A LAYMAN'S GUIDE



iOmniscient
High iQ Recognition

A Layman's Guide:
Facial Recognition in a Crowd

Published by:
iOmniscient Pty Ltd.

Copyright ©2014 iOmniscient Pty Ltd. All rights reserved.
This book or any portion thereof may not be reproduced
or used in any manner whatsoever without the written
permission of the publisher.

Other books published by iOmniscient are:
"iQ-Smart City: Security - Safety - Service", 2012
"Automated Surveillance: A Guide to Intelligent Video Analysis", 2009

TABLE OF CONTENTS

01. INTRODUCTION 8

02. FACIAL RECOGNITION TECHNOLOGIES 12
- 2.1 Facial Recognition for Access Control 13
- 2.2 Facial Recognition for Surveillance and Security 15
- 2.3 Facial Recognition in a Crowd 17
- 2.4 Facial Recognition in a Crowd for Surveillance and Security 18
- 2.5 Facial Detection 21

03. OPERATIONAL CONSIDERATION FOR FACIAL RECOGNITION SYSTEMS IN UNCONTROLLED ENVIRONMENTS 24
- 3.1 Sourcing the Image Database 25
- 3.2 Assessing Image Quality 26
- 3.3 Accuracy of Facial Recognition Systems 26
- 3.4 Impact of Camera Selection on Accuracy of Facial Recognition Systems 28
- 3.5 Other Methods of Increasing the Accuracy of Facial Recognition Systems 29
- 3.6 Dealing with False Alerts 29
- 3.7 Managing Alarms 30

04. FACIAL RECOGNITION FOR CONTROLLED ENVIRONMENTS 32

05. CRITICAL COMPONENTS OF A SURVEILLANCE SYSTEM 36
- 5.1 Comprehensive Surveillance 37
- 5.2 Open & Distributed Architecture, Infigite Scalability and Flexibility 38
- 5.3 Jump to Event 39
- 5.4 Healthcheck 41
- 5.5 Redundancy 43
- 5.6 Universal Connectivity 43
- 5.7 Scheduling 44

5.8 Auto-Archiving 44
5.9 Nuisance Alarm Minimization System (NAMS) 45
5.10 Remote Management 45
5.11 Mobile Management 45
5.12 Networking intelligence 46
5.13 iQ-Hawk 47

06. INTEGRATING WITH BIG DATA 52

07. FACIAL RECOGNITION IN SOCIAL MEDIA 56

08. FACIAL RECOGNITION IN MARKETING 58

09. USING FACIAL RECOGNITION IN A CROWD 60
Example 1: Keeping People Safe in the City 61
Example 2: Passengers without Documentation 61
Example 3: Detecting Culprits after an Event 62
Example 4: Detecting Shop Lifters 62
Example 5: Customer Service for VIP guests 64
Example 6: Tracking in a Crowd 64
Example 7: Queue Management 65
Example 8: Dwell Time Management 66
Example 9: Big Data Analysis 67
Example 10: Social Media analysis 67
Example 11: Access Control 67
Example 12: Private Security 68
Example 13: Policing Applications 69
Example 14: Driver Match 69

10. PRIVACY ISSUES 70

11. IOMNISCIENT'S FACIAL RECOGNITION IN A CROWD TECHNOLOGY 74
11.1 Development Goals 75
11.2 Patent Protection 76
11.3 Multifaceted Facial Recognition 76

12. MAXIMIZING THE RETURN ON INVESTMENT FROM YOUR FACIAL RECOGNITION SYSTEM 78
12.1 Minimizing Cost by Design 79
12.2 Having the Right Cameras and Other Infrastructure 80
12.3 Maximizing Return on Capital 80
12.4 Operational Efficiency 81
12.5 Maximizing Uptime 82
12.6 Overcoming Obsolescence 83

13. IMPLEMENTING A FACIAL RECOGNITION SYSTEM SUCCESSFULLY 84
13.1 Defining System Objectives 85
13.2 Prioritize the Objectives 85
13.3 Disciplined Implementation Process 86
13.4 Frequent Mistakes in Implementation 87
13.5 Questions you should ask your Vendors 88
- System Objectives 89
- Intelligence Levels 89
- Facial Recognition 89
- Jump to Event and Determining the Identity of the Person Involved in an Incident 90
- Automated and Mobile Response for Emergency Management 90
- Big Data Capability 90
- System Cost Effectiveness, Reliability and Efficiency 90
- Software Selection 91
- Hardware Selection 92
- Systems Integrator 92

APPENDIX:
RESULTS OF TESTS ON ACCURACY OF FACIAL RECOGNITION SYSTEMS 93

ABOUT IOMNISCIENT 97



1 INTRODUCTION

The increase in urbanization combined with continuing threats to personal security as well as the ongoing terrorism threats around the world has created a need for face recognition technologies.

Television and Hollywood movies have generated very high expectations on how Face Recognition technologies can be used. However this is complex technology which is evolving very quickly. In order to use it effectively one has to understand how the technology works, its limitations under different conditions and the best and most effective ways to use this technology. This book explains some of the background to the technology and how the use of this technology can be optimized to increase personal safety and security and to enhance personal services.

Facial Recognition can be used for two different purposes and the types of Facial Recognition technologies that are available for each are very different.

The conventional use of Facial Recognition Systems has been for Access Control (e.g. for access into a facility or for passport control). This is a form of biometric control. Such systems require co-operation from the person who is to be recognized and a



good quality, high resolution, full frontal image in a database which is compared with the image of the person at the access point. Given ideal conditions such a system can operate with a high accuracy (See Appendix). A full description of the use of facial recognition systems for Access Control is provided in the book.

The more important use of Facial Recognition systems is for Surveillance to enhance Safety, Security and Service. In these applications the technology needs to work in combination with other video analytic systems. The only alternative to the system is a human operator whose recognition skills are usually excellent for very small databases (less than 20 images). Beyond this the human ability to recall a database and recognize any one from it is close to zero.

The difference between the system used for controlled environments (such as Access Control) and that used for Surveillance is that the former operates with a high resolution image of a co-operative person in a controlled environment and compares their image with a full-frontal, high-resolution image in a database. In surveillance applications the images in the database are of variable (and usually low) quality and the person of interest is usually not co-operative.

For surveillance, a Facial Recognition system that can identify people at significantly higher accuracy than a human is very valuable. It can generate a shortlist for the human who can then make a final decision for further action. The accuracy of such a system will depend on many factors such as the quality of the images in the database. However, even with low quality images the system will keep operating and not breakdown. Its accuracy is always better than the alternative (which is the human operator attempting to recognize people on his own). A hybrid computer/ human system where the Face Recognition system provides a short-list of images to a human who makes the final judgment can generate very good results even with images in relatively low resolution.

This book describes the different kinds of Facial Recognition systems that are available and explains how a Face Recognition System in a crowded area can be used

with optimal effectiveness.

It also discusses how the system can be integrated with Big Data, Marketing and Social Media applications.

Privacy can be a major concern when Facial Recognition technology is used. Current social norms and techniques for protecting privacy are discussed. Finally, the book provides guidance on implementing successful systems.



2

FACIAL RECOGNITION TECHNOLOGIES

Facial Recognition can be used for two very different purposes and the types of facial recognition technologies that are available for each application are very different. Some of these differences are described below.

2.1 Facial Recognition for Access Control

Like other biometric technologies (such as fingerprint or iris matching), Facial Recognition for access control requires the person involved to be co-operative. The person has to be in a controlled environment so that biometric characteristics can be recorded and compared against either a single record (such as an access control card or passport – usually referred to as One-to-One Recognition) or against multiple records (such as a database of criminals – referred to as One-to-Many Recognition).

There are many companies who specialize in such Facial Recognition systems for Access Control. Although the accuracy of such systems is increasing, this technology requires a controlled environment, specialized equipment, high resolution full frontal face images and usually, a co-operative subject.

1. In a controlled environment biometric technologies such as fingerprint recognition and iris recognition can provide a higher level of accuracy than Facial Recognition. Such technologies are usually very useful for Automated Access Control, especially if the subject is co-operative.

2. High quality biometric information is rarely available for suspected persons. Information on potential terrorists for instance may be available in newspaper articles or photographs or a variety of other sources which were not obtained in controlled environments. The quality of such images is variable and they usually do not provide the detail required for the type of facial recognition that is used for Access Control.

Such systems are therefore useful for organizations for restricting entry to their buildings (since the biometric characteristics of each employee can be obtained in a controlled environment) but are not useful if one is trying to apprehend a

suspected terrorist at an airport.

The following case study demonstrates the usefulness of face recognition technology for access control. The technology cannot be compromised by humans, as may have occurred in the case below, where the access control officer may have permitted an unauthorized person to access police headquarters in Western Australia.

Charges laid after police impersonator causes security scare

WA News, By Rhianna King 5 June 2012

A custody officer at the East Perth Watch house has been stood down while police investigate whether he assisted an 18-year-old who posed as a police officer to gain entry to the facility. The 18-year-old Seville Grove man was today charged with 35 offences - including trespass, stealing police equipment and impersonating a police officer - and the incident has prompted a review of security at all police sites across the state.

The 18-year-old - who was in full police uniform at the watch house - was detected on Friday after a supervisor at the facility questioned whether he was in fact a member of the police force. The man was immediately placed under arrest and was charged with several counts of burglary, stealing, impersonating a public officer and trespass. It will be alleged he stole police uniforms, a baton, handcuffs, boots, a notebook and an access card during May.

An ongoing investigation would assess the full extent of access the man had to police facilities, but said it appeared he had no access to police computers or any access to people in custody.

The ABC reported the WA Police Union was seeking assurances from the department that no officer's personal details had been compromised.

2.2 Facial Recognition for Surveillance and Security

Facial Recognition for surveillance and security applications using surveillance systems are different to those for access control.

Information on the person that is to be identified is usually of a very poor quality. Biometric quality data is rarely available hence all the traditional biometric methods for identifying and recognizing the person cannot be used.

Further it is usually necessary to recognize the person in an uncontrolled environment and usually in crowded areas such as when he is disembarking from an airplane or passing through a hotel lobby or other public spaces.

The following case study shows how surveillance footage can be reviewed after an event to trace suspects of a crime and can result in a speedy apprehension and arrest. In this case, manual viewing of CCTV footage after the event was used rather than automated intelligent software systems.

Assassination of Palestinian Leader in Dubai

The Guardian 16 February 2010

A few years ago a Palestinian leader Mahmoud al-Mabhouh was assassinated in Dubai by foreign agents. Facial recognition systems at Dubai airport were able to recognize the agents after the event. Their entry to the country had not been refused because at that stage they had not committed their crime and they were not on any criminal database. According to the detailed but unverified account given on Monday by Dubai police, the killers began arriving in Dubai shortly after midnight on 19 January. Their movement across the city was captured on CCTV footage, much of it released by the Dubai authorities, along with hotel records and flight data they said supports their allegations.

When Mabhouh arrived in Dubai 15 hours later, police believe his assassins, who were using forged European identities, tracked him to room 230 of the luxury Al-Bustan Rotana hotel, in the heart of the city, killed him, and then departed the country. The entire operation was completed in just 19 hours, they said. However, after the event it was possible to identify the agents and to track their movements within hotels and other public areas, resulting in a quick arrest.

Members of the gang almost continuously changed their identities, switching outfits and wearing wigs and glasses as disguises, Dubai police said. The CCTV footage showed one person appeared to have undergone the most extreme transformation, disappearing from the view of a CCTV camera in a hotel lobby as a bald man in a suit, before reappearing with thick black hair and glasses. Dubai police said all 11 alleged killers flew out of the emirate shortly after. Two Palestinians believed to have been involved in the operation are in UAE custody after being handed over by Jordan.

2.3 Facial Recognition in a Crowd

The conventional Facial Recognition systems used for access control cannot work in the security environment where both the image of the target and those in the database are of a relatively poor quality (sometimes as poor as 12 pixels between the eyes).

To address this particular requirement, iOmniscient developed technology to permit Facial Recognition in a Crowd.

This technology is based on very different type of patented algorithms that work with relatively poor quality images of both the target and in the database. It can operate with images and environments where other technologies are unable to operate. The technology is designed to significantly improve the ability of a human operator to recognize a person of interest.

It is important to understand these two very different environments – for access control and surveillance security. The way the technologies are best used in each case are different and so too is the level of accuracy that one can expect.



While Facial Recognition for access control can provide high accuracy, it is often considered an expensive alternative relative to other biometric technologies available for this application.

For Facial Recognition in crowded spaces, the system needs to be as accurate as possible and adequately compensate for variations in pose, angle of presentation to the camera, lighting factors and so on. In addition, the system needs to be capable of fast performance, to detect and recognize a face in a few seconds.

iOmniscient's facial recognition systems are at the forefront of the research in this area and use advanced innovative approaches to solve these problems.

Facial Recognition, when used for security, safety and surveillance can provide a good levels of accuracy and robust results especially when they are used in combination with a human operator to make decisions and act on alerts.

