Tactical and Surveillance Video Quality Experiments

Homeland Security

DHS-TR-PSC-07-03

Department of Homeland Security

Public Safety Communications

Technical Report

Command, Control and Interoperability: Communication, Interoperability and Compatibility

Homeland Security
Science and Technology

Defining the Problem

Emergency responders—police officers, fire personnel, emergency medical services-need to share vital voice and data information across disciplines and jurisdictions to successfully respond to day-to-day incidents and large-scale emergencies. Unfortunately, for decades, inadequate and unreliable communications have compromised their ability to perform mission-critical duties. Responders often have difficulty communicating when adjacent agencies are assigned to different radio bands, use incompatible proprietary systems and infrastructure, and lack adequate standard operating procedures and effective multi-jurisdictional, multi-disciplinary governance structures.

OIC Background

The Department of Homeland Security (DHS) established the Office for Interoperability and Compatibility (OIC) in 2004 to strengthen and integrate interoperability and compatibility efforts to improve local, tribal, state, and Federal emergency response and preparedness. Managed by the Science and Technology Directorate, and housed within the Communication, Interoperability and Compatibility thrust area, OIC helps coordinate interoperability efforts across DHS. OIC programs and initiatives address critical interoperability and compatibility issues. Priority areas include communications, equipment, and training.

OIC Programs

OIC programs, which are the majority of Communication, Interoperability and Compatibility programs, address both voice and data interoperability. OIC is creating the capacity for increased levels of interoperability by developing tools, best practices, technologies, and methodologies that emergency response agencies can immediately put into effect. OIC is also improving incident response and recovery by developing tools, technologies, and messaging standards that help emergency responders manage incidents and exchange information in real time.

Practitioner-Driven Approach

OIC is committed to working in partnership with local, tribal, state, and Federal officials to serve critical emergency response needs. OIC's programs are unique in that they advocate a "bottom-up" approach. OIC's practitioner-driven governance structure gains from the valuable input of the emergency response community and from local, tribal, state, and Federal policy makers and leaders.

Long-Term Goals

- Strengthen and integrate homeland security activities related to research and development, testing and evaluation, standards, technical assistance, training, and grant funding.
- Provide a single resource for information about and assistance with voice and data interoperability and compatibility issues.
- Reduce unnecessary duplication in emergency response programs and unneeded spending on interoperability issues.
- Identify and promote interoperability and compatibility best practices in the emergency response arena.

This page intentionally left blank.

Public Safety Communications Technical Report

Tactical and Surveillance Video Quality Experiments

DHS-TR-PSC-07-03
November 2007

Reported for: **The Office for Interoperability and Compatibility by NIST/OLES**

Homeland Security

Publication Notice

Disclaimer

The U.S. Department of Homeland Security's Science and Technology Directorate serves as the primary research and development arm of the Department, using our Nation's scientific and technological resources to provide local, state, and Federal officials with the technology and capabilities to protect the homeland. Managed by the Science and Technology Directorate, the Office for Interoperability and Compatibility (OIC) is assisting in the coordination of interoperability efforts across the Nation.

Certain commercial equipment, materials, and software are sometimes identified to specify technical aspects of the reported procedures and results. In no case does such identification imply recommendations or endorsement by the U.S. Government, its departments, or its agencies; nor does it imply that the equipment, materials, and software identified are the best available for this purpose.

Contact Information

Please send comments or questions to: S&T-C2I@dhs.gov

This page intentionally left blank.

Contents

This page intentionally left blank.

Abstract

This report describes laboratory studies to investigate the level of quality required for the following public safety video applications:

- Narrow field of view, tactical
- Wide field of view, tactical
- Narrow field of view, live surveillance
- Wide field of view, live surveillance
- Narrow field of view, recorded surveillance
- Wide field of view, recorded surveillance
- Tactical and live surveillance

Requirements for these applications are based on the studies described here, and are given in Section 4 of the Public Safety Statement of Requirements (PS SoR) Volume II [1].

Before conducting the studies described here, a video quality questionnaire was used to obtain an initial estimate for the minimum quality levels that public safety video applications require. After the questionnaire results were examined, the first public safety subjective video quality experiment, PS1, was conducted. Using the analysis results of PS1, two more subjective tests, PS2 and PS3, were conducted.

This report describes the questionnaire and subjective experiments and presents the results from the data analysis.

Key words: measure video quality, narrow field of view, reference model for video performance, subjective testing, surveillance video, tactical video, wide field of view

1 Introduction

Since the use of video in public safety applications is relatively new, a questionnaire was used to gather fundamental video quality information (e.g., frame rates, and luma image size) that could be used to design the first subjective video quality experiments for public safety applications. This questionnaire, conducted from July to September, 2005, was given to 18 public safety practitioners from around the United States.

Practitioners were provided with a list of public safety video applications, and asked to rank their importance. This ranking was used to decide the highest-priority public safety video application that the PS1, 2, and 3 experiments should address. For each application where the practitioner had relevant experience, application-specific questions were asked.

2 Video Quality Questionnaire

The definitions used in this questionnaire to differentiate between the various public safety video applications confused some practitioners. As a result, the video categories have since been refined and the definitions improved. To avoid potential confusion going forward, the original definitions and categories in the questionnaire will not be presented here.

For each specific application (e.g., tactical video, surveillance video), practitioners were asked questions concerning acceptable values for the following video characteristics:

- Video Delay
- Control Lag
- Image Size
- Frame Rate
- Lossless Video Quality Requirement
- Coding Impairment
- Response to Network Impairments
- Color Fidelity
- Lighting Requirements
- Focus Distance
- Transmission Distance (between source and destination ends)
- Camera and Monitor Mobility

For each characteristic, an example level of service was presented using a combination of text, still imagery, and video sequences. High-speed computer hardware was required to correctly play back the high-resolution video sequences, which unfortunately prevented the questionnaire from being more widely distributed. Practitioners were asked to mark each level of service as either "desirable," "acceptable," or "unacceptable." This scale was understood and was well-received.

The following preliminary quality levels were indicated for **tactical video**, which was the application considered as the very highest priority by questionnaire participants. (In the next bullet list, the quality levels listed as "borderline" were marked as "acceptable" or "desirable" for tactical video by 60 percent to 80 percent of responders, and "unacceptable" by 20 percent to 40 percent of responders.) All results obtained from the questionnaire were treated as general indications only, due to confusion regarding the definitions mentioned above and the small number of practitioners that were questioned:

- Image Size: HDTV (high-definition television) is desired, NTSC/525-line (National Television Systems Committee 525 scan-line standard) is acceptable, and CIF (Common Intermediate Format) is unacceptable. (QCIF, or Quarter CIF, which was included as the other decreasing image size, is unacceptable.)
- Video Delay: Real-time is desired, near real-time is acceptable, and from one to several seconds delay is borderline.
- Control Lag: Real-time is desired, near real-time is acceptable, and from one to several seconds delay is borderline.
- Frame Rate: A rate of 30 frames per second (fps) is desired, 10 fps is acceptable, and 5 fps is borderline.
- Lossless Video: Participants have been told they need lossless compression, but probably do not know what this means or why they need it.

- Video Quality: "Imperceptible" decrease in quality levels is desired, "Perceptible but Not Annoying" is acceptable, and "Slightly Annoying" is borderline. (The other two decreasing quality levels that were included, "Annoying" and "Very Annoying", are unacceptable.)
- Response to Errors: Transient drops in quality are undesirable. Consistency was preferred over quality. Video Quality is more important than Video Delay.
- Color: Color is very important, both in terms of accurate color rendition and color display being preferred over black and white.
- Lighting and Distance: All lighting levels are important—indoors, daylight, night with dim light, and no noticeable light. Distances less than 200 yards are important; distances farther away are less important.

As a result of the application rankings, the first application considered was narrow field of view tactical video. The PS1 experiment design focused on 525-line/NTSC video sampled according to ITU-R Recommendation BT.601 (Rec. 601) [5], with a limited number of systems being included that operated at CIF and QCIF image sizes. This PS1 experiment was also limited to a minimum frame rate of 5 fps and to color video (i.e., no black and white). The experiment was designed to have an average quality level that was centered on the quality level indicated by the questionnaire responses.

3 Experiment Design

This report describes three subjective video quality experiments for public safety (PS) communications:

- PS1—Narrow field of view, tactical (Section 4)
- PS2—Wide field of view, tactical; narrow and wide field of view for both live and recorded surveillance (Section 5)
- PS3—Tactical and live surveillance (Section 6)

This section describes design attributes that are common across the three experiments.

3.1 Reference Model for Video Performance Measurements

Figure 1 provides a reference model for specifying video performance measurements. To fully quantify the user-perceived quality of service, the performance of three primary video subsystems must be specified.

1. The Video Acquisition Subsystem—That normally consists of a camera system, and may also include a built-in video coder.
2. The Video Transmission Subsystem—May include a network and its associated interfaces, encryption, etc., and may also include the video coder and decoder and a video storage medium.
3. The Video Display Subsystem—Includes a monitor, playback computer, etc. May also include a built-in video decoder.

The exact demarcation for each of the three subsystems can vary from application to application due to integration of various functions within the end user's equipment. The approach adopted here is to specify performance parameters for the application as a whole (i.e., system parameters), as well as to specify performance parameters that are unique to each video subsystem (i.e., acquisition parameters, transmission parameters, and display parameters). Section 3 of the Public Safety Statement of Requirements (PS SoR) Volume II [1] details system, acquisition, transmission, and display video parameters. Figure 1 depicts a reference diagram for the performance measurements. The letters in the figure denote measurement access points that may or may not be available on all video systems.

Figure 1: Video Performance Measurements Reference Diagram

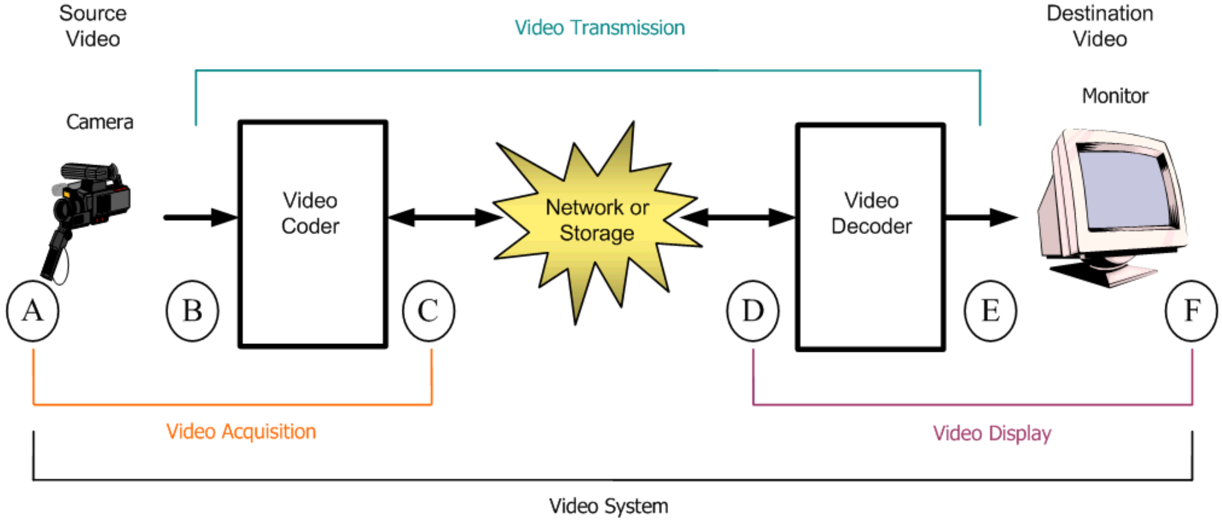
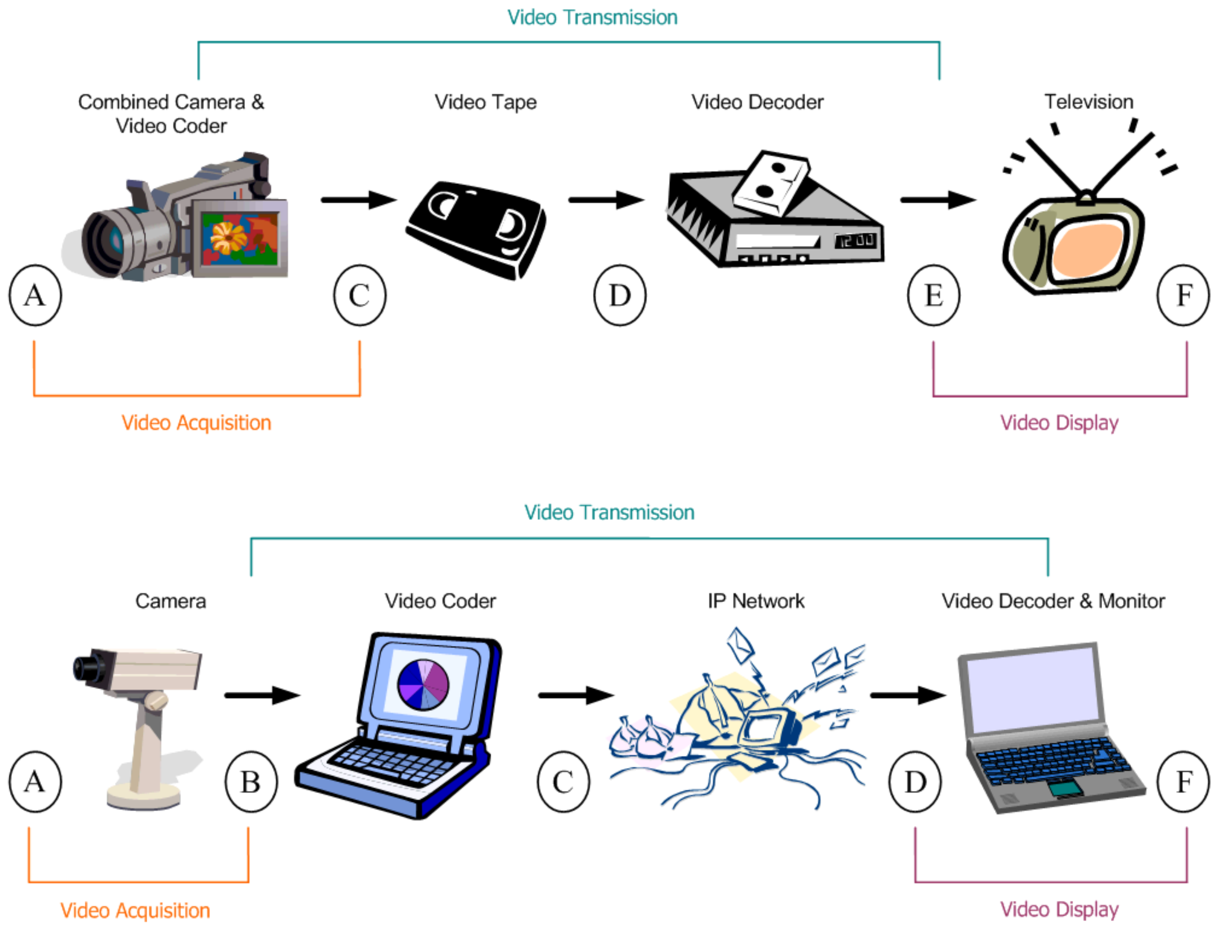


Figure 2 shows two example video systems with reference points identified. In the acquisition system [6], access point B is inside the camcorder and probably not available. In the display system, access point E is inside the computer and may or may not be available. Some video may travel over multiple networks or storage media.

Figure 2: Two Example Video Transmission Systems, with Reference Points Identified



3.2 Video Attributes to Investigate

Hypothetical Reference Circuits (HRCs)[1] were selected for each experiment to investigate the minimal requirements for the video attributes that Table 1 lists.

Table 1: Video Attributes to Investigate

Video Attribute	Applicable Experiment
Coder Bit Rate for H.264 [3] without network impairments	PS1, PS2, PS3
Coder Bit Rate for MPEG-2 [2] without network impairments	PS1
Image Size	PS1, PS2
Frame Rate	PS1, PS2
Packet Loss Ratio	PS1, PS2, PS3
Error Concealment Strategies (including no error concealment)	PS1, PS3, PS3

1. A Hypothetical Reference Circuit (HRC) is an industry-accepted term for a specific configuration of a video transmission subsystem (i.e., fixed configuration settings for the behavior of the video coder, the network, and the video decoder).

Of these video attributes, packet loss ratio and error concealment are the most difficult to characterize using short video sequences. This is because the location of the lost packet within the video transmission stream can significantly affect the reduction in perceived video quality, and the particular error concealment scheme significantly affects the perceptual impact of packet loss. The large number of combinations for the above video quality variables limited what could be examined for any one particular variable.

Investigating the video attributes Table 1 lists, required determining the fundamental characteristics of the video transmission subsystem (see Figure 1) such as image size, frame rate, coder type, coder bit rate, and packet loss ratio. High-quality source video sequences (point B in Figure 2) were used wherever possible so as to not influence the outcome of the experiment. Likewise, a studio-grade television monitor was used to display the resulting videos during the subjective test (points E to F in Figure 2), so that the display would not add any impairments.

Video was always displayed to subjects at Rec. 601 [5] image size. Original video sequences produced with higher resolutions (e.g., high-definition) were down-sampled to Rec. 601 using professional-grade hardware. The video produced by HRCs that used a lower resolution (e.g., CIF) were up-sampled to Rec. 601 prior to use in the subjective test.

Viewing sessions were held in a controlled viewing environment (see ITU-T Recommendation BT.500 [4]). Before participating in the two viewing sessions, subjects were screened for color perception and visual acuity at a distance of 10 feet, and underwent a training session where they were reminded that public safety personnel use tactical video in real time during an incident to make decisions on how to respond to that incident.

3.3 HRC Video Transmission Systems

Each video transmission system sample, referred to as an HRC, or hypothetical reference circuit, by the video quality measurement community, involves taking the original video sequences and changing or distorting them in some way. The HRCs chosen for PS1 can be divided into four categories, three of which also apply to PS2 and PS3:

- Resolution and frame rate modifications

  The original video sequences were modified to reflect different image resolutions and frame rates. (These will be referred to as synthetic HRCs.)

- MPEG-2 IP transmission (PS1 only)

  The original sequences were MPEG-2 [2] encoded, transmitted over an IP (Internet Protocol) network, and then decoded. MPEG-2 was chosen due to the wide prevalence of this established coding scheme.

- H.264 transmission

  The original sequences were H.264 [3] encoded, transmitted over an IP network, and then decoded. H.264 was chosen as a state-of-the-art coding method that is widely regarded as requiring only one-half to one third the bit rate of MPEG-2 for the equivalent quality. Thus, within the next several years, H.264 can reasonably be expected to become the codec of choice for many new services being deployed.

- H.264 transmission with packet loss

  The original scenes were H.264 encoded, packet loss was inserted, scenes with packet loss were transmitted over an IP network, and then decoded.

Given the time constraints for completion of the PS1 experiment, locating MPEG-2 [2] and H.264 [3] video codecs with error concealment proved to be a challenge. As an unfortunate result, most of the video transmission systems included in the PS1 experiment do not implement error concealment. One hardware-based H.264 system with error concealment was included in the experiment. However, this system was limited to 384 kbps and below, and was limited to CIF image size. According to the questionnaire results, the video quality of H.264 at that bit rate and image size was expected to be unacceptable. (However, this did not prove to be the case—see Section 4.5.) A high-priority item for study in future tests will be to include more H.264 codecs with advanced error concealment schemes.

3.3.1 Hardware H.264 Codec

The hardware H.264 codec was set up to optimize motion rather than detail. With this setting, the system down-sampled the image size to SIF (Source Input Format) prior to transmission and coding, and then up-sampled the image size back to Rec. 601 [5] after decoding and error concealment. No other control was possible over the frame rate or coding parameters such as how often I-frames are sent. A network impairment emulator was used to randomly drop packets of variable packet size, with an average size of 460 bytes. The H.264 operating mode of this codec was limited to 384 kbps or lower (this codec was only operated at 384 kbps). The hardware H.264 HRCs implemented error concealment, which appeared as Figure 5 illustrates. (A packet loss ratio of 3 percent was used for this example.)

Figure 3: H.264 HRCs Error Concealment Example

HRC names that used the H.264 hardware codec begin with the character “H”. The following additional abbreviations are used within the H.264 hardware HRC names to denote the bit rate and packet loss ratio:

- 384k — Coded at 384 kbps
- –0 percent — 0 percent packet loss ratio
- –1 percent — 1.0 percent packet loss ratio
- –2 percent — 2.0 percent packet loss ratio
- –3 percent — 3.0 percent packet loss ratio
- –6 percent — 6.0 percent packet loss ratio
- –12 percent — 12.0 percent packet loss ratio

4 PS1 Subjective Video Quality Experiment

The subjective video quality experiment PS1 was conducted from September 2005 to February 2006. PS1 focused on determining the minimum acceptable video transmission quality that is required to support narrow field of view tactical video applications.

During the subjective test, subjects watched a number of short video sequences and judged each of them for quality and acceptability. Subjects rated the quality of each sequence after they had seen it, using the Absolute Category Rating (ACR) grading scale: Excellent, Good, Fair, Poor, and Bad. Subjects were also asked to judge whether the video clip was acceptable or unacceptable for use as narrow field of view tactical video. Subjects had a short time to make their judgment of the video quality and acceptability and mark ratings on a printed score sheet. The first rating session lasted approximately 20 minutes, followed by a break. After the break, there was another similar session.

Subjects watched example video taken by the following equipment:

- A camera carried by a public safety practitioner into a burning building to provide the incident commander with situation information.
- A body-worn camera.
- An aerial camera following a subject on foot.

4.1 PS1 Experiment Design

The PS1 experiment included 47 HRCs to investigate the minimal requirements for the attributes that Table 1 identifies for PS1.

Sixteen source video sequences were selected, each 8 seconds in duration. These video sequences were split into two sets of eight sequences. The HRCs were likewise split into two sets of 25 HRCs.[2] Half of the video sequences were paired with half of the HRCs to produce 400 video sequences (i.e., 8 scenes × 25 HRCs × 2 sets = 400 video sequences). To prevent viewer fatigue, each viewer was asked to rate only half of the video sequences.

2. Three of the 47 HRCs were present in both sets.

The 400 clips were randomized such that no HRC or scene appeared consecutively during a viewing session. Every attempt was made to assure that the same number of viewers saw each of 12 randomized tape orderings. The clips were divided into four session tapes of 100 clips each (1, 2, 3, and 4). Each viewer saw two of these tapes in a randomized order (e.g., 4 then 2, 2 then 1, 1 then 3, etc.).

4.2 PS1 Original Video Sequences

Original video sequences were selected from the following sources:

- Footage shot at a football game with a shoulder-mounted camera that followed police officers as they performed their duties. Footage was shot in HDTV 720p 720p (p = progressive, or non-interlaced) format, and then converted to Rec. 601 [5].
- Footage shot at a firefighter training session using a shoulder-mounted camera. Footage was shot in HDTV 720p format and then converted to Rec. 601.
- In-car camera footage depicting simulated drunk driving stops. Video was shot in HDTV 1080i (i = interlaced, or non-progressive) format, and then converted to Rec. 601.
- Footage of a SWAT team training session. Footage was received on a Video Home System (VHS) tape and digitally sampled in accordance with Rec. 601.
- Footage of an underwater crime investigation. Footage was received on a VHS tape, and digitally sampled in accordance with Rec. 601.

Sixteen video sequences were selected, each 8 seconds in duration. These video sequences were split into two sets of eight sequences. Scene set A was matched with H.264 [3] codec impairments and some simulated impairments. Scene set B was matched with MPEG-2 [2] codec impairments, some simulated impairments, and several H.264 codec impairments.

Video sequences were selected to meet the following needs:

- Match the definition of narrow field of view tactical video, so that subjects could envision using the system in real time during an actual incident.
- Provide a variety of visually different scene content.
- Span a wide range of scene coding difficulty, from easy-to-code (i.e., little motion, little spatial detail) to hard-to-code (i.e., high motion, abundant and intricate details).

The video sequences within each set are described below. The original video format is in parentheses. All video sequences were converted from their original format to Rec. 601 prior to use in the experiment.

Scene Set A

1. Zoom out of a crowd in a football stadium. (720p)
2. Underwater crime scene investigation where a gun is found on a boat. Water contains many floating particles. (VHS recording media)
3. Officer watching a dense crowd of people walk past. (720p)
4. Firefighter trainee squeezing through an opening in a wall. (720p)
5. Camera bouncing while following an officer from car to simulated traffic stop; daytime. (1080i)

6. Nighttime simulated stop, zooming in on officer approaching a stopped car. The driver's face could be seen in the car's driver-side mirror, and the license plate number was visible. (1080i)
7. Walking up an outside fire escape, simulating a shoulder- or helmet-mounted camera. (720p)
8. Scene focused on a fire, which is extinguished by water, forming smoke; then panning back and zooming out to show firefighters spraying water. The camera wobbles a bit. (720p)

Scene Set B

1. Pan across a crowded stadium at a football game. (720p)
2. Night-time footage of a SWAT team deployed around a door, about to knock it down. (VHS)
3. Close-up of a woman undergoing a sobriety test. Camera zooms from showing her entire upper body to just her eyes. (1080i)
4. Firefighter backing up through a large cement pipe. (720p)
5. Camera outside a car window positioned low to the ground as the car drives along a road. (720p)
6. Following officers escorting people out of a football game; some camera flare from bright sunlight. (720p)
7. Inside a driving car with the camera pointed to the driver's side of the car, showing a policeman driving; moving scenery is visible outside the driver's window. (720p)
8. Dumpster fire with background blurred by smoke, panning to show people through the smoke. (720p)

Note that obtaining high-quality original video footage with suitable content is difficult. Due to usage restrictions, sample video frames from most of the video sequences cannot be displayed in this report.

4.3 PS1 HRC Video Transmission Systems

Codecs without error concealment were paired with packet loss ratios of 0, 0.1, 0.5, and 1.5 percent. Significantly higher levels of packet loss ratios (3.3, 5.6, and 11.5 percent) were initially considered for these codecs. However, a visual examination of sample scenes transmitted at these levels of packet loss ratio indicated a level of quality significantly below the minimum quality indicated by the questionnaire responses. Higher packet loss levels (e.g., 2, 3, 6, and 12 percent) were retained for the H.264 [3] codec with error concealment. To ensure that the minimum video quality required by practitioners was presented, the subjective test had to extend significantly beyond unacceptable (at the low-quality end) and acceptable (at the high-quality end).

