To Be Taken At Face Value? Computerised Identification

Michael Bromby
Joseph Bell Centre for Forensic Statistics and Legal Reasoning
Glasgow Caledonian University and University of Edinburgh

Scientific evidence such as fingerprints, blood, hair and DNA samples are often presented during legal proceedings. Without such evidence, a description provided by the victim or any eyewitnesses is often the only means to identify a suspect. With the advent of closed circuit television (CCTV), many crimes are now recorded by cameras in the public or private domain, leading to a new form of forensic identification – facial biometrics. Decisions on how to view and interpret biometric evidence are important for both prosecution and defence, not least for the judge and jury who must decide the case. A jury may accept eyewitnesses as reliable sources of evidence more readily than complicated forensic or scientific evidence. False eyewitness accounts appear reliable when confidently presented to a mock jury. The decision-making process of the judge and jury may be seriously flawed if an eyewitness has made a genuine mistake. Using computerised recognition, the judicial decision of whether to accept an alibi or whether to accept the eyewitness account may be helped by removing the inherent uncertainty of human recognition.

Introduction

This paper looks at the use of CCTV to identify suspects in the UK. By comparing eyewitness identification to computerised facial recognition systems, it will be shown that the latter may be more reliable for suspect matching. The judicial decision to accept eyewitness evidence may result in an erroneous conviction if a witness has made a genuine mistake and unknowingly accused the wrong person. The reliability and admissibility of automatic facial matching may be questioned in court, with or without the aid of expert witnesses, as mistakes may be made in the management of computerised systems. With future technological developments, it may be possible to automatically identify patterns of behaviour through CCTV cameras, aiming to predict whether an incident will arise, thus allowing resources to be managed more effectively. With wide-spread CCTV presence in the UK, the growing potential of facial matching software could lead to the establishment of a national face database that, unlike a DNA database, could extend beyond forensic 'trail' evidence, leading to the identification of fraud, the uncovering aliases, or as a means to validate identity. As an example, Stockton-on-Tees, one of many UK cities to employ CCTV networks, has over 47 cameras in the centre alone and more cameras located in the residential areas of a town with less than 200,000 inhabitants.[1] This mass surveillance demonstrates the quasi-judicial decision made by the government and local authorities to show effective policing and increase public safety by detecting and preventing crime through CCTV monitoring.

[1] www.homeoffice.gov.uk/crimeprev/cctv.htm

Forms of Evidence

During the investigation of a criminal offence and the identification process of a suspect, the accumulated evidence will often fall into one of two categories: scientific forensic evidence, or eyewitness testimony. According to Locard's Exchange Principle (Saferstein, 1988), evidence, or some form of identifier, will be left at the scene of the crime, as every contact leaves a trace. Traditionally the identifier has been a fingerprint whereby improved scientific techniques have made the lifting and matching of prints a routine operation that is scientifically valid and reliable in court, particularly with computerised software for storage and searching. By detecting an identifier, the cause and effect principle can be used to prove that an individual has some physical connection with the crime scene: the effect is observed and the cause can be concluded, perhaps with the aid of further evidence.

With many crimes of assault or sexual abuse, there may be blood or tissue samples either at the crime scene or upon the victim. The discovery of DNA profiling to eliminate suspects from police enquiries, or to provide evidence that a suspect is associated with the crime scene, has become commonplace in today's policing. Fibre analysis is a less personal identifier, but an examination of particularly rare or uncommon fibres may provide additional evidence for corroborating other forms of evidence. Another less personal identifier is connected with ballistics – whereby a recovered bullet can be linked with a firearm, and the possibility that fingerprints can then link the weapon to a person. In the same manner, glass fragment analysis can also act as a link to the crime scene if such fragments are found upon a suspect's clothing.

