Abstract

Identification ensures security. This paper provides an introduction and analysis into the biometric realm of facial recognition security. An in-depth overview of facial recognition, as well as its potential and current uses will be examined. Results of various vendor tests are analyzed along with the many algorithms supporting this technology.

“Absolute anonymity breeds absolute irresponsibility”

-Scott McNealy, Sun Microsystems Inc.

On September 11th of 2001, foreign high-jackers successfully boarded commercial airplanes and crashed them into civilian and government targets. Initially, both citizens and government officials were shocked. The sudden effect soon wore off and we asked ourselves as a nation, how could this happen? The answer is simple: our country was not capable of identifying dangerous individuals. The United States’ intelligence agency failed to properly identify the terrorists when they entered the country, and unfortunately, the lack of identification compromised our security with a tremendous loss of life. The United States will not make the same mistake twice. Through the funding and implementation of new programs and advanced forms of authentication, the government hopes to patch the existing holes in security. One must always provide proof or evidence of identity; however, birth certificates, driver’s
licenses, and other official documents are easily forged. New techniques must be adopted to prevent further breaches of security. Biometrics is one such technique.

Biometrics stands out in the security field because it does not use conventional methods of authentication. Traditional passwords and ID cards are obsolete. Instead, fingerprints, hands, DNA, faces, eyes and voices may be used to identify an individual. The password is an iris, a fingerprint, or measurements on a face. The properties of each technique are very difficult to duplicate whereas conventional passwords are often lost or forgotten. No system is perfect; however, biometrics presents a better solution to a growing problem.

The use of biometrics as forms of identification is a worldwide reality. The U.S. government, in particular, has taken a keen interest in this subject. DARPA is presently developing a HUMAN ID system which can recognize human faces at a distance of up to 150 meters.¹ New, complicated algorithms, coupled with superior technology allows for our society to use our high tech innovations. Biometrics is a new field; however, it has come a long way in a very short amount of time. The U.S. National Biometrics Test Center says the biometric market rose from $6.6 million in 1990 to $63 million in 1999. Cahner In-Stat Group predicts biometrics will reach $520 million by 2006.² A biometric system that stands out is facial scanning technology. Relatively inexpensive, the identification device could be melded into our infrastructure quickly. This paper discusses the potential of using facial scanning technology in closed-circuit television cameras or CCTV cameras. The weaknesses, benefits, and any major conflicts will be analyzed. I will also discuss the inner workings and algorithms of the system.
Facial scanning technology is valuable because it does what hand, fingerprint, and iris scanning systems cannot. An individual under surveillance does not have to willingly participate in a biometric scan. Fingerprinting, hand geometry, and iris scanning involve time and an army of personnel. Each subject at the checkpoint must stop, observe the instructions, and participate in the checkpoint. Lines and delays will naturally result. Our world demands efficiency. If a facial scanning system could scan faces using existent video cameras, would not that solution fit best? Time and energy would be saved. I will point out more opposition to the use of “willing” biometric systems as I proceed through this paper.

A few big name vendors of face biometrics include Imagis, Identix (recently merged with Visionics), Viisage, and BioDentity. Companies claim the use of security cameras as biometric identifiers can almost guarantee accuracy, integrity, and security in almost any situation. A normal, one hundred dollar CCTV surveillance camera may be upgraded with a little hardware to “rip” faces. Our current infrastructure of surveillance will not undergo dramatic reform. Instead, the system would be able to perform additional complicated tasks, with little cost. The facial recognition technology is low cost.

A general overview of a facial scanning process is as follows. A typical facial recognition system using CCTV cameras works as follows. First, the technology automatically targets faces in the field of view of the surveillance camera. Each image is next measured according to the specialized unique algorithm of the product, and processed into a mathematical face-print. The code is encrypted to ensure the integrity of the face-print during its route to the central database. The face-print is compared to the
databases’ watch-list of criminals and terrorists. If the face-print does not match any
other face-print or image in the database, the image is supposedly deleted. However, if
the image surpasses a certain confidence threshold or a complete match occurs, a human
operator is informed and a silent alarm sounds.

