Video security is moving quickly from analog to DSP-based digital systems and from JPEG to MPEG-4 technology.

Advanced DSPs and MPEG-4 Build a Digital Video Security System

By Marc Guillaumet

The world of video security is experiencing extensive changes: reduction of the quantity of data transmitted while keeping the same level of information, flexibility in usage conditions, and interoperability to allow deployment anywhere. As part of these changes, system intelligence is no longer concentrated in one location: digital video cameras not only capture video but also manage audio and video compression, synchronization, and streaming over the network in real time. These cameras take advantage of smart, advanced digital signal processing and video compression algorithms that enable high-quality images to be transferred at low bit rates.

Since the Internet is the universal communications medium, video security systems must be able to stream over IP to be compatible with all network devices, thereby allowing multimedia access from any location. (The supervision application is implemented on a system control station that can be a PC.) Moreover, IP technology allows users to interact with the system in real time, thus reducing operating and on-site maintenance costs.

Because IP infrastructures are developing rapidly, and the available bandwidth increases daily, it's important that systems can be adapted to use the bit rates the network supports.

The new technologies enable you to implement advanced digital features in your video security system: video indexing (dating and referencing), audio and video alarm management, (local) image and video storage, privacy zone management, and automated object recognition and tracking, while providing flexibility. The central elements are the processor and the video compression standard.

THE MULTISTANDARD CAPABILITY OF DSP

The segmentation of this wide market makes it important to choose a powerful and scalable solution. When scalability is key, a DSP enables you to design an appropriate architecture for each application. Unlike an ASIC–based system, the programmable solution offered by a DSP provides flexibility. Because the video and audio processing are done in software, the product can evolve quickly as the compression technologies advance: you can easily change the video and audio codecs. You can also update and maintain the application software remotely.

In addition, a DSP gives you the cost sensitivity of a hardware solution, and your system cost goes down as the cost of the silicon decreases. Also, when you choose a DSP based on Texas Instruments' TMS320C64x™ generation DSP, the

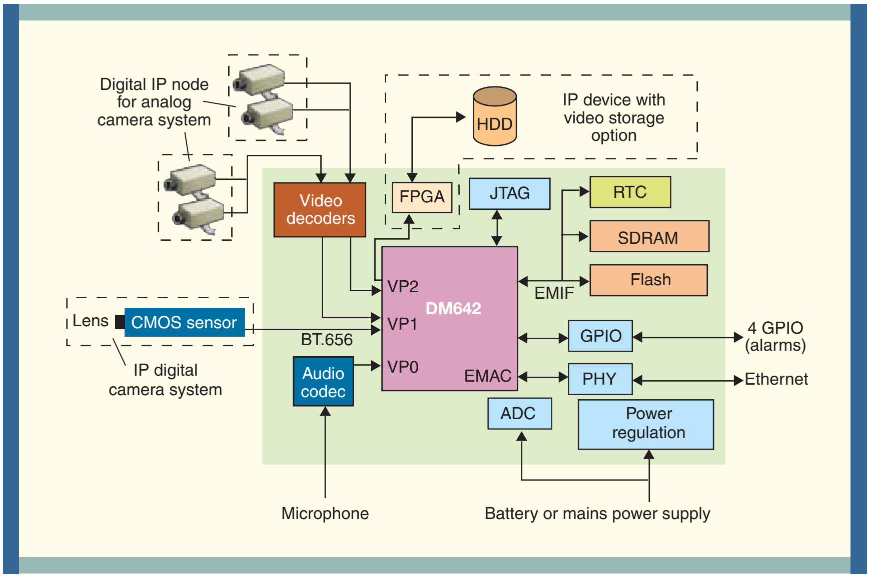

Figure 1: *A hardware platform based on the TMS320DM642 can be used to build an IP digital camera, a digital IP node for an analog multiple-camera system, and an IP device for video storage, thanks to the processor's three integrated video ports and its Ethernet MAC.*

algorithms available have benefited from years of experience being optimized for quality versus performance for the core. You also get comprehensive tools, easing development. Finally, application-oriented reference designs are available, giving you a fast track to market.

Indeed, recent progress in DSP technology, like digital media processors with such integrated features as video ports and an Ethernet interface are making new digital video security systems possible. Consider the example shown in Figure 1. This hardware platform, based on Texas Instruments' TMS-320DM642™ digital media processor can be used to build an IP digital camera, a digital IP node for analog multiple-camera systems, or an IP device for video storage. The DM642 processor is an excellent choice because of its three video ports and its Ethernet MAC. The video ports provide a glueless interface to common video decoders and encoders, support multiple resolutions and video standards, and work either in video capture or display mode.

VIDEO COMPRESSION STANDARDS

The most commonly used video compression standards are JPEG and Motion JPEG (MJPEG), MPEG-2, MPEG-4, and H.264 (see the table). JPEG and MJPEG, however, offer poor compression performance compared with the other standards. H.264 allows more bandwidth reduction, but it's too expensive for video security because of the CPU load needed to encode the video streams. Motion JPEG and MPEG-4, on the other hand, are good candidates for specific video security features.

