In: Book Title
Editor: Editor Name, pp. 1-19
ISBN 0000000000

Chapter 1

FORENSIC FACE RECOGNITION: A SURVEY

***Tauseef Ali*[*]*, Luuk Spreeuwers*[†] *and Raymond Veldhuis*[‡]**
Biometrics Group, Chair of Signals and Systems,
Faculty of EEMCS, University of Twente,
Netherlands

Abstract

The improvements of automatic face recognition during the last 2 decades have disclosed new applications like border control and camera surveillance. A new application field is forensic face recognition. Traditionally, face recognition by human experts has been used in forensics, but now there is a quickly developing interest in automatic face recognition as well. At the same time there is a trend towards a more objective and quantitative approach for traditional manual face comparison by human experts. Unlike in most applications of face recognition, in the forensic domain a binary decision or a score does not suffice as a result to be used in court. Rather, in the forensic domain, the outcome of the recognition process should be in the form of evidence or support for a prosecution hypothesis verses a defence hypothesis. In addition, in the forensic domain, trace images are often of poor quality. The available literature on (automatic) forensic face recognition is still very limited. In this survey, an overview is given of the characteristics of forensic face recognition and the main publications. The survey introduces forensic

[*]E-mail address: t.ali@utwente.nl
[†]E-mail address: l.j.spreeuwers@utwente.nl
[‡]E-mail address: r.n.j.veldhuis@utwente.nl

face recognition and reports on attempts to use automatic face recognition in the forensic context. Forensic facial comparison by human experts and the development of guidelines and a more quantitative and objective approach are also addressed. Probably the most important topic of the survey is the development of a framework to use automatic face recognition in the forensic setting. The Bayesian framework is a logical choice and likelihood ratios can in principle be used directly in court. In the statistical evaluation of the trace image, the choice of databases of facial images plays a very important role.

Keywords: Forensics, face recognition, Bayesian framework, biometrics.

1. Introduction

Face recognition is one of the most important tasks of forensic examiners during their investigations if there is video or image material available from a crime scene. Forensic examiners perform manual examination of facial images or videos to match a trace with an image of a suspects face or with a large database of mug-shots. The use of automated facial recognition systems will not only improve the efficiency of forensic work performed by various law enforcement agencies but also standardise the comparison process. However, until now, there is no automatic face recognition system that has been accepted by the judicial system. A face recognition system must be thoroughly evaluated and verified before it can be utilised for forensic applications. Biometric face recognition has of course been used for secure building access, border control, Civil ID and login verification. However, to date no automatic system exists for identification or verification in crime investigation tasks, such as the comparison of images taken by CCTV with available databases of mug-shots. State-of-the art face recognition systems such as [27, 25] could in principle be used for this purpose, but there are several issues, specific to the forensic domain, which have to be addressed.

First and foremost, the consequences of a wrong decision made by forensic face recognition are far more severe than for most other biometric face recognition applications. Current face recognition solutions [28] are generally not sufficiently robust [18] to the variability in appearance of faces due to variations in pose, lighting conditions, facial expression and caused by imaging systems such as image quality, resolution and compression.

Secondly, a score or binary decision based biometric recognition system is not suitable to the judicial system where the objective is to give a degree of support for one hypothesis against another incorporating the prior knowledge about the case at hand [9, 12].

Finally it should be mentioned that in the forensic scenario the quality of images available is generally low, e.g. images of a crime scene recorded using CCTV. These images usually have a low resolution and depicted faces are often not frontal and may be partly occluded.

The recognition task in the forensic framework can be carried out ”offline” in contrast to other applications where a decision has to be made in real-time, e.g. user access for a building or border control. Forensic face recognition therefore has fewer time constraints and to a certain extent human involvement is allowed and does not effect the overall objectivity of the system.

A related field of forensic facial recognition is forensic facial reconstruction which aims to reproduce a lost or unknown face of an individual for the purpose of recognition or identification [8]. Well known is the approach to reconstruct a face starting from the skull and using pins to model the thickness of the muscle tissue, then filling in the muscle tissue using clay and thus reconstruct the facial surface [26]

In this survey, we review existing literature on forensic face recognition. There are relatively few papers focusing on the forensic application of face recognition as most effort is put into the improvement of the technology itself. However, as the performance of face recognition systems improves the demand for application in the forensic domain also increases and, hence, there is a great need for integration of the technology with the legal system and a uniform framework for application of face recognition technology in forensics.

