White Paper - November 2005

Defining Intelligence for CCTV Based Systems

What is the IQ of your System?

Dr Rustom Kanga & Ivy Li
iOmniscient – Intelligent Surveillance

Executive Summary

The Security industry has traditionally not used advanced technologies. Relative to some other industries, Security has relied on relatively simple mechanistic devices and human intervention.

However as security becomes more important, many players have entered the field with so called “Intelligent Systems”. Some of these systems would be classed as having no intelligence at all. Others have intelligence at various levels.

This document discusses the different types of technologies used to make systems intelligent and helps you to assess the intelligence that you would require in your own environment.

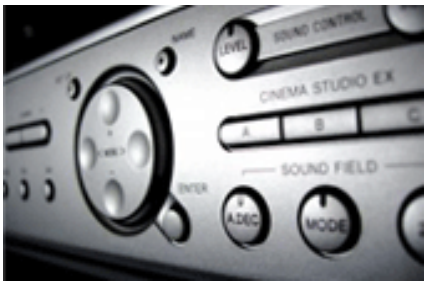
A History of Intelligence (or the lack of it)

The traditional security system consisted of a man and dog team.

People soon discovered that the security guard could not be present everywhere and CCTV cameras were added. This allowed someone in a control room to watch the scene even when the guard was not available to walk around.

Studies have shown that surveillance systems are limited without added intelligence. Watching only 2 monitors an operator misses 45% of action after 10 minutes and after 22 minutes misses 95% of action.

This data is for a relatively empty scene. In a busy scene an operator would see even less action and can virtually never detect a stationary or missing object.



Watching the scene was interesting but then it became apparent that at times the operator may want to review some past event. The system therefore had to “remember” what it had seen. Initially this was done with tape based video recorders (VCRs) and more recently with Digital Video Recorders (DVRs). Finally it was realized that these recorders could be connected up in a network enabling the person in the control room to watch many cameras in many different locations and the Network Video Recorder emerged on the market.

Right through this technological evolution a human was still required to view the video footage and to understand it. The system did not do any thinking on its own and hence all these systems can be viewed as being useful but having no intelligence.

A Surveillance system of CCTV cameras and Digital Video Recorders (DVRs) gives a user the ability to review footage that has been recorded, but how practical is a surveillance system if it is only of use after the event?

Intelligence makes a surveillance system more than a recording device: it becomes an active tool for improving the quality and effectiveness of security and surveillance operations. The primary purpose of any surveillance system should be preventative. By alerting the user to events as they happen, operators are able to understand threats and take immediate action. Intelligent systems enable security officers to react to events in real time and in fact, often to pre-empt disasters. A security system is only as good as the intelligence it provides.

Macro versus Micro Detections

The camera view is dependent on the purpose of the camera. In security applications, cameras may be placed to take a wide angle view of the whole scene thus enabling it to view macro behavior. For instance a person walking on the street may be observed to slip and fall. This is only possible if the whole street and all the people walking on it are in view – what we would call a macro view.

On the other hand there are applications which involve rather close up or micro views of images. Examples of these are in License Plate Recognition systems and for Facial Recognition. Other biometric systems such as those used for Iris Recognition or Fingerprint recognition also require a fairly close up view of the individual or object to be observed.

Macro views are generally useful for doing “detections” while micro views are more helpful when “identifying” a person or object.

Intelligence can be implemented in both types of systems.

Simple Technology Camouflaged as Intelligence

Even unintelligent systems can be made more effective with certain enhancements. For instance the traditional cameras used in security were analogue cameras which provided an analogue output. To connect into the internet these cameras had to be connected into a computer through a “frame grabber card” or with a video server. Many new cameras are IP enabled in that they have an IP address and therefore can become a node on the internet. Also they provide a digitized output which enables the image to be sent over the internet. With such cameras there is no need for video servers or frame grabber cards.

When images are sent over the internet or any other communication line the major constraint is the bandwidth of that line. Different techniques have been developed to compress the image so as to make it cheaper and faster to send the images over communications lines. The original “bitmap” format was compressed to JPEG and then to MPEG4. Each new compression algorithm makes the image “cheaper” to send, reducing the strain on a network’s bandwidth. New algorithms are being created all the time.

