目录视图   摘要视图   RSS 订阅

文章搜索

文章分类

文章存档

阅读排行

Next-generation surveillance system design

分类: video   2008-12-04 11:12   186人阅读   评论(0)   收藏   举报

system   video   encoding   styles   compression   processing

Next-generation surveillance ... gn

Future surveillance systems require high-definition resolution, intelligent encoding with new codecs like H.264 SVC, and the use of video analytics. Here's how to meet these requirements with software configurable processors.

By Mark Oliver, Stretch

The past few years have seen the rise of video surveillance and its widespread adoption throughout the world. This adoption has been driven by a transition from analog to digital systems. This transition is marked by decreasing camera and recording equipment costs, advances in digital sensor and compression technology, and improved IP network infrastructure that reduce the cost of transporting video data over large distances.

While this "cost reduction" approach has been successful in driving rapid growth in the industry, digital surveillance systems have been little more than replacements for their analog counterparts. The true potential of the leading-edge surveillance technology has not yet been realized. This article discusses the technical requirements of future video surveillance systems and the hardware and software changes that are taking place to meet these needs.

Video codecs in a surveillance system have requirements that are very different from those seen in other video applications. For example, latency is relatively unimportant in a broadcast video environment, but it is extremely important in surveillance systems. Surveillance systems often include monitoring personnel who must respond to events in real time. For example, monitoring personnel might track individuals with a Pan Tilt Zoom (PTZ) camera, or they might use an audio link to direct the actions of someone at the scene.

As another example, broadcasters have very little ability to change encoding schemes because any changes they make must be compatible with the receivers already in the field. In contrast, surveillance systems tend to be closed systems. As a result, the operator is free to select standards that best suit his or her needs including frame rate, resolution, and the codecs. Operators may even adjust resolution and frame rate dynamically. This may be

推荐文章

done in response to changes in an observed scene or the capabilities of the consuming device. For example, the system might encode several versions of the same video stream. On stream might go to a high-definition monitor in an observation room, while another goes to a PDA carried by on-site security personnel.

Of codecs available to surveillance operators, most are leveraged from other industries, and do not satisfy the requirements of surveillance applications. Table 1 summarizes the high-level features of various common codecs.

	Typical D1 Bandwidth Requirements	Compute Requirements	Advantages	Disadvantages
MJPEG	30Mb/s	Low	Low complexity Low latency Random access	Low compression High bandwidth
MPEG4 Part 2 Simple Profile	3Mb/s	Medium	High compression Low latency	Complex
H.264 Baseline Profile	1.5Mb/s	High	Very high compression Low Latency	High complexity
H.264 Main Profile	<1.5Mb/s	Very High	Very high compression	Very high complexity High latency

Table 1. High Level codec Comparison.

None of the codecs in Table 1 satisfy all the desired characteristics of surveillance systems: low latency, high compression efficiency, resolution and frame rate flexibility, low complexity, and low cost. H.264 Baseline Profile probably offers the best compromise, but lacks the inherent scalability needed in surveillance applications.

Scalable Video Coding

The Scalable Video codec (SVC) is an extension of the current H.264 standard. SVC was developed with a view to using a single encoded stream to satisfy diverse requirements in terms of bit rate, quality and resolution. SVC supports a high degree of scalability. It scales spatially, allowing for varying display resolutions. It scales temporally, allowing for varying frame rates. And it scales in quality, allowing for varying resulting image quality. For example, an H.264 SVC video stream can be decoded by two different devices with different frame rates and resolutions. In conventional video encoding, if the video stream were to be viewed at a reduced resolution on a portable device, the entire stream would have to be decoded and resized. With SVC, only the portion of the stream yielding the desired resolution and frame rate is decoded.

An SVC decoder's flexibility in how it deals with an SVC bitstream results in many benefits to the user. These benefits include ease of adaptation for different displays; resource-conserving video transmission, storage and display; higher transmission robustness; and ease of heterogeneous network support (for example, simultaneously supporting a number of different transmission networks). An additional benefit of SVC is that the compressed stream can be parsed while stored on a disk. The portions of the files that are used to reconstruct high frame rate or high quality images can be progressively removed over time. This is not possible with conventional codecs where the video data is either

there or it isn't, and one has to select a date upon which the original resolution file will be totally deleted. With an SVC system, the video can be kept for longer periods as storage requirements are gradually reduced.

