

# A Video Analytics Primer for CCTV Users

A Pepperdog Whitepaper

July 2007



## About Pepperdog

Pepperdog develops and supplies advanced Video Analytics software applications for use by organizations that rely upon CCTV video surveillance to protect people and property.

For more information about our products, or to provide feedback on this document, contact us using the information provided below.

## Notices

© 2007 Pepperdog Ltd. All Rights Reserved.

The information in this document is provided in good faith as an aid to decision making. Pepperdog Video Analytics Ltd (“Pepperdog”) cannot accept any liability for the consequences of using this information, howsoever caused.

This document may not, in whole or in part, be copied, photocopied, reproduced, translated, or reduced to any electronic medium or machine-readable form without prior consent in writing from Pepperdog.

Pepperdog may have patents and/or patent pending applications covering subject matter in this document. The furnishing of this document does not give you any license to these patents.

All Pepperdog brands and product names are trademarks or registered trademarks of Pepperdog and other brands and product names may be trademarks or registered trademarks of their respective owners.

## Contact Us

Pepperdog Ltd  
High Beech  
63 Nine Mile Ride  
Finchampstead  
Berkshire  
RG40 4ND  
United Kingdom

office: +44 (0) 1189 736811

email: [info@pepperdog.net](mailto:info@pepperdog.net)

web: [www.pepperdog.net](http://www.pepperdog.net)

**Document Version: 1.41**

# Summary

Video Analytics is a relatively new industry and the technology still evolving. Yet most people might be forgiven for thinking that it is in fact well established, since they encounter vehicle license plate recognition technology on a daily basis; on petrol station forecourts and in congestion charging systems, or in TV programmes that depict advanced video analytics as a routine part of criminal investigations.

Between these two extremes of everyday use and pure entertainment fiction lays today's state of the art.

In exploring the realm of the possible, it is important to understand not only a solutions potential capability, but the degree to which that that is constrained by the operating environment.

Recognizing the licence plate of a moving vehicle is impressive, but since cars can only travel within certain parameters (lanes, speeds, directions), positioning cameras to obtain a clear view of a licence plate is relatively easy. In addition, vehicle motion can be frozen by CCTV cameras running at only modest frame rates so that a static analysis of the image is possible. Finally even with these constraints, specific government legislation was introduced to mandate the format and font of new licence plates to enable easier recognition.

In comparison, people walk through train stations or shopping malls on no fixed path, appearing on CCTV cameras at variable angles and often obscuring one another in doing so.

Our own eyes and brains can interpret these changing images almost without conscious thought, and yet achieving the same task computationally requires dynamic analysis that is many orders of magnitude more difficult to achieve than vehicle licence plate recognition.

Even so, some vendors have been tempted to overstate their capabilities, resulting in a healthy degree of scepticism being introduced into the emerging market for Video Analytics.

In this paper, we seek to help the reader understand the common scenarios for automated computer analysis of CCTV images, along with a critique of its usefulness and limitations.

# The Growing Need For Video Analytics

Everyone is familiar with the growing use of CCTV cameras for video surveillance. As a society we have justified our acceptance of this proliferation on the basis of preventing crime and maintaining public order. If a camera is not immediately visible in a public place, we simply assume it's well hidden, with the average UK citizen caught by one of 4.2 million cameras over 300 times a day!

Yet in practice, video surveillance has failed to live up to its promise of preventing crime. Most CCTV cameras are unmonitored and the vast majority of benefits are in forensic use, after the fact. Yet even here poor quality images coupled with a suspect wearing a simple hooded top can often defeat any useful detection.

It would be much better to try and be proactive, identifying known precursors to suspect and direct security personnel to investigate and intercept before its too late.

This is something that human CCTV operators are trained to do. However, in all but the most secure environments it is simply not economically viable to match the number of CCTV video cameras being deployed to the number of trained operators required to monitor them. And even in the most secure 24 by 7 environments, human performance characteristics limit the number of cameras that can be monitored simultaneously and the time after which viewing fatigue sets in; typically 15-20 minutes.

Video Analytics is the application of computer software to monitor CCTV video images and filter out the 0.1% of events that are of interest from the 99.9% that are not.

Just as with more general Business Intelligence technologies, Video Analytics software aims is to augment the efficiency and effectiveness of human operators rather than replace them entirely.

By continuously monitoring all video images and only alerting operatives to events of interest, Video Analytics applications can optimise the performance of the surveillance infrastructure, converting it from a passive to an active "preventative" system with a far wider range of applications such as in health and safety.

