The myth, hype and new reality of video content analytics

By Geoff Thiel, CEO of VCA Technology

The video analytics industry has not done itself any favours over the last five years or so. Video analytics was billed as the panacea to overcome the problems and fallibilities that people have when monitoring screens, such as the well documented twenty minutes maximum attention span for adults. Video analytics vendors pushed the hype too far, too fast, to suggest that their solutions were using approaches such as artificial intelligence to predict suspicious behaviour. They hinted that we were entering the world of the Tom Cruise film 'Minority Report', where we could identify criminals before they commit crime, the so-called 'pre-crime' state.

This wave of hype from the video analytics industry and the problems and limitations of earlier (and some current) systems have simply turned a portion of the market off. Many early adopters that deployed video analytics over the last five years came away disappointed. They wanted to improve their security by reacting to exceptions flagged up by the system, rather than having their security staff constantly monitoring (and missing) these events. But the core complaint was the huge number of false alarms or 'false positives', which have plagued some analytics systems. Similar to the story of the boy who cried wolf too many times, the faith that an end user has in a video analytics system evaporates very quickly as false alarms increase.

Many end users bought cameras and mainstream solutions purporting to have video analytics, but in reality some solutions were little better than video motion detection (VMD) systems. VMD systems detect changes in pixel brightness between successive video frames, which can be caused by moving objects or individuals, but can equally well be caused by lighting changes, moving tree foliage and other forms of 'nuisance alarms'. These systems, at their crudest, just report any pixel change above a sensitivity threshold to the end user. If these are used outdoors, the rate of alerts is very high and the vast majority of the alerts will be irrelevant.

Zones can be set in a scene to limit motion detection to a specific area of a scene. But the system may still be falsely triggered by a lighting change, cloud shadow or rain within this specified area. With a VMD-level system there is no real video analytics going on at all; all movement and image changes are detected and no real intelligence applied to tracking and classifying moving objects.

At the other end of the video analytics market, enterprise-level solutions are prohibitively expensive and have a bewildering range of features and settings and typically require expert consultants to spend days fine-tuning them before they deliver results. Even then, there can be on-going problems because of changes in the weather and season and the expert has to be called in repeatedly. Due to this complexity and the costs involved in refining a solution, end users may become disillusioned with a product that

never seems to deliver the answers that they need, offering high false alarm rates and delivering poor value for money.

Video analytics – the big myth

It is safe to say that video analytics will not be able to predict crime 'Minority Report'-style any time soon. In the future, substantial increases in computing power and major leaps in the development of artificial intelligence could eventually lead to computers having cognitive abilities resembling the human vision system, but in my estimation this could take 25 years and require computers 1000 times faster than we have now.

So, the first part of the myth is that analytic systems are in any way a match for an (attentive) human being. While it is not possible to make a direct comparison, I would say that current analytics systems have the cognitive level roughly the same as a mouse. To explain the problem a bit further, let's take an example: - if I turn around, I see my colleague's desk with a model car on it and just behind that there is a window through which I can see traffic going around a roundabout with buildings behind. I then see a faint object in the window, which is the reflection of a person passing the open door to my office behind me. Now, for me, all this is easy to understand because my common sense (knowledge of the world that I live in) tells me the size of a desk and hence the size of the model car. I know there is a window because of the window frame and the reflections in the glass. I also just 'know' from experience what is happening outside the window: what is a person, what is a vehicle, and what is a line of people standing at a bus stop.

Analytics systems, on the other hand, are very much like lower animals in that they don't 'see' fully stationary objects and require movement before they can make any sort of decision. In the scene described, an analytics system will be able to identify groups of moving pixels and will be able to infer indirectly what type of object they represent from size and general shape of the group of pixels, or perhaps from the way the groups of pixels move (because people have 'wobbly bits' such as moving arms and legs, whereas vehicles have a fairly constant silhouette). The analytics will not 'see' the desk, model car, or the building beyond; as far as it is concerned these are just parts of the background scene. If the analytics has been built with a long memory of the background scene it will also be able to detect the line of people at the bus stop because where there was once empty pavement, there is now something else.

It is not so much that analytic systems are stupid (for a computer program); it is that our brains are fantastic machines with vast accumulated knowledge of the world around us that enables us to understand what we are seeing. Analytics systems just see groups of changing pixels and only have simple mathematical rules and simple measurements of colour, size and shape for decision-making. This also explains why analytics systems degrade with the level of crowding because if moving objects are close together and crossing in front of each other, the groups of moving pixels merge and split in very confusing ways so that the analytics tracking engine will be unable to track every single object independently. In these conditions, analytics systems might be able to track a single person passing

through a crowd, or perhaps count people passing through a zone by doing a rough head-count, but going much beyond that is not currently possible.

