol{\alpha}$ of $\boldsymbol{x}$. (c) The four background training spectral signatures corresponding to the non-zero entries of $\boldsymbol{\alpha}$.

Similarly, a target pixel $\boldsymbol{x}$ approximately lies in the target subspace spanned by the target training samples $\{\boldsymbol{a}_i^t\}_{i=1,2,\ldots,N_t}$, which can also be sparsely represented by a linear combination of the training samples

$$\begin{aligned} \boldsymbol{x} &= \beta_1\boldsymbol{a}_1^t + \beta_2\boldsymbol{a}_2^t + \cdots + \beta_{N_t}\boldsymbol{a}_{N_t}^t \\ &= \underbrace{[\boldsymbol{a}_1^t \quad \boldsymbol{a}_2^t \quad \cdots \quad \boldsymbol{a}_{N_t}^t]}_{\boldsymbol{A}_t}\underbrace{[\beta_1 \quad \beta_2 \quad \cdots \quad \beta_{N_t}]^T}_{\boldsymbol{\beta}} \\ &= \boldsymbol{A}_t\boldsymbol{\beta} \end{aligned} \tag{9}$$

where N_t is the number of target training samples, $\boldsymbol{A}_t$ is the $B \times N_t$ target dictionary consisting of the target training pixels, and $\boldsymbol{\beta}$ is a sparse vector whose entries contain the abundances

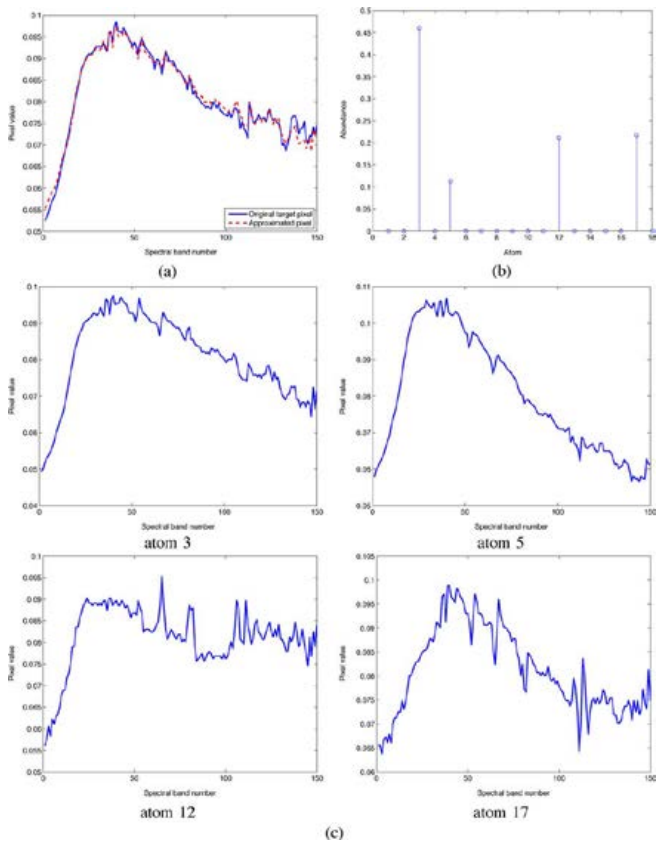


Fig. 2. Example of sparse representation of a target test sample. (a) The original pixel $\boldsymbol{x}$ (blue solid) and reconstructed pixel $\boldsymbol{A}_t\boldsymbol{\beta}$ represented by four training samples in $\boldsymbol{A}_t$ (red dashed). The MSE between $\boldsymbol{x}$ and $\boldsymbol{A}_t\boldsymbol{\beta}$ is 1.70×10^{-6}. (b) The sparse representation $\boldsymbol{\beta}$ of $\boldsymbol{x}$. (c) The four target training spectral signatures corresponding to the nonzero entries of $\boldsymbol{\beta}$.

of the corresponding target atoms in $\boldsymbol{A}_t$. An example demonstrating the effectiveness of this sparse-representation model is depicted in Fig. 2. The target dictionary has $N_t = 18$ training samples. Note that because of the lack of availability of the target spectral signatures, the size of the training dictionary for targets is usually much smaller than that of the background dictionary. Fig. 2(a) shows the original target spectral (blue solid) and its approximation (red dashed) from four training atoms. The sparse vector $\boldsymbol{\beta}$ is shown in Fig. 2(b), and the atoms in $\boldsymbol{A}_t$ corresponding to the nonzero entries of $\boldsymbol{\beta}$ are shown in Fig. 2(c).

In our proposed detection algorithm, an unknown test sample is modeled to lie in the union of the background and target subspaces. Therefore, by combining the two dictionaries $\boldsymbol{A}_b$ and $\boldsymbol{A}_t$, a test sample $\boldsymbol{x}$ can be written as a sparse linear combination of all training pixels

$$\boldsymbol{x} = \boldsymbol{A}_b\boldsymbol{\alpha}' + \boldsymbol{A}_t\boldsymbol{\beta}' = \underbrace{[\boldsymbol{A}_b \quad \boldsymbol{A}_t]}_{\boldsymbol{A}} \underbrace{\begin{bmatrix}\boldsymbol{\alpha}' \\ \boldsymbol{\beta}'\end{bmatrix}}_{\boldsymbol{\gamma}} = \boldsymbol{A}\boldsymbol{\gamma} \qquad (10)$$

where $\boldsymbol{A} = [\boldsymbol{A}_b \quad \boldsymbol{A}_t]$ is a $B \times (N_b + N_t)$ matrix consisting of both background and target training samples, and $\boldsymbol{\gamma} = [\boldsymbol{\alpha}'^T \quad \boldsymbol{\beta}'^T]^T$ is a $(N_b + N_t)$-dimensional vector consisting of the two vectors $\boldsymbol{\alpha}'$ and $\boldsymbol{\beta}'$ associated with the two dictionaries. This model is similar to that of the MSD in (4) where the test sample is assumed to lie in a subspace spanned by training samples from both background and target classes. However, in the case of MSD, the target and background are assumed to have a Gaussian distribution and GLRT is used to develop the detector. In our sparsity-based model, no assumption about the target and background distributions is required. Also, in the MSD signal model, the columns of the background and target dictionaries have to be independent in order to generate the required projection operators. In our approach, the subspace model is more generalized since independence between the training samples is not necessary. The vector $\boldsymbol{\gamma}$ is a concatenation of the two vectors associated with the background and target dictionaries and is also a sparse vector as follows. Since the background (e.g., trees, grass, road, soil) and target (e.g., metal, paint, glass) pixels usually consist of different materials, they have distinct spectral signatures and thus the spectrum of target and background pixels lie in different subspaces. For example, if $\boldsymbol{x}$ is a target pixel, then ideally it cannot be represented by the background training samples. In this case, $\boldsymbol{\alpha}'$ is a zero vector and $\boldsymbol{\beta}'$ is a sparse vector. On the other hand, if $\boldsymbol{x}$ belongs to the background class, then $\boldsymbol{\alpha}'$ is sparse and $\boldsymbol{\beta}'$ is a zero vector. Therefore, the test sample $\boldsymbol{x}$ can be sparsely represented by combined background and target dictionaries, and the locations of nonzero entries in the sparse vector $\boldsymbol{\gamma}$ actually contains critical information about the class of the test sample $\boldsymbol{x}$. Next, we demonstrate how to obtain $\boldsymbol{\gamma}$ and how to label the class of a test sample from $\boldsymbol{\gamma}$.

B. Reconstruction and Detection

This section considers the reconstruction problem of finding the sparse vector $\boldsymbol{\gamma}$ for a test sample $\boldsymbol{x}$, given the dictionary $\boldsymbol{A}$. As discussed above, a test sample can be approximately represented by very few training samples. Given the dictionary of training samples $\boldsymbol{A} = [\boldsymbol{A}_b \quad \boldsymbol{A}_t]$, the representation $\boldsymbol{\gamma}$ satisfying $\boldsymbol{A}\boldsymbol{\gamma} = \boldsymbol{x}$ can be obtained by solving the following optimization problem for the sparsest vector:

$$\hat{\boldsymbol{\gamma}} = \arg\min ||\boldsymbol{\gamma}||_0 \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{\gamma} = \boldsymbol{x} \qquad (11)$$

where $||\cdot||_0$ denotes ℓ_0-norm which is defined as the number of nonzero entries in the vector (also called the sparsity level of the vector). The above problem of minimizing the ℓ_0-norm is a NP-hard problem. If the solution is sufficiently sparse, this NP-hard problem can be relaxed to a linear programming problem by replacing the ℓ_0-norm by ℓ_1-norm, which can then be solved efficiently by convex programming techniques [28], [29]. Alternatively, the problem in (11) can also be approximately solved by greedy pursuit algorithms such as orthogonal matching pursuit (OMP) [30] or subspace pursuit (SP) [31]. Due to the presence of approximation errors in empirical data, the equality constraint in (11) can be relaxed to an inequality one

$$\hat{\boldsymbol{\gamma}} = \arg\min ||\boldsymbol{\gamma}||_0 \quad \text{subject to} \quad ||\boldsymbol{A}\boldsymbol{\gamma} - \boldsymbol{x}||_2 \leq \sigma \qquad (12)$$

where σ is the error tolerance. The above problem can also be interpreted as minimizing the approximation error within a certain sparsity level

$$\hat{\boldsymbol{\gamma}} = \arg\min ||\boldsymbol{A}\boldsymbol{\gamma} - \boldsymbol{x}||_2 \quad \text{subject to} \quad ||\boldsymbol{\gamma}||_0 \leq K_0 \qquad (13)$$

where K_0 is a given upper bound on the sparsity level [32]. In [33], it has been shown that the solutions to the problems in (12) and (13) coincide for properly chosen parameters σ and K_0, and therefore the two problems are in some sense equivalent. In this paper, the greedy SP algorithm [31] is used to approximately solve the sparse recovery problem (13) due to its computational efficiency.

The sparse vector $\hat{\boldsymbol{\gamma}}$ is recovered by decomposing the pixel $\boldsymbol{x}$ over the given dictionary $\boldsymbol{A}$ to find the few atoms in $\boldsymbol{A}$ that best represent the test pixel $\boldsymbol{x}$. The recovery process implicitly leads to a competition between the two subspaces. Therefore, the recovered sparse representation is naturally discriminative. Once the sparse vector $\hat{\boldsymbol{\gamma}}$ is obtained, the class of $\boldsymbol{x}$ can be determined by comparing the residuals $r_b(\boldsymbol{x}) = ||\boldsymbol{x} - \boldsymbol{A}_b\hat{\boldsymbol{\alpha}}||_2$ and $r_t(\boldsymbol{x}) = ||\boldsymbol{x} - \boldsymbol{A}_t\hat{\boldsymbol{\beta}}||_2$, where $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$ represent the recovered sparse coefficients corresponding to the background and target dictionaries, respectively. In our approach, the output of detector is calculated by

$$D(\boldsymbol{x}) = r_b(\boldsymbol{x}) - r_t(\boldsymbol{x}). \tag{14}$$

If $D(\boldsymbol{x}) > \delta$ with δ being a prescribed threshold, then $\boldsymbol{x}$ is determined as a target pixel; otherwise, $\boldsymbol{x}$ is labeled as background.

Fig. 3 shows an example of sparse reconstruction of a background test sample and a comparison to the pseudo-inverse reconstruction. This example illustrates the advantage of ℓ_0-norm in classification problems over the conventional ℓ_2-norm. The pseudo-inverse solution is obtained by solving the following minimum ℓ_2-norm problem:

$$\hat{\boldsymbol{\gamma}}_2 = \arg\min ||\boldsymbol{\gamma}||_2 \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{\gamma} = \boldsymbol{x}. \tag{15}$$

The above problem in (15), for the underdetermined linear system $\boldsymbol{A}\boldsymbol{\gamma} = \boldsymbol{x}$, has a closed-form solution $\hat{\boldsymbol{\gamma}}_2 = \boldsymbol{A}^\dagger\boldsymbol{x}$ with $\boldsymbol{A}^\dagger$ being the pseudo-inverse of $\boldsymbol{A}$. For a test sample $\boldsymbol{x}$ and training dictionary $\boldsymbol{A}$, the minimum ℓ_0-norm vector $\hat{\boldsymbol{\gamma}}$ and minimum ℓ_2-norm vector $\hat{\boldsymbol{\gamma}}_2$ are shown in Figs. 3(a) and (b), respectively. Blue and red represent entries corresponding to the background and target dictionaries, respectively. The original test sample $\boldsymbol{x}$ and the partial reconstructed pixels using only the background dictionary $\hat{\boldsymbol{x}}_b = \boldsymbol{A}_b\hat{\boldsymbol{\alpha}}$, $\hat{\boldsymbol{x}}_{b,2} = \boldsymbol{A}_b\hat{\boldsymbol{\alpha}}_2$ and only the target dictionary $\hat{\boldsymbol{x}}_t = \boldsymbol{A}_t\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{x}}_{t,2} = \boldsymbol{A}_t\hat{\boldsymbol{\beta}}_2$ are shown in Figs. 3(c) and (d). Although the pseudo-inverse solution $\hat{\boldsymbol{\gamma}}_2$ yields perfect reconstruction, we see that it is not sparse and its nonzero entries spread over both classes. Thus, $\hat{\boldsymbol{\gamma}}_2$ cannot be used directly for detection. The minimum ℓ_0-norm solution $\hat{\boldsymbol{\gamma}}$, on the contrary, has all of its nonzero entries concentrated in the background part, which indicates that the test sample lies in the background subspace. Furthermore, with the pseudo-inverse solution $\hat{\boldsymbol{\gamma}}_2$, as seen in Fig. 3(d), neither $\boldsymbol{A}_b\hat{\boldsymbol{\alpha}}_2$ nor $\boldsymbol{A}_t\hat{\boldsymbol{\beta}}_2$ accurately approximates the original pixel, leading to a small difference between the residuals $r_{b,2}(\boldsymbol{x}) = 0.568$ and $r_{t,2}(\boldsymbol{x}) = 0.516$. Hence, the solution $\hat{\boldsymbol{\gamma}}_2$ cannot be used to determine the class of the input solely based on the residuals. On the other hand, the residuals associated with the minimum ℓ_0-norm solution $\hat{\boldsymbol{\gamma}}$ are $r_b(\boldsymbol{x}) = 0.054 \ll r_t(\boldsymbol{x}) = 1$ (i.e., the original pixel $\boldsymbol{x}$ is well approximated by the background dictionary). Clearly, $\boldsymbol{x}$ is a background pixel using the minimum ℓ_0-norm solution.

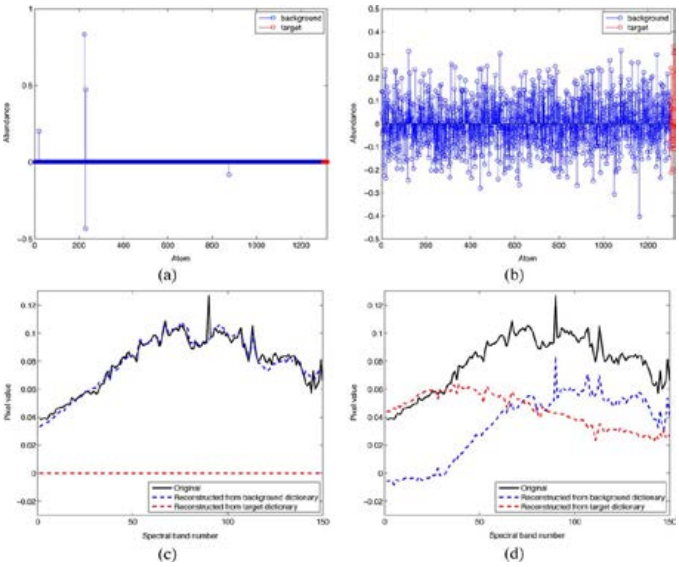


Fig. 3. Example of sparse reconstruction of a background test sample with a comparison to the minimum ℓ_2-norm (pseudoinverse) solution. (a) Minimum ℓ_0-norm solution $\hat{\boldsymbol{\gamma}}$. (b) Pseudo-inverse solution $\hat{\boldsymbol{\gamma}}_2 = \boldsymbol{A}^\dagger\boldsymbol{x}$. (c) Minimum ℓ_0-norm reconstruction from the background dictionary $\hat{\boldsymbol{x}}_b = \boldsymbol{A}_b\hat{\boldsymbol{\alpha}}$ (blue dashed), reconstruction from the target dictionary $\hat{\boldsymbol{x}}_t = A_t\hat{\boldsymbol{\beta}}$ (red dashed), and the original test sample $\boldsymbol{x}$ (black solid). (d) Pseudo-inverse reconstruction from the background dictionary $\hat{\boldsymbol{x}}_{b,2} = \boldsymbol{A}_b\hat{\boldsymbol{\alpha}}_2$ (blue dashed), reconstruction from the target dictionary $\hat{\boldsymbol{x}}_{t,2} = \boldsymbol{A}_t\hat{\boldsymbol{\beta}}_2$ (red dashed), and the original test sample $\boldsymbol{x}$ (black solid).

C. Background and Target Dictionary Construction

Another aspect of the problem that requires careful attention is how to construct appropriate dictionaries $\boldsymbol{A}_b$ and $\boldsymbol{A}_t$. Global dictionaries for target and background can be designed using given training data. However, in target detection applications there is usually a lack of training data especially for the target. The background is often modeled by a subspace by using some random pixels from the test image. Furthermore, a single target spectral signature, as employed in SMF, is often insufficient to represent a target material as the spectrum is affected by environmental conditions (e.g., illumination and atmospheric variations). By using physical models and the MORTRAN atmospheric-modeling program [34], meaningful target spectral signatures can be generated which can capture the target signature appearance over a wide range of atmospheric conditions. For example, in [21] a target subspace was constructed by generating a large number of target signatures using MORTRAN under various atmospheric conditions. A similar idea can be incorporated in our approach to construct a redundant target dictionary which could be invariant to the environmental variations. Furthermore, it can be combined with the idea of frame generation [35], [36] by imposing the constraints on tightness, maximum robustness, equiangularity, etc., to design more desirable overcomplete dictionaries. The K-SVD dictionary design technique [37], which alternately minimizes sparsity of the representation and updates the codebook to better fit the data, can also be used to form the redundant dictionaries to further improve the performance of the proposed sparsity-based algorithm.

In this paper, we use a small global target dictionary constructed by using some of the target pixels on one of the targets in the scene. For the background dictionary, instead of using

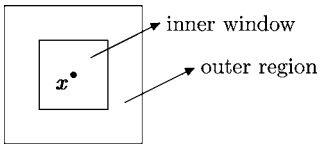


Fig. 4. Dual window centered at test sample $\boldsymbol{x}$.

a fixed global background dictionary containing samples from various background materials (e.g., trees, grass, road, buildings, etc.), we use an adaptive local background dictionary in order to better represent and capture the spectral signature of test sample. Specifically, the background dictionary $\boldsymbol{A}_b$ is generated locally for each test pixel using a dual window centered at the pixel of interest, as shown in Fig. 4. The inner window should be larger than the size of a target. Only pixels in the outer region form the atoms in $\boldsymbol{A}_b$. In this way, the subspace spanned by the background dictionary becomes adaptive to the local statistics. Therefore, if the test sample is a background pixel, it is highly likely that it finds very similar spectral characteristic in the background dictionary. On the other hand, if the test sample is a target pixel, it would be difficult for the pixel to find a good match in $\boldsymbol{A}_b$ since the outer window region does not include any target pixels. The usage of a dual window significantly improves the detection performance over a global background dictionary, as is shown via the simulation results in Section IV.

D. Detection With Smoothing Constraint

In the above process, the sparsity-based target detector is applied to each pixel in the test region independently without considering the correlation between neighboring pixels. Hyperspectral imagery, however, is usually smooth in the sense that neighboring pixels usually consist of similar materials and have similar spectral characteristics where small differences are often due to sensor noise and/or atmospheric variation. In this paper, we assume that there are multiple pixels on the target. Therefore, we propose to incorporate a smoothing penalty term in the proposed sparsity-based detector in order to exploit the spatial correlation between neighboring pixels.

Let $\boldsymbol{x}_1$ be a pixel of interest in a hyperspectral image $\boldsymbol{I}$, and $\boldsymbol{x}_i, i = 2, \ldots, 5$ be its four nearest neighbors in the spatial domain, as shown in Fig. 5. While searching for the sparsest representation of the test sample $\boldsymbol{x}_1$, we simultaneously minimize the vector Laplacian at the reconstructed pixel $\hat{\boldsymbol{x}}_1$, which is a B-dimensional vector calculated as

$$\begin{aligned}\nabla^2(\hat{\boldsymbol{x}}_1) &= 4\hat{\boldsymbol{x}}_1 - \hat{\boldsymbol{x}}_2 - \hat{\boldsymbol{x}}_3 - \hat{\boldsymbol{x}}_4 - \hat{\boldsymbol{x}}_5 \\ &= \boldsymbol{A}(4\hat{\boldsymbol{\gamma}}_1 - \hat{\boldsymbol{\gamma}}_2 - \hat{\boldsymbol{\gamma}}_3 - \hat{\boldsymbol{\gamma}}_4 - \hat{\boldsymbol{\gamma}}_5) \end{aligned} \tag{16}$$

where $\hat{\boldsymbol{x}}_i = \boldsymbol{A}\hat{\boldsymbol{\gamma}}_i$ is the reconstruction of $\boldsymbol{x}_i$ and $\hat{\boldsymbol{\gamma}}_i$ is the corresponding recovered sparse vector. In this way, the reconstructed test sample is forced to have a similar spectral characteristics as its four nearest neighbors; hence, smoothness is enforced across the spectral pixels in the reconstructed image.

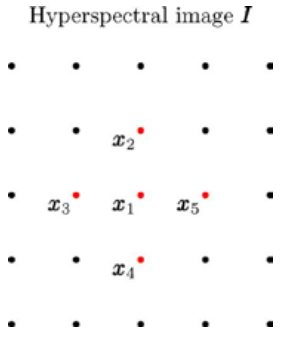


Fig. 5. Four nearest neighbors of a pixel $\boldsymbol{x}_1$.

Let $\boldsymbol{\gamma}_i$ be the sparse vector associated with $\boldsymbol{x}_i$ (i.e., $\boldsymbol{x}_i = \boldsymbol{A}\boldsymbol{\gamma}_i$). The new problem with the smoothing constraint can now be formulated as

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{5} \|\boldsymbol{\gamma}_i\|_0 \\ \text{subject to:} \quad & \boldsymbol{A}(4\boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_2 - \boldsymbol{\gamma}_3 - \boldsymbol{\gamma}_4 - \boldsymbol{\gamma}_5) = \boldsymbol{0} \\ & \boldsymbol{x}_i = \boldsymbol{A}\boldsymbol{\gamma}_i, \quad i = 1, \ldots, 5. \end{aligned} \tag{17}$$

In (17), we aim to find the sparsest vector that approximately satisfies two sets of linear constraints. The first set forces the vector Laplacian of the reconstructed pixel $\hat{\boldsymbol{x}}_1$ to be minimal such that the reconstructed neighboring pixels have similar spectral characteristics, and the second set minimizes reconstruction errors. Now denote the concatenation of $\boldsymbol{\gamma}_i$'s and $\boldsymbol{x}_i$'s by

$$\boldsymbol{\gamma} = \begin{bmatrix} \boldsymbol{\gamma}_1 \\ \vdots \\ \boldsymbol{\gamma}_5 \end{bmatrix} \quad \text{and} \quad \boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_5 \end{bmatrix}. \tag{18}$$

The linear constraints can be written in terms of $\boldsymbol{x}$ and $\boldsymbol{\gamma}$ as

$$\begin{aligned} [4\boldsymbol{A} \quad -\boldsymbol{A} \quad -\boldsymbol{A} \quad -\boldsymbol{A} \quad -\boldsymbol{A}]\boldsymbol{\gamma} &= \boldsymbol{0}, \\ \begin{bmatrix} \boldsymbol{A} & & \boldsymbol{0} \\ & \ddots & \\ \boldsymbol{0} & & \boldsymbol{A} \end{bmatrix} \boldsymbol{\gamma} &= \boldsymbol{x}. \end{aligned} \tag{19}$$

Therefore, the optimization problem in (17) can be reformulated as

$$\begin{aligned} \text{minimize} \quad & \|\boldsymbol{\gamma}\|_0 \\ \text{subject to:} \quad & \tilde{\boldsymbol{A}}\boldsymbol{\gamma} = \tilde{\boldsymbol{x}}, \end{aligned} \tag{20}$$

where

$$\tilde{\boldsymbol{A}} = \begin{bmatrix} 4\boldsymbol{A} & -\boldsymbol{A} & -\boldsymbol{A} & -\boldsymbol{A} & -\boldsymbol{A} \\ \boldsymbol{A} & & & & \\ & \boldsymbol{A} & & & \\ & & \boldsymbol{A} & & \\ & & & \boldsymbol{A} & \\ & & & & \boldsymbol{A} \end{bmatrix} \quad \text{and} \quad \tilde{\boldsymbol{x}} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{x} \end{bmatrix}.$$

The problem in (20) is the standard form of a linearly constrained sparsity-minimization problem and can be solved using the greedy SP algorithm [31]. Similar to the previous case in

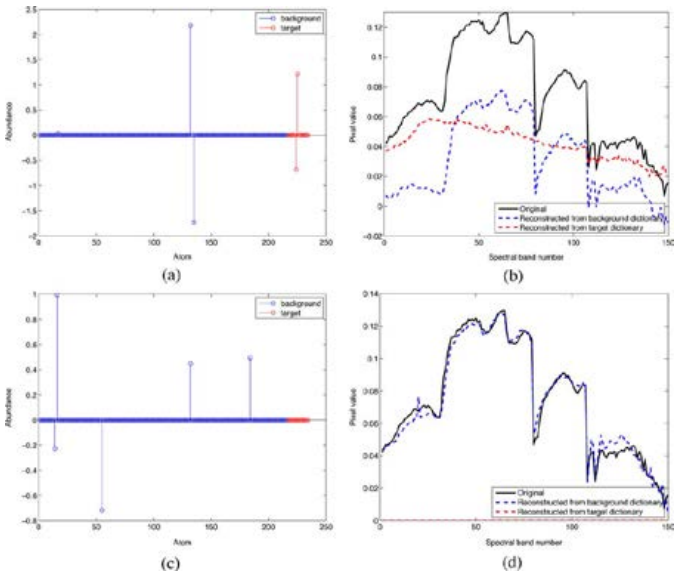


Fig. 6. Example comparing the reconstruction and detection problem for a background test sample without and with the smoothing constraint. (a) Solution to (11) (without smoothing constraint). (b) By solving (11), the reconstruction from the background dictionary (blue dashed), reconstruction from the target dictionary (red dashed), and the original test sample (black solid). (c) Solution to (20) (with smoothing constraint) for the centered test sample. (d) By solving (20), the reconstruction from the background dictionary (blue dashed), reconstruction from the target dictionary (red dashed), and the original test sample (black solid).

