

By  
**Zhengting He,**  
Software Application Engineer,  
Texas Instruments

## **Executive Summary**

The desire for increased security has catapulted the popularity of video surveillance systems, which are now widely deployed in places such as airports, banks, public transportation centers and even private homes. The many problems associated with traditional analog-based systems are influencing the push for digital-based systems. Furthermore, with the growing popularity of computer network, semiconductor and video compression technologies, next-generation video surveillance systems will undoubtedly be digital and based on standard technologies and IP networking.

This paper discusses how to overcome the constraints of a limited network bandwidth with motion estimation (ME) and motion JPEG techniques. In addition, the paper will explore how DSP exercises video codecs for better performance.

# **Video Compression and Data Flow for Video Surveillance**

In a video surveillance system over Internet Protocol (VSIP), hardware handling the network traffic is an integral part of the camera. This is because video signals are digitalized by the camera and compressed before being transmitted to the video server to overcome the bandwidth limitation of the network. A heterogeneous processor architecture, such as digital signal processor (DSP)/general purpose processor (GPP), is desirable to achieve maximum system performance. Interrupt-intensive tasks such as video capturing, storing and streaming can be partitioned to GPP while the MIPS-intensive video compression is implemented on the DSP. After the data is transferred to the video server, the server then stores the compressed video streams as files to a hard disk, overcoming traditional quality degradation associated with analog storage devices. Various standards have been developed for compression of digital video signals, which can be classified into two categories:

- **Motion estimation (ME)-based approaches** – N frames are defined as a group of pictures (GOP) in which the first frame is encoded independently. For the other (N–1) frames, only the difference between the previously encoded frame(s) (reference frame(s)) and itself is encoded. Typical standards are MPEG-2, MPEG-4, H.263 and H.264.
- **Still image compression** – Each video frame is encoded independently as a still image. The most well-known standard is the JPEG. MJPEG standard encodes each frame using the JPEG algorithm.

## **ME Versus Still Image Compression**

Figure 1 on the following page shows the block diagram of an H.264 encoder. Similar to other ME-based video coding standards, it processes each frame macroblock (MB) by macroblock, which is  $16 \times 16$  pixels. It has a forward path and reconstruction path. The forward path encodes a frame into bits. The reconstruction path generates

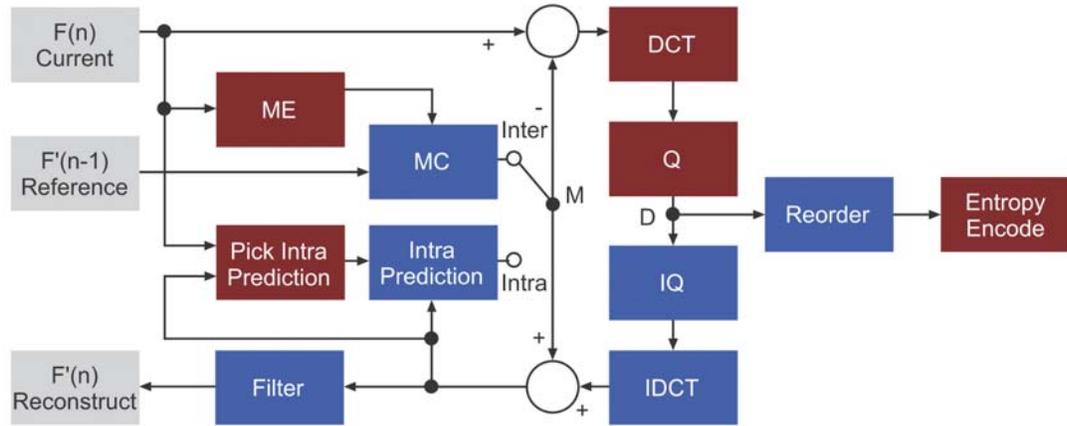


Figure 1. H.264 encoder block diagram.

a reference frame from the encoded bits. Here (I)DCT stands for (inverse) discrete cosine transform and (I)Q stands for (inverse) quantization. ME and MC stand for motion estimation and motion compensation, respectively.