2.4 Facial Recognition in a Crowd for Surveillance and Security

Systems used for Facial Recognition are generally different from access control systems. The main difference is that the person of interest may not be co-operative and may be unaware of the surveillance. The person has to be viewed covertly. The person's expressions may change. The person may never present a full frontal image. The person may be wearing a hat and may attempt to conceal his face with a hand or a scarf. The lighting may vary in the different locations where the person may be viewed.

To perform this type of surveillance, the person has to be viewed from a distance so that he or she is not aware that his image has been captured. Surveillance cameras are not necessarily High Resolution Cameras. The surveillance usually has to be performed with relatively low resolution cameras (4xCIF to 1 megapixel cameras) that are already widely used for security.

The images in the database against which this target is to be compared will usually

not comply with any international standard developed for passport pictures. They may be images taken from a newspaper or from an old family album and possibly when the person was much younger.

Most facial recognition systems for access control do not operate with accuracy in this type of uncontrolled environment. But this is the reality of the requirements of surveillance systems for safety, security and service.

To address this need, iOmniscient which specializes in addressing the problems of crowded scenes, developed its Facial Recognition technology using very different types of algorithms and which operate successfully in crowded environments.

Instead of attempting to get greater detail of the face in order to differentiate one person from another, the system extracts the maximum information from the poor quality images that are available. For example, the images in the database against which the facial images are to be compared may have a low resolution. At best there may be 40 pixels between the eyes. More often it would be around 22 pixels and sometimes as low as 12 pixels between the eyes.



Former homicide squad detective believes the CCTV footage... is the most significant piece of evidence investigators have

Herald Sun, By Wayne Flower, Anthony Dowsley 27 September 2012

Brunswick resident Daniel Gregson, who was wearing a red jumper on the night of Jill's disappearance, told the Herald Sun he called Crime Stoppers after he realized he may be able to help. Before the CCTV was released, he had also gone over in his head multiple times the journey home, to try and compare his late night walk home from the pub, to that of Jill's ill-fated walk home.

"When the CCTV was released to the media yesterday, I reviewed footage and realized that I was in frame on the CCTV. So I called Crime Stoppers afterwards and homicide detectives visited and took my statement yesterday afternoon.'' He said he spoke to the detectives for between half an hour and 45 minutes.

A man was arrested following analysis of the CCTV footage on 27 September 2012.

iOmniscient is a pioneer in the new field of many-to-many facial recognition systems. It designed its systems to operate with the low resolution images available from standard surveillance cameras and where the person may be presented at a variety of angles.

This is a very new and different technology to the one-to-one and one-to-many Facial Recognition systems that have been available for some time from suppliers of biometric equipment. While these suppliers may have many installed systems for access control, there are very few many-to-many facial recognition systems installed anywhere in the world for security applications, especially in crowded places.

Today the technology is still new and is developing rapidly. For Facial Recognition in crowded spaces, the system needs to be more accurate than a human and provide a fast response. The system must also be able to compensate for variations in pose, angle of presentation to the camera, lighting factors and so on. iOmniscient's face recognition systems are at the leading edge of research which such systems.

The alternative to using a Facial Recognition system for surveillance is to use a human operator. Humans are very good at recognizing faces when they have to compare the face with a very small database (one to twenty images). In fact at this level a human can outperform most computer systems in terms of accuracy. However as the database grows in size, the ability of humans to recall faces in the database falls dramatically and with even a slightly larger database a human would have almost zero probability of recognizing a face.

It is at this point when a Facial Recognition System can be useful to the operator. If the system can extract a shortlist of 3 to 10 potential faces and present these to the human operator, he can then use his judgment to make the final determination on whether a target matches a person in the database.

2.5 Facial Detection

Facial Detection, which is the ability to detect a face and to record it, is a valuable tool in surveillance applications. Facial detection can be particularly useful where one needs to know who was present at a certain point in time.

The system will capture the faces of people as they come into the field of view of the camera. If the user requires, all the captured faces can be stored and held for future reference (even if they are not a match against images in the database).

Image capture can be performed with an accuracy of 96% to 100% in the right conditions. The right conditions do depend on the system requirements. For instance if the system is designed to capture images for Face Recognition then faces that meet the criteria for recognition (e.g. both eyes visible, 22 pixels between

the eyes, etc.) would be the ones that are captured. If the objective is to just capture images of every person that is passing through because this may be useful information for future analysis, then it may be necessary to capture every head (even if the person has his back to the camera and the face is not readily visible).



If a significant event takes place, such as an abduction or murder, it would be important for the police and other authorities to know who was present at the scene so that they can be questioned as witnesses. A good Facial Detection system would record the faces of all people present at the event. These people may not be known and hence it may not be fruitful to attempt to match them against a criminal database. However having their pictures would help police use other mechanisms to track them down and get their help.

In an incident at Sydney airport, a gang related incident resulted in a murder within the terminal building. An ideal surveillance system which combined Facial Detection with other video analytics would have detected that a man had fallen down and then activated the Facial Detection system to record the faces of the witnesses. These faces could be automatically recorded in high resolution even if the detection system was operating in low resolution using the patented iQ-Hawk technology described later in this book.

Bikie Brawl at Sydney Airport

Sydney Morning Herald, 24 March 2009

The wild brawl between bikies at Sydney Airport in which a man was bludgeoned to death had started with pushing and shoving 15 minutes earlier in view of security officers, raising serious questions over the adequacy of airport security.

Yesterday the Federal Government ducked calls for an inquiry into security arrangements despite court evidence the fight began about 1.30 pm on Sunday and ended when the Sydney man was bashed to death on the airport floor at 1.45 pm.

The investigation hit a snag yesterday when it was discovered that the quality of the CCTV footage from the terminal was poor. The Herald understands the footage is not clear enough to make identification.

One witness told the Herald: "There was alertness among the security officers at the gate that this bunch of very big, aggressive blokes had just walked past. They were talking quite animatedly into their radios."

The Herald has learnt that the security officers are not able to contact their federal police counterparts directly but must radio their head office, which in turn notifies the police.

A triple-0 call was made from the airport at 1.43 pm and records show New South Wales police attended at 1.47 pm.

The Federal Home Affairs Minister, Bob Debus, said there was no need for an inquiry into airport security because "we have to allow the investigation to run its course without constant speculation about what occurred.

"No one can guarantee events like this will never occur. This was an event which happened very quickly and with no forewarning." He was backed by Mr. Scipione, who said: "You can't have a police officer every three feet or one meter in a terminal."



3

OPERATIONAL CONSIDERATIONS FOR FACIAL RECOGNITION SYSTEMS IN UNCONTROLLED ENVIRONMENTS

The operational steps in a Facial Recognition system are:

1. Capture the images of the faces of people as they approach the camera (if the user desires one can store these faces for future use). The face detection should be done in a covert manner so as not to require the person concerned to consciously stand in front of a camera to have a picture taken. One can expect an accuracy of between 96% and 100% for this exercise with the right combination of camera, lens and lighting. While choke points may be used in an attempt to capture every face in a scene, it is more common to have people randomly moving in the scene

2. Compare the captured image against a database of images which have been previously enrolled into the system from different sources.

3. Raise an alert at a work station in a security room (which may be remote) to enable the officer-in-charge to make a determination on whether the alert requires further investigation.

3.1 Sourcing the Image Database

For good Facial Recognition, it is not only important to capture a reasonably good image from the camera but also to have good quality images for comparison in a database.

For high accuracy, these images too should have at least 22 pixels (and preferably 30 pixels) between the eyes. However, the system can attempt (on a best effort basis) to perform recognition even on images of a much poorer quality.



The database of images for comparison can be from

multiple sources. It should also be possible to compare the captured image against different databases independently and simultaneously.

Note that images of the quality described here is a key feature of "many-to-many" systems. Access control systems, which are "one-to-one" or "one-to-many" systems, usually require images with more than double the number of pixels between the eyes (and often 4 or 5 times more) to achieve high levels of accuracy.

There are several methods for increasing the accuracy of recognition. If the general accuracy of recognition is 50%, then having multiple images of an individual in the database can improve accuracy to higher levels, for example to between 70 and 80%. If it is possible to have a video clip of the person in the database the accuracy can be improved further.

3.2 Assessing Image Quality

iOmniscient has a tool for assessing images as they are being registered in the database which will advise whether a particular image will be useful for accurate recognition. This will depend on the haziness of the image, its resolution, the angle at which the person is looking, whether all parts of the face are clearly visible and so on.

Images from newspapers, magazines or videos can be used to populate the database but obviously the quality of the source will dictate the accuracy with which the recognition can be performed.

3.3 Accuracy of Facial Recognition Systems

Different levels of accuracy can be achieved in one-to-one systems, one-to-many systems and many-to-many systems.

In a one-to-one system which is used for Access Control, Passport Control and other similar systems, one can expect accuracies of up to 99.9% (See Appendix 1). This is because both the capture of the person's image for recognition and the image held



in the database are taken in very controlled environments and with a pre-defined resolution. The person is usually required to stand in a particular position right in front of a high resolution camera and the image for comparison is on an access card or passport or in a central database where the image is pointed to from the card.

In one-to-many systems for access control, the person is still required to stand in front of a camera in a controlled environment. The images in the database also have to be taken in a very controlled environment.

In a many-to-many system, the controls on both parts of the system are lower. Both the image capture for recognition and the images in the database may be in relatively uncontrolled environments.

The relative accuracy of each recognition is dependent on whether the person is seen reasonably clearly, the quality of the image in the database, the number of images in the database and a number of other environmental factors.

For a Facial Recognition system to be useful for surveillance, the system must be more accurate than a comparable human based system because there are no other

alternative technologies. For Facial Recognition in crowded spaces, the system needs to be as accurate as possible and must adequately compensate for variations in pose, angle of presentation to the camera, lighting factors and so on. In addition, the system needs to be capable of fast performance, to detect and recognize a face in a few seconds.

In a many-to-many system, a human would normally not be able to achieve an accuracy of more than 1 or 2%. Tests have shown that a human would be lucky to pick an unknown face in a crowd with a probability of more than 0.001%. In contrast the iOmniscient system will deliver accuracy that is many times higher which gives the security officer a reasonable chance of apprehending the right person without a special purpose camera. The actual accuracy achieved by the system is dependent on the several factors that have been previously discussed.

3.4 Impact of Camera Selection on Accuracy of Facial Recognition Systems

The issue of using standard security cameras rather than special purpose ultra-high resolution cameras or multiple cameras which generate a 3D image is critical because this determines the cost of the system. It is always possible to improve accuracies slightly by using specialized equipment however this can have an extraordinary impact on the cost of the system. This is because most organizations have many security and surveillance cameras installed for purposes other than face recognition. In the ideal scenario these cameras should also be used for Face recognition, thereby reducing the cost of the system and improving its efficiency. The user should always use this question as one of his selection criteria – "Does the system require specialized proprietary cameras?"

At the same time it is important to recognize that not all security cameras are suitable for facial recognition. However the selection criteria should require that the system should be able to operate with cameras from multiple suppliers as long as they meet certain minimum specifications.

3.5 Other Methods of Increasing the Accuracy of Facial recognition Systems

There are other methods for increasing the accuracy of facial recognition systems. For example, one way to increase accuracy even with covert surveillance is to find a way to distract the person so that the person will inadvertently look towards the camera. To achieve recognition, both eyes of the person must be seen. One method of achieving the distraction is by having a monitor positioned close to the camera with some interesting pictures or information. This would draw the person's gaze towards the monitor and hence in the direction of the camera.

Other means are to install a turnstile or pinch point in the path taken by the crowd which forces them to slow down and/or to pass through a barrier in single file where a camera is placed above to capture images more accurately.

Faces in uncontrolled environments may be partially hidden. The person may be wearing a scarf or he may be holding his hand over his mouth or shading his eyes. Such factors can prevent the system from recognizing accurately. A human would fare no better in such a situation.

3.6 Dealing with False Alerts

No discussion on accuracy is complete without a statement on false alarms. As the accuracy of recognition increases, so does the potential for false positive results. Higher levels of accuracy are possible if a higher rate of false alarms can be tolerated. Ultimately there is a trade-off that the user establishes between accurate recognition and false alarms that ensures that the system is practically useful.

In the end, the system is useful if the accuracy it can provide is better than that which can be achieved with an alternative system (e.g. a human system) in a cost effective manner.

3.7 Managing Alarms



In low risk environments it is feasible to have an automated system where a facial recognition system provides an alarm which is then recorded such that it can be reviewed at some later time by the security officer. In an airport, however the whole process needs to be completed in real time to ensure that a suspect does not disappear into the crowds outside before being apprehended. We also need to increase the probability of detecting someone who may be partially disguised.

To achieve this, the security officer must be used to make the final judgment on recognition. If a suspect is detected, the person's image is shown in the monitor or mobile device of the security office immediately along with the matched image from the database. If there is doubt about the recognition the system can also show details of best matches from the database that offer the closest matches to that person. This permits the security officer to use judgment to recognize the person with a higher level of accuracy than possible without such a system.

Generally, a maximum of 5 images are sent to the officer to choose from in situations where there may be doubt. With very large databases the accuracy of recognition can be increased by providing a larger number of suspected matches – even 10 or 20.

A human can make a good judgment in terms of recognition against 10 images while good recognition cannot be done against 1000 images (or even a 100).