(11), this problem can also be relaxed to allow for approximation errors in empirical data and be rewritten as

$$\hat{\boldsymbol{\gamma}} = \arg\min ||\boldsymbol{\gamma}||_0 \quad \text{subject to} \quad ||\tilde{\boldsymbol{A}}\boldsymbol{\gamma} - \tilde{\boldsymbol{x}}||_2 \leq \sigma \tag{21}$$

or

$$\hat{\boldsymbol{\gamma}} = \arg\min ||\tilde{\boldsymbol{A}}\boldsymbol{\gamma} - \tilde{\boldsymbol{x}}||_2 \quad \text{subject to} \quad ||\boldsymbol{\gamma}||_0 \leq K_0 \tag{22}$$

where σ is the error tolerance and K_0 is the sparsity level.

By exploiting the smoothness across the HSI pixels, the detection performance can be significantly improved. Fig. 6 shows an example of a background test sample which is misclassified as a target using (11), but is correctly labeled using (20) with the smoothing constraint. The solution to (11) for the given test sample $\boldsymbol{x}$ is depicted in Fig. 6(a). We see that the nonzero entries of the solution correspond to both background and target training atoms, and the residuals are $r_b = 0.5381, r_t = 0.5088$. In the case with the smoothing constraint, by solving (20), the nonzero entries only concentrate on part corresponding to the background dictionary, and the residuals are $r_b = 0.0538, r_t = 1$. Clearly, $r_b \ll r_t$ and the test sample will thus be correctly labeled as a background sample.

Once the sparse vector in (20) is obtained, detection can be performed based on the characteristics of the sparse coefficients as it was done in Section III-B. We calculate the total residuals obtained separately from the target and background dictionaries

$$r_b(\boldsymbol{x}_1) = \sqrt{\sum_{i=1}^{5} ||\boldsymbol{x}_i - \boldsymbol{A}_b\hat{\boldsymbol{\alpha}}_i||_2^2}$$

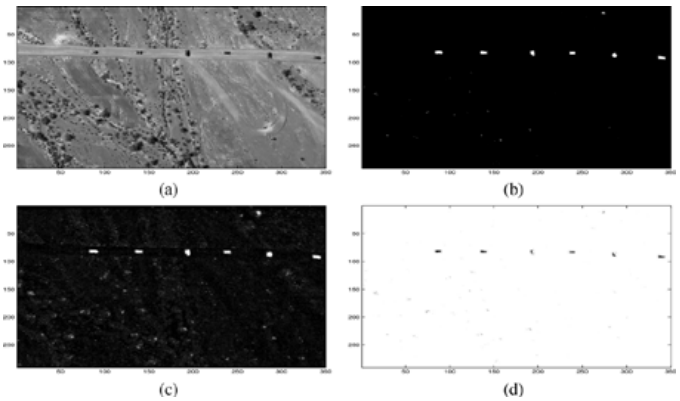


Fig. 7. Results for Desert Radiance II (DR-II) from (20) with the smoothing constraint. (a) Averaged image over 150 bands. (b) Sparsity-based target detector output: difference between r_b and r_t. (c) Residual r_b corresponding to the local background dictionary using the dual-window approach. (d) Residual r_t corresponding to the target dictionary.

and

$$r_t(\boldsymbol{x}_1) = \sqrt{\sum_{i=1}^{5} ||\boldsymbol{x}_i - \boldsymbol{A}_t\hat{\boldsymbol{\beta}}_i||_2^2} \tag{23}$$

where $\hat{\boldsymbol{\alpha}}_i$ and $\hat{\boldsymbol{\beta}}_i$ denote the recovered sparse coefficients for $\boldsymbol{x}_i$ associated with the background and target dictionaries, respectively. The output of the proposed sparsity-based detector for the center pixel $\boldsymbol{x}_1$ is computed by the difference of residuals and the detection decision is made in a similar fashion as in the other algorithms introduced in Section II:

$$D(\boldsymbol{x}_1) = r_b(\boldsymbol{x}_1) - r_t(\boldsymbol{x}_1) \underset{H_0}{\overset{H_1}{\gtrless}} \delta. \tag{24}$$

That is, if the output $D(\boldsymbol{x}_1)$ is greater than a prescribed threshold δ, then the test sample $\boldsymbol{x}_1$ is labeled as a target; otherwise it is labeled as background.

IV. Simulation Results and Analysis

The proposed target detection algorithm, as well as the SMF, MSD, ASD, and SVM, are applied to several real HSI, and the results are compared both visually and quantitatively by the receiver operating characteristics (ROC) curves. The ROC curve describes the probability of detection (PD) as a function of the probability of false alarms (PFA). To be more specific, we pick thousands of different thresholds between the minimal and maximal values of the detector output. The class labels for all pixels in the test region are determined at each threshold. The PFA is calculated by the number of false alarms (background pixels determined as target) over the total number of pixels in the test region, and the PD is the ratio of the number of hits (target pixels determined as target) and the total number of true target pixels.

Two of the images, the desert radiance II data collection (DR-II) and forest radiance I data collection (FR-I), are from a hyperspectral digital imagery collection experiment (HYDICE) sensor [38]. The HYDICE sensor generates 210 bands across the whole spectral range from 0.4 to 2.5 μm which includes the visible and short-wave infrared bands. We use 150 of the 210 bands (23rd-101st, 109th-136th, and 152nd-194th), removing

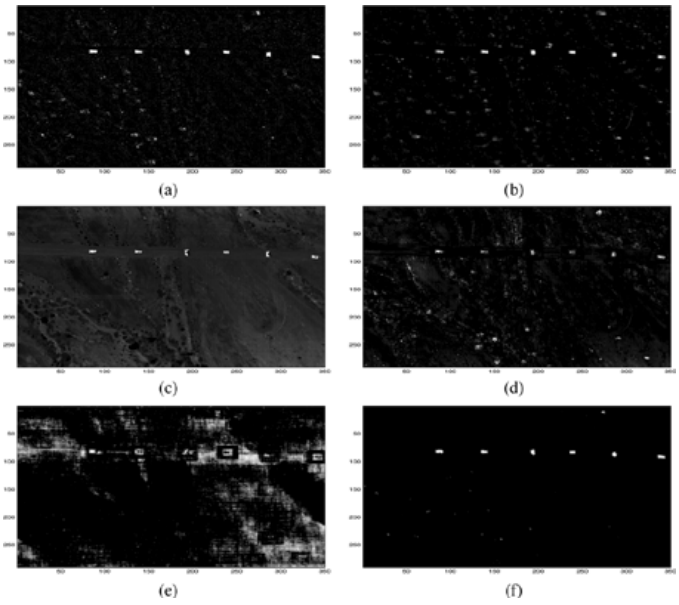


Fig. 8. Output for DR-II using local background dictionary (dual-window approach), with (a) sparsity-based target detector without smoothing constraint using (11), (b) SVM with composite kernel, (c) MSD, (d) SMF, and (e) ASD. (f) We repeat here the result of our proposed sparsity-based target detector with smoothing constraint for visual comparison.

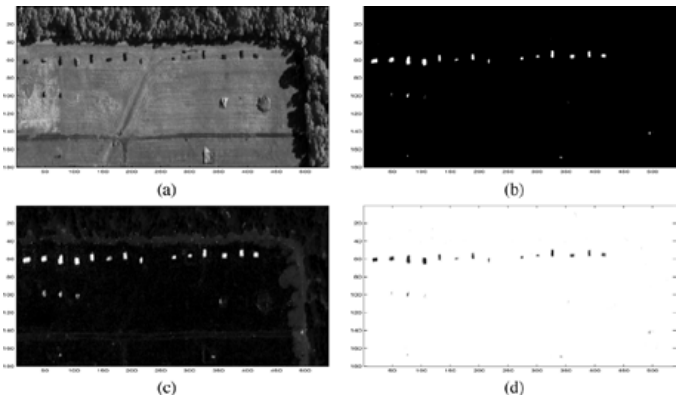


Fig. 9. Results for forest radiance I (FR-I) from (20) with smoothing constraint. (a) Averaged image over 150 bands. (b) Sparsity-based target detector output: difference between r_b and r_t. (c) Residual r_b corresponding to the background dictionary (dual-window approach). (d) Residual r_t corresponding to the target dictionary.

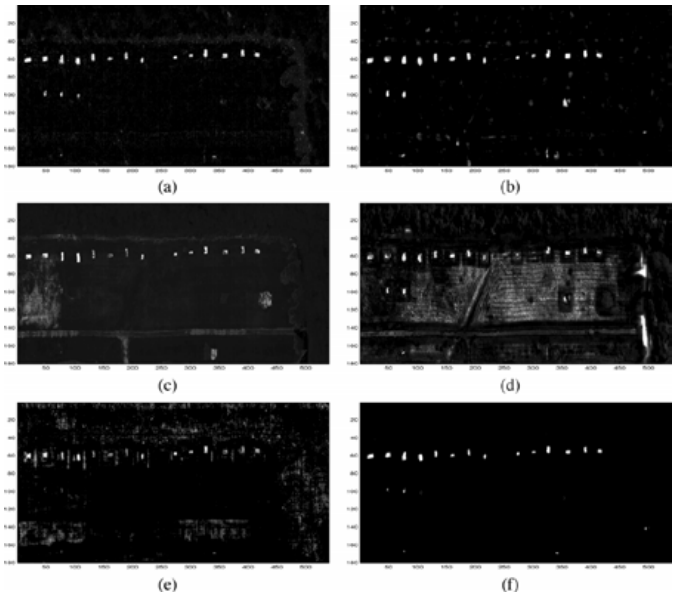


Fig. 10. Output for FR-I using local background dictionary (dual-window approach), with (a) sparsity-based target detector without smoothing constraint using (11), (b) SVM with composite kernel, (c) MSD, (d) SMF, and (e) ASD. (f) We repeat here the result of our proposed sparsity-based target detector with smoothing constraint for visual comparison.

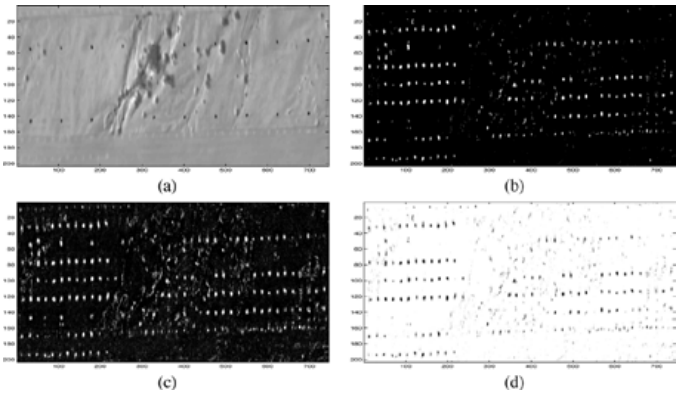


Fig. 11. Results for the mine image from (20) with smoothing constraint. (a) Averaged image over 70 bands. (b) Detector output: difference between r_b and r_t. (c) Residual r_b corresponding to the background dictionary (dual-window approach). (d) Residual r_t corresponding to the target dictionary.

the absorption and low-SNR bands. The DR-II image contains six military target on the dirt road and the FR-I image contains 14 targets along the tree line as depicted in Figs. 7(a) and 9(a), respectively. For these two HYDICE images, every pixel on the targets is considered a target pixel. The third image, collected from the Airborne Hyperspectral Imager (AHI) [39] operating in the long-wave infrared spectrum ranging from 8 to 11.5 μm, contains surface and buried mines as shown in Fig. 11(a), in which every pixel has 70 spectral bands. In this image, there are about 230 mines, each roughly of size 5×5 pixels and each mine is treated as a target when computing the PD.

For DR-II and FR-I, the spectral signatures of the target $\{\boldsymbol{a}_i^t\}_{i=1,\ldots,N_t}$ are collected directly from $N_t = 18$ pixels from the leftmost target in the given hyperspectral data. The background signatures $\{\boldsymbol{a}_i^b\}_{i=1,\ldots,N_b}$ are generated by the pixels in the outer region of a dual window as discussed in Section III. The size of the outer and inner windows are 21×21 and 15×15, respectively, and there are $N_b = 216$ background training samples. The subspace pursuit algorithm [31] is used to solve the sparsity-constrained problems (11) and (20). The results of the proposed detector with the smoothing constraint for DR-II are shown in Fig. 7(b)–(d). Fig. 7(c) and (d) shows the residuals corresponding to the background dictionary $r_b(\boldsymbol{x}) = ||\boldsymbol{x} - \boldsymbol{A}_b\hat{\boldsymbol{\alpha}}||_2$, and the residual corresponding to the target dictionary $r_t(\boldsymbol{x}) = ||\boldsymbol{x} - \boldsymbol{A}_t\hat{\boldsymbol{\beta}}||_2$, respectively, whereas Fig. 7(b) shows the difference between r_b and r_t. In Fig. 7(c), while background pixels are dark, the target pixels are bright due to the fact that for each target pixel the sparsity-constrained optimizer could not find good matches from the background dictionary; therefore, the sparse vector $\hat{\boldsymbol{\alpha}} \approx \boldsymbol{0}$ and the residual associated with the background dictionary is $r_b(\boldsymbol{x}) \approx ||\boldsymbol{x}||$. On the contrary, in Fig. 7(d), the targets are dark while the background are bright. Finally, as shown in Fig. 7(b), the difference between r_b and r_t will further suppress the background and

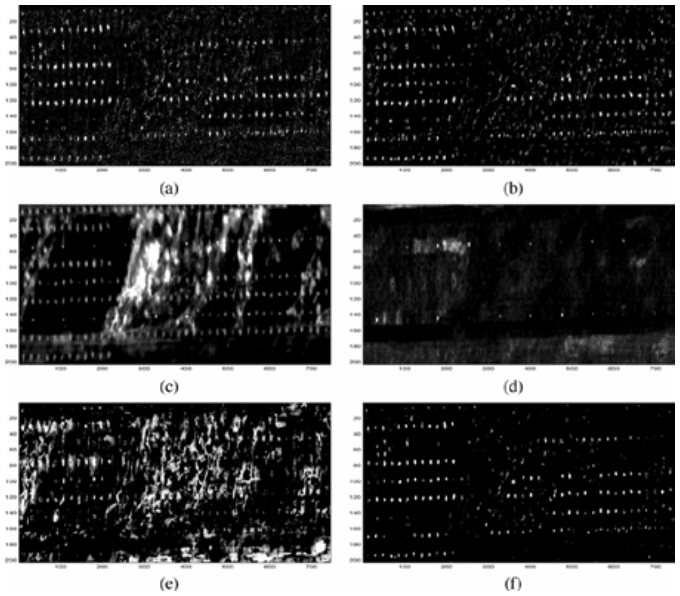


Fig. 12. Output for the mine image using local background dictionary (dual-window approach), with (a) sparsity-based target detector without smoothing constraint, (b) SVM with composite kernel, (c) MSD, (d) SMF, and (e) ASD. (f) We repeat here the result of our proposed sparsity-based target detector with smoothing constraint for visual comparison.

emphasizes the targets, thus yielding better detection performance. Similar results can be seen in Fig. 9(b)–(d) for the FR-I image. In Fig. 9(c) which represents the residual image r_b, although the targets are bright, we can also see the shadow of trees near the upper and right borders of the image has higher magnitude than the other background areas. In Fig. 9(b), the shadow is suppressed and this improves the false alarm rate.

Similar results can be observed in Fig. 11 for the mine image, where the target dictionary $\boldsymbol{A}_t$ is generated from $N_t = 50$ training samples of two mines, each occupying a 5×5 area, outside the test region. Since the targets in this image are smaller than that of the two HYDICE images, the inner window size is chosen to be 9×9 and the outer window size remains 21×21. The background dictionary $\boldsymbol{A}_b$ then consists of $N_b = 360$ samples.

Next we demonstrate the importance of employing a locally adaptive background dictionary. The sparsity-based target detection algorithm is applied to the DR-II and FR-I images using local and global background dictionaries. The local $\boldsymbol{A}_b$ is generated by $N_b = 216$ pixels in the outer region of the dual window centered at the test sample as in Fig. 4, and the global dictionary ($N_b = 1300$ for DR-II and $N_b = 1656$ for FR-I) is generated by randomly collecting background pixels, which can be reduced to a smaller size by an unsupervised clustering algorithm such as K-means. The detection performance is significantly improved by using a local dictionary, as seen in the ROC curves shown in Fig. 13. This is because a fixed global dictionary fails to capture the local similarity between pixels in a small neighborhood. A local dictionary exploits the local statistics and helps to find better resemblance of test samples. We see in Fig. 13 that the detector using local dictionaries outperforms the one using global dictionaries by a large margin for both HYDICE images.

Under the same settings (i.e., same target and background training samples for all detectors), we compare the performance

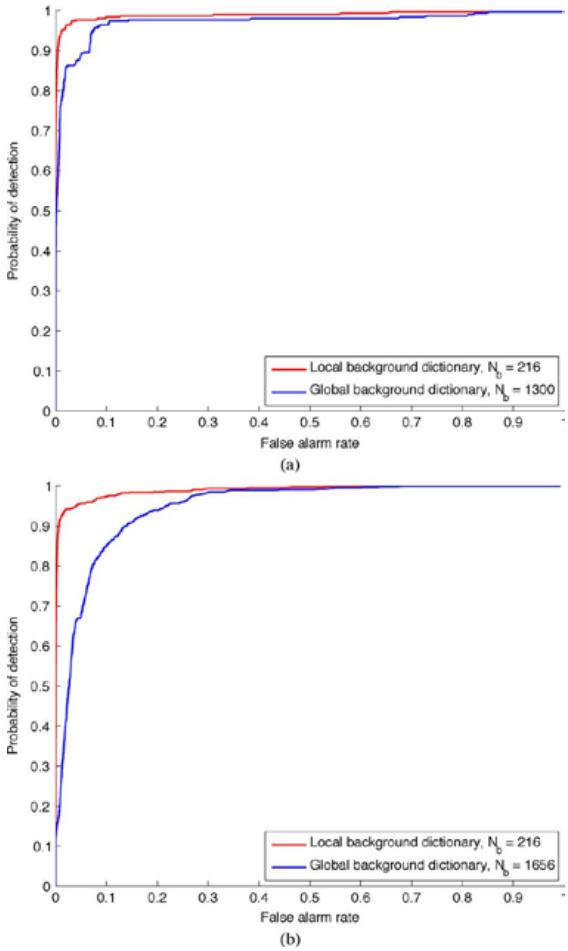


Fig. 13. ROC curves using the sparsity-based target detector with smoothing constraint for (a) DR-II and (b) FR-I with local and global background dictionaries.

of the proposed sparsity-based algorithm to the previously developed conventional classifier SVM and detectors MSD, SMF, ASD using both global and local background dictionaries. Let $\boldsymbol{A}_t = [\boldsymbol{a}_1^t \quad \cdots \quad \boldsymbol{a}_{N_t}^t]$ and $\boldsymbol{A}_b = [\boldsymbol{a}_1^b \quad \cdots \quad \boldsymbol{a}_{N_b}^b]$ be, respectively, the target and background dictionaries used in the proposed sparsity-based algorithm. Note that in the local case, $\boldsymbol{A}_b$ is adaptive and changes for each test pixel. In order to have a fair comparison, in the case of SMF the target signature is the mean of the target dictionary atoms $\{\boldsymbol{a}_i^t\}_{i=1,\ldots,N_t}$ and the background covariance is obtained from the background dictionary $\boldsymbol{A}_b$. In the SMF implementation, a regularization term is added to the background covariance matrix such that the inverse matrix $\hat{\boldsymbol{C}}^{-1}$ in (3) is more stable, as described in [27]. In the case of MSD, the eigenvectors corresponding to the significant eigenvalues of the covariance matrices obtained from atoms in $\boldsymbol{A}_t$ and $\boldsymbol{A}_b$ are used to generate the basis for the target and background subspaces, respectively [23]. For ASD, the basis for target subspace are generated in the same way as in MSD. The ASD noise covariance matrix is computed from the background training samples $\{\boldsymbol{a}_i^b\}_{i=1,\ldots,N_b}$ and a regularization term is added to the noise covariance matrix in order to obtain a stable inverse matrix. In SVM, a model is trained using atoms in $\boldsymbol{A}_b$ and $\boldsymbol{A}_t$ as two different classes using a composite kernel which combines the spectral and spatial feature via a weighted summation, where K_s

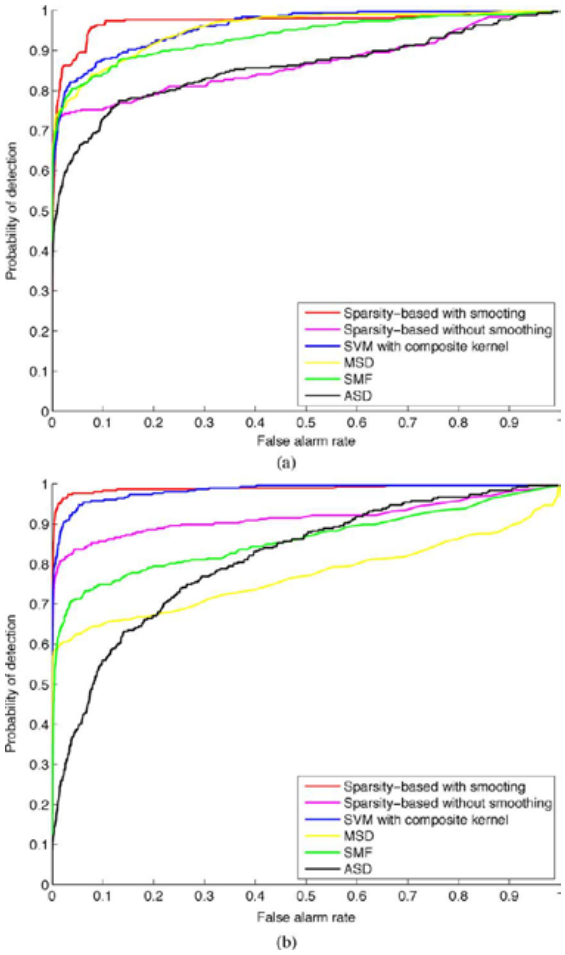


Fig. 14. ROC curves for DR-II. (a) Global background dictionary, $N_b = 1300$. (b) local background dictionary (dual-window approach), $N_b = 216$.

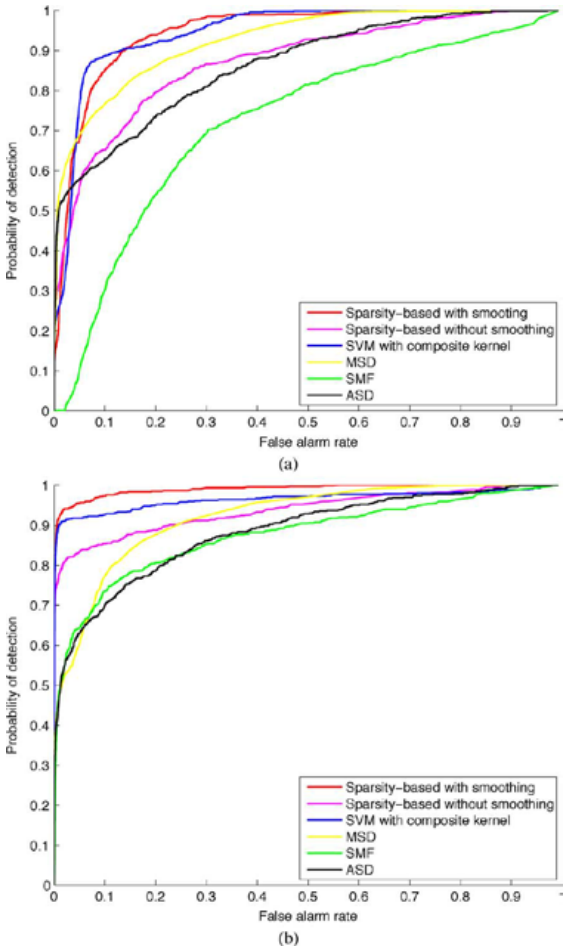


Fig. 15. ROC curves for FR-I. (a) Global background dictionary, $N_b = 1656$. (b) local background dictionary (dual-window approach), $N_b = 216$.

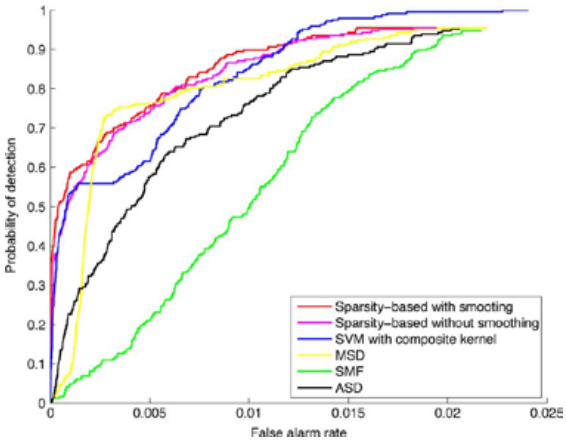


Fig. 16. ROC curves for the mine image using local background dictionary (dual-window approach), $N_b = 360$.

and K_w in (1) are radial basis function kernels [9]. All parameters are adjusted to obtain the best possible performance. Under the current setting of target and background dictionaries, the proposed detector has computational complexity comparable to that of the classical detectors SMF, MSD, and ASD.

