Intelligent Video in 2005

The prospects, players and problems behind smart surveillance

By Mark Sartor, Peter Kalocsai and Nicholas Imparato

Video surveillance has changed. The days of a minimum-wage security guard watching a few cameras, looking (or not) for something unusual are gone. We can't just add "dumb" cameras to increase security; that only results in a need for more guards, who become ineffective after only 20 minutes of watching video monitors. Think about it: After a guard has been viewing surveillance monitors for millions of seconds, the likelihood that he would notice a two-second terrorist event, comprehend it and react effectively is reduced to pure chance.

The good news is that the market for intelligent video, often described as motion, pattern and behavior analysis, is growing. This software is intended to watch everything and send alarms only when it notices "interesting" situations. Guards respond only to what the analysis tools find alarming, suspicious or unusual. Although challenges remain, that's the promise.

Of course, despite the emphasis on technological innovation, the whole configuration is mediated by humans. If there are too many false alarms, for example, guards will turn off the system—no matter how intelligent it is reported to be. That said, there is still much to admire in the progress that the fledgling smart video industry has made over the past year.

Getting Perspective on the Technology

Video-based behavior analysis is the interpretation of events and patterns that are determined by detection and tracking of objects. Once an object is detected, its movement and location within a scene are used to determine if there is an event—say, crossing a threshold or speeding. This might require handoff and tracking of the same object among multiple cameras.

Ideally, the object can be identified through some set of attributes, which may also help in the interpretation of the behavior. Nonetheless, the most basic design choices in the video and optical system—for example, progressive vs. interline scan, field of view, compression, sensitivity and lighting—will make or break the ability of the system to provide such an advanced capability.

The system determines behavior by detecting a series of events that have a pattern or meaning in space or time. Example: A vehicle is detected repeatedly circling a critical building, counter to the normal traffic pattern. In order to provide an effective alert, the system must (1) have a clear enough picture to recognize identifying features of the car across all cameras; (2) track the car and associate these features among any number of cameras; and (3) receive or automatically generate and apply a threshold that determines the normal frequency of the event, and then compare it to the occurrence.

Once the information is detected and translated into data at the event level, the video itself is no longer relevant to the analysis. Now software from other fields, like radar, can be used to interpret the data.

The querying and forensic analysis of pre-recorded video can be just as important as the processing of real-time data. A good, content-based retrieval system allows you to scan through weeks of pre-recorded video within minutes, facilitating the location of critical information about potential security breaches. Ultimately, this can be

used to understand the behavior of the attacker: a critical component of forensic analysis.

The State of the Art Today

The state of the art is rapidly advancing as both government and industry pump funds into developing better analysis tools. Purchasers should breadboard solutions before they proceed into full-fledged implementation. Military agencies are currently scrambling to develop and provide guidance and standards for video-based security systems. Until they succeed, the market will remain a formidable mixture of promise and performance.

There are four major aspects of intelligent video that reflect the effectiveness of the total solution: moving object detection, object tracking, recognition, and behavior and event pattern analysis.

Moving object detection. Offered by all intelligent video vendors with varying performance. Thermal infrared (IR) performance will differ from visible (EO) camera performance. Complications arise when cameras are panning or moving and when video is less than ideal (noisy or compressed). Often objects are detected by motion within the video—for example, pixels change in successive frames—and when they cross a virtual trip wire, loiter or are left behind. However, changes between frames are affected by many parameters, including movement of the camera, noise in the frames, and compression artifacts.

Object Tracking. Solved only at the most basic level: within a single camera and under good circumstances. Tracking software can provide feedback to a PTZ or provide data such as speed or location if the frame position is calibrated. A number of organizations have more sophisticated video-based tracking algorithms that are starting to hit the market, but these are experimental at best. The problem is exacerbated when many objects are being tracked in close proximity and across multiple cameras (although stereo cameras have been used to help mitigate this issue) and when objects disappear and reappear.

