SYBERNAUTIX®

Case Study

CLIENT: MHA, Singapore.

CUSTOMER PROFILE

Ministry of Home affairs, Singapore looks after entire internal security and issue necessary directive and guidelines to the agencies working under its control. Singapore police force is one of such agencies working under its control and looks after law and order and crime control in Singapore. Both of these parties working together planned to implement the criminals moving on the road using the facial recognition technology. They have planned to utilize the existing CCTV surveillance system to be used for the face recognition.

THE BUSINESS ISSUE

Singapore Police Force (SPF) and Ministry of Home Affairs (MHA) with their established security system with CCTV surveillance are unable to identify the suspects.

1. With the recent Terror attacks, they are unable to trace the culprits and stop the attacks that are mainly targeted in the metro railway stations and trains.
2. They are finding difficulty in avoiding the attacks of shop lifters in big shopping malls.

THE SOLUTION

INTELLIGENT CCTV: FACIAL RECOGNITION TECHNOLOGY

A video surveillance system usually consists of number of cameras and some surveillance monitors, whereas security guards are responsible to monitor all those cameras simultaneously. The real time video is channelled and reduced very much such that the likelihood of recognising all real world dangerous situations is low. Video content analysis seems to be the answer to overcome the situation. Face recognition is an additional and superior technology that can improvise video analytics with facial identification.

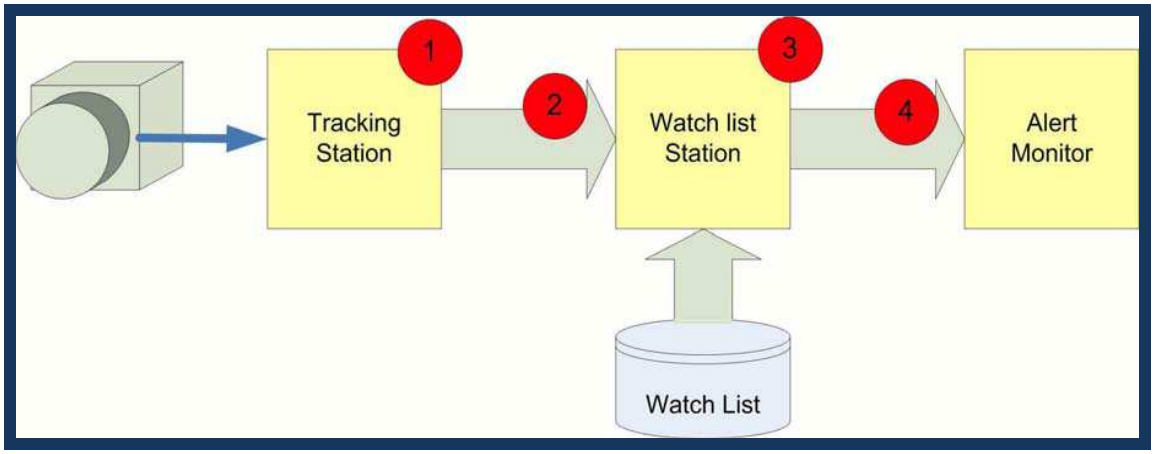
State of art DVR systems already offer some intelligence support to the security staff. Features like motion detection and intelligent monitor switching strategies are emerging. In most cases video surveillance systems are used to observe human beings and their behaviour. It makes sense to move one step further and introduce face recognition in order to identify persons registered in the watch list.

Video surveillance systems can benefit from face recognition in highly populated areas like sports arenas, airports, railway stations, convention centres, shopping malls, casinos etc. The likelihood of identifying suspicious individuals like terrorists, hooligans or known shoplifters is much higher if there is technical system in place that actively supports security guards to focus on most important security alerts.

Face recognition biometrics distinguishes cooperative application scenarios from the uncooperative application scenarios. Video surveillance is always un-cooperative application scenario. Biometric performance in uncooperative scenarios can be significantly lower than biometric performance observed in cooperative scenarios like physical access control. Typically the (true) alarm rate and the false alarm rate are considered to evaluate the system. The (true) alarm rate (also called true-positive identification rate) is the percentage of those evidences from the suspects that cause an alert with the correct match occurring in the match list. The false alarm rate (also called (false identification rate) is

the percentage of those evidences from non-suspects that cause an alert. The higher the false alarm rate the more the security personnel is kept busy in following up false alarms.