The ROC curves in both the global and local cases for DR-II are shown in Fig. 14. We see that the sparsity-based detector with the smoothing term using a local background dictionary outperforms all other detectors. The SMF performs poorly since the target signature is represented by a single vector, while in all other approaches the targets are assumed to approximately lie in a subspace. For visual comparison, the detector outputs for SVM, MSD, SMF, and ASD are also displayed in Fig. 8, where the locally adaptive background dictionary is employed. One can immediately notice that the sparsity-based detector with the smoothing constraint also leads to the best visual quality.

The ROC curves for FR-I are shown in Fig. 15. The FR-I image is more difficult than the DR-II due to the presence of the trees and shadow whose spectral curves have some resemblance to that of the targets. From the ROC plots, the proposed algorithm still leads to the best performance. For visual inspection, the detection results obtained by SVM, MSD, SMF, and ASD are illustrated in Fig. 10. For all detectors in Fig. 10, we can see the bright spots in the shadow area along the tree line. This is alleviated by the proposed detection algorithm, as seen in Fig. 9(b).

The AHI image of mines is the most difficult one among the three test images. The targets include surface mines and buried mines that are invisible. In this case, the ROC curve is obtained slightly differently in that only one pixel on the mine needs to be correctly labeled for the mine to be declared as a target. Therefore, the PD is calculated by the number of hits divided

by the total number of mines in the test region. In this experiment, for all detectors, we use $N_t = 50$ target training samples from two mines outside the test region and $N_b = 360$ background training samples adaptively constructed for each test pixel by the dual-window approach with inner and outer windows of size 9×9 and 21×21, respectively. The ROC curves for the mine image using local dictionaries are shown in Fig. 16. The proposed sparsity-based target detection algorithm still outperforms the other algorithms, especially at low PFA. The outputs for SVM, MSD, SMF, and ASD are displayed as images in Fig. 12. We see that although the MSD yields higher PD at certain PFA, there is a large background area in the middle of the image where pixels have very high magnitude, hence increasing the number of false alarms.

V. Conclusion

In this paper, we propose a target detection algorithm for hyperspectral imagery based on sparse representation of the test samples. In the proposed algorithm, the sparse representation is recovered by solving a constrained optimization problem that simultaneously addresses the sparsity constraint, reconstruction accuracy, and a smoothness penalty on the reconstructed image. Detection decision is obtained from the recovered sparse vectors by reconstruction. The new algorithm consistently outperforms the previously developed detectors in terms of both qualitative and quantitative measures, as demonstrated by experimental results in several real hyperspectral images. Future research includes the construction of better dictionaries. For example, the proposed detector can be improved by generating dictionaries invariant to the effect of atmospheric absorption [21]. We will also investigate the design and exploitation of more discriminative dictionaries learned from the training data [37], [40].

References

[1] M. Borengasser, W. S. Hungate, and R. Watkins, *Hyperspectral Remote Sensing—Principles and Applications*. Boca Raton, FL: CRC, 2008.

[2] D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 29–43, Jan. 2002.

[3] D. W. J. Stein, S. G. Beaven, L. E. Hoff, E. M. Winter, A. P. Schaum, and A. D. Stocker, "Anomaly detection from hyperspectral imagery," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 58–69, Jan. 2002.

[4] M. T. Eismann, A. D. Stocker, and N. M. Nasrabadi, "Automated hyperspectral cueing for civilian search and rescue," *Proc. IEEE*, vol. 97, no. 6, pp. 1031–1055, Jun. 2009.

[5] N. K. Patel, C. Patnaik, S. Dutta, A. M. Shekh, and A. J. Dave, "Study of crop growth parameters using airborne imaging spectrometer data," *Int. J. Remote Sens.*, vol. 22, no. 12, pp. 2401–2411, Aug. 2001.

[6] B. Datt, T. R. McVicar, T. G. Van Niel, D. L. B. Jupp, and J. S. Pearlman, "Preprocessing EO-1 hyperion hyperspectral data to support the application of agricultural indexes," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 6, pp. 1246–1259, Jun. 2003.

[7] B. Hörig, F. Kühn, F. Oschütz, and F. Lehmann, "HyMap hyperspectral remote sensing to detect hydrocarbons," *Int. J. Remote Sens.*, vol. 22, no. 8, pp. 1413–1422, 2001.

[8] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proc. 5th Annu. Workshop Comput. Learn. Theory*, 1992, pp. 144–152.

[9] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York: Springer, 1995.

[10] J. A. Gualtieri and R. F. Cromp, "Support vector machines for hyperspectral remote sensing classification," in *Proc. SPIE*, Jan. 1998, vol. 3584, pp. 221–232.

[11] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1778–1790, Aug. 2004.

[12] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 1, pp. 208–216, Jan. 1992.

[13] D. Manolakis, G. Shaw, and N. Keshava, "Comparative analysis of hyperspectral adaptive matched filter detectors," in *Proc. SPIE Conf. Algorithms for Multispectral, Hyperspectral, Ultraspectral Imagery VI*, Apr. 2000, vol. 4049, pp. 2–17.

[14] L. L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Trans. Signal Process.*, vol. 42, no. 8, pp. 2146–2157, Aug. 1994.

[15] S. Kraut, L. L. Scharf, and L. T. McWhorter, "Adaptive subspace detectors," *IEEE Trans. Signal Process.*, vol. 49, no. 1, pp. 1–16, Jan. 2001.

[16] J. Wright, A. Y. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.

[17] R. Basri and D. W. Jacobs, "Lambertian reflectance and linear subspaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 2, pp. 218–233, Feb. 2003.

[18] J. K. Pillai, V. M. Patel, and R. Chellappa, "Sparsity inspired selection and recognition of iris images," in *Proc. IEEE 3rd Int. Conf. Biometrics: Theory, Applicat., Syst.*, Sep. 2009, pp. 1–6.

[19] X. Hang and F.-X. Wu, "Sparse representation for classification of tumors using gene expression data," *J. Biomed. Biotechnol.*, vol. 2009, 2009, doi:10.1155/2009/403689.

[20] Z. Guo, T. Wittman, and S. Osher, "L1 unmixing and its application to hyperspectral image enhancement," in *Proc. SPIE Conf. Algorithms Technol. Multispectral, Hyperspectral, Ultraspectral Imagery XV*, Apr. 2009, vol. 7334.

[21] G. Healey and D. Slater, "Models and methods for automated material identification in hyperspectral imagery acquired under unknown illumination and atmospheric conditions," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 6, pp. 2706–2717, Jun. 1999.

[22] B. Thai and G. Healey, "Invariant subpixel material detection in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 3, pp. 599–608, Mar. 2002.

[23] H. Kwon and N. M. Nasrabadi, "A comparative analysis of kernel subspace target detectors for hyperspectral imagery," *EURASIP J. Appl. Signal Process.*, vol. 2007, no. 1, pp. 193–193, Jan. 2007.

[24] G. Camps-Valls, L. Gomez-Chova, J. Muñoz-Marí, J. Vila-Francés, and J. Calpe-Maravilla, "Composite kernels for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 93–97, Jan. 2006.

[25] A. Plaza, J. A. Benediktsson, J. W. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gualtieri, M. Marconcini, J. C. Tilton, and G. Trianni, "Recent advances in techniques for hyperspectral image processing," *Remote Sens. Environ.*, vol. 113, pp. S110–S122, Sep. 2009, Supplement 1.

[26] B. Schlkopf and A. J. Smola, *Learning With Kernels: Support Vector Machines, Regularization, Optimization, and Beyond (Adaptive Computation and Machine Learning)*. Cambridge, MA: MIT Press, 2001.

[27] N. M. Nasrabadi, "Regularized spectral matched filter for target recognition in hyperspectral imagery," *IEEE Signal Process. Lett.*, vol. 15, pp. 317–320, 2008.

[28] D. L. Donoho, "For most large underdetermined systems of linear equations the minimal l_1-norm solution is also the sparsest solution," *Comm. Pure Appl. Math.*, vol. 59, no. 6, pp. 797–829, 2006.

[29] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Rev.*, vol. 51, no. 1, pp. 34–81, 2009.

[30] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[31] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2230–2249, May 2009.

[32] J. A. Tropp and S. J. Wright, "Computational methods for sparse solution of linear inverse problems," *Proc. IEEE*, vol. 98, no. 6, pp. 948–958, Jun. 2010.

[33] E. van den Berg and M. P. Friedlander, "Probing the Pareto frontier for basis pursuit solutions," *SIAM J. Sci. Comput.*, vol. 31, no. 2, pp. 890–912, Nov. 2008.

[34] A. Berk, L. Bernstein, and D. Robertson, Geophysics Laboratory, Bedford, MA, "MORTRAN: A moderate resolution model for LOWTRAN 7," Tech. Rep. GL-TR-90-0122, 1989.

[35] J. Kovacevic and A. Chebira, "Life beyond bases: The advent of frames (Part I)," *IEEE Signal Process. Mag.*, vol. 24, no. 4, pp. 86–104, Jul. 2007.

[36] J. Kovacevic and A. Chebira, "Life beyond bases: The advent of frames (Part II)," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 115–125, Sep. 2007.

[37] M. Aharon, M. Elad, and A. M. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

[38] R. W. Basedow, D. C. Carmer, and M. E. Anderson, "HYDICE system: Implementation and performance," in *Proc. SPIE Conf. Algorithms Technol. Multispectral, Hyperspectral, and Ultraspectral Imagery XV*, Apr. 1995, vol. 2480, pp. 258–267.

[39] P. G. Lucey, T. J. Williams, M. Mignard, J. Julian, D. Kobubun, G. Allen, D. Hampton, W. Schaff, M. J. Schlangenand, E. M. Winter, W. B. Kendall, A. D. Stocker, K. A. Horton, and A. P. Bowman, "AHI: An airborne long-wave infrared hyperspectral imager," in *Proc. SPIE Conf. Airborne Reconnaissance XXII*, 1998, vol. 3431, pp. 36–43.

[40] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17, no. 1, pp. 53–69, Jan. 2008.

Yi Chen received the B.Eng. degree in electronic engineering from Tsinghua University, Beijing, China, in 2002 and the M.A.Sc. degree in electrical engineering from the University of Victoria, Victoria, BC, Canada, in 2006. She is currently pursuing the Ph.D. degree in the Department of Electrical and Computer Engineering, The Johns Hopkins University, Baltimore, MD.

Her research interests include the applications of compressed sensing and sparse representations, image processing, wavelets, multirate systems, and filter banks.

Nasser M. Nasrabadi (S'80–M'84–SM'92–F'01) received the B.Sc. (Eng.) and Ph.D. degrees in electrical engineering from Imperial College of Science and Technology , University of London, London, U.K., in 1980 and 1984, respectively.

From October 1984 to December 1984, he was with IBM (U.K.) as a Senior Programmer. From 1985 to 1986, he was with Philips Research Laboratory, Briarcliff Manor, NY, as a Member of Technical Staff. From 1986 to 1991, he was an Assistant Professor in the Department of Electrical Engineering, Worcester Polytechnic Institute, Worcester, MA. From 1991 to 1996, he was an Associate Professor with the Department of Electrical and Computer Engineering, State University of New York at Buffalo, Buffalo, NY. Since September 1996, he has been a Senior Research Scientist (ST) with the U.S. Army Research Laboratory (ARL), Adelphi, MD, working on image processing and automatic target recognition.

Dr. Nasrabadi has served as an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING, the IEEE TRANSACTIONS ON CIRCUITS, SYSTEMS AND VIDEO TECHNOLOGY, and the IEEE TRANSACTIONS ON NEURAL NETWORKS. He is also a Fellow of ARL and SPIE. His current research interests are in Hyperspectral imaging, automatic target recognition, statistical machine learning theory, robotics, and neural networks applications to image processing.

Trac D. Tran (S'94–M'98–SM'08) received the B.S. and M.S. degrees from the Massachusetts Institute of Technology, Cambridge, in 1993 and 1994, respectively, and the Ph.D. degree from the University of Wisconsin, Madison, in 1998, all in electrical engineering.

In July of 1998, he joined the Department of Electrical and Computer Engineering, The Johns Hopkins University, Baltimore, MD, where he currently holds the rank of Associate Professor. His research interests are in the field of digital signal processing, particularly in sparse representation, sparse recovery, sampling, multi-rate systems, filter banks, transforms, wavelets, and their applications in signal analysis, compression, processing, and communications. His pioneering research on integer-coefficient transforms and pre-/post-filtering operators has been adopted as critical components of Microsoft Windows Media Video 9 and JPEG XR C the latest international still-image compression standard ISO/IEC 29199-2. In the summer of 2002, he was an ASEE/ONR Summer Faculty Research Fellow at the Naval Air Warfare Center C Weapons Division (NAWCWD) at China Lake, CA. He is currently a regular consultant for the U.S. Army Research Laboratory in Adelphi, MD.

Dr. Tran was the co-director (with Prof. J. L. Prince) of the 33rd Annual Conference on Information Sciences and Systems (CISS'99), Baltimore, MD, in March 1999. He has served as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING as well as the IEEE TRANSACTIONS ON IMAGE PROCESSING. He was a former member of the IEEE Technical Committee on Signal Processing Theory and Methods (SPTM TC) and is a current member of the IEEE Image Video and Multidimensional Signal Processing (IVMSP) Technical Committee. He received the NSF CAREER award in 2001, the William H. Huggins Excellence in Teaching Award from The Johns Hopkins University in 2007, and the Capers and Marion McDonald Award for Excellence in Mentoring and Advising in 2009.

Digital Holographic Imaging of Aerosol Particles in Flight

Matthew J.Berg and Gorden Videen

Journal of Quantitative Spectroscopy and Radiative Transfer 112 (2011) 1776–1783

Contents lists available at ScienceDirect

Journal of Quantitative Spectroscopy & Radiative Transfer

journal homepage: www.elsevier.com/locate/jqsrt

Digital holographic imaging of aerosol particles in flight

Matthew J. Berg [a,b,*], Gorden Videen [a,b]

[a] Mississippi State University, Department of Physics & Astronomy, Mississippi State, MS 39762, USA
[b] US Army Research Laboratory, RDRL-CIE-S, 2800 Powder Mill Road, Adelphi, MD 20783-1197, USA

ARTICLE INFO

Available online 20 January 2011

Keywords:
Holography
Light scattering
Aerosol
Imaging

ABSTRACT

This work describes the design and application of an apparatus to image aerosol particles using digital holography in a flow-through, contact-free manner. Particles in an aerosol stream are illuminated by a triggered, pulsed laser and the pattern produced by the interference of this light with that scattered by the particles is recorded by a digital camera. The recorded pattern constitutes a digital hologram from which an image of the particles is computationally reconstructed using a fast Fourier transform. This imaging is validated using a cluster of ragweed pollen particles. Examples involving mineral-dust aerosols demonstrate the technique's *in situ* imaging capability for complex-shaped particles over a size range of roughly 15–500 μm micrometers. The focusing-like character of the reconstruction process is demonstrated using a NaCl aerosol particle and is compared to a similar particle imaged with a conventional microscope.

1. Introduction

The *in situ* characterization of small aerosol particles is a persistent objective in applied contexts. Examples include the determination of atmospheric aerosol composition for climate modeling and the detection of biological weapons agents for defense applications. Countless measurements and calculations of single and multiple-particle scattering patterns can be found in the literature. The overall goal of such work is to infer information relating to the particles' physical form, such as size and shape, by analyzing the angular structure of these patterns, e.g. see [1]. Unfortunately, a fundamental limitation of this approach is the absence of an unambiguous quantitative relationship between a pattern and the corresponding particle properties, i.e., the so-called inverse problem. Consequently, the inference of these properties from the patterns has proved to be very difficult in practice, except for the simplest of cases.

Ideally, one would prefer to image the particles directly, thus eliminating the complexity and ambiguity associated with interpretation of the scattering patterns. However, the typical particle size range of interest for many applications is roughly 0.1–10 μm [1,2]. Because of this, direct images are possible in part of this range only with high numerical-aperture (NA) optics and correspondingly small focal volumes. This typically requires collection and immobilization of particle samples, and thus, such imaging is not a practical technique for particle characterization in applications requiring high sample through-put or images of the particles in their undisturbed form, i.e., *in situ* images.

Holography is an alternative technique that combines useful elements of both conventional imaging and scattering. Fundamentally, this is a two-step process: First, an object is illuminated with coherent light and the intensity pattern resulting from the interference of this light with that scattered by the particle is recorded. This pattern constitutes the hologram, from which an image of the object is reconstructed. Traditionally, holograms are

* Corresponding author at: Mississippi State University, Department of Physics & Astronomy, Mississippi State, MS 39762, USA.
Tel.: +1 785 317 3378; fax: +1 662 325 8898.
E-mail address: matt.berg@msstate.edu (M.J. Berg).

recorded with photographic film due to the film's high resolution, which is required to capture the finer features of the interference pattern. The subsequent chemical development of the film is costly and time consuming, and this greatly limits the practical utility of the technique. For this reason, charged coupled device (CCD) detectors are used to record the interference pattern digitally. The resulting so-called digital hologram can then be computationally processed, rather than chemically, to reconstruct an image of the object.

Digital holographic imaging has been demonstrated in multiple small-particle systems and across visible and X-ray wavelengths, see e.g. [3–11]. Examples of work applying holography to aerosols are scarce, and to the best of our knowledge have not yet been reported for *in situ* imaging of aerosol particles in the $0.1-25$ μm size range using visible light. This article will describe the design and implementation of an apparatus that achieves imaging of particles approximately $15-500$ μm in size, and has the potential to image particles as small as 4 μm given further design optimization. The basic concepts involved are briefly reviewed and a validation measurement using ragweed pollen particles is presented. Saharan, Tunisian, and sodium chloride (NaCl) aerosols are used to establish the *in situ* capability of the apparatus. Finally, the microscope-like focusing behavior of the image-reconstruction process is demonstrated using a single NaCl aerosol particle.

2. Digital in-line holography

The apparatus in this work is based on the so-called in-line holographic configuration [3]. Here, the particle, primary optical components, and detector are all co-linearly arranged. The particle is illuminated by a monochromatic spherical wave and the resulting interference pattern formed by this reference wave and the light scattered by the particle is recorded by a CCD detector. Let the source of the reference wave be located at a distance l from the particle and the detector at a distance d. Provided that kl and kd are large enough to satisfy the far-field conditions of [12], both the reference and scattered waves will be transverse and spherical at the detector and can be represented entirely by their scattering amplitudes

$$E^{\mathrm{ref}}(\mathbf{r}) = \frac{\exp(ikr)}{r} E_1^{\mathrm{ref}}(\hat{\mathbf{r}}), \quad E^{\mathrm{sca}}(\mathbf{r}) = \frac{\exp(ikr)}{r} E_1^{\mathrm{sca}}(\hat{\mathbf{r}}), \tag{1}$$

respectively. Then, the intensity of the total wave across the detector's face is [3]

$$I^{\mathrm{holo}}(\mathbf{r}) = \frac{c\varepsilon_o}{r^2} |E_1^{\mathrm{ref}}(\hat{\mathbf{r}}) + E_1^{\mathrm{sca}}(\hat{\mathbf{r}})|^2, \tag{2}$$

where c and ε_o are the vacuum speed of light and electric permittivity, respectively. Expanding Eq. (2) gives

$$I^{\mathrm{holo}}(\mathbf{r}) = \frac{c\varepsilon_o}{r^2} \{ |E_1^{\mathrm{ref}}(\hat{\mathbf{r}})|^2 + |E_1^{\mathrm{sca}}(\hat{\mathbf{r}})|^2 + [E_1^{\mathrm{ref}}(\hat{\mathbf{r}})]^* E_1^{\mathrm{sca}}(\hat{\mathbf{r}}) + [E_1^{\mathrm{sca}}(\hat{\mathbf{r}})]^* E_1^{\mathrm{ref}}(\hat{\mathbf{r}}) \}. \tag{3}$$

The quantity $c\varepsilon_o r^{-2} |E_1^{\mathrm{ref}}(\hat{\mathbf{r}})|^2 = I^{\mathrm{ref}}(\mathbf{r})$ in Eq. (3) is the intensity across the detector when no particle is present, and hence can be considered a known quantity measured before the introduction of an aerosol sample. Subtracting this reference intensity from Eq. (3) and dividing the remaining terms by it gives

$$I^{\mathrm{con}}(\mathbf{r}) = \frac{I^{\mathrm{holo}}(\mathbf{r}) - I^{\mathrm{ref}}(\mathbf{r})}{I^{\mathrm{ref}}(\mathbf{r})} = \frac{|E_1^{\mathrm{sca}}(\hat{\mathbf{r}})|^2}{|E_1^{\mathrm{ref}}(\hat{\mathbf{r}})|^2} + \frac{[E_1^{\mathrm{ref}}(\hat{\mathbf{r}})]^* E_1^{\mathrm{sca}}(\hat{\mathbf{r}}) + [E_1^{\mathrm{sca}}(\hat{\mathbf{r}})]^* E_1^{\mathrm{ref}}(\hat{\mathbf{r}})}{|E_1^{\mathrm{ref}}(\hat{\mathbf{r}})|^2}. \tag{4}$$

Often, the intensity of the reference wave at the detector is much greater than that of the scattered wave. This is especially true in this work where the objects being illuminated are small particles, as opposed to the macroscopic sized objects involved in other applications, see e.g. [13–16]. This means that the term $|E_1^{\mathrm{sca}}(\mathbf{r})|^2 / |E_1^{\mathrm{ref}}(\mathbf{r})|^2$ in Eq. (4) can be neglected, leaving

$$I^{\mathrm{con}}(\mathbf{r}) \simeq \frac{[E_1^{\mathrm{ref}}(\hat{\mathbf{r}})]^* E_1^{\mathrm{sca}}(\hat{\mathbf{r}}) + [E_1^{\mathrm{sca}}(\hat{\mathbf{r}})]^* E_1^{\mathrm{ref}}(\hat{\mathbf{r}})}{|E_1^{\mathrm{ref}}(\mathbf{r})|^2}. \tag{5}$$

This intensity pattern, which is the difference between two measurements – with and without the particle present – is known as a contrast hologram. The key characteristic of I^{con} is its linear dependence on the amplitude of the particle's scattered wave. This means that the phase of the scattered wave over the detector is encoded in the measurement. Consequently, I^{con} can be used to reconstruct unambiguously an image of the particle that closely resembles that obtained from conventional microscopy.

Because there are many references describing the theory behind digital holographic imaging, only a brief description will be given here, see e.g. [17–19]. Basically, the contrast hologram is envisioned as a transmission diffraction-grating illuminated by a normally incident plane wave, i.e., a reconstruction wave. The Fresnel–Kirchhoff approximation is then used to describe the light diffracted from this grating in a parallel plane separated by a distance z from the grating along the z-axis. If z corresponds to the distance between the particle and detector during the hologram measurement ($z=d$) the resulting diffraction pattern in this so-called reconstruction plane yields an image of the particle. The image is essentially equivalent to a conventional microscope image, although the resolution is typically less [3].

The advantage of using the Fresnel–Kirchhoff approximation to calculate the reconstructed particle image is that the approximation's mathematical form is essentially a discrete Fourier transform of the CCD pixel values constituting I^{con}. This enables the use of the fast Fourier transform (FFT) in the calculation, thus substantially reducing the computation time required to render the particle image. This is fortuitous, because in practice d is not known to great enough accuracy to be able to reconstruct an image from a single application of the reconstruction routine. This inaccuracy is due to the variation in particle positions in the aerosol stream as they enter the measurement volume. Consequently, the image-reconstruction stage consists of a focusing-like procedure: First an initial image is reconstructed using an estimate of d based on the experimental layout. Then, the reconstruction plane is scanned along the z-axis in

small steps until the reconstructed image comes into focus. The ability to use the FFT for each of these intermediate steps is thus crucial to the practical implementation of this technique.

The primary drawback to the in-line configuration is that two images of the particle are produced in the reconstruction stage [17]. The in-focus particle image is always accompanied by a blurred twin image that is in-focus in the mirror reconstruction plane, i.e., at $z=-d$. As a consequence, the image quality is degraded. However, as shown in [3], the effect of the twin on the in-focus image becomes negligible if both d and the size of the CCD pixel array are sufficiently large such that an imaging resolution on the order of the wavelength can theoretically be achieved [3,4].

Another drawback of in-line holography is the presence of the zero frequency, or so-called DC, term in the reconstructed image [20]. In the diffraction-grating model, the reconstruction wave is uniform across the hologram since it is planar and normally incident. Upon application of the FFT to I^{con}, this wave then becomes a strong DC contribution in the transform. The result is an unwanted bright spot in the reconstructed image located at the intersection of the optical axis (z-axis) with the reconstruction plane. Fortunately, however, the DC term can be nearly eliminated by subtracting from each pixel value in I^{con} the average value of all the pixels [17]. Notice that in doing this subtraction, the result is a new contrast hologram with both positive and negative values; whereas, its constituent holograms I^{holo}, I^{ref}, and I^{con}, are all inherently positive since they correspond to intensity measurements.

The resolution of the resulting particle images is limited by several factors related to diffraction and the apparatus hardware: the CCD pixel size, CCD pixel-array size w, particle-CCD distance d, and the illumination wavelength λ [17,4]. Given the configuration of the optical elements in this work, the theoretical minimum resolvable length scale is approximately 4 μm following [4]. However, the resolution achieved in practice is in the range of 8−10 μm due to stray-light noise and imperfections in the optical design. Fundamentally, the resolution of this holographic configuration will not exceed what is possible from a conventional optical microscope. However, as discussed earlier, it does provide the substantial advantage of near real-time, *in situ*, and high through-put imaging, which is not typically possible with conventional microscopy.