The remainder of the chapter is organised as follows: in section 2, the techniques and methodologies used by forensic examiners for the purpose of facial comparison are discussed. Section 3 presents a literature review of forensic face recognition. In section 4 we discuss the Bayesian framework and how it can be applied to forensic face recognition. Section 5 discusses reliability and court admissibility issues associated with forensic facial recognition. Section 6 presents conclusions.

2. Forensic Facial Identification

Facial identification refers to manual examination of two face images or a live subject and a facial image to determine whether they are of the same person or not. Facial identification methods generally can be classified into the following four categories:

1. *Holistic Comparison*: In this approach faces are compared by considering the whole face at once.

2. *Morphological Analysis*: In this approach individual features of the face are compared and classified.

3. *Photo-anthropometry*: This approach (sometimes referred to as photogrammetry) is based on the spatial measurements of facial features as well as distances and angles between facial landmarks.

4. *Superimposition*: In this approach, a properly scaled version of one image is overlaid onto another. The two images must be taken from the same angle.

The choice of a specific approach is usually dependent on the face images to be compared and generally combinations of these methods are applied to reach a conclusion. Apart from the above described general categorisation of facial comparison approaches, currently there are no standard procedures and agreed upon guidelines among forensic researchers. The process is very subjective and the opinion of one forensic examiner may vary from those of others.

2.1. Working groups

There are several working groups active in this area the aim of which is to standardise the procedure of forensic facial comparison as well as the proper training of facial comparison experts. One of the best efforts towards developing standards and guidelines for forensic facial identification is currently carried out by the Facial Identification Scientific Working Group (FISWG) [2]. It works under the Federal Bureau of Investigation (FBI) Biometric Center of Excellence (BCOE). FISWG is focusing exclusively on facial identification and developing consensus, standards, guidelines, and best practices for facial comparison. Currently they have developed drafts of several useful documents in this regard which include a description of facial comparison, a facial identification practitioner code of ethics and guidelines for training experts to perform facial comparison. These documents are available for public review and comments [2]. Some other working groups active in developing standards and guidelines for forensic facial comparison include the International Association for Identification [4] and the European Network of Forensic Science Institutes (ENFSI) [1]. The standardisation of the process of facial comparison and specific guidelines which are agreed upon by forensic community is, however, still a largely unsolved problem.

2.2. Manual facial comparison by the forensic expert

In this section we briefly review the forensic experts' way of facial comparison. The discussion is based on the guidelines set forward by the workgroup on face comparison at the Netherlands Forensic Institute (NFI) [5, 13] which is a member of ENSFI [27]. The facial comparison is based on morphological-anthropological features. If possible, for comparison images are used with faces depicted at the same size and with the same pose. The comparison mainly focuses on:

- Relative distances between different relevant features
- Contours of cheek- and chin-lines
- Shape of mouth, eyes, nose, ears etc.
- Lines, moles, wrinkles, scars etc. in the face

When comparing facial images manually, it should be noted that differences may be invisible due to underexposure, overexposure, low resolution, out-of-focus and distortions in the imaging process. On the other hand, due to similar limitations in the image formation process (low resolution, difference in focus and positions of the cameras used to record the images relative to the head and other distortions in the imaging process) may lead to different appearance of similar features in the facial images to compare. Due to the aforementioned effects, which complicate the comparison process, the anthropological facial features are visually compared and classified as: *similar in details*, *similar*, *no observation*, *different* and *different in details*. Apparent similarities and differences are further evaluated by classifying features as: *weakly discriminating*, *moderately discriminating*, and *strongly discriminating*. The conclusion based on the comparison process is a in the form of a measure of support for either of the hypotheses (images show faces of the same person vs. images show faces of different persons) and can be stated as: *no support*, *limited support*, *moderate support*, *strong support* and *very strong support* . The process is subjective and often different experts reach different conclusions. There is a great need to standardise the process. Use of automatic face recognition systems will considerably improve the speed and objectiveness of facial comparison and may also be helpful in standardising the comparison process.