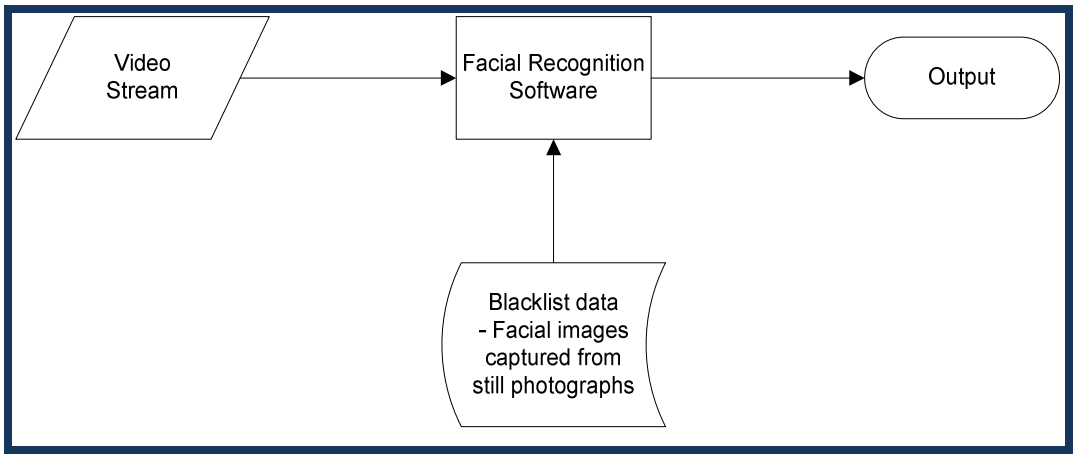
Assuming a customer would like to get 99 suspects out of 100 to be identified. This corresponds to the alarm rate of 99%. The threshold of the system has to be configured to meet the alarm rate. Let's assume further that the threshold set of an alarm rate of 99% will cause a false rate of 20%, for a given watch list, lets say 500 suspects. This would mean that, on an average one out of five alerts would be false.



IMPLEMENTATION DETAILS

The test was performed at Home Team Academy, Singapore. The purpose of the test is to pick out POIs (Persons of Interest) walking by themselves or in a group, within the view of the surveillance camera or CCTV for both indoor and outdoor settings. The face recognition application covers requirements for both security (in a controlled environment) and intelligence (both controlled and uncontrolled environments) applications.

The application is provided with 2 sets of black lists, one list is prepared from still photographs and another list is collected from video footages. The images are enrolled into the database.



Scenario 1:

A video footage of CCTV surveillance took at a close up video of the people at the entrance of an office .The camera is placed at the entrance attached to the top of the gate, so that all the people moving towards the gate are recorded.

Some people may have their hats pulled down onto the face, some may have reflecting glasses, turned heads etc, these cannot be eliminated which reduce possible recognition rates.

We have tested the video with reference to the 2 blacklist images enrolled in the database. The application captures each frame of the video and processes the frame for a face and sends the frame to the watch list station. The results of the identification are displayed as alerts with their identity and matching score on the User Interface.

The results under this scenario are as follows:

Identification: 70%.

Not identified: 30%.

Scenario 2:

A video footage of CCTV surveillance took at a close up video of the passengers passing towards the camera in a metro railway station .The camera is connected to the ceiling in the middle of the station where all the people passing through the view of the camera were recorded.

Some people may have their hats pulled down onto the face, some may have reflecting glasses, turned heads etc, these cannot be eliminated which reduce the possible recognition rates.

We have tested the video with reference to the 2 blacklist images enrolled in the database. The application captures each frame of the video and processes the frame for a face and sends the frame to the watch list station. The results of the identification are displayed as alerts with their identity and matching score on the User Interface.

The results under this scenario are as follows:

Identification: 70%.

Not identified: 30%.