3. Apparatus design and validation

The experimental apparatus, which is shown in Fig. 1, consists of two primary subsystems: aerosol-particle sensing and hologram recording. An aerosol stream is delivered via a nozzle made from a plastic pipettor-tip to the measurement volume where an optical trigger is used to sense the presence of a particle [21,22]. This trigger consists of crossed diode-laser beams, labeled (h) and (i) in Fig. 1. These lasers have different wavelengths of 635 and 670 nm and intersect near the outlet nozzle delivering the aerosol. When a particle passes into this intersection it scatters both wavelengths of light simultaneously. The scattered light is received by two photomultiplier (PMT) modules (Hamamatsu Corp., model H6780-02), (j) in the figure, each sensitive to only one of the two wavelengths. A series of signal-analysis units determines if the signals produced by the PMT modules are coincident. If so, this indicates the presence of a particle at the trigger laser-beam intersection and a fire signal is sent to a pulsed laser for the hologram recording.

The triggered light source is a 70 ns pulsed Nd:YAG laser (Spectra Physics Lasers, Inc., model Y70-532Q), frequency doubled to 532 nm. This light passes through a Glan–Thompson polarizer to ensure linear polarization (a) in Fig. 1. The light is then focused by lens (b) onto a 50 μm diameter pinhole (c). Next the primary lobe of this pinhole diffraction pattern illuminates a second pinhole (d) with a diameter of 25 μm. These pinholes "clean" the beam improving its spatial coherence and enhancing the quality of the hologram. All but the primary lobe of this second pinhole pattern is blocked by iris (e) where lens (f) then collimates the beam, which is brought to a focus by lens (g) at a point approximately 2 mm from the aerosol nozzle outlet. This 2 mm is the distance l in Section 2. In this way, the aerosol particles are illuminated by what is approximately a spherical wave originating from the beam waist. The beam continues until reaching the CCD detector (Finger Lakes Instrumentation, LLC, model ML8300), at which point it expands to fill the entire pixel array (5.4 μm pixel size, 3326×2504 pixel-array size). The separation between the particle stream and the detector is the d discussed in Section 2 and is approximately 8 cm. A small amount of the beam is scattered by the particle (dashed line in Fig. 1), and this light interferes with the remainder of the beam, i.e., the reference wave, to form the interference pattern that becomes the digital hologram I^{holo}.

To test the apparatus and provide a rough calibration of the image-reconstruction procedure, a comparison is made between a holographic and optical microscope image of the same particle. This is done by placing 15.4 μm diameter NIST-traceable polystyrene latex microspheres (Duke Scientific Corp.) on a microscope slide and positioning the slide in the measurement volume at the intersection of the trigger-beams. A hologram is recorded, from which the image-reconstruction procedure of Section 2 is followed. The slide is then transferred to a microscope, where the same spheres are located and imaged. Next, using a 1951 USAF glass-slide resolution target (Edmund Optics), a scale factor is determined relating the microscope-image pixel number to micrometers. Then, by comparing the holographic image of a microsphere to the microscope image of the same microsphere, an additional scale factor is determined relating the hologram pixel number to micrometers. In this way, the holographic images of all subsequent particles can be rendered in calibrated length (micrometers), rather than pixel number. This calibration procedure is approximate, however, because there is ambiguity in determining the hologram pixel-number size of a given microsphere: The contrast between the reconstructed sphere-image and the background is not sharp, which

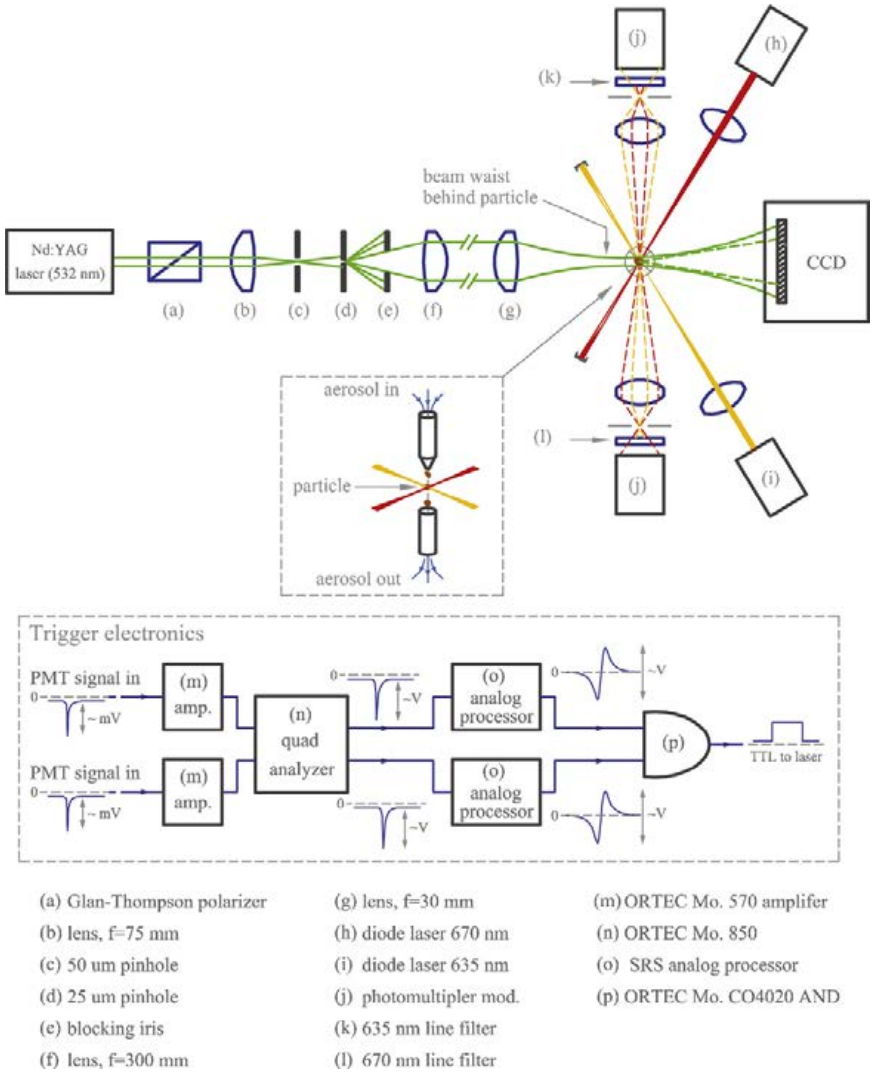


Fig. 1. Diagram of the apparatus. The middle inset shows a schematic of the signal-analysis electronics used in the optical trigger to sense the presence of a particle in the measurement volume. See text for further explanation.

can yield a scale factor that over- or under-determines the particle size.

An example is presented in Fig. 2 demonstrating the comparison between the holographic and microscope images of the *same* particle. Here a cluster of ragweed pollen particles is placed on a microscope slide, then holographic and microscope images of the cluster are obtained. By comparing these images, one can see that the holographic apparatus successfully produces an accurate image of the pollen cluster, with sufficient resolution to discern individual pollen particles and even a faint signature of the single-particle surface roughness seen in the microscope images. This corresponds to a resolution roughly between $8-10\,\mu m$, although a more rigorous resolution analysis is not performed. Referring to the measured and contrast holograms shown in this figure, one can see how subtraction of the incident beam across the CCD, i.e., I^{ref}, removes noise due to imperfections in the incident beam profile. This has the consequence of producing a "cleaner" contrast hologram, which subsequently improves the particle image. Note that the holographic and microscope images of the cluster differ slightly in overall size and detailed structural form. Although it is clearly the same cluster in (c) and (d), the differences are likely due to shifting of the cluster on the microscope slide during transfer from the apparatus to the microscope.

There are several unique aspects to the design of this apparatus. By using the short focal-length lens (g) in Fig. 1 to form a beam waist near the particle, the light illuminating the particle is more intense than it would be if only the pinhole was used for illumination (as is usually done). This results in a relative amplification of the scattered wave at the detector and enhances the interference structure of the hologram leading to improved particle-image quality. Using a pulsed laser permits the

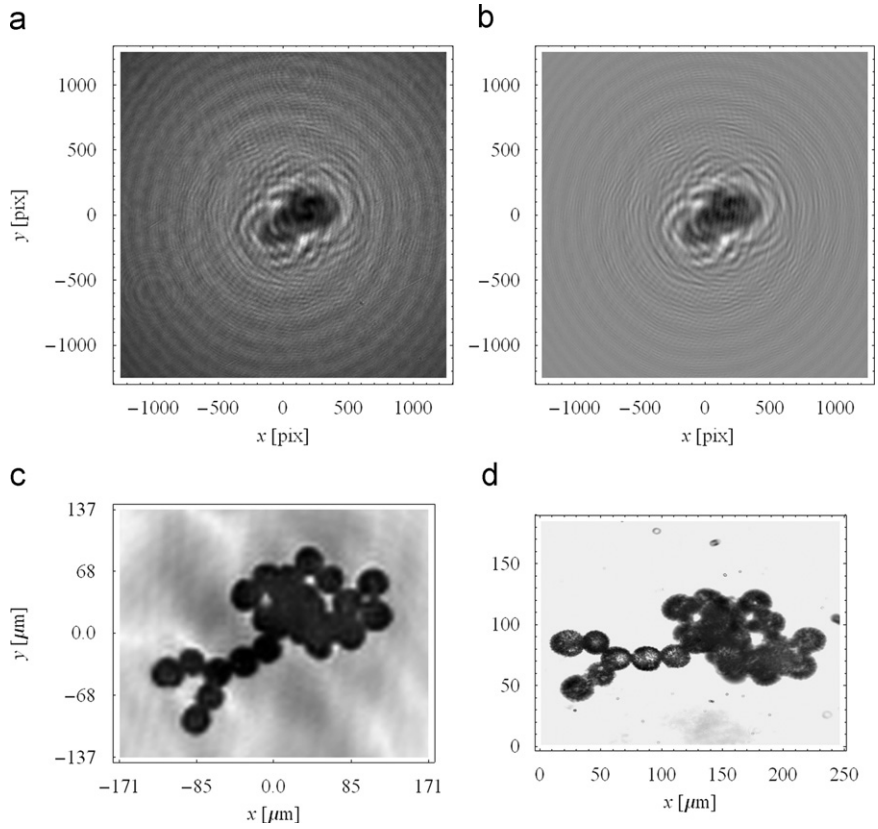


Fig. 2. Validation of the holographic imaging apparatus. Plots (a) and (b) show the measured I^{holo} (digital) and corresponding contrast I^{con} holograms, respectively, for a cluster of ragweed pollen particles on a microscope slide located at the intersection of the trigger-beams, recall Fig. 1. Image (c) shows the reconstructed image resulting from (b) whereas (d) shows a conventional microscope image of the same cluster.

investigation of particle systems in motion. This also greatly relaxes the strict mechanical-stability demands typically required for holographic measurements. There are no optical elements between the aerosol stream and the CCD. This gives the apparatus a working distance of several centimeters, which is substantially greater than the single- to sub-millimeter working distance of a microscope objective. Moreover, the absence of any optical elements between the detector and particle eliminates "noise" resulting from ambient dust that can collect on the optical surfaces.

4. Applications

To further assess the imaging capabilities of the apparatus, several aerosols consisting of complex-shaped particles are examined. The first samples are sieved Saharan and Tunisian sand, which are aerosolized using an Erlenmeyer flask as follows: A small sample of the sand is placed in the flask, then sealed with a stopper. Two aluminum tubes pass through the stopper; one supplies air to the flask, blowing the sand particles around, while the other tube allows some of the airborne particles to exit the flask and be transported to the aerosol nozzle in the apparatus. Fig. 3 shows the contrast holograms along with the resulting particle-image reconstructions for single Saharan and Tunisian sand particles. For comparison, Fig. 4 shows microscope images of these sand samples. One can see that the holographic images provide the same information of overall particle size and morphology as the microscope images. For example, the Saharan particles appear to have less surface roughness than the Tunisian particles. Note that unlike Fig. 2, the particles shown in the holographic reconstructions (Fig. 3) and microscope images (Fig. 4) are not the same sand particles since the holographic images are obtained from flowing particles.

Another unique capability of holographic imaging is that some sense of the three-dimensional form of a particle can be garnered from a single measurement. The basic idea is analogous to the "focusing in" on a particle in conventional microscopy. There, the microscope objective is moved vertically to vary the distance between it and the microscope slide, causing a blurred image of a particle to evolve into a sharp image. If the particle has sufficient thickness and transparence, different depths within the particle can be brought into focus to give a feel for the particle's three-dimensional structure. This same process can be done in digital holography by computationally varying the distance d used in the image-reconstruction stage, as is shown in [3]. The resulting sequence of images gives the same impression of focusing in on the particle as one gets from microscopy. However,

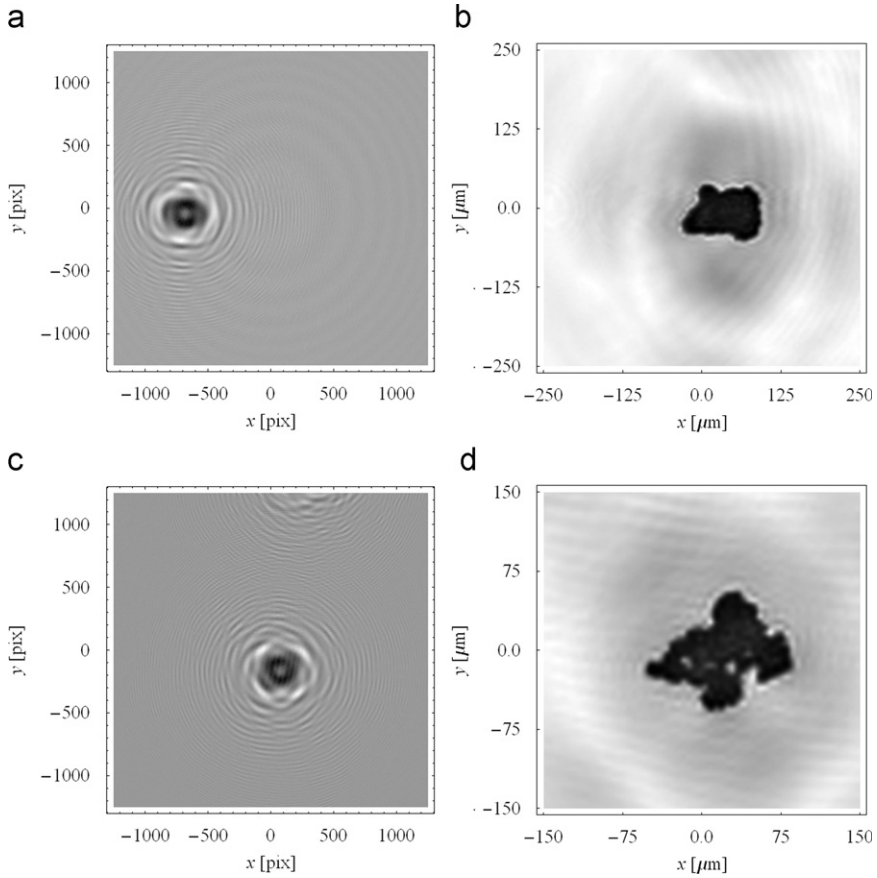


Fig. 3. Saharan and Tunisian sand particles. Images (a) and (b) show the contrast hologram I^{con} and corresponding reconstructed image for a single Saharan sand particle. Images (c) and (d) show the same for a single Tunisian sand particle.

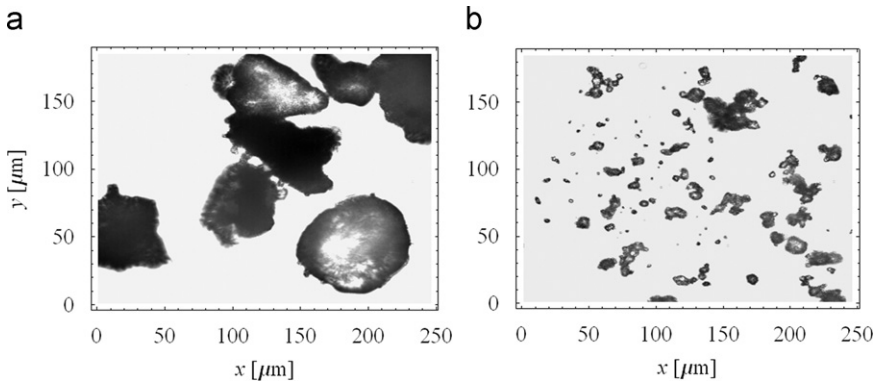


Fig. 4. Microscope images of (a) Saharan and (b) Tunisian sand. The particles seen here are taken from the same sand samples used in Fig. 3, but unlike the ragweed in Fig. 2, these particles are not the exact same particles imaged holographically.

unlike microscopy where an image must be recorded at each "focus depth," the holographic route can obtain a similar image-sequence from the (single) contrast hologram only.

Fig. 5 shows an example of this holographic focusing process. The top row displays conventional microscope images of a NaCl crystal at different focus depths. The bottom row shows a holographic image-sequence for an aerosolized NaCl particle that is produced by scanning the reconstruction plane along the z-axis around $z=d$. The particle in the holographic images is delivered to the apparatus in aerosol form by drying a salt solution on a hotplate and aerosolizing the resulting powder using the Erlenmeyer generator described in Section 3. One can

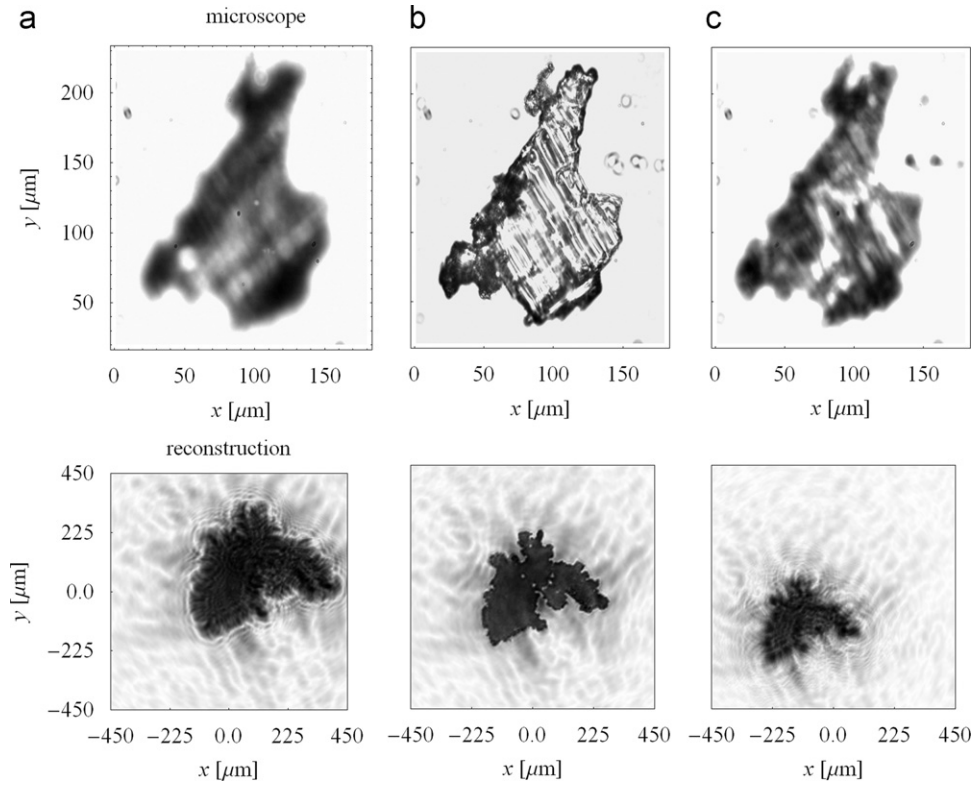


Fig. 5. Focusing behavior of the holographic image-reconstruction process. The top row shows microscope images of a NaCl crystal on a microscope slide at three different focus depths (a)–(c). The bottom row shows the reconstructed images of a NaCl aerosol particle when the reconstruction plane is at three positions for z: $z < d$ for (a), $z = d$ for (b), i.e., in-focus, and $z > d$ for (c).

clearly see the strong similarity in the focusing behavior of the two imaging techniques.

5. Comments

The *in situ* images of aerosol particles presented here are not the only documented examples. Sorensen et al. have obtained images of the particles constituting hydrocarbon-flame soot at various stages in soot formation, i.e., as a function of height in a flame [23,24]. Here a $10\times$-power photomicroscope is mated to a conventional film-camera and a 1.5 μs Xe flash lamp is used for particle illumination. With this arrangement, particles in the range of roughly $5-100$ μm are imaged, which covers the same particle size range considered in our work. One might then wonder what advantage the holographic approach offers over this photomicroscope direct-imaging.

First, the photomicroscope images are obtained photographically, i.e., using film, requiring chemical processing. The holograms, however, are entirely digitally recorded and the resulting images are computationally rendered. Second, and perhaps most important, the photomicroscope images have a very narrow depth of field, and only particles constrained within a narrow volume are in-focus; whereas, for holographic techniques the focusing is done computationally, after the hologram is recorded. This enables the focusing process described in Section 4, which can be used to image multiple particles present at different locations in the measurement volume as demonstrated in [3]. Moreover, this can be done from a single hologram recording. To do this with the photomicroscope would require obtaining a series of exposures with the microscope objective positioned at different distances from the measurement volume. Thus, if the particles are in motion, as they are in flow-through applications, a series of exposures would prevent the imaging of multiple particles present *at a given instant* in the measurement volume.

As mentioned in Section 3, an inherent advantage of the holographic design is that there are no optical elements between the particle and detector. Thus, there are no surfaces for ambient dust to collect on and become sources of stray light, nor are there any lens-based aberrations and multiple reflections. Both of these concerns are present in the photomicroscope approach. The absence of these optical elements in the holographic design is especially advantageous when one wishes to investigate particles that are roughly the same size as ambient dust.

6. Conclusion

This work demonstrates the feasibility of imaging single and multiple aerosol particles *in situ* using digital in-line holography. Imaging is demonstrated on ragweed pollen, Saharan and Tunisian sand, and NaCl particles; a

range of overall particle-size covering approximately 15−500 μm. These images are computationally reconstructed from the digitally recorded holograms and compare well to the corresponding microscope images. Although the resolution of the holographic images is less than those from the microscope, one is able to clearly discern single-particle size and shape. Moreover, the ability to computationally render the images allows the application of numerical operations to improve image quality, whereas the analogs of such operations in conventional optical imaging would be difficult to implement.

Acknowledgments

This work was supported by the Defense Threat Reduction Agency and the US Army Research Laboratory. We are thankful for helpful discussions with Yong-Le Pan, Steve Hill, Chris Sorensen, Dave Ligon, and Jay Eversole.

References

[1] Aptowicz KB, Pinnick RG, Hill SC, Pan YL, Chang RK. Optical scattering patterns from single urban aerosol particles at Adelphi, Maryland, USA: a classification relating to particle morphologies. J. Geophys. Res. 2006;111:D12212.
[2] Hoekstra A, Maltsev V, Videen G, editors. Optics of Biological Particles. Dordrecht: Springer; 2007.
[3] Xu W, Jericho MH, Meinertzhagen IA, Kreuzer HJ. Digital in-line holography of microspheres. Appl. Opt. 2002;41:5367–75.
[4] Garcia-Sucerquia J, Xu W, Jericho SK, Klages P, Jericho MH, Kreuzer HJ. Digital in-line holographic microscopy. Appl. Opt. 2006;45:836–50.
[5] Kreuzer HJ, Jericho MJ, Meinertzhagen IA, Xu W. Digital in-line holography with photons and electrons. J. Phys. Condens. Matter 2001;13:10729–41.
[6] Raupach SMF, Vossing HJ, Curtius J, Borrmann S. Digital crossed-beam holography for in situ imaging of atmospheric ice particles. J. Opt. A: Pure Appl. Opt. 2006;8:796–806.
[7] Fugal JP, Shaw RA. Cloud particle size distributions measured with an airborne digital in-line holographic instrument. Atmos. Meas. Tech. 2009;2:259–71.
[8] Singh VR, Hegde G, Asundi A. Particle field imaging using digital in-line holography. Curr. Sci. 2009;96:391–7.
[9] Thompson BJ. Holographic particle sizing techniques. J. Phys. E: Sci. Instrum. 1974;7:781–8.
[10] Bogan MJ, Starodub D, Hampton CY, Sierra RG. Single-particle coherent diffractive imaging with a soft X-ray free electron laser: towards soot aerosol morphology. J. Phys. B: At. Mol. Opt. Phys. 2010;43:1–13.
[11] Bogan MJ, Boutet S, Chapman HN, Marchesini S, Barty A, Benner WH, et al. Aerosol imaging with a soft X-ray free electron laser. Aero. Sci. Tech. 2010;44:i–iv.
[12] Mishchenko MI, Travis LD, Lacis AA. Multiple scattering of light by particles: radiative transfer and coherent backscattering. Cambridge: Cambridge University Press; 2006.
[13] Schnars U, Juptner W. Direct recording of holograms by a CCD target and numerical reconstruction. Appl. Opt. 1994;33:179–81.
[14] Zhang S. Application of super-resolution image reconstruction to digital holography. EURASIP J. Appl. Signal Process. 2006: 1–7 [article ID 90358].
[15] Ma L, Wang H, Li Y, Jin H. Numerical reconstruction of digital holograms for three-dimensional shape measurement. J. Opt. A: Pure Appl. Opt. 2004;6:396–400.
[16] Javidi B, Kim D. Three-dimensional-object recognition by use of single-exposure on-axis digital holography. Opt. Lett. 2005;30: 236–8.
[17] Kreis T. Handbook of holographic interferometry, optical and digital methods. Germany: Wiley; 2005.
[18] Kim MK. Principles and techniques of digital holographic microscopy. SPIE Rev. 2010;1:018005-1–50.
[19] Ferraro P, De Nicola S, Coppola G. Digital holography: recent advancement and prospective improvements for applications in microscopy. In: Optical imaging sensors and systems for homeland security applications, Advanced Sciences and Technologies for Security Applications, vol. 2; 2006. p. 47-84.
[20] Goodman JW. Introduction to Fourier optics. Englewood: Roberts & Company; 2005.
[21] Pan YL, Holler S, Chang RK, Hill SC, Pinnick RG, Niles S, et al. Single-shot fluorescence spectra of individual micrometer-sized bioaerosols illuminated by a 351- or 266-nm ultraviolet laser. Opt. Lett. 1999;24:116–8.
[22] Berg MJ, Hill SC, Pan Y-L, Videen G. Two-dimensional Guinier analysis: application to single aerosol particles in-flight. Opt. Ex. 2010;18:23343–52.
[23] Sorensen CM, Hageman WB. Two dimensional soot. Langmuir 2001;17:5431–4.
[24] Sorensen CM, Hageman WB, Rush TJ, Huang H, Oh C. Aerogelation in a flame soot aerosol. Phys. Rev. Lett. 1998;80:1782–5.

