

Video Content Analysis, what is it and why would I want it?

White Paper by Nick Hewitson November 2005.

What is Video Content Analysis?

There are a number of terms used in different industries and markets to describe Video Content Analysis:

- Analytics
- Behavior Recognition
- Content Analysis
- Concept Coding
- Intelligent Video
- Object Tracking
- Smart CCTV

They all however describe the real time use of computer vision in a security environment to monitor the CCTV camera feeds and *assist* the guard in his or her decision making process.

The UK is the country with the most CCTV cameras deployed, with over 4 million in use. It's claimed that if you walk through London you will be watched over 300 times, however this is clearly a misconception. While it is probably true that you will be in the field of view of a CCTV camera over 300 times during your walk through London it is certainly not true that you are observed that many times for number of reasons.

Firstly CCTV control rooms have fewer monitors than there are cameras, in many cases a number of cameras are sequentially displayed on a single monitor. If for example five cameras are fed into a single monitor which then you obviously have only a 20% chance of being viewed while in any individual cameras field of view.

Secondly the staff in the control room are often expected to deal with other issues as well as monitoring the CCTV. They will be responsible for issuing keys, badges and permits to both staff and visitors; they are also responsible for monitoring the access control and fire alarm systems, controlling radio communications with both their own foot patrols and possibly the local Police. In addition they will need to be away from their desks for breaks to visit the restrooms etc. During this time they are not monitoring the CCTV images.

Finally the design of CCTV control rooms expects the guard to watch a large number of monitors. According to ASIS International, a human can effectively watch 9-12 cameras for only 15 minutes. Security guard shifts are often 12 hours long so 11 hours and 45 minutes are ineffective monitoring. CCTV Today in November 2005 estimated that the probability of an event being recognised and acted upon if it was clearly in the view of a CCTV camera was less than 1 in 1000. CCTV has historically been a forensic tool not a real time crime prevention system.

The principal of Video Content analysis is to use computers to monitor all of the cameras all of the time and when something unusual happens to alert the security guard to it. For example in a retail shopping centre a person running is unusual. The system can detect that a persons running but is unable to differentiate between a benign event, a teenage girl running over to greet her boyfriend, or a criminal event where someone is running out of a shop with an armload of jeans. However if the running event is drawn to the security guard's attention he is able to make that subjective decision easily and respond appropriately.

Why would I want to use Intelligent Video?

The following scenario is taken from a real test of a behavioural recognition system monitoring access to parked aircraft in the USA. It shows what the advantages of computer vision over human monitoring can be.

There are 8 cameras monitoring a road that passes through a tunnel, above which is an area where commercial aircraft are parked. The system was tasked with looking for cars that stopped under the bridge and people climbing up the slope towards the aircraft. Over 1 month the system reported almost 300 events where vehicles were seen stopping. 298 of these were originally classified as false alarms caused by normal traffic flow problems. One event was due to a “fender bender” accident and one to a breakdown. There were no attempts to approach the parked aircraft.

At first evaluation, it would seem that the value of the system was negligible, all it had produced were 298 false alarms out of 300 events. Previously the cameras were monitored by a guard on conventional CCTV monitors and no events at all had been reported in the month before. It was found that the 300 events would take the guard on average less than 30 seconds each to determine the risk.

So instead of employing staff 24 / 7 for 30 days to monitor the tunnel only two and a half hours man hours were required over the whole 30 day period.

In addition, in the previous month the guard reported no events, given that each of the 300 events reported by the system actually took place in the test month. It is probable that a similar number actually took place previously when the guard was supposed to be watching and he didn't notice them. It is therefore highly likely that if someone had stopped a car briefly to allow a passenger to get out and approach the aircraft, the event would have been missed, while the Intelligent Video system would have caught it.

The “Smart” CCTV system had therefore raised the effectiveness of the monitoring from zero to 100% while reducing the operating costs from 720 man hours to 2.5 man hours of labour.

When the security manager looked at the cost effectiveness on this basis, he had no hesitation in purchasing a system.

Key issues to determine before looking at Video Content Analysis.

What are your operational requirements? As seen above, if it is to have a minimum number of false alarms then the human operator will be more effective, he failed to report any of the traffic events under the bridge, in fact he didn't report anything at all, so his false alarm rate was zero.

What percentage of the cameras is best monitored by computer vision, and what percentage is better monitored by a human operator? In general, today computers do better on the cameras where nothing much happens (and therefore guards get bored) and people do better in busy scenarios where occlusion between people makes it hard for the software.

A good example is an embassy that has a back alley where no one ever goes. This is covered by a CCTV camera and this was the only camera out of about 50 that was implemented at the beginning. The embassy realised that no one paid attention to this camera because nothing ever happened but if someone was in the alley they really needed to know about it fast.

In the majority of applications today, only a percentage of the total number of the cameras are monitored by video content analysis, some are only recorded and some are monitored full time by the security staff. You need to determine what is the specific risk and the most appropriate method of monitoring for each point. Do the risks and scenarios change during the course of 24 hours?

Can you build upgradeability into your plans? In many cases the number of cameras monitored by the software increases as experience of the benefits is gained.

Video Content analysis is a tool that allows you to improve your operational effectiveness. It is not the “all seeing” Big Brother monitoring all activity.

It helps you spot the needle in the haystack; CCTV provides huge amounts of mostly irrelevant data. Video Content Analysis extracts information from that data.

It reduces your costs, manual monitoring is inconsistent and expensive.

It reduces your risk by moving away from the limited human attention span of less than twenty minutes, and screening all of the video streams in parallel.

It allows you to move from a forensic mindset of finding out what happened after the event has taken place towards real time analysis and decision making.

You do still need to employ professional security staff to make the decisions on the information presented to them in a sensible manner.

-End-