Facial scanning has immense potential. If a biometric system were implemented
in the correct manner, everyone in the area could be potentially identified with great
accuracy. Even in a large group, an individual cannot hide. The detection mechanism
detects multiple faces in one frame of video (one frame of video may be 1/32nd of a
second). Currently, fingerprinting is the primary biometric identification method; the
FBI Criminal Justice Information Services Division has over two-hundred and nineteen
million fingerprints cards and forty-two million digital fingerprints on file. The next step
is to compile a completely new federal database of face-prints, or simply match the faces
to the fingerprints such as in the FBI’s Integrated Automated Fingerprint Identification
System (IAFIS). If this IAFIS database becomes operational, it will receive an average
of over 5,000 images a day.³

The camera catches every action and records who is responsible for the action. If
a damaging crime occurs, police could be summoned to the scene. The police could be
briefed on the trip to the crime. The culprit would be caught and prosecuted with the
evidence on the camera. The threat from terrorism would also diminish. Homeland
security, an issue so prevalent today, would have another level of assurance. Known
terrorists could be apprehended in an airport before they board planes, enter schools, or
government institutions. Within a blink of an eye, their picture is snapped, the image of
the face is located, measured and processed, the mathematical face-print compared to a
database. The alarm may be silent as to not stir up any panic among the surrounding citizens. The subjects would simply be apprehended by undercover police, positively identified in another room, and questioned further. The flow of passengers at an airport would not diminish, because the whole pickup is completed in silence. All of these claims, however, assume the biometric system works flawlessly, when in fact, none do.

Cameras are always on and one hundred percent alert. Security guards and police, on the other hand, become fatigued and inattentive. Implementing cameras would cut down on the mistakes and miscalculations of human error. Police forces are currently using Identix face technology to capture the fingerprints and face image of suspects on a mobile device. This information is submitted to federal databases and establishes if the subject is wanted. Racial profiling would not longer be an issue for law enforcement and security agencies as technology would have highlighted the subject and not bias or prejudice.

The potential of this system is enormous; however, sacrifices must be made. The assurance of security does not necessarily overwrite freedom and privacy. A comparison might even be made to the novel 1984 written by George Orwell. Big Brother is a looming threat to those who cherish their freedom. CCTV cameras specially altered to adopt the facial scanning technology could know where someone is, their habits, routines, and current places of destination. Everyone will be constantly tracked and pinpointed at all times. Unfortunately, this information, although kept in federal databases, could possibly fall into the wrong hands. Advertisers and strangers could know all about a person’s private life and tendencies. Personal and sensitive information would be stored
and viewed by others. Is this a breach of the fourth amendment, the right to privacy?

The amendment states:

“The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.”

To answer this question I will use another question. What does “unreasonable search and seizure” constitute? The government defines this term, and it is all too clear they prefer security.

It seems, however, that facial technology sprouted post September 11th, 2001. Before the attack, citizens did not consider homeland security a main issue. Now, homeland security is the only subject people talk about. Facial technology sounds enticing, and more and more companies, politicians, and government and privately owned institutions are adopting its use. The foot is in the door, and we may not be able to get it out. Facial technology is already used in airports, stadiums, and government institutions. Britain alone already has 1.5 million CCTV monitoring streets, banks, buildings, and roads.4 The U.S. uses 2 million cameras for the same purpose. Video surveillance is already fully adopted in our lives, is facial recognition technology next?

A few weeks ago I visited San Francisco International Airport. CCTV cameras encased in a tinted black glass bubble were spread all along every wall. It seemed as though there was a camera for each person in the area. Even without the use of facial recognition, the authorities were fully aware of every action and person. I felt very uncomfortable because no matter where I went, I was being watched.
The right of privacy is not the only fatal flaw of this system. False identifications will always plague any security system. A false identification is a false accusation, and this may damage a reputation severely. Our society should not be paranoid. Not everyone is a terrorist, and therefore, everyone should be assumed to be one. The governments stance is firm, terrorists could be anywhere, and must be sought out everywhere. Bruce Schneider, a computer security expert says it best, “You end up with a society in which the database is more important than the reality”.5

As technology advances, the accuracy of the algorithms and the image quality will improve tremendously. Currently, the technology is neither reliable nor accurate. However, if the technology were reliable, should we still use it? The companies developing this new technology claim the project is not a national ID system and that the images of those not matched are just compared to a database and trashed. But who knows if this is the case; as mentioned earlier, the HUMAN ID project funded by DARPA is seeking to identify all. Congress must make sure the watch-lists occupying the federal databases do not include non-criminal civilians.