In the forward path (DCT to Q), each MB can either be encoded in intra mode or inter mode. In inter mode, the reference MB is found in previously encoded frame(s) by the motion estimation (ME) module. In intra mode, M is formed from samples in the current frame.

The purpose of the reconstruction path (IQ to IDCT) is to ensure that the encoder and decoder will use the identical reference frame to create the image. Otherwise the error between the encoder and decoder will accumulate.

Figure 2 is a JPEG encoder block diagram. It divides the input image into multiple  $8 \times 8$  pixel blocks and processes them one by one. Each block passes through the DCT module first. Then the quantizer rounds off the DCT coefficients according to the quantization matrix. The encoding quality and compression ratio is adjustable depending on the quantization step. The output from the quantizer is encoded by the entropy encoder to generate the JPEG image.

Since sequential video frames often contain a lot of correlated information, ME-based approaches can achieve a higher compression ratio. For example, for NTSC standard resolution at 30 f/s, the H.264 encoder can encode video at two MBPS to achieve average image quality with a compression ratio of 60:1. To achieve similar quality, MJPEG's compression ratio is about 10:1 to 15:1.

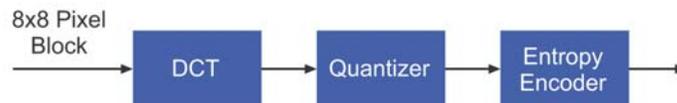


Figure 2. JPEG encoder block diagram.

MJPEG has several advantages over the ME-based approach. Foremost, JPEG requires significantly less computation and power consumption. Also, most PCs have the software to decode and display JPEG images. MJPEG is also more effective when a single image or a few images record a specific event, such as a person walking across a door entrance. If the network bandwidth cannot be guaranteed, MJPEG is preferred since the loss or delay of one frame will not affect other frames. With the ME-based method, the delay/loss of one frame will cause the delay/loss of the entire GOP since the next frame will not be decoded until the previous reference frame is available.

Since many VSIP cameras have multiple video encoders, users can select to run the most appropriate one based on the specific application requirement. Some cameras even have the ability to execute multiple codecs simultaneously with various combinations. MJPEG is typically considered to be the minimal requirement, and almost all VSIP cameras have a JPEG encoder installed.

## **Motion JPEG Implementation**

In typical digital surveillance system, video is captured from sensor, compressed and then streamed to the video server. The video encoder task that is implemented on modern DSP architecture should not be interrupted since each context switch may involve large numbers of register saving and cache throwing. Thus the heterogeneous architecture is ideal so video capturing and streaming tasks can be offloaded from the DSP. The block diagram in Figure 3 illustrates an example DSP/GPP processor architecture used in video surveillance applications.

When implementing Motion JPEG on a DSP/GPP system-on-chip (SoC)-based system, developers should first partition the function modules appropriately to achieve better system performance.

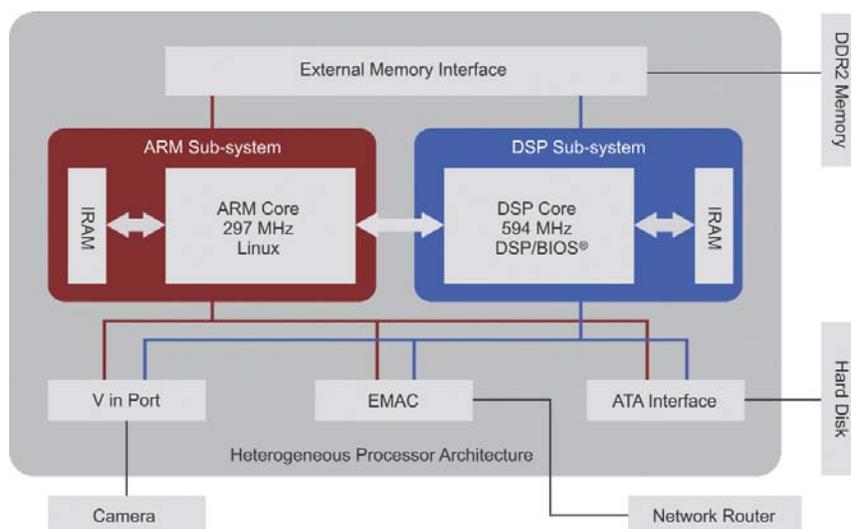


Figure 3: Example of DSP/GPP architecture used in video surveillance applications.

