Keen vision

Analytics technology is providing public-safety agencies with new ways of using surveillance video
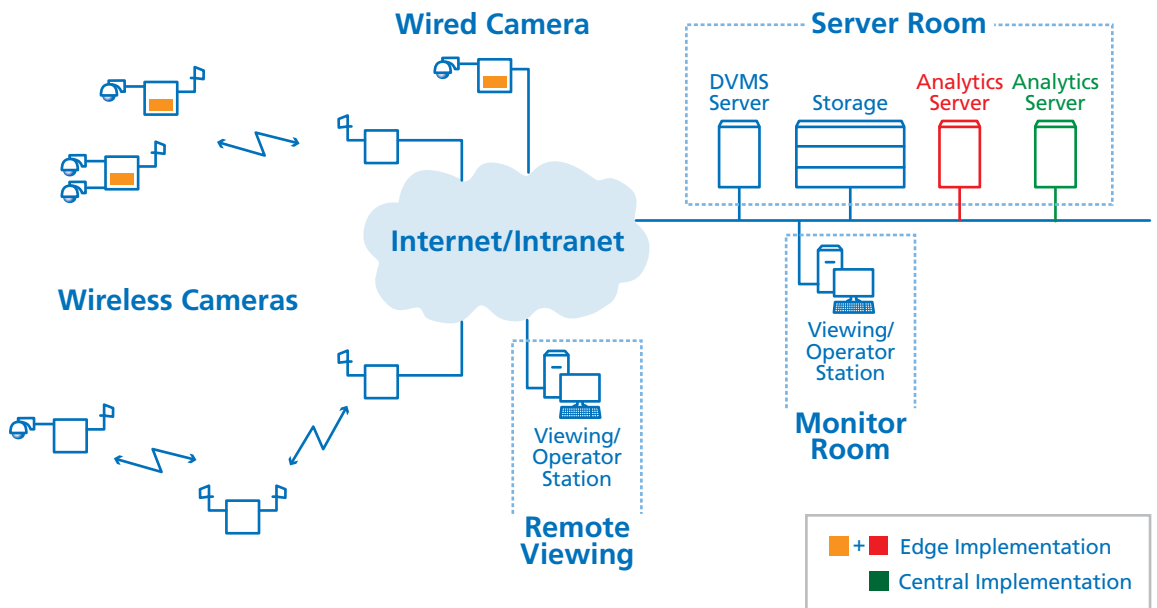
By Leonhard Korowajczuk and Jasper Bruinzeel

The surveillance industry has changed dramatically in recent years. Not that long ago, analog camera systems with analog video-tape recording systems were the industry standard. Then, digital video recorder–based solutions were introduced, allowing the connection of a limited number of analog and/or digital camera feeds, using a proprietary hardware platform. Now, surveillance solutions are managed through so-called network video recorders that feature video management software applications operating on standards-based computer hardware.

IP cameras are connected to the NVR server through IP-based networks. Legacy (analog) cameras can be tied into these digital surveillance systems using encoders. The number of cameras within a single system has virtually no boundaries, and complex, distributed architectures can be supported with multiple server and storage locations. As long as a high-speed connection is available, any networked computer then has access through the NVR servers to any of the cameras for video monitoring and system management purposes.

In parallel, broadband wireless technologies have matured, while the allocations of the 4.9 GHz band for public-safety applications fueled the adoption of citywide surveillance systems to address safety and crime effectively in medium to large cities throughout the United States. Since the wireless platforms today are natively IP based, these implementations logically are digital and NVR-based implemen-

FIGURE 1 EDGE IMPLEMENTATION VS. CENTRAL IMPLEMENTATION

Wired Camera

Server Room

DVMS Server · Storage · Analytics Server · Analytics Server

Internet/Intranet

Wireless Cameras

Viewing/Operator Station

Viewing/Operator Station

Monitor Room

Remote Viewing

■+■ Edge Implementation
■ Central Implementation

||||||||||||||||||||||||||||||||||||||||||||||||||||

5 STEPS FOR BETTER VIDEO SURVEILLANCE

Acquire

Analyze

Alert

Assess

Act

Source: Cernium

tations. Typically, cities also look toward integrating existing, legacy systems into the NVR-based solution, such as in-building security cameras systems.

As cities start building their unified, digital surveillance systems and add more and more cameras, one might wonder how this growing number of cameras can be managed. After the change from analog systems to open standards-based scalable, digital solutions, the next paradigm shift within the surveillance industry very well might be the shift to intelligent surveillance, rather than current evolutionary progress, such as the emergence of H.264 encoding and megapixel cameras.

Research has indicated repeatedly that the concentration span of surveillance system monitoring personnel is limited. As the number of video feeds grows, the ability of monitoring personnel to effectively identify suspicious behavior declines. The purpose of video analytics now is to make surveillance systems "intelligent," but only to a certain extent. There should not be an expectation that the human operator can be replaced, nor that all alarms generated from the video

analytics engine can be considered 100% accurate. But the idea is that such systems analyze the video feeds and alert the operator of situations that would require a closer look; a human assessment. The operator now concentrates primarily on only those few camera feeds it is alerted to by the video analytics engine, and the overload of the operator is no longer an issue.