94-GHz Imager with Extended Depth of Field

Joseph N. Mait, David A. Wikner, Mark S. Mirotznik, Joseph van der Gracht,
Gregory P. Behrmann, Brandon L. Good and Scott A. Mathews

IEEE Trans. Antennas Propag. 57, 1713-1719 (2009)

94-GHz Imager With Extended Depth of Field

Joseph N. Mait, *Senior Member, IEEE*, David A. Wikner, Mark S. Mirotznik, *Member, IEEE*, Joseph van der Gracht, Gregory P. Behrmann, Brandon L. Good, and Scott A. Mathews, *Member, IEEE*

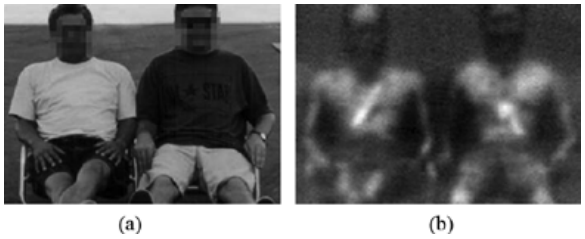
***Abstract*—We describe a computational imaging technique to extend the depth-of-field of a 94-GHz imaging system. The technique uses a cubic phase element in the pupil plane of the system to render system operation relatively insensitive to object distance. However, the cubic phase element also introduces aberrations but, since these are fixed and known, we remove them using post-detection signal processing. We present experimental results that validate system performance and indicate a greater than four-fold increase in depth-of-field from 17" to greater than 68".**

***Index Terms*—Computational imaging, extended depth of field, millimeter wave imaging.**

I. Introduction

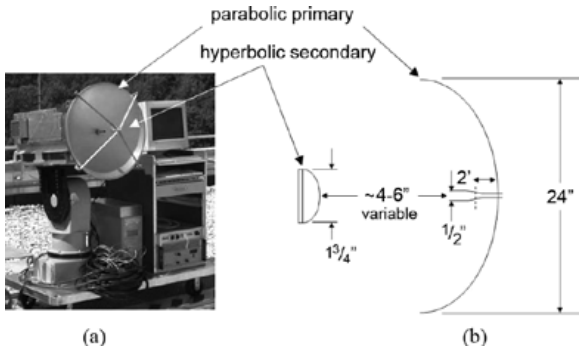
THE ability of gigahertz and terahertz frequencies to penetrate materials that are impenetrable at optical frequencies has prompted recent interest in the development of millimeter wave sources and detectors [1]. Applications of this capability include, for example, the detection of concealed weapons under clothing [2], [3]. However, unlike the stationary figures shown in Fig. 1, a more typical scenario for this application is screening individuals at points of ingress, such as the entrance to a building or the secured portion of an airport. To enhance the performance of screening systems, one would prefer to observe individuals as long as possible as they pass through a volume. This improves the chances of detecting a hidden object. (It might also reduce bottlenecks created at portals.) However, wavelength and system considerations limit focused imaging to only a narrow volume in depth, or depth-of-field. Thus, a screener has only a short amount of time to detect the presence or absence of a concealed weapon. Extending the depth-of-field provides the screener with more time to observe an individual.

A similar problem occurs in iris recognition for security applications, for example, logging-on to a computer system. In this situation the narrow depth of field produces unnatural head movements as a user seeks to place his or her iris in the object plane of the optical system. Extending the depth-of-field for these systems has been addressed using computational imaging techniques [4]–[7]. By computational imaging, we mean an imaging system whose pre-detection optics and post-detection signal processing are designed jointly to achieve a result that is not possible using only optics or only signal processing [8], [9]. For example, placing an optical element with cubic-phase



(a) (b)

Fig. 1. Gigahertz imaging through clothing. (a) Visible image of scene. (b) 94-GHz image of weapons concealed under clothing.



(a) (b)

Fig. 2. 94-GHz scanning imaging system. (a) Image of system. (b) Schematic representation with measured dimensions.

in the pupil plane of the optical system renders system operation relatively insensitive to object distance. However, the cubic phase also generates an aberrated image. But, since the aberrations are known, one can correct them using simple post-detection signal processing. Since the system response is effectively invariant to object location, the combination of optical and electronic processing yields a system with larger depth-of-field than a conventional system.

In this work we describe our application of this technique for extended depth-of-field imaging to a 94-GHz system and present experimental results to verify its performance. In Section II we describe our imaging system and present a mathematical description of its operation with and without extended depth-of-field. We describe design and fabrication of the cubic-phase element in Section III and present experimental results in Section IV. Section V discusses the signal processing required by the cubic phase system to realize the extended depth and presents the results from this processing. We end in Section VI with summary remarks on our approach.

II. 94-GHz Imaging System

Our imaging system, represented in Fig. 2, is a 94-GHz Stokes-vector radiometer used for millimeter wave phenomenology measurements [10]. It is a single-beam system

Manuscript received December 18, 2007; revised June 04, 2008. Current version published June 03, 2009.

J. N. Mait and D. A. Wikner are with the U.S. Army Research Laboratory, Adelphi, MD 20783 USA (e-mail: jmait@arl.army.mil).

M. S. Mirotznik, G. P. Behrmann, B. L. Good, and S. A. Mathews are with the Department of Electrical Engineering, Catholic University of America, Washington, DC 20064 USA.

J. van der Gracht is with Holospex, Columbia, MD 21044 USA.

Digital Object Identifier 10.1109/TAP.2009.2019882

that forms an image by scanning in azimuth and elevation. The radiometer has a thermal sensitivity of 0.3 K with a 30-ms integration time and 1-GHz bandwidth per pixel. A Cassegrain antenna is mounted to the front of the radiometer receiver. It has a 24"-diameter primary parabolic reflector and a 1.75"-diameter secondary hyperbolic reflector. The position of the hyperbolic secondary is variable.

If we model the 94-GHz imager as a linear, spatially incoherent, quasi-monochromatic system, the intensity of the detected image can be represented as a convolution between the intensity of the image predicted by geometrical optics with the system point spread function [11]

$$|i(x,y)|^2 = o_g(x,y) * *h(x,y) \tag{1}$$

where $**$ represents a two-dimensional convolution. The function $o_g(x,y)$ represents the inverted, magnified image of the object that a ray-optics analysis of the system predicts

$$o_g(x,y) = \frac{1}{M^2}\left|o\left(\frac{-x}{M},\frac{-y}{M}\right)\right|^2. \tag{2}$$

If the object and image distances are d_o and d_i, respectively, the magnification M is

$$M = \frac{d_i}{d_o}. \tag{3}$$

For the purposes of geometrical analysis, we can model the system as a single lens imaging system with $d_i = 6''$ (152.4 mm)

$$\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f}. \tag{4}$$

The value of d_i is based on measurements of the antenna. We adjusted the position of hyperbolic element so that nominal operation of the imager is with $d_o = 180''$ (4572 mm). Thus, the effective focal length of the system is $f = 5.81''$ (147.6 mm).

The second term in (1), $h(x,y)$, is the incoherent point spread function (PSF). It accounts for wave propagation through the aperture

$$h(x,y) = \frac{1}{(\lambda f)^4}\left|p\left(\frac{-x}{\lambda f},\frac{-y}{\lambda f}\right)\right|^2 \tag{5}$$

where $p(x/\lambda f, y/\lambda f)$ is the coherent point spread function. The function $p(x,y)$ is the inverse Fourier transform of the system pupil function $P(u,v)$

$$p(x,y) = \mathrm{FT}^{-1}\left[P(u,v)\right]. \tag{6}$$

As a consequence, the optical transfer function (OTF) $H(u,v)$ associated with the PSF is the autocorrelation of the pupil function $P(u,v)$ with frequency axes scaled by λf

$$H(u,v) = P(\lambda f u, \lambda f v) \star\star P(\lambda f u, \lambda f v) \tag{7}$$

where $\star\star$ represents two-dimensional correlation. For example, for a circular aperture of diameter D

$$P(u,v) = \mathrm{circ}\left(\frac{2\sqrt{u^2+v^2}}{D}\right) \tag{8}$$

the PSF $h(x,y)$ is, to within a constant

$$h(x,y) = \frac{\pi^2}{4}\left(\frac{D}{\lambda f}\right)^4\left[\frac{J_1\left(\frac{\pi D\sqrt{x^2+y^2}}{\lambda f}\right)}{\frac{\pi D\sqrt{x^2+y^2}}{\lambda f}}\right]^2. \tag{9}$$

Ray analysis of our system confirmed that the parabolic primary forms the aperture stop, i.e., it defines the location of the pupil plane.

Displacement d_ϵ of an object from the nominal object plane introduces a phase error $\theta_\epsilon(u,v)$ in the pupil function [11]

$$P_\epsilon(u,v) = \exp\left[-j\theta_\epsilon(u,v)\right]\mathrm{circ}\left(\frac{2\sqrt{u^2+v^2}}{D}\right) \tag{10}$$

where

$$\theta_\epsilon(u,v) = \left(\frac{\pi}{\lambda}\right)\left(u^2+v^2\right)\left(\frac{1}{d_o+d_\epsilon} - \frac{1}{d_o}\right). \tag{11}$$

The phase error increases the width of a point response. If the displacement and phase errors are small, the detector (either human or machine) may be unable to resolve the increase and the image is perceived as in-focus.

The distance in object space over which an object can be placed and still produce an in-focus image is the system's depth-of-field DoF

$$DoF = d_{o+} - d_{o-} \tag{12}$$

where formulas for d_{o+} and d_{o-} depend upon system application. Many different definitions exist. For demonstration purposes, we use a conventional definition based on the spatial extent δ of the point response [12]

$$d_{o+} = \frac{d_i}{\left(\frac{d_i}{f-1}\right) - \frac{\delta}{D}} \tag{13}$$

and

$$d_{o-} = \frac{d_i}{\left(\frac{d_i}{f-1}\right) + \frac{\delta}{D}}. \tag{14}$$

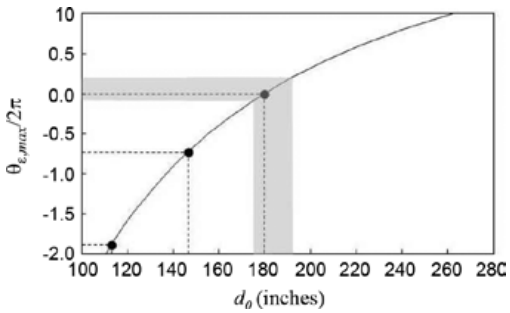


Fig. 3. Maximum relative pupil phase error as a function of object distance. The shaded region indicates a conventional depth of field. The discrete points indicate object distances used in experiments.

Under the assumptions that δ is defined by the Rayleigh criteria, the imager operates in the far-field ($d_o \gg f$), and the lens aperture is large compared to a wavelength

$$DoF \approx (2.44\lambda)\left(\frac{d_o}{D}\right)^2. \tag{15}$$

For a 94 GHz-imager with $D = 24''$ and $d_o = 180''$, $DoF \approx 17.4''$ which ranges from 175.2" to 192.6". Fig. 3 indicates the maximum relative phase error as a function of object distance and also indicates the region we have defined as being in the depth-of-field. (The maximum error for a given plane occurs at the edge of the aperture, $u^2 + v^2 = D^2/4$.)

Equation (15) explains mathematically what any good photographer already knows: one can increase DoF by decreasing ("stopping down") the lens aperture D. However, this reduces throughput and degrades the diffraction limited resolution. Alternatively, it has been shown at optical wavelengths that a cubic-phase element placed in the pupil plane of an imaging system in combination with post-detection processing can also increase DoF but without sacrificing either throughput or resolution.

The cubic phase element $P_c(u,v)$ is

$$P_c(u,v) = \exp\left[j\theta_c(u,v)\right] \operatorname{rect}\left(\frac{u}{W_u}, \frac{v}{W_v}\right) \tag{16}$$

where

$$\theta_c(u,v) = (\pi\alpha)\left[\left(\frac{2u}{W_u}\right)^3 + \left(\frac{2v}{W_v}\right)^3\right]. \tag{17}$$

The phase function is separable in the u- and v-spatial frequencies and has spatial extent W_u and W_v along the respective axes. The constant α represents the strength of the cubic phase. Along one axis the total phase change is $2\pi\alpha$; the phase change across a diagonal is $4\pi\alpha$. In the simulations presented below, we modified the model for $P_c(u,v)$ slightly and included an appropriately sized central obscuration to account for the effect of rays blocked by the secondary mirror.

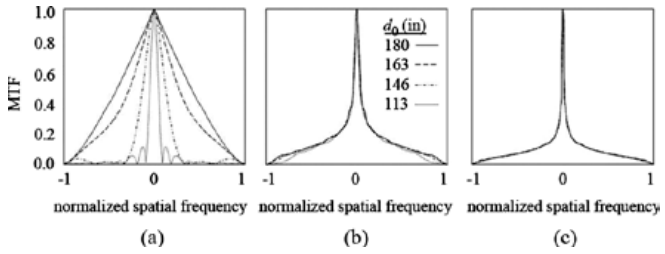


Fig. 4. Modulation transfer function for systems with different strength cubic phase: (a) $\alpha = 0$, (b) $\alpha = 7$, and (c) $\alpha = 20$. The MTFs for four different object distances are represented: 180" (solid line), 163" (dashed line), 146.5" (dot-dashed line), and 113" (dotted line).

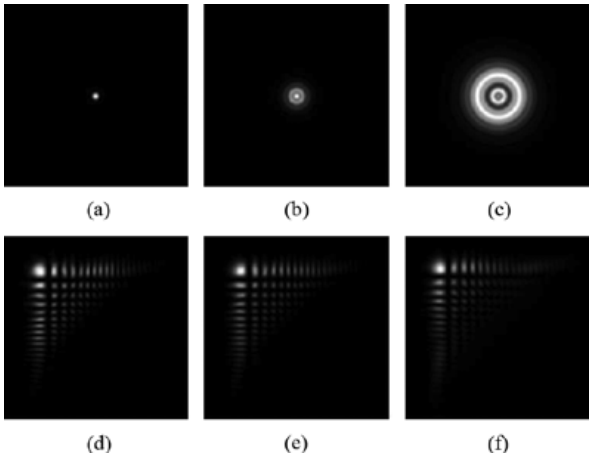


Fig. 5. Simulated point spread functions for conventional imaging and imaging with a cubic phase. Simulated PSFs for conventional imaging system at (a) 180", (b) 146.5", and (c) 113". (d)-(f) Simulated PSFs for an imaging system with cubic phase at the same object distances as (a)-(c).

Representations of the magnitude transfer function (MTF), the magnitude of the OTF $|H(u,v)|$, in cross-section are represented in Fig. 4 for three different values of $\alpha = \{0, 7, 20\}$ and different values of misfocus. (The misfocused planes are located roughly at 1, 2, and 4 times the DoF. The values noted in the figure are measured distances used in our experiments.) A conventional system with no cubic phase ($\alpha = 0$) is represented in Fig. 4(a). Note that the MTFs differ for each value of misfocus. Compensating for misfocus therefore requires *a priori* knowledge of where an object is located. Even if this information were known, due to the presence of zeros in the MTFs, inverting any one of them is ill-posed and will generate noisy results.

In contrast, the MTFs for cubic phase elements with non-zero values of α are relatively constant over an extended range. Note that the larger the value of α the larger the range over which the system is insensitive to object location. However, increasing α reduces the magnitude of the MTF, which is detrimental for applications with low signal to noise ratios. But, because the MTFs do not contain any zeros, their inversion is better conditioned than the MTFs for a conventional system.

Simulations of the point spread functions one can expect from our imaging system with and without a cubic phase element with $\alpha = 7$ are represented in Fig. 5. The response of the cubic phase system is relatively unchanged, whereas the response of the conventional system changes considerably. We address in

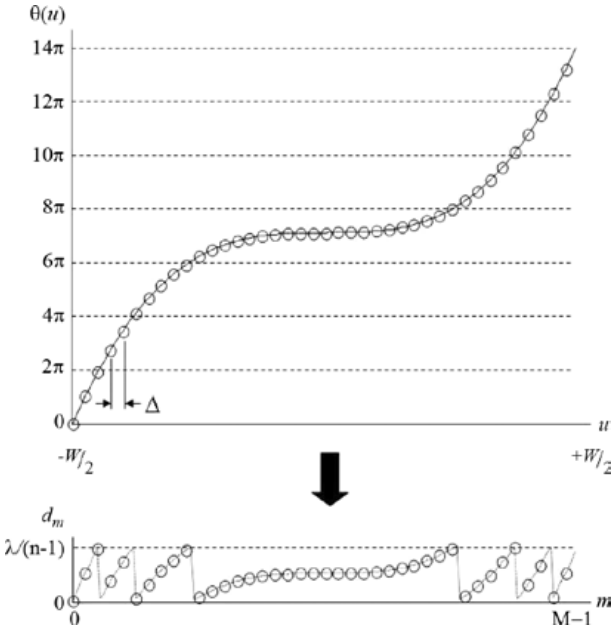


Fig. 6. Representation of the process for converting the cubic phase to a surface depth profile. Only a one-dimensional representation is shown but extensions to two-dimensions are straightforward.

Section V the processing required by the cubic phase system to produce a well-defined spot.

III. Cubic Phase Design and Fabrication

We fabricated a cubic phase element with $\alpha = 7$ from Rexolite using a 3-axis computerized numerical control router. The router has a 0.1969"-minimum feature size (5 mm) and provides 0.0002" ($5\ \mu$m) position accuracy. Element fabrication was a two-step process.

In the first step we machined a continuous surface profile by sampling the cubic phase $\theta_c(u,v)$, converting phase to depth, and using a cubic spline to insure a smooth transition between depth samples. The phase samples were generated according to

$$\theta_{\ell m} = \sum_{\ell=0}^{L-1}\sum_{m=0}^{M-1} \theta_c\left(u - \frac{W_u}{2}, v - \frac{W_v}{2}\right)\delta\left(u - \frac{m}{\Delta}, v - \frac{\ell}{\Delta}\right). \tag{18}$$

We used a sampling distance $\Delta = 0.1''$ (2.54 mm) to insure overlapping features. Given $W_u = W_v = 24''$, $L = M = 240$.

Phase values were converted to depth $d_{\ell m}$ using

$$d_{\ell m} = \left[\frac{\lambda}{2\pi(n-1)}\right][\theta_{\ell m} \text{ modulo } 2\pi]. \tag{19}$$

The modulo-2π operator limits the phase to only a single wavelength and reduces element weight. See Fig. 6. The diffractive characteristics introduced in this conversion have little effect on the response of the element [5], [6]. At 94 GHz Rexolite has a refractive index $n = 1.59$; thus, a depth change of 0.2129" (5.4 mm) in the material generates a 2π-phase change in the wavefront. The second step sharpened the edges at phase discontinuities to within < 0.1 mm. The final element is shown in Fig. 7 mounted to the antenna.

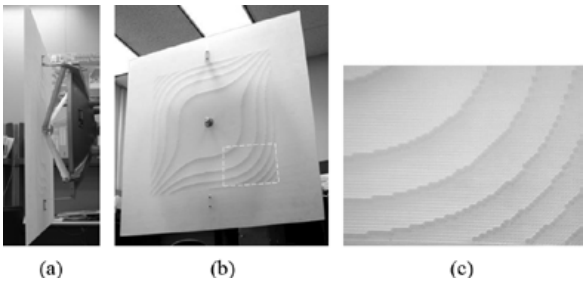


Fig. 7. Fabricated cubic phase element. (a) Side-view and (b) front-view of cubic phase element mounted to imaging system. (c) Detail of fabricated element. The region displayed is in the lower right of the phase element, which is highlighted in (b).

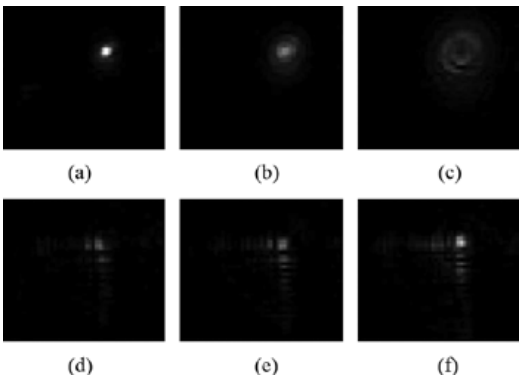


Fig. 8. Measured point spread functions for conventional imaging and imaging with a cubic phase. PSFs for conventional imaging system at (a) 180", (b) 146.5" and (c) 113". (d)-(f) PSFs for a system with cubic phase at the same object distances for (a)-(c).

IV. Experimental Results

To validate the performance of the cubic phase element to extend DoF, we measured the PSF of a conventional system and the cubic-based system at three distances, 113" (2870 mm), 146.5" (3721 m), and 180" (4572 mm). Since the DoF is asymmetric with respect to the object plane and collapses more quickly as the object plane moves toward the system, we measured only this behavior. The out-of-focus object planes correspond to displacements that are twice and four times the calculated DoF. We also imaged an extended object at the same distances using both systems.

To measure the PSF we imaged a point source generated by an open waveguide with dimensions 0.050" $\times$ 0.100". Given that the operating frequency was 94 GHz, the aperture in wavelengths is $0.40\lambda \times 0.80\lambda$. The output power was -14 dBm. The results are represented in Fig. 8. The experimental results agree qualitatively with the simulations presented in Fig. 5. The figures are normalized to the peak value measured, which occurs in Fig. 8(a).

The extended object used in our experiments is represented in Fig. 9(a). The spoke pattern produces 50–50 square waves whose frequencies vary linearly from low values at the circumference to high values in its center. Given that the pattern con-

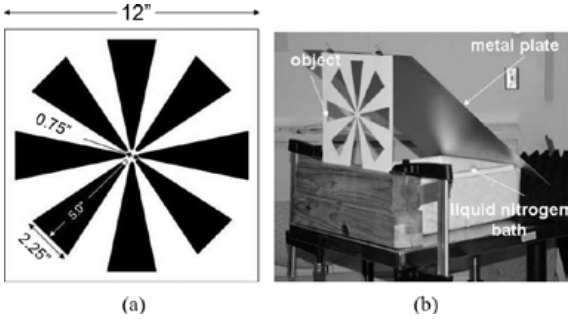


Fig. 9. (a) Representation of the extended object used to compare conventional and cubic-phase imaging. (b) Schematic of object illumination.

TABLE I
PARAMETERS FOR EXTENDED OBJECT IMAGING

d_o (inches)	M	ρ_{min}	ρ_{max}	r_c/r_{max}
180.0	0.0334	9.7127	139.4409	0.5909
146.5	0.0413	7.8448	112.6238	0.4772
113.0	0.0542	5.9768	85.8067	0.3636

tains 8 periods within one rotation, the period S of the imaged square wave as a function of radius r is

$$S = \frac{2\pi r M}{8} \tag{20}$$

and the corresponding spatial frequency ρ is

$$\rho = \frac{1}{S\lambda f}. \tag{21}$$

The maximum radius $r_{\max} = 5.38''$ and the minimum $r_{\min} = 0.38''$. Since the cutoff frequency for a circular aperture ρ_c is

$$\rho_c = \frac{D}{2\lambda f} \tag{22}$$

the corresponding radius r_c is determined by equating (21) and (22)

$$r_c = \frac{8}{D\pi M}. \tag{23}$$

Radii within the range $r_{\max} \geq r \geq r_c$ generate frequencies in the passband of the optical system. Higher frequencies contained in the region $r_c \geq r \geq r_{\min}$ are cutoff. Table I lists the magnification, minimum and maximum spatial frequencies, and the cutoff radius as a fraction of the maximum radius for the three object distances used in the experiments. We note that beyond 304" the system is incapable of resolving the target at all.

The extended object was generated by placing a metal plate cut-out of Fig. 9(a) in front of a metal reflector angled at 45° to a bath of liquid nitrogen. See Fig. 9(b). This arrangement produces a contrast between the surrounding metal reflecting ambient room temperature and the temperature of the liquid nitrogen reflected through the cut-out. The images captured by the system are represented in Fig. 10.

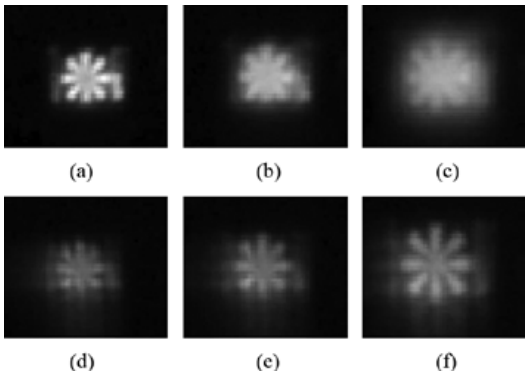


Fig. 10. Images of an extended object for conventional imaging and imaging with a cubic phase. Images from a conventional imaging system at (a) 180", (b) 146.5" and (c) 113". (d)-(f) Images from a system with cubic phase at the same object distances for (a)-(c).

Note that the MTF of the conventional system produces images with significant high frequency loss. In contrast, the entire band of frequencies between $\rho_{\min}$ and ρ_c can be seen in the images captured using a cubic phase element. Even without signal processing these images retain more discernable characteristics of the spoke target than the image from the conventional system.

