

Record



October 24, 2008

Video analytics architecture and intelligent IP cameras at the edge

Advantages of video analytics at the edge include massive bandwidth and storage cost savings, superior DSP-based processing based on high-megapixel images, and easy expansion.

By Richard McClellan, CEO and
Jerry Wesley, Vice President,
Software Development,
Apollo Imaging Technologies,
Inc.

In the year 2008, cameras are nearly everywhere, and wherever they aren't, they will be soon. While the reasons for installation of these cameras vary widely, they generally fall into security and safety categories. These cameras are additional "eyes" that work 24 hours a day, seven days a week. Human monitoring is being replaced by Video Analytics. Recent advances in video processing Digital Signal Processors have enabled [migration](#) of the image analysis task from central computers to autonomous intelligent cameras.

Where it started

Initially, the dominant security and surveillance technology was based on [analog](#) (NTSC/PAL) cameras. Human beings watched one or more monitors; each connected to a single camera, or switched between multiple cameras, looking for something to happen. The next step was to record the video with banks of video tape recorders. Even though this made off-line video review and archiving possible, the mechanics were unappealing. Large installations required dedicated rooms just to store their videotapes. While the video was indeed archived, review and retrieval was awkward; first find the right cassette, put it into a VCR player, then [search](#) using fast forward and rewind.

A significant advancement was the replacement of the tape-based VCR's by Digital Video Recorders (DVRs). While still connected to analog cameras, they had the ability to store the CCTV video data like any other form of digital data hard drives, digital tapes, etc. Network accessible [storage](#) made video storage, retrieval and review much easier.

Current technology for this type of system is based on Network or IP cameras. The entire video system has now been digitized with the introduction of IP Cameras with integrated web servers--stand-alone devices that allow the user to view live, high resolution, full motion video from anywhere on a [computer](#) network (bandwidth permitting), or over the Internet, using a standard web browser. They can be connected either to an ad-hoc or existing [IP](#) network. Images can be viewed and cameras managed via standard web browsers. Network-based storage resources can be set up to record the video output.

Video is high bandwidth, high volume data. A monochrome (eight [bit](#) per pixel) [VGA](#) video stream, uncompressed at 30 frames per second, requires nine megabytes of storage per second of video. Color, higher [resolution](#) and higher [frame](#) rates rapidly multiply this number. The demands on communications, computer MIPS and storage capacity can be very high. Terabytes of information-sparse data can rapidly be created and archiving all of that video may be necessary for the court case, but doesn't do much for responding to a critical situation or developing problem.

The human element

While the collection, management, and storage of all this video has been greatly improved, the analysis has, until recently, been left to humans. Human monitoring is costly and problematic. Humans do not do a very good job at monitoring for low rate of occurrence events. Most of the time, nothing is happening. Watching a video screen of an empty warehouse for an event, which may, should, never occur, is not a desirable job and one that even the most conscientious of employees will not do well. The worst case is the "blink of an eye" event that rarely, if ever, occurs, as there is an extremely high probability that a human will miss it.

Removing the weak link

These problems with human monitoring, both cost and quality, have resulted in the emerging discipline of "Video Analytics", a.k.a. "Machine Vision", "Computer Vision", or "Intelligent Video Surveillance". Video Analytics is the science of analyzing video (as opposed to a single frame) for events of interest. The "event of interest" may be as simple as motion (something in the scene changed) or as complex as detecting the signature of a smoke plume in a warehouse. On detection of the event, appropriate action can be taken--send the video to an actual human being, launch an email, call a cell phone, sound an alarm, stop a machine, etc.

While VA (Video Analytics) has typically been PC based, the advent of high performance video processing DSP systems have opened up the possibility of camera based Video Analytics, or Video Analytics on the Network Edge.

Next: VA architectures

VA architecture

VA implementation requires digitized video accessible by a suitable image processing resource. There are

three predominant architectures in use:

1. Analog Cameras connected to frame grabbers installed in a PC.
2. IP Cameras connected via Ethernet to a PC.
3. Self contained "Smart Cameras", Video Analytics on the Edge (VAE).

