Video and Imaging, 2013-2016

Ruifrok, A.; Geradts, Z.

Publication date
2016

Document Version
Final published version

Published in
18th INTERPOL International Forensic Science Managers Symposium, Lyon, France

License
Article 25fa Dutch Copyright Act (https://www.openaccess.nl/en/in-the-netherlands/you-share-we-take-care)

Link to publication

Citation for published version (APA):
https://www.interpol.int/content/download/33314/426506/version/1/file/INTERPOL%2018th%20IFSMS%20Review%20Papers.pdf

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
18th INTERPOL International Forensic Science Managers Symposium
Lyon, France

11-13 October 2016
Review Papers

EDITED BY:
DR. MAX M. HOUCK, FRSC
MANAGING DIRECTOR, FORENSIC & INTELLIGENCE SERVICES, LLC
ST. PETERSBURG FL USA
MAX@FORENSICINTELLIGENCE.US

The opinions expressed are those solely of the authors and not necessarily those of their agencies, institutions, governments, or Interpol.
# Table of Contents

**Preface**  
Professor Niamh NicDaeid, Chair of the Organizing Committee  

**Criminalistics**  
- Firearms  
- Forensic Geosciences  
- Gunshot residue  
- Marks  
- Paint and Glass  
- Fibers and textiles

**Forensic Chemistry**  
- Fire investigation and fire debris analysis  
- Explosives  
- Drugs  
- Toxicology

**Media Evidence**  
- Audio  
- Video and Imaging  
- Imaging  
- Digital evidence

**Identification Sciences**  
- Fingermarks and other impressions  
- DNA and biological evidence  
- Questioned documents

**Forensic Science Management**
1. Introduction

In this review, the most important developments are presented for the following general fields of expertise: (1) Biometric analysis of image material, (2) Detection of image manipulation, (3) Camera source identification, (4) Video image processing and Image search.

This year we have been requested to add Video to the review. Since there is clearly overlap, we have integrated this field in the review.

Working groups and organizations

The development of forensic image analysis has several international working groups:

- ENFSI DIWG: The ENFSI Digital Imaging Working Group that is focused on methods, techniques, education and training. http://www.enfsi.org
- LEVA : an American group focused on video processing and training: http://www.leva.org
- AGIB, A working group in Germany that is focused on facial image comparison: http://www.foto-identifikation.de/ .
- FISWG, An American group since 2009 that is focused on facial image comparison: http://www.fiswg.org
- OSAC Facial Identification Subcommittee, An American group part of the Organisation of Scientific Advice Committees, with focus on standards and guidelines related to the image-based comparisons of human facial features: http://www.nist.gov/forensics/osac/sub-face.cfm

Since 2003 each year a workshop was organized on Forensic Image and Video processing with handouts on the methods for face comparison, video restoration, 3D reconstruction, length measurement, photogrammetry and image processing. Also each year a scientific session was organized on this field. More information is available on: http://www.aafs.org

2. ENFSI Forensic IT Working Group
The forensic IT working group of ENFSI [2], [3] deals with digital evidence as such. There exist some overlap with the Digital Imaging working group, and for that reason joint events are organized.

Since most CCTV-systems are digital nowadays, often the question of handling the CCTV system itself is a question of digital evidence. Hard drives and other digital media should be handled in a secure way with proper forensic imaging software. The working group organizes training conferences each year. More information is available from http://www.enfsi.eu/.

2.1 Biometric analysis of image material

Biometrics is regularly announced in news items as a panacea against terrorism, security problems, fraud, illegal migration, etcetera. Biometrics, which can be defined as the (automatic) identification or recognition of people based on physiological or behavioral characteristics, is not a single method or technique, but consists of a number of techniques, with each their own advantages and drawbacks. None of the available biometric modalities combines the properties of an ideal biometrics system. We have to acknowledge that biometrics never can be 100% accurate. However, if requirements and applications are carefully considered, biometric systems can provide an important contribution to investigation, authentication and safety.