Video Requirements

High-definition (HD) video is becoming increasingly popular in video surveillance. The increased clarity provided by HD resolution allows for better recognition of, for example, individuals within the scene. A less obvious benefit is the use of electrical pan, tilt, zoom (ePTZ). With ePTZ, a wide angle lens in conjunction with a high definition sensor provides a wide field of view. By moving a region of interest window around within this field of view, a conventional, mechanical PTZ system can be simulated without any moving parts. In extreme cases, "fish eye" lenses can be used in conjunction with digital distortion correction to provide very large ePTZ ranges.

Aside from resolution, another important factor is whether the video is progressively scanned or interlaced. In progressively scanned video, all the horizontal lines of an image are displayed. In Interlaced video, only half of the horizontal lines are displayed at any time (the first field with the odd-numbered lines, followed by the second field with the even-numbered lines). The advantage of interlaced video is that you can double the image refresh rate without increasing bandwidth. The disadvantage is that the vertical resolution for each field (two of which make up a frame), is cut in half. Table 2 shows resolutions and bandwidths for commonly used standard and high definition broadcast video.

Video Source	Resolution (pixels)	Refresh Rate (Hz)	Pixel Clock (MHz)	Bandwidth Req'ts (Mbsp)
480p	720x480 (SD)	60	27	331
720p	1280x720	60	74	884
1080i	1920x1080 (HD)	60	74	995
1080p	1920x1080 (HD)	60	149	1990

Table 2. Bandwidth requirements for common video specifications.

As Table 2 shows, the bandwidth requirements for uncompressed 1080p HD video are very large, exceeding what even 1 Gbit Ethernet can support. Clearly, the compression efficiency of a video codec is an important attribute for surveillance systems.

Video Analytics

Also referred to as intelligent video, video analytics refers to algorithms that detect and track objects of interest to look for possible threats or safety breaches. For example, video analytics might look for a person entering an unauthorized area, or someone leaving a package unattended in an airport lounge for more than a

preprogrammed period of time. Typically, such events, or "triggers," cause the video to be sent to a human observer for further investigation.

Video analytics can be implemented either in the camera or on a central server. If analytics are implemented in the camera, the camera can save network bandwidth by transmitting only video of suspicious activity. In some cases, cameras equipped with video analytics are not required to send video data at all. For example, as illustrated in Figure 1, a camera performing optical character recognition (OCR) on license plates could return only a few bytes of information representing the license plate number,

Figure 1. Analytics can be used to extract information from the scene eliminating the need for operator intervention all together. Image courtesy of IntelliVision.

Processing video on a central server also has advantages. For instance, a human observer can employ analytics to search large amounts of recorded video for possible events of interest.

The benefits video analytics provide surveillance operators are many. Since analytics are real time, they can provide security personnel with immediate notification of threats or safety hazards. They improve the quality of surveillance by filtering out uninteresting activity. And perhaps most importantly, analytics reduce the personnel needed to monitor the system. This is particularly important for systems employing very large numbers of cameras, such as those monitoring public transportation systems or large buildings. Furthermore, analytics reduce fatigue for the required human operators, further enhancing surveillance quality.

Intelligent Encoding

Intelligent video encoding combines situational awareness (via video analytics) with flexible encoding. When video analytics algorithms detect motion, scene changes, or other potentially suspicious activity, an intelligent encoder can not only flag the scene to the appropriate security person, it can also adapt its encoding accordingly. For instance, when suspicious activity is detected, the encoder can increase frames per second, resolution, and the quality of the encoding. (See Figure 2).
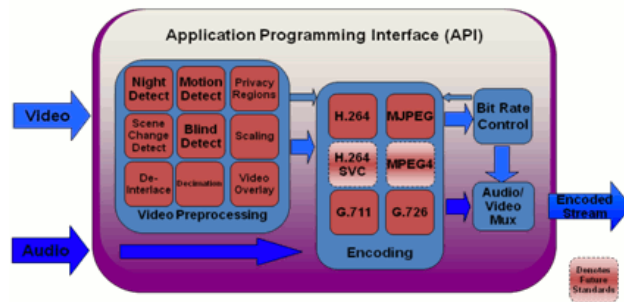
Figure 2. By combining situational awareness with flexible encoding, intelligent video encoding provides decoding flexibility and preserves system bandwidth and storage space.