In both the Analog Camera/Frame Grabber and IP Camera systems, the VA intelligence is located in the PC. The sole function of the camera system is to collect and deliver suitable digitized video to the Video Analytics in the PC. All decisions are made in the PC.

The VAE system, on the other hand, is a fundamentally different architecture. The VAE "camera" does both the video collection and the VA processing/event detection. A PC is not directly involved on a real time basis and in principle; a PC may not even be connected. However, in most applications it performs the Human Machine Interface (HMI), system management functions, and event archival.

Table 1 summarizes some of the key strengths and weaknesses of the three predominate VA architectures, as discussed in the next sections.

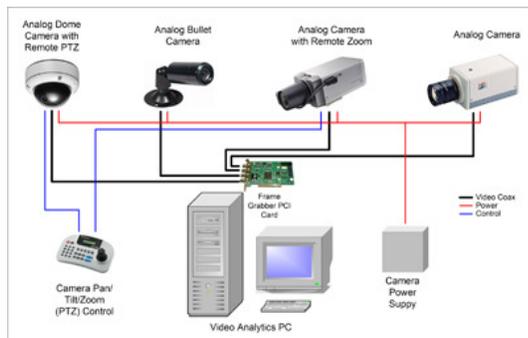
	Analog Camera Network with PC VA	IP Camera Network with PC VA	Video Analytics on the Edge
Cabling (length)	RG-59 (750ft) RG-6 (1000ft) RG-11 (1500ft)	Cat5 (100M directly, greater with inexpensive switches)	Cat5 (100M directly, greater with inexpensive switches)
Separate Power and Control Cables	Yes	No, if POE is used	No, if POE is used
Video Frame Rate	30fps NTSC or 50fps PAL per camera	40 fps uncompressed VGA per network	Sensor Limited (250fps possible)
Network Bandwidth Requirement	N/A	Sufficient for frame rate	Programmable—Only for detected events
Processor Resources for VA	Shared	Shared (plus video decoder load if video is compressed)	Dedicated, in camera
Image Resolution	NTSC/PAL	As required	As required
Processor MIPS Available	PC	PC	2400-9600
Main Memory	PC, easily expanded	PC, easily expanded	16-256MB, not expandable
Rotational Media	Standard PC disk drives	Standard PC disk drives	Not available
Video Format Available to VA	NTSC/PAL Frame Grabber output	Compressed Video (typically)	Raw—VA has direct access to sensor output
Power Consumption	PC	PC	5-10W
Sensor Selection	Wide—CMOS or CCD, but limited to NTSC/PAL	Wide—CMOS or CCD, VGA to MegaPixel	Wide—CMOS or CCD, VGA to MegaPixel including high frame rates
Real Time	No	No	Yes
Development Tools	PC	PC	Embedded
Communication Interfaces	Standard PC	Standard PC	RS242, USB 0, 10/100/1000BaseT
Direct I/O	Requires special I/O PC card	Requires special I/O PC card	Built-In
Size	PC	PC	Compact
Scalability	Available Ports on Frame Grabber card, PC sizing	PC sizing	Easy—each camera is self-contained
Camera Cost	Low, but Frame Grabber port required per camera	Medium	High, but no direct external requirements

[View full size](#)

Table 1: Comparison of video analytics architecture

Analog cameras and frame grabbers

The Analog Camera/Frame Grabber/Computer system shown in Figure 1 has the advantage of mature technology and cost effectiveness. There are a wide variety of analog cameras available, of every conceivable configuration. Frame grabber cards are also a mature technology; while they typically require a slot in the PC, multiple channel units per card are available. While coax cable is required to connect the camera to the frame grabber, it is relatively inexpensive and fairly long lengths (up to 1500ft, depending upon the cable type) can be accommodated without repeaters. Longer cable lengths result in high frequency losses--translating to loss of resolution or focus. The cables between the camera and frame grabber are typically unique and will have to be installed. In addition to the video coax cable, each camera requires its own power wiring. If any camera controls are desired (focus, zoom, aperture, pan/tilt); this will necessitate yet another cable. Potentially, there are three special purpose cables required.



