OFFICE OF THE DIRECTOR OF NATIONAL INTELLIGENCE

Finder
Incisive Analysis Office

LEADING INTELLIGENCE INTEGRATION

Jill D. Crisman, Ph.D.
Finder Proposers' Day, 11 May 2011

Agenda

9:00am – 9:15am	*Welcome and Security Brief*	Mr. John Garlock IARPA Security Chief
9:15am – 9:30am	*IARPA Overview and Remarks*	Dr. Peter Highnam IA Office Director
9:30 am – 10:15am	*Finder Overview*	Dr. Jill Crisman Program Manager
10:15am – 10:30am	*Break*	
10:30am – 11:00am	*Finder Q&A*	Dr. Jill Crisman Program Manager
11:00am – 11:30am	*Contracting*	Ms. Barbara Frantom AFRL/RYWB
11:30am – 1:00pm	Proposers' 5-minute briefings	**Attendees**
1:00pm – 2:00pm	*Lunch*	
2:00pm – 3:30pm	Proposers' Networking and Teaming Discussions	**Attendees**

Disclaimer

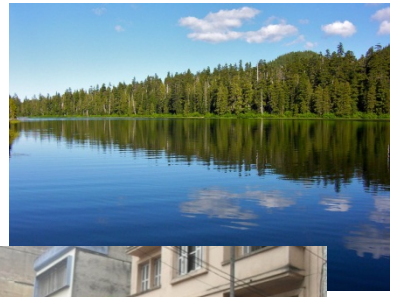
- This presentation is provided solely for information and planning purposes
- The Proposers' Day Conference does not constitute a formal solicitation for proposals or proposal abstracts
- Nothing said at Proposers' Day changes the requirements set forth in a BAA
- BAA supersedes anything presented or said at the Proposers' Day

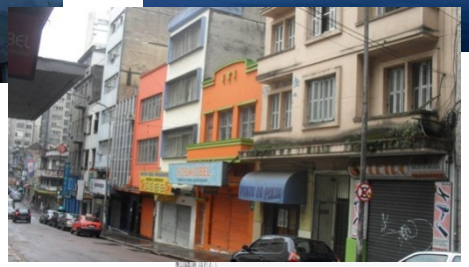
Schedule

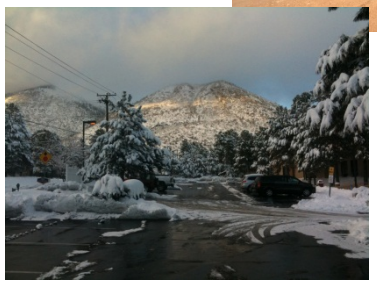
- Once the BAA is released, questions can only be answered in writing on the program website
- Full Proposals are due ~45 days after BAA is published

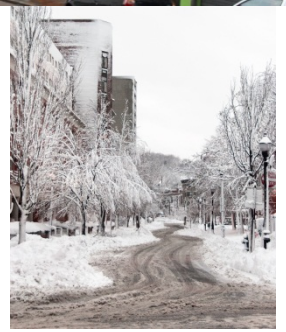
Finder Goals









Where in the world were these photos/videos taken?

- Geolocate ground level outdoor scenes in any region of the world
- Any type of camera
- Any season, time of day, or weather conditions
- Accuracy and speed are important