Eyewitness testimony, however, is not a scientific analysis of physical or material evidence. It is a recall of events or an image of the perpetrator within the memory of the witness, which can be described verbally and incorporated into a written description. This may not be as reliable or as accurate in comparison to the scientific analysis of blood or fingerprints. There are many variables in the human capacity to perceive, store and recollect facial images, as the English case of R v Turnbull[2] highlight. Factors such as distance, duration, time-lapse, lighting conditions and several other factors were judicially noted in the appeal court judgment, which may improve or diminish eyewitness memory performance. This case is extremely significant as a legal precedent was set to ensure that the presiding judge issues a jury direction to inform jurors of the need for caution in accepting evidence when such factors may affect the process of identification. Such a direction has been commonly referred to as the Turnbull Rules and adopted by many common law jurisdictions, not as precedent, but as a highly persuasive case.[3]

The principal legal issue in such investigations is ensuring that only true and correct information is elicited from the eyewitnesses without fear of contamination or suggestion tainting the image stored in the memory. Such information must then be properly presented and interpreted in court. Of course, not all incidents will have witnesses who can go beyond a basic description of age, gender and stature but where observation was prolonged or where the victim was able to vividly recollect the face,

[2] [1977] QB 224; [1976] 3 All ER 549; [1976] 3 WLR 445; 140 JP 648; 63 Cr App Rep 132

[3] See R v Bourchielli [1981] VR 611 in Victoria, Australia; R v Mezzo [1986] 1 S.C.R. 802 at the Canadian Supreme Court.

cognitive interviewing (Geiselman et al., 1985) to obtain a description must be done carefully so as to be of further use in the process or identification. Such interview techniques have been developed with psychologists to ensure completeness and accuracy in the eyewitness's testimony.

Psychological studies have aided our understanding of how faces are recognised. Humans have a remarkable ability to recognise and differentiate between individuals and this is primarily due to the composition of the face. The face, like a fingerprint or a DNA profile must have unique characteristics to enable identification to take place. Therefore, if these characteristics are sufficiently unique, they may be used in a similar way to fingerprints and DNA profiles to identify or eliminate suspects from a criminal investigation.

The Face as a Biometric Identifier

The introduction of the face as a biometric identifier differs from the collectable forms of trail evidence left at the crime scene. As the face is a physical part of the criminal it cannot leave a trail at the scene of crime, therefore it may be seen as an exception to Locard's Exchange Principle. However, the face has possibly been seen by the victim, by other passers-by, or, more importantly, by any Closed-Circuit Television cameras situated in shops, banks, schools and other public places. With the increasing concern for public safety, virtually all UK city centres are under constant observation with a camera on most street corners. A stockpile of recorded footage permits suspicious behaviour to be elicited from the current archive if a criminal investigation so requires. In some criminal cases, the face may often be the only form of identification evidence if fingerprints or biological samples cannot be detected or relied upon.

The unique composition of the face has been used to provide support for analysing CCTV images of robbery or assault through the procedure of facial mapping.[4] Obtaining a similar viewpoint of the suspect in custody enables a comparison of the size and relative position of all visible facial features to be made with the image taken from CCTV footage. It is thought that the location of features within the face is one of the key paths to differentiation and identification in human observation, however, the clarity of images or the slight differences in photographic angle have caused expert witnesses to give conflicting views of opinion. Several cases of facial mapping have proceeded to the Criminal Division of the Court of Appeal, with regard to the admissibility or interpretation of such evidence and the ability of expert witnesses to comment on CCTV images.

Such an example can be seen in the case Church v HMA[5] whereby experts gave conflicting opinions both during the initial trial and during the appeal as to whether a positive identification could or could not be made. Church, the accused, had seen a photograph from a CCTV camera issued by the Crown Office and published in the local press. Church voluntarily attended a local police station, realising the similarity between himself and the robber. He later took part in an identification parade and was identified as the robber by three eyewitnesses. During the proceedings three expert witnesses stated that '*they had found significant differences between the face of the*

[4] Video superimposition was first held permissible as evidence by Stable, J for R v Clarke, Chelmsford Crown Court, 27 August 1993; as noted in the appeal judgment R v Clarke, [1995] 2 Cr App R 425

[5] [1995] SLT 604; [1996] SLT 383

robber and that of the appellant which they could not explain", and a further two experts reported that "*the definition of the robber in the video tapes was so poor as to preclude their use in any biometric analysis and photographic comparison*". The initial verdict was set aside and a retrial authorised which upheld the original conviction.[6]