When scenes contain little or no motion, the encoder can reduce frame rates or resolution, thereby reducing bandwidth and/or storage consumption. This is particularly important for networked systems where bandwidth and storage are most limited.

Intelligent encoding can be further refined using an API that allows the user to determine what coding schemes are used. Additionally, users can also define regions of interest that focus where analytics are applied.

The Advantages of Software Configurable Processors

The processors used in video surveillance are faced with a variety of difficult computational tasks. They must run including increasingly complex tasks in real time such as video analytics algorithms, servicing of high-definition sensors, and network management. Dealing with all of these requires an architecture with a high degree of performance and flexibility. A software configurable processing engine is a good solution for this application.

Stretch's implementation of this type of processor combines a configurable processing engine with a programmable fabric (Instruction Set Extension Fabric, or ISEF). The ISEF is a software configurable compute fabric that enables system designers to extend the processor instruction set and to define new instructions using C/C++ code. These "extension instructions" are then automatically synthesized, placed, and routed into the ISEF. System designers can thus optimize the processor instruction set for specific applications in real time to handle tasks such as video processing, analytics and network management (Figure 3); in other words, the designer can implement portions of a desired algorithm in hardware by using the ISEF, which sits within the processor's data path. With this architecture, the instruction issue logic of the processor, as well as the intelligent compiler, can make full use of these hardware functions and schedule them into the instruction execution flow.
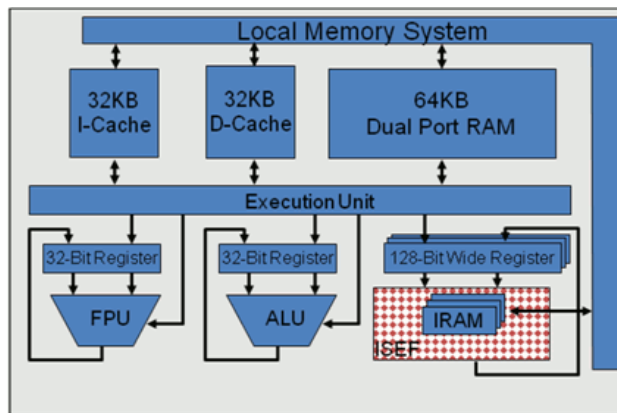
Figure 3. A software configurable processor, combining a processing engine with an Instruction Set Extension Fabric (ISEF) and embedded RAM (IRAM), provides a flexible and powerful processor architecture for video surveillance applications.

This results in a significant improvement in processor performance. The programmable fabric can now execute entire sections of application code in a single instruction. In addition, because the hardware is tightly coupled, the compiler can optimize instruction issues to maximize performance.

At the physical layer, each software configurable processor device can interface with up to four other processors through dedicated 1.2GB/S interfaces. This capability allows system architects to create processor array topologies best suited for their application. To unburden the processor with processor array functions, each processor has a dedicated processor network interface and switch circuitry to accommodate inter-processor communication. At the software layer, programmers can dedicate tasks, establish communication channels between processors, and even share resources between processors.

What's Ahead

The future of video surveillance systems is bright. Emerging video processing and analytics algorithms are raising the capabilities of surveillance systems to ever-higher levels. New scalable codecs are making video streams more readily available for consumption by a multitude of both high definition and hand held devices. And innovative new processor architectures such as software configurable processors are up to the challenge, ensuring that surveillance systems keep pace with ever more stringent security and safety needs in the public and private domain.

About the author

Mark Oliver is the Director of Product Marketing at Stretch. A native of the UK, Oliver gained a degree in Electrical and Electronic Engineering from the University of Leeds. During a ten year tenure with Hewlett Packard, Oliver managed Engineering and Manufacturing functions in HP Divisions both in Europe and the US before heading up Product Marketing and Applications activities at a series of video related startups. Prior to joining Stretch, Oliver managed Marketing for Video and Imaging within the DSP Division of Xilinx.