Table 2 lists the synthetic HRCs that were created to explore the suitability of various frame rates and image sizes. The Scene Set column identifies the set of scenes that were used for each HRC.

Table 2: PS1 Synthetic HRCs of Various Frame Rates and Image Sizes

HRC	Scene Set	Description
original	A, B	Original unimpaired video sequence.
sif	A	Down-sampled by 2 horizontally and vertically to SIF resolution, then up-sampled back to Rec. 601 [5] image size using pixel interpolation.

Table 2: PS1 Synthetic HRCs of Various Frame Rates and Image Sizes (Continued)

HRC	Scene Set	Description
qsif	A	Down-sampled by 4 horizontally and vertically to QSIF, or Quarter SIF, resolution, then up-sampled back to Rec. 601 image size using pixel interpolation.
fps15	B	15 fps video sequence, created by discarding every other frame, and replacing each discard with a duplicate of the previous frame.
fps10	B	10 fps video sequence, created by discarding two of every three frames, and replacing them with a duplicate of the previous frame.
fps5	B	5 fps video sequence, created by discarding five of every six frames, and replacing them with a duplicate of the previous frame.
fps10sif	A, B	Create a 10 fps video sequence by discarding two of every three frames, and replacing them with a duplicate of the previous frame. Then, down-sample by 2 horizontally and vertically to SIF resolution, then up-sample back to Rec. 601 image size using pixel interpolation.
fps10qsif	A, B	Create a 10 fps video sequence by discarding two of every three frames, and replacing them with a duplicate of the previous frame. Then, down-sample by 4 horizontally and vertically to QSIF resolution, then up-sample back to Rec. 601 image size using pixel interpolation.

4.3.1 PS1 MPEG-2 Software Encoder

This MPEG-2 [2] software encoder was operated at constant bit rate (CBR) encoding with a Group of Pictures (GOP) structure equal to "I_BB_P_BB_P_BB_P_BB_P_BB_". This produced Intra-coded frames (I-frames) each second. I-frames limit the propagation of errors since they encode only for spatial redundancy (i.e., no temporal redundancy). Coding was performed using Rec. 601 image size at a frame rate of 30 fps. The MPEG-2 program stream created by the encoder was encapsulated into a transport stream (*.ts), and streamed over IP using UDP (User Datagram Protocol)/RTP (Real-time Transport Protocol) multicast. A network impairment emulator was used to randomly drop packets, 1358 bytes in size. The MPEG-2 decoder did not implement error concealment (EC). Errors generally appeared as error blocks or horizontal strips, as Figure 4 shows (a packet loss ratio of 0.1 percent was used for this example).

Figure 4: PS1 MPEG-2 HRCs with 0.1 Percent Packet Loss (No EC) Example

HRC names that used the MPEG-2 [2] software codec begin with the character "M". The following additional abbreviations are used within the MPEG-2 software HRC names to denote the coder bit rate and packet loss ratio:

- 768k — Coded at 786 kbps
- 1.5M — Coded at 1.5 Mbps
- 3.1M — Coded at 3.1 Mbps
- 6.1M — Coded at 6.1 Mbps
- –0 percent — 0 percent packet loss ratio
- –0.1 percent — 0.1 percent packet loss ratio
- –0.5 percent — 0.5percent packet loss ratio
- –1.5 percent — 1.5 percent packet loss ratio

Table 3 provides a summary description of the MPEG-2 software HRCs that were used in the PS1 experiment, including the scene set used for each HRC.

Table 3: PS1 MPEG-2 Software HRCs

HRC	Scene Set	Description
M768k-0 percent	B	Software MPEG-2, 768 kbps, 30fps, 0 percent packet loss ratio

Table 3: PS1 MPEG-2 Software HRCs (Continued)

HRC	Scene Set	Description
M768k-0.1 percent	B	Software MPEG-2, 768 kbps, 30fps, 0.1 percent packet loss ratio
M768k-0.5 percent	B	Software MPEG-2, 768 kbps, 30fps, 0.5 percent packet loss ratio
M768k-1.5 percent	B	Software MPEG-2, 768 kbps, 30fps, 1.5 percent packet loss ratio
M1.5M-0 percent	B	Software MPEG-2, 1.5 Mbps, 30fps, 0 percent packet loss ratio
M1.5M-0.1 percent	B	Software MPEG-2, 1.5 Mbps, 30fps, 0.1 percent packet loss ratio
M1.5M-0.5 percent	B	Software MPEG-2, 1.5 Mbps, 30fps, 0.5 percent packet loss ratio
M1.5M-1.5 percent	B	Software MPEG-2, 1.5 Mbps, 30fps, 1.5 percent packet loss ratio
M3.1M-0 percent	B	Software MPEG-2, 3.1 Mbps, 30fps, 0 percent packet loss ratio
M3.1M-0.1 percent	B	Software MPEG-2, 3.1 Mbps, 30fps, 0.1 percent packet loss ratio
M3.1M-0.5 percent	B	Software MPEG-2, 3.1 Mbps, 30fps, 0.5 percent packet loss ratio
M3.1M-1.5 percent	B	Software MPEG-2, 3.1 Mbps, 30fps, 1.5 percent packet loss ratio
M6.1M-0 percent	B	Software MPEG-2, 6.1 Mbps, 30fps, 0 percent packet loss ratio
M6.1M-0.1 percent	B	Software MPEG-2, 6.1 Mbps, 30fps, 0.1 percent packet loss ratio
M6.1M-0.5 percent	B	Software MPEG-2, 6.1 Mbps, 30fps, 0.5 percent packet loss ratio
M6.1M-1.5 percent	B	Software MPEG-2, 6.1 Mbps, 30fps, 1.5 percent packet loss ratio

4.3.2 PS1 H.264 Software Encoder

The H.264 software encoder used version 10.1 of the Joint Model (JM), which was developed by the Joint Video Team, a collaborative effort between MPEG and the Video Coding Experts Group (VCEG). Coding was performed using Rec.601 [5] image size at two different frame rates (15 and 30 fps), with one I-frame every second. Encapsulation was done at the video coding layer (VCL) level using the H.264 Network Abstraction Layer (NAL) header option for RTP streaming (see ITU-T Recommendation H.264 [3]). The RTP/UDP/IP streaming was based on the University College London (UCL) package. Packets 600 bytes in size were randomly dropped before decoding.[3] The decoder did not implement any error concealment. Errors appear as dropped blocks of video (i.e., black) as Figure 5 shows (a packet loss ratio of 0.1 percent was used for this example).

3. HRCs that utilized the H.264 software codec were generated by the Wireless Communications Technology Group at the National Institute of Standards and Technology (NIST).

Figure 5: PS1 H.264 HRCs with 0.1 Percent Packet Loss (No EC) Example

HRC names that used the H.264 software codec begin with the letter “S”. The following additional abbreviations are used within the H.264 software HRC names to denote the coder bit rate, the frame rate (only used when two different frame rates were examined for the same coder bit rate), and the packet loss ratio:

- 384k — Coded at 384 kbps
- 768k — Coded at 786 kbps
- 1.5M — Coded at 1.5 Mbps
- 3.1M — Coded at 3.1 Mbps
- A — Coded at 15 Mbps
- B — Coded at 30 Mbps
- –0 percent — 0 percent packet loss ratio
- –0.1 percent — 0.1 percent packet loss ratio
- –0.5 percent — 0.5percent packet loss ratio
- –1.5 percent — 1.5 percent packet loss ratio

Table 4 provides a summary description of the H.264 [3] software HRCs that were used in the PS1 experiment, including the scene set used for each HRC.

Table 4: PS1 H.264 Software HRCs Without Error Concealment

HRC	Clips Set	Description
S384k-0 percent	A	Software H.264, 384 kbps, 15 fps, 0 percent packet loss
S384k-0.1 percent	A	Software H.264, 384 kbps, 15 fps, 0.1 percent packet loss
S384k-1.5 percent	A	Software H.264, 384 kbps, 15 fps, 1.5 percent packet loss
S768kA-0 percent	A	Software H.264, 768 kbps, 15 fps, 0 percent packet loss
S768kA-0.1 percent	A	Software H.264, 768 kbps, 15 fps, 0.1 percent packet loss
S768kA-0.5 percent	A	Software H.264, 768 kbps, 15 fps, 0.5 percent packet loss
S768kA-1.5 percent	A	Software H.264, 768 kbps, 15 fps, 1.5 percent packet loss
S768kB-0 percent	A	Software H.264, 768 kbps, 30 fps, 0 percent packet loss
S768kB-0.1 percent	A	Software H.264, 768 kbps, 30 fps, 0.1 percent packet loss
S768kB-1.5 percent	A	Software H.264, 768 kbps, 30 fps, 1.5 percent packet loss
S1.5M-0 percent	A	Software H.264, 1.5 Mbps, 30 fps, 0 percent packet loss
S1.5M-0.1 percent	A	Software H.264, 1.5 Mbps, 30 fps, 0.1 percent packet loss
S1.5M-0.5 percent	A	Software H.264, 1.5 Mbps, 30 fps, 0.5 percent packet loss
S1.5M-1.5 percent	A	Software H.264, 1.5 Mbps, 30 fps, 1.5 percent packet loss
S3.1M-0 percent	A	Software H.264, 3.1 Mbps, 30 fps, 0 percent packet loss
S3.1M-0.1 percent	A	Software H.264, 3.1 Mbps, 30 fps, 0.1 percent packet loss
S3.1M-1.5 percent	A	Software H.264, 3.1 Mbps, 30 fps, 1.5 percent packet loss

4.3.3 PS1 Hardware H.264 Codec

Table 5 provides a summary description of the H.264 [3] hardware HRCs (with error concealment) that were used in the PS1 experiment, including the scene set that was used for each HRC.

Table 5: PS1 H.264 Hardware HRCs with Error Concealment

HRC	Scene Set	Description
H384k-0 percent	A	Hardware H.264, 384 kbps, error concealment, 0 percent packet loss
H384k-1 percent	A	Hardware H.264, 384 kbps, error concealment, 1 percent packet loss
H384k-2 percent	B	Hardware H.264, 384 kbps, error concealment, 2 percent packet loss
H384k-3 percent	A	Hardware H.264, 384 kbps, error concealment, 3 percent packet loss

Table 5: PS1 H.264 Hardware HRCs with Error Concealment (Continued)

HRC	Scene Set	Description
H384k-6 percent	B	Hardware H.264, 384 kbps, error concealment, 6 percent packet loss
H384k-12 percent	A	Hardware H.264, 384 kbps, error concealment, 12 percent packet loss

4.4 PS1 Viewers

Thirty-five public safety practitioners were recruited to participate in the PS1 subjective video quality experiment. The practitioners came from across the country. Various local jurisdictions (29), state jurisdictions (6), and Federal jurisdictions (2) were represented. Disciplines represented included fire response, law enforcement, and emergency medical services (EMS), with 18, 13, and 9 practitioners, respectively.[4] Three practitioners had less than 10 years of experience, 11 practitioners had 10 to 20 years of experience, 14 practitioners had 20 to 30 years of experience, and 7 practitioners had more than 30 years of experience. A total of 34 practitioners were males, and 1 was female. Roughly 25 percent were in their thirties, about 50 percent were in their forties, and 25 percent were in their fifties.

Each subject participated in two viewing sessions of 100 video clips each. Six sessions of data were discarded (i.e., one subject's scores for one tape) due to missing data (e.g., due to a power outage or missed scores) or extremely low correlation to the overall mean of the other viewer's scores (e.g., possible fatigue or inattention). The remaining data provided 16 viewer scores for each of the 400 video clips.

4.5 PS1 Data Analysis

The PS1 video data set is comprised of 400 separate video clips, which can be further broken down into 16 separate scenes. Each scene has been processed through three different codecs, designated here as M (MPEG-2 [2] software codec), S (H.264 [3] software codec) and H (H.264 hardware codec). The last codec includes error concealment, and was only operated at a coded bit rate of 384 kbps. Each video codec, with the exception just noted, was operated at multiple coder bit rates, and all video streams were subjected to multiple degrees of packet loss. In addition, lossless video streams with different image resolutions and frame rates were also presented to each practitioner.

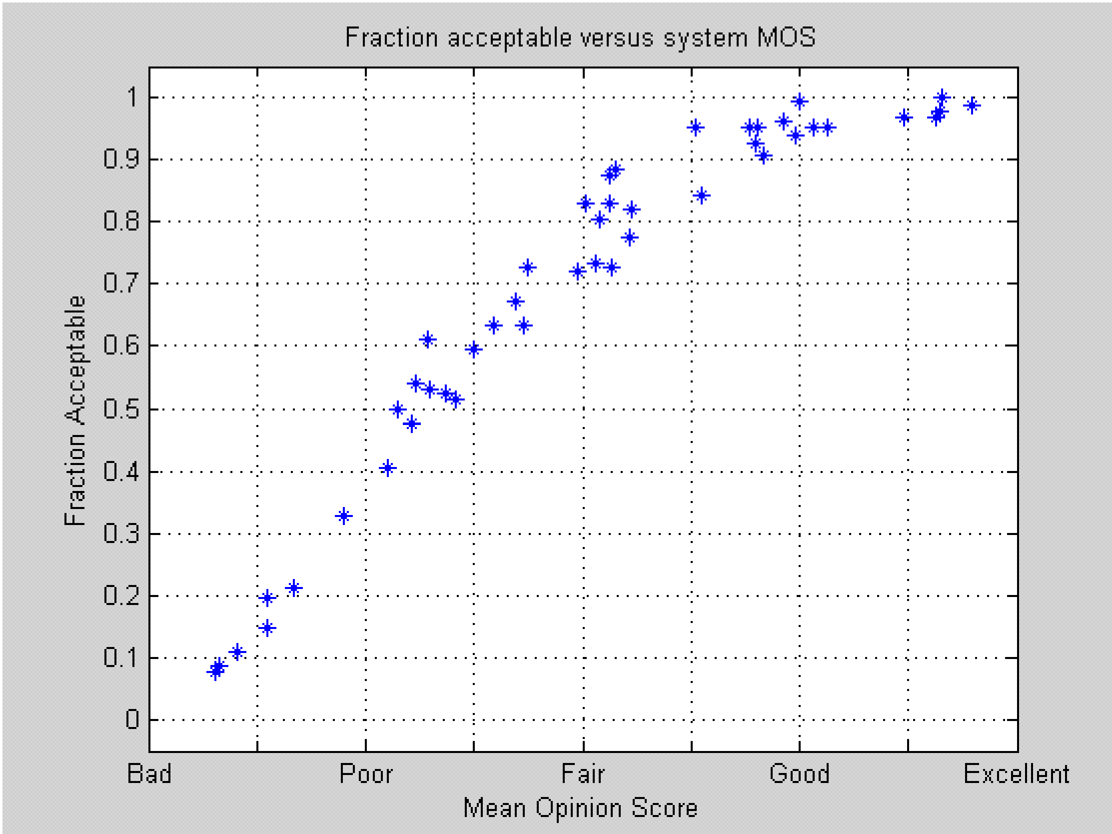
The data consists of 16 observations for each video clip. Each viewer scored each clip for Mean Opinion Score (MOS) on a 1 to 5 point scale (where 1 is "bad" and 5 is "excellent"), as well as acceptability on a 0 (non-acceptable) and 1 (acceptable) point scale. These data were aggregated in multiple ways to arrive at conclusions on viewer preferences in codec, coded bit rate, frame rate, image size, and packet loss tolerance. Some minor conclusions on error concealment are also evident from this experiment, although the paucity of data from the error concealing codec that was available means that such conclusions must be considered tentative at best.

The data presented here have been aggregated over all viewers and all scenes, so that each data point represents the performance of one HRC. Thus, the fraction acceptability values take into account variations among viewers' opinions, and variations due to changing scene content.

Figure 6 is a graph that shows the correlation between the acceptability scale and the MOS scale. The two subjective scales are very well-correlated. Although both acceptability and MOS were measured for each HRC, only acceptability will be used for the remainder of this report.

4. Some practitioners represented more than one jurisdiction and/or discipline.

Figure 6: PS1 Acceptability Scale and MOS Scale Correlation Comparison



The results are grouped into four different sets, presented in the pages that follow:

- MPEG-2 software HRCs
- H.264 software HRCs
- H.264 hardware HRCs
- Synthetic HRCs

4.5.1 PS1 MPEG-2 Software HRCs

The graphs in Figure 7 through Figure 10 give the fraction of acceptable scores for each HRC with the given packet loss ratio. The horizontal axis of each graph represents the coded bit rate. The short red vertical lines that bisect the top edges of each bar graph give the extent or range of the 95 percent confidence interval for the estimate. A table also presents the results for each of the four sets of data.

The first set of graphs and the table that follows them describe the results for the MPEG-2 [2] software HRCs. This MPEG-2 codec does not implement any EC (error concealment); it represents one of the most commonly used video coding technologies.

Figure 7: PS1 MPEG-2 Software HRCs with 0 Percent Packet Loss (No EC)

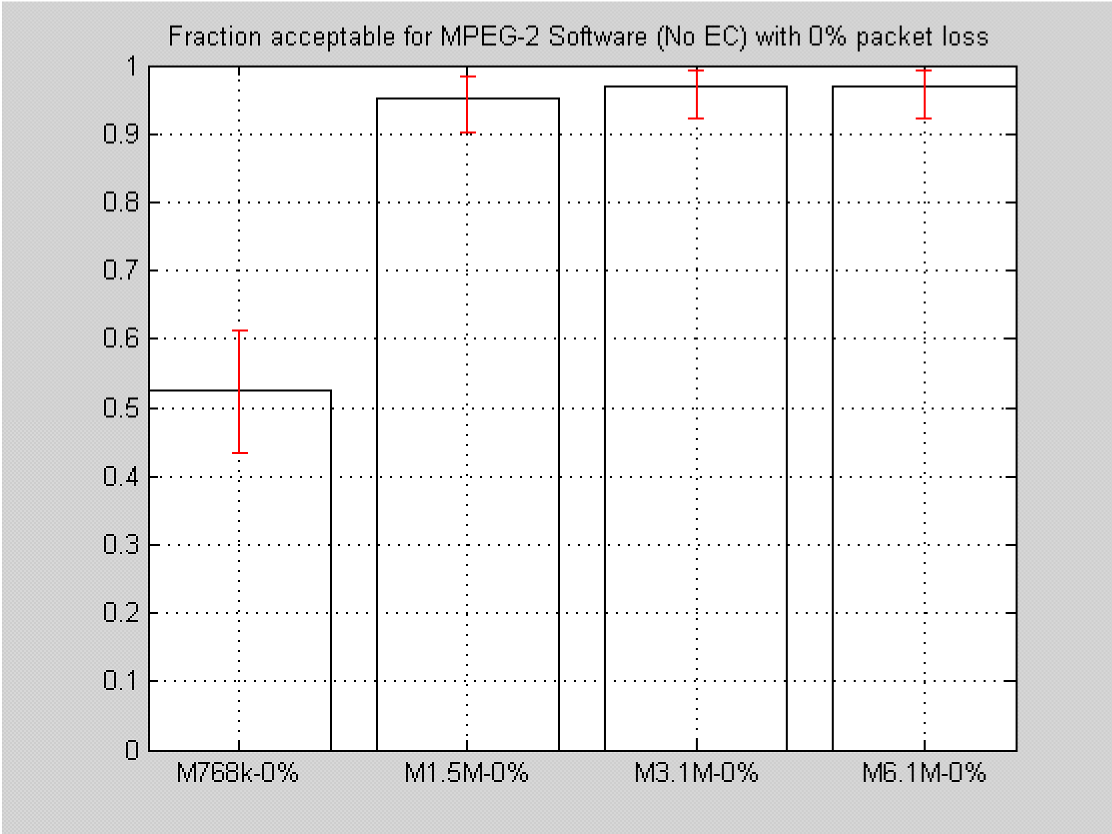


Figure 8: PS1 MPEG-2 Software HRCs with 0.1 Percent Packet Loss (No EC)

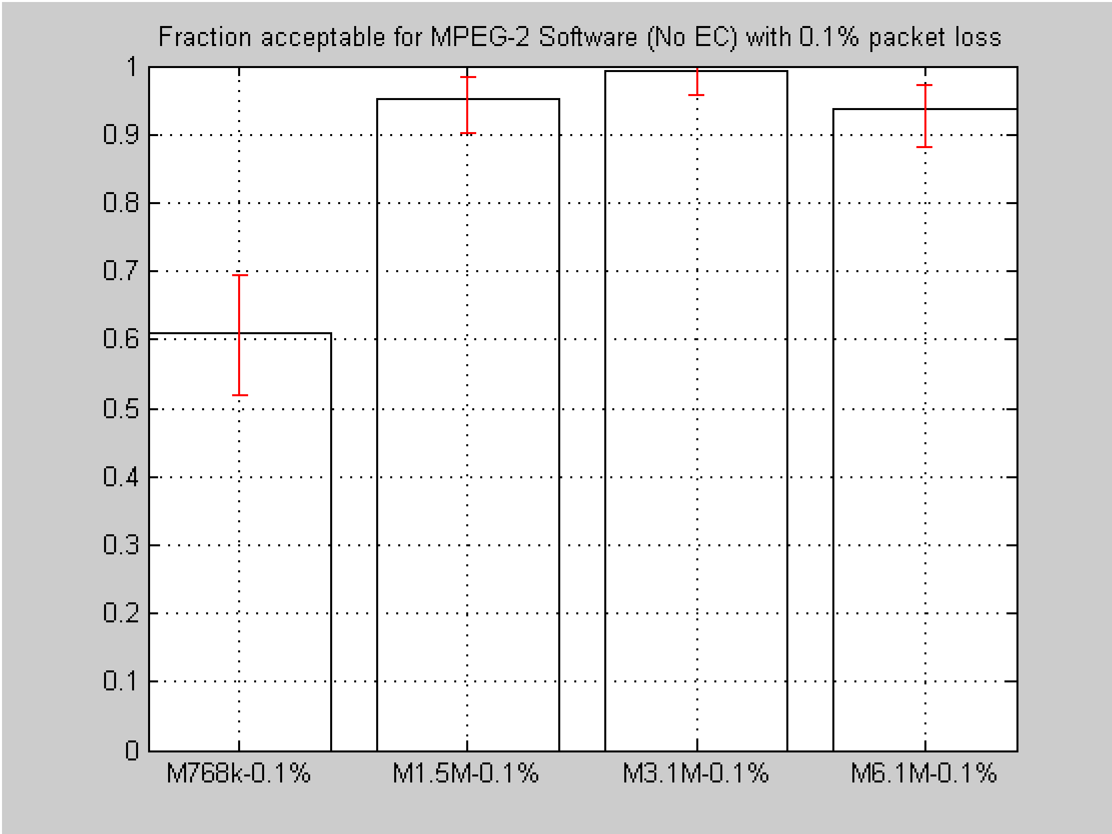


Figure 9: PS1 MPEG-2 Software HRCs with 0.5 Percent Packet Loss (No EC)

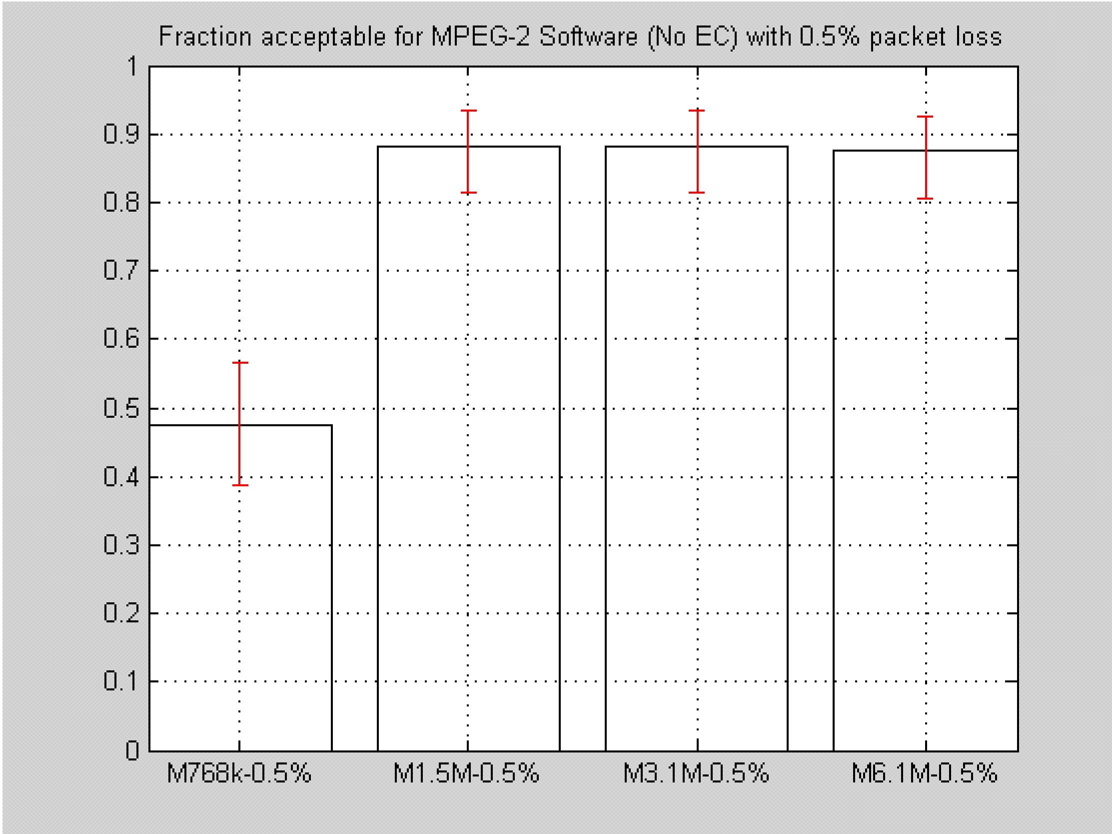


Figure 10: PS1 MPEG-2 Software HRCs with 1.5 Percent Packet Loss (No EC)

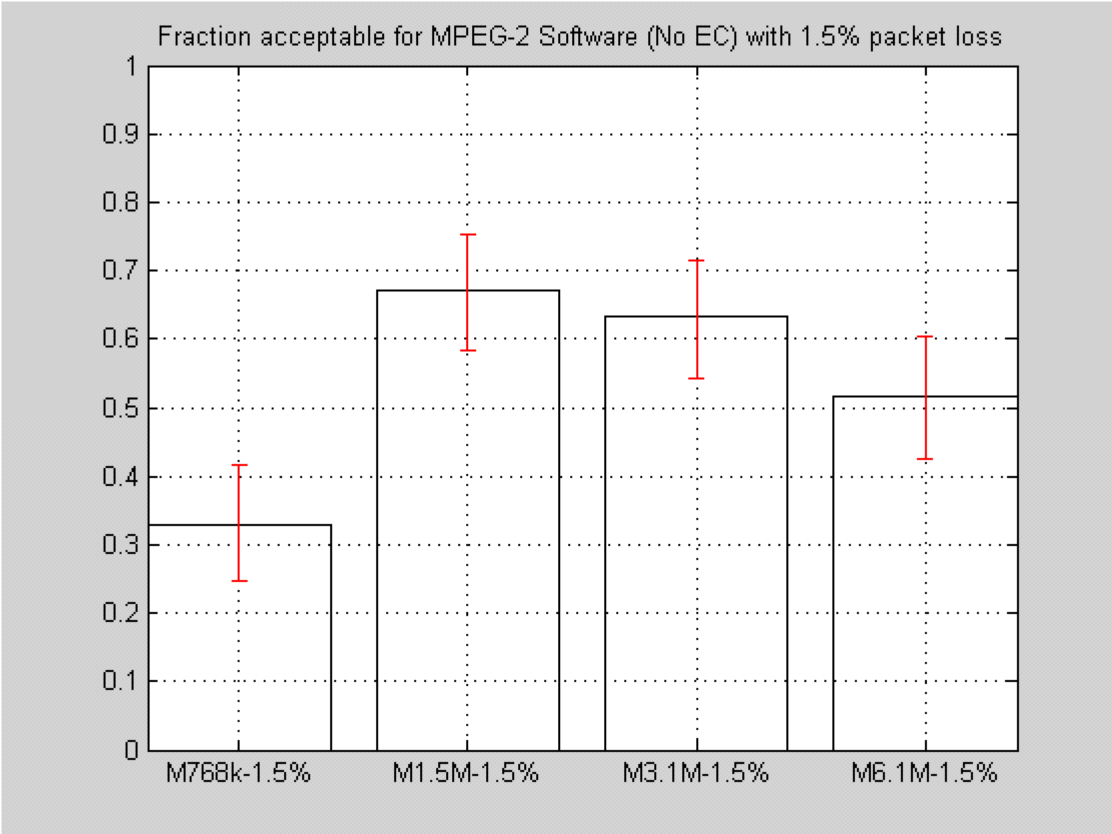


Table 6 summarizes the graphical results for the MPEG-2 [2] software HRCs presented in Figure 7 through Figure 10. The table gives a mean value of the fraction acceptable for a given bit rate and packet loss ratio. Table 6 also gives the upper and lower bounds (indicated in right-side split cells) of the 95 percent confidence interval using a threshold of 0.7 (i.e., a given HRC is declared acceptable only if the lower bound of the confidence interval is greater than 0.7). Since the bottom of the 95 percent confidence bound must be greater than 0.7, the results indicate a requirement for an MPEG-2 coded bit rate of 1.5 Mbps or more, transmitted with a packet loss rate of 0.5 percent or less.

