AN INTRODUCTION TO VIDEO IMAGE COMPRESSION AND AUTHENTICATION TECHNOLOGY FOR SAFEGUARDS APPLICATIONS

Charles S. Johnson
Sandia National Laboratories
Albuquerque, NM 87185

Abstract

The verification of a video image has been a major problem for safeguard applications for a number of years. Various verification schemes have been tried on analog video signals ever since the mid-1970's. These schemes have provided a measure of protection but have never been widely adopted.

The development of reasonably priced complex video processing integrated circuits makes it possible to digitize a video image and then compress the resulting digital file into a smaller file without any noticeable loss of resolution. Authentication and/or encryption algorithms can be more easily applied to digital video files that have been compressed. The compressed video files require less time for algorithm processing and image transmission.

An important safeguards application for authenticated, compressed, digital video images is in unattended video surveillance systems and remote monitoring systems. The use of digital images in the surveillance system makes it possible to develop remote monitoring systems that send images over narrow bandwidth channels such as the common telephone line. This paper discusses the video compression process, the authentication algorithm, and the data format selected to transmit and store the authenticated images.

Introduction

A major problem that has faced unattended surveillance systems is finding a way to verify that the image coming from a video surveillance camera is authenticated and has not been altered or delayed in time. Numerous attempts have been made since 1977 to develop ways to protect the images coming from a video camera. Some early attempts in 1977 were made by Fairchild Corporation in a prototype time-lapse video system which inserted random patterns into the video and checked for those patterns at the recorder. The Surveillance Television and Recording System (STAR) designed by Telemation in the early 1980's under contract to Sandia National Laboratories (SNL) monitored a 3.58 Mhz signal on the video cable for taps and insertions. The development for the first authentication systems occurred in the mid to late 1980's. A video authentication system which sampled the analog video signal in a random manner based on a one-time keypad was designed for the Modular Integrated Video System (MIVS) by SNL. A stand alone Modular Video Authentication System (MVAS) was later designed by SNL and used for protecting any video link. Dr. Neumann Electronik GmbH, working under the German Support Program for the International Atomic Energy Agency (IAEA), designed the Tamper Resistant TV-Link (TRTL) which provides protection for several video systems used by the IAEA. The later two authentication systems utilized digital technology but still had to process the analog video signal. The emerging digital video technology has at last made it possible to design a full digital authenticated video system for use in unattended surveillance systems.

Video Processing

There are two separate domains of video signal processing - analog and digital. Analog processing was developed in the 1930s along with vacuum tube technology and led to the television signals used in most of the world today. A scene before a camera was converted into time varying analog signals that represented the scene. The only pulses present in the analog signal were the vertical and horizontal

synchronization (sync) pulses that indicated how the scanning of the image was accomplished. The first analog cameras were limited primarily to applications in the television broadcast industry. The video processing would have been limited so if the transistor had not been invented. The advent of the transistor permitted the expansion of video signal processing to smaller and less expensive cameras that could be used for various other applications, such as surveillance cameras. Even the availability of the transistor signal processing did not change the actual video signal format. It remained an analog signal with sync pulses and interlaced scanning. The analog signal evolved into three major television standards - NTSC (National Television Standards Committee), PAL (Phase Alternating Lines) and SECAM (Sequential With Memory). NTSC and PAL systems are the two major systems for processing analog video signals.

Digital video processing began to emerge in the 1980s with the development of the personal computer. The analog signals from cameras must be converted for use in the digital domain. The signals are turned into a binary representation of a discrete point or voltage in the analog signal. The conversion process is called "analog to digital conversion " (ADC). Computers can also generate original digital video signals without ever having to go through the analog conversion process.

With only few exceptions, the computer digital image world ignored the television world and developed computer systems without giving much thought to the microprocessor clock speeds and monitor scanning methods. The resulting differences created difficulty as the worlds of video and computers began to converge in recent years. The television industry builds specialized processors to convert their analog signal into digital representations for the purpose of altering the signal, storing it onto magnetic tape, or in the last few years transmitting it by satellite. Computer processing of images is being directed at emerging application areas of communications, entertainment, and multimedia.

Proprietary digital video systems for communications have been around for a decade or more, but none have been widely adopted. New standards for digital video processing and transmission are emerging based on open systems with standardized implementations.

Using technology based on JPEG (Joint Photographic Experts Group) and MPEG (Motion Picture Experts Group) standards, the entertainment industry is beginning to take advantage of digital video processing. The home entertainment market is expected to be very large in the future. "Movies on Demand" will be selected from huge digital servers containing hundreds of movies in a digital format. HDTV (High Definition Television) and interactive video games will also use digital signal processing to bring entertainment to the consumer. To be cost effective, the surveillance developers will have to borrow technology from the main surge of video compression technology and adapt this technology for their limited applications.