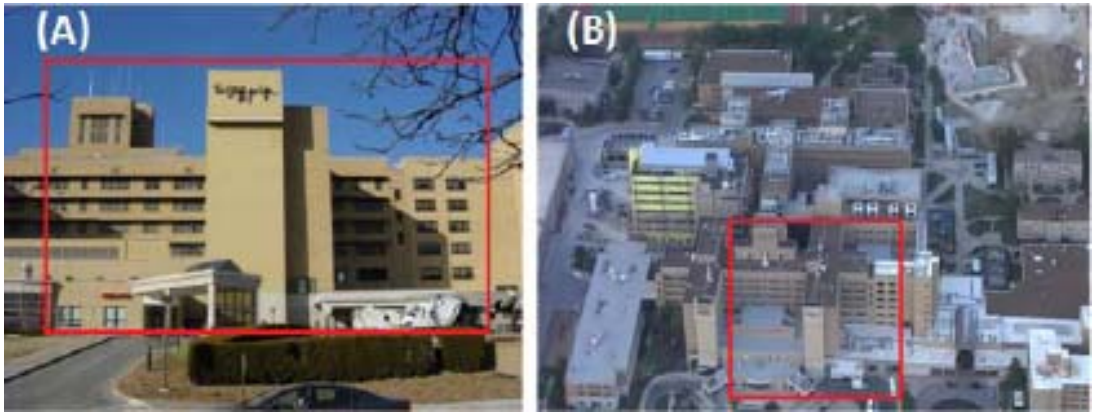
Motivation

- To geolocate a ground-level image, for example, an analyst might:
  - Use provenance of the image to give initial regions of interest (ROIs)
  - Use her understanding of the image content, in conjunction with help from public data sources, to reduce the number and size of the ROIs
  - Use image search and visualization tools to search the ROIs for matching scene(s)

Although this example describes locating an image, many of the same ideas apply to the geolocation of video

Challenges

- This is a "needle in a haystack" problem; time intensive
- What information in the image/video will best reduce the search?
- In much of the world there are few ground-level images, and matching with overhead imagery is problematic



The matching region of the ground level query image (A) to the aerial image (B) is shown in the red boxes

Y.C. Chung, T.X. Han, Z. He, ***Building recognition using sketch-based representations and spectral graph matching***, ICCV 2009

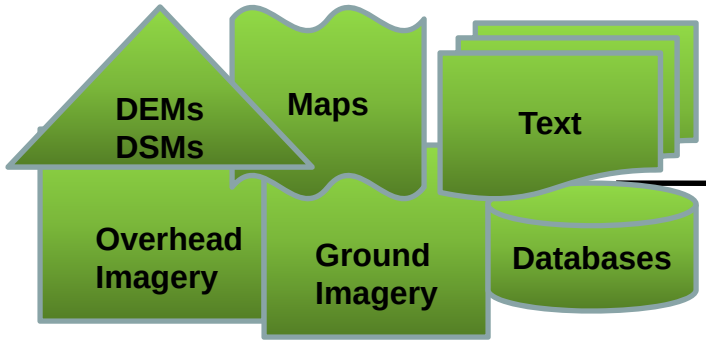
Current Research Limitations

- There are several research efforts attempting the image/video geolocation task
- Each of these efforts has serious limitations, such as:
  - Matches only ground-level imagery to ground-level imagery
  - Only exploits large fixed features like buildings
  - Does not take advantage of vegetation or soil types, cultural cues, etc.
  - Does not fully leverage the analyst's knowledge

Finder

- The Finder program seeks to develop technologies to aid the analyst in geolocating images and videos, in any outdoor terrestrial location in the world, by:
  - Using publicly available reference data and imagery
  - Augmenting analyst knowledge and skill with careful automation

Notional Finder System



Publicly Available Reference Data over the "world"

- Query Photo/Video
- Object Point
- ROIs

Fuse

Knowledge Base

Query

Analyst

Uncertainty

When clicked, each point shows how its probability value was calculated

Camera/Object Probability Maps over ROIs

Fuse (1)

- Goal
  - Discover and integrate multiple sources of reference data to create or update a knowledge base that will be used to support analysts' queries
- Input
  - Conventional data: imagery, digital elevation models, …
  - Diverse data sources such as geology, geography, botany, and anthropology
  - Current knowledge base(s), if available
- Output
  - New or updated knowledge base

Fuse (2)

Challenges include:

- Determine which information sources have value for the geolocation query task
- Organize the information to enable the system to help the analyst reach the best geolocation result as quickly as possible
- Accommodate and maintain uncertainty and data provenance
- Process efficiently