Motion detection and tracking are supported by virtually all vendors in the field at some level. More advanced features further complicate evaluation.

Recognition. Largely unsolved, although there are some products being offered in a range of fields. These products are useful when combined with other technologies to improve performance. Performance can improve dramatically if 3D or other sensory data are used simultaneously.

Behavior and event pattern analysis. Basic pattern analysis capabilities (abandoned object, tailgating, loitering) are offered by a number of vendors with some degree of

Top, camera captures man walking along Amtrak rail bed parallel to the moving train. Center, ObjectVideo VEW detects man as potential threat. Bottom, when man leaves bag, an alert is issued to security personnel. Courtesy ObjectVideo

success. For the most part, available systems recognize simple, pre-programmed patterns and create alerts for operator intervention. Actual analysis or interpretation of behavior is still largely in the R&D space. Several research organizations, including Sarnoff, Roke Manor, SRI and BAE, are developing video analysis tools that learn how to distinguish between normal and unusual scene behavior. This type of analysis is greatly complicated when the event or behavior spans many cameras, many objects, large areas, and/or large time spans.

Products and Testing

Pelco has established the Intelligent Video Testing Laboratory, under the leadership of longtime video expert Jim Dunn, to test the performance of various features offered by different intelligent video products. Vendors voluntarily submit their products for testing, which involves running each vendor's application on the same prerecorded DVD video sequences. Live performance is also tested using the same cameras.

Different products excel in different areas. A unitary judgment or evaluation might not be possible, and this vacuum points to the need for objective standards in the evaluation of video performance. The table below lists major vendors of intelligent video products. "Basic features" indicates video motion detection and basic behavior functions. "Advanced features" indicates behavioral or event analysis. If the products have been tested at Pelco's Intelligent Video Testing Laboratory, this is indicated in the table.

No Revolution Left Behind

There is tremendous interest in research to improve algorithms and performance of video-based systems. Smart camera systems will improve their ability to detect breaches, objects left behind, suspicious or wrong behaviors. Eventually, they will even identify people automatically through biometrics. But at the end of the day, putting the complete picture together in an intuitive way is still a part of the solution that needs to be addressed more vigorously.
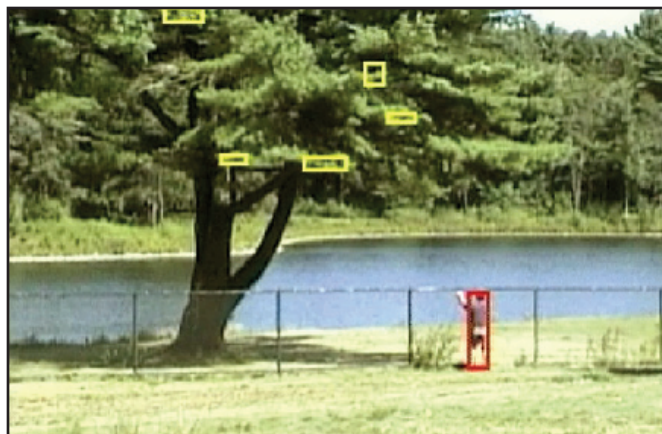
Enter 3-D technology. The creation of a 3-D virtual

Table 1: Notable Vendors with Intelligent Video Solutions (in alphabetical order):

