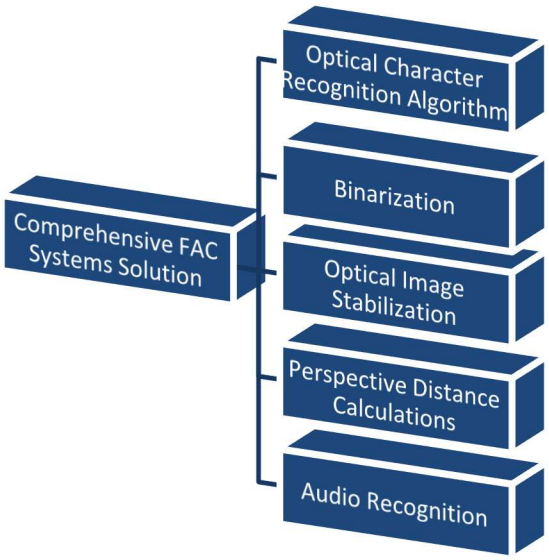
Audio and Optical Character Recognition for Terrorist And Active Shooters Threat Reduction

Farid A. Chouery, PE, SE
FAC SYSTEMS, INC.

Introduction

On April 23 of this year, U.S. Department of Homeland Security Chief John Kelly exclaimed, "I do not know how to stop homegrown terrorist attacks." This statement comes at a time in our history when effects of climate change, resource scarcity, geopolitical instability, and extremist propaganda spreading on social media platforms and through the internet more generally, have come together to create a perfect storm for "lone wolf/home grown" terrorism and active shooters incidences to increase dramatically (as they have over the last decade). Traditional law enforcement methods depend upon the practice of tracking known operatives' intentions and taking necessary action to prevent an attack. This technique has become difficult and antiquated as the terrorist organizations and active shooters now focus their resources on indoctrination of potential moles within a target population. Therefore, a new solution that involves pre-emptive identification of probable threats is necessary. To create a viable solution that can identify a threat quickly and alert appropriate on-site law enforcement, FAC Systems proposes a software development project that harnesses the power of optical character recognition and audio identification patterns.

By modifying existing optical character recognition, binarization, and other technologies to identify abnormal bulges, the shape of a firearm, abnormal infrared heat signatures, and other markers of a possible threat, law enforcement agencies and security officials would enjoy a significantly higher probability of identification of threats over traditional human monitoring practices alone. A sophisticated OCR algorithm and software package designed to identify threats by complementing existing surveillance techniques, camera infrastructure and security systems would give law enforcement the tools they need to preemptively address threats before an attack occurs.



When combined in a software solution, these systems of analysis can aid low-level surveillance staff in identifying potential threats and increasing response time. In the following section, each component of the system is explained in detail.

Solution Component: Optical Character Recognition

OCR is an existing technology and field of continuing research in pattern recognition, artificial intelligence and computer vision. Optical character recognition (optical character reader, OCR) is the mechanical or electronic conversion of images of typed, handwritten or printed text into machine-encoded text, whether from a scanned document, a photo of a document, a scene-photo (for example the text on signs and billboards in a landscape photo) or from subtitle text superimposed on an image (for example from a television broadcast). It is widely used as a form of information entry from printed paper data records, whether passport documents, invoices, bank statements, computerized receipts, business cards, mail, printouts of static-data, or any suitable documentation. It is a common method of digitizing printed texts so that they can be electronically edited, searched, stored more compactly, displayed on-line, and used in machine processes such as cognitive computing, machine translation, (extracted) text-to-speech, key data and text mining.

FAC Systems' solution would use this methodology by redesigning fonts against a set of shapes of interest (e.g. guns, knives, abnormal baggage, abnormal bulges, abnormal clothing, etc.) and deploy this technology in the fight against terrorism. Alone, this solution would be a powerful addition to the toolkit that law enforcement currently uses, and it would create a low-level security enhancement to enable quick decision making by surveillance team personnel. Below, technologies used to carry out OCR are discussed and explored for application suitability. OCR lies at the heart of the solution for identification of objects of interest, but other technologies could play an outsized role in surveillance improvement as well.

Solution Component: Binarization

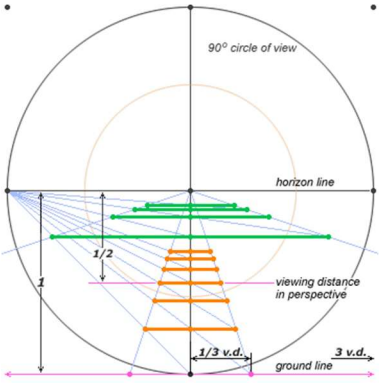
With binarization[1], the font[2] is overlaid with the lighting of the room, or the space on the font itself—assuming a black and white image. In order to accomplish this, the lighting source is set per-camera and captured through a specialized algorithm before any scanning of the image takes place. The created font that the algorithm searches for in the image would ideally be a pre-designed 3D CAD file. This software option with lighting is similar to 3D CAD drawings with light shining on them (i.e. 3DMAX Auto Desk product software). Therefore, when shining a light source on the font, the algorithm can identify an object of interest in three dimensions.