Whilst eyewitnesses may make genuine mistakes, forensic examination of photographs or video stills can be a more reliable source of evidence. The jury's decision to accept incriminating expert evidence and reject eyewitness testimony may overly depend on the how the witness performed under examination. A nervous uncomfortable witness may be alarmed by court procedure and unduly discredited as an unreliable source. Alternatively a witness may make a confident false accusation either as a genuine mistake or with intent to pervert the course of justice. By providing a photographic analysis the judicial decision to accept or reject identification evidence can be made with greater confidence.

Computerised Recognition

As discussed, a human is capable of recognising and differentiating between faces as a common task, and also able to make expert facial comparisons using photographic evidence. However, the errors associated with human judgment, especially when an eyewitness may be subjected to stress and trauma during an event, may reduce the reliability of the witness and their evidence. Computerised facial recognition would eliminate the unaccountable variables associated with human memory. Many studies into computerised recognition have tried to adapt the psychological models of human recognition to work towards a fully computerised system of facial recognition. The adapted models are briefly described below:

Principal Component Analysis developed by Pentland et al. (1993) is based on feature identification: a face is identified and stored, the image is then analysed on the digital composition and the principal components or areas of light and dark are noted. For example, thicker lips will possess a greater surface area and will vary in brightness and contrast between individuals. Areas of light and dark along the edge of the face also serve to identify face shape and relative size. A unique set of data for each individual face is created which may then be used as a template or *Eigenface*[7] to enable the system to recognise the same face or, more correctly, the same set of data in the future.

An alternative model of Graph Matching (Wiskott et al., 1995) relies on the configural identification of a face. This relates to the measurable distances between features and the relative ratios of height and width rather than examining the features themselves. This model is similar to the task of facial mapping performed by imagery experts using CCTV and custody photographs. The eyes can be identified automatically and the locations of the remaining the features can be added in manually if required. A unique algorithm is created from the key points on the face; this algorithm is as unique as a fingerprint or DNA profile. Such systems rely on either automatic location or human input to position facial features, however, with either model there is still sufficient information to recognise and identify faces. The speed by which a result is

[6] 16 March 2001; The High Court, Edinburgh, Scotland

[7] An Eigenface is the term used for a fixed number of facial representations of light and contrast whereby a combination of several Eigenfaces can reconstruct any given face.

obtained would favour the automatic location, although a more thorough and reliable comparison can be made by using human input to locate the facial features.

The computational developments in automatic facial recognition systems are too complex to examine in great detail in this paper.[8] Various support systems have been developed to offer automatic recognition, which have been tested to varying degrees of success. The reliability and testing of support systems will be discussed later in this paper. Before any comment is made regarding reliability, it would be useful to examine the decision support possibilities offered by computerised facial matching software. There are two distinct approaches to employing facial matching techniques. One-to-one matching such as the facial mapping of suspect to criminal image will be discussed in relation to crime detection, and alternatively, one-to-many matching using a facial database will be examined as a means of crime prevention as well as detection. These two techniques have quite different implications for the justice system.

One-to-One Matching

One-to-one matching is a verification process of checking allegations or suspicions, whereby other forms of evidence must be present to suggest the involvement of a specific individual. As stated above, the one-to-one facial mapping technique provides additional scientific evidence to support the existing case. The use of this matching technique lends support to the judicial decision makers who must be persuaded beyond all reasonable doubt. By providing a detailed forensic examination over and above the traditional forms of evidence, the court can be reassured that the jury's decision to convict is secure and less likely to be overturned by a higher court.

If many one-to-one matches are performed then the process of investigation is nearing the alternative one-to-many matching technique, which is not based upon previously obtained evidence, but rather a process of chance investigation examining random individuals without an existing suspicion.