Table 6: PS1 MPEG-2 Software HRCs Results Summary

	Fraction Acceptable (MPEG-2 Software HRCs) Packet Loss Ratio							
Bit Rate	**0.0 Percent**		**0.1 Percent**		**0.5 Percent**		**1.5 Percent**	
768 kbps	0.52	0.61	0.61	0.69	0.48	0.57	0.33	0.42
		0.43		0.52		0.39		0.25
1.5 Mbps	0.95	0.98	0.95	0.98	0.88	0.93	0.67	0.75
		0.90		0.90		0.81		0.58

Table 6: PS1 MPEG-2 Software HRCs Results Summary (Continued)

	Fraction Acceptable (MPEG-2 Software HRCs) Packet Loss Ratio							
Bit Rate	**0.0 Percent**		**0.1 Percent**		**0.5 Percent**		**1.5 Percent**	
3.1 Mbps	0.97	0.99	0.99	1.00	0.88	0.93	0.63	0.72
		0.92		0.96		0.81		0.54
6.1 Mbps	0.97	0.99	0.94	0.97	0.88	0.93	0.52	0.60
		0.92		0.88		0.81		0.43

4.5.2 PS1 H.264 Software HRCs

Figure 11 through Figure 14 and Table 7 give the results for the H.264 [3] software HRCs. This H.264 software codec does not implement any error concealment (i.e., no EC), and no effort was made to spatially decorrelate errors due to packet loss (i.e., errors in the image may appear close to one another on the screen). However, this H.264 software codec represents one of the most advanced video coding technologies currently available.

Figure 11: PS 1 H.264 Software HRCs with 0 Percent Packet Loss (No EC)

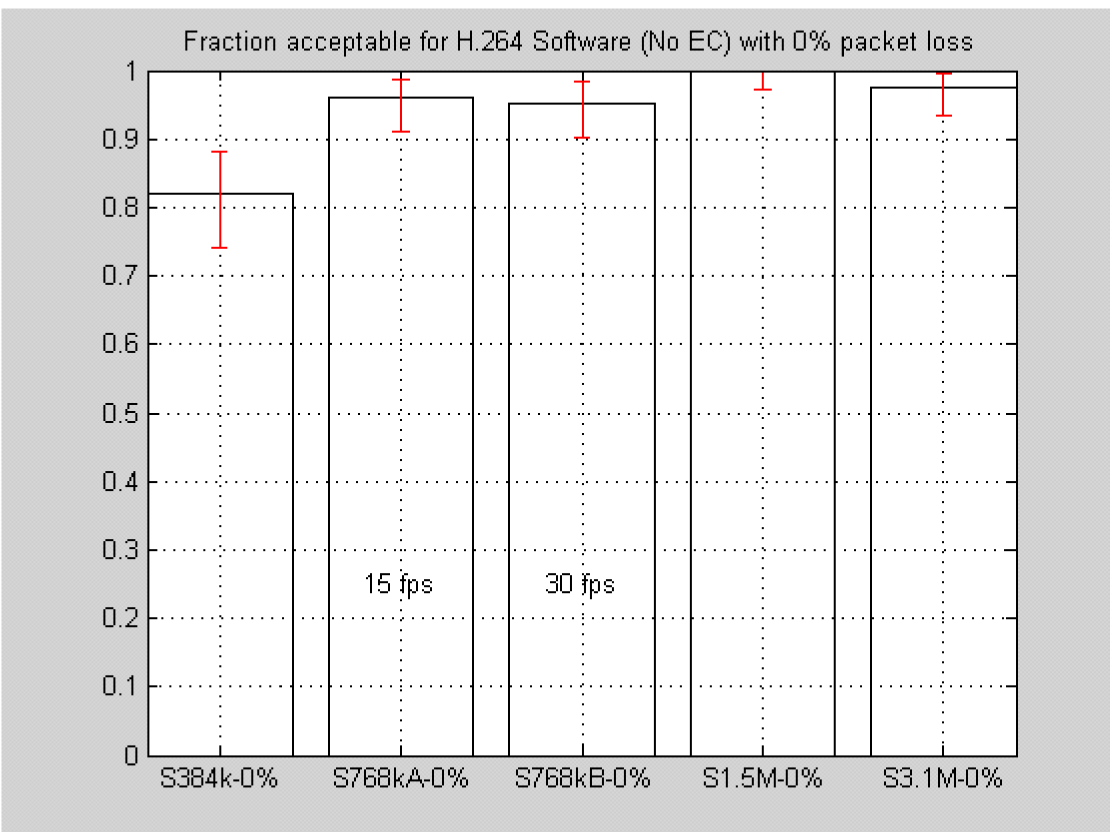


Figure 12: PS1 H.264 Software HRCs with 0.1 Percent Packet Loss (No EC)

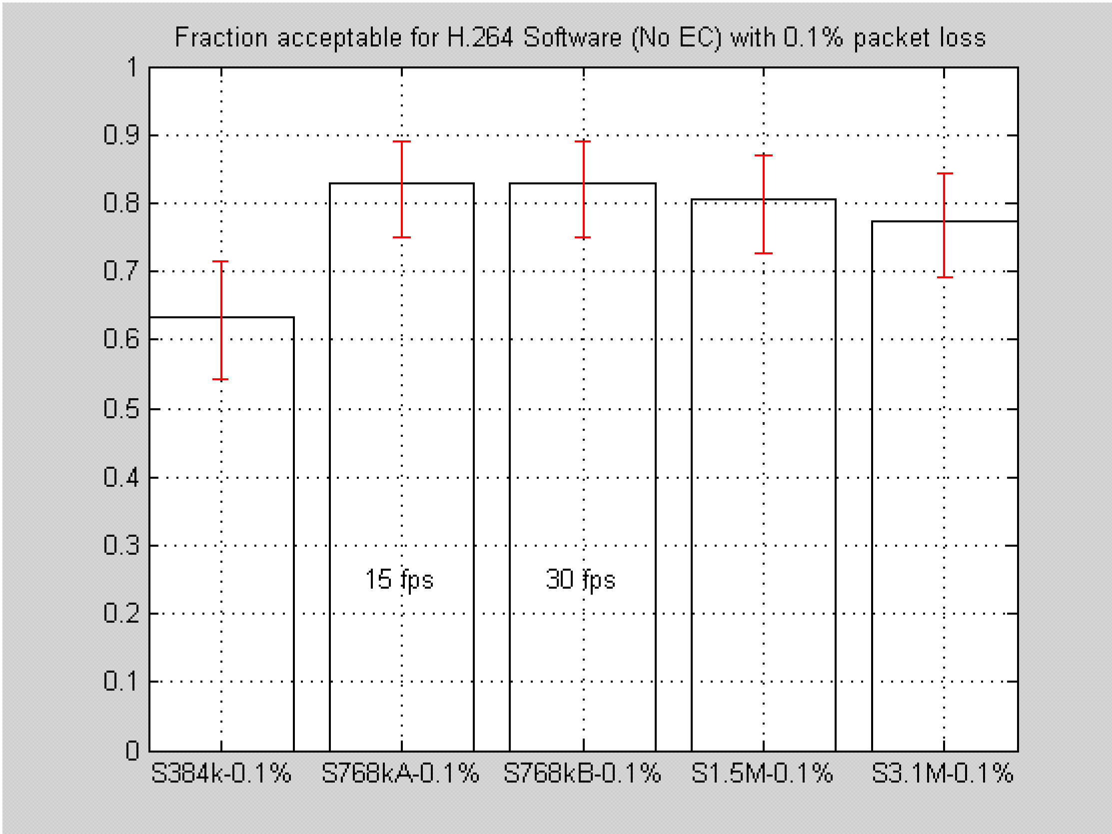


Figure 13: PS1 H.264 Software HRCs with 0.5 Percent Packet Loss (No EC)

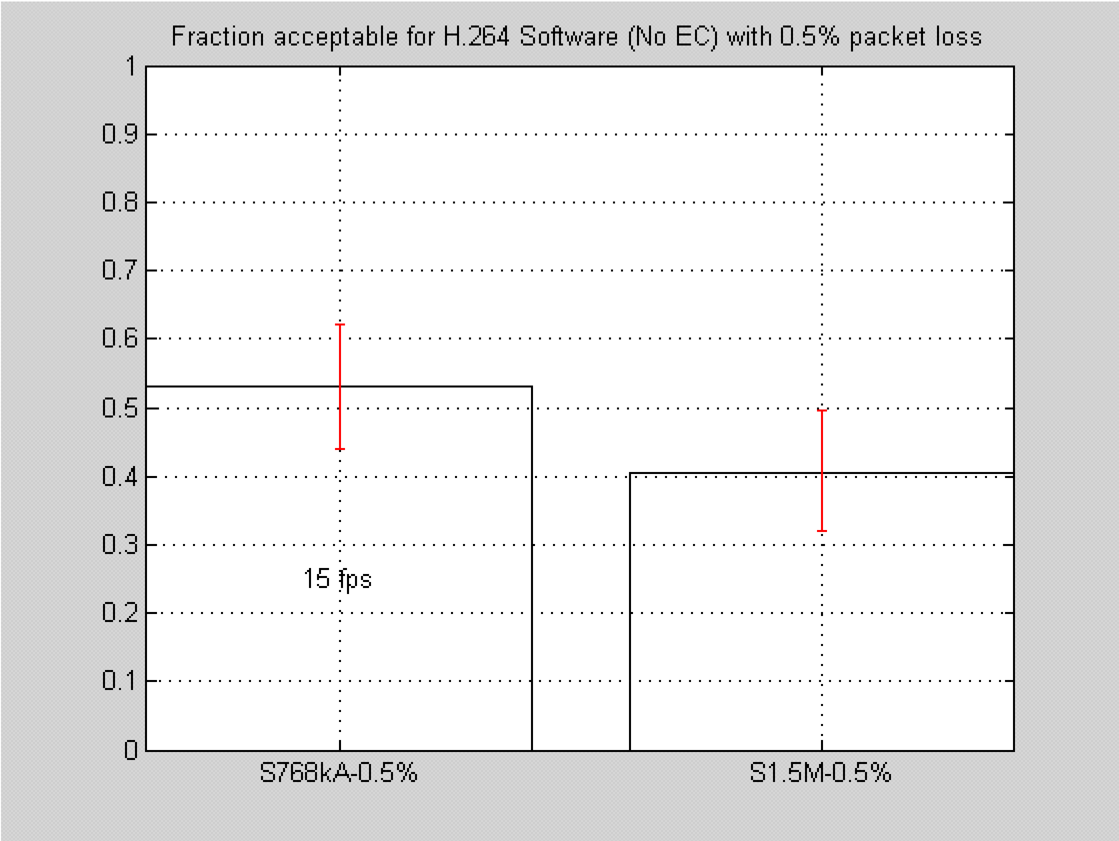


Figure 14: PS1 H.264 Software HRCs with 1.5 Percent Packet Loss (No EC)

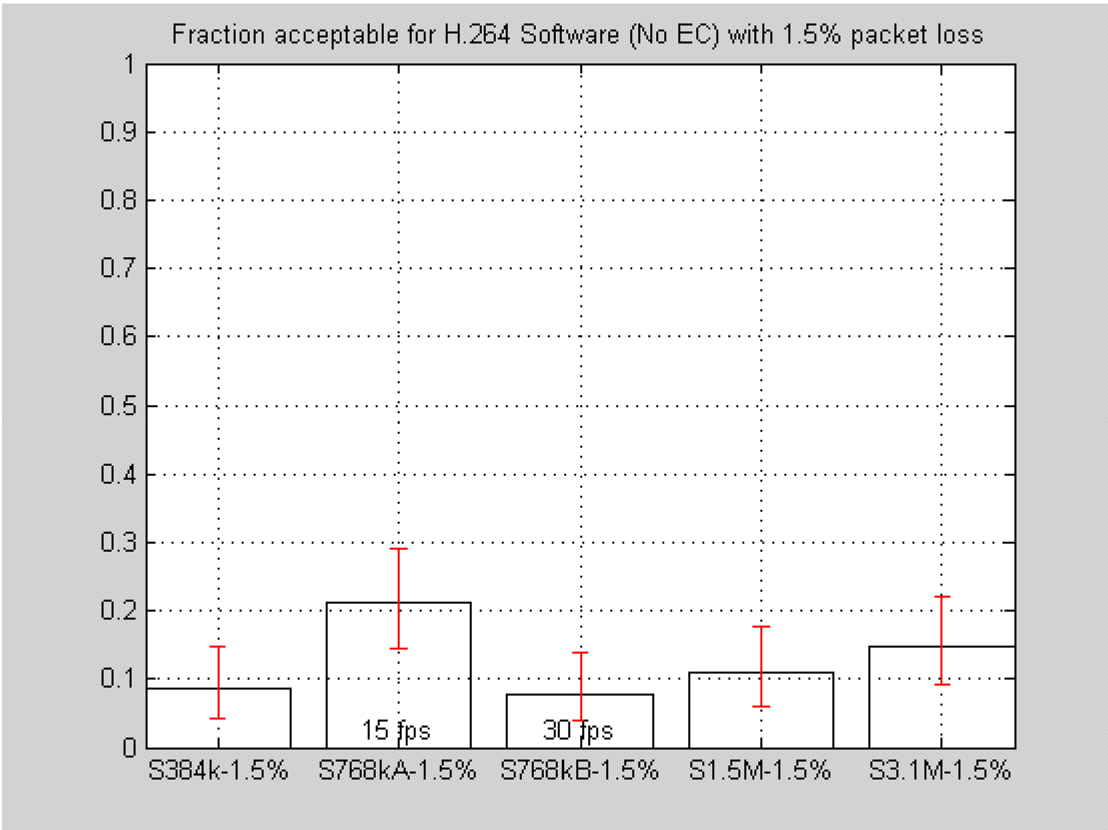


Table 7 summarizes the graphical results for the H.264 [3] software HRCs presented in Figure 11 through Figure 14. The table gives a mean value of the fraction acceptable for a given bit rate and packet loss ratio. Table 7 also gives the upper and lower bounds (indicated in right-side split cells) of the 95 percent confidence interval using a threshold of 0.7 (i.e., a given HRC is declared acceptable only if the lower bound of the confidence interval is greater than 0.7). Since the bottom of the 95 percent confidence bound must be greater than 0.7, the results indicate that a coder bit rate of 384 kbps is acceptable provided the network has no packet loss; whereas coder bit rates of 768 kbps and higher are acceptable if the packet loss ratio is 0.1 percent or less.

Table 7: H.264 PS1 Software HRCs Results Summary

	Fraction Acceptable (H.264 Software HRCs) Packet Loss Ratio							
Bit Rate Frame Rate	**0.0 Percent**		**0.1 Percent**		**0.5 Percent**		**1.5 Percent**	
384 kbps 15 fps	0.82	0.88	0.63	0.72	—	—	0.09	0.15
		0.74		0.54		—		0.04
768 kbps 15 fps	0.96	0.99	0.83	0.89	0.53	0.62	0.21	0.29
		0.91		0.75		0.44		0.14

Table 7: H.264 PS1 Software HRCs Results Summary (Continued)

	Fraction Acceptable (H.264 Software HRCs) Packet Loss Ratio							
Bit Rate Frame Rate	**0.0 Percent**		**0.1 Percent**		**0.5 Percent**		**1.5 Percent**	
768 kbps 15 fps	0.95	0.98	0.83	0.89	—	—	0.08	0.14
		0.90		0.75		—		0.04
1.5 Mbps 30 fps	1.00	1.00	0.80	0.87	0.41	0.50	0.11	0.18
		0.97		0.73		0.32		0.06
3.1 Mbps 30 fps	0.98	1.00	0.77	0.84	—	—	0.15	0.22
		0.93		0.69		—		0.09

4.5.3 PS1 H.264 Hardware HRCs

Figure 15 and Table 8 give the results for the H.264 [3] hardware HRCs. This H.264 hardware codec did implement error concealment (i.e., EC), so higher levels of packet loss ratio were considered.

Figure 15: PS1 H.264 Hardware HRCs with Packet Loss (EC)

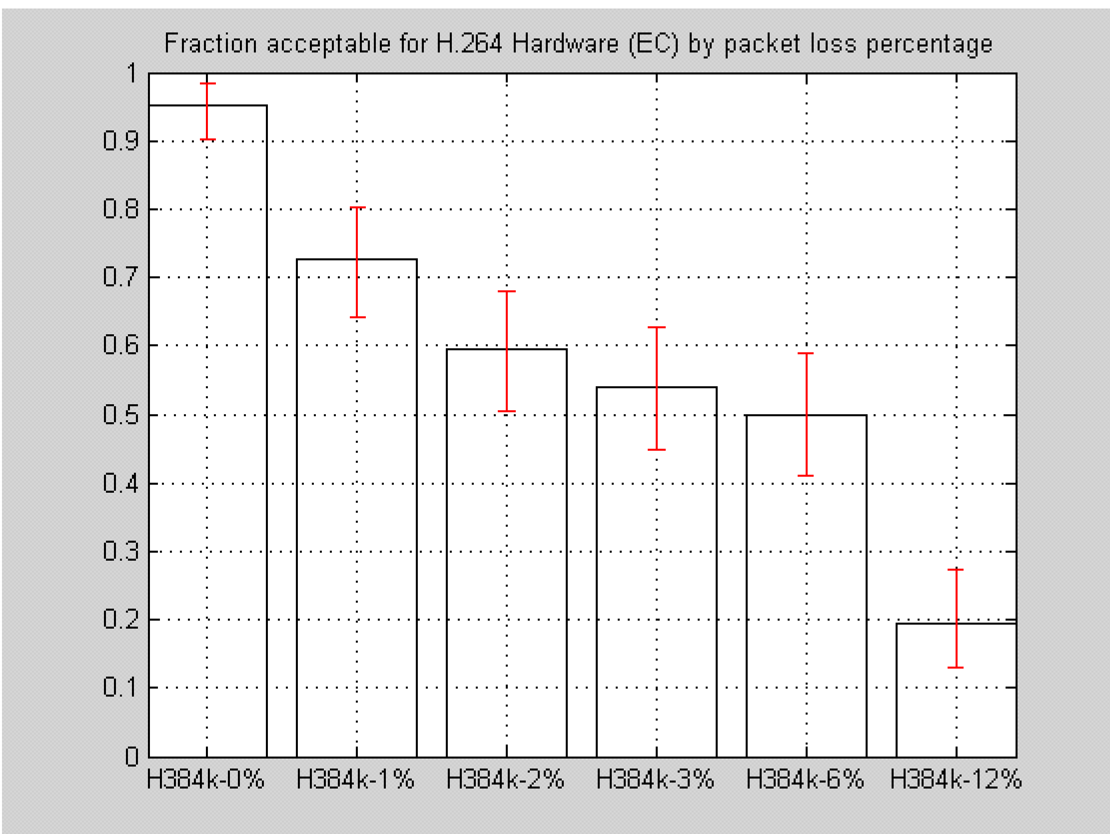


Table 8 summarizes the graphical results for the H.264 [3] hardware HRCs presented in Figure 15. The table gives a mean value of the fraction acceptable for a given bit rate and packet loss ratio. Table 8 also gives the upper and lower bounds (indicated in right-side split cells) of the 95 percent confidence interval using a threshold of 0.7 (i.e., a given HRC is declared acceptable only if the lower bound of the confidence interval is greater than 0.7). Note that the fraction acceptable decreases much more gracefully with increasing packet loss ratio versus the H.264 software HRCs that did not implement any error concealment. Since the bottom of the 95 percent confidence bound must be greater than 0.7, the results suggest that the packet loss ratios must be held below about 1 percent. Further study using higher coder bit rates and error concealment schemes are required.

Table 8: PS1 H.264 Hardware HRCs Results Summary

	Fraction Acceptable (H.264 Hardware HRCs) Packet Loss Ratio											
Bit Rate	**0 Percent**		**1 Percent**		**2 Percent**		**3 Percent**		**6 Percent**		**12 Percent**	
384 kbps	0.95	0.98	0.73	0.80	0.59	0.68	0.54	0.63	0.50	0.59	0.20	0.27
		0.90		0.64		0.50		0.45		0.41		0.13

4.5.4 PS1 Synthetic HRCs

Figure 16 and Table 9 give the results for the synthetic HRCs. These results can be used to establish values for fundamental image quality parameters such as frame rate and resolution. For Figure 16, refer to Table 2 for the definitions of the HRCs.

Figure 16: PS1 Synthetic HRCs for Frame Rate and Resolution

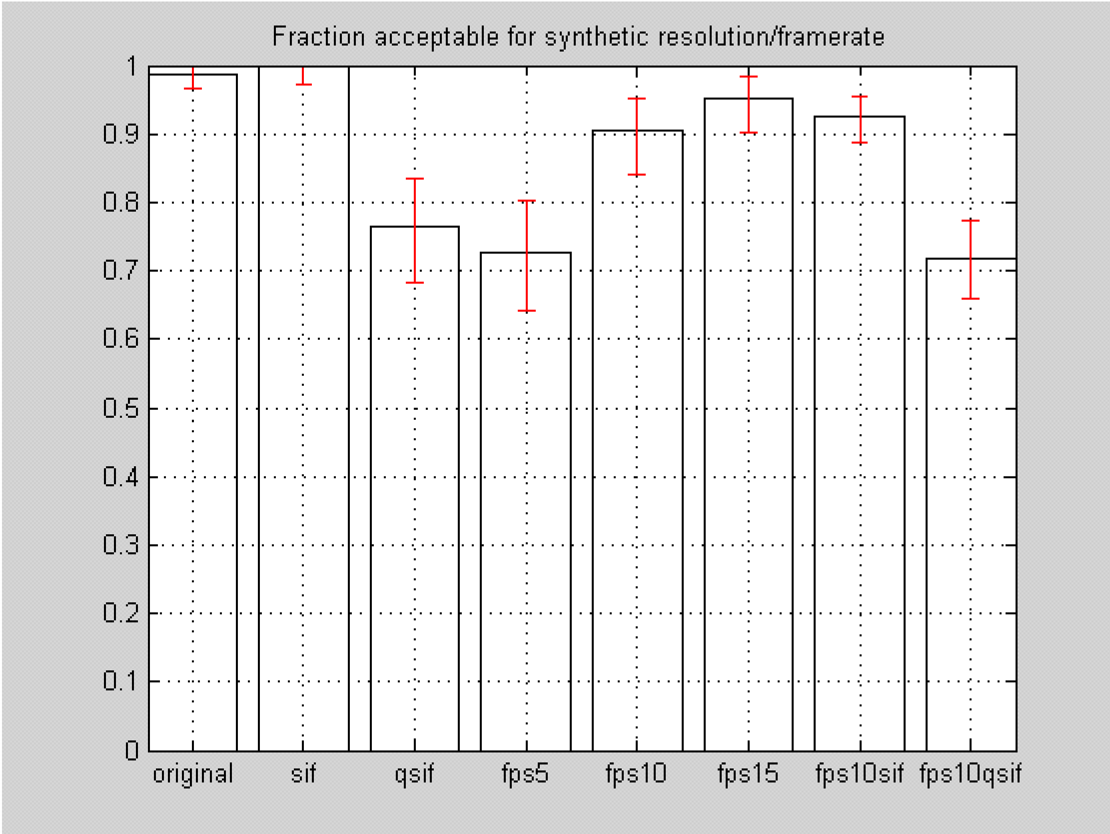


Table 9 summarizes the graphical results for the synthetic HRCs presented in Figure 16. The table gives a mean value of the fraction acceptable for each synthetic HRC. Table 9 also gives the upper and lower bounds (indicated in right-side split cells) of the 95 percent confidence interval using a threshold of 0.7 (i.e., a given HRC is declared acceptable only if the lower bound of the confidence interval is greater than 0.7). Since the bottom of the 95 percent confidence bound must be greater than 0.7, the results indicate that the frame rate should be at least 10 fps, and the image resolution should be at least SIF.

