Making Video Analysis work

By Dr Rustom Kanga,

CEO, iOmniscient Corporation.

iOmniscient is the winner of the Global Security Challenge for Crowded Places awarded by the UK Home Office and BAE. The company is widely regarded as the technology leader in Video Analytics. Dr Kanga, its co-founder and CEO is also one of the authors of "Automated Surveillance" the first book on the use on the practical use of video analysis in industry which sold out its first edition within a month of being published.

Video Analysis seems to be the new buzz word in the Security Industry so everyone feels they must have it. Unfortunately the way most people buy video analysis they are highly likely to be disappointed with the results. Part of the problem is that there are many players in the market (including some otherwise very reputable names) who pass off simple "Motion Detection" technology as Video Analysis. However a large part of the blame must rest with the way users make their buying and implementation decisions.

Let us consider some of these issues in an attempt to understand what it takes to implement a successful video analysis system.

There are three factors that affect the success of the system namely:

1. Knowing what you require from the system
2. Choosing software with the right level of intelligence to meet those objectives
3. Correctly placing the cameras to optimize the performance of the software.

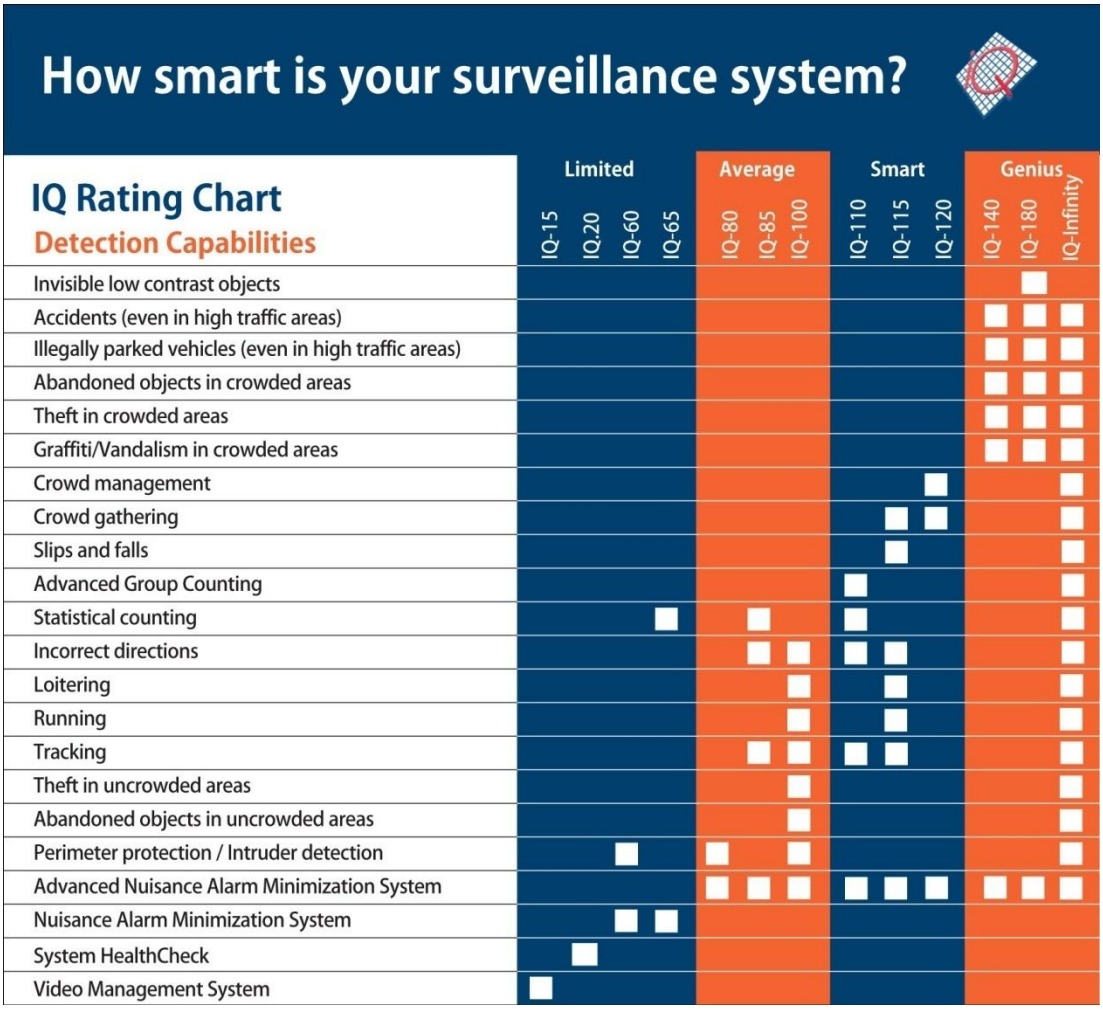
The Commoditization of Video Surveillance