The video analytics process is highly complex and processing intensive. This is why basic video management solutions have not integrated any analytics features beyond basic motion-detection. Specialized video analytics packages are available as overlay implementations to augment the basic video management solutions.

Video analytics engines follow similar steps in analyzing video. First, in the process of segmentation — based on image changes from frame to frame — the video analytics engine identifies the foreground and background pixels. The result of the segmentation process is the identification of a certain number of "blobs," with a blob being a collection of connected pixels.

Next, the process of classification assigns a class to each identified blob. Given a variety of applications, classes might include people, certain types of vehicles, animals and certain static objects.
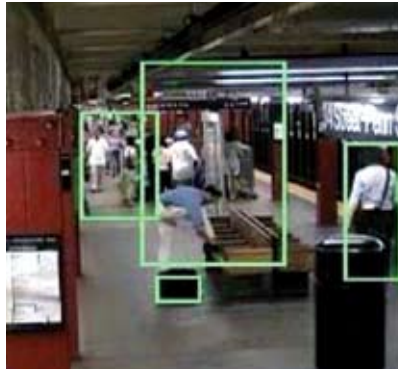
As classified blobs move through the field of view, in multiple frames, tracking algorithms are used to follow each blob. By analyzing the class of the blobs in combination with the tracked movement pattern, activity recognition can be derived to identify the behavior of the blobs, and more importantly, possible suspicious behavior. If multiple people blobs converge, we can conclude that a crowd is forming. If a vehicle blob stops moving in an area where parking is prohibited, this vehicle is illegally parked. Typical behaviors that can be identified with analytics engines include crowd forming, object removed/left behind, crossing of perimeter lines and loitering.

Several years ago, due to processor intensiveness, a central video analytics server would be able to handle only very few camera feeds in terms of analytics, making solutions expensive and impractical to implement. As a result, the concept of distributed analytics processing at the "edge" (i.e., with the camera unit) became popular to allow for better scalability. However, edge-based analytics processing introduces

the issue of compatibility. For new deployments, only certain encoders or IP cameras that allow for running a certain analytics solution can be used. Furthermore, for existing camera systems, significant upgrades might be required to enable the edge processing.

Based on a review of current video analytics solutions, given the strong advances in hardware processing power, the central processor approach today supports many more cameras per server, compared just a few years ago. But strong variance between solutions exists, with a single server able to support anywhere from 12 behaviors in total to as many as 100 cameras with any number of behaviors.

In terms of camera design, it should be clear that current video

Tracking algorithms are used to follow each classified "blob" as it moves through the field of view.

analytics technologies require still images, so that the systems can learn the static background (i.e., background pixels) and identify the foreground blobs. Regarding fixed cameras, the image is inherently "still," as the camera will not move, and the analytics add-on is

therefore straightforward. However, pan-tilt-zoom (PTZ) cameras, commonly used in citywide surveillance applications, might be used in a tour mode of operation and require further consideration. Without adding additional cameras at a location of interest, the tour mode needs to be disabled when analytics are applied to a specific PTZ camera. In other words, a PTZ camera could be programmed to operate in a static, preset position with analytics enabled during certain portions of the day, while during other parts of the day the PTZ camera would resume its tour mode with analytics disabled. For instance, during the day, a regular PTZ tour mode is maintained, but at night the camera points toward a sensitive perimeter line or a high-risk graffiti spot.

Given the advances of server processing power, a central analytics processor implementation provides clear added flexibility compared to an edge implementation, while costs of implementation between a central processor–based and edge processor–based deployment are nearly equal. A central video analytics processor is therefore recommended. In terms of costs, for a 12-camera implementation, the cost per camera roughly would be in the range of $3,000 to $6,000, assuming a standalone project adding video analytics to an existing surveillance system on a turnkey basis.

In terms of application, there currently is not an all encompassing analytics solution that can be expected to alert any suspicious behavior in a citywide deployment model. Available packages, however, do offer a fairly broad range of behaviors, based on classification of people, vehicles, etc. Also, most video analytics companies are very open to working with clients to add behaviors and functionality to their packages, in response to specific needs. Based on the growing size of surveillance systems within cities and the associated need for intelligent video, a fast evolution of analytics packages is expected with possibly a profound impact on the functionality and operation of these video surveillance systems.

Already today, analytics packages should be able to provide benefits for specific applications, and the timing is right for cities to start testing and implementing analytics, at least on a small scale for specific uses. Particularly, with a central processor–based approach, implementation is simple, and a specific analytics license can be transferred from one camera to another. This allows cities to test analytics in a pilot setup at different locations in a network, without spending a large budget right away. As cities take their first steps into video analytics, by forming partnerships with their integration partners and analytics vendors, video analytics might soon develop into a mandatory part of surveillance systems from day one of deployment and fundamentally change the mode of operation. ∎

Leonhard Korowajczuk is the chief executive officer of CelPlan Technologies. Jasper Bruinzeel is the company's vice president of marketing and sales.