[View full size](#)

Figure 1: Analog cameras with frame grabber in analytics PC

Each camera/cable/frame grabber port is a standalone system--communication bandwidth is not shared. When another camera is added, the frame rate up through the frame grabber to PC interface is not affected. Since all VA processing takes place in the PC, PC processor power and resources are shared. VA can be quite processor intensive, resulting in an aggregate total frame rate from all connected cameras that can be serviced by the PC. If more cameras are added, or additional or a more processor intensive VA are installed,

the PC will either have to be upgraded or the system restructured to add additional PCs. If less than the aggregate maximum frame rate is installed, the PC is oversized.

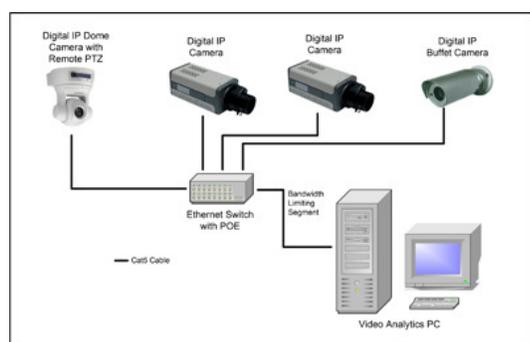
Analog cameras are inherently limited to NTSC/PAL resolutions, which are about 0.4 megapixel, approximately VGA. This is problematic in today's world; [CMOS](#) sensors approaching 10 megapixels are readily available. Higher sensor resolution makes possible either better image quality, or wider fields of view. These capabilities are not easily obtained in the analog camera world.

While Analog cameras alone appear quite cost effective, to the camera cost needs to be added the cost of a frame grabber channel per camera and the special purpose coax cable. Especially if cabling and installation is considered, the analog camera may no longer be cost effective.

Next: IP cameras

IP cameras

The more modern PC centric architecture is based on IP cameras and is shown in *Figure 2*. Connected via Cat5 cable they are no longer constrained to NTSC/PAL resolutions and frame rates; IP security cameras of 5 megapixels are readily available. It can be expected that resolutions will follow consumer cameras with some time lag. While more expensive at the camera level than analog cameras, cabling is less daunting. The building might already have Cat5 cable available to some or all of the camera sites. Power over Ethernet (POE) is becoming near universal in this category of application so that no separate power cabling is required. Camera control, including lens control is done over the network. Instead of the (potentially) three cables of the purely analog camera (coax, power and control), one Cat5 cable, which might already exist, is all that is required.



[View full size](#)

Figure 2: IP cameras with video analytics PC

While CAT 5 cabling is more attractive than the special purpose coax, it comes at a price. [100BaseT](#) is 100Mbps--about 40 fps (theoretical) at VGA resolution. Where multiple camera feeds join to a single network (e.g., the [switch](#) connected to the Ethernet interface on the VA PC), this frame rate is shared among all the cameras. Video [Compression](#) in the camera can significantly increase the bandwidth, and the frame rate, at an increased camera cost and the possibility of spatial or temporal resolution loss. If the video is required to be evidentiary in nature, particularly encoders subject to temporal artifacts (MPEG2 or H.264, for instance) may not be permissible. Additionally, before any VA processing can be done in the PC, the arriving compressed video must be uncompressed, putting further demands on the PC; and since the compression/decompression cycle is lossy, it may hamper the VA from functioning or may even cause false positives.

[Cat5](#) cable is considered to have a direct connection length limit of 100 meters--or about 328 ft. Switches and routers can be used to increase the total distance and to allocate bandwidth, but the network must be planned accordingly. On the other side, being digital, the problem of loss of resolution with [cable](#) length doesn't exist.