The major developers of face scanning biometric systems claim their facial recognition system is reliable. According to their data, a suspect will be accurately identified sixty to ninety percent of the time. These numbers are calculated using the preexisting mixed images (some good quality, others not) in the FBI databases. Joseph J. Attick, CEO of Identix, further claims that the face-print

“…contains the physical measurements of your skull and your face and is identity-specific. It is unique to you, it does not change with aging and it is not affected by viewing conditions and also not affected by superficial disguises. If you put facial hair, mustache, beard, change your hairstyle, that is not what the facial print is doing.”
However, these claims are nothing more than marketing tactics. The true data comes from the real world. Very few civilians know that facial scanning technology was already in use prior to September 11th. PISCES, or Personal Identification Secure Comparison and Evaluation System, compares passengers with a collection of images of terrorists. The system was operation and did not recognize a single terrorist; it failed completely. Third party companies and organizations have conducted many tests on the current facial recognition system in order to test measuring algorithms, identify future areas of research, and assess the state of the art technologies. The first of such tests include FERET, or the Facial Recognition Program, which included a database of over 14,000 images of over 1,000 individuals. However, FERET’s successor, the Facial Recognition Vendor Test 2000 (FRV 2000 sponsored by DoD, Technology Development Program Office, the National Institute of Justice, and the Defense Advanced Research Projects Agency), is the new standard for assessment. Their results: the systems simply do not work. Two measurements are crucial in this conclusion: the False Acceptance Rate (FAR), or percentage of imposters allowed, and False Rejection Rate, or the percentage of valid users rejected, also known as “False Positives.” Factors such as lighting, disguises, distance of camera, facial angle, expression, resolution of the image, and large groups of people all contribute to the overall failure including high FAR and FRR rates. The results of one case study from the FRV 2000 test are as follows.

In this case study, a camera was placed at a baggage claim. As passengers retrieved their baggage, an image was taken of them and compared to a watch-list database. Passengers knew a camera was watching them.
<table>
<thead>
<tr>
<th>Factor</th>
<th>Expected % if only this factor is variable</th>
<th>Watch-list image Database size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expression</td>
<td>95%</td>
<td>225</td>
</tr>
<tr>
<td>Illumination between mugshot and image</td>
<td>95%</td>
<td>227</td>
</tr>
<tr>
<td>Angle of Face</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15-25 deg.</td>
<td>99%</td>
<td>200</td>
</tr>
<tr>
<td>40 deg.</td>
<td>88%</td>
<td></td>
</tr>
<tr>
<td>60 deg.</td>
<td>58%</td>
<td></td>
</tr>
<tr>
<td>Resolution</td>
<td>98%</td>
<td>101</td>
</tr>
<tr>
<td>Distance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 meters</td>
<td>60%</td>
<td>185</td>
</tr>
<tr>
<td>3m</td>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>5m</td>
<td>30%</td>
<td></td>
</tr>
</tbody>
</table>

To analyze these results, I must point out the significant flaw of these statistics. Only one factor is a variable, that is, all other factors are ideal. In the real world, almost all of these factors are variable because nothing is controlled. A person is not instructed to look into the camera or take off his or her tinted sunglasses. Therefore, the actual real life statistics are far worse. The marketing divisions of facial scanning producers advertise using statistics similar to these.

The demonstration of facial-recognition at Palm Beach Airport completely supports the failure of this technology. The test database included a set of 250 images. The system was to match the database against the 5000 passengers and employees who traveled through the airport a day. According the American Civil Liberties Union (ACLU), “the system failed to match volunteer employees who had been entered into the database fully 503 out of 958 times, or 53 percent of the time.” The director of the
National ACLU’s Technology and Liberty Program, Berry Steinhart, even went as far as to say, “Under real world conditions, Osama Bin Laden himself could have easily evaded a face recognition system.” Furthermore, the false rejection rate (FRR) averaged two an hour, which is unacceptable. The conclusion, facial technology is not ready to use in a real world setting because it is unreliable (FAR too high), inaccurate, and burdensome (FRR).

