APPLICATION OF DIGITAL VIDEO TECHNOLOGY TO THE AAR TRAINING-FEEDBACK PROCESS

George A. Scherer
NATIONS/AST
Orlando, FL

William J. Curran
NAWCTSD, Orlando, FL

ABSTRACT

The National Training Center at Fort Irwin provides Force-on-Force and Live-Fire training of US Army soldiers in an environment that is as close to real-world conditions as possible. Training exercises are conducted and recorded, with important sequences edited and played back for evaluation of the soldier's performance - a process known as "After Action Review" (AAR). During a simulated battle using live players, material is gathered from numerous sources including video, audio, tactical communications, instrumentation system graphics and data. Actions involving Armored and Mechanized Tactical Maneuver units, Light Infantry, Air and Fire Support units are recorded on videotape for subsequent incorporation into the AAR.

The increased workload imposed by the Advanced Warfighter Experiment (AWE) created a need for better and faster AAR preparation. To address the additional requirements of AWE, a new approach, involving the application of digital video technology on a previously unheard of scale, was designed and implemented at the NTC. The new system includes one theater presentation, nine video logging, and four edit/post-production workstations. The architecture is based on four networked video servers, each supporting four workstations, providing shared access to 810 GBytes of RAID 7 storage. This system cuts AAR preparation time in half, allows simultaneous digitizing, sharing, and editing of up to thirteen video sources, and eliminates the cumbersome storage of hundreds of videotapes from each rotation, while maintaining broadcast quality video.

AUTHOR BIO

George A. Scherer is a Senior Systems Engineer with Advanced Systems Technology, Inc. (NATIONS/AST), providing technical support to the US Army STRICOM Logistics Directorate in Orlando, Florida. He is currently responsible for the design and integration of professional audio and video systems at the Combat Training Centers worldwide, as well as providing computer systems engineering support.

William J. Curran is an Engineer in the Research and Engineering Competancy department in the Science and Technology Division at the Naval Air Warfare Center Training Systems Division (NAWCTSD). He is currently providing technical support to the US Army PM-TRADE, as a multimedia consultant.

APPLICATION OF DIGITAL VIDEO TECHNOLOGY TO THE AAR TRAINING-FEEDBACK PROCESS

George A. Scherer
NATIONS/AST
Orlando, FL

INTRODUCTION

STRICOM, the US Army Simulation, Training, and Instrumentation Command, operates the National Training Center at Ft. Irwin, CA. In this remote installation in the Mojave desert, STRICOM conducts Brigade level training exercises, known as a rotation, typically once per month. All players (humans and vehicles) are instrumented, outfitted with Multiple Integrated Laser Engagement system (MILES), and observed over a period of approximately two weeks. The activities and performance of each unit are recorded and evaluated. Tactical position and status data, video battle footage, and narration are all edited into the "After Action Review" (AAR) package. The AAR is the process of evaluation of the soldiers' performance. During the live battles, between the visiting soldiers (BLUEFOR) and the resident (OPFOR) personnel, information and material is gathered and recorded from numerous sources that include video, audio, tactical communications, instrumentation system graphics and instrumentation data. Battle actions involving Armored and Mechanized Tactical Maneuver units, Light Infantry, Air and Fire Support units are recorded on videotape for subsequent incorporation into the AAR Battle Action Summary (BAS) tape. The AAR process, which normally lasts up to two hours, offers each soldier an opportunity to discover their own mistakes and learn how to improve real battlefield performance. The sooner the individual can get training feedback, the greater the value of the AAR.

