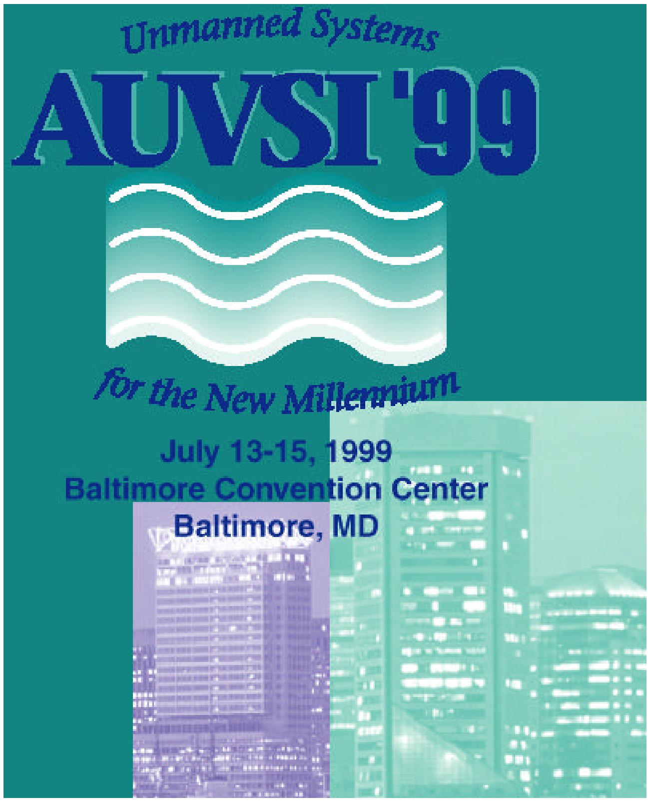


AUVSI, Baltimore, July 14th, 1999

"Sensors to the Desktop"

Robin Rowe, CEO/COO
MovieEditor.com
San Diego, California
619-232-5880
www.MovieEditor.com
Robin.Rowe@MovieEditor.com

Abstract

At a time when imagery and video play an increasing role in military operations the number of analysts to study such sensor output is actually decreasing. Who is going to look at all this sensor data from UAV's and UGV's, and UUV's?

Many would like to be users of sensor video, but don't have high bandwidth video channels available to carry such data to them. The solution is a client/server video-pull approach that keeps most of the content on the server, and extracts just the brief snippets of video that a user really wants to see. For example, a Predator UAV can stay in the air for 40 hours, but only a few seconds or minutes of that footage is of interest to any particular user. Bandwidth conservation enables users to access video over existing LAN, ISDN, modem, or radio links.

VERA (Video Exploitation Remote Access) provides a low bitrate interface to video streams to the desktop via a Java applet (Internet Explorer or Netscape Navigator). Development is sponsored by DARPA ISO and DISA through the Joint Program Office.

As part of this presentation an untrained volunteer from the audience operates VERA running on a palmtop Windows computer.

The Problem

The number of intelligence analysts has been declining, a trend that is expected to continue. At the same time, technological advances are making available an increasing number of motion imagery sensors. Unmanned aerial vehicles (UAV) are proliferating. Joint Strike Fighters (JSF) are expected to serve double duty as a sensor platform. Unmanned ground vehicles (UGV) will provide new sensors. Drug interdiction uses sensors based on aircraft and unmanned underwater vehicles (UUV). And, traditional cameras on U-2 aircraft and other platforms continue to be important because film provides such high resolution.

Recent figures from the British Ministry of Defense indicate that 70% of the value of imagery is missed in the first analysis. Even though without further study most of the value of imagery is lost, such study is often difficult to accomplish. Much of our military video imagery is "lost" in mountains of tapes or film reels without convenient indices or accessibility.

The Predator UAV has a wingspan of 48 feet and can stay aloft for 40 hours at a time. We have all seen Predator video on the television news showing troop movements in the recent conflict in Kosovo, not to mention prior conflicts in Bosnia and Iraq. One Predator flight is potentially twenty two-hour video tapes. Predator video has a data track that includes not only geospatial information but also aircraft systems data such as engine temperature, the type of data held on expensive flight recorders in commercial aircraft. If a Predator ever fails the prior tapes are analyzed to look for indications of a mechanical flaw that led to the failure. Consequently, every video tape ever recorded by a Predator, even test flights, has been saved.

DARPA recently had an interest in muzzle flashes of tanks. The question was posed, "Would it be possible to create software that detects muzzle flashes automatically?" So DARPA asked, "Do we ever fire tanks that are observed by Predator?" The answer was in the affirmative, that this happens routinely. DARPA asked for a tape showing muzzle flashes from any exercise. They were told it would be easier to fly a new mission than to find existing footage.