Related articles:

Discuss This Article

2 message(s). Last at: Apr 1, 2008 1:34:22 PM

jimhoerricks

PSS1

commented on Apr 1, 2008 12:23:28 PM

A very well stated overview of your product with one very important omission - video surveillance systems as generators of evidence that can end up in court.

If the purpose of your system is to monitor an area, without recording - then the system as described would work great. It's when the record button is pressed and a crime is "witnessed" by one of these systems that confusion begins.

It all starts with the designer of the system. What is the system's purpose? Is it observation or monitoring of an area? Is to help in the recognition that some activity is occurring within an area? Is it there to aid in the identification of individuals and objects for internal purposes? With these questions in mind, will these recordings ever be turned over to the police?

Each of the above questions will yield a slightly different system design. You simply will not be able to identify someone/something at a range of 250m with a 4.5mm lens. You will have difficulty identifying someone at 100' at 1CIF. But you can observe activity at these resolutions and distances.

The state where the installation takes place will have it's own evidence code, governing statutes, and case law. These all need to be taken into consideration when designing a system and selection a codec where the data will be used to prosecute offenders. There is an interesting trend in the courts where MPEG4 based video is becoming more of a problem in a prosecutor's case than a help.

Here's a question: Can a "B" or "P" frame be considered a "true and accurate representation of a scene?" Many states, and the Federal Rules of Evidence, have specific guidance as to how this "true and accurate" clause is to be interpreted. If the B and P frames are only representations of the change that the computer predicts between I frames, then how could they be considered "true and accurate" under the rules of evidence - thus summarizing the defense's objection to the evidence as such?

With that in mind, how frequent are I frames generated? 1 second? 5 seconds? It varies by manufacturer and by installation. How reliable are the rates in practice vs. what is published by the manufacturer? In one famous case in Florida, there was such a variance that the prosecutor could not use the video - prompting the dropping of charges and a counter suit for false imprisonment.

Hopefully, some consideration will be given for the fundamental change that occurs when the "record button" is pressed - for the potential problems that exist when multimedia data becomes multimedia evidence.

Jim Hoerricks

Forensic Image Analyst

http://forensicphotosho.blogspot.com

Kenton

Site Editor

commented on Apr 1, 2008 1:34:22 PM

Thanks for the great comment. The analysts I've talked to disagree; they say that technical concerns like these are not a major issue when it comes to forensics. According to the folks I've talked to, a B frame is as good as a P frame is as good as an I frame.

I'd love to hear more of your views on this issue. Feel free to write me at kentonwilliston@yahoo.com.

上一篇　　Tutorial: The H.264 Scalable Video Codec (SVC)

下一篇　　iLBC编解码相关知识

主题推荐　　　　interface　　　broadcast　　　application　　　performance　　　scalability

猜你在找

查看评论

暂无评论

发表评论

用 户 名： oDavid12345678922

评论内容：

提交

* 以上用户言论只代表其个人观点，不代表CSDN网站的观点或立场

核心技术类目

全部主题　　Hadoop　　AWS　　移动游戏　　Java　　Android　　iOS　　Swift　　智能硬件　　Docker　　OpenStack
VPN　　Spark　　ERP　　IE10　　Eclipse　　CRM　　JavaScript　　数据库　　Ubuntu　　NFC　　WAP　　jQuery
BI　　HTML5　　Spring　　Apache　　.NET　　API　　HTML　　SDK　　IIS　　Fedora　　XML　　LBS　　Unity
Splashtop　　UML　　components　　Windows Mobile　　Rails　　QEMU　　KDE　　Cassandra　　CloudStack　　FTC
coremail　　OPhone　　CouchBase　　云计算　　iOS6　　Rackspace　　Web App　　SpringSide　　Maemo
Compuware　　大数据　　aptech　　Perl　　Tornado　　Ruby　　Hibernate　　ThinkPHP　　HBase　　Pure　　Solr
Angular　　Cloud Foundry　　Redis　　Scala　　Django　　Bootstrap