The system can produce a report of all the recognition information for a specific period. The default settings of such systems only hold information about those persons who are accurately recognized with a high level of confidence. The system can be customized to show and store the results of the best matches of those who act suspiciously.



4

FACIAL RECOGNITION FOR CONTROLLED ENVIRONMENTS

All systems for facial recognition involve an analysis of certain characteristics of the face. The suppliers of Facial Recognition systems for controlled environments such as for access control, require very detailed images of the face. The accuracy of facial recognition systems increases with the level of detail that is available; this means that very high definition images are required.

Highly specialized equipment is needed to obtain these high definition images. The image below, of the a typical passport control system, shows how a person is required to stop at a precise point where the light conditions are well controlled or walk through a specially lit gate to ensure constant lighting.



Typical set up for capturing high quality image for use in a database for Access Control. Note that the subject always has to be co-operative.

The image database against which the person's image is to be compared also has to be of a high quality. It has to be a full frontal image usually with at least 300 pixels between the eyes. For passport control the image database consists of passport photographs which are compared with the image of the person at the gate.

Most importantly, the facial recognition for these access control systems are

one-to-one face recognitions, which means the image on the passport is compared to the image taken at the access control gate.

In addition, a one-to-many identification may be undertaken where the image of the person at the access control point is compared with a database of images, of say employees in the company, or known persons of interest at an airport.

Facial Recognition for Access Control	Facial Recognition for Surveillance
Target	
The target will be aware that his image is being captured.	The image may be captured covertly.
The target must co-operate in presenting a proper full frontal, emotionless view to the camera.	There is no restriction on the target. The target is normally not aware of the presence of the camera.
Image Database	
Images must have high resolution, typically at least 90 pixels between the eyes and preferably 300.	Images can be of any resolution with good results achievable with around 22 pixels between the eyes. The minimum resolution can be as low as 12 pixels.
Images must be relatively recent to avoid "aging". Images must be taken in a light controlled environment.	This system is less sensitive to the age of the image. Images may be taken in different lighting conditions.
The images should be full frontal.	Any image is acceptable as long as both eyes are visible.
The person in the image must not show undue emotions or pull a face.	There is no restriction on the emotions or facial expressions of the person in the image.
The person must not wear a hat.	Hats, other headgear and even spectacles do not affect the accuracy of the recognition as long as both eyes are visible.
Camera Used	
The camera is usually restricted to a megapixel camera generating a single image in a format such as MJPEG.	The camera may be any security camera (even a D1 camera), as long as it meets certain functional specifications. Any industry standard protocol is acceptable.
Camera Placement	
The camera must be placed directly in front of the target.	The camera may be placed at a height as with most surveillance cameras.
The camera must be placed close to the target.	The camera may be placed at a distance - the maximum distance being determined by the resolution and the lens of the camera.
Accuracy	
To be useful, the accuracy has to be better than the alternatives namely fingerprint or iris recognition.	To be useful, the accuracy has to be better than the alternative which is a human operator.
The target will be aware that his image is being captured.	The image may be captured covertly.



5 CRITICAL COMPONENTS OF A SURVEILLANCE SYSTEM

A camera based system which is used for surveillance and security, must have certain key architectural and functional characteristics. These are described below.

5.1 Comprehensive Surveillance

For instance, there may be no requirement for facial recognition of every person who passes by a particular camera. It may be more important to trigger the recognition when a person exhibits certain behaviours. These behaviours may be individual actions such as running or loitering or it may involve a combination of behaviours such as a crowd gathering where someone has fallen down. Such events can be detected by advanced detection systems such as those available from iOmniscient.

Facial Recognition in a crowd system should be an integral part of this comprehensive detection and identification capability.

iOmniscient's Detection in Crowded Scenes



Can detect an abandoned object in a crowded scene with significant obscuration and long detection times (minutes / hours). Internationally patented.



Accurate counting even in crowds and with non-overhead views.

Other Systems



Can detect left Objects in a relatively empty scene with very short detection times



Simple counting. Requires overhead view and relatively inaccurate. Unable to cope with groups.

Examples of iOmniscient's detection capabilities relative to others.

It is important that these behaviours are recognizable in a crowd. Whereas some systems can detect an abandoned bag in an empty scene, it is important to be able to detect such objects in a crowded area.

While some systems can count individuals as they pass a camera one at a time, it is important to be able to count number of people in a crowd. The images demonstrate the difference between the simpler detection systems and the more sophisticated ones available from iOmniscient.

Since many events such as terrorist actions take place in crowded scenes and public spaces, the ability to perform both detection and facial recognition in a crowd are critical elements of a comprehensive intelligent surveillance system.

5.2 Open & Distributed Architecture, Infinite Scalability and Flexibility

A good comprehensive surveillance system must be designed with totally open architecture which means that the user can choose any supplier for the camera and computer hardware. The user must be free to purchase such hardware from any supplier as long as it meets certain minimum specifications.

The system must also be scalable, i.e. its speed and accuracy should not be affected by the number of cameras required for the application.

The system should also be able to operate in a distributed configuration, which means it should be possible to distribute different parts of the system to different locations to optimize network and storage usage.



The system should have infinite scalability and the flexibility that can come from being widely distributed rather than centralized.

5.3 Jump to Event

The old method of using Post Alarm video review to determine what happened during an incident has become obsolete because in today's world the incidents themselves are becoming much more complex.

Information retrieval is a key part of an intelligent comprehensive surveillance system. When an event occurs, it is critical for the operator to be able to immediately find the video footage for the event when it first started. It is possible that the event started some time back. For instance a person abandons a bag and walks away. Several hours later the bag is discovered. The security officer needs to understand who abandoned the bag and when the event occurred.

The system must have a "Jump to Event" capability that goes back (at the click of a single button) to the start of the event. In the above case the system would bring up the video of the person with the bag entering the scene. This ability to retrieve information about an event at very short notice with little effort is important in ensuring a fast response.



iOmniscient's "Jump to Event" capability ensures that the required information can be available in real-time. The "Jump to Event" capability takes information retrieval to a new level of sophistication.

As an example, an airport had a bomb threat. The authorities had to take it seriously and hence were forced to evacuate the airport while they studied the videos to determine if indeed there was an object abandoned somewhere. There were several hundred cameras and the authorities had to search through many hours of video.

The exercise was an extremely costly and time consuming one. It turned out to be a hoax.

A comprehensive video analytic system would have been able to advise the security management of every such event on any camera. And for each such event, the operator could have retrieved all the information on who brought in the object at the press of a button.

If such a system included facial recognition, capture of the image of the person's face and a comparison with images in a database would have also assisted with recognition and apprehension.

Even if the person's image was not in the database, the face image would be available for distribution to the police.

In simpler systems a few seconds of video may be retrieved showing the moments just before and just after the event.

iOmniscient has developed a unique Jump to Event capability associated with its patented non-motion detection system that goes way beyond this. It enables the user to jump back to a pre-defined time before the event started. Hence in our example of the abandoned bag, 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 and gain important information on who brought

the bag and the person's current location if he is still within camera view.

Now consider a scene where there might be ten suspicious bags in the area being viewed. The Jump to Event function should enable the user to 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 discarding the alarm information.

None of the simpler data storage and retrieval systems has such a capability as it is very dependent on the intelligence required to find the bag in the first place. Only a system that uses the patented non-motion detection can actually have a sophisticated Jump to Event function similar to the one described here. And this is the domain in which iOmniscient has got international patents.

5.4 Health Check

In any large system about 20 percent of the cameras tend to be out of operation for various reasons. Unfortunately the user does not know which ones are not operating and can only be reactive when such cameras are discovered.

In a recent case of the theft of the Van Gogh painting from a Cairo Museum, it was discovered that the cameras were not operational after the event.

An intelligent surveillance system should have a comprehensive capability to check on the operational availability of every camera that is installed on the system.

iOmniscient has a Healthcheck technology which provides such a capability.

This technology makes the operator aware immediately if any camera is not operational for any reason. If the camera is sabotaged, is moved or if it goes out of focus, the operator receives an alert.

A good system which checks on camera availability may go beyond advising that a particular camera is not functional but also provide an alert if the camera is unable to "see", for example due to fog or heavy rain or due to sabotage. Such a system will warn the operator if the cameras cannot adequately see what they are supposed to see. This capability greatly increases the productivity of the users of the system.



Spectacular Art Heist in Paris – Surveillance Cameras not working

Reported in SchadenSpiegel, Munich Re, 21 May 2010

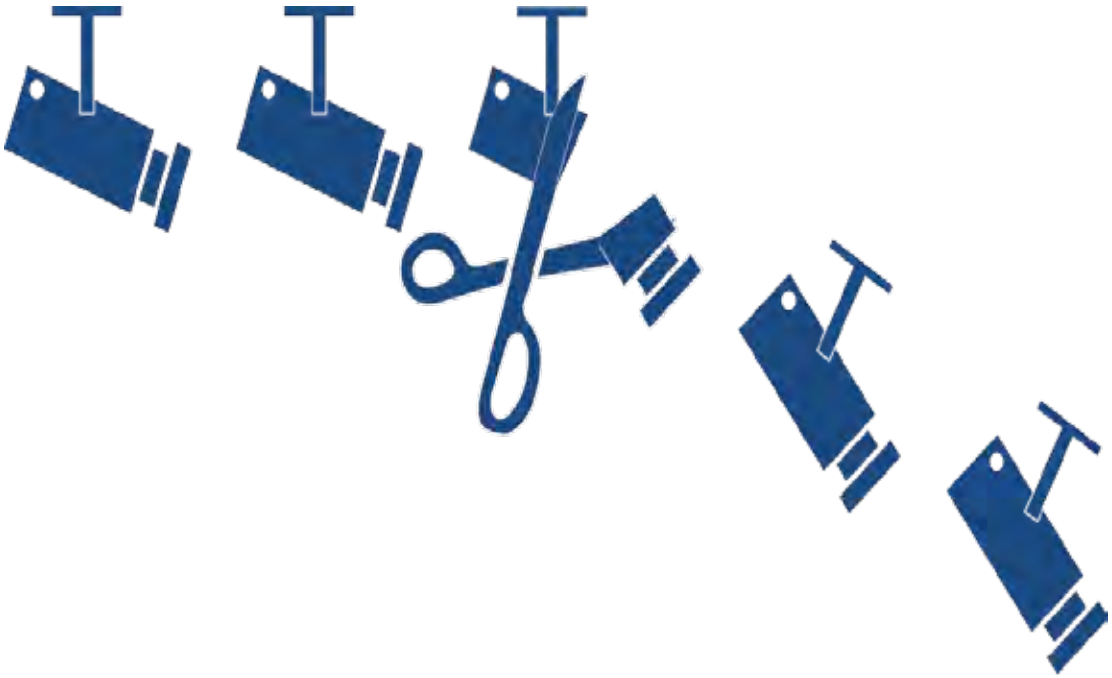
On 20 May 2010, a lone thief broke into the Museum of Modern Art in Paris and stole five masterpieces by Picasso, Braque, Modigliani, Léger and Matisse. The deputy culture secretary at Paris city hall put the value of the works at €100m. The Picasso alone was worth €23m and the painting by Matisse around €15m.

The burglary was discovered early on the morning of 20 May. Museum staff established that a window had been broken and a padlock smashed. A section of the criminal investigation department specializing in art thefts conducted the investigations. As the paintings are effectively un-saleable, the police assume this to be a " theft to order".

It was also reported that the surveillance cameras were not working. After apprehension, the thief admitted that the "theft to order" was for one painting but he was tempted to steal five paintings because of the ease of theft and lack of surveillance.

5.5 Redundancy

Redundancy is an overused and misunderstood term. Often users will build significant redundancy into their computing environments while ignoring the most likely critical points of failure. These critical points of failure are often the cameras and the network rather than the computer.



A truly reliable system will have two features. It will have a Healtcheck system which can advise the operators when some part of the system has failed.

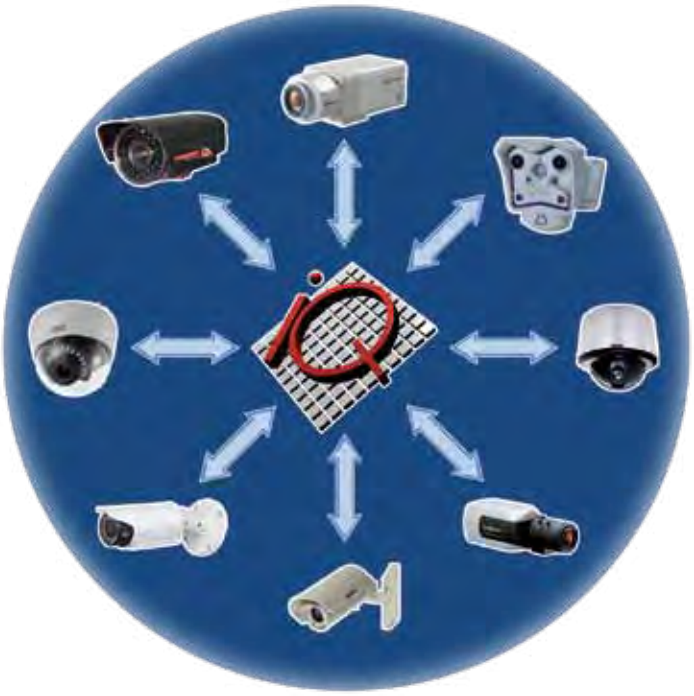
It will also have built-in redundancy for those aspects of the system which are more prone to failure and as noted before, one has to look beyond the computers in designing for redundancy.