Table 9: PS1 Synthetic HRCs Summary

Fraction Acceptable (H.264 Hardware HRCs) Packet Loss Ratio															
Original		**SIF**		**QSIF**		**5 fps**		**10 fps**		**15 fps**		**10 fps SIF**		**10 fps QSIF**	
0.99	1.00	1.00	1.00	0.77	0.84	0.73	0.80	0.91	0.95	0.95	0.98	0.93	0.95	0.72	0.77
	0.97		0.97		0.68		0.64		0.84		0.90		0.89		0.66

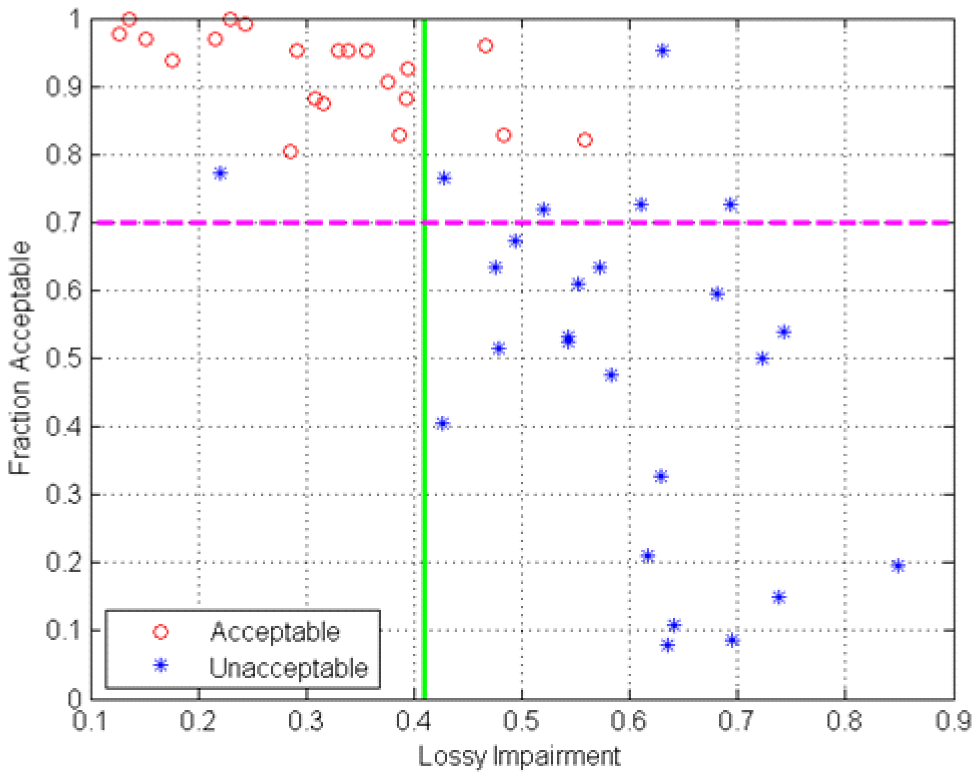
4.5.5 Fraction Acceptable Versus Lossy Impairment Metric

Figure 17 gives a plot of Fraction Acceptable versus the Lossy Impairment metric (see Section 3.5.2.2 of the PS SoR Volume II [1]) for the 47 HRCs in the PS1 experiment. If an acceptability threshold (see Section 3.3.7 of the PS SoR Volume II) of 0.7 is used, that is, the pink horizontal line in the plot, the 47

HRCs can be categorized as either acceptable (the red circles) or unacceptable (the blue asterisks). Please recall that the acceptability threshold is applied to the lower bound of the 95 percent confidence interval. This is so that an HRC can have a fraction acceptable that is greater than 0.7, but still not be acceptable because the lower bound of the 95 percent confidence falls below 0.7.

The vertical green line represents a measured lossy impairment value of 0.41. HRCs that fall to the left of the green line would be acceptable, whereas HRCs that fall to the right of the green line would be unacceptable.

Figure 17: PS1 Fraction Acceptable Versus the Lossy Impairment Metric



The Lossy Impairment metric misclassifies four of the 47 HRCs. The first of these errors is S3.1M-0.1 (see Table 4) located at (Lossy Impairment = 0.22, Fraction Acceptable = 0.77). The Fraction Acceptable indicates this HRC as being unacceptable, while the Lossy Impairment metric indicates it being acceptable. The lower confidence bound for HRC S3.1M-0.1 is 0.69, which is just below the 0.7 acceptability threshold. Thus, S3.1M-0.1 was almost categorized as acceptable by subjective testing, and this particular error is benign. Generally speaking, however, this type of error is to be avoided, because errors of this sort would lead public safety practitioners to purchase inadequate equipment. Thus, the recommended value for the Lossy Impairment metric should be chosen conservatively (i.e., lower) to avoid this type of error.

For the other three errors, the Fraction Acceptable indicates acceptable, while the Lossy Impairment metric indicates unacceptable. The first is S384k-0 located at (0.56, 0.82), where the lower confidence bound was 0.74. The second is S768kA-0 located at (0.47, 0.96), and the lower confidence bound was 0.91. The third is S768kA-0.1 located at (0.48, 0.83), and the lower confidence bound was 0.75. Thus, two of these errors involve a lower confidence bound that is fairly close to the 0.7 acceptability threshold. Errors of this sort will cause public safety practitioners to avoid purchasing equipment that may have adequate video quality for their application. These errors are considered less problematic than the previous type of errors.

As a note, the Lossy Impairment metric has been approved by ANSI for general use in objectively quantifying the perceptual impact of errors due to video codecs. That is why this particular metric has been placed in Section 3.5.2 of the PS SoR Volume II [1]. The Lossy Impairment metric has not been generally approved for use when network transmission errors (like packet loss) are present. However, the Lossy Impairment metric is currently one of the best available metrics for objectively measuring the perceptual video quality of digital video systems. The results presented above provide some support that the metric could have value in more general settings (e.g., as in the presence of network transmission errors).

4.6 PS1 Conclusions

The results presented in this report suggest that SIF image resolution should be the minimum specification for narrow field of view tactical video applications, particularly if any coding or transmission loss is present. The minimum frame rate should be at least 10 fps. The choice of codec is dependant upon available network capacities.

If capacities of 1.5 Mbps or higher are available, then an MPEG-2 [2] codec (with no error concealment) could be chosen, provided the packet loss ratio is kept to 0.5 percent. For networks with lower capacities (768 kbps), an H.264 [3] codec is the better choice as long as the packet loss ratio can be held to 0.1 percent or less (for H.264 codecs with no error concealment). H.264 systems operating at 384 kbps were acceptable for a packet loss ratio of 0 percent (i.e., no packet loss).

Preliminary results from this experiment indicate that acceptable packet loss ratios for H.264 codecs that perform error concealment are on the order of 1 percent, but further study is required in this area. Since error concealment algorithms are constantly being improved, these further studies should include state-of-the-art error concealment algorithms.

The proposed value for the Lossy Impairment metric (0.41) produces a fairly clean classification of the acceptable and unacceptable HRCs for an Acceptability Threshold of 0.7.

Results from the questionnaire (see Section 2) indicate that the one-way video delay for narrow field of view tactical video systems be kept to a maximum of 1 second.

5 PS2 Subjective Video Quality Experiment

The subjective video quality experiment PS2 was conducted in from August to September in 2006. Five tests constitute the subjective video quality experiment PS2. Table 10 breaks down how three of the five tests used the same HRCs and the same number of scenes (Group 1), while two of the five tests used the same HRCs and the same number of scenes (Group 2). Note that in the sections that follow “Observed,” e.g. “Observed surveillance,” is the same as “Live.”

Table 10: PS2 Experiment Test Groups

Group	Test Name	Description	Reference
1	PS2TW	Tactical wide field of view	See Section 5.5.
	PS2ON	Observed surveillance narrow field of view	See Section 5.6.
	PS2OW	Observed surveillance wide field of view	See Section 5.7.

Table 10: PS2 Experiment Test Groups (Continued)

Group	Test Name	Description	Reference
2	PS2RN	Recorded surveillance narrow field of view	See Section 5.8.
	PS2RW	Recorded surveillance wide field of view	See Section 5.9.

5.1 PS2 Experiment Test Design

Nine 20-minute viewing sessions were split among the five PS2 subjective tests. All tests were performed on 525-line NTSC video. During the tests, subjects watched a number of short video sequences and judged each of them for quality and acceptability. Subjects were given seven seconds to rate the quality of each sequence after they had seen it, using the ACR grading scale: Excellent, Good, Fair, Poor, and Bad. Subjects were also asked to judge whether the video clip was acceptable or unacceptable for use. Subjects had seven seconds to make their judgment of the video quality and acceptability and indicate ratings on a computerized form. The first rating session lasted approximately 20 minutes, followed by a break. A second 20-minute rating session followed the break.

Table 11 lists the examples given to subjects, and the tests to which the examples apply.

Table 11: PS2 Experiment Test Groups

Camera Example	Applicable Test
A camera carried by a public safety practitioner into a burning building to provide the incident commander with situation information A body worn camera during a SWAT raid An aerial camera following a subject on foot A camera mounted on a robot being used to dismantle a bomb	PS2TW
Monitoring an entry way for suspicious activity Cameras covering a football stadium that allow a police officer to watch for disturbances remotely Cameras in a bank lobby that can show people’s location and activities	PS2ON, PS2OW
Documenting routine incidents such as a prisoner transfer or traffic stop Video from an ATM camera that records a crime Video from a convenience store that has been robbed	PS2RN, PS2RW

Original video sequences were selected from the following sources:

- Footage shot at a football game with a shoulder-mounted camera that followed police officers as they performed their duties. Footage was shot in HDTV 720p format, and then converted to Rec. 601 [5].
- Footage shot at a firefighter training session using a shoulder-mounted camera. Footage was shot in HDTV 720p format and then converted to Rec. 601.
- In-car camera footage depicting simulated drunk driving stops. Video was shot in HDTV 1080i format, and then converted to Rec. 601.

- Footage of a SWAT team training session. Footage was received on a VHS tape and digitally sampled in accordance with Rec. 601.
- Footage of an underwater crime investigation. Footage was received on a VHS tape, and digitally sampled in accordance with Rec. 601.
- Footage that was purchased from video footage clearing houses.
- Footage that was shot on HD cameras at a mountain wildfire training.

5.2 PS2 Group 1

The following sections present information specific to Group 1 PS2 tests, which consisted of the PS2TW, PS2ON, and PS2OW tests (see Table 10).

5.2.1 Group 1 Test Design

Group 1 PS2 tests included 17 impaired HRCs plus the unimpaired sources to investigate the minimal requirements for the attributes. HRCs were selected for each experiment to investigate the minimal requirements for the video attributes lists that Table 1 identifies for PS2.

Sixteen source video sequences were selected for each test in Group 1, each 8 seconds in duration. These video sequences were split into two sets of eight sequences. Two sets of HRCs were developed—one set of 8 HRCs involved frame rate and resolution variations; the other set of 10 HRCs involved compression and packet loss variations. One set of sequences was paired with one set of HRCs to produce 144 video sequences (i.e., 8 scenes × 18 HRCs = 144 video sequences). The original sources were added to the scene pools for a total of 160 clips. To prevent viewer fatigue, each viewer was asked to rate only half of the video sequences.

Video sequences were selected to meet the following needs:

- Match the definition of wide field of view tactical video, narrow field of view observed surveillance video, and wide field of view observed surveillance video, so that subjects could envision using the system in real time during an actual incident.
- Provide a variety of visually different scene content.
- Span a wide range of scene coding difficulty, from easy-to-code (i.e., little motion, little spatial detail) to hard-to-code (i.e., high motion, abundant and intricate details).

The 160 clips were randomized such that no HRC or scene appeared consecutively during a viewing session. The clips were divided into four session tapes (1, 2, 3, and 4) of 40 clips each. Each viewer saw two of these tapes in a randomized order (e.g., 4 then 2, 2 then 1, 1 then 3, etc.). Every attempt was made to assure that the same number of viewers saw each of 12 randomized tape orderings.

5.2.2 Group 1 Synthetic HRCs

Table 12 lists the synthetic HRCs that were created to explore the suitability of various frame rates and image sizes. The Scene Set column identifies the set of scenes that were used for each HRC.

Table 12: Group 1 PS2 Synthetic HRCs of Various Frame Rates and Image Sizes

HRC	Scene Set	Description
original	A, B	Original unimpaired video sequence.

Table 12: Group 1 PS2 Synthetic HRCs of Various Frame Rates and Image Sizes (Continued)

HRC	Scene Set	Description
sif	A	Down-sample by 2 horizontally and vertically to SIF resolution, then up-sampled back to Rec. 601 [5] image size using pixel interpolation.
qsif	A	Down-sample by 4 horizontally and vertically to QSIF resolution, then up-sampled back to Rec. 601 image size using pixel interpolation.
fps15	A	15 fps video sequence, created by discarding every other frame, and replacing each discard with a duplicate of the previous frame.
fps10	A	10 fps video sequence, created by discarding two of every three frames, and replacing them with a duplicate of the previous frame.
fps5	A	5 fps video sequence, created by discarding five of every six frames, and replacing them with a duplicate of the previous frame.
fps15sif	A	Create a 15 fps video sequence by discarding two of every three frames, and replacing them with a duplicate of the previous frame. Then, down-sample by 2 horizontally and vertically to SIF resolution, then up-sample back to Rec. 601 image size using pixel interpolation.
fps10sif	A	Create a 10 fps video sequence by discarding two of every three frames, and replacing them with a duplicate of the previous frame. Then, down-sample by 4 horizontally and vertically to SIF resolution, then up-sample back to Rec. 601 image size using pixel interpolation.

5.2.3 Group 1 Hardware H.264 Codec

Table 13 provides a summary description of the H.264 [3] hardware HRCs (with no error concealment) that were used in the PS2TW experiment, including the scene set that was used for each HRC.

Codecs without error concealment were paired with packet loss ratios of 0, 0.1, 0.5, and 1 percent. To ensure that the minimum video quality that practitioners required was presented, the subjective test had to extend significantly beyond unacceptable (at the low-quality end) and acceptable (at the high-quality end).

Table 13: Group 1 PS2 H.264 Hardware HRCs with Error Concealment

HRC	Scene Set	Description
CIFx128	B	Hardware H.264, 128 kbps, no error concealment, SIF resolution, no packet loss
CIFx384	B	Hardware H.264, 384 kbps, no error concealment, SIF resolution, no packet loss
CIFx786	B	Hardware H.264, 786 kbps, no error concealment, SIF resolution, no packet loss
CIFx1500	B	Hardware H.264, 1.5 Mbps, no error concealment, SIF resolution, no packet loss
NTSCx786	B	Hardware H.264, 786 kbps, no error concealment, NTSC resolution, no packet loss

Table 13: Group 1 PS2 H.264 Hardware HRCs with Error Concealment (Continued)

HRC	Scene Set	Description
NTSCx1500	B	Hardware H.264, 1.5 Mbps, no error concealment, NTSC resolution, no packet loss
NTSCx1500p01	B	Hardware H.264, 1.5 Mbps, no error concealment, NTSC resolution, .01 percent packet loss
NTSCx1500p05	B	Hardware H.264, 1.5 Mbps, no error concealment, NTSC resolution, .05 percent packet loss
NTSCx1500p10	B	Hardware H.264, 1.5 Mbps, no error concealment, NTSC resolution, 1 percent packet loss
NTSCx3000	B	Hardware H.264, 3 Mbps, no error concealment, NTSC resolution, no packet loss

5.3 PS2 Group 2

The next sections present information specific to Group 2 PS2 tests, which consisted of the PS2RN and PS2RW tests (see Table 10).

5.3.1 Group 2 Test Design

Group 2 PS2 tests included 12 impaired HRCs. HRCs were selected for each experiment to investigate the minimal requirements for the video attributes lists that Table 1 identifies for PS2. The HRCs were divided into two sets of six each.

Ten source video sequences were selected, each 8 seconds in duration. Half (5) of the video sequences (scene sets A and B) were paired with each HRC set to produce 30 video sequences (i.e., 5 scenes × 6 HRCs × 2 sets = 60 video sequences). The original sources were added to the scene pools for a total of 70 clips. To prevent viewer fatigue, each viewer was asked to rate only half of the video sequences.

The 70 clips were randomized such that no HRC or scene appeared consecutively during a viewing session. The clips were divided into two session tapes (1, 2) of 35 clips each. Each viewer saw two of these tapes in a randomized order (e.g., 1 then 2, 2 then 1). Every attempt was made to assure that the same number of viewers saw each of randomized tape orderings.

Scene set A was matched with simulated impairments. Scene set B was matched with several H.264 [3] codec impairments and network impairments (packet loss).

Video sequences were selected to meet the following needs:

- Match the definition of recorded surveillance video, so that subjects could envision using the system in real time during an actual incident.
- Provide a variety of visually different scene content.
- Span a wide range of scene coding difficulty, from easy-to-code (i.e., little motion, little spatial detail) to hard-to-code (i.e., high motion, abundant and intricate details).

5.3.2 Group 2 Synthetic HRCs

Table 14 lists the synthetic HRCs that were created to explore the suitability of various frame rates and image sizes. The Scene Set column identifies the set of scenes that were used for each HRC.

Table 14: PS2 Group 2 Synthetic HRCs of Various Frame Rates and Image Sizes

HRC	Scene Set	Description
original	A, B	Original unimpaired video sequence
sif	A	Down-sample by 2 horizontally and vertically to SIF resolution, then up-sampled back to Rec. 601 [5] image size using pixel interpolation
qsif	A	Down-sample by 4 horizontally and vertically to QSIF resolution, then up-sampled back to Rec. 601 image size using pixel interpolation
fps15	A	15 fps video sequence, created by discarding every other frame, and replacing each discard with a duplicate of the previous frame
fps10	A	10 fps video sequence, created by discarding two of every three frames, and replacing them with a duplicate of the previous frame
fps5	A	5 fps video sequence, created by discarding five of every six frames, and replacing them with a duplicate of the previous frame
fps1	A	1 fps video sequence, created by discarding 29 of every 30 frames, and replacing them with a duplicate of the previous frame

5.3.3 Group 2 Hardware H.264 Codec

Table 15 provides a summary description of the H.264 [3] hardware HRCs (with no error concealment) that were used in the PS Group 2 experiment, including the scene set that was used for each HRC.

Codecs without error concealment were paired with packet loss ratios of 0, 0.1, and 1 percent. To ensure that the minimum video quality required by practitioners was presented, the subjective test had to extend significantly beyond unacceptable (at the low-quality end) and acceptable (at the high-quality end).

Table 15: PS2 Group 2 H.264 Hardware HRCs with Error Concealment

HRC	Scene Set	Description
CIFx384	B	Hardware H.264, 384 kbps, no error concealment, SIF resolution, no packet loss
NTSC786	B	Hardware H.264, 786 kbps, no error concealment, NTSC resolution, no packet loss
NTSCx1500	B	Hardware H.264, 1.5 Mbps, no error concealment, NTSC resolution, no packet loss
NTSCx3000	B	Hardware H.264, 3 Mbps, no error concealment, NTSC resolution, no packet loss
NTSCx3000p01	B	Hardware H.264, 3 Mbps, no error concealment, NTSC resolution, 0.1 percent packet loss

Table 15: PS2 Group 2 H.264 Hardware HRCs with Error Concealment (Continued)

HRC	Scene Set	Description
NTSCx3000p10	B	Hardware H.264, 3 Mbps, no error concealment, NTSC resolution, 1 percent packet loss

5.4 PS2 Viewers

Twenty-four public safety practitioners were recruited to participate in the PS2 subjective video quality experiment. The practitioners came from across the country. Various local jurisdictions, state jurisdictions, and Federal jurisdictions were represented. Disciplines represented included fire response, law enforcement, and EMS.

Subjects were screened for color perception and visual acuity at a distance of 10 feet. Each subject participated in two viewing sessions.

The next sections describe each PS2 test.

5.5 PS2TW Tactical Wide Field of View Test

PS2TW focused on determining the minimum acceptable video transmission quality that is required to support wide field of view tactical video applications. Before the test, viewers underwent a training session where they were reminded that public safety personnel use tactical video in real time during an incident to make decisions on how to respond.

5.5.1 PS2TW Original Video Sequences

Sixteen video sequences were selected, each 8 seconds in duration. These video sequences were split into two sets of eight sequences. Scene set A was matched with simulated impairments. Scene set B was matched with several H.264 [3] codec impairments. These HRC sets are described in Table 12 and Table 13.

The video sequences within each set are described below. All video sequences were converted from their original format to Rec. 601 [5] prior to use in the experiment.

Scene Set A

1. Outside, police pull over car, license plate visible
2. From inside truck, drives around with lights on
3. Zoom in on freeway, follow truck with flashing lights
4. Pan of mock airplane disaster, victims on blankets
5. Officers escort football team through crowd
6. Riot police at protest, sweep of incident scene
7. Helicopter flying at night, spotlight on highway, building, and officers
8. Wildfire, huge, movement of firefighter and truck

Scene Set B

1. EMS technicians load ambulance, night, and flashing lights
2. Real fire, night, assist people out of burning building
3. Pan parking lot, zoom in on white van
4. Forest mountain rescue, descend rope
5. House collapse, firefighters dig out vehicle
6. Night time stop, man's face clearly visible in mirror
7. Medic aids firefighter
8. Wildfire and truck fire, wide angle of many firefighters

5.5.2 PS2TW Data Analysis

The data presented here have been aggregated over all viewers and all scenes, so that each data point represents the performance of one HRC. Thus, the fraction acceptability values take into account variations among viewers' opinions, and variations due to changing scene content.

The results are grouped into two different sets, presented in the pages that follow:

- H.264 hardware HRCs
- Synthetic HRCs

5.5.2.1 PS2TW H.264 Hardware HRCs

Figure 18, Figure 19, and Figure 20 give the results for the H.264 [3] hardware HRCs. (CI in these figures = confidence interval.)

Figure 18: Fractional Acceptance PS2TW H.264 Hardware HRCs vs. Throughput Rate for NTSC

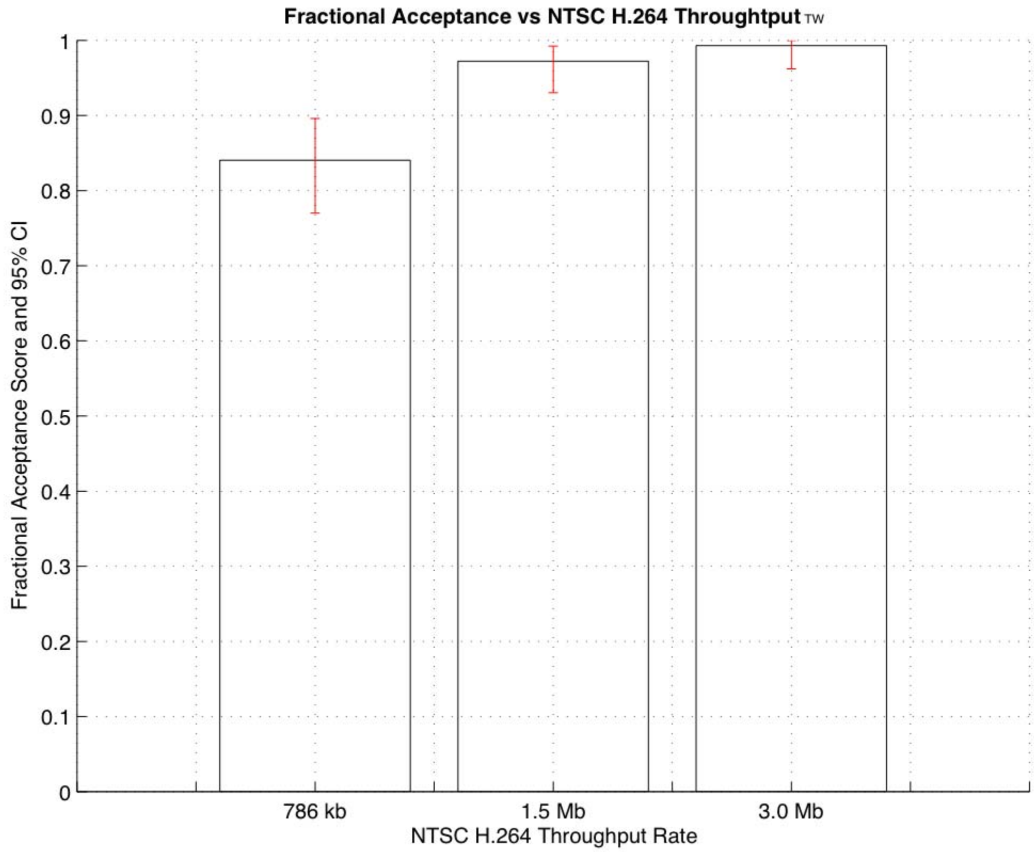


Figure 19: Fractional Acceptance PS2TW H.264 Hardware HRCs vs. Throughput Rate for CIF