More sensors and less eyeballs. How do we find more eyeballs to look at motion imagery?

Existing Approaches

Actually, there are many eyeballs that would like to see video today who don't have it available to them. Video uses much more bandwidth than conventional data, and can create an overwhelming load if placed on network links. Dedicated broadband video links are very expensive. B2PA, and its follow on, the Global Broadcast System (GBS), bring Predator feeds to the Pentagon Joint Staff and other locations. Still, these links don't reach many of those who would like to monitor and search Predator footage, particularly in the field. Although bandwidth continues to get cheaper, the cost of dedicated satellite or land links is enough to discourage its use for wide delivery of reconnaissance video in the near future.

What if we had more pipes to serve more eyeballs? We would like to offer video to personnel who don't get access to raw video now, such as forward deployed units, ships at sea, staff on conventional TCP/IP networks in offices, and in some cases even to the Internet where hundreds of thousands of individuals might view it.

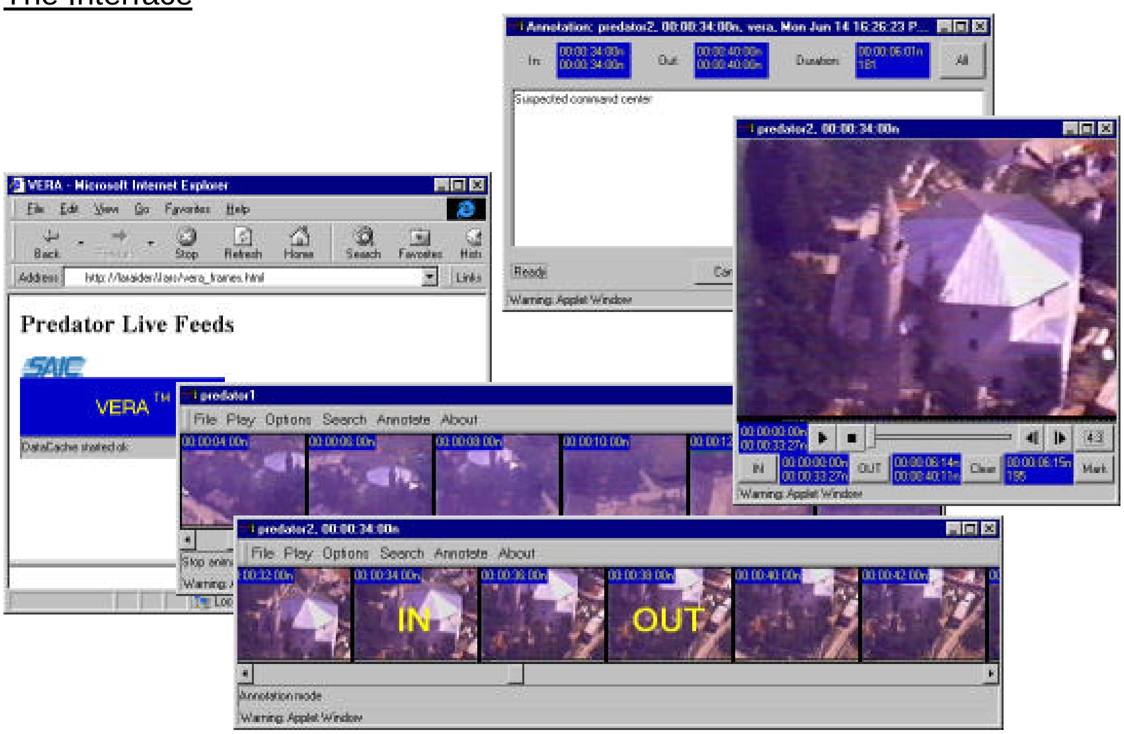
Video compression conserves bandwidth, but even compressed video requires huge pipes. For conventional video applications the requirement might be an OC-3/STS-3/STM-1 link at 155 mbs, capable of carrying about 10 simultaneous 10-bit digital analog video signals without audio.[i] That sort of bandwidth is simply unavailable to typical users. That's a hundred times the bandwidth of a T-1 line. You won't see any OC-3 lines hanging off the back of a ship at sea. But, if we could somehow bring sensors to the desktop over existing pipes then we would have a solution.

A Solution Based on Video Browsing

In any Predator flight it is typical that only a few minutes of motion imagery is actually of interest to a particular observer. If we could somehow avoid transmitting the "uninteresting" video we would see a huge bandwidth conservation of 800 to 1, dropping the flight's video duration from 2,400 minutes to maybe three. Using MPEG-1 compression at 1.15mbs, this represents the difference between transmitting 24GB to a user and just 30MB. Using replicated databases, scaled bandwidth, and video pull technology, and a browser applet, we bring full resolution video clips over existing low bandwidth pipes.