Scenario 3:

A video footage of CCTV surveillance taken at the top of an Escalator, so that persons moving up on the escalator and passing through the escalator are covered .The camera is placed at the starting point of the escalator where the people leave.

Some people may have their hats pulled down into the face, some may have reflecting glasses, turned heads etc, these cannot be eliminated which reduce possible recognition rates.

We have tested the video with reference to the 2 blacklist images enrolled in the database. The application captures each frame of the video and processes the frame for a face and sends the frame to the watch list station. The results of the identification are displayed as alerts with their identity and matching score on the User Interface.

The results under this scenario are as follows:

Identification: 50%.

Not identified: 50%.

Scenario 4:

A video footage of CCTV camera which is placed outside of a shopping mall, which covers the people moving in front of the mall .The camera, is positioned at the top of a wall outside the mall,

Some people may have their hats pulled down onto the face, some may have reflecting glasses, turned heads etc, these cannot be eliminated which reduce possible recognition rates.

We have tested the video with reference to the 2 blacklist images enrolled in the database. The application captures each frame of the video and processes the frame for a face and sends the frame to watch list station. The results of the identification are displayed as alerts with their identity and matching score on the User Interface.

The results under this scenario are as follows:

Identification: 30%.

Not identified: 70%.

The variations in the percentage of identification from those of not identified are due to the poor video footages. The higher the quality of the video footage, the more is the percentages of identification. The enrolled facial samples should be of good quality without any interlacing and blurred images. The video footages should also be of good quality without interlacing.

HOW DOES SYBERFACEID-ALERT WORK?

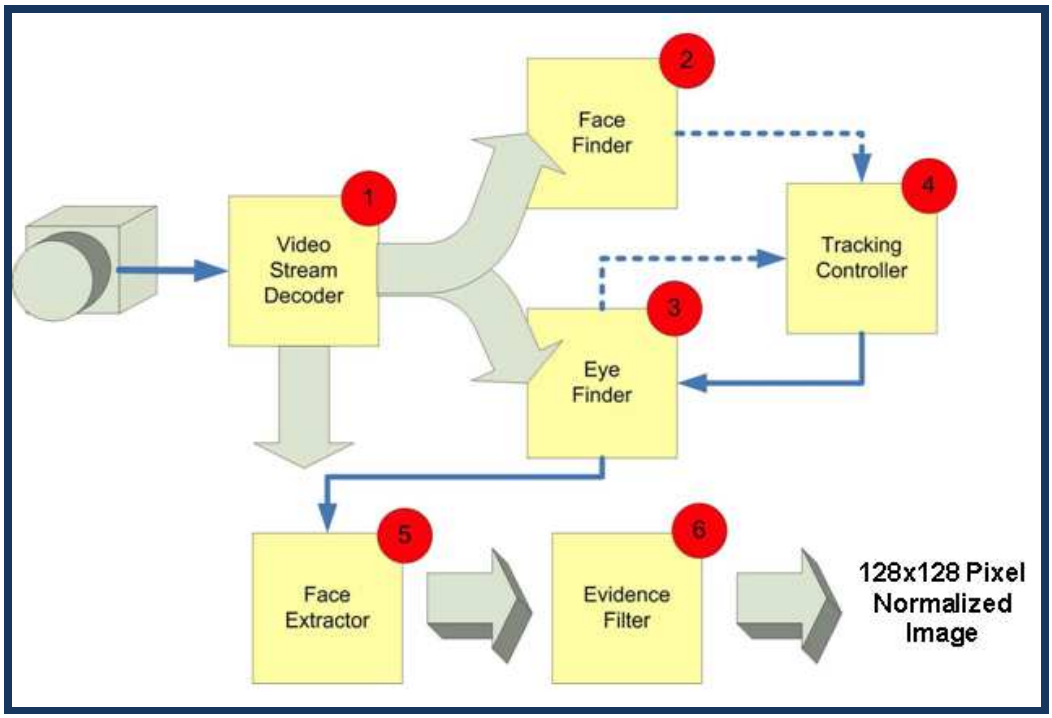
The principle is simple: Every camera generates a real-time video stream.