One-to-Many Matching

One-to-many identification can lead to the identification of a previously unconsidered suspect by searching through archive databanks. The police may then investigate the results to either eliminate or increase suspicion of the nominated suspects. This technique may be compared to the established investigative procedure of collecting fingerprints and searching large databases for a matching set. The counter argument to this comparison is that fingerprint matching must meet the minimum criteria of a point matching scale. This, therefore, shows that the reliability and accuracy of the matching procedure is more significant than the choice of biometric identifier.

The difference between verification of suspicion and actual searching for identification purposes has important legal implications. One-to-many matching would require an extensive database of facial images collec-ted either from police custody records or created from non-criminal records such as the face images held by the Passport Office or the DVLA. A fully comprehensive national database of all adult facial images obtained from non-criminal records would not be in accordance

[8] For further in-depth analyses of facial recognition systems, see the FERET and FRVT2000 literature quoted below.

with the data protection offered under the legislation governing the use of data. The Second Data Protection Principle[9] does allow disclosure of data from the data controller to a third party on the grounds of 'prevention or detection of crime'. This would not extend to the entire records of the passport office for one-to-many matching. The Principle applies to gathering data relating to a specific criminal act that would lead to one-to-one matching of single suspects as previously discussed. Relying on images obtained from previously convicted individuals would analogise facial comparison to fingerprint and DNA matching.

One-to-many matching as a means of crime prevention will be discussed in relation to the real-time surveillance capabilities of facial recognition systems below.

Reliability and Testing

In-house testing of facial recognition systems by software companies can be extremely subjective, with varying aims, goals and test data depending on the actual use and requirements of the tasks that the algorithms were developed to perform. Accuracy and reliability can only be assessed by comparing a product with standardised references or samples and further analysis by independent bodies. In 1996, the Californian Welfare Fraud Prevention Squad commissioned a report to examine the current techniques for biometric identification, with particular attention given to accuracy, closed and open search requirements and data collection error rates.[10] The report rejected facial recognition along with iris scan biometrics due to the lack of research and development at that time. Fingerprint analysis was recommended to the Californian authority, having open and closed search abilities, low data collection errors and a high user acceptance rate. The measurement of accuracy was conducted in such a way that not only could two facial biometric applications be compared, but such systems could also be assessed against other systems employing hand geometry, retinal and iris scans, fingerprint or other biometric algorithms.

To assess accuracy, the rate of false acceptance was calculated from the number of occurrences where an individual is wrongly accepted as a match, and secondly, the rate of false rejects was recorded when the system failed to recognise a legitimate match. Creating a graph of both values as a function of the threshold for recognition, the point where the two error rates intersect was taken as the accuracy of the system: the greater the cross-over value, the greater the inherent accuracy of the system. This type of assessment could then compare different techniques to select the most appropriate approach to combat a given problem. This independent analysis was also free from the bias and propensity of company sales literature so commonly accompanying in-house testing.

In 1996 the FERET Verification Testing Protocol for Face Recognition Algorithms was devised to provide an accurate and independent assessment of the reliability and accuracy of existing facial recognition systems as reported by Phillips et al., (1996). It also served to promote research in facial biometrics in academic and public / private sector industry, sponsored by the U.S. Department of Defense Counter-drug Technology Development Program. A Target set of 'known individuals' and a Query

[9] Data Protection Act 1998. Part I of Schedule 1.
[10] Prepared by The Biometric Consulting Group. http://www.bioconsulting.com/bio.htm

set of 'unknown faces' were presented to participating software developers in September 1996. Two versions of testing were administered: the first assessed automatic facial location, and the second version provided eye co-ordinates to assess the recognition performance of manual input systems. Enrolment and test data was collected according to strict guidelines to enable a fair comparison to be made. A scoring procedure was devised based on ROC graphs, originally devised for SONAR false recognition rates (Phillips et al., 1997).

Receiver Operating Characteristic (ROC) curves employ similar 'failure to enrol' and 'failure to acquire' rates to the Californian report. False match rate and false non-match rate were plotted on log-log scales as shown in figures 1 and 2. Further tests were performed until March 1997, when the Department of Defense published details of their results for the whole project. (Phillips et al., 1998) A significant increase in performance was seen for the general field of facial biometric comparison and for each individual algorithm-based system. Strengths and weaknesses of each algorithms were highlighted to facilitate further research to promote and improve the use of facial biometrics. However, it was evident that further research was still required if facial biometrics were to compete with other forms of biometric identification such as fingerprints, even though progress had been made in these areas over six months.