3. Literature Overview

In this section we briefly review existing literature on forensic face recognition. This review focuses on work discussing forensic aspects rather than on work describing techniques for biometric face recognition. Surveys on the latter subject can be found in [28, 19].

3.1. Forensic biometrics from images and videos at the FBI

Forensic Biometrics from Images and Videos at the Federal Bureau of Investigation (FBI) is described in [20]. The paper gives a description of FBI's Forensic Audio, Video and Image Analysis Unit (FAVIAU) and the forensic recognition activities that they perform. Many of these activities are performed manually. Types of manual tasks include voice comparison, facial comparison, height determination, and other side by side image comparisons. Two types of examinations that involve biometrics are photographic comparisons and photogrammetry [7]. Currently, in both cases, the forensic examinations are performed manually. Photographic comparison means a one-to-one comparison of a trace facial image to facial images from suspects. The characteristics used in photographic comparison can be categorised into class and individual characteristics [24]. Class characteristics such as hair colour, overall facial shape, presence of facial hair, shape of the nose, presence of freckles, etc. place an individual within a class or group. Individual characteristics such as the number of and locations of freckles and scars, tattoos, the number of and positions of wrinkles etc. are unique to an individual and can be used to individualise a person. Photogrammetry [7] determines spatial measurements of objects using photographic images. It is used to determine e.g. the height of a subject or the length of a weapon used in a crime. In [20] several current and past research projects in the field of forensic recognition are discussed and also directions for future research on forensic recognition are proposed.

3.2. Facial comparison by experts

In [21] the need for facial comparison experts, their role in biometric face recognition development and their training are described. The paper describes the need for facial comparison experts to verify the results of future automatic forensic face recognition systems. It emphasises the systematic training of experts who will be working with these systems. For any future application of an automated face recognition system, the ultimate judgment will be the manual verification of the outcome of the system. Because the implications of an incorrect decision are severe the verification of the outcome of an automated system by an expert is very important. In case of fingerprint technology, there are many experts available working in association with the automated process. Compared to fingerprint technology, forensic application of face recognition is still immature and, therefore, requires even more this manual verification of the results by experts. This means in the near future more experts will have to be trained in order to use automatic face recognition systems. Comparison of images taken under controlled conditions such as passport photos or photos for arrest records requires less expertise compared to images taken under uncontrolled conditions such as snapshots and images from surveillance cameras. The experts also need training in legal issues because they will be working in the judicial system and will present their conclusions in court. The facial image examiners should be trained in three main areas:

1. General background on facial recognition approaches, which includes the history of person identification, current methods in biometrics, underlying principles of photographic comparison [24] and basic knowledge of image formation and processing.

2. Specific knowledge regarding the properties of the face such as the aging process, temporary changes (e.g., makeup and hair change), permanent changes (e.g. formation of scars, loss of hair, cosmetic or plastic surgery), structure of bones and muscles, facial expressions and the involved muscle groups and comparison of ears and iris.

3. Understanding of the judicial system, awareness of the implications of a testimony, admissibility issues of facial comparison in court, presentation of facial comparison results and processes in court and to laymen.

3.3. Forensic individualisation from biometric data

In [14] basic concepts of forensic science are reviewed. Also a general forensic face recognition framework is proposed based on the Bayesian likelihood ratio approach. Although this work is a comprehensive review of forensic concepts and provides a general description of the system, there is no experimental work described to prove the effectiveness of the proposed framework.

In forensic literature there is confusion between the terms identification and individualisation. If the class of individual entities is determined to be the source, it is called identification or classification. If a particular individual is determined to be the source, it is called individualisation. In the former case, the identity is called *qualitative identity* while in the later case the identity is called *numerical identity*.

In forensic science, the individualisation process is usually considered as a process of rigorous deductive reasoning, as a syllogism constituted of a major premise, a minor premise, and a conclusion. The major premise here in forensic face recognition context is the general principle of uniqueness applied to the source face and trace face. However, it is based on inductive reasoning which cannot be considered as a form of rigorous reasoning, because what is true for one instance is not necessarily true for all. While the demarcation criteria of empirical falsifiability reject the uniqueness of properties used for individualisation from face, this does not imply that face recognition cannot be used in forensic individualisation. It rather just puts a limit on the reliability depending on the quality of the images and method used.