Anticipated Reference Data

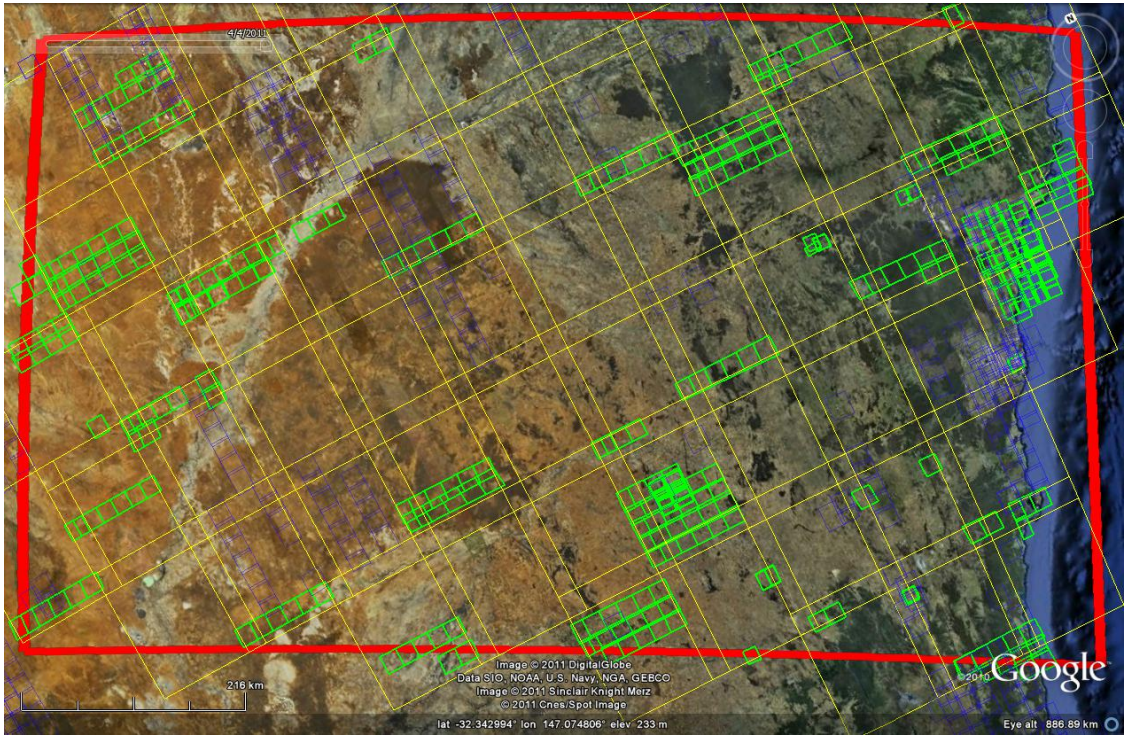
- Performers are free to use any data sets with verification that the data are publicly available and lawfully obtained
- Government will supply certain commonly requested data such as overhead imagery to amortize costs (GFI)
- Outside scope of program:
  - Collecting additional data in the field

Potential GFI Reference Data (1)

- Baseline:
  - 30m Digital Elevation Model (NASA/ASTER)
  - 30m Multispectral Satellite Imagery (NASA Global Land Survey)
  - 125m GIS Map Data (e.g. roads, utilities, coastlines) (1:250,000, VMAP1)
- As available:
  - Higher resolution DEM (e.g. 3m)
  - Panchromatic/MSI/HSI Satellite Imagery (varying resolutions)
  - 0.5m Stereo Panchromatic Satellite Imagery
  - Aerial imagery (e.g. 1m, <0.5m)
  - <1m GIS Map Data including urban roads (1:2000)

Now is the time to propose and discuss data sets!

Potential GFI Reference Data (2)



- GFI in a region of interest (shown in red) will contain:
  - Baseline of data everywhere (shown with yellow boxes)
  - Patchwork of various types and resolutions of imagery (notionally shown with green boxes)

Query (1)

- Goal:
  - Geolocate the image/video using a knowledge base
- Input:
  - Query: image/video (and a frame number if a video)
  - Object point: pixel location on an object in the scene
  - Input ROI: geographic boundaries of the region-of-interest
- Output:
  - Camera and/or Object Point Geolocation Probability Map over ROI; each cell in the map has a value
  - Uncertainty value(s) for the probability map
  - Rationale for the probability value in each cell of the map, in terms of contributing data

Example Query Image and Probability Map

The probability map is a grid placed over the ROI, with each grid cell containing a real number between 0.0 and 1.0 representing the probability of the specific location being the query location (either the camera geolocation or the object point's geolocation). The sum of the probability map cells is 1.0. The resolution of the grid will be provided, for example in the open scene above, 100m x 100m may be appropriate.