Within the context of person identification (individualization), different processes can be defined. Within different areas of science, different terminologies are used for the same process, and sometimes the same terminologies are used for different processes. Therefore, a clear definition of the different terms as used in this text is important and made explicit here.

Recognition can be defined as the process of identifying or matching a person, his/her photograph or image with a mental image that one has previously stored in long term memory. Recognition requires observation and retention of a person’s features and the process of comparison of the retained information with an external image whether it be the life person, a photograph or composite image. The word recognition is important for investigation as well as witness statements. Recognition is within the forensic community also used for the automated searching of a facial image in a biometric database (one-to-many), typically resulting in a group of facial images ranked by computer-evaluated similarity.

Identification is the most contentious term because this most often used term can mean several things in different context, like the automated searching of a facial image in a biometric database (one-to-many) in biometrics, the examination of two facial images or a live subject and a facial image (one-to-one) for the purpose of determining if they represent the same person in forensics, or the assignment of class or family name in biology and chemistry. Therefore, the authors of this paper prefer not to use the term identification unless the meaning is unambiguous within the context.

Recall is here defined as the process of retrieving descriptive information of a person from long term memory in the absence of the person, his/her photograph or other image. Recall requires observation, retention and reproduction of a person’s features. Recall is essential for the production of composite images, as produced by a police artist for investigational purposes. However, these images can only be used as investigative tools, and can never be used as proof of identity.

2.2 Face
On top of the list of preferred, and in most travel documents required, biometric modalities is the face. The face has always been the most important personal feature on travel documents. The most important change the last decade is that the face is now also stored digitally in passport, and is optimized for automatic facial recognition. The automated systems are still very sensitive to ageing of the person depicted [Agrawal 2015, ]; The latest test results indicate that higher resolution and well controlled images may result in a 10-fold better performance than using average passport like images.

Facial image comparison is defined as the visual examination, by a human operator, of the differences and similarities between two facial images or a live subject and a facial image (one-to-one) for the purpose of determining if they represent the same person. In biometrics the one-to-one comparison is termed verification. The Facial Imaging Scientific Working (FISWG) group also uses the term Facial Identification for the same process. However, the authors of this review prefer to use the term facial image comparison, because that exactly describes the process, and cannot be confused with the use of the word identification as used in other contexts.

Facial Reconstruction is used in two different meanings:
1. The process of reconstructing three-dimensional facial (computer) models of individuals from their 2D photographic images or video sequences.
2. The process of recreating the face of an individual (whose identity is often not known) from their skeletal remains through an amalgamation of artistry, forensic science, anthropology, osteology, and anatomy.

These two different uses of facial reconstruction may meet when three-dimensional computer models are used to recreate the face of an individual based on skeletal remains. The current review does not consider facial reconstruction as a means to recreate the face of an individual based on skeletal remains.

Facial composite is a graphical representation of an eyewitness’s memory of a face, as recorded by a composite artist, also sometimes termed facial sketch. A facial composite may be based on combination of images of facial features (photo composition) as well as actual drawings by a composite artist.

2.3 Facial image recognition
Biometric systems that can search databases with facial images, using automatically extracted facial features, are still being developed further. Face recognition error rates have declined massively in the two decades since initial commercialization of the various technologies [4]

However, one of the complicating issues is that the images of the unknown person often differ from the target images in the database with respect to the orientation of the head, the distance to the camera, the illumination and the image resolution. Recently, increased attention has been given to the existence of quality measures for face recognition [5][6]. New approaches focus on better acquisition techniques in order to get better images, from which as many facial features as possible can be extracted for comparison to images in the database.

2.4 Unconstrained images
Forensic material in general is not acquired under controlled circumstances. This means that a number of confounding factors may influence the effectiveness of face recognition. A number of academic studies try to tackle the different confounding factors, like pose issues,
resolution issues, and occlusions. How different quality features may influence the performance of face recognition systems has been studied by Dutta et al [6].