The 3D cad drawing overlaid by the lights in the room rotates to all options per pixels with aid of zoom and using an OCR algorithm. In this case, every pixel has a binary location number similar to a mouse location or coordinates in CAD, which uses a mathematical algorithm to extract fonts.

Solution Component: Perspective Distance Calculations

The perspective distance calculations determine the relevancy and probability of a match by object of interest (created font) per database. (To explain, we will use a "gun font" as an example."

Because of the relative distance between the subject in the image and the camera can vary greatly, there are limitations of the size of the gun to be determined. The size of the gun will not occupy the entire screen. The gun is proportioned to someone standing in the room and its size can be estimated for faster identification. The area where you expect for a gun to be in the image is predetermined. Also, relating to a camera: if the picture is enlarged or being smaller the perspective calculations change in the computer.



[1] The task of binarization is performed as a simple way of separating the text (or any other desired image component) from the background. This is done by using a greyscale and lighting to distinguish objects from background images.

[2] Here, Font refers to an object of interest, such as a weapon or piece of clothing that could be used to identify a suspect.

The computer scans the image in black and white vertically and horizontally and at every two pixels that are changing on the greyscale, the computer adds a pixel colored in red nearby. Thereafter, the image is screened to show only red and all of the other outlines or contours of the fonts will be shown. With this solution component, the selection can be designed to eliminate unnecessary lines or outlines. Generally speaking, this component idea would be faster because it includes less pixels to compare with the font or the outline or boundary of the threat, as such, a line drawing in 2D CAD for a gun or a rifle can now be used and identified. This same strategy could be employed to identify anomalies in body size by using finite element analysis to measure biomass and test for deflection difference. Another option that maybe faster is to extract all contours and cross-examine against database information to identify patterns. Every pixel has a binary location number similar to mouse location or coordinates—just as in CAD, which uses a mathematical algorithm to extract fonts and contours. For a face, the extracted contours may be combined to form a new font for a face. The solution would allow for measurement of the location of a contour of pixels and we can recognize the patterns through a mathematical algorithm. A complement to this process would be to utilize optical sorting image processing, which would take over and manipulate the raw data to extract and categorize information about specific features.

Solution Component: Red Pixel Image Augmentation

Another component of the solution would be to assign a number that relates to a position on the screen for every positioned red pixel superimposed over the image post-binarization. Then, use a mathematical algorithm relating the relation of the red pixel to the remainder to identify an object of interest (i.e. to convert to vector space in CAD). The algorithm can be designed as such: if several pixel lines have the same slope or radius, then a line equation or a circle equation is identified and the computer can then skip all of the other outlines that do not satisfy the equations. The CAD database of pre-designed "objects of interest" can be easily configured to match the mathematical equation. This procedure can be done by selecting a local origin for the red pixel and fitting it into the vector space equation. (A similar algorithm is used to convert from a PDF drawing to CAD.) This set of instructions may slow the computer down due to the number of mathematical computations, but accuracy of identification would likely increase.

Solution Component: Audio Recognition

In addition to the technologies discussed in previous paragraphs, in some cases audio would be monitored. It would be advantageous to also detect the sound of gun shots and explosions for post-attack forensics. An early proposed solution by others for shootings and explosions was to insert at least three microphones in some city. When a shooting or an explosion happened, the computer would interpolate the audio signal and locate the exact location of the shooting or explosion. In the previous section the element of surprise was taken away by alerting the authorities of terrorist's activities in time. In this section, the audio component is used post-

attack. Even as an attack as not been prevented using this technology, it would be incredibly helpful in diffusing ongoing attacks, such as the Paris attacks that lasted for hours before police ended the violence. Many lives could be saved by implemented such technology on a wide scale.

Speech recognition is also a component of the solution that is already developed by the National Security Agency (NSA). The technology could also be used in venues where large amounts of financial transactions take place. When keywords and statements are used by an assailant, such as, “put the money in the bag,” “get on the ground,” etc. the threat could be immediately recognized by a computer which would relay the threat to authorities.

To use all of these components in a single software solution would require performance in discreet time every millisecond. Therefore, the more scanning at milliseconds the more successes will be achieved. It is possible to scan first using Binarization to find the boundaries then resume using Optical Character Recognition to detect where the geometry is already identified and there are less pixels to compare. The speed of identification is very important and allows one computer to scan more than one camera and reduce costs of having one computer per camera. It is assumed that the database of all possible fonts would be in the thousands and hopefully not in the millions.

Conclusion

This solution would overcome multiple problems with current methods of identification of suspects and potential terrorists and active shooters. Right now, incident investigation is all reactive – the technologies listed do currently exist in their component parts and are used independently of one another to investigate post-attack. It is very difficult to discern a suspected terrorist and active shooters from a civilian while using current technology in real time, and methods for front-end detection are dependent upon the staff that are monitoring equipment. Current technology uses facial recognition, and audio identification, but does not combine OCR algorithms, font audio recognition, optical image stabilization, and other recommended technologies into one package where the technologies can work in concert with each other to give an edge to surveillance personnel. Front-end detection using the suite of technologies listed in the FAC Systems solution would be a cutting edge predictive analytics platform and would change the way that law enforcement performs security services.