To describe the likelihood ratio approach based on the Bayes theorem, two mutually exclusive hypotheses, the prosecution hypothesis (H_p) and defence hypothesis (H_d), can be defined as the set of all possible hypotheses for the inference of the identity of the source of a trace. Let I represent the background information about the case at hand and E the evidence. The likelihood ratio approach requires computation of E, *between-source variability* (BSV) and *within-source variability* (WSV). Fig.1 and 2 show how to incorporate the likelihood ratio approach into forensic individualisation as described in [14]. A more detailed description of the Bayesian framework and its application to forensic face recognition is presented in section 4.

3.4. Automatic forensic face recognition from digital images

Automatic forensic face recognition from digital images is addressed in [16]. This paper describes small scale experimental work carried out by the Forensic

Science Service in the UK, exploring the performance of an existing automatic face recognition system [3] in the forensic domain. The paper investigates the application of the Bayesian framework for forensic facial comparison and decision making. Experiments are carried out using the Image Metrics OptasiaTM [3] software package for face recognition.

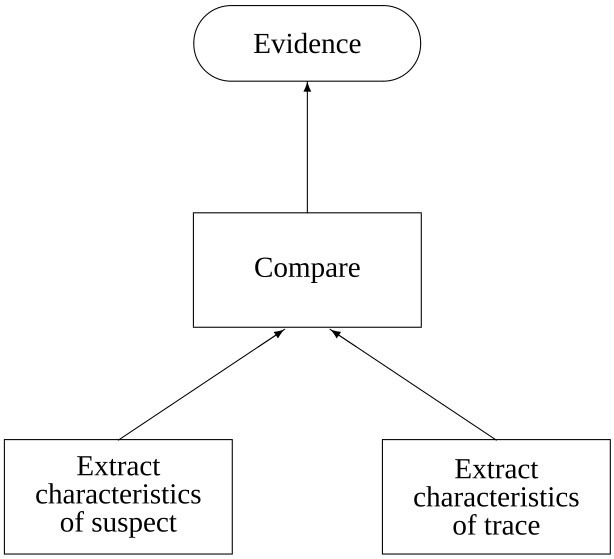

Figure 1. Obtaining evidence in the Bayesian framework.

The approach of the Image Metrics OptasiaTM software used for experiments is straightforward. Active shape and appearance models [22], based on a general dataset of faces, are fitted to a new facial image. The fitted model consists of local information around landmark points in the facial image and forms a face template. To compare two faces, the similarity of the two face templates is determined. In [16] the similarity is expressed in a percentage (0-100%) and is called *recognition probability*. Given a database of n facial images, then a query image results in n recognition probabilities. Query images of persons included in the database are presented to the system and for each query image all n similarity scores are computed. The authors carried out three tests for evaluation of the system.

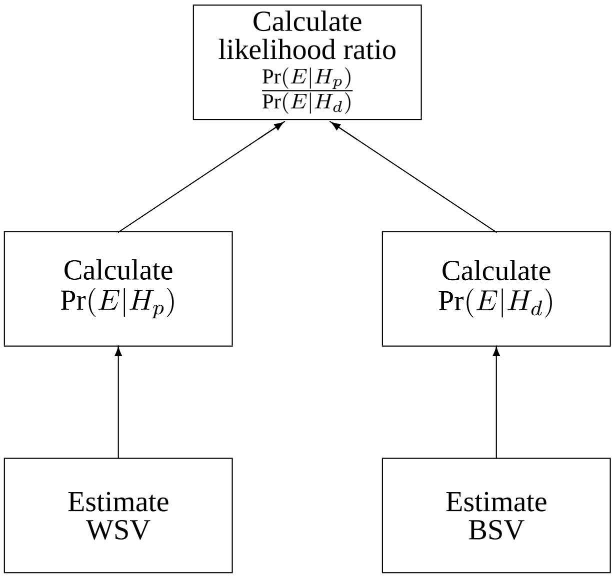

Figure 2. Using the evidence to calculate the likelihood ratio.