Many parts of the security business are being commoditized. Commoditization involves a process where all the products are essentially the same and competitors are forced to compete on minor features, service and price. This is true for cameras which have been around for a while but as yet it is not the case for video analysis.

The intelligent surveillance industry is still at an early stage in its development. In this phase of an industry one tends to have a very small number of players who have effective technology and a large number of players who have jumped into the market to take advantage of the growth but with very limited technology. This makes it difficult for users to determine whether a solution that is proposed can really deliver what is promised.

One method that has been used to determine the intelligence level of the software for different requirements is to understand its IQ Rating. The Intelligence Quotient or IQ is a measure of intelligence used for humans. An IQ of 100 is the average of the population so half the population will have an IQ above 100 and the other half will be below.

A similar rating system has been developed for intelligent software.



Systems with very limited intelligence (IQ up to 80) can detect simple intrusions and perform some other simple detections.

Systems with a more average level of intelligence (e.g. IQ 100) can do tracking and various simple behavioural detections such as Loitering or Running. At this level one can detect objects left in an EMPTY scene.

At IQ 110 one can count people or vehicles. At IQ 120 one can do this in crowds. The genius level systems (IQ140 can detect abandoned bags in a CROWDED scene. Since iOmniscient has international patents on this capability no one else operates in this space. At the IQ 180 level objects that are invisible to the human eye can be detected as seen in the image below.

The most difficult scenes are the crowded ones.

Image of detection of abandoned bag that is invisible to the human eye as it is a black bag on a black background. This requires iOmniscient's patented IQ 180 technology.

Even at a particular level of intelligence there can be different levels of sophistication. The images below show the type of counting that can be performed by the counting system from iOmniscient (on the left) relative to the capability of more simple systems on the market at the same IQ level (on the right).

Systems even at the same intelligence rating can perform differently. The left hand side shows what can be achieved by counting systems from iOmniscient. The image on the right shows the normal view required for most generally available counting systems.

Detection and Identification

When suppliers talk about Video Analytics they talk about either Detection or Identification because virtually all of them can only do one or the other.

Detection involves using a wide and distant view to detect behaviours such as intrusion and loitering while Identification involved a much closer view sufficient to identify a vehicle by its number plate or a person from his face.

In the past Detection and Identification required different kinds of cameras and different camera placement. Facial Recognition for instance could only be done in tightly controlled environments where one got a close up of the face (usually requiring around 300 pixels between the eyes for good recognition). Recently iOmniscient announced a Facial Recognition system for Crowded Scenes that will operate with 22 pixels between the eyes in an uncontrolled environment allowing facial recognition to be performed using standard megapixel surveillance cameras.

View in which iOmniscient's new Facial Recognition system for a crowd would detect Faces in an uncontrolled environment.

Further the same camera can be used for Detection and Identification. iOmniscient's patented IQ Hawk technology allows multiple detections (eg for different events happening at the same time) to trigger identifications (to identify the person or vehicle involved). This technology also significantly reduces the storage and bandwidth requirements normally associated with megapixel cameras. For a 5 megapixel camera for instance the storage and bandwidth requirements can be reduced by 200 times. This means that if the use of traditional technology cost $200, the use of IQ Hawk reduced that cost to $1.