Vendor	Web Address	Pelco Tested	Basic Features	Adv. Features
ActivEye	www.activeye.com	•	•	•
Cernium	www.cernium.com	•	•	•
DynaPel	www.dynapel.com/index.shtml	•	•	
GE Interlogix	www.geindustrial.com/ge-interlogix/kalatel/videoiq.html		•	
Genex Technologies	www.genextech.com/pages/612/Intelligent_Algorithms.htm		•	
Guardian Solutions	www.guardiansolutions.com		•	
IntelliVision	www.intelli-vision.com	•	•	•
ioIMAGE	www.ioimage.com	•	•	
iOmniscient	www.iomniscient.com/index.htm	•	•	•
L-3 Com	www.l-3com.com		•	•
Northrop Grumman	www.tasc.com/solutions/ip/alertvideo.html	•	•	•
ObjectVideo	www.objectvideo.com	•	•	•
PerceptiVU	www.perceptivu.com/PerceptiVU.html	•	•	
Radiant	www.dvmd.com		•	
Radionics (Bosch)	www.radionicsinc.com		•	
Roke Manor (Siemens)	www.roke.co.uk	•	•	•
Safehouse Group	www.clarityvi.com/index.html		•	•
Sarnoff*	www.sarnoff.com		•	•
Verint	www.verint.com/video_solutions		•	•
Vidient	www.vidient.com		•	•
Vistascape	www.vistascape.com		•	•

* Sarnoff recently sold its Vision Alert, Video Flashlight and Hawk technologies to L-3 Com

environment of the secured area into which all the video and alert information can be synchronously imported and displayed, and with which all data can be played back and analyzed at any time or point in the virtual space, will be a key advance.

Through such an environment, patterns of behavior can displayed or played back for the user to interpret in the context of the scene. Alarms can be generated or used to slew the 3-D display to a bird's-eye view of the problem while displaying the GPS coordinates of the alarm. This type of environment provides an intuitive interface and

The SmartCatch 2.0 from Vidient features a user-friendly GUI for viewing and managing alerts. Courtesy Vidient

The VideoIQ from GE's security business uses Concept Coding™ technology to recognize people by their behavior.

the ability to synchronously view, record and play back all of the video from any perspective at any designated time. The benefit? More effective response with fewer guards or responders, as well as faster training cycles.

Measuring What We Think We Know

As more companies enter the market with intelligent video applications, integrators and end users find it more difficult to differentiate among them. The current testing of solutions is normally qualitative at best, leaving space for highly subjective evaluations. Quantitative testing of video content and behavior analysis tools has only recently been coming into its own.

Several organizations are active in this area. On the government side, these include the USAF Force Protection Battlelab, the Army Corps of Engineers and Sandia National Labs. The commercial world has also become engaged, as noted before with Pelco's Intelligent Video Testing Lab. Pelco is currently in the process of developing benchmark test sets and sophisticated statistical tools

that will increase objectivity in the assessment process. In addition, the company is providing an Application Porting Tool to facilitate easier integration into its test environment and product line. The goal of benchmark testing is to provide quantitative comparisons among candidate applications so that hardware customers get the software product that best fits their needs.

Accurate and automatic video-based assessment of events and behavior is a relatively young industry that is rapidly growing and changing. Users need to anticipate an iterative process while the software is customized, refined or otherwise tweaked.

Public Policy

Privacy concerns will grow in tandem with advances in the technology. The chief privacy officer at the Department of Homeland Security had grown her staff to 400 by the beginning of this year. Hers is the only CPO position in government that is statutorily required. The group's mandate, which includes an assessment of how new security-oriented technologies align with current privacy legislation, is a popular one.

Buyers will and should ask questions about a technology's profile in terms of public expectations and priorities. Although a number of industry executives sit on various public-private coordinating task forces and committees, more involvement will be needed in the future. Chief executives of intelligent video firms will have to put more resources into public policy compliance, and that will make things interesting. More important, it will produce a list of winners and losers that can't be easily predicted today. **ST⊕D**

Mark Sartor, PhD., is head of Sarnoff's Vison and Visualization Business unit. He has worked for more than 23 years in the image and video processing fields.

Peter Kalocsai, PhD. (pkalocsai@pelco.com), is a senior research scientist at Pelco's R&D department, focusing on face and object recognition, intelligent video applications and statistical analysis.

Nicholas Imparato, PhD. (imparato@usfca.edu), is a professor at the School of Business, University of San Francisco, a Research Fellow at the Hoover Institution, Stanford University and an industry advisor.