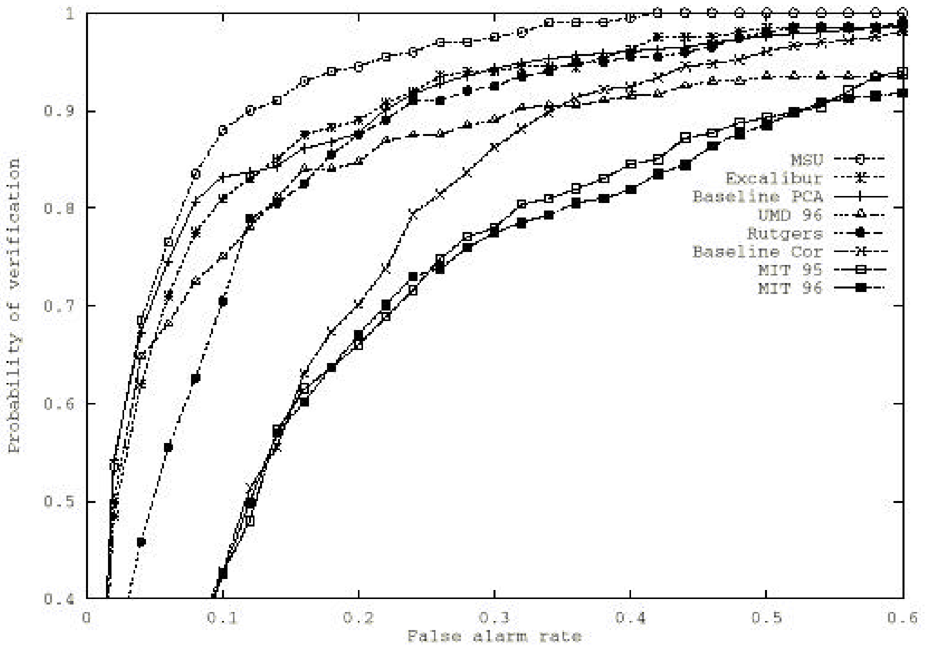

Figure 1. A ROC curve from September 1996 (Rizvi et al., 1998).
Note that to achieve a recognition probability of >90%, the system will exhibit a higher false alarm rate: the best system having a 10% false alarm rate, with others varying from 30 to 60%. By reducing the false alarm rate, the general trend was a reduction in verification probability.

A major fault of face recognition algorithms appeared to be sensitivity to variations in illumination caused by the change in sunlight intensities throughout the day. Changing the illumination may result in a significant performance drop. For some algorithms, this drop was equivalent to comparing images taken over the course of a year and a half apart. In addition, the position of the target face may also affect performance. A 15-degree difference in position between the query image and the database image creates a significant difference in the recognition threshold, and a difference of 45 degrees renders the system ineffective for recognition. These results give facial biometrics a disadvantage over other identifiers; however, by identifying the weaknesses further improvements may be seen in future developments. Despite the limitations from luminosity and face position, the process of identification was proven as successful and provides some useful resources to the field of law enforcement.