Several techniques are being developed to use unconditioned images and media collections for unconstrained face recognition [7] and partially occluded faces [8]. Though issues surrounding pose, occlusion, and resolution continue to confound matchers, there have been significant advances made in face recognition technology to assist law enforcement agencies in their investigations [9].

Park et al [10] report improvements in facial image retrieval of partially occluded images (sunglasses and scarfs), resulting in more than 90% retrieval of partially occluded faces.


2.5 Pose variation
Pose is the “orientation of the face with respect to the camera, consisting of pitch, roll, and yaw”. An optimal frontal pose may be considered as 0° in all directions. Variations to the optimal pose can be due to photographing a physical subject who can move freely during the capture process, or misalignment of the camera. As images are a 2-dimensional representation of the 3-dimensional world, pose of a subject has a major influence on the image captured by a capturing device. As a result of this the appearance and position of facial features can change depending of the pose of the person and the position of the camera at the moment of capture. This is, together with inter and intra observer variability of landmark annotation, one of the main causes of the limited value of landmark measurements on photographs [12]. However, development of pose detection and automatic landmark detection has been reported to result in almost 90% identification accuracy in side view positions [13].

For predicting face recognition performance in a video, it was observed that face detection confidence and face size serve as potentially useful quality measure metrics [14].

2.6 3-dimensional face comparison
The most promising approach to the complicating issues of pose and illumination is the use of 3 dimensional models for pose an illumination correction. Since the previous review, there has been an increase in reports on development of methods that are based on the use of 3-dimensional computer models of faces. A number of 3d-acquisition systems are now available for the acquisition of these models. Most 3d-cameras work with a configuration of 1 or more normal digital photo cameras, a flash and the projection of a pattern on the face. These models can be used in two ways. A 3d-facial model of a suspect can be compared to a 3d-model of an unknown person, or the 3d-model of a suspect is used to compute an image that can be compared to an image of an unknown person. Since there are many sources of images and video in practice, a number of studies are focused on the (partial) reconstruction of 3d-models from 1 or more images or video streams. Van Dam et al [15] developed a model 3-D face reconstruction algorithm based on 2D landmarks. He 3D landmark reconstruction algorithm simultaneously estimates the shape, pose and position of the face, based only on the fact that all images in the sequence are recorded using a single calibrated camera.

2.7 Deep learning
With the further development of computer technology, neural network approaches for facial recognition have gained renewed interest. Alignment and the representation of the face by
employing explicit 3D face modelling have resulted in improved accuracy of face recognition in unconstrained environments [16][17] [18] [19]

2.8 Facial image comparison

The result of facial image recognition is often the selection of 1 or more target facial images that could be matched with the image of the unknown person. In practice, however, this often leads to hit lists with multiple possible matches to the query image, and the correct target not necessarily on top of the hit list. In such cases, the decision has to be made by a forensic anthropologists or forensic image analysts. Since the previous review, more studies and proficiency tests have been reported on the performance of facial image comparison by lay people and experts, showing that there is a reason for concern, and that better methods and technology are needed. A number of institutes have published documents that describe their procedures for performing facial image comparison. These procedures show that measures are being taken to limit the influence of subjective judgments and that there is a need for quantitative statistical data. The FBI has started a working group in 2009 for facial image comparison that is expected to stimulate the development of better methods and technology (FISWG).