HISTORY

The key to training effectiveness is the quality and timely preparation of the AAR, which provides immediate feedback to the trainees. In the past, only two video sources were ever recorded at one time. The video, captured by Sony Betacam cameras and broadcast back from the field, was recorded on a Sony Betacam master videotape, which was then shared among all eight Training and Analysis Facilities (TAF). At the same time, during the battle, the TAF analyst would manually note the desired video cuts on paper. Digital and Hi-8mm tape recordings are also hand-carried in from the field. The video cut log would then be delivered to a Voice/Video Collection and Editing (VVCEC) operator, who would search the master tape, and copy the cuts to another tape. Since each TAF conducts their own AARs, with differing video requirements, each TAF would have to wait until the master tape copy was available. Four linear tape A-B roll edit suites were shared by all TAF groups, used to prepare the 7-minute AAR Battle Action Summary (BAS) tape. The entire process could take as long as six to eight hours. Post-rotation VVCEC activity includes preparing VHS copies of the BAS tape for inclusion in a take-home package.

PROBLEM DEFINITION

The Advanced Warfighter Experiment (AWE), held at the NTC in March of 1997, introduced additional multiple video sources to record. These included Appliqué graphics monitor snapshots, and up to eight mobile camera feeds at one time, rather than the usual two. Appliqué is a tank commander's battlefield awareness and information management system. Additionally, the time allotted to prepare the AAR tape was reduced to two hours. This increased workload imposed by the AWE created a need for better and faster AAR preparation. STRICOM desired to provide the NTC with a new system, one that would handle the increased load imposed by AWE, and produce better quality products in less time. The videotape-based linear system would not be able to keep up with the increased number of video sources and shortened AAR development times, due to the time required for linear A-B tape searching, editing and compositing, increased tape wear, tape generation losses, and the time lost waiting for the master tapes.

SYSTEM OVERVIEW

To address the additional requirements of AWE, and eliminate the limitations of tape-based systems, it was decided to implement a disk-based, non-linear, digital video system at the NTC. An approach using networked digital video systems was designed.