However compression is just an image formatting technique and does not constitute intelligence. In fact it may often be more “intelligent” to have a system that can provide alarms of events and then to transmit only the footage for those events instead of compressing and sending hours of images of a scene with no significant event in it.

Effective image retrieval is also useful. If video is stored in a data base it is important to be able to extract the appropriate footage based on some criteria. The simple retrieval systems that have been widely adopted by DVR manufacturers retrieve information based on a date/ time stamp or on event type are useful but involve simple programming and would not constitute “intelligence”.

Suppliers of such systems do promote their systems as being intelligent causing much of the confusion for customers.

Gimmicks that pretend to be Intelligence

There are also many products that are based on gimmicks. They give the client the impression that they are intelligent systems but in fact they often diminish the client's ability to protect his environment.

A prime example of this is the so called ability to use Motion Detection to pan a camera so as to "follow the intruder". When a blob is recognized in the system a mechanism in the camera is activated to follow the blob. Using this simple interface between Video Motion Detection system and the Pan-Tilt-Zoom controller of the camera, such a system appears very smart initially. However if two intruders enter the scene the camera could be zoomed in on one of them while other would not even be seen. In this sense such a system would provide a client with a false sense of security while missing a major event. Intruders who understand the technology can beat such a system just by knowing that the system is blind when it is following the first intruder.

There are better ways to solve this problem as discussed in the Section on Technology Convergence.

Rating Intelligence

In view of the considerable confusion associated with the definition of intelligence in the surveillance industry and the level of intelligence of various capabilities, iOmniscient came up with a rating system for its own portfolio of products. The criteria used to define the level of intelligence essentially related to how much "thinking" the system had to do and how difficult and complex this thinking was. This is a similar system to that used for measuring human intelligence. Just for reference, the average IQ of the human population is 100 with 50% being above that and of course the other 50% being below.

In the next section we describe the type of technologies used to do the "thinking".

Intelligence Rating	Level of Intelligence	Potential Capabilities
IQ-60	Very Limited	Video Motion Detection; Perimeter and Intruder Detection
IQ-100	Average	Object Tracking; Object Detection in relatively empty scene
IQ-110	Average	Counting in a relatively empty scene
IQ-115	Smart	Behavior Analysis (Slips and Falls)
IQ-120	Smart	Counting in a Crowd; Complex Behavior Analysis
IQ-140	Genius	Object Detection in a Crowded Scene;
IQ-180	Genius	Object Detection where the object is virtually invisible to the human eye

The complexity of the thinking also has a direct correlation with the number of companies in the market place who can provide the capability. Hence at the IQ-60 level there are around 2000 companies around the world who can do this (and the number is probably growing daily).

At the IQ-100 level (Complex Object Tracking) there are about 50 companies who can provide a robust solution.

As we get to IQ-120 there are only 3 or 4 companies around the world who can provide some of these applications. Up to this level the technology is primarily in the public domain. Some companies have managed to exploit the technology more effectively and can do more sophisticated things and some of these enhancements are protected as trade secrets but the core technology is readily available.

At IQ-140 and above only one company (iOmniscient) has the ability to do the job. At this level the technology is protected by international patents and hence no other company can provide this capability.

iOmniscient's product portfolio spans the entire range of intelligence and it competes effectively right across the spectrum. This rating system is now widely used by some consultants and security analysts as a guideline for measuring the intelligence of CCTV systems.

Types of Technologies used in Intelligence

Video Motion Detection (VMD)

At the IQ 60 level the technology used is Video Motion Detection. This involves one image being compared with the next one to see if there is a difference. If there is a difference it is inferred that there was motion. The system usually draws a blob to show the area of the change. Simple size filters can be used to eliminate blobs of a certain size. Hence the system can be made more or less sensitive. In a video environment it is possible to draw a box around the blob and follow it around. If the blob crosses a line drawn in the image it can be assumed that the object (represented by the blob) has crossed a perimeter.

The system is very simple and can be useful for applications such as perimeter protection when the conditions are benign (e.g. there are few light variations). In complex environments a system based on this technology alone can suffer severe false alarms.