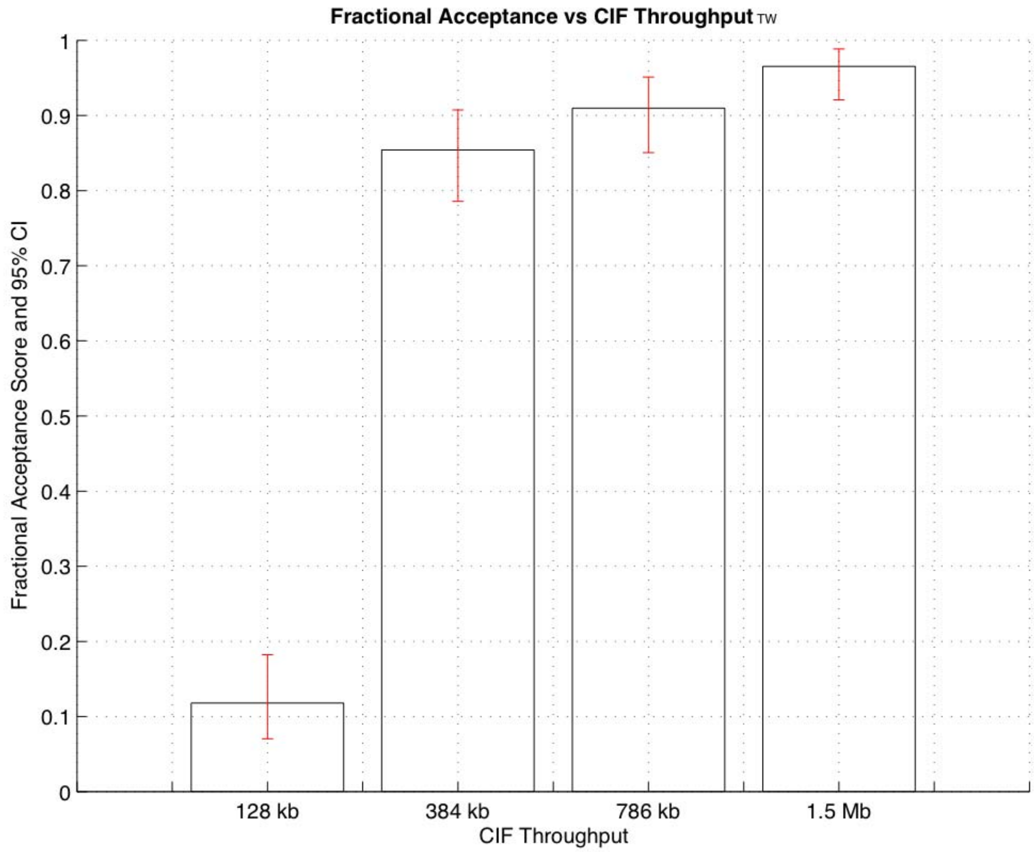
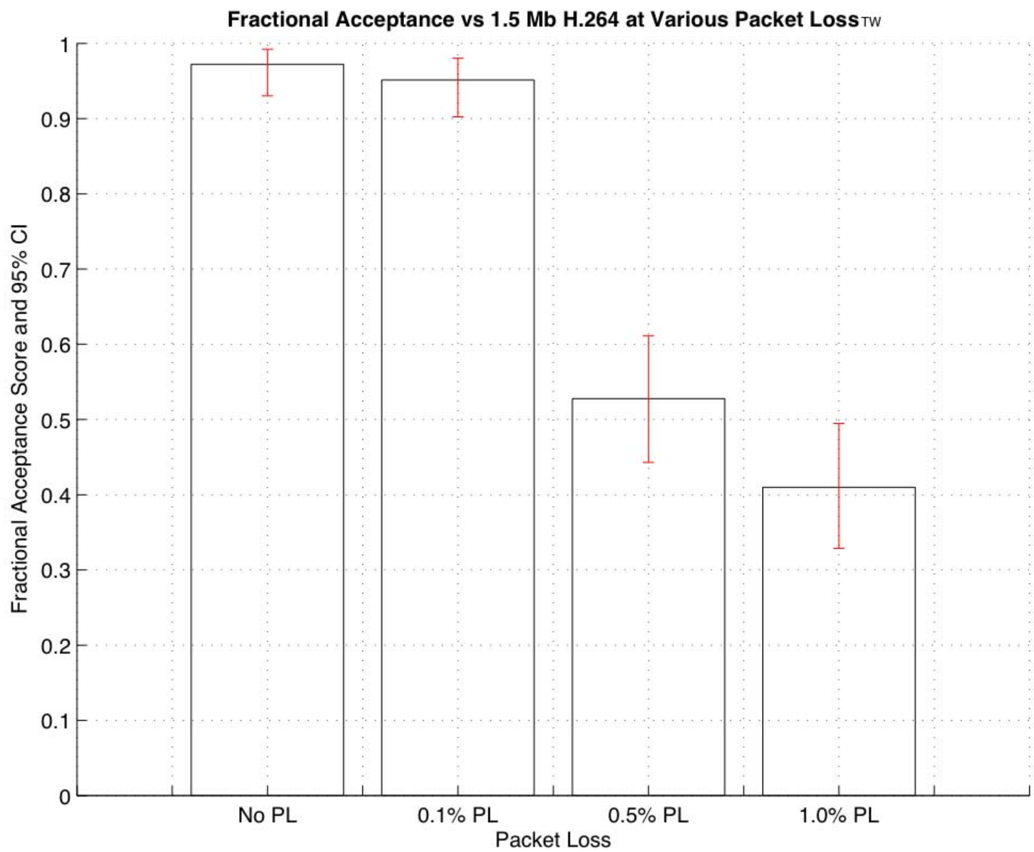


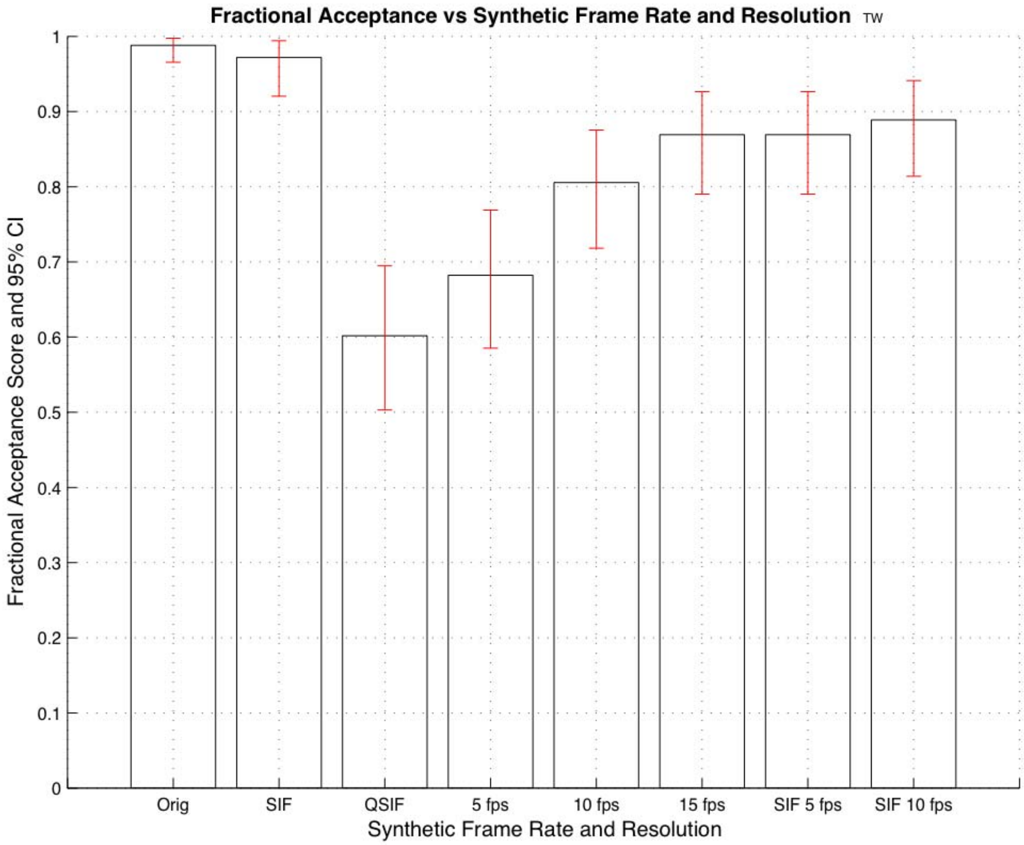
Figure 20: Fractional Acceptance PS2TW H.264 Hardware HRCs with Packet Loss



5.5.2.2 PS2TW Synthetic HRCs

Figure 21 gives the results for the synthetic HRCs (CI in this figure = confidence interval). These results can be used to establish values for fundamental image quality parameters such as frame rate and resolution.

Figure 21: PS2TW Fractional Acceptance vs. Synthetic HRCs for Frame Rate and Resolution



5.5.3 PS2TW Conclusions

The results in this report suggest that SIF image resolution should be the minimum specification for wide field of view tactical video applications, particularly if any coding or transmission loss is present. The minimum frame rate should be at least 10 fps. The choice of codec is dependant upon available network capacities.

H.264 [3] systems operating at 384 kbps were acceptable for a packet loss ratio of 0 percent (i.e., no packet loss).

Preliminary results from this experiment indicate that acceptable packet loss ratios for H.264 codecs that do not perform error concealment are on the order of 0.1 percent.

5.6 PS2ON Observed Surveillance Narrow Field of View Test

PS2ON focused on determining the minimum acceptable video transmission quality that is required to support observed surveillance narrow field of view video applications. Before the test, viewers underwent a training session where they were reminded that public safety personnel use surveillance video during an incident to investigate and make judgments about that incident.

5.6.1 PS2ON Original Video Sequences

Eight video sequences were selected, each 8 seconds in duration. These video sequences were split into two sets of eight sequences. Scene set A was matched with simulated impairments. Scene set B was matched with several H.264 [3] codec impairments. These HRC sets are described in Table 12 and Table 13.

The video sequences within each set are described below. All video sequences were converted from their original format to Rec. 601 [5] prior to use in the experiment.

Scene Set A

1. Stabbing, under tree with bright background, coming closer
2. Follow officer to car, look in window, camera bounce
3. Control room, person in charge talking on radio
4. Red-and-blue-hatted men talk, the red-hatted man looks concerned
5. Firefighters practice extinguishing fire, smoke
6. Smoke-filled hall, civilians escape building
7. Two officers talking with civilians, zoom out
8. Underwater camera observing boat and occupants

Scene Set B

1. EMS technicians load ambulance, night, and flashing lights
2. Firefighters standing, coordinating, comparing notes
3. One firefighter directing civilian car
4. Firefighters and hose, with fire in background
5. Night time stop of driver by police, blue flashing lights
6. Firefighters practice use of flame-protective bags
7. Pro-war rally
8. Stationary officer, watches crowd walk past

5.6.2 PS2ON Data Analysis

The data presented here have been aggregated over all viewers and all scenes, so that each data point represents the performance of one HRC. Thus, the fraction acceptability values take into account variations among viewers' opinions, and variations due to changing scene content.

The results are grouped into two different sets, presented in the pages that follow:

- H.264 hardware HRCs
- Synthetic HRCs

5.6.2.1 PS2ON H.264 Hardware HRCs

Figure 22, Figure 23, and Figure 24 give the results for the H.264 [3] hardware HRCs. (CI in these figures = confidence interval.)

Figure 22: Fractional Acceptance PS2ON H.264 Hardware HRCs vs. Throughput Rate for NTSC

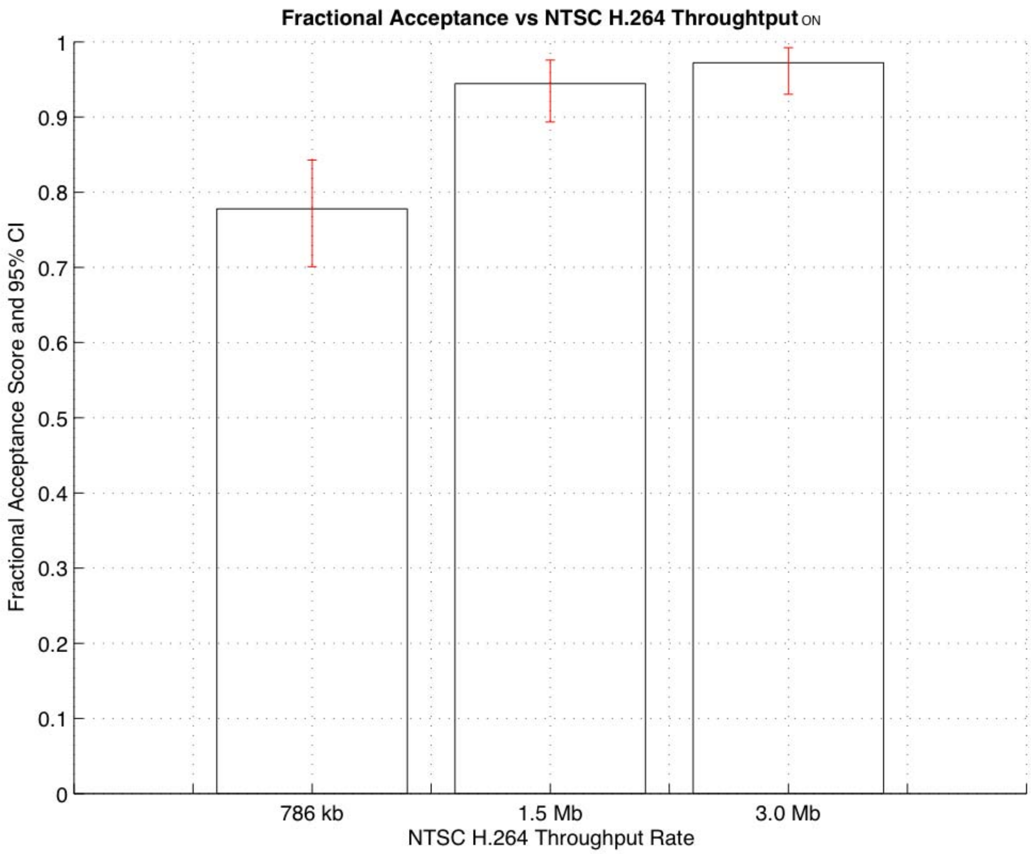


Figure 23: Fractional Acceptance PS2ON H.264 Hardware HRCs vs. Throughput Rate for CIF

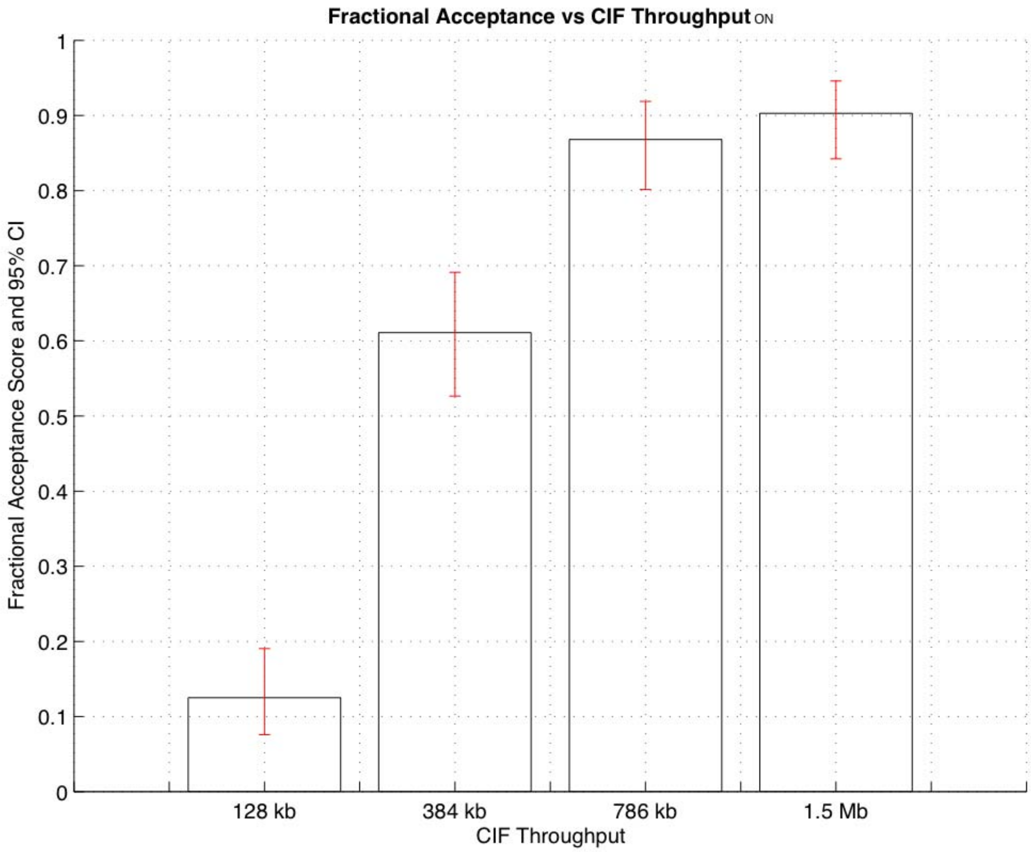
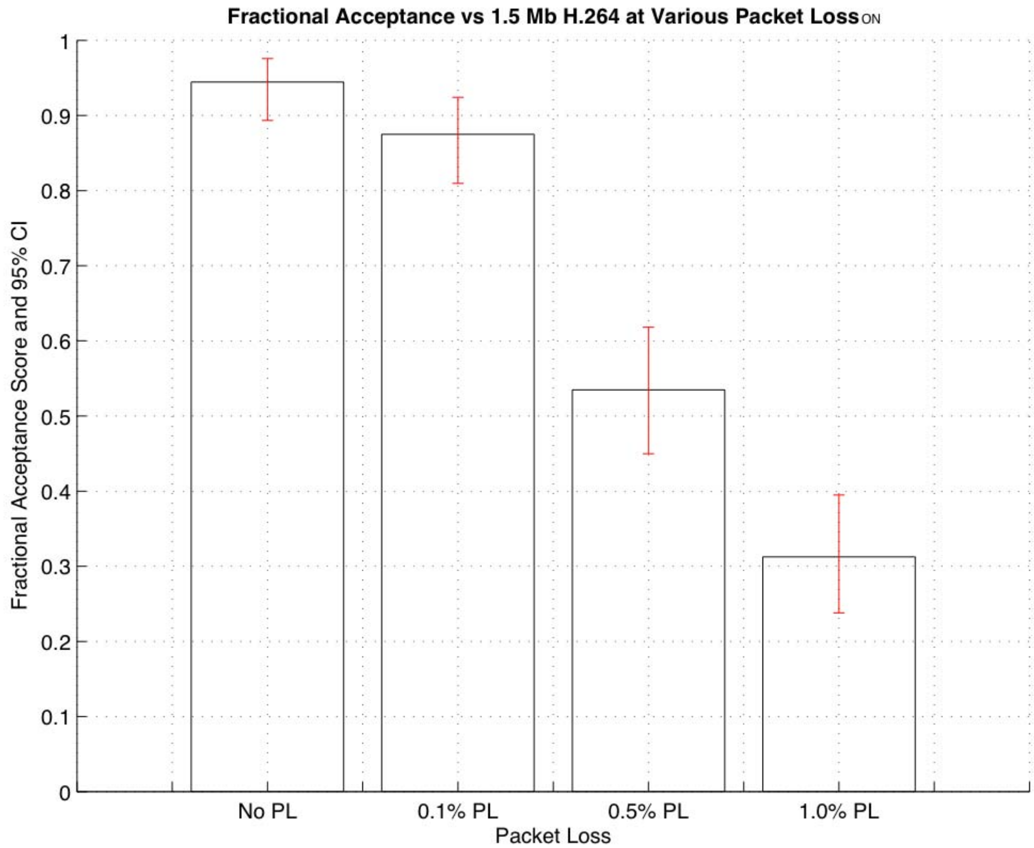


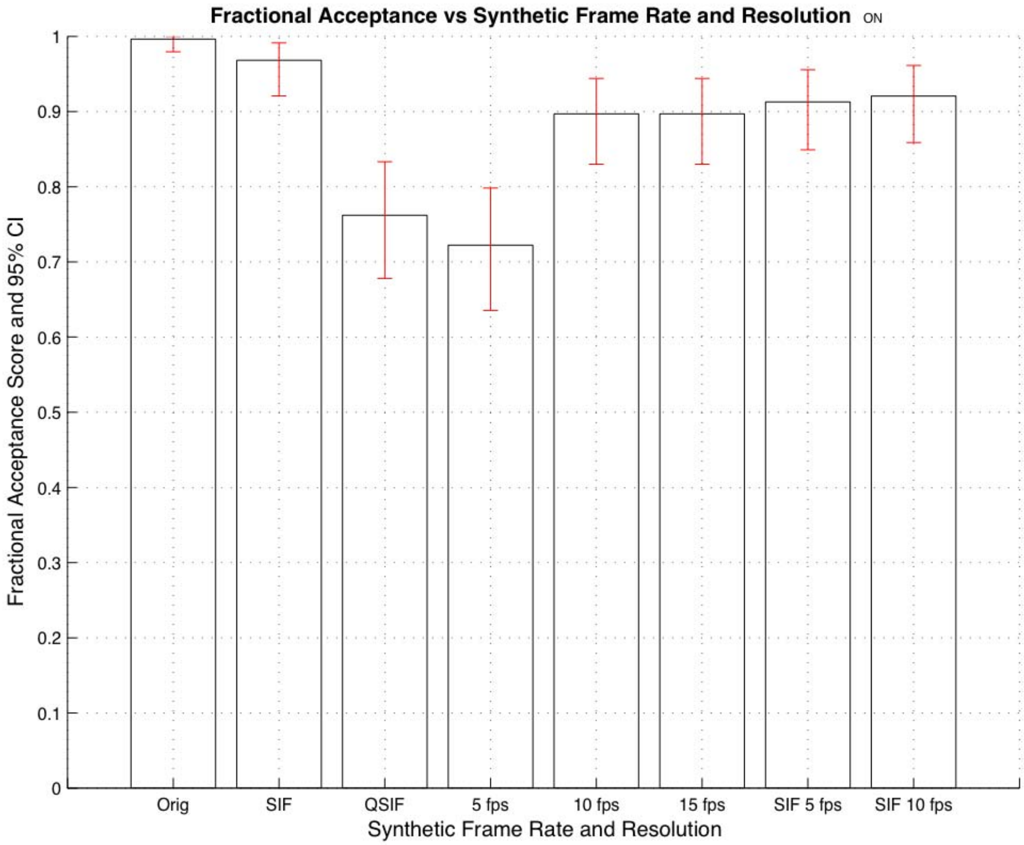
Figure 24: Fractional Acceptance PS2ON H.264 Hardware HRCs with Packet Loss



5.6.2.2 PS2ON Synthetic HRCs

Figure 25 gives the results for the synthetic HRCs (CI in this figure = confidence interval). These results can be used to establish values for fundamental image quality parameters such as frame rate and resolution.

Figure 25: PS2TW Fractional Acceptance vs. Synthetic HRCs for Frame Rate and Resolution



5.6.3 PS2ON Conclusions

The results in this report suggest that SIF image resolution should be the minimum specification for narrow field of view observed surveillance video applications, particularly if any coding or transmission loss is present. The minimum frame rate should be at least 10 fps. The choice of codec is dependant upon available network capacities.

H.264 [3] systems operating at 786 kbps were acceptable for a packet loss ratio of 0 percent (i.e., no packet loss).

Preliminary results from this experiment indicate that acceptable packet loss ratios for H.264 codecs that do not perform error concealment are on the order of 0.1 percent.

5.7 PS2OW Observed Surveillance Wide Field of View Test

PS2OW focused on determining the minimum acceptable video transmission quality that is required to support observed surveillance wide field of view video applications. Before the test, viewers underwent a training session where they were reminded that public safety personnel use surveillance video during an incident to investigate and make judgments about that incident.

5.7.1 PS2OW Original Video Sequences

Eight video sequences were selected, each 8 seconds in duration. These video sequences were split into two sets of eight sequences. Scene set A was matched with simulated impairments. Scene set B was matched with several H.264 [3] codec impairments. These HRC sets are described in Table 12 and Table 13.

The video sequences within each set are described below. The original video format is listed in parenthesis. All video sequences were converted from their original format to Rec. 601 [5] prior to use in the experiment.

Scene Set A

1. A person being choked, seen from doorway cam, action very small
2. Zoom out, inside crowded stadium
3. Truck drives around with lights on, from inside
4. Pan left of many men working
5. EMS technicians attend heart attack victim, overhead and close
6. Side view, line of civilians frisked as they enter gates
7. Pan of mock airplane disaster, victims on blankets
8. Riot police at protest, sweep of incident scene

Scene Set B

1. Police confront man and woman, camera flare from sun
2. Civilians enter ticket gate
3. Observe crowd walking in multiple directions
4. Firefighter rides all-terrain vehicle with lights flashing into and away from camera
5. Fish-eye lens, people walk down hallway intersection
6. Night-time pan of airport and outbuildings
7. Driver of yellow truck spins wheel, lots of shadows
8. Different vans and trucks cross diagonally

5.7.2 PS2OW Data Analysis

The data presented here have been aggregated over all viewers and all scenes, so that each data point represents the performance of one HRC. Thus, the fraction acceptability values take into account variations among viewers' opinions, and variations due to changing scene content.

The results are grouped into two different sets, presented in the pages that follow:

- H.264 hardware HRCs
- Synthetic HRCs

5.7.2.1 PS2OW H.264 Hardware HRCs

Figure 26, Figure 27, and Figure 28 give the results for the H.264 [3] hardware HRCs. (CI in these figures = confidence interval.)