After all, preventing accidents or crimes has much greater value that simply recording them.

# Common Monitoring Scenarios

Video Analytic systems vary in cost and sophistication but it is important to understand what you are buying if you are going to get what you pay for. This section outlines the most common surveillance modes prevalent in the industry today.

## Video Motion Detection (VMD)

Simple systems rely on motion detection sensors to trigger the recording of significant events. Whilst useful in the simplest of cases such as an intruder in a secure indoor area, they have limited effectiveness in outdoor environments where an unexpected rabbit or tree swaying in the wind can raise a false alarm. And although more sophisticated systems can tune their sensitivity to avoid being set off by small objects, they cannot cope with expected and authorised movements, such a security worker on a scheduled patrol.

## Intrusion Detection

More sophisticated than VMD, this type of system raises an alarm when an object moves into a predefined area of a monitor. The system may also allow for rules relating to the direction of travel, for example to detect persons going up on a down escalator.

## Tripwire

More sophisticated than Intrusion Detection, Tripwire systems allow lines to be set in a cameras field of view to allow for monitoring complex areas with perspective, to raise an alarm only if a line is crossed and the line break is of a certain size and in a certain direction of travel.

Whilst in common use, these systems are notoriously difficult to set-up so as to avoid generating too many false alarms. This also has implications should any CCTV cameras inadvertently change position after routine maintenance.

## Background Learning

Background Learning is a technique used upon system initiation to 'learn' the scene that a camera is directed at, and subsequently to filter out fixed elements of the scene that have no or only limited movement, such as trees that sway in the wind. It is also a useful technique to identify and raise an alert of a camera has been moved.

## Left Object and Museum mode

Left Object detection raises an alarm when an object, such as a suitcase, is left in a controlled region. Museum Mode is the

opposite of Left Object in that an alarm is raised only when an object is removed.

These modes build on simple Background Learning, using it to determine if a scene has changed with some measure of adjustment for the size of a change being possible.

Although these surveillance modes are often used as the prototypical examples of automated video surveillance, both modes are easily compromised when designated areas are obscured - such as in crowded environments.

Furthermore, in Left Object mode it is often difficult to distinguish a legitimate object - such as a cleaning trolley or luggage set down temporarily by a traveller - from one that is not, i.e. an unattended luggage bomb threat.

For all these reasons more sophisticated Video Analytics techniques will be required if the organisation is not to incur the inconvenience and expense of false alarms.

### Tagging & Tracking

This mode allows operators or the system itself to identify and tag objects and then track their progress within a single camera's field of view, or across a network of many cameras.

The accuracy of this capability is fundamental to the dependability and functioning of more sophisticated analytic modes.

### Object Detection

Most video - and some radar - analytics systems will identify objects and classify them into sizes and possibly speeds. For example a small object moving slowly may be a person, whilst a larger object travelling at speed may be a motorbike based on the probabilities determined by analyzing the contextual environment.

### People Counting

A specialization of Object Detection where a count is kept of the number of People objects moving through a certain area (e.g. doorway, gate or corridor) for health and safety or statistical purposes. This mode can easily be extended to cars or other distinguishable objects.

### Object Recognition

Sophisticated Video Analytics systems will use classifiers to identify any objects detected. Regular objects such as cars are relatively easy to classify. People present themselves in a variety of directions and angles making reliable identification

from every camera angle, in every situation, an immensely challenging task.

### Image Recognition

Of all the modes discussed this is often the most problematic with *False Acceptance* and *False Recognition* rates ranging from almost perfect for established Automatic Number Plate Recognition (ANPR) technologies, through to little more than 50% for Facial Recognition systems in uncontrolled environments, where grabbing and recognizing faces from partial views taken at variable camera angles in changing lighting conditions is a very complex problem.

Yet this area of passive biometrics is understandably of intense research interest for security applications where the use of active biometrics is unacceptable or impractical, for example in mass transit systems.

### Behaviour Profiling

The most sophisticated mode of operation and currently something of a Nirvana for the Video Analytics industry.

Behaviour Profiling draws upon elements from all of the modes already examined combined with advanced statistical analysis. The aim is to build models of behaviour based on video surveillance images combined with other sources of information such as geospatial and other more traditional sources of Business Intelligence data.