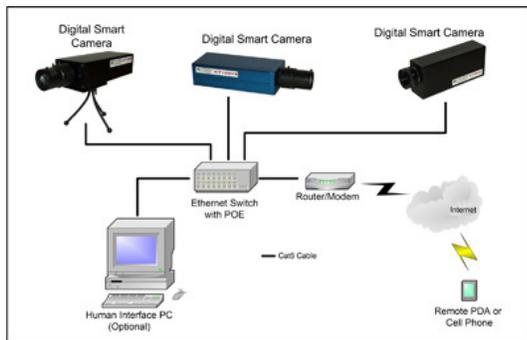
A variation on the IP camera network involves the use of Video Servers, which are boxes that connect to IP Networks and convert the signal from analog cameras to digital format. Video Servers contain onboard processors and web [server software](#) that essentially turns CCTV cameras into IP Network Cameras. As a hybrid, it shares some disadvantage of both--it has the resolution limitation of the Analog camera and the frame rate limitation of the IP network, but it does get the video into a digital, network accessible format. The primary reason for Video Servers is to protect an investment in analog cameras and the installation.

Although the IP camera network is a much cleaner way to collect the video data, the VA still runs on the PC. Processor resource requirements have the same problem as the analog/frame grabber system--the PC will be either oversized or undersized. Video compression in the camera will trade off communication bandwidth to the PC for processing bandwidth in the PC.

Next: The VAE advantage

The VAE advantage

The Video Analytics on the Edge solution is shown in *Figure 3*. This fundamentally different architecture distributes the intelligence to each individual camera with PCs or other network devices optionally serving as HMI or archival devices. All real-time activities are camera based and with all the event detection taking place in the camera, the video sent back over the network is only necessary as it relates to a specific application or installation. As an example, suppose a camera has a motion detection VA installed. The camera could be programmed such that when an event is detected, it also flags the pre and post event video frames located in its video buffer. The result would be an event message sent out over the network with the sequence of images that precede, include, and follow the motion event detected. During normal, non-event conditions, nothing needs to be sent out and therefore no network bandwidth is required. Because all of the "heavy lifting" is done in the camera, the monitoring device that receives the message can even be one with limited resources, such as a cell phone.



[View full size](#)

Figure 3: Video analytics on the edge

At least theoretically, the camera can be autonomous—other than configuration; a PC may not have any function. In practice though, most systems will have a PC on line for archival and management purposes. For instance, in the motion detection example, if archiving is required for system assurance purposes, the camera can be programmed to send time stamped compressed video frames to an archiving server located anywhere on the network. During non-event times, this can be done at low frame rates (once per second) and then change to high frame rates during event detection. This not only saves archival storage space, but network bandwidth. Because the archiving server does not have to process the video, it could even be a simple network hard drive.

One of the major advantages of Video Analytics on the Edge vs. Video Analytics done in a PC is that the VA algorithm has full access to high-resolution, high frame rate, raw uncompressed video as input for its algorithms. After the VA has processed the frame, the camera can then optionally annotate, highlight those areas of interest in each frame, or even reduce the frame size, before compressing the video and sending it out. The camera can be programmed to manage bandwidth and resources based upon the events taking place. It can even change from a temporal type of compression (MPEG2 or H.264) to a lossless static type (JPEG2000) based upon the current status.

Another advantage of VAE systems is system expansion. Because each "camera" is self-contained, adding another can be as simple as plugging it into the network and configuring it. This compares to analog/frame grabber systems where video capture ports must be considered and for VA PC based systems where the total processing power available is critical. Exceeding either limit may necessitate adding another or replacing the VA PC.

It should be noted that in high-speed Video Analytics applications, the VAE architecture might be the only practical one. 250 frames per second VGA CMOS sensors are economical and readily available, but it requires a network bandwidth of over 600 Mbps to transmit this; even if Gigabit Ethernet was installed, most Windows based PC's could not keep up. To be sure, the VAE camera can only do elemental VA at that frame rate, but it is more feasible than sending the high frame rate video data over a network for processing.