1. The video stream is analyzed in real-time by a so called tracking station regarding the visibility of human faces.
2. Together with time and location information, the extracted facial image is sent to a watch list station for checking.
3. The watch list station has access to a set of facial images of suspect persons, the so called watch list, and compares every incoming facial image with that watch list.
4. In case of a match, an alert is raised and connected alert monitors that are used by security guards will receive and display those alerts.

The information processing chain can be perceived as a sequential filter that first filters for any persons showing up with their faces visible and next filtering for very specific persons who are known as suspects.

The Tracking Station

The tracking station provides the ability to determine faces and their exact position in a frame at a specific timestamp. i.e., the tracking station analyses the content of the video stream and delivers information on facial objects.



Internally, the tracking station is designed for conducting parallel tasks and optimized for speed. In future, those tracking components will be part of intelligent cameras and facial image streams can be delivered as a specific MPEG 4 channel.

1. The incoming video stream is decoded and the frames are extracted for processing. All frames get labeled by a timestamp and, a provided camera identifier for reference purposes.
2. As often as configured and as often as computing resources allow, frames will be exhaustively analyzed as to where faces are located. Position and face rotation angle is delivered by estimating both eye positions.
3. Based on the roughly estimated eye positions, the near neighborhood is analyzed for the most likely eye positions and those positions will be the output.
4. There is a strategy in place considering the tracking. Eye positions will be provided by either eye position indication coming from the face finder or by eye positions as computed from the previous frame. Timing and object movement speed is considered throughout this strategy.
5. The found face is cropped from the original frame, scaled and optionally rotated to a size of 128 x 128 pixels and converted to grayscale. The result is called normalized facial image.
6. Because people who move very slowly in front of the camera will appear multiple times within a sequence of frames, it is not necessary to flood the network and the Watch List Station with all found evidences. The evidence filter takes care such that only the most prominent normalized images, i.e. the most suitable normalized images for face recognition will be delivered.

The Watch List Station

The watch list station checks if the received objects, i.e. normalized facial images match with one of the faces that are registered in the watch list (database). The tracking station identifies faces by their appearance in space (camera identifier and position in a frame) and time (timestamp), whereas the watch list station generates alerts in case of a match with an identity registered in the watch list. Although the match will never be 100% identification, an alert will help to nab the wanted person.

The watch list station is a central station and is designed for scalability, robustness and high availability.

1. Incoming normalized images get queued for encoding such that no resources are blocked.
2. Encoders generate the biometric template from the given “normalized facial image”.
3. The evidences’ templates get queued for matching with the match list.
4. A set of matchers compares every evidence template with the appropriate set or subset of the watch list templates (reference templates are created during watch list enrollment).

Encoders (2) can be deployed on multiple dedicated servers in order to scale up with increased throughput.

Matchers (4) can be deployed on multiple dedicated servers in order to scale up with increased throughput and also with larger galleries.

Load control, queuing (1 and 2) and alert rising is done on the watch list station. In 24x7 operational scenarios, there is a need to have hot spare resources available. The Watch List Station can be extended by a standby Watch List Station that takes over automatically in case the active Watch List Station fails.

Encoder and matcher can also be extended by hot spare servers, such that any hardware failure will cause an automatic failover.

TECHNICAL CHALLENGES

From a biometric point of view, video surveillance represents the most challenging scenario because

- Observed persons are not cooperative
  - best case would be that they behave as if no video surveillance existed
  - worst case would be that they actively try to hide from the video surveillance system
- The ambient conditions are not optimized as the environment is pre-existing and most often cannot be changed. Video surveillance only provides additional support.
- Reference images that are used to setup the watch list might also be of bad quality. However, it is always better to have a bad image of a suspect than none.

Although this might appear low performance, there are scenarios where even worse equal error rates are sufficient to improve security. This is because an electronic system will be used as a facilitating tool. It never experiences attention slowdown and lets human inspectors focus on more relevant evidences compared to a video surveillance system w/o face recognition.