At the IQ-100 level the blob that results from the Motion Detection can have other attributes attached to it. For instance, if the blob persists in one location for a significant period of time it could be interpreted as a left object. Based on this many companies claim they can do Left Luggage detection. Using this simple method for identifying left luggage is prone to false alarms due to more complex environmental factors and obscuration by passers-by.

If two people walked into a room it is possible to follow their blobs around and with technologies available in the public domain determine which blob belonged to which person. Even when the blobs cross each other, by making predictions on the direction and the speed of the blobs it is possible to keep track of the two people and follow them around. As multiple people enter the scene this problem gets more complex.

It is important to note that this level of technology only works effectively in a relatively empty scene. If a hundred people were walking around in a railway station every person would constitute a blob and these blobs would merge and split at an enormous rate. In fact there may be many moments when there is only one big blob in the scene.

Hence companies who attempt to use this technology to identify left luggage can only do it in a relatively empty scene with one or two people present. To cope with crowded scenes needs other technologies described later.

At the IQ-110 level an attempt is made to relate each person or object in the scene with a blob. Counting can be done by counting the number of blobs that pass a certain line in the scene. Again the technology works well for clearly separated individuals. In a more crowded scene one needs to be able to understand the characteristics of the person or objects represented by the blobs to separate them. Object characterization techniques described later in this paper can be used by the more sophisticated systems to get more accurate counts. Even so there is a limit to the level of crowding that can be coped with by such a system. One has to move to IQ-120 to really count in a crowd.

Neural Networks

At the IQ-115 level one begins to see an analysis of behavior. The traditional approach for analyzing behavior was to use an Artificial Intelligence technique called Neural Networks. The concept here is to let the system run for a long period of time and learn what normal behavior is, and then any behavior outside of that behavior would be considered abnormal. For example if

everyone walked at a certain speed then that would be considered normal. If someone started running that could be considered abnormal. Similarly, if everyone walked in one direction through a corridor that would be considered normal, but if someone then walked in the opposite direction that would be considered abnormal.

Neural networks have two major disadvantages. First they require huge databases and massive computing power to cope with learning and remembering the many types of behavior that are considered normal. Secondly most environments are not simple enough to allow adaptive learning. For instance if one wanted to catch a slip and fall event on 100 different cameras, then one would have to let each of those cameras run for a long period of time till each one had a statistically significant number of such events recorded. Assuming these events are quite rare for any specific camera, this adaptive learning process would take a lot of time and resources.

If the camera angle and lighting changed for that particular scene, the adaptive learning would need to be initiated again.

Understanding Object Characteristics

An alternative Artificial Intelligence used by iOmniscient which is far more robust and very light in terms of computing usage is one where one understands the characteristics of the object. Hence if one is attempting to identify a dog, rather than trying to match against a very large data base of all possible dogs taken from every possible angle, one would use characteristics of dogs. One would see that it had a wagging tail and a tongue hanging out and one would know it was a dog. That's how humans would recognize them.

The same technique is used in behavior analysis to give a highly accurate result.

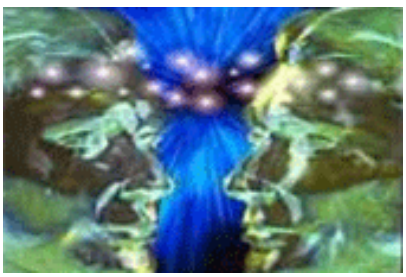
This technique is extended in IQ-120 to understand the characteristics of crowds which then allows the product to do Counting in a very crowded scene.

Understanding what the object is from its characteristics enables the system to operate with extremely high levels of accuracy.

This technique also permits the system to be able to detect very fast moving objects without unnecessary false alarms.

Heuristic Artificial Intelligence

Finally at the IQ-140 level and above a totally different type of algorithm known as a Heuristic Algorithm is required. In this case a statistical approach is used to determine an event. These algorithms are internationally patented and hence are totally unique to iOmniscient.