The technology is currently implemented throughout the world, especially in mugshot police systems. Many airports are also using the same technology. Iceland International Airport and Oakland Airport are experimenting with the technology. During SuperBowl XXXV in Tampa Bay, a facial detection system was used to prevent terrorism. Over 100 casinos nationwide use the technology to prevent blacklisted gamblers and card sharks from gambling. The United States Army catalogs all the prisons in the current war with Iraq using biometrics. The Army is also pushing for the use of facial scanning technology in news cameras that videotape in Iraq. The idea is to look for wanted individuals in a crowd or rally. Oakland International Airport has invested over $2.5 million to install two Imagis technologies: the ID-2000 and CABS, or Computerized Arrest and Booking System. The West Virginia Department of Motor Vehicles checks for duplicate license registrations using the technology. Fresno Airport in California employs a different technology. The technology is similar to a metal detector. The subject stands in front of a large, aluminum-like tube. The machine snaps a picture of the subjects’ facial skeletal structure. Twenty six points on the face are converted to numbers and then compared. The use of facial scanning technology is not solely a U.S. based phenomenon. Israel uses facial recognition at the Israel-Palestine
border to protect the commuting workers. Criminal alert systems are in place throughout England. The English soccer stadiums have even gone as far as to prevent hooliganism using this technique. In Mexico, facial detection was used to deter people from registering more than once in the last presidential election.

Facial extraction of video surveillance is crucial. Therefore, the algorithms used to identify a subject from an image must be very accurate and quick. I will only discuss a few of the methods of facial finding and calculating general measurements. There are far too many calculations involved in these algorithms, so I will explain the basic foundation and concepts of these algorithms. It is important to note that many of these algorithms use vectors. An image is really just a two dimensional array of pixels for grayscale (0 to 255 or black to white) and a three dimensional array for color (RBG, red blue, and green).

The first algorithm discussed in the nonlinear detection filter. In general, filters are used for the facial finding aspect. The algorithm’s primary task is to apply a filter to the image and detect deep energy valleys. The peaks and valleys are also known as nodal points. This valley energy (how deep is the valley), is calculated by the following:

$$\text{Energy}_v(x_0) = \max_{x=1}^{k} \{ dR(x_0,x) + dL(x_0, x) - \alpha \cdot \text{abs}(dR(x_0,x) - dL(x_0,x)) \},$$

where $dR(x_0,x) = f(x_0 + x) - f(x_0), \ dL(x_0,x) = F(x_0 - x) - f(x_0)$. 

11
The filter is extremely important because after application, the landmarks (nose, mouth, cheek bones, and eyes) are easy to detect. However, after computing the various energies, certain unwanted areas appear as landmarks, such as hair. By obtaining certain known values within a specified range, the system may quickly discover the eyes and mouth. After identification of the two primary landmarks energies of these regions are computed. The sum of the valley energy, the edge magnitude, and the image intensity define the energy. However, this is still not good enough, for unwanted parts of the face might be in the defined range. Therefore, one may take an estimation of the region containing the facial landmarks. The region of interest is determined by first estimating the shoulder line which was set to be two-thirds the width of the largest foreground width, which is located on the lower part. Then an upper line is estimated by finding a position above which the image intensity values are low and below which the values are high. Finally, the face is estimated by taking the mean width of the foreground region above the shoulder line. Basically, the face is divided into a grid. The prime region of
interest is from temple to temple and down to the mouth because this area tends not to be
affected by facial hair, eyeglasses, or age.

As I said earlier, images are treated as multi dimensional pixel arrays.
Eigenfaces, eigenvectors, and their corresponding eigenvalues, are then computed. These
peaks and valleys of the person’s facial features are translated into a mathematical
vector. The concept behind Eigenfaces is as follows. ‘A set T1, T2, ..., Tm are used as
the training images. A vector of length N represents each image where N is the number
of pixels of the image. The average face is defined using the following formula: \( A = \frac{1}{M} \sum_{n=1}^{m} T_n \)  
PCA, or Principle Component Analysis, is used to find a set of M
vectors. The covariance of the matrix is:

\[
C = \frac{1}{M} \sum_{n=1}^{M} XX^T = \frac{1}{M} * YY^T
\]

where \( Y = [X_1 * X_2 * ... * X_m] \)
To determine the eigenvectors, an MxM matrix \( L=YY^T \) is constructed where \( L_{mn}=X_m^T X_n \).  The M eigenvectors \( (u_l) \) are found for \( L \). The eigenfaces are then computed using:

\[
u_l = \sum_{k=1}^{M} v_{lk} X_k, \quad l = 1, ..., M
\]
These eigenfaces are ranked according to the associated eigenvalues. Faces are projected
into “face space” using \( w_k = u_k^T(T-A) \). These vectors may then be used in a standard
pattern recognition algorithm such as the nearest neighbor algorithm using Euclidean
distance.