Replicated databases have been around for many years, but as a technology it lacks the glamour to be something that most people would be familiar with. Replication is often used in transaction processing, for example banking, where it is important to have an alternate database server should a system go offline. In case of failure the secondary database server, containing exactly the same data as the primary server, immediately comes on line. (Replication servers are often separated geographically for reliability.) Applying a twist to the replication approach, our browse recorders create "imperfect" copies of video. These low-resolution video pictures are just good enough that a user can decide what to skip over viewing. When the user finds something of interest he can mark the low-res imagery to indicate which detailed high-resolution imagery he wishes to download.

The Interface

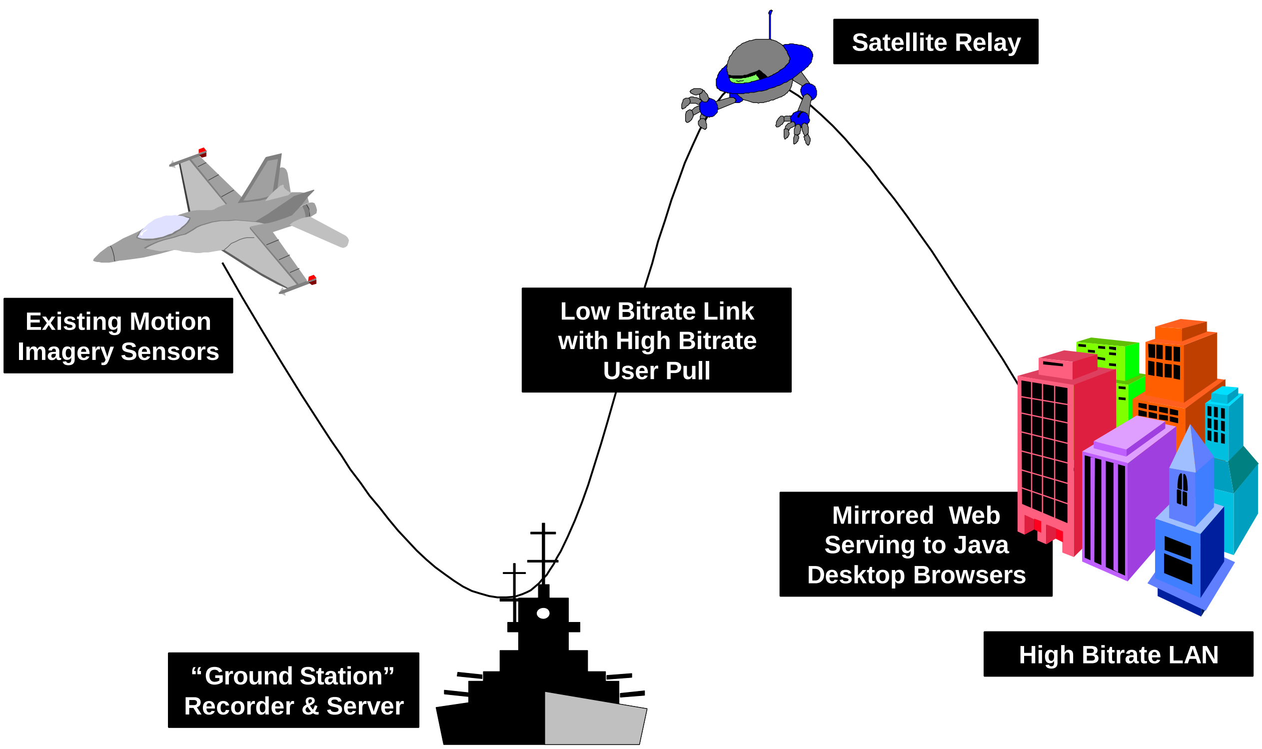


A user logs into our Web server using their browser, that is, Internet Explorer or Netscape. A Java applet pops up and interrogates the browse stream server to find out what video streams are available. Note that these are not video streams in the usual sense. Absolutely nothing is transmitted to the client until it is requested. An unwatched stream consumes zero bandwidth.

In the user interface the filmstrip continuously slides from right to left as new images are taken from the stream. When a user sees something interesting the thumbnail icons, which each represent two seconds of video, he marks IN and OUT points to carve out a discrete video clip that can be viewed. The user can "rewind" a filmstrip using the slider control to skim video he missed.