In the first test they used the same images as those in the database for benchmarking to get an idea of the the maximum performance of the technique. Twenty pictures chosen at random from the database were used as query images and a similarity score of greater than 95% was obtained for the correct match for each of the query images. The recognition probability sharply drops after the nearest match.

The second test was a feasibility test. For five persons in the database, new images, not present in the database, were obtained and captured exhibiting variation in pose, illumination, age, facial expression, resolution and image quality and used as query images to the system. In this experiment, illumination turned out to have the strongest effect on the recognition probability. The other variations had smaller but significant effects on the recognition probability.

Finally, for evaluation testing, the applicability to the forensic framework was investigated. To be able to calculate likelihoods, the WSV and BSV are needed. Five people of whom images were present in the database were photographed under similar conditions as those used to record images for the database in order to estimate the WSV and the BSV of the database. Of each person 10 images were recorded, resulting in a set Q if 50 images. From this set Q, the WSV for each person was determined from the matching scores resulting from comparing the templates of the person to the template of the same person in the database. The BSV was obtained by matching all images in the set Q to all images in the database. Using the WSV and BSV, the likelihood ratio for a matching score can be calculated. For the set Q in 58% of the cases the comparison to the correct person in the database resulted in the highest likelihood.

The evaluation test provides a small scale, very limited assessment of the expected value or performance of the system in the forensics domain. There is no discussion on how the population size may influence the results.

3.5. Face matching and retrieval using soft biometrics

Although it does not directly focus on the forensic aspects of face recognition, the techniques and methodology proposed in [15] seem very attractive for forensic application of face recognition. Soft biometrics (ethnicity, gender and facial marks), if combined with a traditional face recognition system such as [23, 10] can improve the recognition accuracy as well as the ease of use and interpretation of the outcome in the forensic domain.

In [15] first facial landmarks are detected using an Active Appearance Model (AAM) [22]. Using these landmarks primary facial features are extracted and excluded in the subsequent facial marks detection process. First the face image is mapped to a mean facial shape to simplify the subsequent processing. The Laplacian of Gaussian (LoG) operator is utilised to detect facial marks. Each detected facial mark is classified in a hierarchical fashion as linear vs. not linear and circular vs. irregular. Furthermore, each mark is also classified based on its morphology as dark vs. light. In this way, each of the facial marks can be classified as a mole, freckle, scar etc.

Although the demonstrated performance of the proposed approach, using facial marks detection is not robust, facial marks nevertheless give a more descriptive representation of facial recognition accuracy compared to the numerical values obtained from traditional face recognition systems. This representation may be particularly useful in forensic applications. In such an approach semantic based queries can be issued to retrieve a particular image from a database. Furthermore, the facial marks can be used for facial comparison of partly occluded faces, which are quite common for surveillance cameras, and may even allow differentiation of identical twins. In [15] experimental results are presented, based on the FERET [17] database and a mug-shot database that show that using the soft biometrics in combination with existing face recognition technology can improve the overall performance of the system and is more useful to forensic applications.

4. A Bayesian Framework for Forensic Face Recognition

The aim of a forensic biometric system is to report a meaningful value or expression in court to assess the strength of forensic evidence. The output of a biometric system cannot be used directly in forensic applications as discussed in detail in literature on forensic speaker recognition [9, 12, 6]. Systems using a simple threshold to decide between two classes resulting in a binary decision are not acceptable in the forensic domain [6].

For the purpose of forensic applications, the likelihood ratio framework is agreed upon as a standard way to report evidential value of a biometric system. This framework has been discussed in detail in the speaker recognition domain [9, 12] and the theory presented here benefits from it However, unlike for forensic speaker recognition, there are very few published works which focus on the forensic aspects of face recognition and there is a serious need for reliable facial comparison and recognition systems which can assist law enforcement agencies in investigation and can be used in courts.