With applications ranging from footfall analysis in shopping malls to augment Customer Relationship Management data, through to identifying pre-cursors to criminal activity through identifying repeat patterns of movement and loitering.

A specific security example might be to correlate ANPR data with face grabs at maritime ports or airports to build a statistical model of the frequency of vehicle and individual visits with other reported incident data.

# System Design Considerations

Security systems are comprised of many components and their physical architecture will have a considerable impact on the way that Video Analytics can be introduced into the system.

## Analogue v. Digital Systems

Traditional CCTV installations used co-axial cable to connect site-based cameras to Cathode Ray Tube (CRT) monitors and Video Cassette Recorders (VCRs) located in a central control room. Images were sent encoded as analogue signals.

VCR's with video tape have since been superseded by Digital Video Recorders (DVR) that encode video signals into digital form. This can then be stored on a computer hard disk drive, either within the DVR or on some Network attached storage solution (NVR).

More recently, CCTV cameras themselves have started to encode images digitally as well as provide interfaces that allow them to be directly attached to an IP (Internet Protocol) based network. Whilst this allows for convergence between security and IT infrastructure, it creates the potential for competing network bandwidth as well as conflicts in other areas.

It is this transition of video infrastructure to the digital domain that is enabling the use and growth of Video Analytics. Yet the changeover to digital is far from complete. As a result, many Video Analytics vendors have been providing systems that convert CCTV camera analogue video into a digital format that can be analyzed by their tools. In computing terms this is a very costly process and places limitations on the number of cameras a system can support, as well as having considerable computing hardware costs.

To overcome this scalability problem, more sophisticated solutions use dedicated DSP (Digital Signal Processing) hardware to encode the video stream, sending digital images to other PC systems for actual scene analysis.

However, in addition to the requirement for sophisticated computing hardware, installation involves re-cabling the camera to the DSP either directly or via a local loop. Both of these factors tend to lead to these solutions being expensive and time consuming to implement.

## Centralised v. Edge Systems

As well as being analogue, traditional systems are highly centralised with all video feeds being cabled back to a central control room for viewing and recording. It is therefore not surprising that many CCTV installations have preserved this traditional system architecture in the digital domain and centralised all of their Video Analytics functions.

Whilst understandable, real world experience suggests that this approach does not scale due to the volume of video image information that has to be sent across the network versus the bandwidth available with typical DSL based installations.

There are lessons to be drawn from more general IT solutions, where the drivers for change from Client/Server to Service Oriented Architectures included the requirement to avoid sending large amounts of data across networks unnecessarily.

To this end, more enlightened Video Analytics vendors have realised that it must be possible to deploy data intensive analytic functionality at the networks edge, collocated with CCTV and DVR devices if required. The control room is then only sent alarms based on configured scenarios.

Some vendors have tackled this by integrating Video Analytics into CCTV devices using onboard DSP's, however there is little provision for data storage so low frame rates are used until an alarm is triggered. This is often of limited value where continuous recording is required to establish an entire scenario.

A comprehensive Edge solution must allow for the possibility of operators retrieving all relevant scenes pre-and post an alarm event. This means integrating the Video Analytics function with control of a DVR/NVR.

## Conclusion

In this primer we have outlined the common CCTV application scenarios for which Video Analytics solutions are increasingly being considered and some of the key system design issues.

Given the level of maturity of the market and the many technologies involved, we recommend adopting the following strategy when investigating a new Video Analytics solution:

- Evaluate any proposed technology in the exact same environment that the deployed solution is expected to operate in.

If you are going to monitor a busy street with changing lighting conditions and leaves or rubbish being blown across the scene on a regular basis, then this is the environment you should evaluate. Is the system robust enough to cope?

Canned footage, controlled demonstrations and even reference visits can not reflect the subtle specifics of a given environment.

- How easy is the system to deploy, configure and re-configure as requirements change?

Video Analytics technology is complex rocket science, but a CCTV operator has a job to do; can they use the system too? Or will any changes to the environment require expensive consultancy?

- Video is “fat” data. Can your infrastructure cope?

As the transition from analogue to digital systems gathers momentum, the security and IT worlds begin to converge. Can your current servers deliver enough PC horsepower? Can your network cope with the amount of data that a given solution needs to pump over it? Or will you need to purchase additional PC and networking hardware?

Are your security and IT departments both on the same page?

Video is the next great data frontier and represents a tremendous opportunity for organizations to exploit this underutilized asset.

Video Analytics technology will enable this change in the way organisations view and relate to video data.