The continuous video stream is not pushed at the user, rather the clip is selected and pulled (or streamed) on demand. Once a user marks a video clip he can open an annotation window to write notes about that particular clip. A note that the clip contains "tanks" becomes a searchable entry in the database. A pointer to the video clip can be emailed for other users who can add their annotations. Or, the user can download the marked clip as a discrete MPEG-1 file to be attached to an email or embedded in a PowerPoint presentation.

The Architecture



The recorder digitizes the analog video (NTSC or PAL) into MPEG format that can be played on any PC. The browse stream recorder also extracts a sub-sample of JPEG thumbnail images snapped at two second intervals from the video stream. These small images are 120x80 pixels, and about 2KB each in file size. At a transmission rate of 1KB per second that is slim enough to go over a 9600 baud radio link or cell phone modem. If bandwidth is still too constrained, the client software can elect to pull every other thumbnail, cutting the bandwidth consumed in half.

For browsing we make two replicas of the video: one in MPEG-1 format at 1.1mbs (about 10 MB/min) and another as JPEG thumbnails at 8kbs. There is no limit to how many times this replication strategy can be repeated.

MPEG

MPEG-2 is a popular CODEC used by DVD and satellite broadcast TV. MPEG-2 is true broadcast quality video (or better). We are adding support for MPEG-2, even though that format is too big to go over many communication links. The idea is to provide video to the user in whatever format desired at whatever bitrate they can handle. It's worth pointing out that the highest bitrate video always ends up requiring the lion's share of the server storage. The cost to store replica copies is not a significant impact. A year's worth of JPEG thumbnails is about 32GB, but the same space can only hold about 13 hours of MPEG-2 video.

MPEG-2 is actually multiple CODEC's in various "profiles," including HDTV. MPEG-1 has roughly the picture quality of VHS EP-mode video, and isn't broadcast quality like MPEG-2. However, at 1.1mbs MPEG-1 has 1/4 the typical bitrate of MPEG-2 and doesn't require a hardware decoder for playback on a computer. At 1.1mbs MPEG-1 consumes 10MB/min. MPEG-4, which isn't really a standard yet, is better suited to restricted bandwidth Internet delivery. Its data rate is typically 250kbs and below. Microsoft AVI and Apple Quicktime are other popular video formats we are working to support.

HDTV

The Navy has been demonstrating some very interesting advancements with the HDTV 720 fps progressive 60 fps video format. Meetings of the NIMA Video Working Group in Washington D.C. have provided the Navy with an opportunity to show off some impressive examples of this format, which has stunning picture detail.

Today, U-2 reconnaissance uses film because conventional video quality is judged to be inadequate. As a consequence, U-2 footage is typically flown to Washington for distribution and many who want access can't view it. Even for those who get the film there is the time delay of processing, duplication, and distribution.

HDTV 720 progressive format at 60 fps is capable of replacing film in most military reconnaissance uses. But, HDTV is a 1.5gbs signal when uncompressed or 360mbs under the proposed SMTPE 259M-D SDI compression standard. Even compressed 720p content is practically 100 times the already massive bitrate of DVD-quality MPEG-2. We're talking about video files that are 3GB/min in size. Even so, we are researching adding support for this format. We have a solution for accessing this content over existing TCP/IP networks without overwhelming them.

Recall that we are carving out short video clips from continuous video streams so as not to overwhelm users' networks. We can't possibly bring a HDTV movie to most users desktops without a significant upgrade in their network capability. However, analysts who examine video are often looking for just one frame that contains the best view of the item they are interested in. Using the Java applet interface they can already single frame through the selected MPEG-1 clip. With HDTV support we will be able to extract any frame they select from that video, convert it to JPEG still image in high definition, and pop it open in another Java window.

Using a "Madden" buffer interface, analysts will be able to scribble markings on that image which they can share with other users. Incidentally, for users who are interested in receiving HDTV 720p movies it is worthwhile to note that Microsoft has announced upcoming support for this format in Windows NT. Commercial video board makers are busy creating inexpensive hardware intended for the mass market to playback 720p on computer monitors. This technology is predicted to debut within the next year at a cost in the hundreds of dollars.

HDTV presents interesting challenges to sensor platforms. For instance, if you place an HDTV camera on a UAV how do you acquire enough bandwidth to transmit the signal off the aircraft in real time? For such situations the browse server may actually be placed on the aircraft, with mostly low-resolution imagery being provided in real time to the ground. A burst of high definition imagery could be pulled down in parallel using the bandwidth conserved by using the low-resolution browse stream. This high-resolution data would be somewhat delayed, since it would exceed the real-time bitrate. An analyst might get a cup of coffee while waiting for his HDTV clip to download.