JPEG and MJPEG are appropriate for court usage because they're based on motion estimation. For that reason, they don't impose a heavy CPU load. In addition, frames are compressed independently, thereby allowing error resilience, robustness, low latency, and easy implementation of fast-forward and fast-backward. However, because of their poor compression performance, JPEG and MJPEG aren't suited to video security applications that need high resolution and high-frame-rate streaming and storage. In those cases, MPEG-4 is better adapted.

MPEG-4 enables a low bit rate and a high-quality image. In comparison, MPEG-2 would require almost twice the bandwidth for an equivalent image quality. In addition, MPEG-4 is widely used, thus ensuring system interoperability. The feature richness of MPEG-4 technology also allows a flexible solution: the stream structure, resolution, and bit rate regulation mode can be adjusted by the application software.

MPEG-4

MPEG-4 is an ISO/IEC standard developed by the Moving Picture Experts Group (www.m4if.org). It consist of six parts: System, Visual, Audio, Conformance, Reference Software, and Delivery Multimedia Integration Framework (DMIF)

MPEG-4 Video technology covers a wide range of applications. The low bit rate and error-resilient coding allow for robust communication over limited-rate wireless channels, which is useful for cell phones, videophones, and space communications. MPEG-4 is also suited to video surveillance data compression, since it's possible to have a very low or a variable frame rate.

In addition to the six parts, many profiles have been defined. These profiles give content-specific tools for visual processing.

The Simple Profile provides efficient, error-resilient coding of rectangular video objects, making it suitable for mobile network applications. The Advanced Simple Profile provides advanced error-resilient coding techniques for rectangular video objects using a back channel and improved temporal resolution stability with low buffering delay. It is suitable for real-time coding applications—for example, videophones, teleconferencing, and remote observation.

Like Motion JPEG, MPEG-4 coding consists in the successive coding of I (intra) frames—also called key frames—with each frame coded independently. And like JPEG coding, I frame coding is based on texture coding.

The Simple Profile also uses P (predicted) frames. A P frame is coded by taking into account the contents of the previous frame (motion estimation data). This method significantly reductes the amount of data, since only the differences between two consecutives frames are coded.

The MPEG-4 Advanced Simple Profile (ASP) uses I, P, and B (bidirectional) frames (Figure 2). "Bidirectional" means that the encoding takes into account motion estimation data from the previous and the following frame.

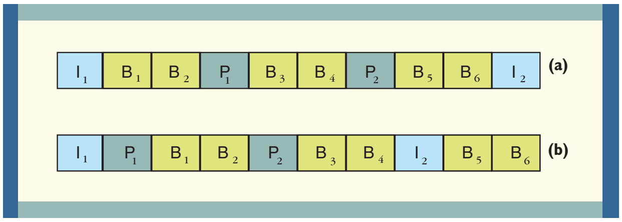

Figure 2: *In the Advanced Simple Profile encoding process, each I frame is encoded independently, each P frame is encoded taking into account the preceding I or P frame, and each B frame is encoded taking into account the preceding and the following I or P frame. Hence for the data stream shown (a), P1 is encoded in terms of the information in I1, P2 is encoded in terms of P1, B1 and B2 need the information in I1 and P1 to be processed, and so on (b).*

THE IMPACT OF THE STREAM STRUCTURE

In the MPEG-4 Advanced Simple Profile the periodicity of I frames is specified by the Group of Picture (GOP). A long GOP decreases the bit rate because encoded I frames are larger than P and B frames. And because the ASP includes B frames, it allows a lower bit rate than the Simple Profile—roughly 15 to 20 percent, depending on the video complexity. However, because a B frame is encoded from the preceding and the following I or P frame, the use of B frames increases latency. On the other hand, because I frames are independently encoded, decoding an I frame allows video restoration from any bit stream error.

In addition, a short GOP yields a short restoration period, enabling the video system to change channels quickly. Thus choosing a GOP value involves a compromise among bit rate, robustness, and channel-hopping concerns.

It's important to choose the appropriate resolution for the targeted application.

CIF resolution (352x240 for NTSC, 352x288 for PAL) doesn't require a great deal of processing power. Therefore it can be useful for multichannel applications. It also allows low bit rates, thereby preserving storage space and transmission bandwidth. It's well suited to low-cost applications with mosaic display, for example.

Full D-1 resolution (NTSC: 720x480, PAL: 720x576) is more expensive in terms of processing power and bit rate, but it can meet customers' expectations in terms of quality. Thus it's an appropriate choice for high-end applications, where detail is needed.

An intermediary resolution, 2CIF (NTSC: 704x 240, PAL: 704x288) lowers the system cost compared with full D-1 resolution but still provides a good visual aspect: users who want full D-1 for visual comfort can zoom out for the higher resolution.

THE IMPACT OF THE BIT RATE REGULATION MODE

Three bit rate regulation modes can be implemented in MPEG-4 video compression: variable, constant, and average bit rate.