A real world example is axonX's (<http://www.axonx.com/>) SigniFire IP camera based upon an Apollo Imaging design. The SigniFire IP operates autonomously utilizing the Texas Instruments TMS320DM642-720 Digital Media Processor running axonX's proprietary software. Detection events such as flames, smoke, motion and fault conditions (no picture, camera out of focus, low-light) are reported to the Digital Video Management (DVM) system over a standard network connection and/or communicated to a Fire Alarm Control Panel through a configurable dry contact relay mounted onto the backplane of the camera. These relay closures are programmable and can even be optionally configured with a delay to allow human video verification before the Panel is notified, thus reducing or eliminating nuisance alarms.

Table 1 summarizes some of the key strengths and weaknesses of the three predominate VA architectures.

	Analog Camera Network with PC VA	IP Camera Network with PC VA	Video Analytics on the Edge
Cabling (length)	RG-59 (750 ft) RG-6 (1000 ft) RG-11 (1500 ft)	Cat5 (100M directly, greater with inexpensive switches)	Cat5 (100M directly, greater with inexpensive switches)
Separate Power and Control Cables	Yes	No, if POE is used	No, if POE is used
Video Frame Rate	30fps NTSC or 50fps PAL per camera	40 fps uncompressed VGA per network	Sensor Limited (250fps possible)
Network Bandwidth Requirement	N/A	Sufficient for frame rate	Programmable—Only for detected events
Processor Resources for VA	Shared	Shared (plus video decoder load if video is compressed)	Dedicated, in camera
Image Resolution	NTSC/PAL	As required	As required
Processor MIPS Available	PC	PC	2400-9600
Main Memory	PC, easily expanded	PC, easily expanded	16-259MB, not expandable
Rotating Media	Standard PC disk drives	Standard PC disk drives	Not available
Video Format Available to VA	NTSC/PAL Frame Grabber output	Compressed Video (typically)	Raw—VA has direct access to sensor output
Power Consumption	PC	PC	5-10W
Sensor Selection	Wide—CMOS or CCD, but limited to NTSC/PAL	Wide—CMOS or CCD, VGA to Megapixel	Wide—CMOS or CCD, VGA to Megapixel including high frame rates
Real Time	No	No	Yes
Development Tools	PC	PC	Embedded
Communication Interfaces	Standard PC	Standard PC	RS242, USB 2.0, 10/100/1000BaseT
Direct I/O	Requires special I/O PC card	Requires special I/O PC card	Built-in
Size	PC	PC	Compact
Scalability	Available Ports on Frame Grabber card, PC sizing	PC sizing	Easy—each camera is self-contained
Camera Cost	Low, but Frame Grabber port required per camera	Medium	High, but no direct external requirements

[View full size](#)

Table 1: Comparison of video analytics [architecture](#) (repeated)

Next: VAE implementation

VAE implementation

The PC based VA have available to them the entire gamut of PC based hardware and software based resources with all the associated merits and demerits. Camera based VAE system are more limited as far as development and hardware/software resources are concerned, but this is changing rapidly with all the embedded development being done today.

Most VAE systems are Digital Signal Processor (DSP) based because DSPs are software programmable, have the processing power required for VA and compression algorithms, directly provide the sensor and network interfaces required, and do all this for a competitive price. Additionally, they are available in small, low power packages that require no special heat dissipation considerations so they can be easily fit into compact "camera" packages. Embedded real-time [operating systems](#) and powerful and full-featured Integrated Development Environments (IDE's) are now available from multiple sources including the DSP manufacturer. Software modules such as video compression encoders and other image processing tools are available either from the manufacturer or third parties. With today's optimizing compilers, image-processing algorithms written in C have proven readily portable to the DSP [platform](#) with little or no assembly language required.