V. POST-DETECTION SIGNAL PROCESSING AND RESULTS

Removing the artifacts of the aberrations introduced by the cubic phase element requires post-detection electronic processing. We assumed a linear process

$$i_p(x,y) = |i\,(x,y)|^2 * *w(x,y) \tag{24}$$

and implemented $w(x,y)$ as a Wiener filter in Fourier space

$$W(u,v) = \frac{H_c^*(u,v)}{|H_c(u,v)|^2 + \frac{K^{-2}\hat{\Phi}_N(u,v)}{\hat{\Phi}_L(u,v)}}. \tag{25}$$

The optical transfer function (OTF) $H_c(u,v)$ associated with the cubic phase element was estimated from the experimentally measured point response images. The parameter K is a measure of the signal-to-noise ratio. The functions $\hat{\Phi}_L(u,v)$ and $\hat{\Phi}_N(u,v)$ are the expected power spectra of the object and noise, respectively. Research has shown that Wiener power spectra are good assumptions for natural scenes [13]. We adjusted the mean spatial detail parameter to produce restored PSFs with widths comparable to that of the experimentally measured focused point. Further, we assumed a flat noise spectrum with $K = 50$.

Reconstructed PSFs are represented in Fig. 11. In Fig. 11(a)–(c), the cubic OTF was estimated from the PSF measured experimentally at 146.5" and the subsequent reconstruction filter applied to all the images. In Fig. 11(d)–(f),

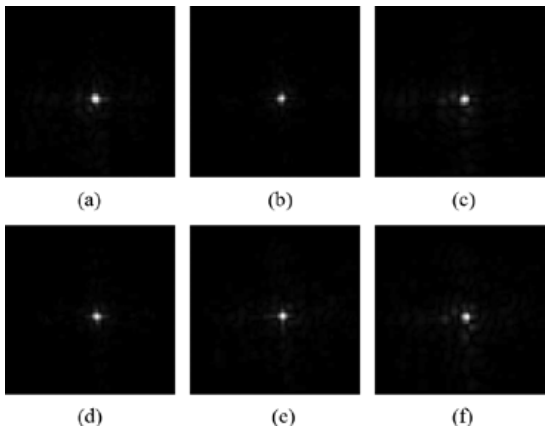


Fig. 11. Processed point responses from a system with cubic phase. Images processed using the PSF for $d_o = 146.5''$ at (a) 180", (b) 146.5" and (c) 113". (d)-(f) Images processed using the PSF for $d_o = 180''$.

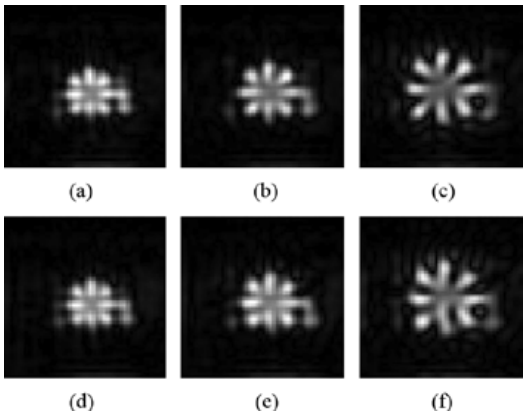


Fig. 12. Processed images from a system with cubic phase. Images processed using the PSF for $d_o = 146.5''$ at (a) 180", (b) 146.5" and (c) 113". (d)-(f) Images processed using the PSF for $d_o = 180''$.

the cubic OTF was estimated from the experiment PSF at 180". From Fig. 4, data collected at 113" has significant high frequency loss in comparison to that collected at both 146.5" and 180". We, therefore, did not use this data to create a reconstruction filter.

In both instances the best reconstruction occurs for the image matched to the filter and the noisest reconstruction is at 113", which one would expect. Nonetheless, the reconstructions of the point response are comparable in terms of spatial scale to the in-focus response of the conventional system. Compare Fig. 11(a), (b), (d), and (e) to Fig. 8(a). Thus, we have extended the region over which the system generates a diffraction-limited spot from 5" in front of the focal plane to 34". Since we expect similar behavior for objects beyond the focal plane, the depth of field has been expanded in that direction by at least 34" as well but we expect considerably more.

This behavior is reflected also in the reconstruction of the extended object represented in Fig. 12. For both reconstructions, not only are the images at 180" comparable to the focused image from the conventional system but the images at 146.5" are comparable as well. In comparison to the image at 146.5" from the conventional system, the reconstructed images display higher contrast and higher resolution. Note especially the reflection off the metal plate on the right-hand side, which is clearly visible in the reconstructions but is apparent only in the focused conventional image in Fig. 10.

VI. Summary

We applied a computational imaging technique for extending depth-of-field at optical frequencies to a millimeter wave imaging system. The technique requires inserting a cubic phase element in the pupil plane of an imaging system and subsequent post-detection signal processing. We designed and fabricated the cubic phase element in Rexolite and validated its performance experimentally.

In some applications, *a priori* range information can be used to improve estimated PSFs and thereby improve restoration. Further, non-linear image restoration techniques incorporating *a priori* knowledge of the scene can improve restoration relative to linear restoration.

A critical difference between the performance of millimeter wave imaging systems and imaging systems for optical and infrared wavelengths is the underlying phenomonology and availability of technology, especially detector arrays. Millimeter wave systems image temperature contrasts. Careful analysis of noise and contrast in such systems is necessary to assess the impact of inserting an element into the optical train whose amplitude transfer function, although flat, is relatively low. A more in-depth analysis should also consider the coherence and spectral bandwidth of the illumination.

In addition, in terms of practical application, one needs to consider the scale of the optical system and the lack of large arrays of millimeter wave detectors on system design. Whereas one can design an optical staring imager, millimeter wave systems will continue to be scanning systems until detector array technology matures. Even so, given the physical constraint on the security system mentioned in the introduction, it is unlikely that a such system will have a detector array larger than 200×200. Nonetheless, the applicability and advantage of computational imaging techniques to millimeter wave systems has been demonstrated.

References

[1] A. Luukanen, A. J. Miller, and E. N. Grossman, "Passive hyperspectral terahertz imagery for security screening using a cryogenic microbolometer, Passive millimeter-wave imaging technology VIII," in *Proc. SPIE*, R. Appleby and D. A. Wikner, Eds. Bellingham, WA: SPIE, 2005, vol. 5789, pp. 127–134.

[2] L. Yujiri, M. Shoucri, and P. Moffa, "Passive millimeter-wave imaging," *IEEE Microw. Mag.*, vol. 4, pp. 39–50, Sep. 2003.

[3] R. Appleby and H. B. Wallace, "Standoff detection of weapons and contraband in the 100 GHz to 1 THz region," *IEEE Trans. Antennas Propag.*, vol. 55, pp. 2944–2956, Nov. 2007.

[4] E. R. Dowski , Jr and W. T. Cathey, "Extended depth of field through wave-front coding," *Appl. Opt.*, vol. 34, pp. 1859–1866, 1995.

[5] J. van der Gracht, E. R. Dowski, W. T. Cathey, and J. Bowen, "Aspheric optical elements for extended depth of field, novel optical system design and optimization," in *SPIE Proc.*, J. M. Sasian, Ed. Bellingham, WA: SPIE, 1995, vol. 2537, pp. 279–288.

[6] J. van der Gracht, E. R. Dowski , Jr, M. G. Taylor, and D. M. Deaver, "Broadband behavior of an optical-digital focus-invariant system," *Opt. Lett.*, vol. 21, pp. 919–921, 1996.

[7] S. Bradburn, W. T. Cathey, and E. R. Dowski, "Realizations of focus invariance in optical-digital systems with wave-front coding," *Appl. Opt.*, vol. 36, pp. 9157–9166, 1997.

[8] W. T. Cathey and E. R. Dowski, "New paradigm for imaging systems," *Appl. Opt.*, vol. 41, pp. 6080–6092, 2002.

[9] R. Athale, D. M. Healy, D. J. Brady, and M. A. Neifeld, "Reinventing the camera," *Opt. Photon. News*, vol. 19, no. 3, pp. 32–37, Mar. 2008.
[10] D. A. Wikner and G. Samples, "Polarimetric passive millimeter-wave sensing, passive millimeter-wave imaging technology V," in *SPIE Proc.*, R. M. Smith and R. Appleby, Eds. Bellingham, WA: SPIE, 2001, vol. 4373, pp. 86–93.
[11] J. W. Goodman, *Introduction to Fourier Optics*, 3rd ed. Englewood, CO: Roberts and Company, 2005.
[12] F. L. Pedrotti and L. S. Pedrotti, *Introduction to Optics*. Englewood Cliffs, NJ: Prentice-Hall, 1987, p. 125.
[13] C. L. Fales, F. O. Huck, and R. W. Samms, "Imaging system design for improved information capacity," *Appl. Opt.*, vol. 23, pp. 872–888, 1984.

Joseph N. Mait (S'78–M'84–SM'03) received the B.S.E.E. degree from the University of Virginia, Charlottesville, in 1979 and the M.S.E.E. and Ph.D. degrees from the Georgia Institute of Technology, Atlanta, in 1980 and 1985, respectively.

Since 1989, he has been with the U.S. Army Research Laboratory, Adelphi, MD, where he is currently Senior Technical Researcher (ST) for Electromagnetics. He was an Assistant Professor of electrical engineering at the University of Virginia from 1984 to 1989 and an Adjunct Associate Professor at the University of Maryland, College Park, from 1997 to 2005. He has also held visiting positions at the Lehrstuhl für Angewandte Optik, Universität Erlangen-Nürnberg, Germany, and the Center for Technology and National Security Policy, National Defense University, Washington, DC. His research interests include the application of optics, photonics, and electromagnetics to sensing and sensor signal processing.

Dr. Mait is a Fellow of SPIE and OSA.

David A. Wikner received the B.A. degree in physics from Ohio Wesleyan University, Delaware, in 1986 and the M.S. degree in physics from Stanford University, Palo Alto, CA, in 1988.

Currently, he leads the Millimeter Wave Phenomenology Team, Army Research Laboratory, Adelphi, MD. He is responsible for the exploration of MMW imaging techniques and signature analysis, system design, and new device technology. He has worked extensively with 94-GHz radar systems for vehicle and ground clutter measurements. He has also designed and built a W-band Stokes-vector radiometer for fundamental phenomenology measurements. He has published 20 papers on the subject of MMW phenomenology and technology. In the last several years, his work has focused on efforts to create affordable MMW imaging systems for U.S. Army applications and techniques for creating enhanced imaging through computational optics.

Mark S. Mirotznik (S'87–M'92) received the B.S.E.E. degree from Bradley University, Peoria, IL, in 1988 and the M.S.E.E. and Ph.D. degrees from the University of Pennsylvania, Philadelphia, in 1991 and 1992, respectively.

Since 1992, he has been a faculty member in the Department of Electrical Engineering, The Catholic University of America, Washington, DC, where, in 1997, he was promoted to Associate Professor. He also holds the position of Senior Research Engineer for the Naval Surface Warfare Center (NSWC), Carderock Division. His research interests include applied electromagnetics and photonics, computational electromagnetics and bioelectromagnetics.

Joseph van der Gracht received the Ph.D. degree in electrical engineering with an emphasis in modern optics from the Georgia Institute of Technology, Atlanta, in 1991.

Since beginning his professional career in 1979, he has worked for Texas Instruments, Mead and Rockwell International on a variety of optoelectronics related technologies. In 1991, he was a Principal Investigator at the U.S. Army Research Laboratory, Adelphi, MD. He is currently a Senior Engineer at HoloSpex, Inc., Columbia, MD, responsible for digital holography and imaging. His main research efforts are directed toward investigating new imaging modalities that optimize the information content of a scene through the use of novel optical elements, unorthodox detector geometries and post-detection signal processing.

Gregory P. Behrmann received the B.S.M.E. degree from Drexel University, Philadelphia, PA and the M.S.M.E. degree from The Johns Hopkins University, Baltimore, MD. He is currently working toward the Ph.D. degree at the Catholic University of America, Washington, DC.

Previously, he has held senior engineering positions at the U.S. Army Research Laboratory, Potomac Photonics, Inc., and Corvis Corporation.

Brandon L. Good is currently working toward the M.S. degree at the Catholic University of America, Washington, DC.

Scott A. Mathews (M'06) received the B.S. degree in physics, the M.S. degree in materials engineering, and the Ph.D. degree in materials engineering from the University of Maryland, College Park, in 1988, 1993, and 2001, respectively.

In 2003, he joined the Electrical Engineering Department, Catholic University of America, Washington, DC, where he is an Assistant Professor. Prior to joining Catholic University, he did research and development in the private sector for 15 years, including work on solid state radiation imaging systems, optical data storage systems, and laser micro-fabrication of microelectronic and microfluidic structures. He is interested in miniature, electronic detection systems for chemical and biological warfare agents.

Experimental Demonstration of Coherent Beam Combining Over a 7 km Propagation Path

Thomas Weyrauch, Mikhail A. Vorontsov, Gary W. Carhart, Leonid A. Beresnev, Andrey P. Rostov, Ernst E. Polnau and Jony Jiang Liu

Optics Letters 2011, 36 (22), 4455-4457.

Experimental demonstration of coherent beam combining over a 7 km propagation path

Thomas Weyrauch,[1,*] Mikhail A. Vorontsov,[1,3] Gary W. Carhart,[2] Leonid A. Beresnev,[2] Andrey P. Rostov,[3] Ernst E. Polnau,[1] and Jony Jiang Liu[2]

[1]*Intelligent Optics Laboratory, School of Engineering, University of Dayton, 300 College Park, Dayton, Ohio 45469–2951, USA*

[2]*Intelligent Optics Laboratory, Computational and Information Sciences Directorate, United States Army Research Laboratory, 2800 Powder Mill Road, Adelphi, Maryland 20783, USA*

[3]*Optonicus, 711 East Monument Avenue, Suite 101, Dayton, Ohio 45402, USA*

**Corresponding author: thomas.weyrauch@udayton.edu*

Received August 9, 2011; revised October 11, 2011; accepted October 12, 2011; posted October 13, 2011 (Doc. ID 152671); published November 15, 2011

We demonstrate coherent combining (phase locking) of seven laser beams emerging from an adaptive fiber-collimator array over a 7 km atmospheric propagation path using a target-in-the-loop (TIL) setting. Adaptive control of the piston and the tip and tilt wavefront phase at each fiber-collimator subaperture resulted in automatic focusing of the combined beam onto an unresolved retroreflector target (corner cube) with precompensation of quasi-static and atmospheric turbulence-induced phase aberrations. Both phase locking (piston) and tip–tilt control were performed by maximizing the target-return optical power using iterative stochastic parallel gradient descent (SPGD) techniques. The performance of TIL coherent beam combining and atmospheric mitigation was significantly increased by using an SPGD control variation that accounts for the round-trip propagation delay (delayed SPGD).

OCIS codes: 140.3298, 010.1285, 010.1080.

Coherent combining of laser beams that originate from a fiber-based multichannel master oscillator power amplifier (MOPA) laser system at a remotely located target after propagation through the atmosphere requires adaptive compensation of both random phase shifts introduced by the MOPA system and atmospheric turbulence-induced phase aberrations [1,2]. Coherent beam combining, also referred to as phase locking, has been demonstrated in several laboratory-based experiments (see, e.g., [3–7]) and over a 408 m long distance in an outdoor experiment with a cooperative target [8].

In this Letter, we report the results of the first (to our best knowledge) successful coherent beam combining and turbulence mitigation experiments over an extended-length atmospheric propagation path in a target-in-the-loop (TIL) setting with a noncooperative target using adaptive control of the piston (subaperture-averaged phase) and tip and tilt corrections at each fiber-array subaperture. The round-trip propagation delay issue—a major obstacle for TIL adaptive optics techniques—was overcome by utilizing the recently proposed "delayed" stochastic parallel gradient descent (SPGD) wavefront control technique [9], which allowed the duration between wavefront control updates to be shorter than the round-trip propagation delay and resulted in a significant increase of compensation bandwidths.

The setup used in the experiments (Fig. 1) consists of the following major subsystems: (i) a seven-channel master oscillator power amplifier (MOPA) system based on single-mode, polarization-maintaining (PM) fiber elements; (ii) a fiber-collimator array with built-in capabilities for electronic control of wavefront phase tip and tilts at each fiber-collimator subaperture; (iii) an unresolved target (a corner-cube retroreflector) located at 7 km distance; (iv) a receiver telescope for measurements of the target-return optical wave power, referred to as the power-in-the-bucket (PIB) metric, J; and (v) a control unit that includes piston (phase-locking) and tip–tilt phase control subsystems.

In the MOPA system, the light from a narrow-linewidth (~5 kHz) fiber laser with wavelength $\lambda = 1064$ nm and single-mode PM fiber output is divided into seven channels using a fiber splitter with integrated, electrically controlled phase-shifting elements from EOSPACE [10]. The MOPA system output fibers, each with a mode field diameter of 7 μm, are connected to a fiber-collimator array (Optonicus INFA 7C [11]). In the fiber array, the tip of each output fiber is placed in the focus of the corresponding collimating aspheric lens with a clear aperture diameter of $d = 33$ mm and a focal distance of $f = 174$ mm. The closest center-to-center distance between the collimating lenses in the array is 37 mm, and the entire fiber-array aperture is 107 mm. The output fibers are mounted inside special fiber-positioner devices with piezo-actuators that can independently displace the fiber tips within a $\pm 35\,\mu$m range in two lateral directions [11,12]. These fiber-tip displacements result in controllable deviations of the propagation directions of the outgoing beams anywhere within a ± 0.2 mrad solid angle about the optical axis and were used to provide precise overlapping of the outgoing beams at a remote target (electronic beam focusing) as well as precompensation of wavefront phase tip and tilt static and dynamic aberrations [4].

The outgoing beams with a combined optical power of 12 mW were transmitted through a window located in the Intelligent Optics Laboratory at the fifth floor of the University of Dayton's College Park Center building (15 m above ground) and propagated toward the corner-cube retroreflector (50 mm aperture) located in a shed on the rooftop of a 40 m high building 7 km away. The laboratory double-glass window introduced significant phase aberrations with a peak-to-valley (PV) amplitude of $\sim 1.0\lambda$ over the fiber-array aperture and $\sim\lambda/4$ PV over fiber-array subapertures. The impact of these quasi-static

aberrations was partially mitigated using the adaptive tip–tilt control system.

The optical wave returning from the target entered a receiver telescope (aperture 20 cm) located near the fiber-array transmitter, as shown in Fig. 1(b). The received light was divided between a CCD camera and a photodiode for telescope pointing (target imaging) and the received light power measurements, respectively. The photodiode output signal was used as the performance metric, J, for both the phase locking and the tip–tilt control subsystems.

The two parallel operating control subsystems were both based on the maximization of the PIB metric, J, using an asynchronous SPGD control technique with significantly different (~48 times) iteration rates [13]. The tip–tilt control subsystem with 14 control channels (two per fiber collimator) utilized a personal computer with analog input/output cards and a set of high-voltage amplifiers (±70 V) for generation of the control voltages $\{u_j^{(n)}\}$ $(j = 1, \ldots, 14)$, which were applied to the piezo-actuators. At each tip–tilt iteration, (n), a control voltage update was performed using the conventional SPGD algorithm [13]

$$u_j^{(n+1)} = u_j^{(n)} + \gamma \delta u_j^{(n)} [J_+^{(n)} - J_-^{(n)}], \qquad (1)$$

where γ is the gain coefficient and $\{\delta u_j^{(n)}\}$ is a set of 14 small-amplitude random control voltage changes, denoted as perturbations. The perturbations in the form $\{\delta u_j^{(n)}\}$ (positive) and $\{-\delta u_j^{(n)}\}$ (negative) are applied between two sequential updates of the control voltages. In Eq. (1), $J_+^{(n)}$ and $J_-^{(n)}$ are the measured PIB metric values that correspond to the positive and negative perturbations. The characteristic time τ_{SPGD} (SPGD cycle time) between sequential control voltage updates is given by

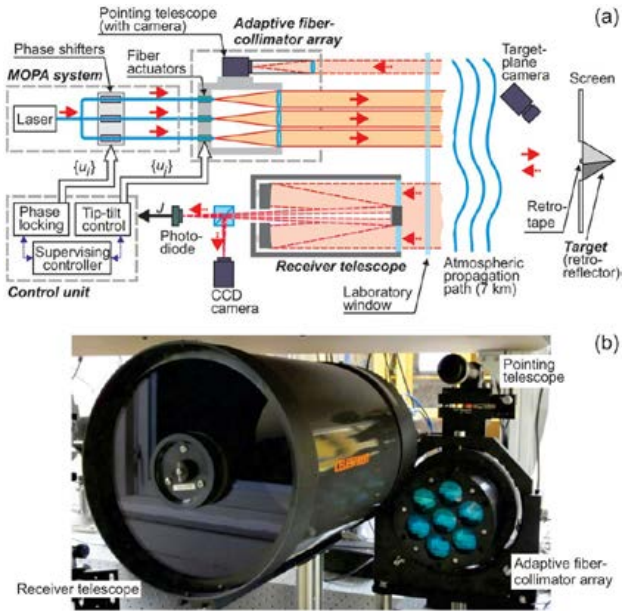


Fig. 1. (Color online) (a) Schematic of the experimental setup used for coherent beam combining over a 7 km atmospheric propagation path. (b) Photo of the fiber-array transmitter with the pointing telescope (right) and the receiver telescope (left).

$$\tau_{\text{SPGD}} = 2(\tau_{\text{pert}} + \tau_{\text{resp}} + \tau_J + \tau_{\text{delay}}), \qquad (2)$$

where τ_{pert} is the time required to perturb the control voltages, τ_{resp} is the delay between a control voltage change and the corresponding optical phase response, τ_J is the PIB metric measurement time, and τ_{delay} is the delay between an induced wavefront phase variation and the corresponding metric change. The last term in Eq. (2) is the double-pass delay $\tau_{\text{delay}} = 2L/c$ caused by the optical wave propagation over the distance L with the speed of light, c (in the experimental setting $L = 7$ km and $\tau_{\text{delay}} = 46.7\,\mu$s). The tip–tilt SPGD cycle time, τ_{SPGD} in Eq. (2), is mainly limited by the time response of the piezo-actuators, $\tau_{\text{resp}} \approx 120\,\mu$s, which is significantly longer than τ_{delay} and $\tau_J \approx 20\,\mu$s. The resulting tip–tilt subsystem SPGD iteration rate $f_{\text{SPGD}} = 1/\tau_{\text{SPGD}}$ was $f_{\text{SPGD}} \approx 3$ kHz.

The piston phase control subsystem utilized the fiber-integrated phase shifters of the MOPA system, which have a short response time of $\tau_{\text{resp}} < 10$ ns so that the limiting factor for increasing the SPGD control iteration rate is the double-pass delay time τ_{delay}. Considering $\tau_{\text{delay}} = 46.7\,\mu$s, the piston-phase control SPGD cycle time is at least ~100 μs and thus $f_{\text{SPGD}} \le 10$ kHz. Note that the SPGD + CU 8D controller from Optonicus used in the experiments can provide much higher iteration rates (up to ~250 kHz) [10]. Therefore, the propagation delay imposed the limit on the operational bandwidth of the conventional SPGD-based piston-phase control subsystem and its capability for mitigation of atmospheric turbulence-induced aberrations.

In order to overcome this problem, we utilized in the piston-phase control subsystem the recently proposed delayed-SPGD wavefront control algorithm, where the iterative procedure of the control voltage update during each iteration cycle (n) can be described by the following rule [9]:

$$u_i^{(n+1)} = u_i^{(n)} + \gamma [J_+^{(n)} - J_-^{(n)}] \delta u_i^{(n-\Delta n)}, \qquad (i = 1, \ldots, 7). \qquad (3)$$

Here the integer number $\Delta n > 0$ is the delay parameter that accounts for the double-pass propagation time. In Eq. (3), Δn links the variation of the metric $\delta J^{(n)} = [J_+^{(n)} - J_-^{(n)}]$ measured during iteration (n) to the control signal perturbations $\{\delta u_i^{(n-\Delta n)}\}$, which caused the metric change. The delay parameter can be calculated as the closest integer number to the ratio $\tau_{\text{delay}}/\tau_{\text{SPGD}}$. With the SPGD cycle time $\tau_{\text{SPGD}} = 7.0\,\mu$sec (iteration rate $f_{\text{SPGD}} \approx 143$ kHz) and $\tau_{\text{delay}} = 46.7\,\mu$s, we obtain $\Delta n = 7$.

During the experiments, the fiber-collimator array control system was repeating 50 sequences of 5.25 s long trials comprising three operational states of 1.75 s each. These stages are indicated in Fig. 2 as "feedback off" (all control loops were off), "piston control on," and "piston and tip–tilt control on." In the "piston control on" state, the piston-phase (phase-locking) control system was turned on. During the last state, both the piston and tip–tilt control subsystems were switched on. Values for the PIB metric, J, were recorded for all 50 trials by the supervising controller at a rate of about 10 k samples/s.

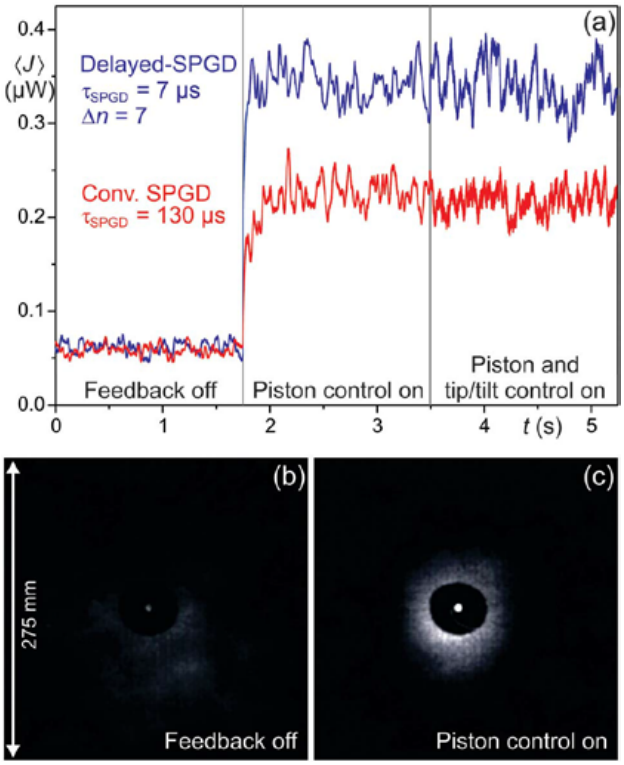


Fig. 2. (Color online) Experimental results from the coherent beam combining experiment: (a) average PIB metric evolution curve, $\langle J \rangle$ and (b), (c) averaged irradiance distribution at the target plane with feedback off (b) and piston control on (c).