Low Bitrate Video

We've been talking about high bandwidth HDTV. Let's go to the other end of the spectrum. How low can we go?

Navy SEAL teams can carry sensors ashore that would be very interesting to monitor from ships at sea and all the way back to the command center. Transmission of forward deployed sensors presents difficulties both for power conservation and for detection through emitted signals. We are experimenting with inexpensive COTS megapixel still cameras that we control from a PC through a serial port interface. The goal is to create a sensor that has modest power requirements and minimal signal emissions under most circumstances, but high resolution on demand.

When something interesting is observed happening in the live thumbnail filmstrip a control signal is transmitted to the still camera and it snaps a megapixel image for transmission. A browse stream "ground station" located on a ship at sea, perhaps a submarine, will store and relay the images captured. A repeater ground station on shore will archive and serve the imagery to a wider audience. An asset at the scene carrying a palmtop computer will be able to mark up images as they are captured and immediately receive back images marked by analysts in the rear. Aircraft in flight will be able to monitor video on the ground in real time.

Relevant Applications

There are many useful scenarios for scaled bandwidth capabilities beyond traditional reconnaissance. A downed airman can show his situation and enable those coming to his aid to suppress or avoid hostile forces. A biological weapons defense team can deploy cameras that are in harm's way. Micro Air Vehicles (MAV) deployed to fight terrorists would benefit from minimal power transmissions. Plus, a broad area of interest is security video.

Secure government facilities deploy cameras, as do commercial buildings. Security cameras seem ubiquitous. However, facilities have the same problem of a decreasing number of eyeballs. The new San Diego courthouse, for instance, has integrated the jail in the same building. Judges would look upon it as a bad thing if prisoners broke out of jail and into their courtrooms. Manpower reductions, however, mean that fewer people than ever are watching these cameras to make sure nothing bad is happening. A video browse system would enable more eyeballs to keep track of things.

Video cameras are being deployed in police squad cars, as anyone who watches cops footage on television these days can't help but notice. In San Diego county, police cars have a 19kbs radio data link and an onboard PC. We are researching bringing together this video from police cars as well as helicopters. U.S. Customs and JIATF (the U.S. agency that handles drug interdiction coordination with foreign countries) are another law enforcement reconnaissance application.

Even guided munitions, such as smart bombs and Tomahawk missiles, could be enhanced with video browse streams. Using scaled bandwidth in hostile military situations is of particular interest because of the problem of jamming. The less signal you have to transmit the less it can be interfered with. And, with less signal you can devote more transmission power to breaking through jamming.

Future Directions

Image processing research has interested us for many years, but we've found progress hampered for lack of a powerful manual system that works operationally. With a deployed video browse system it will be possible to use real analysts' output to define a baseline for metrics against which to judge an automated image processing system. As image processing systems compete with humans (preferably a lot of humans) the algorithms will improve. We want to solve the image processing challenge problem of, "Find everything that looks like a Scud missile." In working with our other automated video systems that monitor broadcast television we program pagers to instantly react to events in the news. We look forward to doing the same with Predator footage.

Our development was sponsored by DARPA (www.darpa.mil). We are also grateful to DISA (http://www.disa.mil) and the Joint Program Office for ongoing support to transition our prototype into the Battlefield Awareness and Data Dissemination (BADD) component of the Global Command and Control System (GCCS). We have tested our software at sea on the flagship USS Coronado (http://www.comthirdflt.navy.mil/c3f/agfhome.htm). It has been installed at SPAWAR SSC-SD and the Pentagon Video Recording Facility.

[i] William C. Philipson and Steven L. Storozum, "Transporting Production/Mastering-Quality Digital Video Signals Through the Public Network Using SONET/SDH and ATM Technology," SMPTE Journal, February 1999.

SMPTE is the Society of Motion Picture and Television Engineers (www.smpte.org). MPEG-1 and MPEG-2 are standards of the Moving Picture Experts Group (www.mpeg.org). JPEG is a standard of the Joint Photographic Experts Group (www.jpeg.org).

Robin Rowe is the CEO/COO of MovieEditor.com. During his tenure as chief technologist at SAIC, a Fortune 500 IT company, he designed VERA. Before joining SAIC he was a research scientist at the Naval Postgraduate School in Monterey where he worked on virtual weather in VR battlefield simulation for the army, created software for experiments in night vision for ONR, and taught Advanced C++. He was president for seven years of a software consulting company based in Chicago and Redmond, and taught C++ at night school at the University of Washington Extension. He helped build the robotic studios at Chicago NBC station WMAQ-TV. He was technical director of broadcast news at WICD-TV, an NBC station in central Illinois.