5.6 Universal Connectivity

Many suppliers attempt to maintain their business by developing products that are proprietary. Once the user has bought a particular product, they are locked into

that supplier for a long time.

The system must have a universal interface that will allow its software to connect with any camera from any manufacturer as long as it meets some very basic standards.



5.7 Scheduling

In a complex environment, a camera may be required to perform different functions at different times or indeed to be set up with different configurations at different times. For instance from 9am till noon the system may be used for counting but from then on it may be required for intrusion detection including the ability to recognize the face of the intruder. The ability to schedule different functions in a robust manner is critical for the effectiveness of a good system.

5.8 Auto-archiving

Normally if an event occurs it should be brought to the attention of the operator who can make a decision on whether to archive that particular footage for later review. If there is no operator available it should be possible to set the system up so that it can automatically archive any event footage. The amount of time that lapses before the system realizes that there is no human around to intervene and automatically archive events should of course be configurable by the user.

5.9 Nuisance Alarm Minimization System (NAMS)

The minimization of False Alarms is critical for any video based system. Every iOmniscient product is armed with an Artificial Intelligence based NAMS module. In a recent comparison test against a major US competitor that lasted over a week, both sides achieved similar accuracies for detection. However, iOmniscient was differentiated by its ability to cope with the false alarms. The competitor was getting 200 false alarms each night compared to zero from iOmniscient.

Of course it is not possible to eliminate every single false alarm and humans themselves can often be fooled by light changes and mirages. However, a good NAMS module is critical for any good Video Analytics System.

5.10 Remote Management

The control room for a surveillance system may be located in fairly remote locations. It may not be possible to locate the required resources locally.

For such situations the iOmniscient system has been designed to enable implementation, diagnostics and maintenance to be performed remotely.

The remote access capability can be used for more than diagnosis of problems. It can be used for configuration when the system is being implemented.

5.11 Mobile Management

In many situations it may not be possible to continuously monitor the system. To ensure that people who need to receive information on events receive this immediately, iOmniscient systems all come with Mobile Client systems that operate on Smart Phones. So the operator and senior management can monitor events even when they are not inside a control room.

5.12 Network Intelligence

So far we have only talked about intelligence as it relates to the images from a particular camera.

The next generation of video management goes beyond the analysis of single images. It involves using the information from multiple cameras to provide the whole network with intelligence.

The best way to visualize network intelligence is with an example. Consider a large theme park. At such a venue they have to manage long queues for their various rides and activities. Some of these queues are very long. They could, for example, start at the entrance of one building, wind through underground corridors and emerge at a different building. No single camera can see every part of the queue. There are often many entrances to the queue and there may be many points where people leave the queue (perhaps out of frustration).

In order to manage these queues, management requires information like the Average Waiting Time for a person who enters the queue. They also need to know where the queue ends or the length of the queue, - especially when all parts of the queue are not visible.

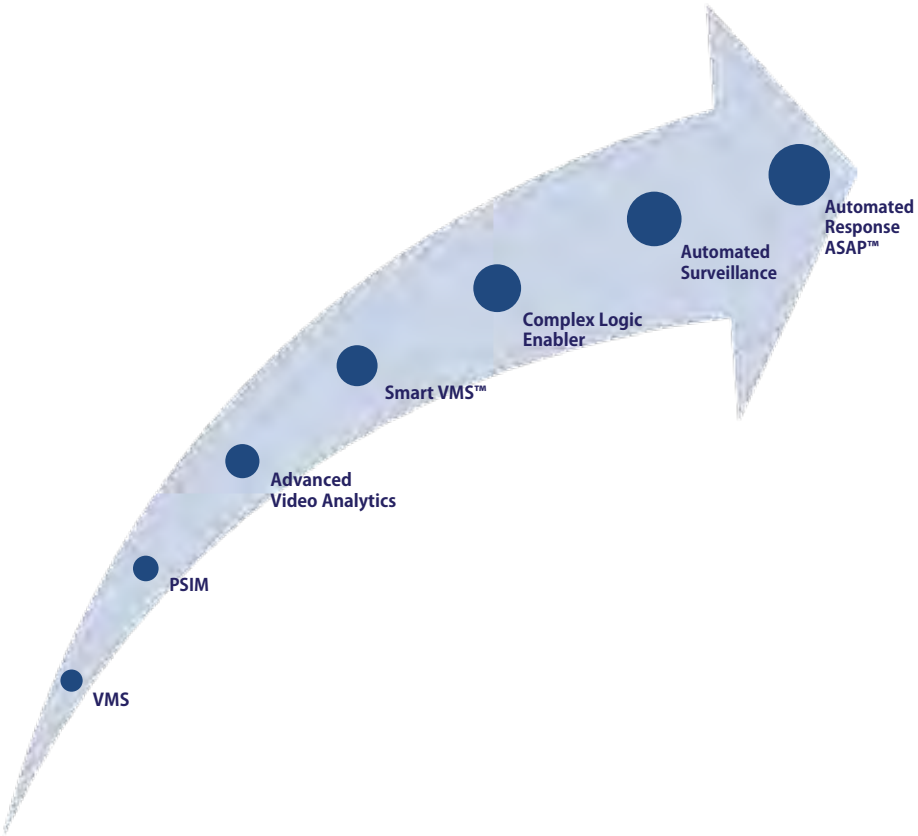
This is a very good example of an application that requires intelligence that goes beyond the single camera. This type of application requires cameras to be placed strategically at all the entrance and exit points for the queue. Every camera is then used to count the number of people that pass the point. The cameras is also used determine where the queue ends. The information from all the cameras is pooled together to provide the information that the park management requires. This allows management to open up more service points and to put up electronic signage advising customers of the expected waiting time for their ride.

Some suppliers of video analysis systems are still focused on analyzing the information from a single camera. Network Intelligence is only available from the most sophisticated providers of video analysis.

5.13 iQ-Hawk

When the industry first started it focused on recording videos. As recording became more distributed this technology was called VMS (Video Management System).

Using the video to analyze and detect certain types of events became known as Video Analytics. In the last few years the most sophisticated video analytics companies also provide recognition systems for vehicles and faces.



iOmniscient is already recognized as the technology leader in Video Analytics. However, it has taken the technology several steps forward.

First it introduces the concept of a Smart VMS. This patented technology enables the system to store important information such as a face in high resolution while the rest of the image is stored in low resolution.

Next, its Complex Logic Enabler allows the system to understand a situation better by understanding the sequence and combination of events.

The patented Automated Surveillance capability allows for the automatic recognition of persons or vehicles involved in an accident and finally the patented Automated Response functionality enables the system to find the nearest first responder and send him information about the event.

Conventional techniques for recognizing a person

In the past, Pan Tilt and Zoom (PTZ) cameras were used when one attempted to do two things at the same time. The camera would monitor the total scene. Let us say an event occurred – someone jumped over a fence and entered the area that is being monitored. The PTZ camera could then be zoomed in onto the intruder and with the close-up view he can be identified.

Initially this maneuver was done manually. Once detection occurred, the operator would manually zoom in on the individual who had jumped over the fence. It was soon discovered that manual zooming was prone to error. It was very difficult for an individual to accurately zoom onto the target. Even for a skilled operator it was not easy to make the camera close in directly on the targeted person. The task is equivalent to taking aim with a rifle sight. A slight movement of the camera meant that the person is not in the view.

This method was so difficult to use that when a capability was announced for cameras to automatically zoom in on their target, this was readily embraced. With this capability, the camera would detect an intrusion and then it would zoom in the PTZ camera to the co-ordinates of the intruder.

This method has one very significant limitation. It is easy to defeat by anyone who is familiar with PTZ cameras. A decoy can be sent in. The camera will zoom in on the decoy and the real intruder can enter the scene from the other side.

The next step in this evolution was to use two cameras. A fixed camera was used for

detection. If an intrusion occurred, a separate PTZ camera was zoomed onto the intruder to perform identification. However, the first camera would continue detecting and if a second person intruded he too was detected. Unfortunately the second intruder would not be identified as the PTZ camera would be busy with the first intruder.

The PTZ camera was usually programmed to follow the largest object in the scene. If a bird flew across the camera view, the PTZ camera would follow the bird (because it is closer and appears larger) and it would ignore the intruder.

Many users of video surveillance systems have rejected the use of PTZ cameras for this type of detection and zooming because of these limitations.

Automated Surveillance using Higher Resolution cameras

A really effective solution has evolved through the introduction of a totally different technology – mega-pixel cameras. As mega-pixel cameras became more widely used, it was evident that by recording an image at a resolution of say 2 mega-pixels, the same image could be used for both detection and identification. Multiple detections and identifications can be done on the same image view. However, mega-pixel cameras have their own disadvantage. They are incredibly expensive to use because of the computing power required to process the images and the high bandwidth required for transmission. And the large, dense images are expensive to store.

The standard camera image with around 384 x 288 pixels is equivalent to a 0.1 mega-pixel image. Using a 1 mega-pixel camera requires 10 times the computing power over that of the standard camera. A five mega-pixel camera requires 50 times the computing power. One would also require 50 times the storage capacity and 50 times the bandwidth for transmitting images.

iOmniscient finally solved the issue of the convergence of detection and identification technologies using a patented approach which it called iQ-Hawk.

This revolutionary technology involves using a mega-pixel camera. It optimized the way detection and identification is performed at different resolutions to achieve several significant goals.

iQ-Hawk allows detection to be performed at 1 x CIF. When an event occurs (such as a person seen abandoning a bag or falling down), the system provides an automatic Digital Zoom to allow the system to capture the face of the person involved.

This technique has the advantage that it permits the user to receive high resolution images but only when there is a significant event of interest. This results in a huge saving in bandwidth and storage. When using a 2 megapixel camera the saving is 20 times. In other words, if the cost of the transmission was going to be $20, using iQ-Hawk it would reduce to $1.

Thus with this breakthrough technology:

1. The system cannot be fooled by decoys.

2. Unlike standard megapixel camera based systems it does not require massive computing, storage and networking resources. (When using a 2 mega-pixel camera the system requires 20 times less computing resources than an equivalent system processing images in the traditional way).

3. It can perform all types of behavior analytics and identification (both License Plate Recognition and Facial Recognition) at the same time on the same scene using the same camera.

And all this is performed automatically. If 10 events occur concurrently in the scene the system will recognize the people and vehicles associated with all of them.

At last, detection and identification can finally be achieved on the same camera in a commercially viable way. And it can be done such that it cannot be defeated by someone with a cursory knowledge of how these systems work.



iQ-Hawk (above) enables detection of multiple events and identification of those invilved on the same camera. A PTZ camera (top) could not be used as it is easily defeated by the use of decoy.

iQ-Hawk reduces the overall cost of the system by reducing storage and network bandwidth requirements by a huge factor.



6

INTEGRATING WITH BIG DATA

Big Data is an emerging technology where unstructured data from many systems can be accessed and manipulated to draw out useful information. Previously this has only been possible with text data. Meta data (which essentially means the data about the data) from videos and other sensors converts information contained in the images into text.

Meta data is like the envelope which carries the address of the recipient and sender and the postman. The data is analogous to the contents of letter. Assembling envelopes to a particular person or organization can convey useful information even if the data is not available.

The task is difficult because most meta data is obtained from relatively unintelligent systems and hence there can be a large quantity of meaningless data to be manipulated. iOmniscient's systems generate what is called **Meaningful Multisensor Meta Data (MMM or M3)** which is meta data resulting from pre-processing video to understand what is actually happening in the scene. The more intelligent the Video Analysis, the more meaningful the Meta Data. iOmniscient prides itself on having the most advanced and intelligent video analytics in the industry.



The analysis of Big Data is not currently performed by iOmniscient. iOmniscient works in partnership with the major companies offering Big Data analysis to achieve these objectives. This data can be analyzed to provide further intelligence for organizations to forecast future requirements and plan and optimize their resources.

For example, incident data from traffic monitoring systems can be used to plan and develop road networks, forecast public transport requirements and manage traffic flows. This can also enable cities to be smarter, by reducing congestion, reducing noxious emissions and improving the quality of life for residents. It will be possible for cities to levy congestion taxes and estimate expected revenues based on forecasts of traffic flows in various sections of the city at different times of the day and different days of the week.

Crowd management information can be used to plan the patterns of crowd flows by day of the week, time of day and even seasons of the year. This can be used to improve security, provide better facilities such as public toilets and plan public transport by re-locating bus stops, for example.

There are organizations that can extract meta data from video. Indeed the extraction and the provision of such meta data is enshrined in the ONVIF standard, being developed by a consortium of organizations involved in different aspects of the manufacture and use of CCTV systems. The difference between organizations is based on the quality of the analysis that is performed to generate the meta data. For instance, if the analytics software that generates the meta data is only able to detect motion, it will generate a message every time something moves (even though the movement may only be a cloud or a tree waving in the breeze). Such a system would not generate any meaningful information, especially if the scene is crowded as everything in the scene would be moving. There would be little analysis of what is happening in the scene and the meta data would be large in volume but not meaningful.