As shown in Fig. 1(a), the retroreflector at the target plane was mounted behind a hole in a cardboard screen and a small piece of retroreflecting tape (~6 mm diameter) was glued onto the center of the retroreflector's cover glass. An near-IR camera with a wide-angle objective was placed about 1 m in front and 20 cm to the side of the retroreflector and used to image the irradiance pattern (beam footprint) on the screen and the retrotape.

Figure 2(a) shows the time dependence of the trial-averaged PIB metric $\langle J \rangle$ for two different settings of the piston-phase controller: the first curve (red, lower) corresponds to the conventional SPGD algorithm (1) and the second curve (blue, upper) to the delayed algorithm (3). In comparison to the open loop state, the average PIB metric, $\langle J \rangle$, increased 3.7-fold for the conventional and 5.6-fold for the delayed-SPGD control.

Recorded target-plane beam footprints (averages of 270 frames) for the cases with feedback off and piston-phase control on can be seen in Figs. 2(b) and 2(c), respectively. The dark annular region in the center corresponds to the circular opening for the retroreflector with the retrotape spot in the center. A comparison of these two images demonstrates the higher concentration of the beam energy at the retroreflector when phase control is on and proves that the PIB metric maximization locks the beam phases at the target plane.

The experimental results in Fig. 2(a) correspond to atmospheric turbulence conditions characterized by a path-averaged refractive index structure parameter $C_n^2 = 6 \times 10^{-16}\,\mathrm{m}^{-2/3}$ (measured by a Scintec BLS2000 scintillometer [14]) and a normalized standard deviation of metric fluctuations $\sigma_J/\langle J \rangle = 0.92$ (open loop). Piston control resulted not only in the increase of the average metric value, but also led to a decrease in the metric fluctuation level down to $\sigma_J/\langle J \rangle = 0.52$ for the conventional and to 0.42 for the delayed SPGD controllers.

Note that the tip–tilt control subsystem, which was turned on during the last state of the adaptation trials, did not result in a further metric increase (and caused only a slight change in metric fluctuations due to the tip–tilt perturbations). This can be explained by taking into account the 48-fold faster updates of the piston-phase control system, which can provide a partial mitigation of overall wavefront phase tip and tilt aberrations using a stepwise (piston) approximation prior to a reaction of the tip–tilt subsystem. However, our experiments showed that efficient coherent combining with piston-phase control was only possible if the transmitted beams overlap well at the target, which was achieved by turning on the tip–tilt control subsystem for a few seconds in addition to piston control. In the experiments described above, tip–tilt control voltages were fixed at the end of each adaptation trial and provided sufficient overlapping during the piston control stage of the next adaptation cycle. Without a tip–tilt control phase in each trial, we observed a slow (on the order of 100–200 s) decline in coherent beam combining efficiency, indicating that static tip–tilt control voltages do not maintain efficient overlapping of the outgoing beams at the target over a longer time period, mostly due to thermal expansion-induced system misalignments.

This work was performed in the frame of collaborative agreement W911NF-09-2-0040 between the United States Army Research Laboratory and the University of Dayton.

References

1. M. A. Vorontsov and S. L. Lachinova, J. Opt. Soc. Am. A **25**, 1949 (2008).
2. S. L. Lachinova and M. A. Vorontsov, J. Opt. Soc. Am. A **25**, 1960 (2008).
3. H. Bruesselbach, S. Wang, M. Minden, D. C. Jones, and M. Mangir, J. Opt. Soc. Am. B **22**, 347 (2005).
4. M. A. Vorontsov, T. Weyrauch, L. A. Beresnev, G. W. Carhart, L. Liu, and K. Aschenbach, IEEE J. Sel. Top. Quantum Electron. **15**, 269 (2009).
5. V. Jolivet, P. Bourdon, B. Bennaï, L. Lombard, D. Goular, E. Pourtal, G. Canat, Y. Jaouën, B. Moreau, and O. Vasseur, IEEE J. Sel. Top. Quantum Electron. **15**, 257 (2009).
6. X. L. Wang, Y. X. Ma, P. Zhou, H. T. Ma, X. Li, X. X. Xu, and Z. J. Liu, Laser Phys. **19**, 984 (2009).
7. B. Pulford, T. Shay, J. Baker, A. Flores, C. Robin, C. Vergien, C. Zeringue, D. Gallant, A. D. Sanchez, C. Lu, and A. Lucero, in *Conference on Lasers and Electro-Optics*, OSA Technical Digest (CD) (Optical Society of America, 2010), paper CThO5.
8. H. Bruesselbach, S. Wang, M. Minden, C. Jones, and M. Mangir, in *2005 Conference on Lasers & Electro-Optics* (IEEE, 2005), Vol. 1, pp. 746–748.
9. M. A. Vorontsov and G. W. Carhart, Invention disclosure ARL No. 12-03 (U.S. Army Research Laboratory, 2011).
10. www.eospace.com.
11. www.optonicus.com.
12. L. A. Beresnev and M. A. Vorontsov, Proc. SPIE **5895**, 58950R (2005).
13. M. A. Vorontsov and G. W. Carhart, J. Opt. Soc. Am. A **23**, 2613 (2006).
14. www.scintec.com.

Turbulence-Free Ghost Imaging

Ronald E. Meyers, Keith S. Deacon and Yanhua Shih

Applied Physics Letters 2011, 98 (11).

Turbulence-free ghost imaging

Ronald E. Meyers,[1,a)] Keith S. Deacon,[1] and Yanhua Shih[2]
[1]*US Army Research Laboratory, Adelphi, Maryland 20783, USA*
[2]*Department of Physics, University of Maryland, Baltimore County, Baltimore, Maryland 21250, USA*

(Received 11 January 2011; accepted 25 February 2011; published online 18 March 2011)

Atmospheric turbulence is a serious problem for satellite and aircraft-to-ground based classical imaging. Taking advantage of the natural, nonfactorizable, point-to-point correlation of thermal light, this experiment demonstrated turbulence-free ghost imaging, which will be extremely useful for these applications. In addition, this observation suggests that the nontrivial intensity-intensity correlation of thermal light cannot be caused by the statistical correlation of intensity fluctuations.
 [doi:10.1063/1.3567931]

One of the most surprising consequences of quantum mechanics is the nonlocal correlation of a multi-particle system measured by joint-detection of distant particle detectors. Ghost imaging is one such phenomena. Recently Meyers *et al.*[1,2] performed ghost imaging of remote objects by measuring reflected photons, thereby making ghost imaging practical for applications. Two types of ghost imaging have been experimentally demonstrated since 1995. One type of ghost imaging uses entangled photon pairs as the light source[3] and another type of ghost imaging uses chaotic thermal light.[1,4] The nonlocal multiphoton interference nature of ghost imaging determines its peculiar features: (1) it is nonlocal; (2) its imaging resolution differs from that of classical.

Recently there has been increased interest in ghost imaging through turbulence and related index of refraction distortions as demonstrated by theoretical[5] and experimental papers.[2,6] However, thermal light ghost imaging through turbulence has not been previously reported. In this letter, we wish to report a recent ghost imaging experiment with thermal light which demonstrated another peculiar yet useful feature of ghost imaging: "turbulence-free," i.e., any index of refraction fluctuation of turbulence in the optical path will not affect the quality of the ghost image. This important feature will be useful for applications like satellite and aircraft-to-ground based distant imaging, for which atmospheric turbulence is a serious problem. We present the main result from the ghost imaging experiment performed at the US Army Research Laboratory (ARL), namely that the ghost image is virtually free of the adverse effects of turbulence. We also highlight the two-photon interference nature of the ghost imaging as the primary cause of this turbulence-free effect. We expand on the effects of turbulence on ghost imaging in another letter.[7]

A schematic of the experimental setup is shown in Fig. 1. It is a typical thermal light lensless ghost imaging setup,[1] except for the addition of heating elements to produce laboratory atmospheric turbulence. It uses secondary ghost imaging which helps make ghost imaging practical for applications. In this experiment, turbulence is introduced by adding heating elements at 550 °C underneath any or all optical paths as illustrated in Fig. 1. Heating of the air causes temporal and spatial fluctuations on its index of refraction that makes the classical image of the object jitter about randomly on the image plane causing a "blurred" picture. As in our early experiment,[1] the light source is a typical chaotic pseudo-thermal source, which contains a laser beam and a fast rotating ground glass diffuser. The chaotically scattered laser beam, with a fairly large size (11 mm diameter) in transverse dimension, is split into two by a 50%–50% beam-splitter. One of the beams illuminates an object located at z_1, such as the letters "ARL" shown in Fig. 1. The photons scattered and reflected from the object are collected and counted by a "bucket" detector, which is simulated by the right-half of the charged coupled device (CCD) in Fig. 1. The other beam propagates to the ghost image plane of $z_2 = z_1 \simeq 1.4$ m and the path from the target to the detectors over heating elements is ~1.7 m. Like our early demonstration of ghost imaging, placing a CCD array on the ghost image plane, allows it to capture the ghost image of the object if its exposure is gated by the bucket detector.[1] Our CCD array imaged the target and reference planes located on a sheet of paper where one half is glossy white and the other half contains the target. The scattered and reflected light from the glossy white half of the paper, which contains the reference spatial information for the ghost image, is then captured by the left-half of the high resolution CCD camera operating in the photon counting regime. The CCD camera is focused onto the ghost image plane and is gated by the bucket detector for the generation of the secondary ghost

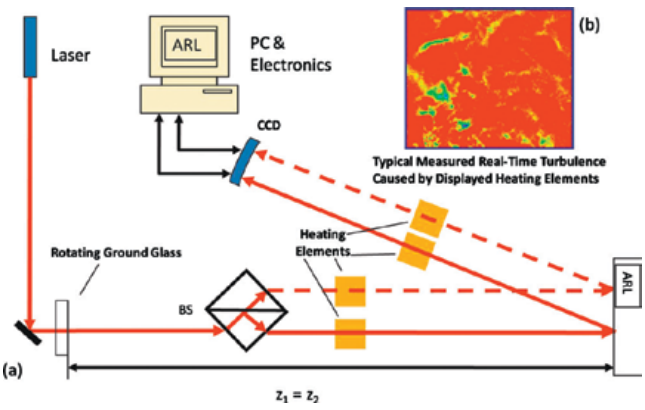


FIG. 1. (Color) (a) Schematic setup of a typical lensless ghost imaging experiment with thermal light in which significant turbulence is introduced in its optical paths. Dashed arrows indicate the optical path to the "bucket" and solid arrows indicate the optical path of the reference image. (b) The inset depicts typical turbulence structures measured during the experiment.

[a)]Electronic mail: ronald.meyers1@us.army.mil.

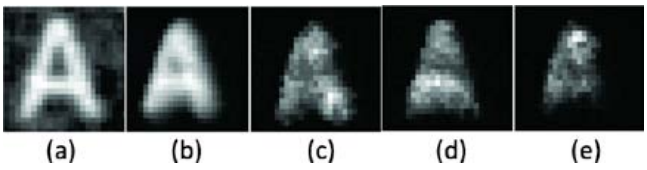


FIG. 2. (Color online) From left to right: (a) a ghost image of the letter "A," (b) an averaged image of the letter "A" followed by three classical image "shots" of letter "A" with turbulence (c)–(e). These images were taken through the bucket detector arm. Notice the turbulence induced jittering and distortion of the images from one "shot" to another "shot." Both the ghost image and the averaged classical image came from the same turbulence. The ghost image is not adversely effected by the turbulence while the average of the classical images shows significant degradation.

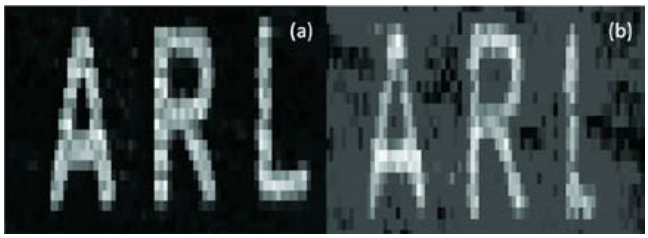


FIG. 3. (Color online) A ghost image of "ARL" without turbulence (a) is compared with a ghost image of "ARL" with significant turbulence (b). In this measurement turbulence was introduced into all optical paths as shown in Fig. 1. There is virtually no difference when turbulence is introduced.

image. Analogously, Valencia *et al.*[4] focused light from the source onto the imaging plane to be measured with a scanning fiber tip. In our special setup each half of the CCD camera can play the role of an independent classical camera in its "normal" ungated operation. The hardware circuit and the software program monitor the outputs of the left-half and the right-half of the CCD individually, as two independent classical cameras, and simultaneously to monitor the gated output of the left-half of the CCD as a ghost camera. The classical image and the secondary ghost image of the object were captured and monitored simultaneously when turbulence was introduced to any or to all of the optical paths.

The effect of turbulence on a classical image is easily seen in Figs. 2(c)–2(e), in which three sequential classical image "shots" of the letter "A" were taken by the right-half of the CCD in its normal operation as a classical camera. Integrating a number (10 000) of sequential images with an exposure time of 1 ms results in a blurred image, Fig. 2(b). Turbulence is a result of strong stochastic space and time variations in the fluid properties such as velocity components $u_i(\mathbf{x},t)$ and index of refraction $\eta(\mathbf{x},t)$. Our experiments used realtime imaging, Fig. 1(b), to extract properties for objective characterization of the anisotropic inhomogeneous turbulence.[8] Velocity probability density functions, velocity correlations, $\langle u_i(\mathbf{x}_1,t_1)u_j(\mathbf{x}_2,t_2)\rangle$, and their time and space scales[8] were computed. The turbulence velocity correlation space scales were found to be 1–2.5 mm and the turbulence velocity correlation time scales were 2.5–5 ms. While each beam distorts and spreads due to turbulence, the point-to-point correlation between the image plane and the object plane is maintained. The small space turbulence scales indicate that the reference and bucket beams experience independent turbulence deviations, and the small time correlation scales indicate that the image frames at different times also experience different turbulence realizations. Optical turbulence is the variability of light propagating through $\eta(\mathbf{x},t)$ fluctuations, and is often characterized by the structure parameter $\langle[\eta(\mathbf{x})-\eta(\mathbf{x}+r)]^2\rangle=C_n^2\mathcal{F}$ where $\mathcal{F}$ is a scaling function[9] that is often set to $r^{2/3}$. C_n^2 is a standard means of characterizing both laboratory and atmospheric optical turbulence and has dimensions of length$^{-2/3}$, rendering the structure function dimensionless. Using standard methods[9] we determined that the images in Fig. 2 were taken in high turbulence with $C_n^2=1.5\times10^{-12}$.

This experiment demonstrates that neither the spatial resolution nor contrast of the ghost image were affected significantly by the turbulence present in the bucket and reference optical paths. Figures 2(a) and 2(b) compare a ghost image (a) with a simple average image (b) under the same turbulence conditions as in Figs. 2(c)–2(e). It is clear that the ghost imaging resolution surpasses the resolution of the "average image" in turbulence. In Fig. 3, a ghost image of "ARL" with the same high turbulence described above is compared with another ghost image of "ARL" without turbulence. It is difficult to see the difference. Visibilities in Figs. 3(a) and 3(b) after normalization were 75% for the nonturbulence case and 33% for the turbulence case. Subtracting out the nearly constant backgrounds in both the turbulence and nonturbulence cases, yielded nearly 100% visibility. The quality of the ghost image is virtually unaffected by turbulence even though the turbulence acts to scatter the energy along the optical paths to the target and detection planes. Further experiments were performed with exposure times as short as 1 μs and yielded results similar to those presented above. Resolution was quantified between Figs. 2(a) and 2(b) by applying a Gaussian point spread function (PSF) to the initial "A" target to approximate Figs. 2(a) and 2(b), respectively. PSF standard deviations in each dimension were 1.6 pixels to match the ghost image of Fig. 2(a) and 3.2 pixels to approximate the classical average of Fig. 2(b), which still had unaccounted for aberrations. In summary, the ghost image appears relatively undistorted by turbulence, i.e., turbulence-free.

The quantum theory of photodetection[10] gives a reasonable interpretation to the turbulence-free ghost imaging of thermal light. In Glauber's theory, a joint detection of two independent point photodetectors measures the probability of observing a joint-detection event of two photons at space-time points $(\mathbf{r}_1,t_1)$ and $(\mathbf{r}_2,t_2)$:

$$G^{(2)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)=\langle\langle E_1^{(-)}E_2^{(-)}E_2^{(+)}E_1^{(+)}\rangle_{QM}\rangle_{Ensemble}, \tag{1}$$

where $E_j^{(\pm)}(\mathbf{r}_j,t_j), j=1,2$ are the negative and positive field operators at $(\mathbf{r}_1,t_1)$ and $(\mathbf{r}_2,t_2)$. We have proven that the quantum expectation is the result of a superposition[1]

$$\langle\langle E_1^{(-)}E_2^{(-)}E_2^{(+)}E_1^{(+)}\rangle_{QM}\rangle_{Ensemble}=|g_2(\vec{\rho}_2,z_2;\vec{\kappa})g_1(\vec{\rho}_1,z_1;\vec{\kappa}') + g_2(\vec{\rho}_2,z_2;\vec{\kappa}')g_1(\vec{\rho}_1,z_1;\vec{\kappa})|^2, \tag{2}$$

where $g_j(\vec{\rho}_j,z_j;\vec{\kappa})$ is the Green's function, which propagates the field from the source to the jth photodetector in the image and object arms.[11,12] Equation (2) indicates an interference between two quantum amplitudes, corresponding to two alternatives, different yet indistinguishable, which lead to a joint photodetection event. This interference involves both arms of the optical setup as well as two distant photodetection events at $(\vec{\rho}_1,z_1)$ and $(\vec{\rho}_2,z_2)$, respectively. Figure 4 schematically illustrates the two alternatives for a pair of

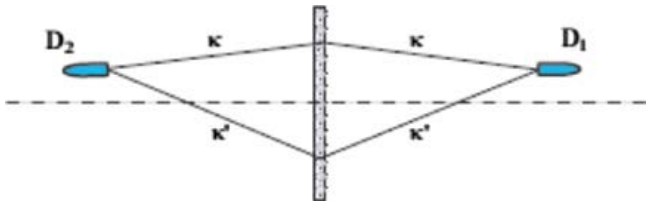


FIG. 4. (Color) Schematic illustration of two-photon interference: a superposition between $g_2(\vec{\rho}_2,z_2;\vec{\kappa})g_1(\vec{\rho}_1,z_1;\vec{\kappa}')$ and $g_2(\vec{\rho}_2,z_2;\vec{\kappa}')g_1(\vec{\rho}_1,z_1;\vec{\kappa})$. The two-photon amplitudes $g_2(\vec{\rho}_2,z_2;\vec{\kappa})g_1(\vec{\rho}_1,z_1;\vec{\kappa}')$ and $g_2(\vec{\rho}_2,z_2;\vec{\kappa}')g_1(\vec{\rho}_1,z_1;\vec{\kappa})$ will superpose constructively at $\vec{\rho}_1 \simeq \vec{\rho}_2$ and $z_1 \simeq z_2$.

modes κ and κ' to produce a joint photodetection event.

Now, we introduce an arbitrary phase disturbance $e^{i\varphi_1(\vec{\rho}_1)}$ into the image arm and another phase disturbance $e^{i\varphi_2(\vec{\rho}_2)}$ into the object arm to simulate the turbulence, where $\Delta\varphi_1(\vec{\rho}_1)$ and $\Delta\varphi_2(\vec{\rho}_2)$ add random phases onto the radiation at transverse coordinates $\vec{\rho}_1$ and $\vec{\rho}_2$, respectively. The quantum expectation is thus

$$\begin{aligned}&\left|g_2(\vec{\rho}_2,z_2;\vec{\kappa})e^{i\Delta\varphi_2(\vec{\rho}_2)}g_1(\vec{\rho}_1,z_1;\vec{\kappa}')e^{i\Delta\varphi_1(\vec{\rho}_1)}\right.\\&\quad\left.+\,g_2(\vec{\rho}_2,z_2;\vec{\kappa}')e^{i\Delta\varphi_2(\vec{\rho}_2)}g_1(\vec{\rho}_1,z_1;\vec{\kappa})e^{i\Delta\varphi_1(\vec{\rho}_1)}\right|^2\\&=\left|g_2(\vec{\rho}_2,z_2;\vec{\kappa})g_1(\vec{\rho}_1,z_1;\vec{\kappa}')\right.\\&\quad\left.+\,g_2(\vec{\rho}_2,z_2;\vec{\kappa}')g_1(\vec{\rho}_1,z_1;\vec{\kappa})\right|^2.\end{aligned}\tag{3}$$

Notice that the phase disturbances introduced by the turbulence have a null effect on the second-order correlation function $G^{(2)}(\vec{\rho}_1,z_1;\vec{\rho}_2,z_2)$ of Eq. (1). The normalized nonfactorizable point-to-point image-forming correlation $g^{(2)}(\vec{\rho}_1;\vec{\rho}_2)$ of thermal light is thus turbulence-free. The two-photon symmetric wave function conditions established in the theory show the aberration cancelling effect of turbulence-free ghost imaging.

As proven earlier[1] and demonstrated here in turbulence, the image forming correlation $g^{(2)}(\vec{\rho}_1;\vec{\rho}_2)$ of thermal light is a nonfactorizable point-to-point intensity-intensity correlation that comes about from quantum superposition of two-photon amplitudes instead of classical correlation of intensity fluctuations. See Agarwal *et al.*[13] on nonclassical interference. Significantly, the nonlocal nonfactorizable property of thermal light that we demonstrated could be useful as a potential resource for a quantum information processing thermal qubit.

By comparison, a classical simulation of ghost imaging was proposed by Gatti *et al.*,[14] in which two classical imaging systems are used to image the speckles of the light source onto the object plane and the image plane, respectively, to form a trivial speckle-to-speckle correlation. Under turbulence, the object and image plane speckles would be blurred in a random manner as would the factorizable speckle-to-speckle correlation.

In conclusion, we have demonstrated the peculiar turbulence-free feature of ghost imaging with thermal light, which can be extremely useful for applications such as distant imaging. The turbulence-free ghost imaging phenomena is the result of nonlocal two-photon interference which cannot be simulated classically by factorizable intensity-intensity correlations.

R. Meyers and K. Deacon thank ARL for support. Y. H. Shih thanks Army Research Office Quantum Imaging MURI, W911NF-05-1-0197 and the Air Force Office of Scientific Research, FA9550-09-1-0489, for support.

[1]R. Meyers, K. Deacon, and Y. Shih, Phys. Rev. A **77**, 041801(R) (2008).
[2]R. Meyers, K. Deacon, and Y. Shih, J. Mod. Opt. **54**, 2381 (2007).
[3]Type-I ghost imaging uses entangled photon pairs as the light source: T. Pittman, Y. Shih, D. Strekalov, and A. Sergienko, Phys. Rev. A **52**, R3429 (1995); D. Strekalov, A. Sergienko, D. Klyshko, and Y. Shih, Phys. Rev. Lett. **74**, 3600 (1995).
[4]Type-II ghost imaging uses thermal radiation as the light source: G. Scarcelli, V. Berardi, and Y. Shih, Phys. Rev. Lett. **96**, 063602 (2006); A. Valencia, G. Scarcelli, M. D'Angelo, and Y. Shih, *ibid.* **94**, 063601 (2005).
[5]F. Wang, Y. Cai, and O. Korotkova, Proc. SPIE **7588**, 75880F (2010); J. Cheng, Opt. Express **17**, 7916 (2009).
[6]R. Meyers, K. Deacon, and Y. Shih, Proc. SPIE **7092**, 70920E (2008).
[7]R. Meyers, K. Deacon, and Y. Shih (unpublished).
[8]R. Meyers and E. O'Brien, Combust. Sci. Technol. **26**, 123 (1981); R. Meyers, W. Ohmstede, and R. Cederwall, *Seventh Symposium on Turbulence and Diffusion* (American Meteorological Society, Boston, 1985), pp. 349–342.
[9]L. Andrews and R. Philips, *Laser Beam Propagation Through Random Media* (SPIE, Bellingham, WA, 2005), Vol. PM152; A. Tunick, Opt. Express **16**, 14645 (2008).
[10]R. Glauber, Phys. Rev. **130**, 2529 (1963); **131**, 2766 (1963).
[11]M. Rubin, Phys. Rev. A **54**, 5349 (1996).
[12]J. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, New York, 1968).
[13]G. Agarwal, J. von Zanthier, C. Skornia, and H. Walther, Phys. Rev. A **65**, 053826 (2002).
[14]A. Gatti, E. Brambilla, M. Bache, and L. Lugiato, Phys. Rev. A **70**, 013802 (2004).

Biographies of ARL Authors

William A. Beck is a research physicist in the EO/IR Materials and Devices branch of the U.S. Army Research Laboratory (ARL). He holds Ph.D. and M.S. degrees from the University of Maryland, College Park, and a B.S. from Rensselaer Polytechnic Institute. He uses finite-difference, time-domain (FDTD) and rigorous coupled wave analysis (RCWA) techniques to design optical coupling struc**tures for solar cells and quantum well infrared photodetectors (QWIPs). He also performs material-level (bandstructure) and device-level modeling and design. Prior to joining ARL in 1999, Dr. Beck worked for 21 years at Martin Marietta Laboratories on a variety of infrared detector materials and devices including LiTa2O3, PbSnTe, HgCdTe, QWIPs, and analyzed FPA performance usi**ng 3-D noise analysis. He also developed the mobility spectrum technique for extracting carrier properties from semiconductor structures with multiple carrier mobilities. He has authored and coauthored more than 50 refereed papers and holds one patent.