Identix Faceit© ARGUS technology employs the use of a facial template. Two
types of facial templates may be chosen. The first is the vector template, which is only
88 bytes and is created in 0.8 seconds (using a Pentium III CPU) using only two
Megabytes of memory. This scheme is only used for one-to-many searching because it
is fast. The other facial template is an intensive template. The template is only 4.5 Kilobytes in size, and is used to compare against a smaller set of images, on the order of N%, normally 5% or less, or the ordered matches following the vector template search in a 1:N matching application. These vectors are also used in verification operations. Templates are also useful because they can not be reverse engineered, so the facial image may not be reconstructed. FaceIt© utilizes the Local Feature Analysis developed by Joseph J. Attick, Dr. Paul Griffin, and Dr. Normal Redich (all of Identix). The identification process first locates the face (using the above technique). The next step is for the system to identify the patterns of points (such as the nose, mouth, etc) which will differentiate and distinguish the image from others images in the database. Next, the system creates patterns and compares them to the patterns of the subject’s face. The human face has eighty nodal points; however, the recognition only requires the use of 32 to 50.

Another company, Viisage Technology, uses a completely different method of comparison. The images are compared to one-hundred and twenty eight models of faces on its system. The image is then given a number according to its similarity to the models. The numbers are compiled to make this version of the face-print.

Another leader in this industry, Imagis, uses a completely separate process. Spectral analysis and 3-D modeling constitute the main methods of face extraction. Whereas Identix only uses a small quantity of nodal points, Imagis’ ID-2000 uses approximately 700 facial descriptors. After the location of the facial descriptors, the face is transformed into a surface model of the head; with values of pitch, roll, and yaw.
included. Once facial rendering is completed, another algorithm is used to create a unique wavelet.\textsuperscript{20} The face-print size of this system is 1400 bytes.

Other algorithms compare the images, not the face-prints. The first is the Hamming Distance Algorithm. In this scheme, images are superimposed on one another and the image is viewed as multiple polygons. The area of the union of two overlaying polygons is computed. If the region measured is small, the accuracy is high. This distance algorithm is not accurate unless the two images it receives are similar in lighting, size, and facial angle. Another method is the medial-Axis transformation. Also known as thinning algorithms, this concept attempts to extract a skeleton of the objects and compare the two using tree topology.

A few drawbacks of the algorithms are that they are affected by variables. The major causes of facial recognition failure are as follows. The pose of the subject must include both eyes. Any pose beyond 35 degrees results in significant loss in accuracy. Eyeglasses are another major problem. The glare of the eyeglasses and tinted sunglasses significantly reduce the algorithm’s effectiveness. The lighting is optimal if it is diffused ambient. However, if the lighting is considerably different between the image and the comparison image, then failure may occur. The image color must be color or on gray-scale. The image depth must be 24-bit and the resolution should be 320x240 at a minimum. The actual size of the head (in pixels) is another factor. The minimum face must be 20x30 pixels, but 80x120 is preferred.\textsuperscript{21}

Facial recognition technology takes into account liveness. By monitoring the subject over a few images, the image may be seen to blink or change facial expression. Fooling the system with a picture or painting will not work. This test usually takes 2-3
seconds. However, a single frame image test is also present, but not as accurate. Another intelligent idea incorporated in the system is the ability to detect poor images and alert a human operator that another image must be snapped.

After these algorithms convert the image into a “face-print” or equivalent, the image is compared through a database at a rate of 10 million comparisons a minute. That is 166,666 comparisons a second! Matching the face-print against a database is not complicated. A certain confidence threshold is established, so if two face-prints resemble each other, a list of potential matches is created. The technology is flexible, as the database may be any Oracle, SQL Server, MS Access, or ODBC-compliant database. The database has no size restrictions. The memory to execute the algorithms is relatively small, only 5 Megabytes for each image.

Facial Recognition technology is in a transition. The technology offers deterrence, but not much more. The systems are not currently accurate enough for deployment; however, they offer flexibility. The software is compatible with other widely used operating systems and database software and the requirements for computing power and memory are small. Only a small piece of hardware must be attached to the CCTV camera. The concept and ideas fueling the algorithms are increasing by leaps and bounds. More funding and resources are available which were not once present. In short, this technology will be ready for broad use in the future, but not today.
References

6. Phillips
10. Piller
14. Puckle
16 FaceIt© Frequently Asked Technical Questions
17 FaceIt© Frequently Asked Technical Questions
19 Rutherford
20 http://www.imagisttechnologies.com/Product/ID2000FRS.htm
21 FaceIt© Frequently Asked Technical Questions