With variable bit rate (VBR), the bit rate varies over time and the quantization ratio is constant. This mode is well suited to video security applications, since the bit rate is minimized when nothing happens in the acquired video sequence, thus reducing storage space, and

When streaming compressed video, if the network can support variations in the bit rate, the MPEG-4 encoder can be configured for variable-bit-rate operation.

the quality remains good when movement occurs. With constant bit rate (CBR), the bit rate is constant over time, but the quantization ratio varies, so changes occur in the quality along the compressed stream. And with average bit rate (ABR), the size of the bit stream varies; as a result, the bit rate and the quantization ratio aren't constant. The quantization ratio varies with the video contents so as to have a predictable encoded stream size while maintaining good quality.

When streaming compressed video, if the network can support variations in the bit rate, the MPEG-4 encoder can be configured for variable-bit-rate operation. If not, the constant-bit-rate mode can be used.

Additionally, two types of tool have been developed to improve video quality and lighten the encoding processing load: pre- and post-processing filters.

FILTERS

Video preprocessing filters are used to reduce noise. Noise in the video source means more (unwanted) information in the image to be encoded. As a result, the bit rate is higher, more processing power is needed for the encoder, the quality of the display is lower, and postprocessing is necessary on the decoder side. Of course, preprocessing requires processing power, so different levels of filtering exist that can be activated or not independently. The main preprocessing filters are antidust and low-pass.

On the decoding side, postprocessing filters aim to reduce artifacts due to the compression algorithm. The two used most are deblocking and de-ringing filters. A deblocking filter smoothes the block frontiers, allowing a smart display with much more compressed bit streams, but it requires processing power on the decoder side. A de-ringing filter also requires processing power on the decoder side, but it removes undulations due to compression.

AUDIO

Audio can be very useful in a security system, since sounds can provide clues. For example, an audio edge detection feature can be added to a video security system so that video recording starts when a sound

Comparing Video Compression Standards

	JPEG/Motion JPEG	MPEG-2	MPEG-4	H.264
Bit rate for a given quality (typical bit rate for D-1)	- (15–20 Mb/s)	+ (5–6 Mb/s)	++ (3–4 Mb/s)	+++ (2–3 Mb/s)
CPU load[1]	100%	200%	200%	500%–800%
Fast-forward, fast-backward	++	-	-	-
Camera remote control (latency)	++	VBR: + CBR: -	VBR: + CBR: -	VBR: + CBR: -
Error resilience	++	-	+	+
Court usage (as evidence)	Yes	No	No	No
Channel-hopping delay	++	+/-[2]	+/-[2]	+/-[2]

Note: *VBR = variable bit rate, the mode that allows a constant quality factor. CBR = constant bit rate, the mode that allows a compressed-stream constant bit rate.*

***1.** The CPU load values are based on the Motion JPEG reference and are relative.*
***2.** Depends on the GOP size.*

alarm is detected.

With audio, the aim is to get a sufficient quality to preserve surrounding sounds while minimizing the CPU load on the DSP. Adaptive differential pulse code modulation (ADPCM) is the best candidate for the audio compression algorithm, as it allows for this compromise. Less than 1 percent of the DSP's resources is required for CPU tasks for acquisition, compression, and audio streaming. ◆

Marc Guillaumet is the marketing director at ATEME in Paris. He joined the company in 1992 as employee number 10. After seven years managing technical staff and development projects, he created and developed ATEME's Products business unit. He has a leading role in defining the product development roadmap and heads the technical and marketing aspects of the unit.

Tips for Your Development

How do you get started? What traps should you be aware of? How can you dramatically reduce your time to market?

A traditional design cycle comprises three stages: architecture definition, prototype realization, and production.

In the first stage, architecture definition, you must choose the compression algorithm and the target processor. You must also evaluate the application performance on the chosen target to design the right system architecture.

When developing the application prototype, you must design a hardware platform, develop the DSP application software, and integrate and validate it with the host software. Then you can validate the solution you've chosen using functional and performance tests.

The way to reduce those two steps, saving time and money, is to use an open application development platform, like VSIP, a complete reference design kit, or application development platform, that includes a board based on the TMS320DM642™ digital media processor. Furthermore, if you're moving from another technology to DSP technology, a complete application development platform substantially reduces the risk.

When the platform meets the fundamental requirements of the final product, architectural definition becomes unnecessary, since the compression algorithms (MPEG-4, ADPCM), the processing libraries (motion detection, preprocessing, audio event detection, and the like), the streaming libraries, and the the target processor have already been chosen and the architectural design already done. In addition, end-product demos are available for evaluation purposes, using PC software to control and display the DSP application "results."

As for prototype realization, the initial hardware design becomes unnecessary because the development kit provides the operational hardware platform. In addition, all the other tools needed for development—DSP and FPGA development tools and evaluation versions of the video and audio compression algorithms—are included in the kit, and the manufacturer provides technical software and hardware support.

An open application development platform also smoothes the production stage by providing reference design schematics and software that is easy to port.

IMPORTANT NOTICE