Query (2)

Challenges include:

- Limited "wall clock time" for the entire query
- Limited human interaction time, which could be with a technician for basic tasks at the start, followed by time with an analyst who may have additional data
- Must use decision-theory and human factors techniques to produce high-value interactions with the analyst
- Imperfect information from the analyst
- Description/depiction of the full rationale for each location's probability score, for an analyst to review

Query Metrics

- Measure number of detections and false alarms per square kilometer
  - System must automatically compute the probability threshold for each query
  - If the true-geolocation has a probability ≥ threshold, this is a detection; if the true-geolocation has a probability < threshold, this is a missed detection. The detection percent is the number of detections divided by number of queries
  - Number of locations with a probability above the threshold that is not the true-geolocation is a false positive. This can be computed for each query and averaged over all queries
- Measure the "list position" of the true-geolocation
  - Number of incorrect locations with a probability ≥ the true-geolocation's probability
  - This represents the amount of work that an analyst would have to do to examine all of the higher scoring incorrect locations
  - This measure is a work in progress
- Quality of the rationale and its depiction as determined by analysts in government test

Out of Scope

- Any new data collecting in the field
- Face recognition or person tracking
- New language processing or identification algorithms in audio tracks
  - Can use existing algorithms, but this program is not developing new algorithms in this area
- Crowd sourcing of query or fusion
  - Use of publicly available crowd sourced information is okay

Phases

- Ongoing research will be incorporated into the development and evaluation of prototypes
- Phases become more difficult as terrain types are added and the area of the input ROI increases

	Length	Terrain	Terrain types in ROI	Area of the Largest ROI
Phase I Base Period	18 months	Coasts Desert	1	10,000 sq km
Phase I Option Period 1	12 months	Any	1	10,000 sq km
Phase II Option Period 2	12 months	Any	2-3	50,000 sq km
Phase II Option Period 3	12 months	Any	All	500,000 sq km

Performance Goals

Metric	Phase I BP	Phase I OP1	Phase II OP2	Phase II OP3
CONSTRAINTS				
Largest ROI area	10,000 sq km	10,000 sq km	50,000 sq km	500,000 sq km
Fusion Time	< 5 days	< 5 days	< 3 days	< 3 days
Analyst Interaction Time	< 20 min	< 20 min	< 15 min	< 10 min
Query Time	< 4 hours	< 4 hours	< 2 hours	< 1 hour
MEASURE				
Percentage of detections	> 60%	> 70%	> 80%	> 85%
Number of False Alarms	< 0.01 / sq km	< 0.002 / sq km	< 0.0004 / sq km	< 0.00005 / sq km
List position of ground-truth geolocation	TBD	TBD	TBD	TBD

Test and Evaluation

- At the beginning of each period, IARPA will provide information that represents the challenge for that period, including:
  - Test region and corresponding GFI data
  - Set of training query images/videos
- Throughout each period, IARPA will review research and progress at regular intervals
- Near the end of each period IARPA will provide the challenge problem (new GFI data over a new test region)
- Evaluation will include knowledge base construction time and performance on query images/videos run by government team surrogate analysts

Eligibility Information

- Other Government Agencies, Federally Funded Research and Development Centers (FFRDCs), University Affiliated Research Centers (UARCs), and any other similar type of organization that has a special relationship with the Government, that gives them access to privileged and/or proprietary information or access to Government equipment or real property, are not eligible to submit proposals under this BAA or participate as team members under proposals submitted by eligible entities
- Non-US organizations and individuals may be able to participate.
  - Must comply with Non-Disclosure Agreements, Security Regulations, Export Control Laws, etc. as appropriate
  - Specific guidance for non-US participation will be provided in the BAA

Proposal Guidance

- Your proposal should include a full discussion of the technical approach that will be used to meet the program goals.
- Programmatic issues to be addressed in the proposal:
  - Your team's current technical capabilities
  - Key resources needed (not currently available to your team), to include capital equipment and special expertise (teaming will likely play an essential role in providing special expertise). The risk in acquiring these key resources, and mitigation strategies, should be indicated as well
  - A teaming plan along with the roles and responsibilities of each member of the research team
  - End of phase and some intermediate milestones are set, but it is expected that other intermediate milestones that are on the critical path of the proposed approach will be offered
  - A schedule of all milestones including a clearly charted description of the various risk mitigation strategies that will be undertaken to achieve program goals

Proposal Evaluation Criteria

- Overall Scientific and Technical Merit
- Effectiveness of Proposed Work Plan
- Relevance to IARPA Mission and Finder Program Goals
- Relevant Experience and Expertise
- Cost Realism

Evaluation criteria will appear in the BAA.