Figure 26: Fractional Acceptance PS2OW H.264 Hardware HRCs vs. Throughput Rate for NTSC

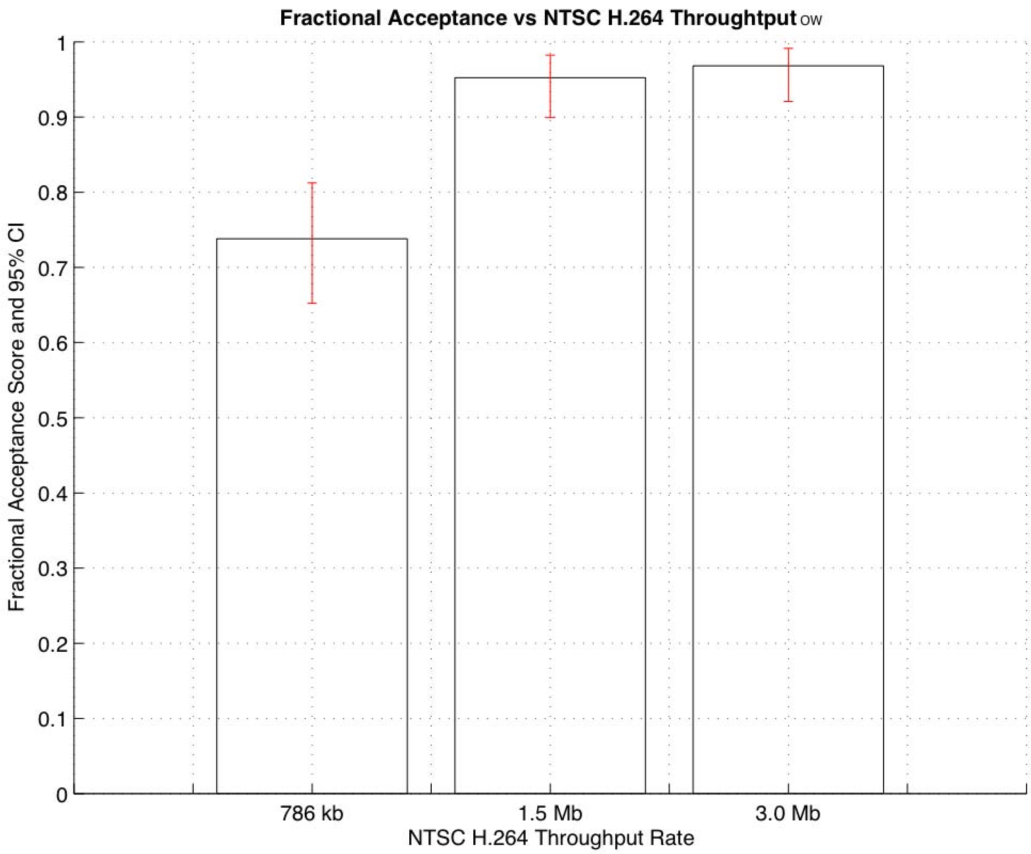


Figure 27: Fractional Acceptance PS2OW H.264 Hardware HRCs vs. Throughput Rate for CIF

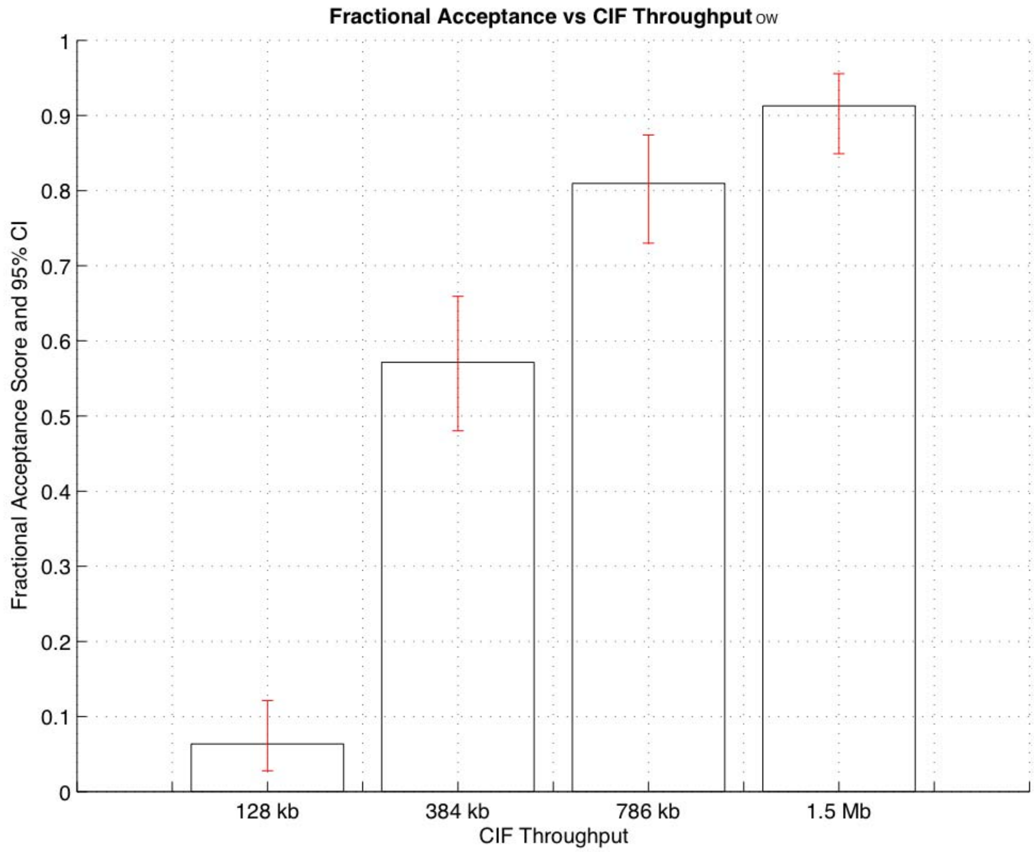
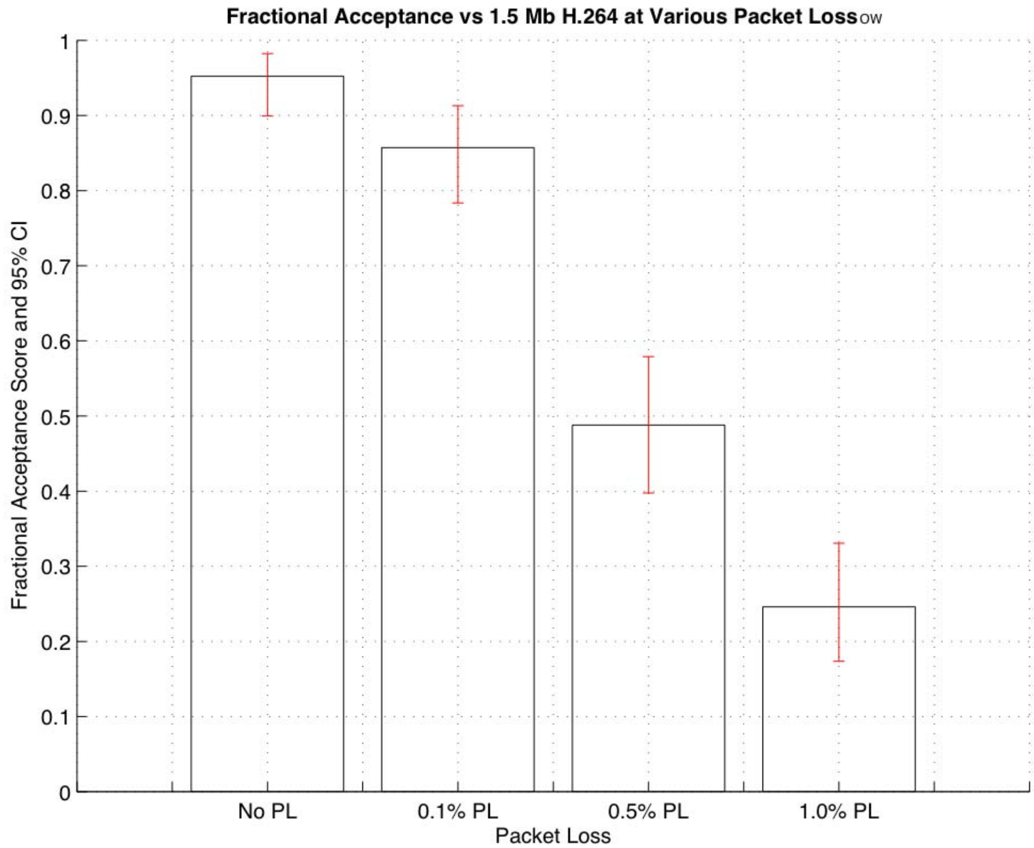


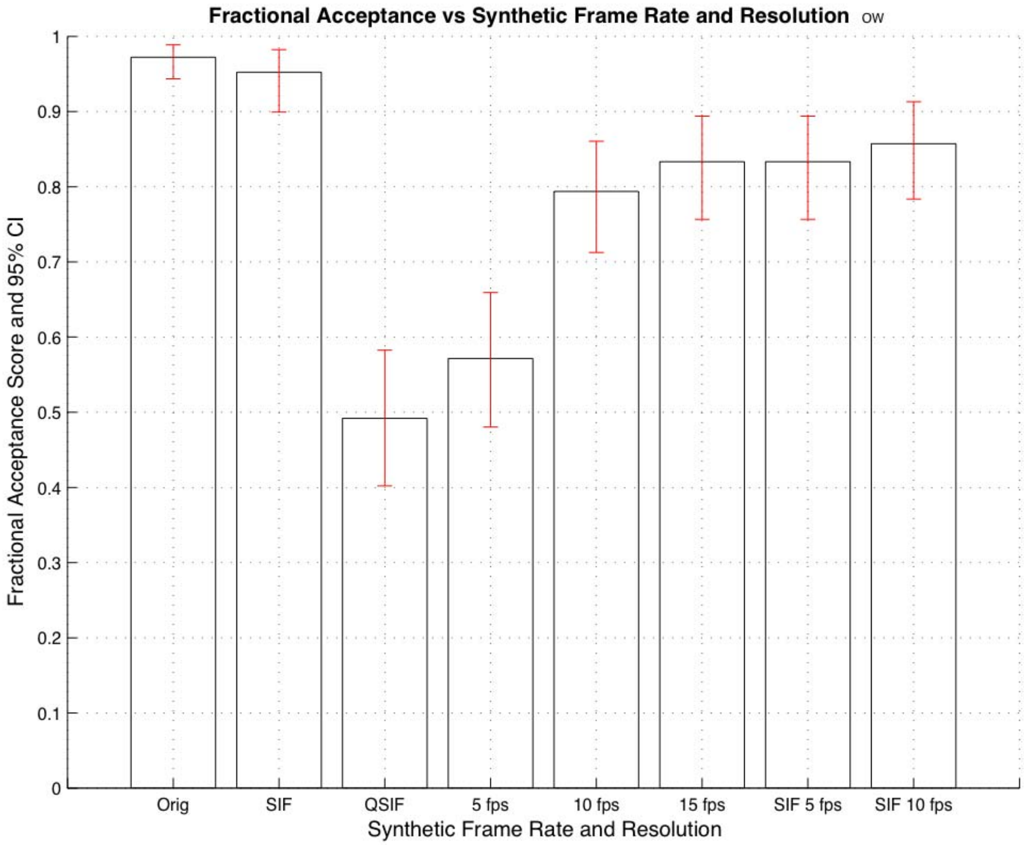
Figure 28: Fractional Acceptance PS2OW H.264 Hardware HRCs with Packet Loss



5.7.2.2 PS2OW Synthetic HRCs

Figure 29 gives the results for the synthetic HRCs (CI in this figure = confidence interval). These results can be used to establish values for fundamental image quality parameters such as frame rate and resolution.

Figure 29: PS2OW Fractional Acceptance vs. Synthetic HRCs for Frame Rate and Resolution



5.7.3 PS2OW Conclusions

The results in this report suggest that SIF image resolution should be the minimum specification for wide field of view observed surveillance video applications, particularly if any coding or transmission loss is present. The minimum frame rate should be at least 10 fps. The choice of codec is dependant upon available network capacities.

H.264 [3] systems operating at 786 kbps were acceptable for a packet loss ratio of 0 percent (i.e., no packet loss).

Preliminary results from this experiment indicate that acceptable packet loss ratios for H.264 codecs that do not perform error concealment are on the order of 0.1 percent.

5.8 PS2RN Recorded Surveillance Narrow Field of View Test

PS2RN focused on determining the minimum acceptable video transmission quality that is required to support recorded surveillance narrow field of view video applications. Before the test, viewers underwent a training session where they were reminded that public safety personnel use surveillance video following an incident to investigate and make judgments about that incident.

5.8.1 PS2RN Original Video Sequences

Five video sequences were selected, each 8 seconds in duration. These video sequences were split into two sets of five sequences. Scene set A was matched with simulated impairments. Scene set B was matched with several H.264 [3] codec impairments. These HRC sets are described in Table 14 and Table 15.

The video sequences within each set are described below. The original video format is listed in parenthesis. All video sequences were converted from their original format to Rec. 601 [5] prior to use in the experiment.

Scene Set A

1. Stabbing, under tree with bright background, coming toward camera, laugh at end
2. Police put prisoner into back seat
3. Drunk test, close up of woman's eyes
4. Night-time stop, blue flashing lights
5. Underwater crime scene search, officer pulls a gun from some wreckage

Scene Set B

1. Three EMS technicians strap victim to gurney, automatic iris lens of camera goes dark
2. Police escort man and woman out gate, walk
3. Firefighters spray water, fire extinguished
4. Officer frisks prisoner
5. Zoom in on jeep, follow

5.8.2 PS2RN Data Analysis

The data presented here have been aggregated over all viewers and all scenes, so that each data point represents the performance of one HRC. Thus, the fraction acceptability values take into account variations among viewers' opinions, and variations due to changing scene content.

The results are grouped into two different sets, presented in the pages that follow:

- H.264 hardware HRCs
- Synthetic HRCs

5.8.2.1 PS2RN H.264 Hardware HRCs

Figure 30 and Figure 31 give the results for the H.264 hardware HRCs.

Figure 30: Fractional Acceptance PS2RN H.264 Hardware HRCs vs. Throughput Rate for NTSC

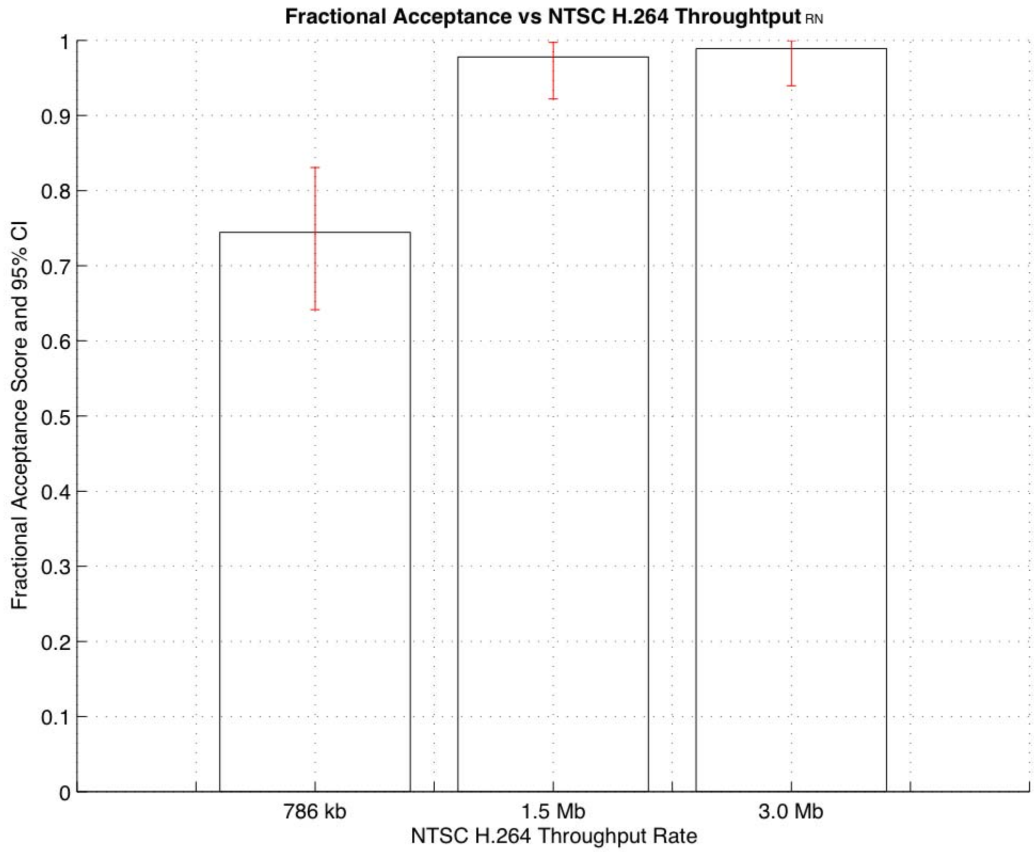
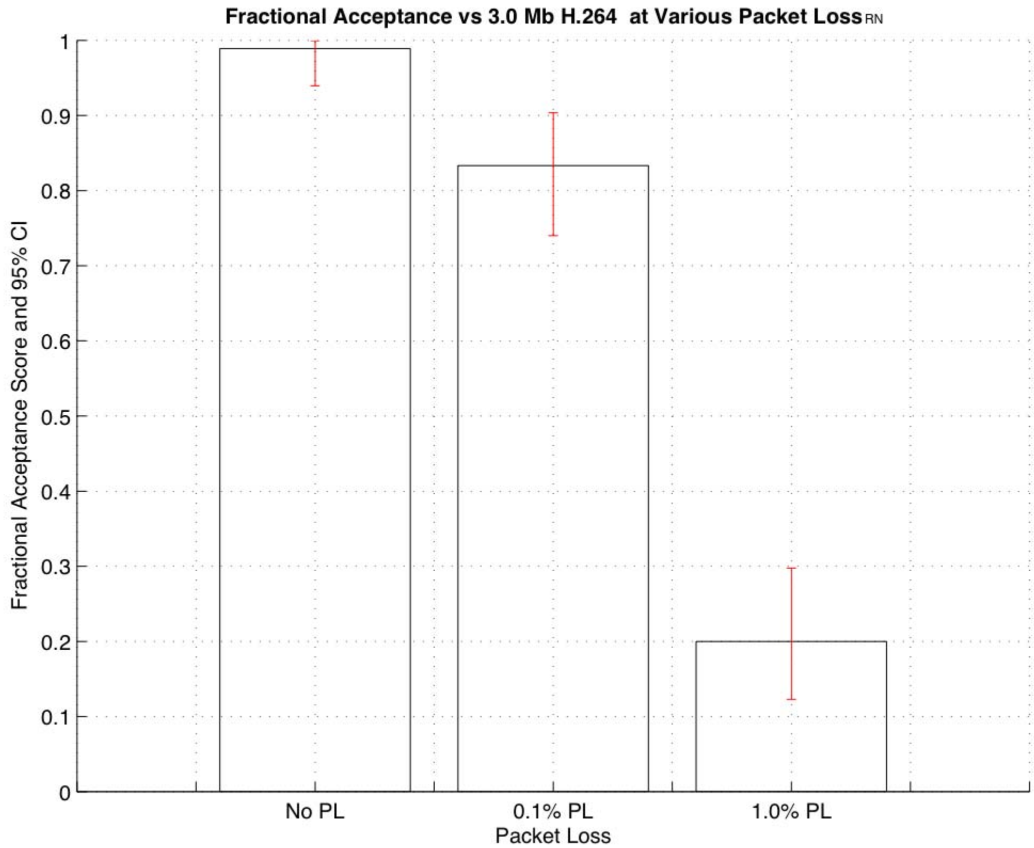


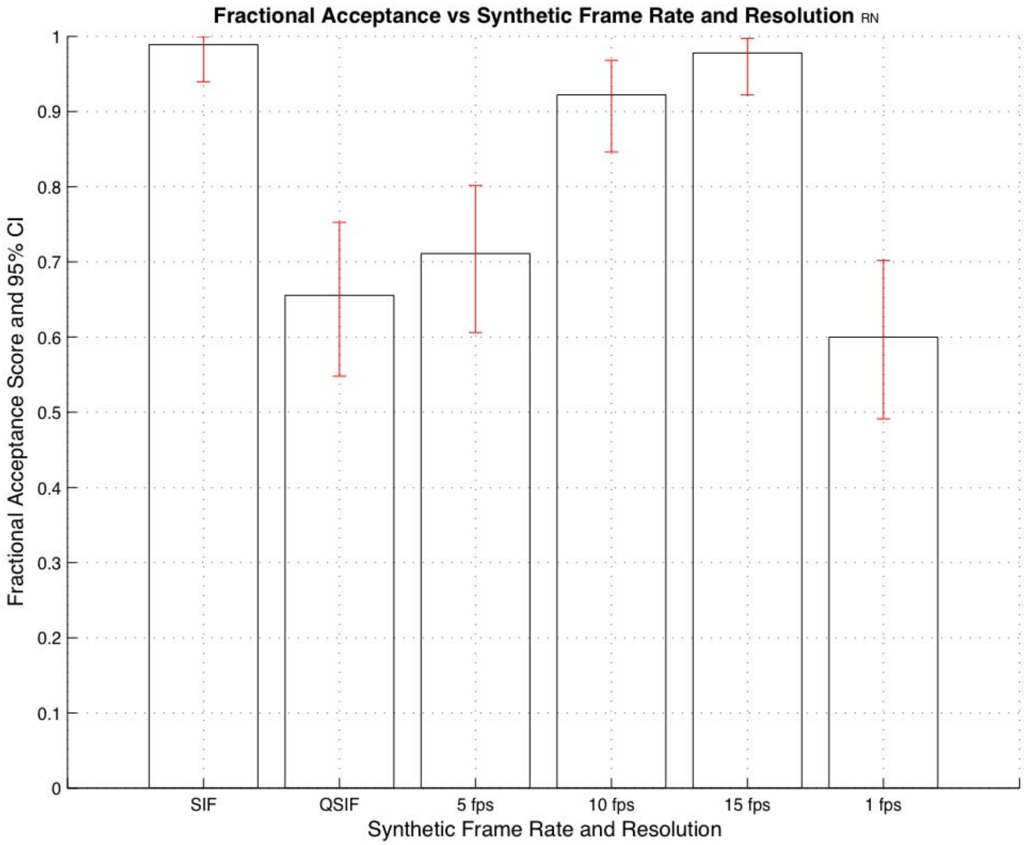
Figure 31: Fractional Acceptance PS2RN H.264 Hardware HRCs with Packet Loss



5.8.2.2 PS2RN Synthetic HRCs

Figure 32 gives the results for the synthetic HRCs. These results can be used to establish values for fundamental image quality parameters such as frame rate and resolution.

Figure 32: PS2RN Fractional Acceptance vs. Synthetic HRCs for Frame Rate and Resolution



5.8.3 PS2RN Conclusions

The results in this report suggest that SIF image resolution should be the minimum specification for wide field of view observed surveillance video applications, particularly if any coding or transmission loss is present. The minimum frame rate should be at least 10 fps. The choice of codec is dependant upon available network capacities.

H.264 [3] systems operating at 1.5 Mbps were acceptable for a packet loss ratio of 0 percent (i.e., no packet loss).

Preliminary results from this experiment indicate that acceptable packet loss ratios for H.264 codecs that do not perform error concealment are on the order of 0.1 percent, but there was such a precipitous change between the acceptability at 0.1 percent and 1 percent packet loss, that the true number lies somewhere between, and will be the subject of further study.

5.9 PS2RW Recorded Surveillance Wide Field of View Test

PS2RW focused on determining the minimum acceptable video transmission quality that is required to support recorded surveillance wide field of view video applications. Before the test, viewers underwent a training session where they were reminded that public safety personnel use surveillance video following an incident to investigate and make judgments about that incident.

5.9.1 PS2RN Original Video Sequences

Five video sequences were selected, each 8 seconds in duration. These video sequences were split into two sets of five sequences. Scene set A was matched with simulated impairments. Scene set B was matched with several H.264 [3] codec impairments. These HRC sets are described in Table 14 and Table 15.

The video sequences within each set are described below. All video sequences were converted from their original format to Rec. 601 [5] prior to use in the experiment.

Scene Set A

1. House collapse, firefighters dig out vehicle
2. Night crime scene, many officers and cars
3. Officers escort football team through crowd
4. Night flying, spotlight on highway, building, and officers
5. Wildfire and truck fire, wide angle of many firefighters

Scene Set B

1. One man attacks another with a stick, moderate distance then far, movement towards a building
2. Real fire, night, assist people out of burning building
3. EMS technicians attend heart attack victim, overhead and close
4. Side view, line of civilians frisked as enter gates
5. Wildfire, huge, movement of firefighter and truck

5.9.2 PS2RW Data Analysis

The data presented here have been aggregated over all viewers and all scenes, so that each data point represents the performance of one HRC. Thus, the fraction acceptability values take into account variations among viewers' opinions, and variations due to changing scene content.

The results are grouped into two different sets, presented in the pages that follow:

- H.264 hardware HRCs
- Synthetic HRCs

5.9.2.1 PS2RW H.264 Hardware HRCs

Figure 33 and Figure 34 give the results for the H.264 hardware HRCs. This H.264 hardware codec did implement error concealment (i.e., EC), so higher levels of packet loss ratio were considered.