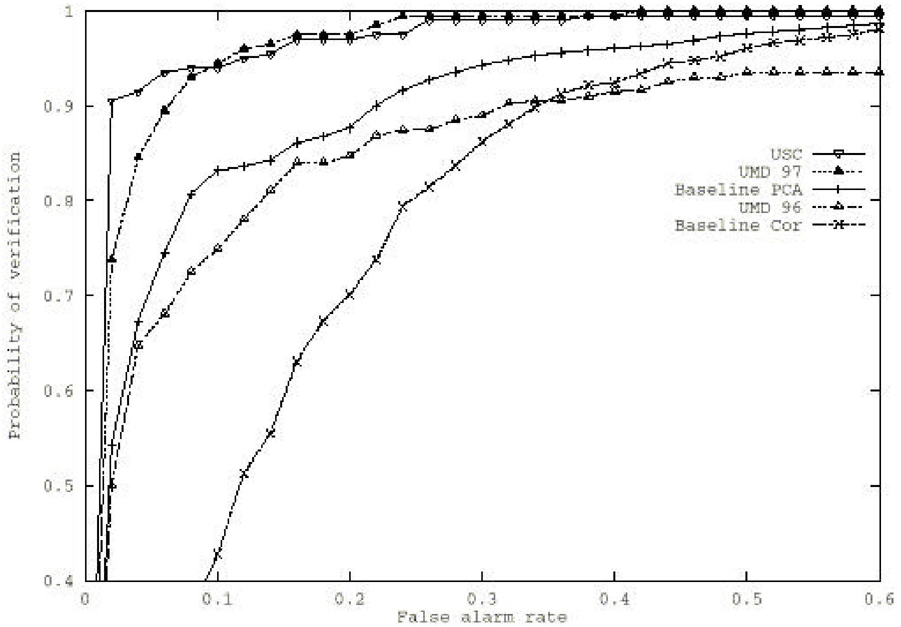

Figure 2. A ROC curve from March 1997. (Rizvi et al., 1998).
Over six months, the >90% recognition rate using the same FERET protocol exhibited a general reduction in false alarm rates. The best system exhibited < 2% false alarm rates, with others varying from 20 - 30%. This gave the system an increased recognition probability with no significant increase in false alarm rates.

With significant improvement in the development of facial algorithms for computerised recognition since the FERET programme, the capabilities of modern technology is in need of re-assessment. A new set of evaluations was set up by the Department of Defense as the Facial Recognition Vendor Test (FRVT) 2000 (Blackburn et al., 2001). Administered in May - June 2000, the test will provide the most up-to-date analysis of facial recognition. Significant improvement is expected for individual algorithms and computer systems, but it remains to be seen whether facial biometrics will be as effective or accurate as fingerprint analysis. With recognition performance and product usability being tested, it is likely that facial recognition software will receive a hallmark of approval and will cease to be dismissed from reports such as the Californian Welfare Fraud Prevention investigation into biometric techniques.

Employing Facial Recognition Systems

By combining automatic recognition technology and criminal databases of known offenders, computer systems to alarm law enforcement agencies as to the real-time presence of a known criminal have been developed. The first CCTV and facial recognition system in the UK was instigated by the Metropolitan Police in Newham, East London as reported by Thomas (1998). In the face of pressure from many civil liberties groups, the Mandrake system is able to examine every passing face and alert the police when an individual is recognised from the hit-list database. Despite analysing every single face in a crowd, information is only stored when a match is made, data from inconclusive analyses are discarded.[11] The system relies wholly on a graph matching system, analysing the area around the eyes and the nose, which can be converted into an algorithm without any manual intervention. This is a means of crime prevention as a potential criminal can be observed and intercepted before a crime is committed. Unwarranted surveillance in anticipation of any crime occurring by chance is not permitted under sections 28 and 29 of the Regulation of Investigatory Powers Act 2000. However, the selection of faces to be recognised and the location of the CCTV cameras may permit facial recognition systems to be used for crime prevention.

By placing a surveillance system in a unique area and attaching a database specific to known criminals who would operate in that area, a reasonable successful hit rate can be achieved without infringing the general privacy of the public. From the example put forward by Newham Council, other locations may be highlighted as target areas for particular types of offenders. Airports are prime examples of sites that are frequented by drug traffickers who will usually be operating in such locations. Security cameras are a regular feature of airports and may easily be linked to a recognition system containing an appropriate police database. Alternatively, a facial image bank of registered paedophiles may be used for one-to-many matching in restricted areas such as schools and playgrounds. Other such locations may also provide legitimate reasons for intrusive surveillance of particular offenders. The identification of known hooligans at sports matches may prevent violence and disruption at both national and international levels.