Teaming

- Because of the many challenges presented by this program, both depth and diversity will be beneficial for overcoming these challenges
  - Throughput: Consider all that you will need to do, all the ideas you will need to test. Make sure you have:
    - Enough people and expertise to do the job
    - Sufficient resources to follow critical path while still exploring alternatives – risk mitigation
  - Completeness: teams should not lack any capability necessary for success, e.g., should not rely on enabling technology to be developed elsewhere.
  - Tightly knit teams
    - Clear, strong, management; single point of contact
    - No loose confederations
    - Each team member should be contributing significantly to the program goals. Explain why each member is important, i.e., if you didn't have them, what wouldn't get done?
    - No teaming for teaming sake
- Remember, you may be very accomplished, but can you do it all?

Additional Information

- Email dni-iarpa-baa-11-05@ugov.gov with additional questions
- Finder BAA will be posted on FedBizOpps website (www.fedbizopps.gov)
- Q&As will appear after the BAA. See http://www.iarpa.gov/solicitations_finder.html

Questions?

Image Citations

Photo Name	Photographer Name	Flickr Username	Flickr User ID	Flickr Photo ID	URL to Photograph	License	License info URL
Thingvellir	Stig Nygaard	Stig Nygaard	10259776@N00	4125634966	http://www.flickr.com/photos/10259776@N00/4125634966/	Creative Commons - Attribution	http://creativecommons.org/licenses/by/2.0/
		WalkingGeek	14481478@N00	164546766	http://www.flickr.com/photos/14481478@N00/164546766/	Creative Commons - Attribution	http://creativecommons.org/licenses/by/2.0/
Farm in Benton Center, near Penn Yann, NY	Jamie Lantzy	jlantzy	18762235@N00	699855371	http://www.flickr.com/photos/18762235@N00/699855371/	Creative Commons - Attribution	http://creativecommons.org/licenses/by/2.0/
Silverton NSW	Ian Sutton	ibsut	22616984@N07	5103806365	http://www.flickr.com/photos/22616984@N07/5103806365/	Creative Commons - Attribution	http://creativecommons.org/licenses/by/2.0/
20090809_0925-DSC_0791	Germany	Tscherno	34073445@N00	5584265233	http://www.flickr.com/photos/34073445@N00/5584265233/	Creative Commons - Attribution	http://creativecommons.org/licenses/by/2.0/
20090813_1005-DSC_0134	Germany	Tscherno	34073445@N00	5586487314	http://www.flickr.com/photos/34073445@N00/5586487314/	Creative Commons - Attribution	http://creativecommons.org/licenses/by/2.0/
20100813_1325-DS3_2852	Germany	Tscherno	34073445@N00	5586627510	http://www.flickr.com/photos/34073445@N00/5586627510/	Creative Commons - Attribution	http://creativecommons.org/licenses/by/2.0/
elk lake_087		eddiecoyote	42280592@N02	5314958717	http://www.flickr.com/photos/42280592@N02/5314958717/	Creative Commons - Attribution	http://creativecommons.org/licenses/by/2.0/
Day one 2010	Bianca	Bianca's Lens	46081989@N06	4245025438	http://www.flickr.com/photos/46081989@N06/4245025438/	Creative Commons - Attribution	http://creativecommons.org/licenses/by/2.0/
Flagstaff	Nathan Borror	nathanborror	72897141@N00	5516562129	http://www.flickr.com/photos/72897141@N00/5516562129/	Creative Commons - Attribution	http://creativecommons.org/licenses/by/2.0/
SAM_3045	Luiz Fernando Reis	Luiz Fernando / Sonia Maria	7477245@N05	5584540876	http://www.flickr.com/photos/7477245@N05/5584540876/	Creative Commons - Attribution	http://creativecommons.org/licenses/by/2.0/
Topas Ecolodge	David McKelvey	David McKelvey	94735786@N00	5057647092	http://www.flickr.com/photos/94735786@N00/5057647092/	Creative Commons - Attribution	http://creativecommons.org/licenses/by/2.0/
Santiago despues de la tormenta	Nicolas Schubert	nischg	95521687@N00	552907440	http://www.flickr.com/photos/95521687@N00/552907440/	Creative Commons - Attribution	http://creativecommons.org/licenses/by/2.0/
slushy s 2	Mark Zastrow	Mark Z.	96146111@N00	5353596101	http://www.flickr.com/photos/96146111@N00/5353596101/	Creative Commons - Attribution	http://creativecommons.org/licenses/by/2.0/