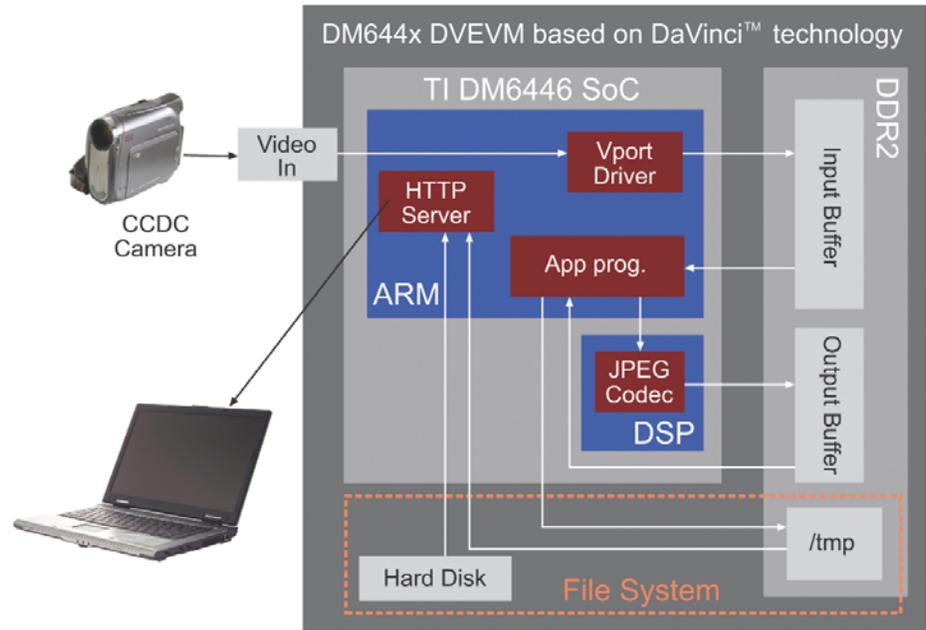


Figure 4: Data flow of Motion JPEG demo on TI's TMS320DM6446 Digital Video Evaluation Module based on DaVinci™ technology used in video surveillance systems.

The EMAC driver, TCP/IP network stack and HTTP server, which work together to stream compressed images to the outside, and the video capture driver and ATA driver should all be implemented on the ARM® to help offload DSP processing. The JPEG encoder should be implemented on the DSP core since its VLIW architecture is particularly good at processing this type of computation-intensive task.

Once the video frames are captured from the camera via the video input port on the processor, the raw image is compressed by exercising the JPEG encoder and then the compressed JPEG image files are saved to the hard disk on the board.

Typically PCs are used to monitor a video scene in real time, by retrieving the streams in the video server and decoding and displaying them on the monitor. Encoded JPEG image files can be retrieved on the board via the Internet. Multiple streams can be monitored in a single PC. The streams can be also watched simultaneously from multiple points in the network. As a huge benefit over traditional analog systems, the VSIP central office can contact the video server through the TCP/IP network and can be physically located anywhere in the network. The single point of failure becomes the digital camera, not the central office. The quality of the JPEG images can also be dynamically configured to meet varying video quality specifications.

## **Optimizing JPEG Encoder**

Out of the three main function modules in a JPEG encoder, DCT and quantizer are computationally intensive. The performance difference between highly optimized assembly code and un-optimized C code for these two modules can be dramatic. Thus, optimizing these two modules is necessary.