The very first step in having a successful implementation is knowing what one wants the software to achieve. Understanding the capabilities of different types of software ensures that the user can define the correct level of intelligence required for his project.

Asking for the Right Solution

Recently, Jeff Bardin, a consultant released a document translated from a Jihadi social networking site. It advised terrorists on how and where to place bombs. The core of the advice was that bombs had to be placed in crowded areas where they could cause maximum damage. This may seem like common sense. Why would a terrorist place his bomb in an empty location.

However when users call for tenders, whether it be for an airport passenger terminal or for a bank lobby, they rarely specify that their system needs to work in such a crowded environment.

Since they don't ask for this capability they get responses from suppliers who specifically cannot work in a crowded scene.

When these systems are implemented whether they are for detecting abandoned luggage or for recognizing people or for the detection of a number of particular behaviours, they will only work when the scene is empty – which means they are useless when they are really needed.

This is true for many airports, railways and city surveillance systems around the world where they have got what they have asked for. However because they were not precise in what they asked for they have now got systems that will not do the job that they require done.

So the first step in getting a useful system is to ask for what is required in very precise terms. Consider the following two extracts from two tender documents.

Tender Document 1:

Quote 1 - The video analysis system must be able to detect abandoned luggage.

Quote 2 - The system must be able to do facial recognition.

Tender Document 2:

Quote 1 - The video analysis system must be able to detect abandoned luggage in a crowded scene even when the luggage in question is partially or totally obscured for up to 50% of the time but is still present and abandoned up to 15 minutes after it has first been recognized as having been abandoned. The image of the object may be a maximum of 4x4 pixels in size.

Quote 2 - The system must recognize faces of people approaching the camera in a crowded environment even at distances up to 20 meters away where the person may not be looking directly at the camera and where the environment has variable lighting.

One can see from the above that the user that wrote the first tender would have received a large number of responses from people who could do simple things like detecting a bag when the scene was totally empty or recognize a face when the person looks straight into a megapixel camera in a very controlled environment. In the image on the right hand side below one can see the type of performance they can expect from systems which comply with their requirement. On the left hand side one sees the performance one can expect from a system that meets the requirements stated in the second Tender document.

The user who wrote the second document would have got responses that would have resulted in a more useful system. (1)

The image on the right is the typical system that one would receive in response to a poorly worded tender document (as in Document 1 above). The abandoned bag shown can be detected but only because the scene is empty. The image on the left shows what the patented iOmniscient technology can provide if the user requested this in his requirements (as in Document 2 above).

The Curse of the False Alarm

The biggest challenge with any video analysis system is not what can be detected but the number of false alarms that are generated in attempting to make the detections.

A video based system can be prone to false alarms for a number of reasons. First of all light variations (shadows and shards of light) can be mistaken by a system for a real person or object. Moving clouds, swaying trees, ripples of water and extraneous birds and small animals can all cause a system to mistake something that is quite innocuous for a real threat.

A system that relies on simple motion detection is likely to give a very high number of false alarms and if there are too many such alarms the operators will stop using the system.

How can one define whether a system is good at eliminating false alarms? It is difficult to set a standard. Even a human may make a mistake or see a mirage in certain difficult situations. What users should know however is that there are suppliers who have put a significant focus in addressing this issue and there are others who have not. Those who have addressed this requirement have built a module called a Nuisance Alarm Minimization System (NAMS).

From the users' perspective it is important to know whether a system comes armed with NAMS or whether the supplier has not even thought about the concept.

Even with NAMS algorithms there are simple ones which can cope with some basic types of false alarms and very sophisticated ones. Ultimately false alarms are the biggest cause of irritation with these systems and users would be well advised to get a system with the best NAMS capability that is possible (2).

Making the Final Selection

With literally thousands of companies offering video analysis let us summarize how a user can choose the system he requires?

The first step is to decide what applications and hence what level of intelligence is required. Does the system have to work in a crowded location? Does the system have to cope with difficult light variations? What is the IQ level of the system that is needed. Based on this assessment the user has to choose suppliers that offer products with the required level of intelligence.