The Bayesian framework is a logical approach and can be applied to any biometric system without change in the underlying theory. The likelihood ratio (LR) assessed from a score based biometric system can be used directly in court. While in commercial biometric systems, the objective is to present a score or decisions in a binary form, in forensic applications, the objective is to find the degree of support for one hypothesis against the other. Using the Bayes theorem, given the prior probabilities, the posterior probabilities can be

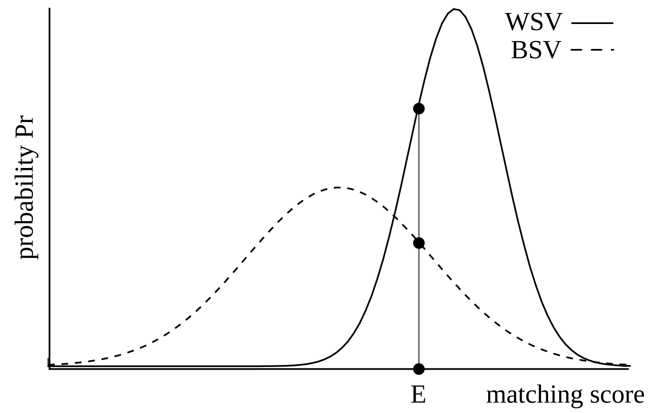

Figure 3. Calculation of the LR from the WSV and the BSV. The solid curve represents the WSV or $\Pr(E|H_p, I)$ and the dashed curve the BSV or $\Pr(E|H_d, I)$. If a trace results in a matching score or evidence E, the LR is obtained by dividing the values of $\Pr(E|H_p, I)$ by $\Pr(E|H_d, I)$. Here the LR would be about 2.

calculated as:

$$\Pr(H_p|E, I) = \Pr(E|H_p, I)\Pr(H_p|I) \quad (1)$$

$$\Pr(H_d|E, I) = \Pr(E|H_d, I)\Pr(H_d|I) \quad (2)$$

where H_p and H_d are the prosecution and defence hypotheses respectively and E represents forensic information (evidence), while I is background information on the case at hand. The prosecution hypothesis H_p states that the suspect is the source of the trace (in this case a facial image) while the defence hypothesis H_d states that someone else in a relevant population is the source. Equations 1 and 2 give the posterior odds required by judicial systems given the prior odds (background knowledge on the case) and likelihood ratio of the evidence E. The likelihood ratio

$$\text{LR} = \frac{\Pr(E|H_p, I))}{\Pr(E|H_d, I))} \quad (3)$$

gives a measure of degree of support for one hypothesis against the other, taking into consideration the circumstances of the case (background information I), and the result of the analysis of the questioned face. It calculates the conditional probability of observing a particular value of evidence with respect to

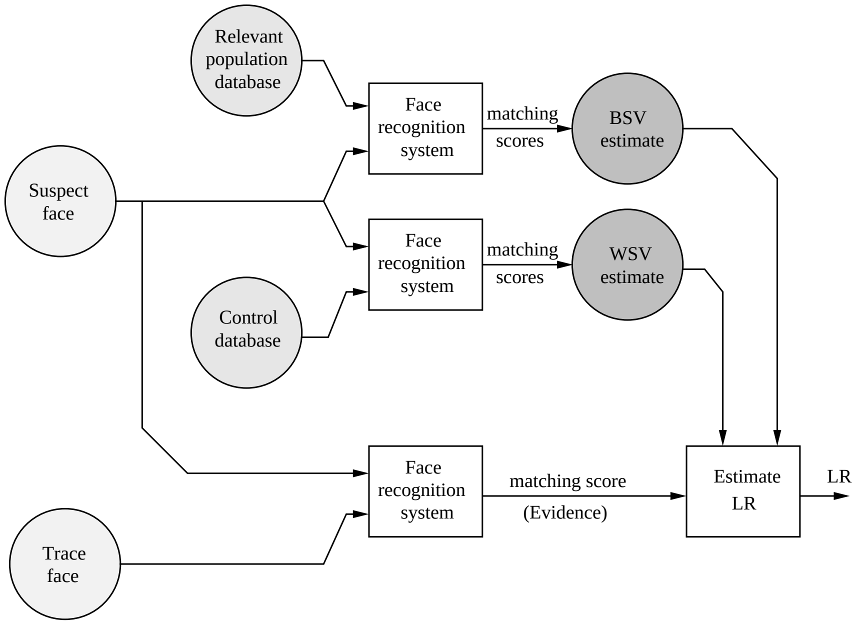

Figure 4. Estimation of the LR. First the WSV and BSV are estimated using a Control database and a Relevant population database with images recorded under the same circumstances as the suspect facial image. Then the LR can be computed by comparing the trace facial image with the suspect facial image and using the WSV and BSV.

two competing hypotheses [11]. The numerator of the LR requires the WSV while the denominator requires the BSV to be calculated. This calculation of the LR from the WSV and BSV for a given matching score or evidence E is illustrated in Figure 3.