Optimizing the two-dimensional (2D)  $8 \times 8$  DCT function reduces the number of additions/subtractions and multiplication by removing the redundant computations in the original equation. Many fast DCT algorithms have been published among which Chen's algorithm is widely accepted by the industry. For 2D  $8 \times 8$  DCT, Chen's algorithm requires 448 additions/subtractions and 224 multiplications.

These additions/subtractions and multiplications can further be partitioned to multiple function units in the DSP core to achieve parallel instruction execution, which achieves better performance. Highly optimized DSP assembly code can finish a 2D DCT within 100 cycles, excluding the overhead. Other fast DCT algorithms require even less amount of computations. However, they often require more buffer to save intermediate computation results. For modern DSP with pipelined VLIW architecture, loading/storing data from/to the memory takes more cycles than a multiplication. Thus, it is important for developers to consider the idea of balancing computations and memory accessing when optimizing the algorithm.

Quantizing each pixel requires a multiplication and an addition. The computation typically only requires 16-bit precision while the size of DSP registers is 32 bits. The first way to optimize the quantizer module is to pack two pixels into one register and perform additions and multiplications on a pair of pixels. The second option is to also use multiple DSP function units in parallel. Since the DSP core in TMS320DM6446 SoC has two multipliers and two adders, up to four pixels can be quantized simultaneously. The last but not least alternative is to take advantage of the pipelined DSP architecture. When the DSP core is quantizing the current four pixels, the next four pixels can be loaded from memory so that data can be fed to the multipliers and adders every cycle. The first two options usually have to be realized by developers themselves writing optimized C code or assembly code. Pipelining the code can rely on the DSP compiler.

Other than optimizing each function module, a PING-PONG buffering scheme needs to be deployed to optimize the JPEG encoder at system level. The DSP core accesses data residing in internal RAM (IRAM) at much faster speed compared to accessing data in external DDR2 memory. However, the precious IRAM has very limited size and is not large enough to fit the whole input frame. Thus, a portion of the blocks are processed at a time

in IRAM. When the PING(PONG) set of blocks are being processed, the PONG(PING) set of blocks are transferred by DMA from DDR2 to IRAM so that the DSP core can start processing the next set immediately after completing the current set.

It is clear that the move to digital video surveillance systems is well on its way. Understanding video compression, system partitioning and codec optimization are key to developing next-generation video surveillance systems to meet the escalating demand.

Important Notice: The products and services of Texas Instruments Incorporated and its subsidiaries described herein are sold subject to TI's standard terms and conditions of sale. Customers are advised to obtain the most current and complete information about TI products and services before placing orders. TI assumes no liability for applications assistance, customer's applications or product designs, software performance, or infringement of patents. The publication of information regarding any other company's products or services does not constitute TI's approval, warranty or endorsement thereof.

DaVinci is a trademark of Texas Instruments. All trademarks are the property of their respective owners.

## IMPORTANT NOTICE

Texas Instruments Incorporated and its subsidiaries (TI) reserve the right to make corrections, modifications, enhancements, improvements, and other changes to its products and services at any time and to discontinue any product or service without notice. Customers should obtain the latest relevant information before placing orders and should verify that such information is current and complete. All products are sold subject to TI's terms and conditions of sale supplied at the time of order acknowledgment.

TI warrants performance of its hardware products to the specifications applicable at the time of sale in accordance with TI's standard warranty. Testing and other quality control techniques are used to the extent TI deems necessary to support this warranty. Except where mandated by government requirements, testing of all parameters of each product is not necessarily performed.

TI assumes no liability for applications assistance or customer product design. Customers are responsible for their products and applications using TI components. To minimize the risks associated with customer products and applications, customers should provide adequate design and operating safeguards.