Authentication

Authentication has always been a difficult task for video images, especially the analog images from a surveillance camera. The present day video technology makes it very easy to substitute and manipulate video images. There is very little assurance that an image being reviewed by an inspector is a true image. As previously mentioned, there have been three analog video authentication systems designed and used in a limited number applications; but there is still an urgent need to upgrade video image protection techniques. Changing technology has caused the direction of development to move from protecting analog video images to determing how digital video images can be protected.

Since the digital image must be protected from the time it leaves the camera enclosure, it is essential that the system developed must be small enough to be mounted inside the camera enclosure along with the camera. The small size requirement indicates the need for a highly specialized integrated circuit approach. System engineering choices must be made about how the design of the digital video module. One approach is to digitize the video image and authenticate the resulting large file of 400 to 500 Kbytes. The second approach is to compress the video image to a small file of about 40 Kbytes and then authenticate the image. The latter approach of applying a compression algorithm to

the digital video image before authenticating the image file is preferred.

Digital Image Compression Technology

Digital image compression technology is the science of converting a large digital image file representing a scan image into a smaller file that still represents the same image. Image files must be compressed so that more images can be stored on any selected digital storage media and that more images can be sent through any given transmission link. There are four ways to reduce the size of the resulting digital files associated with video images. The ways are:

- Reduce resolution
- Reduce frame rate
- Reduce color fidelity
- Reduce redundancy

Video processing circuitry must be designed to perform one or more of these tasks. If an image is digitally sampled at lower rates, the resolution of the image will be reduced; and the resulting image file will contain fewer bytes. A byte is generally used to represent the digitized value of an image pixel (picture element). A byte allows the value of the pixel to be represented by 256 shades of gray. Some early compression systems used 4 or 6 bit values when the cost of solid state memory was very high. Now some video signals are digitized to 10 and 12 bits per pixel. Eliminating pixels to compress the file size is accomplished by sampling each line of the image at a slower rate and skipping lines in the image. The result of reducing the number of bytes by reduced sampling of the available pixels is to reduce the resolution of the image. Improper sampling ratios can also introduce visual artifacts into the compressed image. (The under sampling approach was the only practical method available for many years to reduce the size of image files.) A reduction of image resolution also occurs with redundancy processing of the video image.

Frame rate reduction is one way to lower data rates when motion images are being transmitted or played from stored media. The standard NTSC image displays 30 frames per second in order to reduce flicker and give the illusion of smooth motion. Some teleconferencing systems will display only 12 to 15 frames per seconds in order to save bandwidth. Such rates are acceptable for this application since people do not tend to move about rapidly. Frames below 12 per second will quickly become objectionable. Frame rate reduction is a natural part of most surveillance systems since only single frames are recorded between relatively long time periods.

All video systems use the reduction of color fidelity as part of their design. The eye is not very sensitive to color information so the color signals are separated from the luminance signal and sampled at lower rates. The video processor should be designed to handle both monochrome and color video signals. Computers handle color information in the RGB color space. All the colors created are defined by using different proportions of red, green, and blue signals. All the television systems use color spaces to conserve bandwidth. The NTSC composite video signal uses a YIQ color space, while PAL composite video signal uses a YUV color space. The "Y" stands for the luminance signal which is made of specified proportions of RGB signals. The "IQ" and "UV" parts are chrominance signals that are inserted into the luminance signal. The reduced bandwidth chrominance signals are mixed with a carrier signal to form a quadrature double side band amplitude modulated carrier. The bandwidth of the chrominance signal can be reduced because the detail information is in the luminance signal and the human eye is less sensitive to color detail. By using the YIQ/YUV color spaces instead of RGB, the information in the chrominance channels can be sampled at lower rates without losing much visible resolution. So the first step taken in a compression algorithm decodes the composite video signal to separate RGB signals and then uses a straight mathematical transformation to convert RGB to YIQ or YUV.

The most technically demanding task for video processing circuitry is the problem of reducing redundancy. The major thrust of most redundancy reduction algorithms is directed toward transform coding and other forms of image processing. High speed integrated circuits can now perform the functions which required a laboratory computer for a few years ago. Video can only be compressed because its original design left temporal and frequency space in its signal structure. An analysis of the common video signal shows a lot of low frequency energy (the luminance part of the signal) with remaining information concentrated in high frequency

sidebands around harmonics of the scanning frequency.