Promises come easy. Delivering on the promise is not as easy. This is true for any new technology. It was true in the pharmaceutical industry; it was true in telephony and it is true today for video analysis.

The user should consider every promise from a supplier and extract from it what has not been said. If an abandoned bag can be detected – can it be done in a crowded place? Can it be detected even it has been left there for a long period. Can it be detected if it cannot be seen by a human? Has the supplier conveniently avoided mentioning these aspects of their solution.

Once one has eliminated the suppliers that cannot meet the user's requirements, the next step is a shootout. Here the user positions the key suppliers against each other and lets them demonstrate their capabilities. Even for this it is important to ensure the tests reflect the real world in which the systems are going to be used. If a company says that it can provide a certain performance it is fair to let them prove it. As these tests can be very expensive the user may need to consider paying the suppliers their costs for participating in the tests. The most competent suppliers are in high demand and they may otherwise decline to participate in a trial. The user will then be left testing a number of less effective systems.

The best system will not necessarily be the one that is offered at the lowest price. If the cheapest system cannot deliver the required results then it is too expensive at any price. Unfortunately we see users repeatedly burnt by picking the system that comes at the lowest price and realizing too late that it cannot deliver on its promise.

The Importance of Camera placement

The software is the brains. The cameras are the eyes. Having selected the right software it is still important to ensure that the cameras can see what it needs to see to make good judgements.

Installers have been installing cameras for years and they generally believe they know exactly how and where cameras should be placed. Unfortunately computers are much less forgiving than humans in terms of what they would accept as a valid input for their decision making and it is therefore important to understand how a particular video analysis product works to ensure that the camera is ideally placed for the necessary software to work at its best. Not all installers are sufficiently knowledgeable on video analysis to know how and where to place their cameras.

Its not that video analysis systems alone require special camera placement. Humans too require correct camera placement if they are to view a scene and make effective judgements.

At one of our customer sites video analysis was installed and it was found that at a particular time of the day the system could not operate as it was blinded by the sun shining straight at the camera. The customer was initially upset. According to him the camera had been in that location for ten years and as far as he knew they had not had any problems with it - but when the video analysis system was installed then it just could not operate effectively at that particular time of the day. At our suggestion they then talked to the guards who had been using the system before the video analysis software was implemented. The guards confirmed that they could not see anything on that camera when the sun shone at it. The system had always been ineffective at that time of day but no one had mentioned it and hence they were under the delusion that it worked perfectly.

The need to ensure correct camera placement is not surprising. With any technology it has to be used in the environment it was designed for. A Formula 1 racing car will not perform well in narrow village streets and a video analysis system will only work well when the camera view provides an ideal image to the system.

The lesson is very simple. If you want a working video analysis system, make sure that your installer has received sufficient instruction and information from your video analysis supplier. Otherwise there is a high risk that your system will not work as it should.

Summary

There are many successful video analysis systems that have been implemented. Unfortunately there are an even larger number of unsuccessful ones. If a user wants to ensure that his system is successful then he has to do a few things first.

The more clearly he can articulate the objectives of each cameras the more chance there is of the resultant system being competent to meet the objectives.

The user must then select the software that can meet his objective.

Finally the cameras must be placed optimally for video analysis.

The video analytics must always be selected before one chooses the cameras. Camera suppliers have developed some amazing technology but not every camera is suitable for every application. If one selects the camera first there is a high risk that the camera may not deliver the type of image required. Facial Recognition systems require very different cameras to License Plate Recognition systems. A counting camera will operate very differently to a camera required for detecting abandoned objects in the dark.

By rigorously setting out objectives, choosing the right software to meet those objectives and placing the cameras appropriately for the selected software you have a high chance of making your system one of the successful ones.

Notes:

(1) For more details on how to write tender documents that result in useful systems refer to the appendix of “Automated Surveillance” by Dr Rustom Kanga et al available from www.amazon.com or directly from info@iomniscient.com
(2) Specifying a NAMS capability is also addressed in the book mentioned in Note 1 above.