With these algorithms the status of each pixel is examined on a statistical basis. It is compared with the status of neighboring pixels till a picture is formed on what is happening in the scene. This technique enables the user to do Non Motion Detection and find objects that get left (or are removed) from a crowded scene. The technique also enables Nuisance Alarms to be minimized. The computer power required by this approach is very low compared even to Video Motion Detection.

A statistical approach inevitably involves creating histograms. The unique Heuristic Algorithms analyze the shape of various histograms that are created to understand the nature of the object being detected.

Combined with some of the earlier techniques this technology can provide the user with a very good view on what is happening in a scene.

Non Motion Detection ignores motion which is precisely what all other intelligent video systems are based on. Counter-intuitively iOmniscient's technology focuses on static or stationary objects. The system learns the background of a live scene in a few minutes (there is no need to stop the normal activity in the scene during the learning process). After that, any object placed in that scene or removed from it can activate an alarm in a timely fashion. The system learns continuously and updates the background even while it ignores motion.

The Heuristic approach allows the system to cope with significant obscuration from passerby as well as recognize and ignore the multitude of false alarms that can arise from light variations, water movement and other environmental factors.

Advantages

In a busy scene, where many people are in motion, Non Motion Detection provides results which traditional video motion detection is incapable of. Traditional VMD has proven very problematic in lively outdoor settings as is typical in Train Stations, Airports and public venues where one can expect crowds. Non Motion Detection provides a solution for these complex and crowded environments and has revolutionized intelligent video surveillance.

Handling Obscuration

The Non Motion Detection (NMD) algorithm used by IQ-140 and IQ-180 has the ability to handle obscuration – that is to detect an object even when people are walking in front of it.

The system can cope with hundreds of people moving around (e.g. in a busy lobby area). Not all parts of the object have to be visible at the same time (no need to stop the normal scene activities during learning). It is sufficient for the detection algorithm if different parts of the object are visible at different times, each for about 50% of the detection period.

Working with Tiny Objects

As objects in the scene get smaller it can become very difficult to detect them effectively and consistently. At the IQ 180 level the system uses all the Non Motion Detection capabilities of the IQ 140 but in fact it can operate with very tiny objects which may be camouflaged against a background of a similar color (such as a black bag left on the black marble tiles in a shadow). The system can detect objects which the human eye may not see and it can do this accurately and consistently with minimal false alarms.

NAMS - Nuisance Alarm Minimization Systems

All CCTV systems can suffer from nuisance alarms. These can be caused by light variations but also by other aspects of the environment such as trees waving in the breeze or waves rippling on the water. Even real objects can cause nuisance alarms such as birds flying across the scene (when the user is only looking for people).

The simpler systems attempt to eliminate nuisance alarms by masking out areas of the scene such that if a tree is causing the false alarms that whole area would be ignored. This has the significant disadvantage of then not being in a position to detect any real alarms that occur in that area.

The more sophisticated systems such as the IQ Series by iOmniscient attempt to understand the cause of the alarm. A human viewing the scene would ignore a bush waving in the breeze while still noticing a person walking in front of it because he recognizes the tree for what it is and ignores it while still looking for the human intruder. Similarly the more sophisticated systems would recognize the tree by its characteristics and ignore it while still detecting the human intruder. (A separate White Paper is available on the technology behind Nuisance Alarms).

Even the most sophisticated Nuisance Alarm system is only as good as its human counterpart. If a human cannot tell whether an alarm is a real one or a false one it is very unlikely that the system could do this.

All iOmniscient products have incorporated in them the sophisticated Nuisance Alarm Minimization System (NAMS) which helps minimize false alarms. If correctly set up NAMS can virtually eliminate all false alarms without affecting detection.

Perspective

A human being viewing any scene has the advantage of having two eyes and hence he is able to see perspective. This can be achieved in CCTV systems by using more than one camera to view a scene but having multiple cameras can add to the overall cost of the system. iOmniscient's products can understand perspective in any scenario using a single camera. It therefore knows that an object viewed in the foreground would look larger than if it was further away in the scene.

If an object in the foreground is moved to the rear of the scene the system can tell that although it is smaller it is still an object of the same size. Unlike other products in the market, the measurement can be done automatically using the image collected from a single CCTV camera.