A fully featured open VAE development platform is Apollo [Imaging Technologies](#)' AIT100 (www.apollo-image.com). In a compact 46 cubic inch metal enclosure it features:

- Texas Instruments TMS320DM642 DSP (720 MHz/5760 MIPS)
- 256MBytes SDRAM
- 32MBytes Flash
- Serial EEPROM
- Battery backed Real Time Clock
- RS-232 (RS-422 available)
- USB 2.0 (optional)
- General Purpose I/O; Four Inputs, Four Outputs (Form-C relay and differential I/O available)
- Composite Video [Output](#) (NTSC/PAL)
- Audio Input/Output with full-featured voice CODEC.
- Multiple programmable LED's.
- 10/100BaseT Ethernet
- Xilinx CPLD
- JTAG interface
- Choice of 3.0 megapixel color or B/W CMOS, or 1.3 megapixel smart color CMOS sensor
- C or CS mount (with adapter) Lens mounting
- Optional IR-Cut filter available for true-color video

To speed development of VAE applications, the AIT100 is available with an SDK that includes an IP Camera demo application with a Windows client, a complete Board Support Package (BSP), and a [JPEG](#) encoder. The IP Camera demo application is DSP/BIOS based and requires Texas Instruments' Code Composer Studio and a compatible emulator. The windows [client](#) application is a Microsoft [.Net](#) application and will require Microsoft Visual Studio if any modifications are required.

For production, you can use the hardware package as is, or Apollo Imaging will design a custom version to your specifications.

VAE in summary

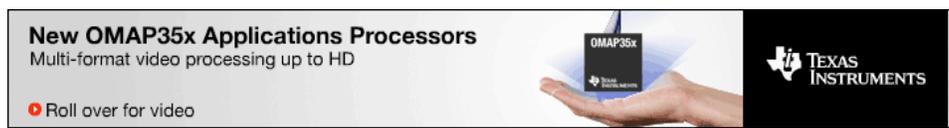
VAE cameras are an emerging technology well suited to inherently distributed safety and security applications. Each camera brings with it the computation, [memory](#) and I/O resources necessary for the Video Analytics. VAE cameras enable responsive, distributed systems that are inherently scalable. High-speed video is more easily handled in the VAE camera. Network traffic can be vastly reduced--only events of significance need be transmitted, and those not in real time. PC requirements and network loading can be greatly reduced, or possibly eliminated.

About the authors

Rich McClellan is CEO of Apollo Imaging Technologies; his duties are best described as Vice President in Charge of Miscellaneous. Prior to founding Apollo with Dan Grantz and Jerry Wesley, he was CEO of Compumotor, Prolog, and RJS. He has a PhD in Electrical Engineering from the University of Arizona where his research interests were in Artificial Intelligence and Pattern Recognition.

Jerry Wesley is the Vice President of Software development for Apollo Imaging Technologies. He is responsible for all of the DSP software for the Apollo Intelligent Vision Platform, including the Board Support Package (BSP) and demo applications. Additionally, he developed the SDK which also includes a Windows client. Prior to joining Apollo, he worked as a real-time embedded engineer and/or manager for various companies including Wind River, Doctor Design, Uniden, Simpack Associates, Hughes Aircraft, and NASA. He has over 25 years of experience and has a BS in Electrical Engineering from New Mexico State.

They can be reached at contactAIT@apollo-image.com.



New OMAP35x Applications Processors
Multi-format video processing up to HD

Roll over for video

OMAP35x

TEXAS INSTRUMENTS

TechOnline Communities

Audio DesignLine | Automotive DesignLine | CommsDesign | Digital Home DesignLine | DSP DesignLine | EDA DesignLine
eeProductCenter | [Green SupplyLine](#) | Industrial Control DesignLine | Mobile Handset DesignLine | Planet Analog | Power Management DesignLine
Programmable Logic DesignLine | RF DesignLine | Teardown.com | TechOnline | Video/Imaging DesignLine | Wireless Net DesignLine

EE Times

United States | Asia | China | Europe | France | Germany | India | Japan | Korea | Taiwan | United Kingdom | EE Times Supply Network

Additional Network Sites

Analog Europe | Automotive Designline Europe | DeepChip | Design & Reuse | Elektronik iNorden | Embedded.com
Industrial Control Europe | Microwave Engineering Europe | Portelligent | Power Management Designline Europe
RFID World | Semiconductor Insights

All materials on this site Copyright © 2008 TechInsights, a Division of United Business Media LLC All rights reserved.
[Privacy Statement](#) | [Your California Privacy Rights](#) | [Terms of Service](#)