ENVIRONMENTAL CONSIDERATIONS AND RECOMMANDATIONS

A proper camera setup and environment is required to optimize the biometric performance. There are minimal requirements as well for reliable face recognition. Minimal Requirements Optimal Environment

- Eye distance larger than 32 pixels
- Pose deviation (Pitch, Yaw, Roll) less than (15°, 15°, 30°)
- 64 different levels of dynamic range in face region

For optimal performance the image characteristics should be

- Eye distance larger than 60 pixels
- Pose deviation (Pitch, Yaw, Roll) > (5°,5°,30°)
- 64 different levels of dynamic range in face region
- No glasses; no reflections, no shadowed areas
- Neutral facial expression (eyes open; mouth closed)
- Progressive scan signal video

Note: Faces may appear small and anywhere in the scene so that the Auto-exposure and auto-focus control of the camera won't adjust to the face region.

The more frontal the facial image taken within a surveillance situation, the higher the chances of getting a match .Clever camera installation can be a big advantage and achieve better recognition rates · i.e. people are walking towards the camera in a long hallway or go down an escalator · i.e. people get attracted by an advertising or information screen with a camera mounted close by. Changing lighting conditions, such as sun light at times of the day, prove to be the most complex challenge when it comes to illumination. Dim-outs or artificial lighting can affect the conditions in a positive way.

Difficulties such as hats pulled down into the face, reflecting glasses, shades, hands covering up half of the face, turned heads etc. cannot be eliminated and reduce possible recognition rates intrinsically.

Cameras and Digitizer

Of course the only way to gather real time video surveillance information is through cameras and digitizers. Every device in the signal processing chain contributes to the overall quality of the system. I.e. the better the image stream quality delivered to the face recognition system, the better the recognition rate. Some remarks from the face recognition perspective may help to get a better understanding about which characteristics are crucial for face recognition.

- Video Signal Transmission: The signal flow from the camera sensor to the face tracking process should be as little as possible disturbed. Ideally, the digital raw data as produced by the (CCD or CMOS) sensor will be used for tracking. Any conversation to analogue and back to digital representation through a frame grabber, any compression any density reduction will lower the biometric performance and potentially add more hardware needs (e.g. computing resources required for image decompression)
    1. Digital cameras designed for machine vision are most appropriate. The GigE Vision protocol for machine vision pushed by the AIA (Automated Imaging Association) solves this problem. GigE Vision cameras can be connected to SyberFace-Alert and are highly recommended.
- High resolution cameras: Due to the nature of video surveillance, faces almost never completely fill the picture frame. The minimal optimum distance between the eyes for faces to be found and identified is 60 pixels. An image taken at a resolution of 640 x 480 pixels

probably does not allow capturing more than 3 persons in a quality sufficient for a face recognition system.

- Professional cameras with C-Mount or even SLR allow finding an appropriate lens.
- Camera attributes: good white balance and low noise are helpful towards qualitative input
- Frame rate: a good surveillance system should deliver 25 frames per second. Information might get lost at anything below this rate.
- High dynamic range cameras can better deal with the visible light power spectrum of up to 130 dB (sun light) than standard cameras. It is important that not only the sensor supports more than 8 bit of density information, but also the complete transmission line from sensor to the tracking software. Analogue video signal transmission will not support that but the GigE Vision based transmission over Gigabit Ethernet does.
- Color cameras are not necessary for computer based face recognition. Though Security guards who handle the alerts and watch video streams, prefer having color information available, black and white cameras will deliver the same quality of facial recognition

.

ABOUT SYBERNAUTIX

Sybernautix is a leading provider of biometric security software products and expertise focused on meeting the needs of governments and commercial organizations worldwide. Sybernautix supports customers and system integrators in building enterprise solutions requiring the highest level of security, performance, scalability, reliability and privacy.

Governments, systems integrators, commercial enterprises, land, sea and air ports choose Sybernautix to meet their critical identity assurance requirements because, Sybernautix provide the most flexible, scalable and secure platform for managing identities in border management environments, civil identity programs, traveller ID applications and employee credentialing solutions.

We develop new software platforms and integrate solutions to provide better public access to data, promote mission success and improve internal productivity. Our biometric solutions are developed in conjunction with select partners and specialists that are experts in their fields. These will offer a biometric solution to a pressing business concern that affects our customers.

For further information please visit www.sybernautix.com or mail to info@sybernautix.com