A more sophisticated system should be able to understand particular behaviors such as someone falling down and pass such meaningful information on to Big Data systems.

iOmniscient's systems can provide Big Data engines with inputs that are meaningful and represent the results of intelligent analysis. Further since the iOmniscient system can analyze video, sound and smell, multisensor information is now available for Big Data analysis.

The multi-media Big Data from iOmniscient is a breakthrough for companies that have not been able to previously access information from these sources.

When humans attempt to understand their environment they use all their senses. Similarly using vision, sound and smell allows the iOmniscient system to better understand the environment and to provide better inputs to the Big Data system. The old adage "garbage in, garbage out" is still relevant in the world of Big Data analysis. By helping to improve the inputs of the system iOmniscient helps to improve the results.

The type of questions that can now be answered include:

- We know Fred and Jack are criminals. How often have they been seen together?
- Identify all the people Fred has met over in the last week.
- Show all the incidents of people falling down.
- Show all incidents where people fall down after a gun shot was heard.
- Show all incidents of smoke or fire in a railway tunnel.
- Show all incidents of smoke or fire in a railway tunnel caused by electrical faults as determined by the smell of burning insulating plastic being detected first before the fire.

There are many sources of information that Big Data users have access to. In certain countries various police and government agencies have access to social media. They are able to collect information from social media sites and make connections such as A is a friend of B. To date they could only analyze the information that is in text form. With iOmniscient's ability to provide them with Meaningful Multisensor Meta Data (MMM or M3) they access information from pictures and video. For instance, they can recognize a person. If that person is with someone else they can recognize their friend and so on.

This raises serious issues with respect to privacy. However for those agencies authorized to deal with such information, the technology is now available.



7

FACIAL RECOGNITION IN SOCIAL MEDIA

It was mentioned earlier that Secret Service agencies in some countries are able to collect data from Social Media. A more common application would be the use of social media for private purposes.

Individuals may want to check out if their friends have uploaded pictures of themselves on their sites such as on Facebook. They may want to know if they are portrayed in a video in applications such as YouTube.

They may want to find older pictures of themselves and their friends from their own records. The applications are numerous and only limited by the imagination of the user.

Today, Facial Recognition applications are available from several social media companies but these are only available for high resolution images taken in relatively controlled environments. The iOmniscient Facial Recognition system allows these applications to operate with low resolution images in uncontrolled environments where the person may not be looking directly at the camera or where the lighting may be low. See example in chapter 9.





8

FACIAL RECOGNITION IN MARKETING

Marketers have an interest in getting information on the effectiveness of their advertising programs. For this they do not need information on specific individuals but rather on how people generally react to their advertisements and also on traffic flows through various shopping areas.

Understanding statistics on traffic flows within a shopping area can be achieved with video analytics based counting systems. These systems can provide information on how many people enter a store and how many travel in different directions. However, such systems cannot provide information on the specific path that individuals take. For instance, the counting system can tell you that 60% turn left on entering the store and 40% turn right. Such a system will not tell you that on average an individual is first going to stop at the vegetables and then spend 10 minutes in the biscuit section before moving on to the chocolates.

This latter information can be achieved using Facial Recognition where individuals are automatically "enrolled" in the system as they enter the store. The system does not know who they are but they can be given an identity number. As they move through the store they can be recognized and it is possible for the system to establish exactly where the person has travelled and how long is spent in each area.

Marketers like to understand whether people looked at their advertisements and posters. A counting system can tell how many people passed a particular poster or location. A Facial Detection system could tell how many people turned and looked at the poster.

Marketers also like to understand demographics. Determining the sex and age range of an individual is now possible and easier than to actually recognize the face. This is valuable information for the marketer. Based on this type of information, the marketer can determine what products should be promoted in a particular space and also whether the promotions are reaching the intended target audience. See examples in Chapter 9.

Multiple Faces

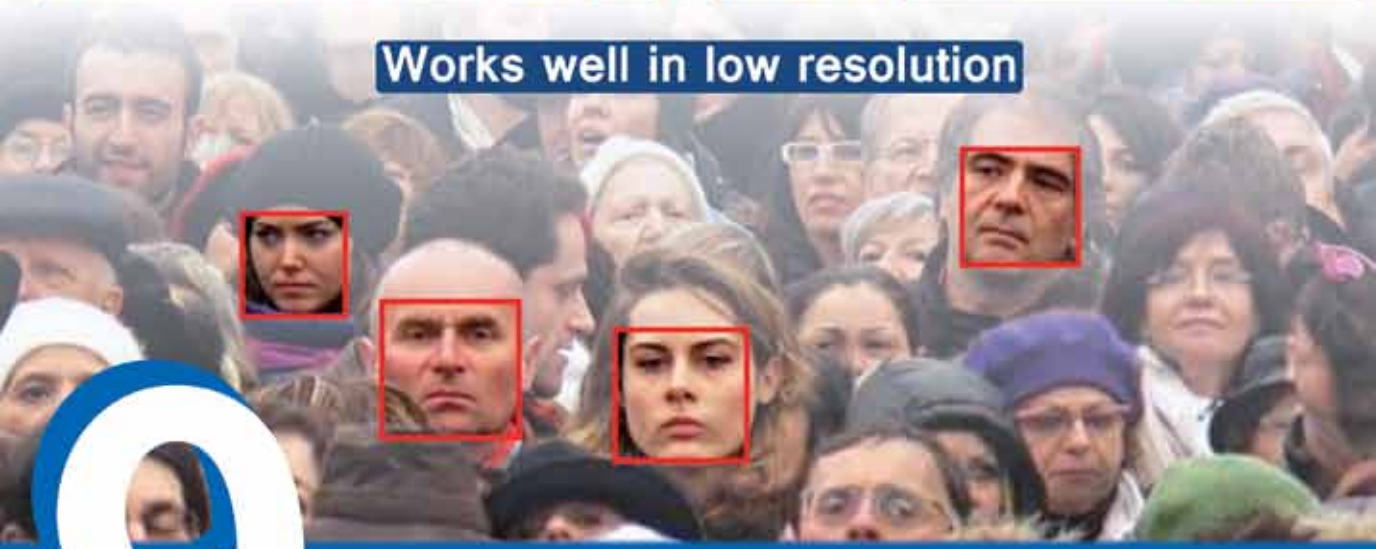
Distant View

Wider Face Angle

22 Pixels between the eyes

Convergence with iQ-Hawk

Works well in low resolution



9

USING FACIAL RECOGNITION IN A CROWD

There are many applications for Facial Recognition in a crowd.

As will be seen from the examples below, these are very different from the access control and passport checking applications.

Example 1: Keeping People Safe in the City

The most important focus of city surveillance systems is on the people themselves. The objective is to keep people safe and prevent accidents or criminal activities. Not all activities are easy to detect. For example, it is very difficult to detect if two people are fighting as this can take many different forms. However many associated behaviours can be detected. For instance, if a person falls down as a result of a fight, this can be detected. If a person begins to run suddenly – this too can be detected. If crowds gather suddenly this could be a symptom of some unusual activity as well. Having a system that can monitor various types of human behavior can help the city's officials ensure the safety of their citizens.

The use of facial recognition when such events occur is very useful to determine if there are known offenders involved and increases the rate of apprehension.

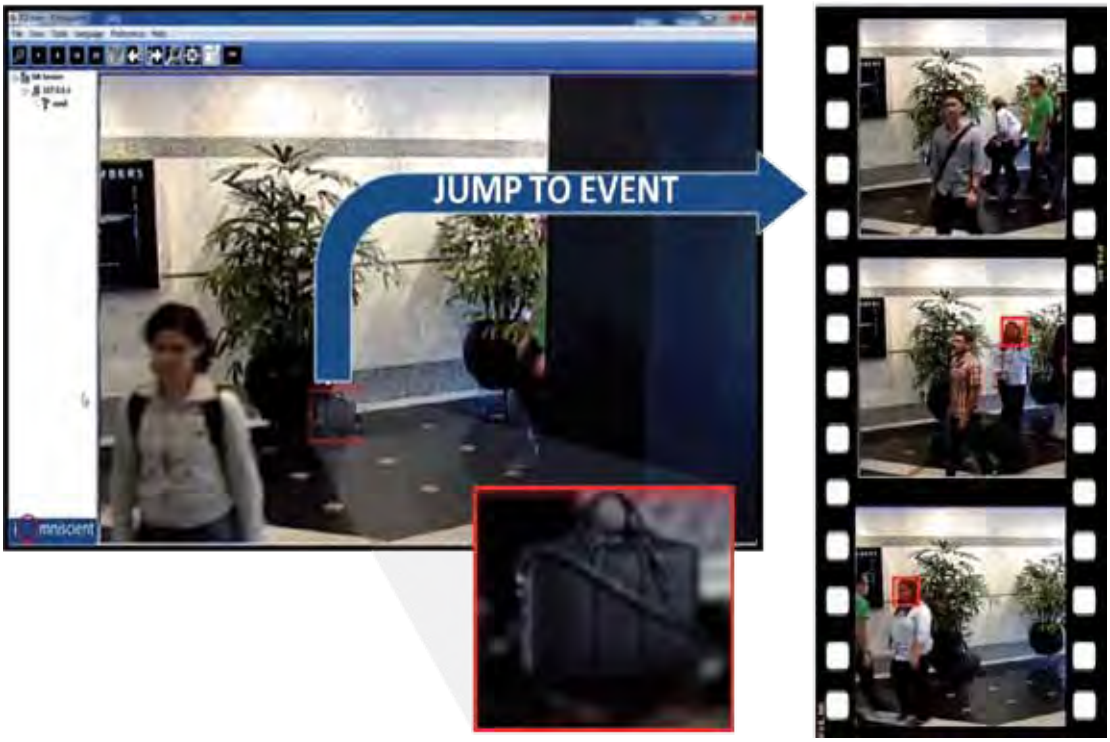
Example 2: Passengers without Documentation



Often in airports, a person may arrive at the immigration counter of an airport and declare that he is a refugee. He will usually have no papers or admit to where he has come from.

Cameras at the air-bridges (where people get off the plane) can capture the images of all people coming off every plane. When a person behaves as described above, their image can be supplied to the system to determine the passenger's incoming flight number and original port of embarkation.

Example 3: Detecting Culprits after an Event



If an event occurs (e.g. someone abandoning a bag or someone falling down for whatever reason), iOmniscient technology (using iQ-Hawk) can automatically zoom in and recognize the person involved if the image is already in the database.

The system will also record the faces of all those who are present at the event even if they are not involved so that the police have a record of all witnesses.

Example 4: Detecting Shoplifters

Shoplifters and thieves, especially repeat offenders in a particular public space such as a shopping mall can be detected and apprehended using iOmniscient's "many-to many" facial recognition system.

For example, a shopping mall was plagued by shoplifters. It therefore produced a blacklist of known shoplifters (people who had been caught before in that mall) and gave that to their security guards. There were over 200 people on this black list. The



security guard would keep a little booklet with pictures of these 200 known shop-lifters in his pocket and his job was to identify these individuals if they entered the mall and to keep a close eye on them. Such a system was totally ineffective. It is very difficult for an individual to memorize a face from a picture unless if there were some very glaring features (like a scar across the face). The accuracy level of this manual system was close to zero.

The mall then implemented iOmniscient's "many-to-many" facial recognition system. The system watched people as they entered the mall and captured their image if they matched a person on the database. Matched images were sent to the officer in the field on his wireless device. The officer could review the images and make a judgment on whether the match was a good one and whether the person in view was actually one of the persons in the database. Humans are good at making such judgments against a small number of images even though they are not good at comparing thousands of people against hundreds of images in a database in real time.

If this hybrid system is only 80% accurate, it is still much more accurate than the

alternative, which was a human performing the task without help. If 20% of the potential culprits are missed, it still means that 80% are apprehended. The technology certainly makes the human operator much more effective by complementing the capabilities of the system.

Example 5: Customer Service for VIP guests

A system similar to the one used in the shopping mall has been used by hotels to greet VIP guests.

When a VIP customer enters the hotel, the system identifies the person and brings this to the attention of the receptionist. The system might provide a short list of matches for the receptionist to make the final selection. The customer's profile can then be shown automatically on the screen. When the guest arrives at the desk, the receptionist is able to greet the person by name pleasantly surprising the customer who may not have visited the hotel for a year or more.