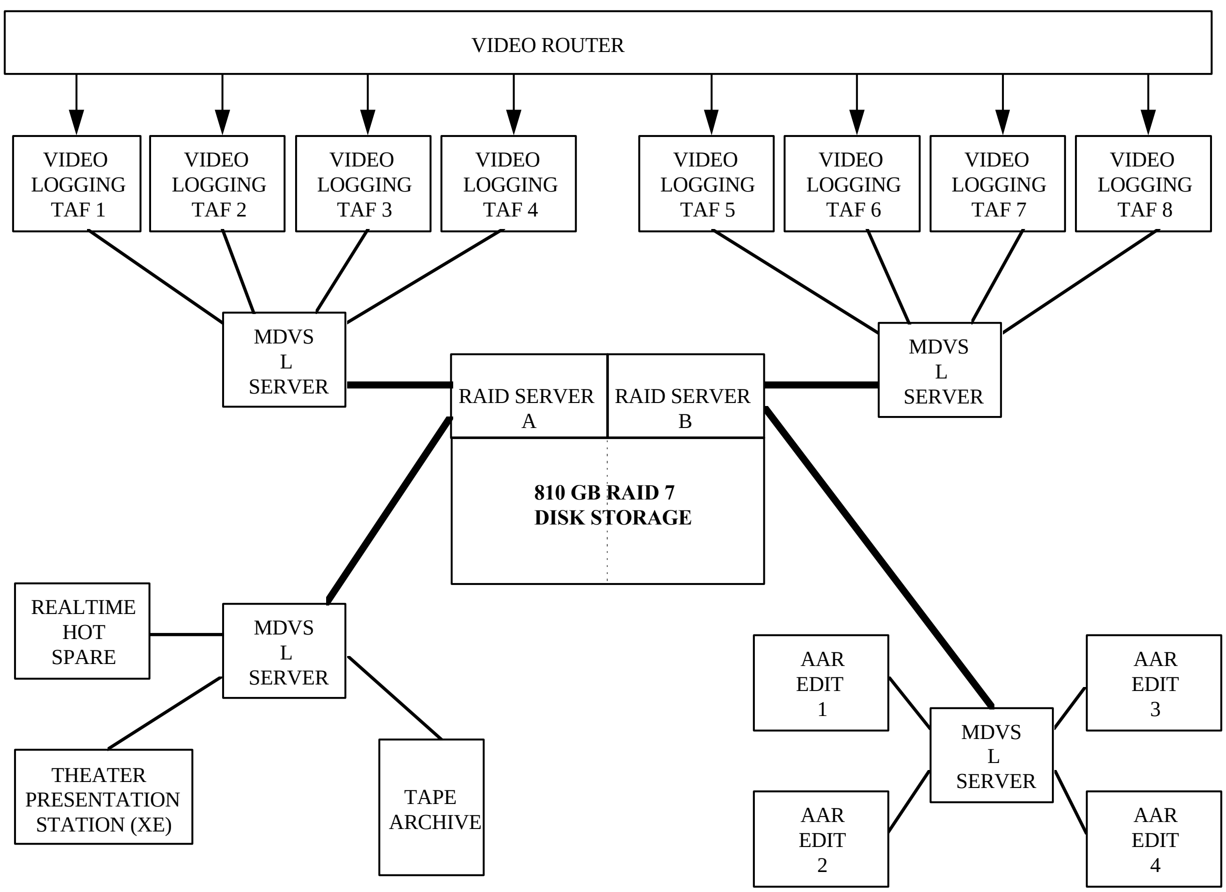

Figure 1. System Architecture

Architecture

The NTC digital video system consists of one theater presentation station, eight video logging stations, four edit/post-production workstations, and one hot spare on-line station, which can also be used to log video cuts. The architecture is based on four networked MDVS video servers, each supporting four workstations, thereby providing shared access to 810 GBytes of centralized RAID-7 storage. Fast/Wide SCSI, which will eventually be upgraded to Fibre Channel, is used to move video between MDVS servers and the RAID storage servers, and also between each client workstation and its local MDVS video server. The RAID storage, provides approximately 100 hours of on-line video and 375 hours of video archived on digital tape. 10 Megabit ethernet is used for the control interface from the MDVS video server to the client. The theater presentation station is used to present AAR program material digitally, directly from the server to the main theater screens.

Storage Requirements

It was determined that 8:1 compression (70-80 KB/frame) is acceptable for the training applications at NTC, which involve video recording of low-contrast desert scenes. The Storage Requirement was determined by using the following formula:

100 hours of video, at 70 KBytes/frame =
70 KB/frame x 30 frames/s x 3600s/hr x 100hr = 810 GBytes total storage

Ninety (90) 9GB hard drives were used to meet this requirement, plus an additional 6 hot spare (on-line) drives. Obviously, the storage capacity varies with the compression rate selected by the user. Each workstation can use a different rate of video compression. In this way, the 100 hour requirement could be met, using the maximum acceptable compression of 70 KB/frame.

Throughput Requirements

90 KB/frame was used as the baseline compression rate for throughput calculation, to allow for better quality video. The Throughput Requirement was calculated in the following way:

90 KB/frame = 2.7 MB/s x 4 workstations/ server = 10.8 MB/s throughput for one server.

10.8 MB/s x 4 servers = 43.2 MB/s total system throughput required.

A total of 15 workstation clients are required (8 Logging, 4 Edit, 1 Theater presentation, 1 hot spare, 1 Archive), which means 4 video servers are required.

Actual system throughput was measured, using test software, to be 43.6 Mbytes/s for Write only, and 62.5 Mbytes/s for Read only, with 14 stations active simultaneously.

SYSTEM DESIGN

Selection of Video Format

Many tape-based video formats are available. VHS is a low quality consumer tape format, unacceptable for professional use. S-VHS tape, or Hi-8mm tape, make use of Y/C format, where the color and luminance information in the video signal is separated for better performance. Component tape (analog and digital) has long been the professional industry standard. This includes Sony Betacam SP (analog component) and D-1 (digital component). Newer digital component tape formats are Digital Betacam, DVCPRO, DVCam, Digital S, and D-5. All of the DVC (digital) formats offer great promise, however, there are compatibility problems and it is still a tape-based format.