Figure 33: Fractional Acceptance PS2RW H.264 Hardware HRCs vs. Throughput Rate for NTSC

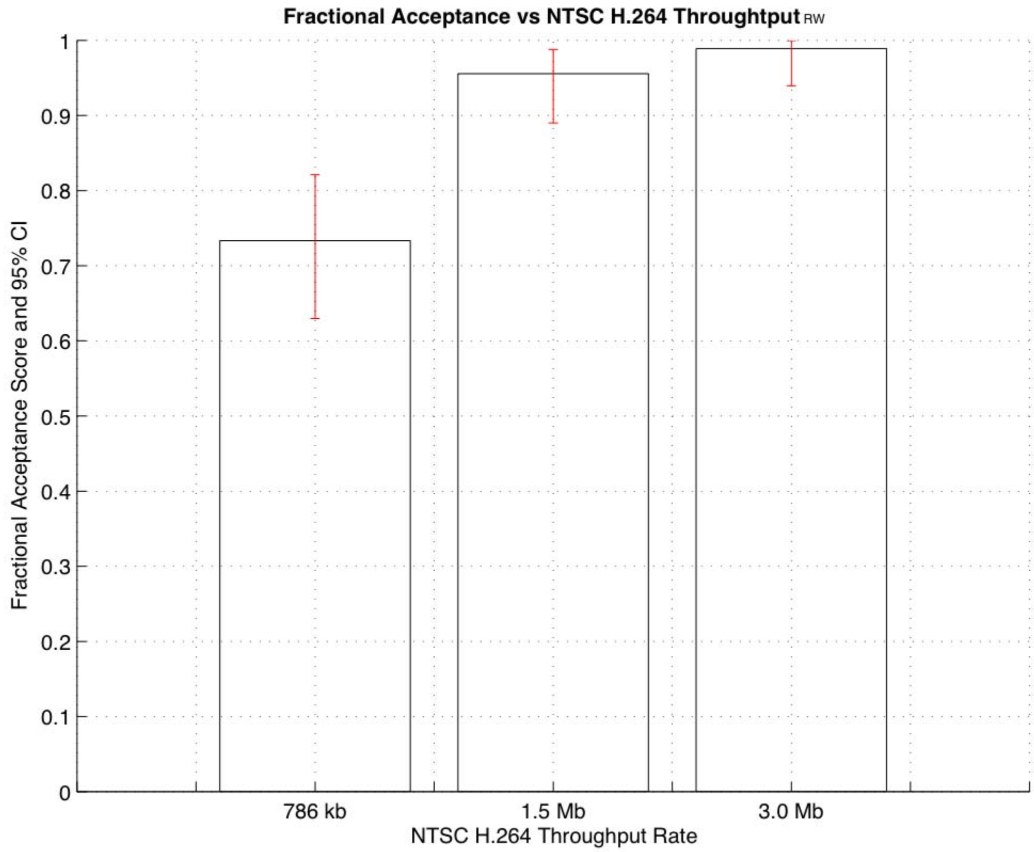
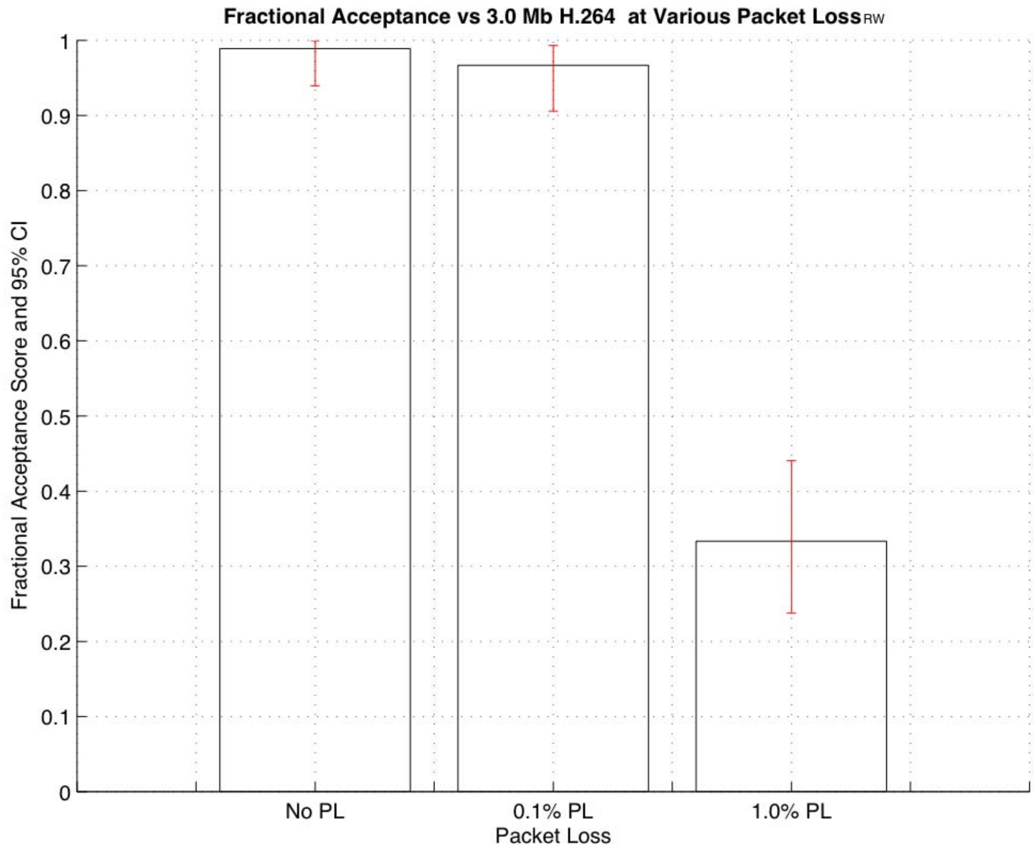


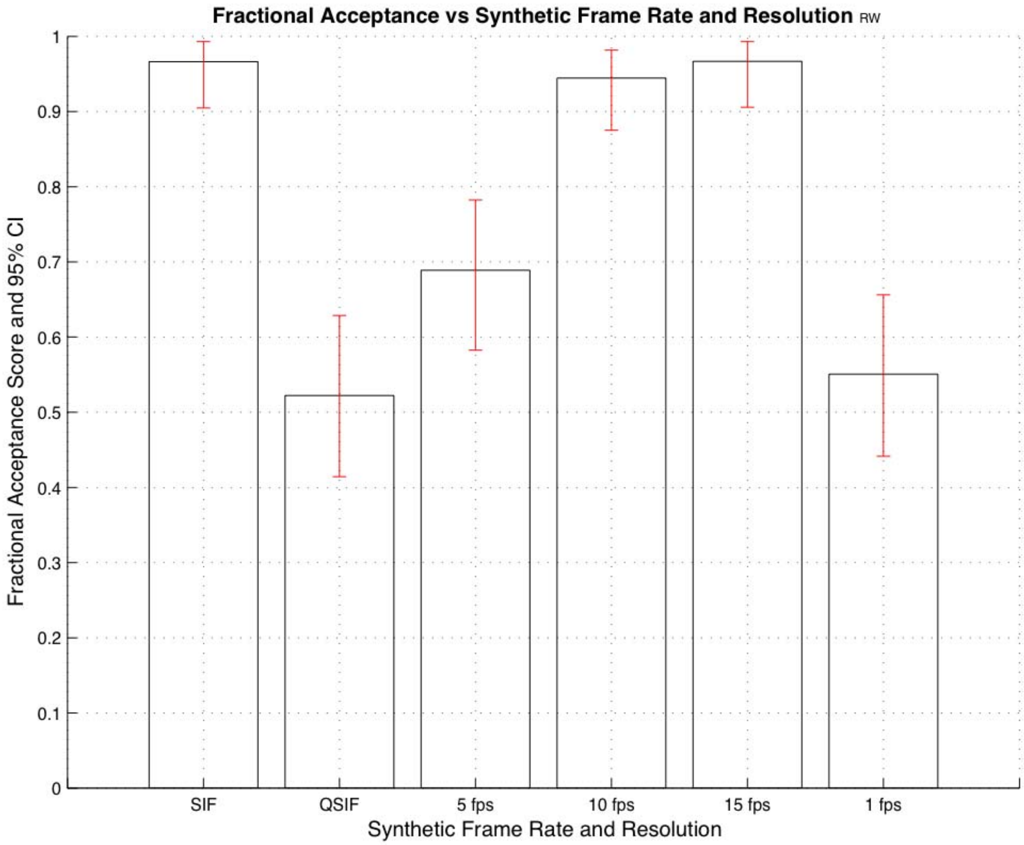
Figure 34: Fractional Acceptance PS2RW H.264 Hardware HRCs with Packet Loss



5.9.2.2 PS2RW Synthetic HRCs

Figure 35 gives the results for the synthetic HRCs. These results can be used to establish values for fundamental image quality parameters such as frame rate and resolution.

Figure 35: PS2RW Fractional Acceptance vs. Synthetic HRCs for Frame Rate and Resolution



5.9.3 PS2RW Conclusions

The results in this report suggest that SIF image resolution should be the minimum specification for wide field of view observed surveillance video applications, particularly if any coding or transmission loss is present. The minimum frame rate should be at least 10 fps. The choice of codec is dependant upon available network capacities.

H.264 [3] systems operating at 1.5 Mbps were acceptable for a packet loss ratio of 0 percent (i.e., no packet loss).

Preliminary results from this experiment indicate that acceptable packet loss ratios for H.264 codecs that do not perform error concealment are on the order of 0.1 percent, but there was such a precipitous change between the acceptability at 0.1 percent and 1 percent packet loss, that the true number lies somewhere between, and will be the subject of further study.

5.10 PS2 Conclusions

In studying the results of Group 1 and Group 2, there were very similar conclusions for each test within each group. However, the two groups differed significantly enough in bit rate requirements that it was determined they should remain as separate studies. Within each group it was decided that the tests could be combined, since requirements were so similar within the group.

Therefore, recorded surveillance video was separated from the rest of the testing, and the design of PS3 was developed by combining the TW, ON, and OW tests together.

6 PS3 Subjective Video Quality Experiment

The subjective video quality experiment PS3 was conducted from February to March of 2007. PS3 focused on determining the minimum acceptable video transmission quality that is required to support tactical and live surveillance video applications.

Before the test, viewers underwent a training session where they were reminded that public safety personnel use tactical video in real time during an incident to make decisions on how to respond to that incident.

During the subjective test, subjects watched a number of short video sequences on an LCD monitor, and judged each of them for quality and acceptability. Following each sequence, a computerized form displayed choices for rating the video using the ACR grading scale: Excellent, Good, Fair, Poor, and Bad. Subjects were also asked to judge whether the video clip was acceptable or unacceptable for use as wide field of view tactical video. The viewer controlled the pace of the session, and could take as long as he or she wanted to rate the sequence. The first rating session lasted approximately 20 minutes, followed by a break. A second 20-minute rating session followed the break.

6.1 Tactical and Live Surveillance Video Examples

Public safety personnel use tactical video in real time during an incident to make decisions on how to respond to that incident. Live surveillance video is used in real time to make decisions about whether a situation requires a response.

Examples of tactical video include:

- Video used to provide the incident commander with situation information, such as 1) a camera carried by a public safety practitioner, looking for victims, into a burning building; 2) a camera worn on an officer's body during a SWAT raid; and 3) an aerial camera following a suspect on foot
- Close-up videography from a camera on a robot being used to dismantle a bomb
- Aerial videography during wildfire suppression
- Aerial video during pursuit of an automobile

Examples of observed surveillance video include:

- Monitoring an entry way for suspicious activity
- Cameras covering a football stadium that can pan and zoom, which allow a remote police officer to look for disturbances
- Cameras in a federal judge’s residence that is under protective surveillance
- Cameras in a bank lobby that can show the location and activities of people

6.2 PS3 Experiment Design

PS3 included 48 impaired HRCs. HRCs were selected for each experiment to investigate the minimal requirements for the video attributes lists that Table 1 identifies for PS3.

Four pools of eight source video sequences were selected, each 8 seconds in duration. These video sequences were split into two sets of four sequences. The HRCs for each circuit were created as a full matrix of all combinations of three H.264 [3] bit rates and four dropped package percentages, for a total of 12 HRCs per circuit.

6.3 PS3 Original Video Sequences

Original video sequences were selected from the following sources:

- Footage shot at a football game with a shoulder-mounted camera that followed police officers as they performed their duties. Footage was shot in HDTV 720p format, and then converted to Rec. 601 [5].
- Footage shot at a firefighter training session using a shoulder-mounted camera. Footage was shot in HDTV 720p format and then converted to Rec. 601.
- In-car camera footage depicting simulated drunk driving stops. Video was shot in HDTV 1080i format, and then converted to Rec. 601.
- Footage of a SWAT team training session. Footage was received on a VHS tape and digitally sampled in accordance with Rec. 601.
- Footage of an underwater crime investigation. Footage was received on a VHS tape, and digitally sampled in accordance with Rec. 601.
- Footage purchased from video footage clearing houses
- Footage shot on HD cameras at a mountain wildfire training
- Footage shot on HD cameras of acted scenes

The video sequences within each set are described below. All video sequences were converted from their original format to Rec. 601 prior to use in the experiment.

Scene Set A

1. Police hot stop at night, flashing lights, approach prone suspect
2. Daytime police stop, driver back of car, passenger flees
3. Camera bouncing while following an officer from car to simulated traffic stop, daytime
4. DUI test, close up of face
5. Pan of crowded stadium, team
6. Helicopter view, pan parking lot, zoom in on white van
7. Security camera records start of a robbery
8. Red-and-blue-hatted men talk, red-hatted man looks concerned

Scene Set B

1. Daytime police stop, driver back of car, passenger flees
2. Wide shot of fire incident, civilians escape
3. Fireman close-up, fire inside house

4. Officer frisks prisoner near a police car
5. Pan of crowded stadium, team
6. Entrance area to football game, wide angle entering
7. Fireman climbs fire ladder
8. Mugging, situation needs response

Scene Set C

1. In car, officer receives a radio call that needs response
2. Wide shot of wildfire incident, pan and jiggle
3. Fireman close-up, dim hall with fire
4. Night time arrest scene, zoom to close-up on officers
5. Pan of crowded stadium, team
6. Simulated airplane crash, medical triage being performed
7. Remote robot operation
8. Two people fighting, then officer breaks it up

Scene Set D

1. In car, officer receives a radio call that needs response
2. Mountain rescue, descend on rope
3. Two officers escorting a person out of stadium
4. Two officers talk separately, zoom/pan
5. Pan of crowded stadium, team
6. Search rubble of house, hurricane
7. Helmet-cam, driver of car gets directions from a firefighter
8. Load person into ambulance, flashing lights

6.4 PS3 HRC Video Transmission Systems

Codecs without error concealment were paired with packet loss ratios of 0, 0.1, 0.2, and 0.4 percent. Higher packet loss levels (0.5, 10 and 20 percent) were retained for the H.264 [3] codec with error concealment. To ensure that the minimum video quality required by practitioners was presented, the subjective test had to extend significantly beyond unacceptable (at the low-quality end) and acceptable (at the high-quality end).

6.4.1 PS3 Hardware H.264 Codec

Table 16 provides a summary description of the H.264 [3] hardware HRCs (with error concealment) that were used in the PS3 experiment, including the scene set that was used for each HRC.

Table 16: PS3 H.264 Hardware HRCs with Error Concealment

HRC	Scene Set	Description
256p00ec	A	Hardware H.264, 256 kbps, error concealment, 0 percent packet loss
256p02ec	A	Hardware H.264, 256 kbps, error concealment, 0.2 percent packet loss
256p04ec	A	Hardware H.264, 256 kbps, error concealment, 0.4 percent packet loss
256p1ec	A	Hardware H.264, 256 kbps, error concealment, 1 percent packet loss
384p00ec	A, B	Hardware H.264, 384 kbps, error concealment, 0 percent packet loss
384p02ec	A	Hardware H.264, 384 kbps, error concealment, 0.2 percent packet loss
384p04ec	A	Hardware H.264, 384 kbps, error concealment, 0.4 percent packet loss
384p1ec	A	Hardware H.264, 384 kbps, error concealment, 1 percent packet loss
512p00ec	A, B	Hardware H.264, 512 kbps, error concealment, 0 percent packet loss
512p02ec	A	Hardware H.264, 512 kbps, error concealment, 0.2 percent packet loss
512p04ec	A	Hardware H.264, 512 kbps, error concealment, 0.4 percent packet loss
512p1ec	A	Hardware H.264, 512 kbps, error concealment, 1 percent packet loss
384p5ec	B	Hardware H.264, 384 kbps, error concealment, 5 percent packet loss
384p10ec	B	Hardware H.264, 384 kbps, error concealment, 10 percent packet loss
384p20ec	B	Hardware H.264, 384 kbps, error concealment, 20 percent packet loss
512p5ec	B	Hardware H.264, 512 kbps, error concealment, 5 percent packet loss
512p10ec	B	Hardware H.264, 512 kbps, error concealment, 10 percent packet loss
512p20ec	B	Hardware H.264, 512 kbps, error concealment, 20 percent packet loss
1Mp00ec	B	Hardware H.264, 1 Mbps, error concealment, 0 percent packet loss
1Mp5ec	B	Hardware H.264, 1 Mbps, error concealment, 5 percent packet loss
1Mp10ec	B	Hardware H.264, 1 Mbps, error concealment, 10 percent packet loss
1Mp20ec	B	Hardware H.264, 1 Mbps, error concealment, 20 percent packet loss
384p00	C	Hardware H.264, 384 kbps, no error concealment, 0 percent packet loss
384p01	C	Hardware H.264, 384 kbps, no error concealment, 0.1 percent packet loss
384p02	C	Hardware H.264, 384 kbps, no error concealment, 0.2 percent packet loss

Table 16: PS3 H.264 Hardware HRCs with Error Concealment (Continued)

HRC	Scene Set	Description
384p04	C	Hardware H.264, 384 kbps, no error concealment, 0.4 percent packet loss
512p00	C, D	Hardware H.264, 512 kbps, no error concealment, 0 percent packet loss
512p01	C, D	Hardware H.264, 512 kbps, no error concealment, 0.1 percent packet loss
512p02	C, D	Hardware H.264, 512 kbps, no error concealment, 0.2 percent packet loss
512p04	C, D	Hardware H.264, 512 kbps, no error concealment, 0.4 percent packet loss
768p00	C, D	Hardware H.264, 768 kbps, no error concealment, 0 percent packet loss
768p01	C, D	Hardware H.264, 768 kbps, no error concealment, 0.1 percent packet loss
768p02	C, D	Hardware H.264, 768 kbps, no error concealment, 0.2 percent packet loss
768p04	C, D	Hardware H.264, 768 kbps, no error concealment, 0.4 percent packet loss
1Mp00	D	Hardware H.264, 1 Mbps, no error concealment, 0 percent packet loss
1Mp01	D	Hardware H.264, 1 Mbps, no error concealment, 0.1 percent packet loss
1Mp02	D	Hardware H.264, 1 Mbps, no error concealment, 0.2 percent packet loss
1Mp04	D	Hardware H.264, 1 Mbps, no error concealment, 0.4 percent packet loss

6.5 PS3 Viewers

Twenty-one public safety practitioners were recruited to participate in the PS3 subjective video quality experiment. The practitioners came from across the country. Various local jurisdictions, state jurisdictions, and Federal jurisdictions were represented. Disciplines represented included fire response, law enforcement, and EMS.

Subjects were screened for color perception and visual acuity at a distance of 10 feet. Each subject participated in two viewing sessions.

6.6 PS3 Data Analysis

The data presented here have been aggregated over all viewers and all scenes, so that each data point represents the performance of one HRC. Thus, the fraction acceptability values take into account variations among viewers' opinions, and variations due to changing scene content.

The results are represented by each circuit, showing results of packet loss on each bit rate.