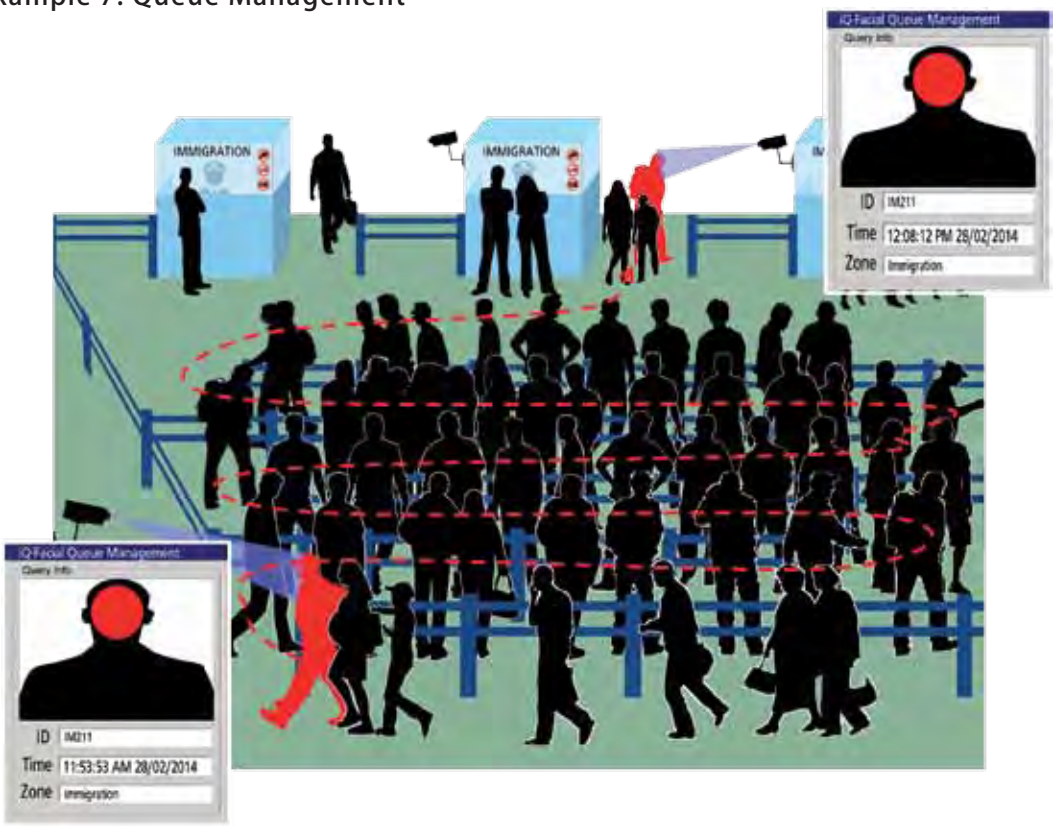
The system does not have to be 100% correct. The receptionist can make the final judgement as to whether the person is indeed the customer shown in the database. If the hotel can provide this service to 80% of their important customers, it is far better than not being able to do it at all.

Example 6: Tracking in a Crowd



Tracking an individual in a crowd can be very difficult. iOmniscient's Facial Recognition in a Crowd technology provides information on exactly when and where an individual has travelled (e.g. seen on Camera 1 at 10.00 and then on Camera 16 at 10.10 and on Camera 22 at 10.15). Even if the person of interest is missed on some cameras, other cameras will pick up the information.

Example 7: Queue Management



Facial Detection can be used to manage queues. For example, as people get off a plane, their image may be captured, they are allocated a number and their image placed in a dynamic database.

As people reach the immigration counter their images are captured again. The system still can recognize a person and check his image is in the dynamic database. The system can now calculate how long it took the person to travel from the air-bridge to the immigration counter. The person can be tracked right through the airport.

For privacy reasons the images can be encrypted and not visible to the operator. Once the person has left the airport his record in the dynamic database is deleted.

Such systems can help authorities manage queues, understand waiting times and develop strategies for improving the level of service.

Example 8: Dwell Time Management

The Facial Detection technology can be used to detect "dwell times". Dwell Time Management is a variation of Queue Management. The system can determine how long a person spends in different parts of a mall or different parts of a particular store.



The system does not need to recognize individuals. It simply enrolls them in a database when they are first seen and calculates how long they spend between two points.

Example 9: Big Data Analysis



Big Data analysis can be used to analyze data from various sources. For example, a bomb detonates in several metro stations in a major city. From the video footage police investigators are able to recognize the perpetrators. But were they acting alone? Using Big Data Analysis the secret service branch can access the social media used by the perpetrators. From their telecommunications record they can determine the numbers they have called most often in recent days. A Facial Recognition system can pull out all records of pictures or videos of the perpetrators and their contacts from social media or conventional media sources.

Example 10: Social Media Analysis

Jill rings Jack to tell him that she saw his picture in a newspaper. This makes Jack curious. Who else has exhibited pictures of him?

He can use a Facial Recognition system to do a quick search on conventional and social media to discover all the occurrences of his image uploaded by his friends and others. This is just a simple example of the use of Facial Recognition in social media.

Example 11: Access Control

Access to important public buildings (and car parks) may need to be controlled. Standard access control systems can be easily defeated. Access cards can be used by unauthorized personnel. Or an unauthorized person may tail-gate behind an authorized person.

In such situations, iOmniscient's Facial Recognition in a Crowd system can be used to authenticate that the holder of the access card is indeed its owner. Further tail-

gating can be detected and the faces of the culprits recorded using iOmniscient's Facial Detection system.

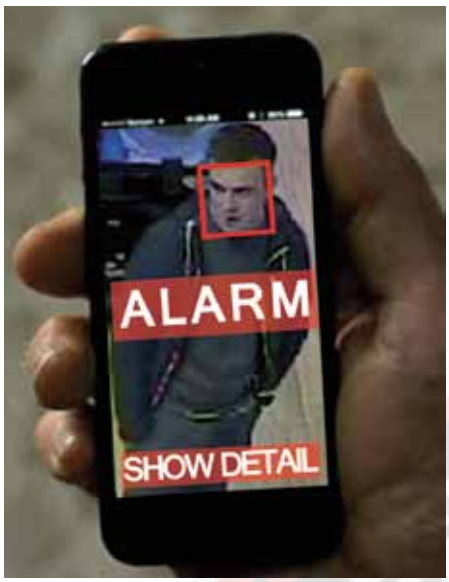
While Facial Recognition in a crowd is not useful as the primary method of identification for access control, it can be very useful as a secondary identifier.

A major manufacturer employed large numbers of contract laborers. Over ten thousand of them would arrive on buses each morning. Each had an access card that they swiped to gain entry to the site. However, it was discovered that cards could be swapped. It was possible for a person to give his card to a friend who might come in and work on his behalf. Also he might swipe a number of cards (and hence collect multiple wages).

A Facial Recognition system was installed to provide secondary verification. When a card was swiped, the person's face was matched against the image associated with the card and access was permitted or denied based on the result. As a secondary verification, the system was sufficiently accurate and fast to ensure that the traffic through the gate was not slowed.

Example 12: Private Security

A certain person travels extensively for her work. She wants to ensure that her family is safe at home. She puts in a security system at her house with a camera at her front door. When a person visits, the system checks the person's identity with those of known friends and service people. The system itself runs somewhere in "the cloud". If an unknown person turns up at her door she receives an immediate alarm on her smart phone with the face of the visitor. If she is concerned about the visitor she can then take immediate action if necessary.



Example 13: Policing Applications

A policeman sees a person behaving suspiciously. He can take a short video of the person from his body camera and send it back to a central database for matching. Within seconds he can have information about the matches found in the system and know details about the person's identity and background.

Such systems make policing easier and make cities safer for their citizens.



Example 14: Driver Match

In public car parks vehicles may sometimes be driven out by someone other than the driver who brought the vehicle in. Very often this is a symptom of a vehicle theft. Experience shows that it is often insufficient to use humans at checkpoints to detect such theft of vehicles. The Driver Match system recognizes the number plate of the vehicle being driven as well as the face of the driver and checks that the driver was the one who previously drove that vehicle.



10 PRIVACY ISSUES

The increasing accuracy of facial recognition systems and the number of applications for this technology have raised the issue of privacy of people whose face images are being captured and analyzed.

Facial images are just one of various categories of data that is captured and stored by governments, commercial enterprises and social media. Some of this is provided voluntarily via social media, others are captured in the process of accessing and using commercial information such as telephone and internet records, travel information and medical data.

Social norms on privacy differ from country to country. What may be considered normal in one country may be considered intrusive in another. They may also change over time. When there is a high concern about security and safety, privacy issues may receive less attention than otherwise. Further while some organizations such as passport, driving license issuing bodies and the police may be authorized to capture and use images in certain circumstances, others may not be permitted to do so.

Each society makes its own tradeoffs between rights of citizens to privacy and society's right to information which can enhance its safety and security. Some societies may even legislate on how long certain types of images may be kept and for what purpose.

It is always the responsibility of the implementer of such technologies to ensure that he complies with the rules and expectations of the society within which he operates.

Of course, there are techniques available to protect the privacy of individuals even while the system is being used for its intended purpose.

The most common method is to restrict access to information from the system only to those who need to know what is in the system. If it is a system owned by the police, they can restrict its use to authorized trusted officers.



The second method is to use the information but to ensure it is not available for view. Hence the faces of all individuals could be automatically pixilated. The system could view the information, but an operator viewing the video would not see individual faces unless if they were authorized to do so.

Video is often stored for long periods of time. There is often no requirement for this. There would be fewer concerns about privacy if the information was deleted as soon as it was no longer needed.

For instance consider a shopping mall that is interested in apprehending known shoplifters. Their system can watch all those who come into the store. It can match individuals against their list of people of interest. For those not on the list, the video containing their images can be deleted as soon as the system has finished the matching. For those who are matched against images within the database the mall may wish to provide the information to their security officers in the mall so that they can observe any unusual behavior.

All information has the potential to be abused by those who have it. Each society has to make its own judgment on how it ensures the right balance between its need

for protecting its citizens and managing their environment and their concerns about privacy.

As is usually the case, technology and innovation are ahead of the law. There is clearly a benefit to the community for law enforcement agencies to improve surveillance and thereby provide improved security and safety for its citizens. With increasing security concerns and global terrorism, the community has accepted that there is a trade-off between increased security and some loss of privacy. Commercial enterprises are also interested in utilizing CCTV technology to increase the level of service they provide to their customers.

While this technology continues to evolve, self-regulation by developers and commercial enterprises and legislative controls by national government will be needed to ensure that the information is collected and used for legitimate purposes and properly disposed of when not required. The communities of developers and users have a joint responsibility to ensure that appropriate safeguards are put into place.

Multi-Sensor
Multi-Analytics
Multi-Recognition



11

IOMNISCIENT'S FACIAL RECOGNITION IN A CROWD TECHNOLOGY

When the iOmniscient Facial Recognition in a Crowd system was first launched, it won international awards around the world.

The company won the "Global Security Challenge for Crowded Scenes in 2010" and in 2011 it won the coveted "Best CCTV Technology of the Year" at the IFSEC Exhibition (this was the second time that the company had won this award). The awards reflected the fact that Facial Recognition had never been able to achieve the type of results offered by this new system. The technology continues to win awards and in 2014 it was declared to be the best technology at the Australian Security Exhibition.



11.1 Development Goals

In developing the Facial Recognition in a Crowd system the company had several design goals which it achieved.

1. iOmniscient is renowned for the ability to analyze video in crowds. Its existing Video Analytics systems can detect abandoned objects in crowds,

theft in a crowd and vandalism and graffiti in a crowd. The system can count in a crowd and understand crowd behavior. The Facial Recognition system also works in a crowd.

2. The system is primarily designed to work with readily available relatively low resolution surveillance cameras. It does not require any specialized equipment.

3. It works in an uncontrolled environment. The target of the surveillance may not be aware of being watched and hence the system recognizes the person irrespective of behavior or expressions.

4. Information on the target can come from many public sources, so it must be possible to collect the images in the gallery or database for matching from uncontrolled environments. Images can be collected from family albums, holiday videos and even newspaper clippings.

11.2 Patent Protection

The system is protected by a number of international patents on the core capability which ensures that competitors have great difficulty in achieving similar results. To date no other company has achieved results that are even close in an uncontrolled surveillance environment.

11.3 Multifaceted Facial Recognition

The three main differentiators of iOmniscient's Facial Recognition system are:

- The system can operate optimally with a resolution that delivers only 22 pixels between the eyes. In fact the system can operate at even lower resolutions (down to 12 pixels) but accuracy is increasingly sacrificed as the resolution drops below 22 pixels. No other company has provided similar results even at twice this resolution. Both the target and the matching gallery or database can include low resolution images.

- The system can cope with uncontrolled environments. Both the target and the matching gallery can be uncontrolled.

- The system is tightly integrated into the very advanced iOmniscient Video Analytics system. This provides the user with the ability to use it in a preventive more to analyze behavior and to recognize faces in response to certain behaviors even before the person commits a crime. This type of Automated Surveillance is also a patented capability.

The database may exist already.

Alternatively the database may be generated dynamically. For instance, a person may be detected abandoning a bag. His face is automatically enrolled in the database. He can then be tracked as he passes other cameras.

Or there may be no database at all such as where the system is used to determine how long a person dwells in a particular area. The person may be automatically enrolled in a database for a few moments. His information is deleted when it is no longer required for the calculation.

In addition, the system has all the attributes ascribed to surveillance oriented Facial Recognition systems in this booklet.

The facial recognition system is open, scalable and flexible. It is unaffected by pose, expression, spectacles and facial hair. It provides a fast real time response even with thousands of people in the matching database and a large number of targets appearing in the scene simultaneously.