Non-linear (disk-based) digital video has the capability to exceed the quality of component tape. Non-linear digital video editing is significantly faster than linear tape editing, with no generational loss. The encoding of the analog video signal into digital storage for professional non-linear systems is primarily accomplished using one of two algorithms - Motion JPEG, based on the still-frame JPEG format, and MPEG. Other non-professional encoders do exist (AVI), however, these were not considered acceptable for this application.

Motion JPEG (MJPEG) is an industry accepted standard that specifies I-frame, 30 frames/s, dual field encoding. This provides the frame-accurate editing that is a necessity in the modern video editing environment. Every single frame is encoded, with compression variable from 2:1 to 120:1. Higher compression yields lower fidelity with lower storage requirements. Since the way MJPEG handles audio is not addressed in the standard, several different implementations are used by different vendors.

MPEG was designed to be a delivery mechanism for transport of video for wide area distribution. MPEG interpolates the data between I-frames. The MPEG algorithm compresses the data that did not change between selected I-frames. The compression rate is variable, being determined by scene complexity, content, and motion (i.e. background is compressed, person walking is not). The number of frames analyzed in one pass is variable as well, with increased spacing between I-frames yielding lower fidelity. If a frame has been interpolated away, then it is not possible to perform a cut on that particular frame. The frame data must be reconstructed before editing. It is extremely difficult to reconstruct the data after full compression. Most MPEG algorithms use 2-pass compression. The analog video is compressed using an MJPEG type algorithm, then the digitized video can be edited, then MPEG re-compresses to produce a final product. The final video product is saved as a separate file. All very high quality MPEG encoding is done using non-real-time encoders, with typical times of 3x-10x real-time to encode. Commercial DBS (satellite) technology uses non-real-time MPEG-1 encoding, which can be converted to MPEG-2 for distribution.

No frame-accurate MPEG based editors, with professional features are currently available. (Note: recently Sony announced a 15 frame I-frame/B-frame MPEG-2 editing system capable of every-other frame accuracy, which was not available at the time this digital video system was designed. The current availability of this system is unknown).

Composite (NTSC) is the lowest quality video format, where chroma and luminance are mixed into a composite one-wire signal. Y/C is a two-wire signal, where the where chroma and luminance information are separated. Component is the industry standard; it uses a three-wire signal path. Each channel of video information is stored separately. The result is the ability to resolve finer detail, far lower noise in both the color and black and white portions of the signal. Any professional digital video system must be able to support component signal processing.

Server Technology

The Mercury Computer Systems 1200R server that was selected is a Windows NT platform with 20 Mbytes/second Read/Write throughput. It is capable of supporting 300 KBytes/frame (2:1 compression) for two clients, or 150 KBytes/frame for four clients. A maximum of four clients per server are supported. It uses SCSI Fast/Wide interfaces to move video between clients and the central storage, and 10 Megabit ethernet for the control interface to the client. Client connections are limited to 100 feet. The Mercury 1200L server is capable of 80 Mbytes/second throughput (40 for Read, 40 for Write). It makes use of 100 Megabit ethernet for connections to the clients up to 1000 feet away, and it can support Fibre-channel interfaces. Operator intervention is not required for normal server operation. The operator would only use the server console to modify the table that controls user access to specific RAID storage logical volumes, or for maintenance.

As previously mentioned, it was determined that 8:1 compression (70-80 KB/frame) is acceptable for the training applications at the NTC. At the time the system was designed, the 1200L server was not available. The current system will be upgraded to Mercury 1200L Fibre-channel servers, when they become available. MPEG servers were evaluated and found to be unsuitable for this application.

Storage Technology

The Mercury RAID-7 storage subsystem consists of a two-part RAID-7 central server (A and B), with associated storage drives, with the two halves cross-connected to make all storage space available to all clients. The central server contains a total of 810 Gigabytes disk space, using 9GB drives. This provides 100 hours of video storage (@70KB/fr), with an upgrade path to a maximum of 600 hours (4.8TB using 9Gbyte disk technology). Proper environmental cooling must be provided, as heat generation is a consideration with such a large array of disks. Disk storage was partitioned into 12 volumes, one for each TAF user. Only one user at a time can write to a volume. Four of the volumes are used for special projects only.