The task of a forensic scientist is to evaluate the LR which is then used by the judicial system to reach a conclusion. In order to use a score based biometric face recognition system, in order to calculate the LR we thus need the following:

- The evidence E, a score obtained by comparing a trace face and a suspect face.
- A distribution of matching scores obtained by comparing facial images of the suspect taken under similar conditions to that of suspect facial image (control database): the within-source variability (WSV). The WSV

is then used to estimate the numerator,Pr$(E|H_p, I)$ of the likelihood ratio.

- A distribution of matching scores obtained by comparing facial images of a relevant population taken under similar conditions as that of the suspect facial image (relevant population database): the between-source variability (BSV). The BSV is then used to estimate the denominator, Pr$(E|H_d, I)$ of the likelihood ratio.

Figure 4 illustrates the complete procedure.

5. Reliability and Court Admissibility Issues

The reliability of forensic face recognition is more critical compared to biometric face recognition where an incorrect decision at most results in e.g. a denial of access for a person to a building or a login restriction, the consequences of which are usually not very serious. In the forensic case, however, the consequences are far more severe as an incorrect decision can convict a person as a criminal while he is innocent. While it is agreed upon that future application of automatic facial recognition systems must be assisted by human experts for the final verification [21], the reliability of these systems themselves is also very important as it will reduce manual efforts and help standardise the process of facial comparison. In order to assess the reliability of forensic face recognition systems, several factors such as lighting conditions, facial expressions and pose etc., which are widely explored in the biometric domain should be considered here as well. If the Bayesian framework is used, other factors such as the number of images used to compute the BSV and WSV must also be taken into consideration. Apart from the sizes of the databases, it should also be ensured that the databases sufficiently cover the variations in imaging conditions (such as lighting conditions, image quality etc.) and facial appearance (such as pose, facial expressions etc). The Bayesian framework is the most logical framework, however it is sensitive to the methods used to determine the BSV and WSV. In particular, the distribution of H_p and H_d scores are probability density functions and their estimation is sensitive to the underlying mathematical model chosen. Therefore, a different modelling method can easily lead to different likelihood ratio values.

As a general rule, in order for the evidence extracted from forensic face recognition to be admissible in a court of law, the employed technology must be thoroughly tested and evaluated. In the United States this was ensured by application of the ”Frye rule”. It states that the judges should be acting as ”gatekeeper” to asses if the technology on which the evidence is based is generally accepted in a relevant scientific community or not. Nowadays, in the United States, mostly a revised version of Frye rule called ”Daubert” is in practice. It ensures that, in addition to general acceptance of the technology, the employed technology is tested and can be challenged in some objective way, the technology or theory must be peer-reviewed and a description of the error rate of the technology must be available. Finally, the technology must be maintained and adhere to standards.

In the European judicial system, there is no specific admissibility rule described regarding the scientific evidence. The judges are responsible for the evaluation of the scientific evidence pertaining to the case at hand.

6. Conclusions

At the moment there are no generally accepted standards for facial comparison by human forensic experts. However, several working groups are working on documents for standardisation and training.

Although much research and effort is put into improving state-of-the-art face recognition systems performance, far less effort is devoted to integrating face recognition technology with the legal system of court and justice. Only few papers have been published on the subject of how automatic face recognition can be adopted to forensic purposes. The output of a biometric face recognition system is not suitable for use in forensic applications and the output of a conventional score based biometric system must be processed in order to render it more useful and acceptable by the court.

Although the likelihood ratio value is subjective since it depends on the databases used for the estimation of the WSV and BSV and the underlying mathematical model of their distribution, it still provides the most logical framework for the judicial system to incorporate biometric evidence and background information on the case to reach a conclusion. There is an urgent need for ”tuning” and integration of face recognition systems or development of new systems which can fulfil the requirements of law enforcement agencies and legal systems of court and justice.