TI does not warrant or represent that any license, either express or implied, is granted under any TI patent right, copyright, mask work right, or other TI intellectual property right relating to any combination, machine, or process in which TI products or services are used. Information published by TI regarding third-party products or services does not constitute a license from TI to use such products or services or a warranty or endorsement thereof. Use of such information may require a license from a third party under the patents or other intellectual property of the third party, or a license from TI under the patents or other intellectual property of TI.

Reproduction of TI information in TI data books or data sheets is permissible only if reproduction is without alteration and is accompanied by all associated warranties, conditions, limitations, and notices. Reproduction of this information with alteration is an unfair and deceptive business practice. TI is not responsible or liable for such altered documentation. Information of third parties may be subject to additional restrictions.

Resale of TI products or services with statements different from or beyond the parameters stated by TI for that product or service voids all express and any implied warranties for the associated TI product or service and is an unfair and deceptive business practice. TI is not responsible or liable for any such statements.

TI products are not authorized for use in safety-critical applications (such as life support) where a failure of the TI product would reasonably be expected to cause severe personal injury or death, unless officers of the parties have executed an agreement specifically governing such use. Buyers represent that they have all necessary expertise in the safety and regulatory ramifications of their applications, and acknowledge and agree that they are solely responsible for all legal, regulatory and safety-related requirements concerning their products and any use of TI products in such safety-critical applications, notwithstanding any applications-related information or support that may be provided by TI. Further, Buyers must fully indemnify TI and its representatives against any damages arising out of the use of TI products in such safety-critical applications.

TI products are neither designed nor intended for use in military/aerospace applications or environments unless the TI products are specifically designated by TI as military-grade or "enhanced plastic." Only products designated by TI as military-grade meet military specifications. Buyers acknowledge and agree that any such use of TI products which TI has not designated as military-grade is solely at the Buyer's risk, and that they are solely responsible for compliance with all legal and regulatory requirements in connection with such use.

TI products are neither designed nor intended for use in automotive applications or environments unless the specific TI products are designated by TI as compliant with ISO/TS 16949 requirements. Buyers acknowledge and agree that, if they use any non-designated products in automotive applications, TI will not be responsible for any failure to meet such requirements.

Following are URLs where you can obtain information on other Texas Instruments products and application solutions:

<b>Products</b>		<b>Applications</b>	
Amplifiers	<a href="http://amplifier.ti.com">amplifier.ti.com</a>	Audio	<a href="http://www.ti.com/audio">www.ti.com/audio</a>
Data Converters	<a href="http://dataconverter.ti.com">dataconverter.ti.com</a>	Automotive	<a href="http://www.ti.com/automotive">www.ti.com/automotive</a>
DSP	<a href="http://dsp.ti.com">dsp.ti.com</a>	Broadband	<a href="http://www.ti.com/broadband">www.ti.com/broadband</a>
Interface	<a href="http://interface.ti.com">interface.ti.com</a>	Digital Control	<a href="http://www.ti.com/digitalcontrol">www.ti.com/digitalcontrol</a>
Logic	<a href="http://logic.ti.com">logic.ti.com</a>	Military	<a href="http://www.ti.com/military">www.ti.com/military</a>
Power Mgmt	<a href="http://power.ti.com">power.ti.com</a>	Optical Networking	<a href="http://www.ti.com/opticalnetwork">www.ti.com/opticalnetwork</a>
Microcontrollers	<a href="http://microcontroller.ti.com">microcontroller.ti.com</a>	Security	<a href="http://www.ti.com/security">www.ti.com/security</a>
RFID	<a href="http://www.ti-rfid.com">www.ti-rfid.com</a>	Telephony	<a href="http://www.ti.com/telephony">www.ti.com/telephony</a>
Low Power Wireless	<a href="http://www.ti.com/lpw">www.ti.com/lpw</a>	Video & Imaging	<a href="http://www.ti.com/video">www.ti.com/video</a>
		Wireless	<a href="http://www.ti.com/wireless">www.ti.com/wireless</a>

Mailing Address: Texas Instruments, Post Office Box 655303, Dallas, Texas 75265  
Copyright © 2007, Texas Instruments Incorporated