Leonid A. Beresnev is a senior research scientist at the U.S. Army Research Laboratory (ARL). He received his M.S. in Radio-Physics and Electronics from the Moscow Institute of Physics and Technology in 1971 and a Ph.D. in Solid State Physics from the Institute of Solid State Physics of Russian Academy of Sciences in 1979. His extensive research experience includes research at the Organic Intermediates and Dyes Institute, Chemical Industry, Moscow; Institute of Crystallography, Russian Academy of Sciences; Darmstadt University of Technology and German Telekom, Darmstadt, Germany; Center of Bio-Molecular Science and Engineering at NRL, Washington DC; and Intelligent Optics Lab at ARL. Dr. Beresnev has 27 U.S. patents and over 150 scientific publications. His current research activities are in development of fast control devices for adaptive optics, imaging, and directed energy applications; development of fast fiber optic positioning devices for optical telecommunications and beam combining in harsh environment; development of arrays of fiber optics collimators for laser beam combining; integration of fiber laser devices and arrays into systems for image processing, free-space optical communication, and directed energy applications.

Matthew J. Berg received his B.S. in Engineering Physics in 2003 from the Colorado School of Mines and his Ph.D. in Physics in 2008 from Kansas State University. His dissertation work focused on theoretical and computational electromagnetic scattering. Dr. Berg then held a National Research Council Chemical and Biological Defense postdoctoral fellowship at the U.S. Army Research Laboratory (ARL) from 2009–2010, where he performed experiments in light scattering from aerosol particles. He is now a tenure-track Assistant Professor in the Department of Physics and Astronomy at Mississippi State University. Dr. Berg has published 21 first-author peer-reviewed journal articles, one book chapter, one ARL technical report, and is coinventor on one patent application. In recognition of his research achievement, in 2011 he was awarded the Young Scientists Award from the Journal of Quantitative Spectroscopy and Radiative Transfer.

Gregory Brill received his Ph.D. in Physics from the University of Illinois at Chicago in 2001 studying surfactant-mediated growth mechanisms of Germanium nucleated on Silicon substrates by molecular beam epitaxy and characterized principally by Raman Spectroscopy. Joining the U.S. Army Research Laboratory (ARL) that same year, his responsibilities included establishing the II-VI MBE growth facilities and initiating the HgCdTe in-house research effort. Currently, he has over 17 years of experience related to thin film material growth and characterization—with the majority of that effort focused on studying and developing Hg-based II-VI compounds for infrared applications. His research has touched on numerous topics including large lattice mismatched epitaxy, surface and interface science, nanotextured studies on stepped surfaces, and dislocation dynamics and annihilation processes through annealing. Dr. Brill has presented his work at numerous national and international conferences and is the lead author or coauthor on over 50 peer-reviewed journal papers and also has authored several ARL technical reports. Additionally, he has one book chapter and one patent related to his area of expertise. Currently, Dr. Brill is the Associate Chief for Science and Technology for the Electro-Optics and Photonics Division within the Sensors and Electron Devices Directorate of ARL.

Gary W. Carhart attended the University of New Mexico and New Mexico State University studying astronomy, physics, computer science, electrical and computer engineering, and optics. He has worked for the Air Force Weapons Laboratory as a Computer Engineer trainee, System Industries Inc. as a Field Service Engineer, and the New Mexico State University as a Research Assistant and Systems Technologist. Currently, he is an Electronics Technician in the Computational and Information Sciences Directorate of the U.S. Army Research Laboratory (ARL). He has published over 75 papers on the subjects of optical correlation and pattern recognition, parallel image processing and correction, adaptive optics and imaging through turbulence, lucky region fusion for image reconstruction, adaptive free-space laser communication, and tiled phased optical fiber arrays for laser communication and laser-based weapons systems. Mr. Carhart has extensive experience in adaptive optics control systems development and integration and is a codeveloper (along with M. Vorontsov, et. al.) of the now popular Stochastic Parallel Gradient Descent (SPGD) control system architecture. Honors include the ARL Achievement Award for Best Publication (2006) and Best Paper Award, 27th Army Science Conference (2010).

Alex Lipchen Chan was born and raised near Kota Kinabalu, the capital of the State of Sabah, Malaysia. He received the B.S. in Electrical Engineering from Chung Yuan Christian University (Chung-Li, Taiwan). He also received the M.S. in Industrial Engineering and the M.S. and Ph.D. in Electrical and Computer Engineering from the State University of New York at Buffalo. He began his association with the U.S. Army Research Laboratory (ARL) at Adelphi, MD as a graduate technical assistant in 1995. From 1997–2003, Dr. Chan continued his research at ARL as a postdoctoral fellow under the American Society for Engineering Education Program. In 2003, he became an Electronics Engineer in the Image Processing Branch of ARL/SEDD. Since 2008, he has been leading the Image Understanding Team in this branch to perform a wide range of research and development using EO/IR and Lidar imageries. His research interests include automatic target recognition, computer vision, and artificial neural networks. He was the Perception Factor Lead during the source selection process of the current Robotics CTA program and Context From Imagery Lead of the I-DECIDE ATO program. He is actively involved in the ARL Data-to-Decisions (D2D) efforts, DARPA Computer Science Study Group (CSSG) program, IARPA Aladdin Video program, and several ARO MURI programs. Dr. Chan is a senior member of the IEEE.

Yuanping Chen received B.S. in Semiconductor Physics from Xiamen University in 1982 and started his career as a research engineer, studying the growth of HgCdTe on CdTe by liquid phase epitaxy. In 1990, he resumed his graduate study and received his Ph.D. in physics from the University of Illinois at Chicago in May 1995, with research focused on the heteroepitaxial growth of II-VI compound semiconductors on Si and GaAs by molecular beam epitaxy (MBE). He continued his research as a postdoctoral research associate at the Microphysics Laboratory of the University of Illinois at Chicago until 1996 and as a senior scientist at EPIR Technologies until April 2002. He joined the U.S. Army Research Laboratory (ARL) as a Physicist and civilian employee in May 2002, with a continuing focus on the research of MBE growth of II-VI and IV-VI materials on alternate substrates—such as Si, GaAs, InSb, and GaSb. Currently his main focus is on the MBE growth of CdSeTe, ZnTe, and ZnSeTe on Si(211) and Si(100), as well as the growth of HgCdSe on ZnTe/Si. He is also actively pursuing the growth of ZnTe nanopillars on Si and various dislocation reduction techniques to be applied to the heteroepitaxial systems of ZnSeTe/Si and CdSeTe/Si. Dr. Chen has over 30 years of research experience on epitaxial growth of II-VI compound semiconductors with over 70 archival articles in refereed publications and one book chapter and numerous presentations in national and international conferences.

Kwong-Kit Choi is a Senior Research Scientist for the Physical Sciences in Sensors and Electron Devices Directorate at the U.S. Army Research Laboratory (ARL). His research focus is devoted to the invention and development of the next generation of advanced sensors and imagers for a full spectrum of Army operations—such as, night vision, missile and unmanned aerial systems defense, and helicopter navigation. He also works with NASA in space applications. Dr. Choi received his B.S. from Hong Kong University in 1979 and Ph.D. from Yale University in 1984. After working at Princeton University and AT&T Bell Laboratories, he joined ARL in 1988. Dr. Choi is a Fellow of ARL, American Physical Society, IEEE, and a member of SPIE. He is an inductee of the NASA Space Technology Hall of Fame and a recipient of a Distinguished Presidential Rank Award.

Keith Deacon is a Research Mathematician in the Special Projects Branch of the U.S. Army Research Laboratory's (ARL) Computational and Informational Sciences Directorate. In November 2009 he received the Department of the Army Research and Development Achievement Award for Technical Excellence. He is an Army veteran who, over the past 25 years, has developed scientific algorithms and computer programs for the government while serving in various research positions. Currently, his sophisticated computer programs control the apparatus used in quantum experiments and provide data-analysis tools for the Quantum Imaging and Information Science and Technology (QUIIST) project at ARL. These programs control the lasers, photon detectors, imagers and modulating equipment, and entangled photon coincidence equipment. Earlier, Mr. Deacon's research included turbulence, nonlinear equations, and transport and diffusion. He developed many algorithms and computer programs to solve turbulence atmospheric flow problems over complex 3-D structures and terrain. Many of the modeling and simulation programs he developed use high-performance computers and parallelization. In 1998, he was recognized along with Mr. Ronald Meyers by the CIA and Director of ARL for modeling a simulated bio-agent attack on Washington, DC. Keith Deacon is coauthor of numerous peer-reviewed publications and is presently a cochair of the SPIE Quantum Communications and Quantum Imaging Conference. He is a coinventor on 17 U.S. patents and is named as coinventor on a number of patent applications. Additionally, he is a regular reviewer for several scientific journals.

Charles Dietlein is an Electronics Engineer in the Sensors and Electron Devices Directorate of the U.S. Army Research Laboratory (ARL), where his research focuses on novel uses of the millimeter-wave and terahertz spectrum (50–1000 GHz) for defense and security applications. He graduated with a B.S. in Electrical Engineering from Seattle Pacific University in 2003, and received the M.S. and Ph.D. degrees, also in Electrical Engineering, from the University of Colorado at Boulder in 2005 and 2008, respectively. Before joining ARL in 2010, he was a member of the Quantum Information and Terahertz Technology Project at the National Institute of Standards and Technology in Boulder, CO, where his research efforts helped produce the first video-rate 100–2000 GHz superconducting camera for standoff personnel imaging applications, along with the first quasi-optical

absolute calibration standard for the millimeter-wave/terahertz spectral regime. Dr. Dietlein is a member of the Government Microcircuit Applications and Critical Technology (GOMACTech) technical program committee, a member of IEEE, a full member of USNC-URSI Commission C, and has over 50 technical publications.

Kevin Doyle received his B.S. degree in Physics from Trinity University (San Antonio, TX) and his M.S. and Ph.D. from West Virginia University in 2008 and 2013, respectively. In 2009 he worked as a Visiting Scholar at Texas State University (San Marcos, TX) and in 2011 he began performing research at the U.S. Army Research Laboratory (ARL) in Adelphi, MD. His graduate research focused on the growth of II-VI compounds via molecular beam epitaxy and the electrical characterization of various II-VI and III-V materials. In 2013 Dr. Doyle became a Postdoctoral Research Associate at the Oak Ridge Associated Universities, under which he continues his research on Hg-based compounds at ARL. He has authored or coauthored 11 papers and presented at several national and international conferences.

Melvin Felton received his B.S. in Mathematics from Morehouse College (Atlanta, GA) in 2000. In 2003, he received his M.S. in Physics with a concentration in atmospheric science from Hampton University (Hampton, VA). For the past 10 years, he has been employed as a physicist in the Battlefield Environment Division of the U.S. Army Research Laboratory (ARL) where his research interests have included boundary layer classification using lidar remote sensing, environmental impacts on thermal polarimetric signatures, and infrared spectroscopy.

Kriston Gurton received his Ph.D. in Physics from New Mexico State University in 1994, where he specialized in Applied Optics/ Electromagnetic theory. Prior to his appointment at the U.S. Army Research Laboratory (ARL), he was an associate fellow with the National Academy of Sciences where he conducted experimental research to determine various optical properties of man-made, and naturally occurring, aerosols. He is currently an Experimental Physicist in the Computational and Informational Sciences Directorate of ARL and is an active member in the electro-optic and aerosol research groups. Dr. Gurton is currently investigating various applications dealing with the phenomena of polarized thermal emission in both the MidIR and LWIR. During his tenure at ARL he has authored in excess of 60 research papers, awarded 11 DoD and academia related research grants, and has seven U.S. patents, and two pending. He is an advisor for the National Research Council (NRC) and sits on the following scientific journal review committees: Applied Optics, Optical Engineering, Int. Journal of Infrared and Millimeter Waves, Journal of Quantitative Spectroscopy and Radiative Transfer, and serves as Advisor.

Abigail S. Hedden received a B.A. degree in Physics from Carleton College (Northfield, MN) in 2001, and a Ph.D. degree in Astronomy from the University of Arizona (Tucson, AZ) in 2007. She was a submillimeter array postdoctoral fellow with the SMA Receiver Lab at the Center for Astrophysics (Cambridge, MA). In 2010, she joined the Millimeter-Wave Sensor Technology team in the Sensors and Electron Devices Directorate at the U.S. Army Research Laboratory (ARL). Her current research interests include developing millimeter-wave/terahertz instrumentation, phenomenology, and experimentation.

Shuowen (Sean) Hu received his B.S. in Electrical and Computer Engineering from Cornell University in 2005, and his Ph.D. in Electrical and Computer Engineering from Purdue University in 2009. Both his undergraduate and graduate focus was on digital signal and image processing. He was awarded the Andrews Fellowship to study at Purdue University, conducting research on biomedical signal processing of functional magnetic resonance imaging (fMRI) data. Following graduation, he joined the U.S. Army Research Laboratory (ARL) as an Electronics Engineer in the Image Processing Branch. His specific areas of interest are in automatic target recognition, cross-modal face recognition, activity recognition, and image enhancement. He received an Army R&D award in 2012, and has over 30 publications in image processing and fMRI analysis.

John W. Little received his B.S. degree in Physics from Tennessee Technological University in 1976 and his Ph.D. degree in Physics from the University of Tennessee, Knoxville in 1982. After a year as a postdoctoral fellow at Oak Ridge National Laboratory in Tennessee, he moved to Maryland and began his career in electro-optics and photonics at the former Harry Diamond Laboratory and Martin Marietta Laboratory. Dr. Little is currently Team Leader of the Infrared and Energy Conversion Devices Team in the Electro-Optics and Photonics Division at the U.S. Army Research Laboratory (ARL), and his research interests are in experimental and theoretical characterization of high-performance infrared detectors, photovoltaic devices, and novel electro-optical systems.

Jony Jiang Liu is the leader of the Intelligent Optics Team in the Atmospheric Sensing Branch, Battlefield Environment Division of the Computational and Informational Sciences Directorate of the U.S. Army Research Laboratory (ARL). He received his Ph.D. in Solid State and Laser Physics in 1991 from Pennsylvania State University. He joined ARL in 1994 as an NRC research associate. Dr. Liu has been leading and working on ARL's internal mission programs and customer projects (DARPA, HEL-JTO). He has worked on developing the ultrahigh bandwidth optical interconnects, semiconductor electro-optical sensors and transceivers,

free-space laser communications, atmospheric imaging and processing, directed energy laser systems, advanced munitions sensing, and countermeasure technologies. Dr. Liu has authored and coauthored over 100 technical papers and book chapters in the U.S. and internationally circulated scientific journals and proceedings. He is a senior member of IEEE, and holds six U.S. patents. Dr. Liu has served as a Technical Assistant to the Director of ARL and also as an S&T LNO to the U.S. Army G-4 Logistics in the Office of Deputy Chief of Staff.

Joseph Mait received his B.S. in Electrical Engineering from the University of Virginia in 1979 and his graduate degrees from the Georgia Institute of Technology; his M.S. in Electrical Engineering in 1980 and Ph.D. in 1985. Since 1988, Dr. Mait has been with the U.S. Army Research Laboratory (ARL), where he is presently a Senior Technical (ST) Researcher. Early in his career, Dr. Mait was an assistant professor of Electrical Engineering at the University of Virginia. He was also an adjunct associate professor at the University of Maryland, College Park, and adjunct professor at Duke University. He has held visiting positions at the Lehrstuhl für Angewandte Optik, Universität Erlangen-Nürnberg, Germany and the Center for Technology and National Security Policy, National Defense University, Washington DC. Dr. Mait's research interests include sensors and the application of optics, photonics, and electro-magnetics to sensing and sensor signal processing. Particular research areas include diffractive optic design and computational imaging. He also had an unexpected sojourn into autonomous systems, where for six years he led ARL's program on micro-autonomous systems and technology. He is currently Editor-in-Chief of Applied Optics. He is a Fellow of SPIE and OSA, and a senior member of IEEE. Additionally, he is also a member of Sigma Xi, Tau Beta Pi, and Eta Kappa Nu, and is a Raven from the University of Virginia.

Ronald Meyers has been a PI Research Physicist at the U.S. Army Research Laboratory (ARL) since 1982. In 2013 he was recognized as the Army's Top Innovator by Heidi Shyu, Assistant Secretary of the Army at the Army's acceptance of the Thompson-Reuters Top 100 Innovators Award. He received the Assistant Secretary of the Army Coin for Excellence. He currently has 18 U.S. patents, and for inventing Quantum Ghost Imaging of remote objects he received the 2009 Department of the Army R&D Achievement Award. He has authored more than 100 scientific papers and coauthored the physics book, From Instability to Intelligence, 1997. He performed research at Army laboratories in the 1970s while on active duty. As an Army research liaison at DOE, he set up research programs at the National Laboratories and worked with the White House on scientific programs focused on drug detection. Later at Brookhaven National Laboratory, he developed the first comprehensive U.S. model of air pollution and acid rain for the White House. He also developed the physics and chemistry method of choice model for computing turbulent combustion in jet engines. Ronald Meyers established the ARL Quantum Imaging and Information Science and Technology (QUIIST) laboratory. In 1996, he was recognized for outstanding support of the Suite of Integrated Infrared Countermeasures and Common Missile Warning System by the PM, Aviation Electronics Combat, and he received a NASA Innovator Award. In 1998, he was recognized by the CIA and Director of ARL for modeling a bio-agent attack on Washington, DC. He also inspected and upgraded crisis centers for Army chem-bio storage facilities for GEN Thompson, CG AMC. He performed a reach-back, real-time, downwind hazard forecast for the U.S. forces attack of Salmon-Pac, the Iraqi bio-center. In 2005 he was recognized by MG Nadeau, CG RDECOM. He is founding and current chair for the SPIE Quantum Communications and Quantum Imaging Conference.

Predrag Milojkovic received his Ph.D. in 2001 from George Mason University for his work on novel free space optical interconnects for multiprocessor systems. Dr. Milojkovic joined the U.S. Army Research Laboratory (ARL) in January 2010 after working at Northrop Grumman on new electro-optical and infrared imaging concepts, systems, and devices. His research at ARL is focused on computational imaging, optical super-resolution with active illumination, gradient index optics, and optical analysis and modeling of high resolution, wide field-of-view imagers. Dr. Milojkovic has published over 30 publications in peer-reviewed journals and conferences, and holds one patent. He currently serves as Program cochair for OSA's Computational Optical Sensing and Imaging conference, is Guest Editor for a Special Issue of Optical Engineering on Gradient Index Optics, and serves on the Program committee for the Military Sensing Symposia Passive Sensors conference.

Nasser M. Nasrabadi received the B.S. in Engineering and Ph.D. degree in Electrical Engineering from Imperial College of Science and Technology (University of London, London, England) in 1980 and 1984, respectively. From October–December 1984, he worked for IBM (UK) as a senior programmer. During 1985 and 1986 he worked with Philips Research Laboratory in New York, NY as a member of the technical staff. From 1986–1991 he was an assistant professor in the Department of Electrical Engineering at Worcester Polytechnic Institute (Worcester, MA). From 1991–1996 he was an associate professor with the Department of Electrical and Computer Engineering at State University of New York at Buffalo. Since September 1996 he has been a Senior Research Scientist (ST) with the U.S. Army Research Laboratory (ARL) working on image processing and automatic target recognition. Dr. Nasrabadi has served as an associate editor for the IEEE Transactions on Image Processing; the IEEE Transactions on Circuits, Systems and Video Technology; and the IEEE Transactions on Neural Networks. He is a Fellow of ARL, SPIE, and IEEE.

Kimberley A. Olver received her B.A. degree in Chemistry/Biochemistry from Goucher College (Towson, MD) in 1981. She currently works as a Physical Scientist at the U.S. Army Research Laboratory (ARL) located in Adelphi, MD. As part of the Sensors and Electron Devices Directorate, her main interests are the processing and characterization of both II-VI and III-V material systems. Prior to joining ARL in 1999, she was employed as a Senior Engineer by the Martin Marietta Laboratories, now Lockheed Martin Corporation, and was part of the Advanced Infrared Technology Group. Her work included IR Focal Plane Array packaging and array characterization. Her previous experience has also included surface and failure analysis research utilizing Auger, XPS, and SIMS.

Guifu (Jason) Sun received an M.S. degree from the University of Kansas in Physics. He is a Physicist at the U.S. Army Research Laboratory (ARL), Adelphi, MD. He has experience in a wide range of optoelectronic and RF device physics and fabrication. He conducted research and development on tunable RF capacitor design and characterization, silicon microbolometer uncooled FPAs and thermal electric cooler design and fabrication. He has many years of experience in research on conventional and high-Tc superconductors and expertise in high-Tc superconductor epitaxial thin film growth and characteristic test and analysis. His current interest is QWIP FPA research and fabrication. He has four patents, has authored and coauthored more than 50 publications, and holds a world record for the highest critical temperature of high-Tc superconductors.

Brenda VanMil received the B.S. degree in Physics from Montana State University (Bozeman, MT) in 1994, and the M.S. and Ph.D. degrees in Physics from West Virginia University in 2002 and 2005, respectively. Her graduate work focused on MBE growth of II-VI and III-V semiconductors. Dr. VanMil was an ASEE Postdoctoral Fellow at the U.S. Naval Research Laboratory. She then worked for Amethyst Research, Inc., Ardmore, OK starting in 2008 on infrared and thermoelectric material growth and characterization. During this time she was also a visiting scientist with the II-VI Materials and Devices Team at the U.S. Army Research Laboratory (ARL). Dr. VanMil joined General Technical Services in 2012, where she is now a contractor to ARL. At ARL, she grows Hg-based semiconductors by molecular beam epitaxy as part of the II-VI Materials and Devices Team and has recently begun growing III-V based devices for infrared and energy applications. Dr. VanMil has published over 50 refereed journal articles and conference proceedings and has three patents pending.

Gorden Videen received his Ph.D. in Optics at the University of Arizona in 1992 studying the light scattering by rough surfaces and contaminants on mirrors. He spent nearly three years researching irregularly shaped atmospheric aerosols in postdoctoral research at Dalhousie University, followed by two more years developing aerosol detection techniques at White Sands Missile Range as an NRC postdoctoral fellow and at New Mexico State University. He has worked as a Physicist at the U.S. Army Research Laboratory (ARL) since 1997, developing unique theories to calculate the light scattering from irregular particle systems. He has published some 200 peer-reviewed articles dealing with experimental, theoretical, and computational aspects of light scattering. He also has received 11 patents, edited 11 books, coauthored 14 book chapters, and has cochaired eight international conferences—including two NATO Advanced Research Workshops and one NATO Advanced Study Institute. During 2005–2007 he undertook studies at the University of Amsterdam Astronomy Department, where he developed and optimized algorithms to calculate light scattering from irregular particles. Since 2011, he also has served as the acting Atmospheric Sciences Program Manager at the Army Research Office. He is an ARL Fellow.

Priyalal S. Wijewarnasuriya received his Ph.D. in Physics from the University of Illinois at Chicago and is at present leading the development of the next generation infrared materials and devices for the U.S. Army Research Laboratory (ARL) at Adelphi, MD. At ARL he is the Team Leader for the II-VI Materials and Devices Team, EO/IR Materials and Devices Branch of SEDD at ARL. Before joining ARL, he was a member of the Technical Staff at the Rockwell Scientific Center, CA. At Rockwell, Dr. Wijewarnasuriya was a member of the team dedicated to the demonstration of novel, large-format infrared focal plane arrays for tactical and strategic military applications, as well as astronomy. Dr. Wijewarnasuriya has authored or coauthored over 73 papers in the open technical literature, two book chapters and has presented his work at numerous national and international conferences.

David A. Wikner is the Millimeter-Wave Sensor Technology Team Leader in the Sensors and Electronic Devices Directorate of the U.S. Army Research Laboratory (ARL). He has worked in the area of mm-W radar and imaging for 25 years. His work has involved mm-W sensor design, the exploration of unique phenomenology at these frequencies, and the exploitation of novel device technology. He has cochaired the Passive Millimeter-Wave Imaging Technology Conference at SPIE's Defense and Security Symposium since 2002. He is a full member of URSI Commission C. He and his coauthors were awarded the 2011 IEEE Antennas and Propagation Society Harold A. Wheeler Applications Prize Paper Award. He received the M.S. degree in Physics from Stanford University in 1988 and the B.A. degree in Physics and Mathematics from Ohio Wesleyan University in 1986.

S. Susan Young received the B.S. degree in 1984 and the M.S. degree in 1987, from South China University of Technology (Guangzhou, China) and the Ph. D. degree in 1995 from State University of New York at Buffalo (Amherst, NY) all in Electrical and Computer Engineering. In 1995, she became a Research Associate at the Department of Radiation Medicine at Roswell Park Cancer Institute, Buffalo, NY, where she worked on medical imaging problems for radiation therapy. From December 1996–June 2002, she was a Senior Research Scientist at Health Imaging Research Laboratory, Eastman Kodak Company, Rochester, NY. She worked on multiresolution object recognition and image analysis on medical images, and medical image compression. Since June 2002, she has been with U.S. Army Research Laboratory (ARL), Adelphi, MD working in the areas of automatic target recognition (ATR), image enhancement technology for various Army weapon applications, thermal modeling and perception, and biometric recognition. Dr. Young serves as an associate editor of Optical Engineering. Dr. Young has published over 70 papers in reputable journals and conferences. She holds eight patents for inventions that are related to medical diagnostic imaging, image compression, image enhancement, pattern analysis, and image super-resolution. She coauthored a text book on Signal Processing and Performance Analysis for Imaging Systems that was published by Artech House, 2008. She is a recipient of 2007 and 2012 U.S. Army Research and Development Achievement Awards.


U.S. ARMY
RDECOM®