The trick of any compression algorithm is how to extract only the essential information. The major approaches used to reduce redundancy are intraframe, interframe, and predictive coding. Intraframe coding removes redundant information within a single frame and is applicable to both single images as well as a series of images or motion pictures. In the situation of a series of images creating motion, then interframe coding can be used. Interframe coding keeps track of the pixels that change from frame to frame and transmit only the changes. Predictive coding go one step forward and attempts to predict where moving objects will be in the next frame.

Each of the coding techniques can be implemented separately, or they can be combined to produce a compression algorithm. There are measurements that can be used to measure the best algorithm, but the final judgment is always somewhat subjective. The best indicators are picture resolution, color fidelity (for color signals), motion handling ability and overall frame capability. The latter two indicators only apply to algorithms that attempt to handle live motion video.

As with any data compression scheme, there are always many tradeoffs. Anytime video compression is perform, performance parameters will have to be comprised. An example is the difference between image quality from video tape recorders. The picture from a broadcast Type C one-inch machine is greatly superior to the picture from a home VHS video recorder. The performance parameters that were sacrificed to build the VHS machine were bandwidth and color fidelity. The changes in picture quality is evident as the level of compression is increased. Different artifacts will appear with different algorithms. Transform coding algorithms using discrete cosine transform functions will begin to show blocking or tiling; others will begin to get fuzzy. Spark-like points will begin to form halos around the edges of objects. Blurring of moving objects will occur for motion video algorithms.

<u>Compression Algorithms</u>

There is a large field of research on image compression technology, but only a few are being commercially implemented and still fewer have had the specialized integrated circuits fabricated to permit the algorithm utilization on a cost effective basis. Some of the compression techniques are having standards officially developed under the auspices of organizations such as the International Telephone and Telegraph Consultative Committee (CCITT) and International Organization for Standards (ISO).

The major compression techniques are:

- Px64 Group/H.261
- JPEG (Joint Photographic Experts Group)
- MJPEG (Motion JPEG)
- MPEG (Motion Picture Experts Group)
- Wavelets
- Fractal

A number of specialized and proprietary compression techniques have been introduced that are based on the basic concepts. Some of the more widely used techniques are:

- Indeo (Intel Video)
- DVI (Digital Video Interactive)
- Truemotion
- Cinepak

The two leading compression algorithms for use in safeguards systems are JPEG and Wavelets. The JPEG algorithm is the most used algorithm today since it can be implemented both in software and hardware. Special compression integrated circuits (IC) have designed by a number of different semiconductor manufacturers. ICs have not yet been marketed to perform the Wavelet algorithm, but they probably will be within the next year. At that time the use of a particular algorithm will be determined primarily by the compression ratio required for a given quality of image and the type of artifacts that are acceptable. Recent developments seem to indicate that wavelets may permit higher compression ratios for the same subjective quality of images.

Several digital video systems using compression technology are now in various phases of development, field test, or implementation for safeguards application. Hymatom in France has developed the EMOSS which is being used by Euratom. Euratom and the IAEA are using the GEMINI which was developed by Aquila Technology Group. The German Support

Program to the IAEA is presently funding the development of a Video Data Authentication and Encryption (VDAE) device which will be installed in camera housings to digitize, compress, authenticate, and encrypt video images. The VDAE is being developed by Dr. Neumann Electronik GmbH. SNL is using digital compression in the Remote Monitoring Systems that are being field tested under the International Remote Monitoring Project. SNL will shortly field test systems using the Image Compression and Authentication Module (ICAM). The ICAM can be installed in a camera housing to provide authentication on the transmission link from the camera. All the previously mentioned systems all use some form of JPEG as the compression algorithm. Los Alamos National Laboratories (LANL) has developed an Experimental Inventory Verification System (EVISystem) which uses the Wavelet algorithm for unattended image collection.

Authentication Algorithms

Safeguards monitoring involves the acquisition of data and images to be used to verify compliance with an agreement. In unattended monitoring scenarios, equipment is deployed for long periods of time without knowledge of its status. Data acquired by these unattended monitoring stations is vulnerable to tampering during transmission or transportation to the inspectorate. Data authentication is a cryptographic technique which can be used on digital files to ensure that the compliance data received by an inspectorate is exactly the same data that was produced by the monitoring system. Public-key data authentication using a digital signature algorithm allows the inspectorate to ensure that the data has not been modified after the authentication process. It identifies the origin of the data and provides the ability to prove when and where the data originated.