By collaboration with international law enforcement agencies, databases or sets of images of terrorists or drug traffickers can be used to combat crimes such as the use of aliases and false documentation. Automatic examination of faces is a less intrusive form of surveillance in comparison to taking fingerprints or blood samples from all airline passengers and can be done either as a one-to-one match to verify passenger identity thus creating a transparent security check involving overt passenger co-operation; or covertly as a one-to-many search to identify known suspects or potential criminals.

Conclusions

Computerised facial recognition systems must be seen as decision support tools rather than decision-making tools. It would not be acceptable to allow facial recognition software to make final decisions regarding identity at such an early stage in their development. The role of such recognition systems is to aid law enforcement agencies by suggesting possible matches that may be verified by human skill. Future advances may allow facial recognition systems to be as accurate and reliable as fingerprints or DNA testing. Existing systems can provide substantial support to investigations.

In comparison to eyewitness statements, the reliability of an automatic system can be verified and shown to be accurate to a measurable degree of certainty. When

[11] Therefore compliant with the Fifth Data Protection Principle ‘Personal data … shall not be kept for longer than is necessary…”

presenting eyewitness testimony in court, the reliability and accuracy of the eyewitness can never be determined unless some support is given in the form of photographic or video evidence. This will be accompanied by a one-to-one match with the suspect indicating the degree of similarity between the assailant and the accused.

The one-to-many matching techniques are of significance to the judicial process when considering how a crime has been detected, or prevented, and whether the correct procedures were adopted in relation to the Data Protection Act [1998] and Regulation of Investigatory Powers Act [2000]. Their use in situations such as airports and football grounds can allow checks to be made where the use of traditional identifiers such as fingerprints would not be feasible.

The supporting role of facial comparisons can aid the decision making process for the law enforcement agencies in deciding whether sufficient evidence is present to form a prosecution case. The judicial decision making process within the court by the judge and jury as to the admissibility, reliability and the weight of such evidence can also be aided by facial comparison evidence.

References

Blackburn, D., Bone, M., & Phillips, P. (2001) Facial Recognition Vendor Test 2000 Evaluation Report, **DoD Counterdrug Technology Development Program Office**, Defense Advanced Research Projects Agency, (USA, National Institute of Justice)

Geiselman, R.E., Fisher, R.P., MacKinnon, D.P. & Holland, H.L. (1985) Eyewitness Memory Enhancement in the Police Interview: Cognitive Retrival Mnemonics Versus Hypnosis, **Journal of Applied Psychology**, 70, pp. 401-412

Pentland, A., Moshaddam, B. & Starber, T. (1993) View-based and Modular Eigenfaces for Face Recognition. **Proceedings of the IEEE Conference on Computerised Vision and Pattern Recognition 1994**, pp. 84-91

Phillips, P., Moon, H., Rizvi, S. & Rauss., P. (1997) The FERET Evaluation, in: H. Wechsler (Ed.) **Face Recognition from Theory to Applications**. (Berlin, Springer-Verlag)

Phillips, P., Rauss, P., & Der, S. (1996) FERET (FacE REcognition Technology) Recognition Algorithm Development and Test Report, **U.S. Army Research Laboratory**, ARL-TR-995

Phillips, P., Wechsler, H., Huang, J., & Rauss, P. (1998) The FERET Database and Evaluation Procedure for Face Recognition Algorithms, **Image and Vision Computing Journal**, 16 (5), pp. 295-306

Rizvi, S., Phillips, P. & Moon, H. (1998) The FERET Verification Testing Protocol for Face Recognition Algorithms. **National Institute of Standards and Technology. Technical Report 6281**

Saferstein, R. (1988) **Criminalistics: An Introduction to Forensic Science** 6th Ed., (Upper Saddler River: Prentice Hall)

Thomas, R. (1998) As UK Crime Outstrips The US, A Hidden Eye Is Watching: Police switch on a camera that recognises your face, **The Observer**, 11 October, pg. 5.

Wiskott, L., Fellous, J., Kruger, N. & von der Malsburg, C. (1995) Face Recognition and Gender determination, **Proceedings of the International Workshop on Automatic Face and Gesture Recognition. Zurich 1995**