Video Logging Stations

NTSC composite video is distributed from the main video/audio router to each TAF workstation (and the Edit stations) for encoding of video. Two centrally located Panasonic WJ-420 Color Quad splitters are used to provide a 4-in-1 video display for TAF analysts to monitor four channels at once. The desired composite video is then selected and digitized at the logging station. SMPTE Linear Timecode (LTC) is distributed to all logging stations from a Fast Forward timecode generator.

Each logging station consists of the following:

1 - Sony BVW-75 Betacam SP recorder/player
1- Macintosh 9500 computer platform with 17”monitor
1- Media 100 VINCENT digital encoder/decoder card
1- Sony PVM-1354Q color video monitor
1- LTC to 422 timecode converter
1- Pesa remote control for the central audio/video matrix router

This system can be used two different ways to digitize video at the logging station:

1. Digitize directly into the server -
This method gives the best video quality and is fastest, since the clips are available immediately for editing. The user must be careful to digitize only what is needed, or excessive disk space will be used up. Also, the battle action can be lost, if the user is not attentive.

2. Record on videotape, then auto-digitize -
Live video is recorded on Betacam tape while TAF analyst logs cuts. Only the cuts log is saved to the server. The user then uses the Media 100 automatic tape digitizing function to batch digitize only the desired video clips, without user intervention. In this way, no battle action is lost, even if the user is not attentive to the system.

3. The logging station is also used to review tapes.

AAR Post-Production/Edit Stations

Operators edit the AAR package together at the Edit station. Composite video can be digitized at the Edit station as well as at the Logging station. 1280x1024 graphical map snapshots are taken from a SUN instrumentation display system, and are input to the Folsom Research 9700XL scan converter, which outputs Y/C video to the Media 100 for digitizing. The Sony BVW-75 VCR is connected to the Media 100 via component I/O. The two computer graphics monitors are used as an expanded virtual desktop, which makes non-linear video editing a lot easier. Audio narration, provided by a TAF analyst in the edit room, is recorded digitally to the central storage as well. Audio is also imported digitally off of CD Rom and digitized from cassette (Tactical audio). When the operator has a completed AAR program, it is mastered to Betacam tape automatically. Once finished, The AAR program can be presented in the main theater digitally.

Each Edit station consists of the following:

1 - Sony BVW-75 Betacam SP recorder/player
1- Macintosh 9500 computer platform with two Sony 17" monitors, CD Rom
1- Media 100 VINCENT digital encoder/decoder card
1- Sony PVM-1954Q color video monitor
1- Folsom Research 9700XL scan converter
1- Pesa remote control for the central audio/video matrix router
1- Inline IN3592 composite video auto-switcher (2x1)
1- Tektronix vectorscope/waveform monitor
2- Powered speakers
1- Sony 8x2 audio mixer
1- JL Cooper MCS2 Jog/Shuttle Control
1- Nakamechi MR1 audio cassette deck
1- Microphone

The Media 100 workstations use a true CCIR-601 4:2:2 Component YCrCb Digital Video Architecture. The 4:2:2 component signal is maintained throughout the digital editing process. Component, Y/C, and NTSC composite I/O is provided. The Media 100 also provides 8 audio tracks, 20-20Khz @ 44.1/48 Khz sampling, with an 8x2 internal (software) mixer, connected to 2 ch I/O (bal/unbal) audio. All editing is performed on the original digitized video clip, by using Edit Decision List (EDL) pointers. No multiple copies of a clip need to be stored (saves space).

UPGRADE PATH

The current system will be upgraded to Fibre-channel servers, when they become available. Storage capacity can be increased by simply purchasing additional 9GB hard drives. Maximum storage capacity is at least 4.8 Terabytes (600 hours @ 8:1 compression). Additional servers and edit workstations can be added to a Fibre Channel system to increase AAR preparation throughput. The Mercury servers are open architecture, fully compatible with many different vendor's products.