References

[1] European network of forensic science institutes, http://www.enfsi.eu/.

[2] Facial identification scientific working group, http://www.fiswg.org/.

[3] Image Metrics OptasiaTM operating instructions, 2002, www.image-metrics.com.

[4] International association for identification, http://www.theiai.org/.

[5] Netherlands forensic institute (nfi), http://www.forensicinstitute.nl/.

[6] C.G.G. Aitken. *Statistics and evaluation of evidence for forensic scientists*. John Wiley & Sons, 1997.

[7] D. Alspaugh. *A brief history of photogrammetry*, pages 1–14. 2004.

[8] M. Y. Aulsebrook, W. A. Iscan, and J. H. Slabbert. Superimposition and reconstruction in forensic facial identification: a survey. *Forensic Science International*, 75:101–102, 1995.

[9] C. Champod and D. Meuwly. The inference of identity in forensic speaker recognition. *Speech Communication*, 31:193–203, 2000.

[10] http://www.cognitec-systems.de Cognitec Systems GmbH. Facevacs software developer kit.

[11] F.i. Botti et al. An interpretation framework for the evaluation of evidence in forensic automatic speaker recognition with limited suspect data. In *Proc. ODYSSEY*, pages 63–68, 2004.

[12] J. Gonzalez-Rodriguez, J. Fierrez-Aguilar, and J. Ortega-Garcia. Forensic identification reporting using automatic speaker recognition systems. In *Proc. ICASSP*, 2003.

[13] Netherlands Forensic Institute. General method of facial comparison. addendum #:1. internal document.

[14] D. Meuwly. Forensic individualization from biometric data. *Science & Justice*, pages 205–213, 2006.

[15] U. Park and A. K. Jain. Face matching and retrieval using soft biometrics. *IEEE Transactions on Information Forensic and Security (TIFS)*, pages 406–415, 2010.

[16] C. Peacock, A. Goode, and A. Brett. Automatic forensic face recognition from digital images. *Science & Justice*, pages 29–34, 2004.

[17] J. Phillips, H. Wechsler, J. S. Huang, and P.J. Rauss. The Feret database and evaluation procedure for face recognition algorithms. *Image and Vision Computing*, pages 295–306, 1998.

[18] P. J. Phillips, W. T. Scruggs, A. J. OToole, P. J. Flynn, K. W. Bowyer, C. L. Schott, and M. Sharpe. FRVT 2006 and ICE 2006 large-scale results. *IEEE Transactions on PAMI*, 2010.

[19] A. Samal and P. A. Iyengar. Automatic recognition and analysis of human faces and facial expressions: A survey. *Pattern Recognition*, page 6577, 1992.

[20] N. A. Spaun. Forensic biometrics from images and video at the federal bureau of investigation. In *Proc. BTAS*, pages 1–3, 2007.

[21] N. A. Spaun. Facial comparisons by subject matter experts: Their role in biometrics and their training. In *Proc. ICB*, pages 161–168, 2009.

[22] Cootes T.F., Edwards G.J., and Taylor C.J. Active appearance models. *IEEE Transactions on PAMI*, pages 681–685, 2001.

[23] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, pages 71–86, 1991.

[24] H. Tuthill and G. George. *Individualization: Principles and Procedures in Criminalistics, 2nd Ed.* Lightning Powder Company, Inc., Jacksonville, FL, 2002.

[25] F. Wang, J. Wang, C. Zhang, and J. T. Kwok. Face recognition using spectral features. *Pattern Recognition*, pages 2786–2797, 2007.

[26] C.M. Wilkinson and R.A.H. Neave. Skull re-assembly and the implications for forensic facial reconstruction. *Science & Justice*, 41(3):233–234, 2001.

[27] L. Wiskott, J.-M. Fellous, N. Kruger, , and C. von der Malsburg. Face recognition by elastic bunch graph matching. *IEEE Transactions on PAMI*, pages 775–779, 1997.

[28] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face recognition: A literature survey. *ACM Computing Surveys*, pages 399–458, 2003.