It can operate in a Crowd.



12

MAXIMIZING THE RETURN ON INVESTMENT FROM YOUR FACIAL RECOGNITION SYSTEM

Any intelligent system is a major investment for the stakeholders. There are several factors that affect how the return can be maximized.

The normal assumption is that minimizing the cost of the system is the best way of maximizing the return. Unfortunately getting the cheapest system usually results in the implementation of a system that does not meet the objectives of the stakeholders. It is better to have no system than to have spent a significant amount of money on a system that is ineffective.

12.1 Minimizing Cost by Design

And an effective system does not have to be more expensive. It is only more expensive if it is not well designed.

For instance, as explained already, the use of iQ-Hawk can reduce the cost of storage and network bandwidth by such a significant amount that even after adding the cost of the software, the overall cost of the solution would be much lower. Of course to achieve this benefit the whole system has to be designed by taking into account the advanced capabilities of iQ-Hawk. If the system is designed using conventional assumptions, then these cost reduction benefits would not be realized.

Therefore the lowest cost software may not result in the lowest cost system. The overall cost of the system is what is important rather than the cost of any one component.

Other architectural assumptions can also affect the cost of the system. Systems are often designed as Centralized or Decentralized and this is usually based on the technology provided by the vendor. If the vendor provides an Edge based solution there will be a propensity to design a decentralized solution and vice versa. Since the iOmniscient solution is available both in a server based centralized architecture and in a Super Edge based decentralized one, the actual design can be one that is most appropriate for meeting the objectives of the stakeholders. The most cost effective systems are hybrid systems such as those available from iOmniscient as they provide the flexibility for different user groups to use different architectures to

meet their objectives in the most cost effective manner without compromising the overall design of the system.

As we have seen, some tasks are better performed centrally and others work better when distributed. If the architecture is not the constraining factor then the overall cost can also be minimized.

Irrespective of the architecture, the level of intelligence available on the system will have a drastic impact on the effectiveness of the system. A system that is incapable of working in a crowded system will merely give an unacceptably high number of false alarms in that type of situation. Selecting software with the right level of intelligence for the task to be performed is the most critical factor in determining the value of the system.

12.2 Having the Right Cameras and Other Infrastructure

Users will often spend significant amounts on cameras which may not be appropriate for the job at hand and these may be placed in positions that are not appropriate either. There are numerous examples of cameras being installed with inappropriate views. And there are many examples of systems failing because cameras were not appropriate. This does not mean that the cameras were of an inferior quality. Rather they did not have the characteristics required to generate the right image for the application. Networks have also been put in place with insufficient bandwidth to transfer the video generated by the installed cameras. This results in significant wasted investment. In the next chapter we will discuss the right process to ensure that such waste is minimized.

12.3 Maximizing Return on Investment

The difference in price between the most effective product and one that barely works may not be significant. But the difference in utility is infinite. If one feels that one cannot afford to purchase a product that can do the job effectively it may make more sense not to purchase a product at all, as ineffective products just provide an illusion of security and safety.

A well designed system can achieve multiple objectives using different types of sensors and indeed can achieve multiple objectives on each camera.

The effectiveness of a product cannot always be measured in absolute terms. If due to poor camera placement or a very difficult environment a product can achieve 70% accuracy in a particular situation, it may still be far more accurate than the alternative which may be a human attempting to achieve the same objectives.

There are many components in a Facial Recognition system and the intelligence software constitutes a very small proportion of the total cost. It is however the core of the system. It is critical to select the best available core as the return on the entire system can be compromised if the core does not work effectively.

12.4 Operational Efficiency

The capital costs of the system are not the only costs that the users will incur. There are many operational costs in using a sophisticated system.

If it takes a long time for the system to provide answers required by the user it reduces his efficiency. If the system cannot function effectively when there is no one around in the control room that can affect efficiency. If a large part of the system is not operational for any reason, that can affect efficiency. Reduced operational efficiency reduces the value of the system to the user.

iOmniscient has designed several features just to improve operational efficiency. The Jump to Event function is designed to answer the question about when an event first started and who was involved.

The Automated Surveillance Capability of iQ-Hawk provides immediate information about people and vehicles involved in an incident. Without such a system, operators may spend hours (or even days) attempting to extract the data from unrelated systems.

The Auto-archive function and the availability of information on mobile devices

ensure that the system can continue to operate even when the control room is not manned.

The Automated Response Capability can enable the police and other emergency teams to arrive at the scene of an accident or crime much faster than would be otherwise possible improving efficiency and possibly having an impact in a situation where a few minutes can be the difference between life and death for the people involved.

The iQ-Health Check system can advise the user if part of the system stops functioning.

12.5 Maximizing Uptime

Every time the system is down and not operational, the user is not receiving the value that he should be expecting from the system. Therefore maximizing uptime is critical. To achieve this, the software and hardware has to be reliable and it is important to know when it is not working so that it can be immediately fixed.

Reliability comes from using good quality hardware and software. It can be enhanced through the use of redundant components.

Knowing that parts of the system are not functioning because they have broken down can be established fairly easily. The iQ-Health Check system can make the user aware that the system is not functioning correctly, even if this is due to external factors such as too much rain, a spider's web blinding the camera or movement of the camera due to vibrations.

The user should reasonably expect both the Systems Integrator and the suppliers to offer a level of service and ongoing support for fixing problems that is commensurate with the requirements.

12.6 Overcoming Obsolescence

The technology is advancing rapidly. Computing hardware and cameras can become obsolete and unusable in just a few years. Software improves even more rapidly and within a year a system can be overtaken by new advances.

For this reason iOmniscient offers an ongoing product update service to ensure that the system is continuously upgraded. In ten years' time the hardware may have collapsed and been replaced several times but the software should be as new as if it was bought that very day.

The key to achieving this goal is that software must always be both forward and backward compatible with itself. It must be able to interface with products with which it could interface before while developing new interfaces for new devices and new products. This can only be achieved through strategic design. Very few products can achieve this goal. iOmniscient products are designed to be forward and backward compatible across versions because the objective is to eliminate obsolescence.

Throughout the book we have emphasized the importance of openness and pointed out the negatives associated with products that attempt to lock users in with proprietary hardware. Nowhere is this more important than when one is attempting to reduce the risk of obsolescence. If a user is locked in to one supplier it is impossible to take advantage of the latest improvements in technology. These are not the monopoly of any one supplier. A commitment to open architectures and interfaces is the foundation to ensuring a system can take advantage of various technologies as they improve.



13

IMPLEMENTING A FACIAL RECOGNITION SYSTEM SUCCESSFULLY

For a Facial Recognition system to be successful, the objectives of the user must be clearly defined at the start of the project.

A successful system will provide the right information in a timely manner, automating all standard events and allowing sufficient time for humans to make a judgment on the more unique, special events.

13.1 Defining System Objectives

The best method for defining the system objectives is to perform a systematic risk assessment to ensure that the right problems are being solved. There are numerous examples of cameras that have been installed pointing in the wrong direction or indeed where they are in the wrong place. The priority risks lay unattended while systems are installed to solve the wrong problem.

There is now an international standard for Risk Analysis (ISO 31000:2009). iOmniscient can provide a service for a comprehensive Risk Analysis for any environment.

Once the key risks have been identified and their consequences understood, it is important to prioritize them.

13.2 Prioritize the Objectives

Prioritization of risks and therefore the objectives of the system ensure that the investment is made to address the most important objectives first.

Once the user has an understanding of the risks to be addressed, it is important to select the most appropriate risk mitigation strategies. These may not involve the use of cameras. There may be risks that can only be addressed by other methods such as the presence of an armed guard or a lock on a door.

The systematic risk assessment will assist the user in deciding where surveillance and automation technology is required and the reasons for this decision. This will ensure that the most appropriate technology is used to achieve specific purposes.

In security the onion principle is important. According to this principle, it is best to have layers of protection just like the layers within an onion. If one layer of protection is defeated, there are still several levels of defense that have to be overcome. No single technology will solve all problems. However many technologies can work together to provide a secure and safe environment.

Having prioritized the risks, it is important to determine the types of event where an automated response is appropriate and where human intervention may be necessary. Many situations which were previously responded to by humans are simple to automate leaving the human to address more complex tasks.

Once the objectives have been determined, there are two important elements to implementing a successful system.

1. Following a disciplined implementation process.
2. Selecting the right suppliers for software systems, hardware and implementation services (the Systems Integrator).

13.3 Disciplined Implementation Process

To implement a successful system one has to follow a four step process as follows:

1. Select the appropriate software to meet the objectives.
2. Select the cameras and other hardware infrastructure.
3. Select the Systems Integrator.
4. Implement the system with appropriate camera placement and test the outcome to ensure the objectives are met.

Step 1: Select the appropriate software to meet the objectives of the system

It is important to identify and select the software that would be most suited to addressing the objective of the system.

Step 2: Select the cameras and other system hardware infrastructure

As a next step, the user should identify the sensors that should be used. These include the cameras. Without going through this process the organization may select inappropriate cameras for the required task and place them incorrectly for the applications. Of course the software would then not work optimally.

Once the cameras have been selected the next step is to select the computers, storage and networking infrastructure and ensure it is appropriate to the software to be used. Selecting the right software in step 1 can greatly reduce the cost of the other hardware elements.

Step 3: Select the Systems Integrator

Next, it is important to select the Systems Integrator who has the capability of bringing all the components of the solution together.

Step 4: Implement the System

The final step is to proceed to implement the system. Cameras and sensors need to be placed as per the directions of the supplier of the Video Analytics (not the camera vendor).

It is also important for the user to establish a test plan to ensure that all the objectives of the system are met once the system has been implemented.

13.4 Frequent Mistakes in Implementation

Unfortunately many users of such systems have started with Step 2. First they have selected the cameras and even installed them without determining if they are the appropriate cameras and whether they are placed correctly for the application.
Some users have selected the Systems Integrator at Step 3. They have hoped that this organization will implement a working system. However without going through the previous steps, the Integrator has little chance of implementing a

system that actually meets the needs of the user.

When a system fails to deliver the results the fingers are pointed in all directions – except at the core problem which is that the user has not followed the disciplined approach that has been prescribed.

13.5 Questions You Should Ask Your Vendors

Intelligent systems are large and complex with many suppliers involved. The key suppliers are:

1. The software suppliers
2. The hardware suppliers
3. The System Integrator

Most Facial Recognition systems are implemented by a Systems Integrator. These companies acquire the software and hardware for the system and manage the integration of these to implement the system. It is important that the purchaser understands and specifies the requirements of their system.

The user needs to ensure:

1. The system requirements are clearly specified. iOmniscient can provide sample specification documents to help the user with this exercise.
2. The selected software and hardware solution will meet these requirements.
3. The selected software and hardware is delivered and implemented (not a cheaper less effective alternative).

The user must have a test plan to ensure that the system delivers the required capabilities in a realistic complex environment and the objectives of the system are met.

Sample test plans are available from iOmniscient.

System Objectives

1. Does the system achieve the objectives of the key stakeholders of the system?

Intelligence Levels

1. Is the software so smart at optimizing resources that a system with the software can be less expensive than a system without the software?
2. How robust is the system? This can be determined by establishing if similar systems have been operational for long periods of time.
3. How committed is the supplier to enhancing capabilities in future to ensure the system does not become obsolete?
4. Is it necessary to implement record all video all the time? Most advanced systems may record video only for events and for their verification. In very few situations does all video have to be recorded and archived for long periods of time.
5. Can a single system provide all the capabilities required (including Facial Recognition in a crowd, License Plate Recognition, Automated Response and Audio and Smell Analytics? Remember that interfaces between different suppliers are often the weakest link in the system.

Facial Recognition

1. Can the system perform Facial Detection and Recognition in crowded and complex environments?
2. Can the system recognize people based on the events that it has detected?
3. Can the system track people from camera to camera even when their views do not overlap?

Jump to Event and Determining the Identity of the Person Involved in an Incident

1. For complex events that might take some considerable time, can the system go back to before that event commenced at the press of a single button and recognize the person who was involved? (Consider an event such as a person abandoning a bag for ten minutes. Can the system recall the video of the person bringing the bag into the scene)?
2. Once the start of an event has been determined can the system automatically identify the person or vehicle (or other object) involved in that event?

Automated and Mobile Response for Emergency Management

1. Can the system automatically show the human operator where events are happening on a map of the environment?
2. Can an operator view all events on a Smart Phone and can these be archived and managed from this device?
3. Can the system locate the nearest responder (e.g. police car, ambulance, fire brigade) and automatically advise them of an incident?

Big Data Capability

1. Does the software generate Meaningful Meta Data?
2. Can this data be used to provide reporting, analysis and forecasting capabilities?
3. Can the system pull together unstructured information from a variety of different sources and use it to understand the environment and predict future events and automatically advise them of an incident?