Media 100 will have an add-on card available late 1997 that will provide digital I/O:

SMPTE 259M (SDI)
DVCPRO
D1
D5
ASEBU audio
SPDIF audio

SMPTE 259M is a widely accepted industry standard for digital video distribution (uncompressed). The ability to export MPEG-2 audio/video will need to be supported in the future.

CONCLUSIONS

The digital video age has arrived. The world's largest networked digital video system, based upon broadcast quality MJPEG algorithms has been designed and implemented at the National Training Center. This system cuts AAR preparation time to a third of the time it previously took, allows simultaneous digitizing, sharing, and editing of up to thirteen video sources, and eliminates the cumbersome storage of hundreds of videotapes from each rotation, while maintaining broadcast quality video. This advanced digital video system has proven to be a success during the AWE. The days of tape-based editing are over.

FUTURE APPLICATIONS

A similar system is being investigated for possible installation at Hohenfels, Germany (CMTC), and a more advanced system is under consideration for Ft. Polk, LA (JRTC).

REFERENCES

"Designing Digital Systems", The Grass Valley Group, Inc. TP3488-00, Issue B, February 1994.

Martin Bock, "RAID meets video - the ultimate matchup", Feb. 96 edition of Data Storage.

B. Beck and C. Stakutis, "Breakthrough Technology for Cost-Effective Video Editing Server" white paper, 1996, Mercury Computer Systems.

"Networks for Post Production Studios", Prisa Networks information paper.

Neil Weinstock, "Network Wars", June 95 edition of POST.

"MPEG Digital Compression Systems", International Telecommunication Union, Radio Communication Study Groups Document # 11-1/6-E, 11-3/15-E, 9 August 1994.

Guy Cherry, "Compressed Television Systems", Tektronix Video Systems information paper.

"Ethernet/IEEE 802.3", Cisco, internet url at http://cio.cisco.com/univercd/data/doc/cintrnet/ito/55771.htm#HDR2.

ABOUT THE AUTHORS

George A. Scherer is a Senior Systems Engineer with Advanced Systems Technology, Inc, providing technical support to the US Army STRICOM Logistics Directorate. He is currently responsible for the design and integration of professional audio and video systems at the Combat Training Centers worldwide, as well as providing computer hardware and software engineering support. He holds a degree in Electrical Engineering from Fairleigh Dickinson University (BSEET). Previously, Mr. Scherer was a Senior Software Engineer at Computer Sciences Corporation, where he spent six years designing vehicle monitoring instrumentation systems for the USAF space program at Cape Canaveral, FL. He also designed aircraft simulators and performed instrumentation of LANTIRN electro-optical systems at Martin Marietta for four years. He has also designed professional audio systems as an independent contractor.

William J. Curran is an Engineer in the Research and Engineering Competancy department in the Science and Technology Division at the Naval Air Warfare Center Training Systems Division. He is currently providing technical support to the US Army PM-TRADE, as a multimedia consultant on the JRTC-IS and JRTC MOUT IS programs. In addition, he provides professional audio and video systems engineering support to US Army STRICOM Logistics Directorate, regarding future generations of the Combat Training Centers. He holds degrees in Electrical Engineering from the University of Miami (BSEE) and the University of Florida (MSEE), and has completed coursework toward a Masters degree in Ocean Engineering at the University of Miami Marine Laboratory. Previously, he was the Principal Investigator in the Advanced Sensor Laboratory, where he developed AN/APS-137 ISAR simulation, with concomitant table-top trainers. He was Project Engineer on the Standard Ocean Model (acoustics), and the Passive Acoustics Analysis Tabletop Trainer, lead engineer on Ship Tactical Team Trainer.

Footnotes

[1] Throughout this paper, “Service” is used to refer to the military service components.

[2] The Army, Navy, Air Force, and Marines have implemented one or more distance learning programs with use of one or more of the following technologies: satellite broadcast, videoteletraining, and computer-based instruction on CD-ROM.