Authentication is needed to meet the requirement to verify the correctness of any data or image used in safeguards application. In the case of similar images, it is very important to know the date and time the images were recorded. The date and time must become part of the image data file in order to protect against substitution of a genuine image that occurred at an earlier time. The actual image can be used to identify the origin of the image since it must be able to match all earlier reference images. Encryption can be used to provide authentication functions but is generally not preferred since host countries may want to examine the images. Public-key authentication algorithms are the best approach to authentication since they permit sharing the verification key without revealing the key which was used to produce an authentication signature for the image file.

Digital Video Format

The challenge for the future for digital video surveillance systems is not the technology but the standards for the data formats. While it is true that computers can be programmed to handle most any digital data, it would be far easier if a basic digital video format could be developed as a standard. The number of different formats already being used is an indication that all the developers will tend to travel their own paths unless there is a standard to follow. Work has been done under the sponsorship of the IAEA in a recent conference to put forth a format for consideration. The format is shown in figure 1 for reference.

Field Description	Bytes	Definition
Serial Sync Pattern	4	Serial Sync Pattern
Facility ID	2	Facility Number
Camera Number	1	Number assigned to a camera in facility
Image File Storage Type	2	JPEG, TIG, GIF, MPEG, etc.
Compression Type	2	Information about compression algorithm
Compression Ratio	2	Or Quality Factor
Image Standard	1	NTSC; PAL; SECAM
Digital Sampling	1	Bits per Sample (8, 10, 12 bits)
Color Processing	1	Y, U, V; RGB; Y, I, Q; Y, R, B ; HIS
Pixels - Horizontal	2	320; 640; 720
Pixels - Vertical	2	250; 256, 480; 512,
Window Coordinate "X"	2	Beginning "X" Coordinate of sampled window
Window Coordinate "Y"	2	Beginning "Y" Coordinate of sampled window
Sampling Rate	4	Sample frequency used for video signal
Frame Rate	1	Single Frame to 25/30 frames/sec
Scan Format	1	4:3; 16:9
Authentication	1	Type of authentication algorithm
Camera Event	1	State of Health/Tamper Status
Frame Number	4	Number of a transmitted video frame.
Trigger ID	1	Identification of the Cause of Image Storage
Date	8	Year, Month, Day
Time	8	Hour, Minute, Second, Tenths of Seconds
Image Frame Size	4	Number of Bytes in Video Frame
Digitized Image Frame	Var.	Output of the Image Processor
Authentication Signature	40	Authentication Signature for Image Frame
Link Field Indicator	1	"0" if link data does not exist
Link Field Size	4	Number of bytes in link field
Link Field Data	Var.	Special data about video image

Figure 1. Digital Video Format

The Future

The technology for digital video systems is developing at a very rapid rate today. Digital video systems will offer new capabilities to handle and process images from unattended remote surveillance systems. The processing power of the computer will bring about increased accuracy with less chance for human error. The combination of compressed digital video and authentication techniques provides the framework today for designing and building the secure remote monitoring systems of tomorrow.

References

1. Charles S. Johnson, "Optical Surveillance Equipment for the Late 1990's", *INMM 31st Annual Proceedings,* July 1990, Vol.XIX, pp 548-551.
2. H. G. Wagner, "Trends in Digital Video Surveillance at Euratom Safeguards", *INMM 34th Annual Meeting,* July 1993, Vol. XXII, pp 723-730.
3. Charles S. Johnson, "Application of Digital Compression Techniques to Optical Surveillance Systems," *INMM 32nd Annual Proceedings,* July 1991, Vol.XX, pp 841-845.
4. Charles S. Johnson, and Julian Whichello, "Adapting Digital Video Technology to the Surveillance Requirements of the IAEA", *INMM 32nd Annual Proceedings,* July 1991, Vol.XX, pp 846-851.
5. Cheryl Rodriguez, J. E. Brown, Peter Chare, John Goerten, and Hans Wagner, "Unattended Digital Video Surveillance: A System Prototype for Euratom Safeguards", *INMM 35th Annual Proceedings,* July 1994, Vol.XXIII, pp 1024-1029.
6. K. J. Gartner, J. Whichello, "Digital Image Surveillance for IAEA Safeguards", *INMM 35th Annual Proceedings,* July 1994, Vol.XXIII, pp 1018-1023.
7. Majid Rabbani, and Paul W. Jones, "Digital Image Compression Techniques", SPIE Optical Engineering Press, Bellingham, WA. (1991).
8. William B. Pennebaker, and Joan L. Mitchell, "JPEG Still Image Data Compression Standard", Van Nostand Reinhold, New York, NY. (1993).

This work was supported by the United States Department of Energy under contract DE-AC04-94AL85000.

DISCLAIMER