Figure 36: PS3 H.264 Circuit A, 256 kbps vs. Packet Losses

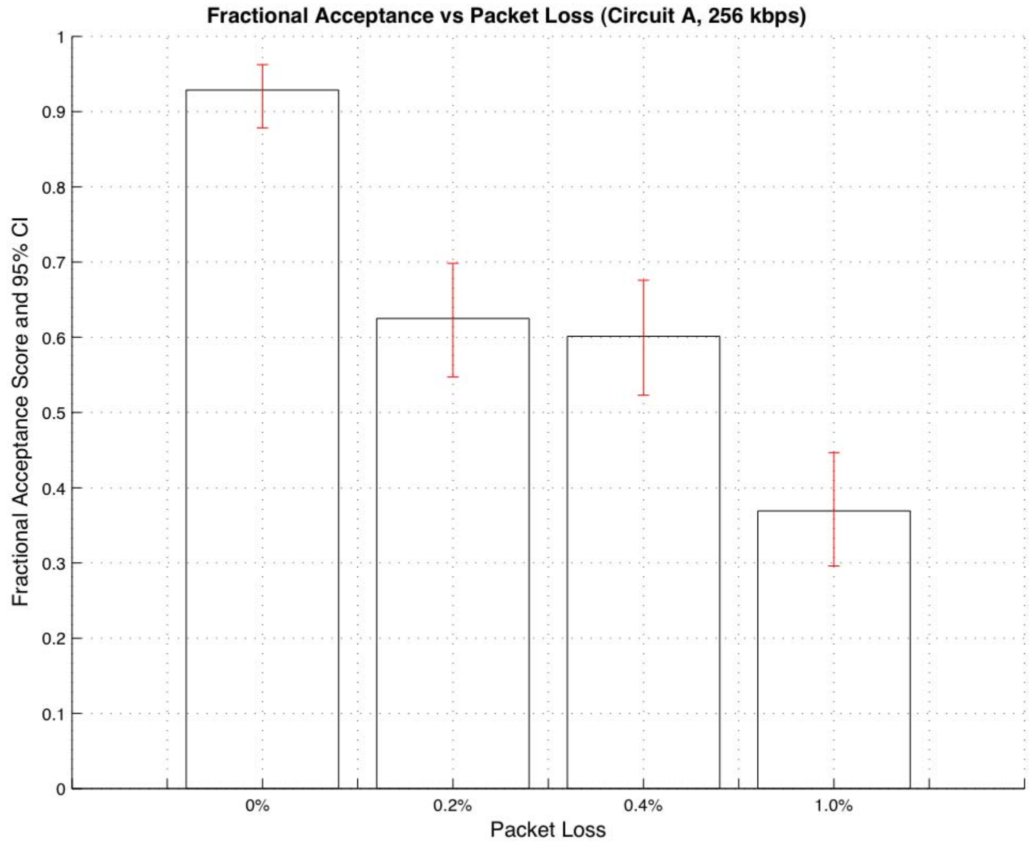


Figure 37: PS3 H.264 Circuit A, 384 kbps vs. Packet Losses

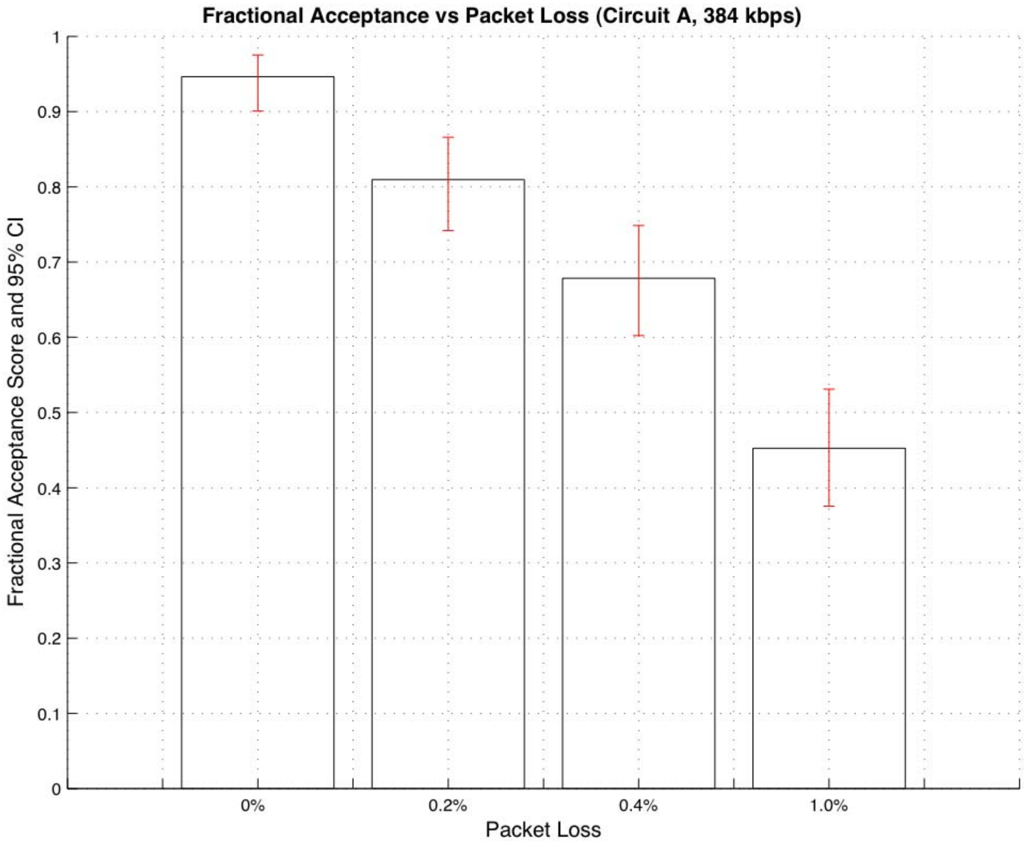


Figure 38: PS3 H.264 Circuit A, 512 kbps vs. Packet Losses

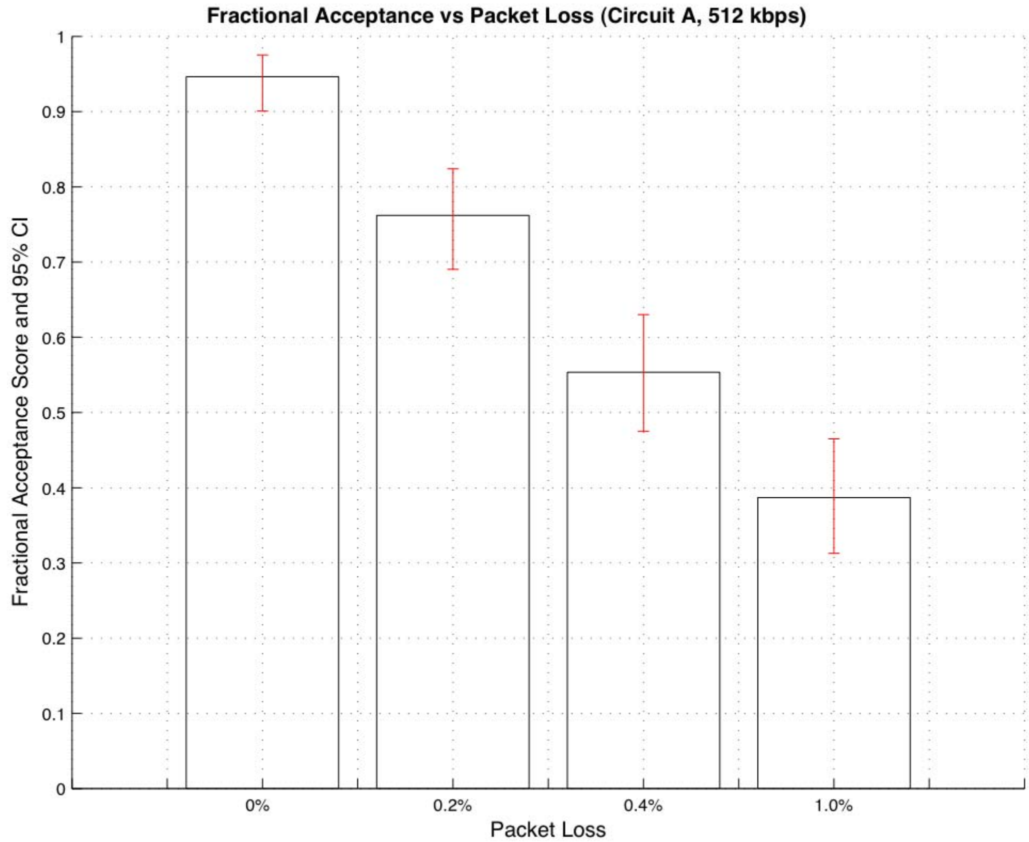


Figure 39: PS3 H.264 Circuit B, 384 kbps vs. Packet Losses

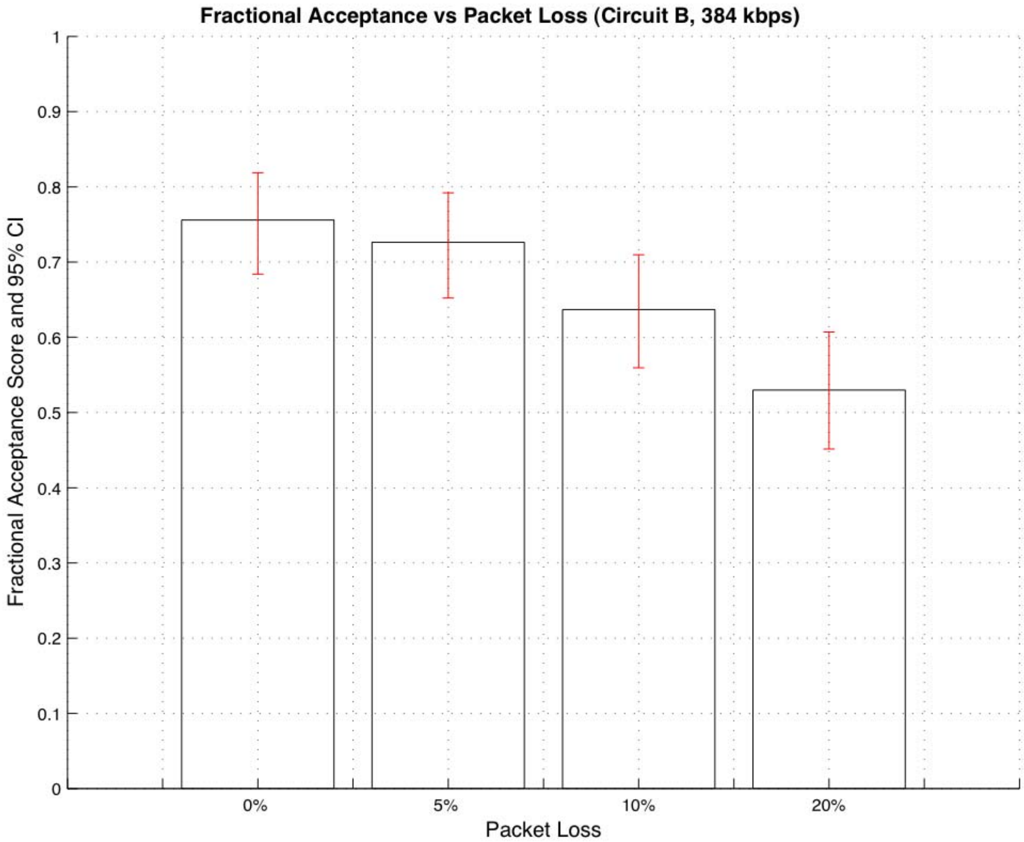


Figure 40: PS3 H.264 Circuit B, 512 kbps vs. Packet Losses

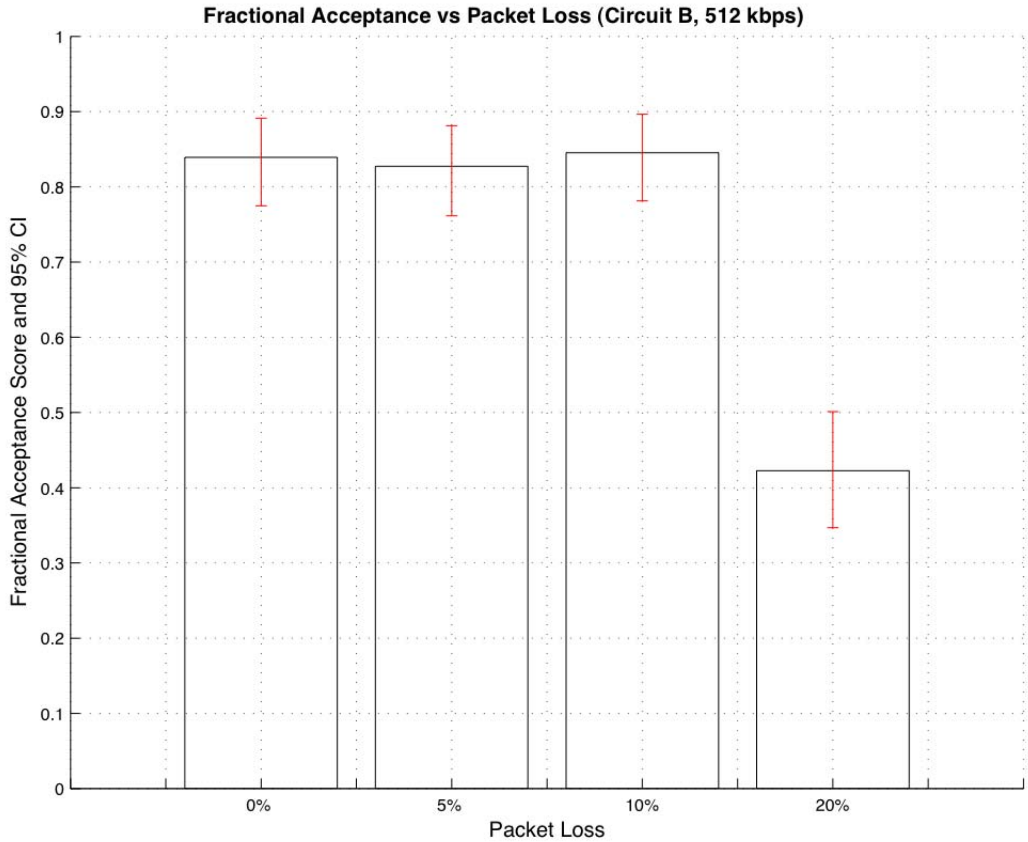


Figure 41: PS3 H.264 Circuit B, 1 Mbps vs. Packet Losses

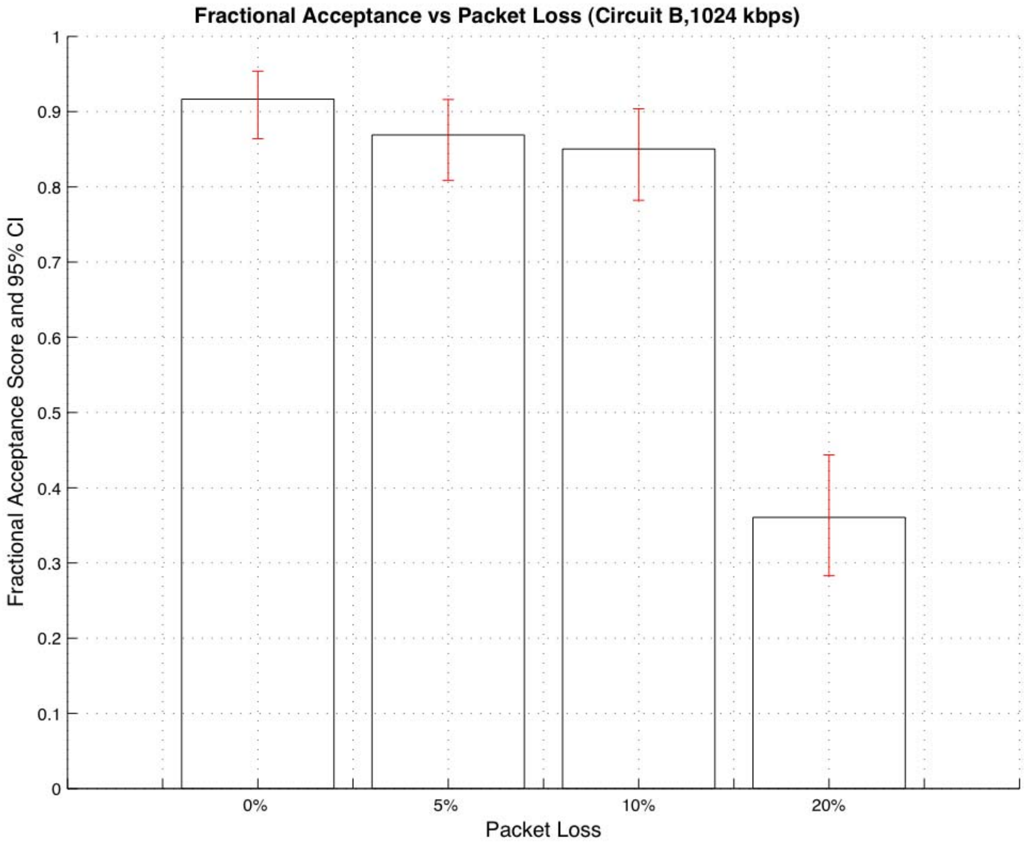


Figure 42: PS3 H.264 Circuit C, 384 kbps vs. Packet Losses

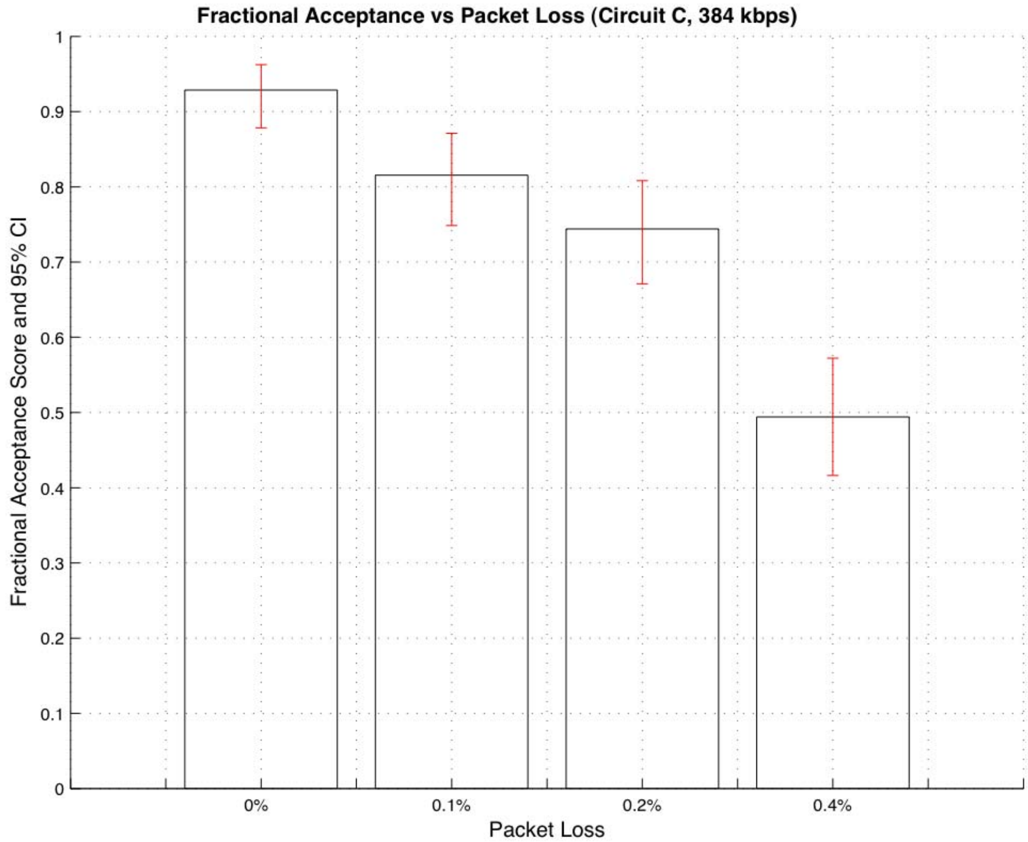


Figure 43: PS3 H.264 Circuit C, 512 kbps vs. Packet Losses

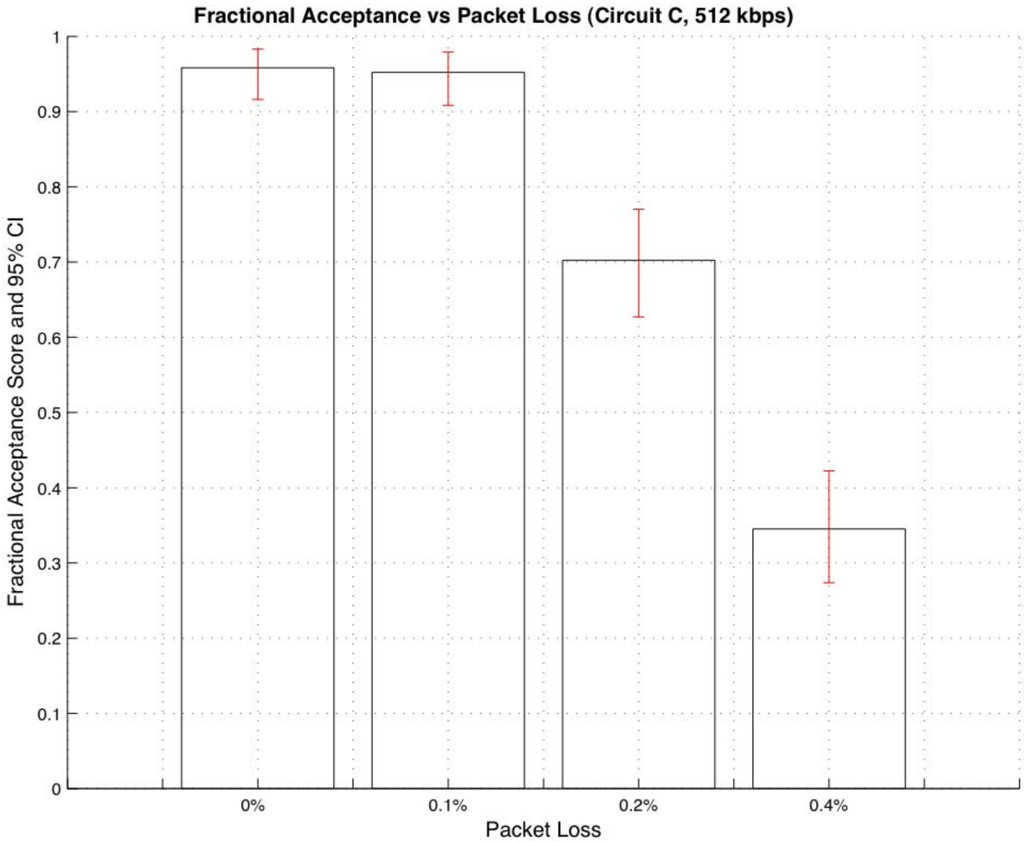


Figure 44: PS3 H.264 Circuit C, 768 kbps vs. Packet Losses

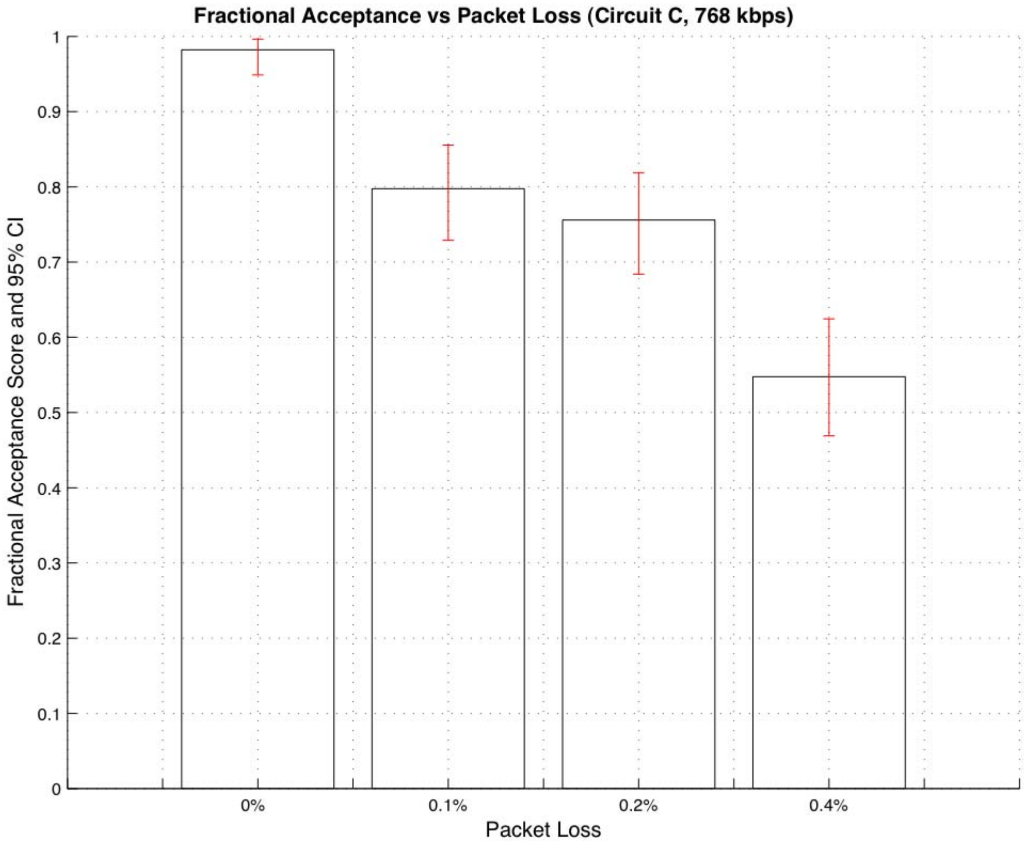


Figure 45: PS3 H.264 Circuit D, 512 kbps vs. Packet Losses

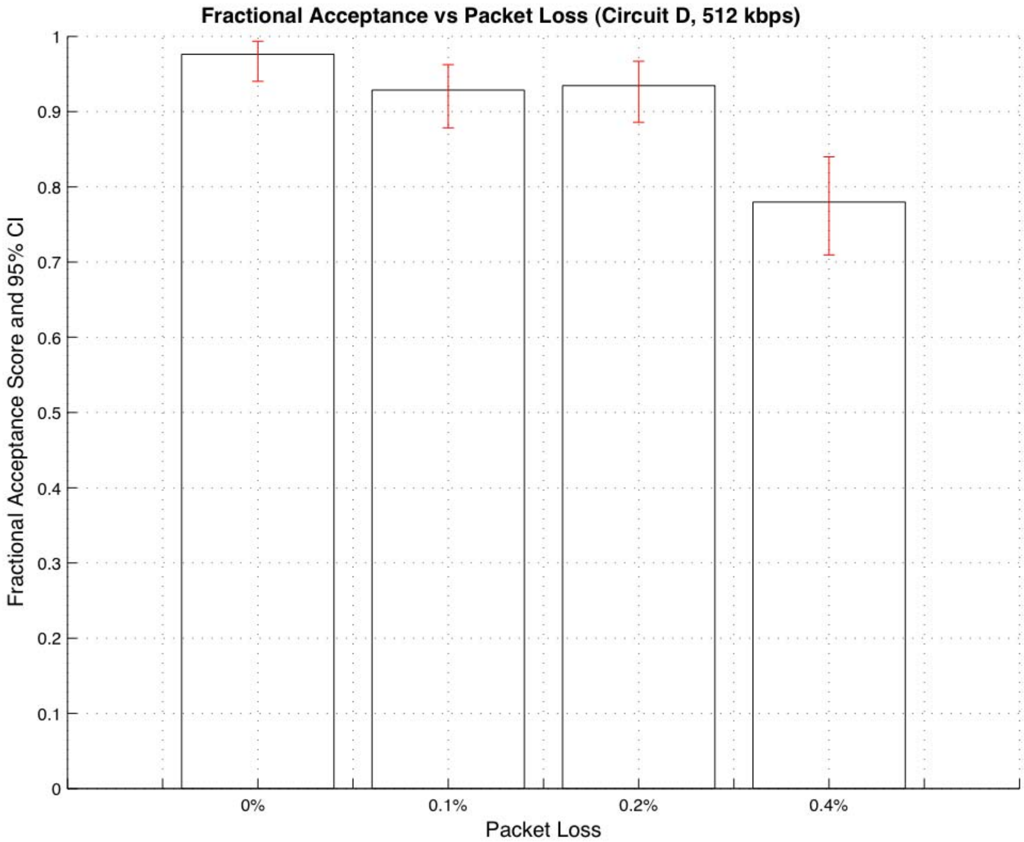


Figure 46: PS3 H.264 Circuit D, 768 kbps vs. Packet Losses

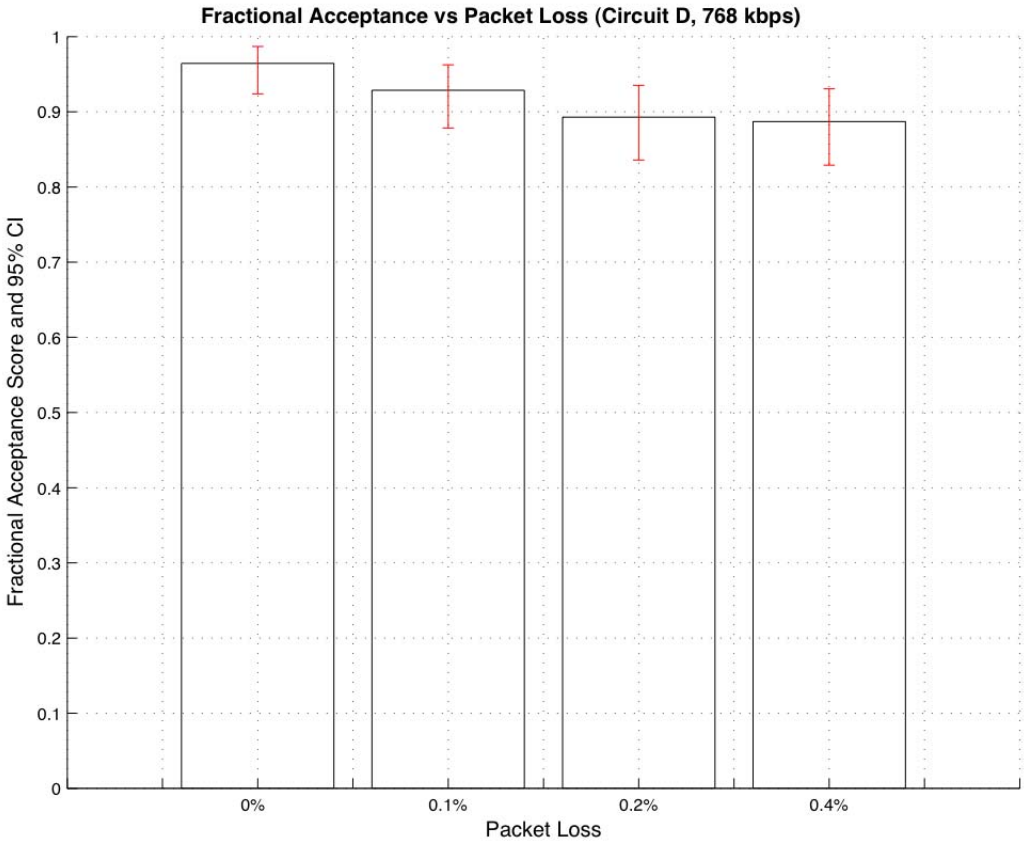
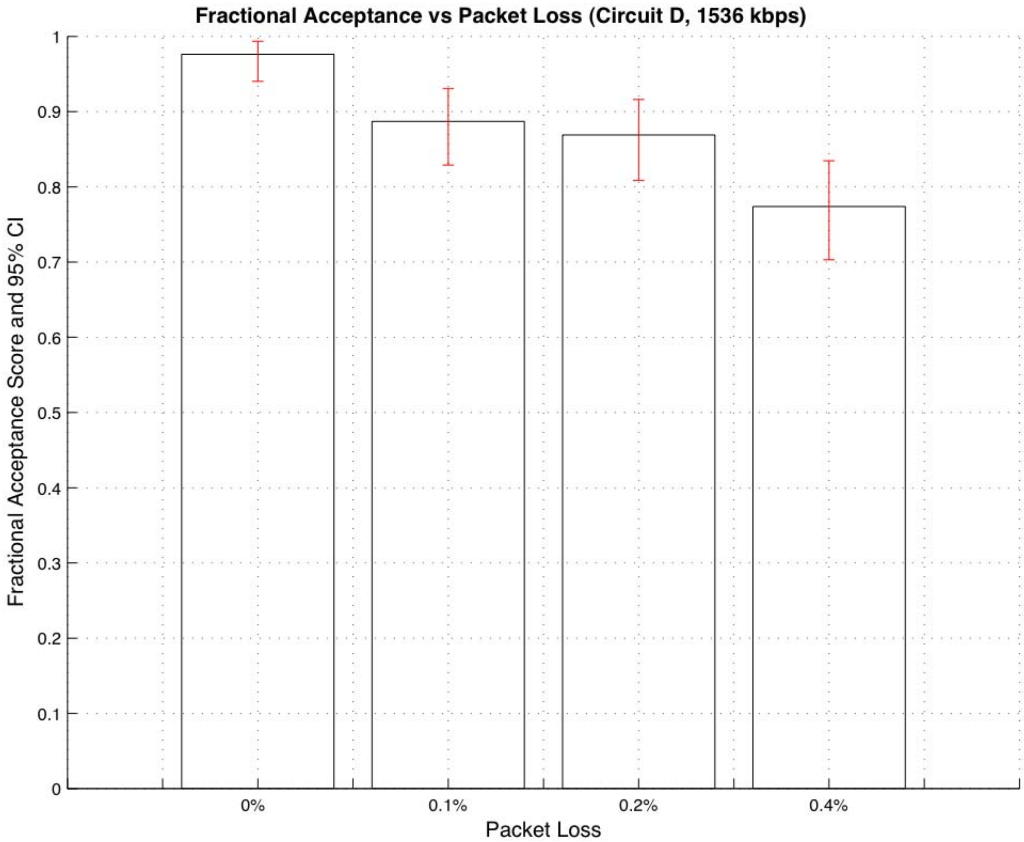


Figure 47: PS3 H.264 Circuit D, 1 Mbps vs. Packet Losses



6.7 PS3 Conclusions

H.264 [3] systems operating at 512 kbps without error concealment were acceptable for a packet loss ratio of between 0.1 and 0.4 percent. System A's error concealment method could only withstand packet losses of 0.2 percent at 384 kbps. System B's error concealment method, however, could tolerate packet losses of up to 10 percent at 512 kbps. Error concealment methods are under development and constantly being improved; the results of PS3 show that these efforts are effective and should be strongly encouraged.

7 References

[1] *Public Safety Statement of Requirements for Communications & Interoperability*, Volume II: Quantitative, Version 1.1, November 2007.

[2] International Organization for Standardization, ISO/IEC 13818 (commonly known as MPEG-2), "Information Technology—Generic Coding of Moving Pictures and Associated Audio Information," 2000.

[3] Recommendations of the International Telecommunication Union, Radiocommunication Sector, ITU-T Recommendation H.264, "Advanced Video Coding for Generic Audiovisual Services," 2005.

[4] Recommendations of the International Telecommunication Union, Radiocommunication Sector, ITU-R Recommendation BT.500, "Methodology for the Subjective Assessment of the Quality of Television Pictures," 2002.

[5] Recommendations of the International Telecommunication Union, Radiocommunication Sector, ITU-R Recommendation BT.601, "Studio Encoding Parameters of Digital Television for Standard 4:3 and Wide-Screen 16:9 Aspect Ratios," 1995.

[6] Public Safety Communications Technical Report, "Video Acquisition Measurement Methods," Department of Homeland Security, DHS-TR-PSC-07-02, 2007.