As a very rough 'rule of thumb', analytics may only be suitable in about 50% of potential applications. If the people and vehicles in a scene remain fairly separate, then all the analytics functions will work well, but if scenes are more crowded, then analytics may only provide broad statistical measures such as approximate counts and area of coverage.

Video analytics – the positive reality today

For those of you that are getting depressed at this point don't be. Analytics may only work in some situations, but in these cases it has huge advantages over us humans. First of all it is cheap and for the price of a few hours of a person's wages you get a machine that is never off duty, and can provide you with real-time alerts and information for many years. Secondly, analytics is getting easier to use; with self adjusting algorithms, the tracking engines just work 'straight-out-of-the box'.

So what type of situations really suit analytics? First and foremost is the basic intruder situation of detecting someone who shouldn't be there. In this type of situation the added 'intelligence' of analytics can be used to reduce false alarm rates to acceptable levels. For example, the system can be set to only detect people and vehicles so that small animals, litter and other moving objects do not cause an alarm. Also, in the staff car park you may want to ignore people leaving the building or going about their normal business, but if an individual lingers for a long time or someone approaches the doorway at night you might want to check on them, or at least trigger a recording system. There are additional uses in transport where this ability to discriminate direction of movement and object type is very useful, for example, to detect intruders crossing a railway line that also has moving or stopped trains in the same zone. With suitable system design this sort of system can detect intruders on the track reliably while achieving false alarm rates of just one per week.

Another whole class of tasks falls into the category of 'abandoned object'. Cases that work well are vehicles that park where they shouldn't – this could be on a no parking zone in a car park or it could be a broken down vehicle on a level crossing. Alternatively, it can be used to detect criminal activity on an ATM cash machine by detecting the presence of a foreign device such as a card skimmer, keyboard overlay or spy camera.

I would finally like to mention the numerous tasks that can be achieved in more crowded conditions provided that the level of information required is reduced to just counting or estimating coverage of the background. Take a retail situation and combine the point of sale terminal and analytics outputs to detect cash resister transactions when no customer is present. Such incidents can indicate staff theft particularly if they happen frequently. Another application is to count the rate that people come into a store hour by hour during the week; this can help a manager with planning the weekly staff rota. Alternatively, a higher than usual flow rate of customers entering the store indicates that there will be extra pressure on the checkouts soon afterwards and so the store manager can re-deploy staff even

before the checkout lines build up. And lastly, the marketing department can get objective feedback on their promotional displays by looking at the occupancy information for the floor zones in front of the displays.

Analytics can be highly effective, offering robust solutions which can benefit not only security professionals but managers, marketers, indeed anyone that can use real-time information in their business to make it operate more efficiently, more safely or more conveniently. All companies are facing constant pressure to reduce their costs and increase effectiveness, and video analytics solutions can really help. A new breed of video analytics vendor is emerging, of which VCA Technology is one, which believes the future of analytics is to provide solutions which are simple to set up, easy to use and will deliver concrete results.

This new breed of video analytics has moved on a long way from the VMDs of ten years ago, especially around tackling the core problem of false alerts. Our VCAsys solution has a robust tracking engine which will realistically track many moving objects at the same time. It has the ability to differentiate objects of interest such as people and vehicles from other scene movement caused by moving clouds, tree foliage movement, small animals, lighting changes and weather related effects. With a robust tracking engine at its core, the user can then simply instruct the system for the behaviour that they are looking for by setting up detection areas and rule(s) that apply in each area. A zone and its rule can be set up in as little as 6 mouse clicks. Think about what you need to monitor: individuals or vehicles entering or leaving a zone; traffic speeds; vehicles stopping, individuals lingering longer than they should, vehicles making a wrong turn; an object being abandoned. By using a suite of scene filters you can monitor direction, stopping, dwell-time, appear, disappear, enter and exit, abandoned object, removed object or filter on object class such as person, group of people, vehicles etc. The ease of use of the system is a major step for video analytics. It finally enables the non-expert end user to sift out what he/she wants to spot and analyse further and discard all the time in between where nothing of interest happens.

Now is the time to re-examine video analytics. The costs are at a level that really makes it accessible to all end-users to try out – with OEM starting prices of VCAsys under $25 per system, the technology is robust, simple to install, use and maintain. The technology can be embedded on a DSP chip, uploaded into a network camera or encoder, or run in a PC – it is becoming highly accessible. This is a far cry from the previous mega-budget systems.

In conclusion, analytics takes CCTV well beyond simply being a system for providing video evidence after incidents have happened. It opens up a real-time data feed that innovative managers can use to improve their businesses. It is our hope that within the next few years, companies that have invested in cameras already will wonder how they ever lived without the use of video analytics to sweat these assets.