Understanding perspective in a two dimensional scene is critical to understanding what an object is.

Jump to Event

The "Jump to Event" capability takes information retrieval to a new level of sophistication. In traditional DVRs it is possible to retrieve information based on the fact that an alarm occurred at a certain time. One can usually do the retrieval based on the time of the event or on the fact that some event occurred.

The Jump to Event capability goes way beyond that. It enables the user to jump back to a pre-defined time before an event or alarm occurred. If an alarm or event occurs, such as one caused by an abandoned bag in an airport, at the press of a button the user can jump back in time not just before the bag was detected but before it was first brought into the scene and abandoned. The user can immediately review both the event and the preceding period of video footage.

Further, let us consider a scene where there might be ten suspicious bags in the area being viewed. The user can click on any one of these bags and "jump back" to the moment when that particular bag was brought into the scene and abandoned.

The user then has the option of archiving the event for later review or to discard the alarm information. All the IQ Products incorporate this Jump to Event feature.

Object Recognition

Having identified object in a scene the more sophisticated systems are able to advise the customer on what the object is.

Certain situations will require the user to ignore humans and target vehicles, while others will require the user to target humans and ignore small animals or birds. iOmniscient systems have a sophisticated object recognition system that allows the operator to hone in on exactly what he is trying to detect.

The system can differentiate between people, animals and vehicles, allowing for a diverse range of customer specific solutions. It can categorize types of objects (vehicles of various sizes; adults/children).

Counting only cars while ignoring trucks (or vice versa) passing a gate is a good example of how object differentiation can be useful.

At the IQ-140 level and above the Object Recognition can get even more sophisticated. At this level a pushcart/ airport trolley can be differentiated from a bag that may have been left abandoned. In the real world the user wants to know even more. He would like to know not only whether an abandoned item is a bag or a pushcart but, if it is a pushcart, whether it is empty or loaded with bags.

In the intelligent surveillance industry no-one other than iOmniscient has such a refined Object Recognition capability. The system is capable of detecting abandoned loaded pushcarts/ trolleys in order to treat them as suspicious objects while avoiding nuisance alarms caused by pushcarts/ trolleys with no luggage on them being left in an area.

Mass Customization

Many companies have not realized the importance of customization. They have fallen into the trap of assuming that to achieve an economy of scale and hence an affordable price the product has to be extremely standard and mass produced.

However every user's problem is different. iOmniscient has pioneered the concept of mass customization. Using these techniques the company can provide every client with a system that exactly suits his needs while at the same time achieving the benefits of a very affordable price.

Technology Convergence

Many different technologies are converging together. For instance a person could be detected using an IQ 100 type of Object Tracking system and a Mega-pixel camera. The mega-pixel camera can be used to do a digital zoom to see a close up of the person involved. That image can then be used to do some Facial Recognition. The advantage of such a system is that the camera can maintain a wide angle image of the whole scene even while the image is being zoomed into.

Another example of technology convergence is where as system like that above were used to detect cars entering a car park. Combined with a License Plate Recognition system it would be able to detect vehicles stopping in a car park for longer than the permitted period, then recognize the vehicles number plates and issue tickets even while the system continues to monitor that entire scene.

Convergence is happening at all levels and a classic example is the combination of IQ 140 with IQ 100. In this case not only is one able to recognize that a bag was left in the scene. Once can then track the person who left it there.

Summary

There are many players in the Intelligent Surveillance industry and each promotes their own capabilities. It is important for a user of CCTV systems to understand the level of sophistication of the technologies offered by various players as well as to understand their own needs. These are rated using an IQ rating system very similar to the one used for humans.

A User having a very simple requirement may be quite happy with a system with a low IQ. On the other hand if they feel that their environment justifies a sophisticated technology they need to understand the technologies that are available and choose suppliers who can meet their total requirements comprehensively.

*** * ***

Please contact info@iomniscient.com if you are interested in obtaining more information about iOmniscient's ability to operate in complex, live and busy environments involving light variations and other sources of nuisance alarms. Video clips showing how its Nuisance Alarm Minimization System (NAMS) operates are available.

Other whitepapers on iOmniscient technology architecture are also available.