System Cost Effectiveness, Reliability and Efficiency

Some questions to ensure that the system operates cost effectively and with a minimum of human intervention are:

1. Can the system intelligently store information of interest or does it simply record all video footage?
2. Can important information (such as faces) be stored automatically in higher resolution than other irrelevant details?
3. Does the system know if it is not working or if the cameras cannot see properly?
4. Can the system be scheduled to perform different functions on a camera at different times?
5. Does the system have a Nuisance Alarm Minimization System (NAMS) capability?
6. Does the system have a Universal Connectivity Module?
7. Can the system be configured, implemented and maintained remotely?
8. Can the system operate as an Intelligent Network rather than just as a smart camera?

Software Selection

Some questions to ask when selecting the software vendor are:

1. Can the technology actually meet the stated objectives?
2. Is the proposed solution comprehensive enough so that the number of interfaces to other vendors is minimized?
3. Is Video Analytics their core business? This is a very specialized field and many organizations and many software suppliers are relatively inexperienced. Ensure you are working with an expert in the field who has been around for a long time.
4. Is the software supplier knowledgeable? Make sure your supplier is recognized as knowledgeable in the field. Have they published significant books, guides or other information that demonstrates their knowledge?
5. Is the software supplier committed to openness? Engaging a vendor with proprietary interfaces which do not work openly with others will greatly restrict the user in the long term.
6. Is the software robust? Has it been implemented in a large number of different circumstances? There are many new players who are essentially attempting to commercialize university projects. The resultant products tend not to be sufficiently robust in meeting the real user requirements.

Hardware Selection

Some questions to ask when selecting the hardware vendor are:

1. Can the camera vendor commit that the selected cameras are suited for the application?
2. Is the vendor able to provide the other sensors needed to enhance the capability of the overall system?
3. Does the computer hardware meet the specification required for the Video Analytics and is there sufficient network bandwidth for the system to perform adequately?
4. Is the vendor able to supply hardware required for the mobility and Automated Response applications?
5. Has all the hardware been certified by the supplier of the analytics as being fit for use

Systems Integrator

Some questions to ask when selecting the Systems Integrator are:

1. Is the Systems Integrator experienced in the implementation of facial recognition systems?
2. Make sure their experience is not limited to implementing simple VMS/DVR systems.
3. Has the Integrator developed a test plan that has been approved by all the stakeholders and which meets all the objectives for the system?
4. Has the Integrator been certified as trained by the supplier of the Video Analytics?
5. Are training and documentation provided?
6. Does the user have a commissioning plan that covers everything that has to be implemented including the advanced video analytics, Automated Surveillance and Automated Response?

Appendix :

Results of Tests on Accuracy of Facial Recognition Systems

Facial recognition systems are new technology and international standards are still evolving.

The US National Institute for Standards and Technology (NIST) has tested different Facial Recognition Systems in controlled environments and published the results over the years. Over the past decade, there has been a great improvement in the accuracy of facial recognition systems.

The latest controlled tests by NIST, were conducted in 2013 and compared the accuracy of facial recognition software from a number of commercial providers. NIST tested the accuracy of these systems for high quality mug shot images, such as those used for driving licenses, as well as for poor quality webcam images. The systems were tested for the accuracy of one to one verification, such as for e-passport holders and for the issue of driving licenses and one to many searches, such as in a criminal database. The tests did not cover many to many recognition, nor did the database used in the test include facial images in video sequences or with varied pose and angle. In the tests both the images of faces tested and the images in the database were full frontal images.

The NIST tests compared mugshot facial images to a database of 1.6 million images, similar in quality to those used by law enforcement agencies and which complied with the ANSI/NIST ITL 1-2011 Type 10 standard. In addition, higher quality images that are used for visa applications and that meet the ISO/IEC (International Organization for Standardization/ International Electrotechnical Commission) standard were included in the database. The study also included 140,000 webcam poorer quality images that do not comply with any standard.

The NIST tests showed that, using a database of mugshot images from 1.6 million individuals, the most accurate algorithm, failed to yield the correct match (of an image with one in the database) in rank one position, in 4.1% of cases. For webcam images, which are generally of poorer quality, but are full frontal images nevertheless, the algorithms failed to yield the correct match in rank one position in 11.3 % of cases. The level of accuracy decreases as more

identities are enrolled into a biometric system, as the possibility of a false positive increases due to lookalike faces For example, the rate of false positives for full frontal high quality images reduces to around 3% if the size of the database is reduced to 160,000.

iOmniscient did not participate in the tests as they were focussed on controlled environments and not on Face Recognition in a Crowd which is its specialization.

The NIST tests showed that the greatest accuracy was obtained with relatively high-quality images that comply with the ISO/IEC standard and which are usually collected for passport, visa and driving license applications. The NIST study did not find any supplier that submitted its software for testing, performed well with poorer quality webcam images. Search failure rates for those images were around three times greater than for the higher quality images.

The tests were performed with very high resolution images in controlled environments.

Even for access control systems such conditions are rarely available. In surveillance applications where full frontal high-resolution images are rarely available, much lower accuracy can be expected from the systems that were tested.

Since the NIST tests are targeted at the type of environment that might exist in an Access Control situation, what level of accuracy can be expected using facial recognition systems?

Ideally, the false acceptance and rejection rates should be zero.

The lower the rate of false acceptances, the more secure the application. The lower the rate of false rejections, the greater the comfort for users and the lower the workload for human operators.

One way of increasing accuracy is to use a human to adjudicate from a list of likely facial matches. The tests performed by NIST demonstrated that human adjudication reduced false positive matches significantly. For example, if a list of up to 20 images

is provided for human adjudication in order of likelihood, there was up to 50 percent improvement in accuracy.

The NIST tests also indicated that accuracy of recognition increases with an individual's age. Babies and young children are the most difficult to recognise while older individuals can be recognised with greater accuracy.

The report suggests that the accuracy of facial recognition systems can be improved by improving the quality of the facial images and in particular conforming to the ISO/IEC Standard 19794-5 for images. Of course this is not feasible for surveillance type applications.

The NIST study recommends that human adjudication be used to improve the accuracy of facial recognition systems. However to ensure accuracy, NIST recommends that the image should be at least 800 pixels. Such high resolution full frontal images are not available in surveillance type applications. This is why iOmniscient's systems are designed to operate with around a third of the resolution recommended by NIST and with varying pose and expression.

The NIST study recognised that the quality of the algorithms used in facial recognition is crucial to the accuracy of the system. The best image capture systems will have poor levels of accuracy if they are not supported with the effective and intelligent software for facial recognition.

Most facial recognition systems can only tolerate limited variations in the portrait. For example, if a person allows his beard to grow, then the system will recognize him with almost the same reliability, as if his beard had not changed. However they tend to be less tolerant to changes in pose, weight and other factors.

For example systems dependent on high resolution images suffer from the "aging" of the picture. As people grow older they develop new lines on their face and they might gain weight. The high level of detail used for Facial Recognition systems that are used for access control begins to work against these systems when such changes occur and accuracies can fall significantly.

While the NIST tests were conducted in an operational environment relevant to access control systems, there are many factors which may affect a system's accuracy in practice. In addition to orientation, resolution and lighting, the accuracy of face recognition systems is affected by problems due to physical changes and changes in appearance: expression, changes of hair, changes in weight, spectacles, hats, ageing, injuries, illness, etc. Hence the actual accuracy achieved are usually lower than in the NIST test.

For further information see:

National Institute of Standards and Technology (NIST), Performance of Face Identification Algorithms, NIST Interagency Report 8009, May 2014, available from
http://www.nist.gov/customcf/get_pdf.cfm?pub_id=915761

National Institute of Standards and Technology (NIST), Performance of Face Recognition Algorithms on Compressed Images NIST Interagency Report 7830, December 2011, available from
http://www.nist.gov/customcf/get_pdf.cfm?pub_id=908515

National Institute of Standards and Technology (NIST), Report on the Evaluation of 2D Still Image Face Recognition Algorithms, NIST Interagency Report 7709, August 2011, available from
http://www.nist.gov/customcf/get_pdf.cfm?pub_id=905968

About iOmniscient

iOmniscient has been the technology leader in providing Video Analytics and Surveillance Solutions since its inception in 2001. It provides the most comprehensive suite of Detection, Recognition and Automated Surveillance Systems. To enhance its capabilities, iOmniscient has also added Sound and Smell Analytics to its comprehensive portfolio. It provides multi-media and multi-sensor analytics capabilities.

iOmniscient's advanced systems have an unmatched capability in analyzing realistic, complex and extremely crowded scenes using its internationally patented technologies such as Non-motion Detection that are not available from anyone else. The systems provide numerous applications from simple intrusion and counting systems to unique object detection in a crowd, multi-lingual License Plate Recognition, the world's first non-co-operative Facial Recognition for crowded scenes, the detection of unusual behaviors and much more.

The analytics capabilities are integrated to provide a Smart City Solution for enhancing safety and security of citizens and improve the efficiency of services. iOmniscient is at the forefront of implementing smart city systems around the world.

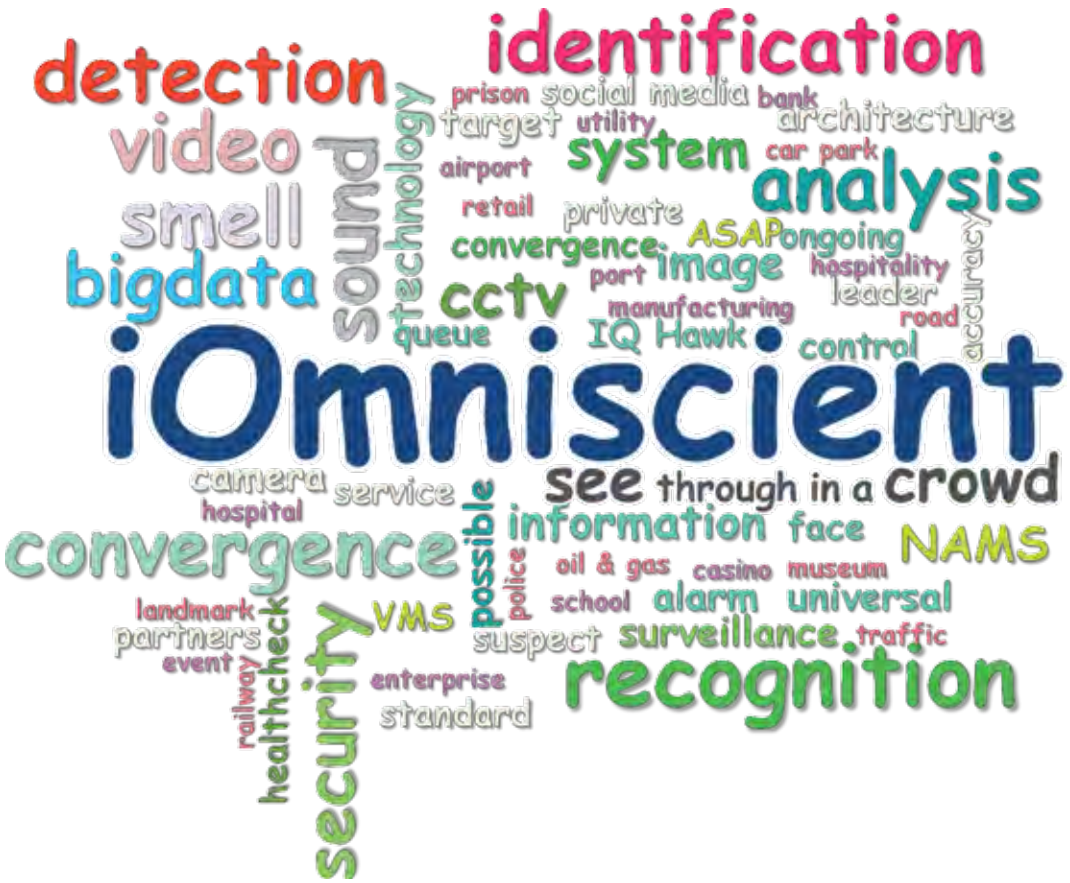
All iOmniscient systems are armed with an Artificial Intelligence based Nuisance Alarm Minimization System (NAMS) to help eliminate false alarms while still maintaining detection accuracy.

With its extensive industry experience, the company has developed a portfolio of unique solutions for more than 30 industries ranging from Airports and Railways to Oil Refineries and Road Traffic Management.

iOmniscient has won many international awards for its intelligent and unique products, including the IFSEC Award for the "Best CCTV System of the Year" for its Facial Recognition in a Crowd and the Global Security Challenge for Crowded Scenes.

As a pioneer in Video Analytics with over 14 years of commercial experience, iOmniscient is dedicated to continue to develop revolutionary intelligent surveillance technology and to maintain its global lead in this field.

For more information, please visit
www.iomniscient.com
www.iqsmartcity.com


identification
detection
prison
social
media
bank
target
utility
architecture
video
sound
technology
system
car park
airport
analysis
smell
retail
private
ASAP
ongoing
convergence
image
hospitality
bigdata
port
leader
cctv
manufacturing
road
accuracy
queue
IQ Hawk
control
iOmniscient
camera
service
see
through in a
crowd
hospital
possible
information
face
NAMS
convergence
police
oil & gas
casino
museum
landmark
school
alarm
universal
VMS
partners
suspect
surveillance
traffic
event
enterprise
recognition
railway
healthcheck
security
standard



info@iomniscient.com | www.iomniscient.com

iOmniscient

High iQ Recognition