Human and computer performance has been systematically compared as part of face recognition competitions, with results being reported for both still and video imagery. Analysis of cross-modal performance shows that for matching frontal faces in still images, algorithms are consistently superior to humans. For video and difficult still face pairs, humans are superior [20]

People doing facial image comparison can be found in four different kinds of professions: forensic photographers, forensic anthropologists, video investigators and imaging scientists. Knowledge of anatomy and physiology of the face is needed to get a good interpretation of differences and similarities in facial features. Similarities or differences in such images can often be explained by differences in the imaging conditions, pointing to the importance of knowledge about optics. Small facial details can be distorted, and artifacts looking like small details introduced due to noise, pixel sampling and compression, requiring knowledge about image processing for the proper interpretation of observations. Changes in image quality, pose and position, lighting and facial expression greatly influence the comparison process. Therefore, it is strongly recommended that one acquire reference images of the suspect and a number of other people with the same video camera in the same situation under similar lighting conditions. While the techniques of facial image comparison are generally accepted within their practitioner communities, they are not tested, and their error rates are unknown. On that basis, the methods of facial image comparison would appear not to meet the anticipated standards [21]

It is well-established that matching images of unfamiliar faces is rather error prone. Experimental studies on face matching underestimate its difficulty in real-world situations. Photographs of unfamiliar faces seem to be unreliable proofs of identity, especially if the ID documents do not use very recent images of the holders [22]

Existing scientific knowledge of face matching accuracy is based almost exclusively, on people without formal training. Human performance curtails accuracy of face recognition systems, potentially reducing benchmark estimates by 50% in operational settings. Mere practice does not attenuate these limits [23], and some training methods may be inadequate [24]. However, large individual differences have been reported, suggesting that improvements in performance could be made by emphasizing personnel selection [25]. White et al [26] also have shown that forensic facial examiners outperformed untrained participants and computer algorithms on challenging face matching tests, thereby providing
the first evidence that these examiners are experts at this task. Notably, computationally fusing responses of multiple experts produced near perfect performance.

### 2.9 Eyewitness identification / Facial composites

In most of the criminal investigations of a crime, one of the first steps is to interview eyewitnesses. In these interviews the witnesses are asked to provide a description of the perpetrators. For investigational purposes this description may be made into an image by a (police) sketch artist. The sketch artist can also help the witness to recall the face of the perpetrator by showing multiples examples of facial features. Instead of sketches, it is also possible to create photo compositions using examples from databases with facial images. As not always images of perpetrators are available, matching of composite sketches with facial photographs (e.g. mugshots) is of interest. Matching performance of composite or forensics sketches against photo galleries are promising but still considerably lower than photo matching performance of commercially available systems [27][28]

### 2.10 Other biometrics

#### 2.10.1 Ear

Even though current ear detection and recognition systems have reached a certain level of maturity, their success is limited to controlled indoor conditions. In addition to variation in illumination, other open research problems include occlusion due to hair, ear symmetry, earprint forensics, ear classification, and ear individuality [29]. The experimental results show that ear recognition may achieve an average rank-one recognition accuracy of more than 95% [30] Current studies are directed towards more robust automated methods for ear detection, landmark localization and ear recognition using 2D and 3D techniques [31], [32] [33].

#### 2.10.2 Body geometry and gait

With the standardisation of photographs, identification primarily occurs from the face. However, results consistently show that less body measurements are needed to find no duplicates when compared to the face. With the combination of eight body measurements, it is possible to achieve results comparable with fingerprint analysis [34]. Thicker garments produce higher inaccuracies in landmark localisation, but errors decrease as placement is repeated. Overall, comparison to truth reveals that on average statures can be predicted with accuracy in excess of 95% [35].

Also lower leg shape, sometimes the only body part consistently depicted in images, has been reported as “an effective biometric trait” [36]. Recent studies have shown that when face identification fails, people rely on the body but are unaware of doing so [37]. Bouchrika et al [38] reported a method to extract gait features for different camera viewpoints achieving an identity recognition rate of 73.6 % processed for 2270 video sequences. Furthermore, experimental results confirmed the potential of the proposed method for identity tracking in real surveillance systems to recognize walking individuals across different views with an average recognition rate of 92.5 % for cross-camera matching for two different non-overlapping views.


#### 2.10.3 Soft biometrics

Soft biometric information extracted from a human body (e.g., height, gender, skin color, hair color, and so on) is ancillary information easily distinguished at a distance but it is not fully distinctive by itself in recognition tasks. However, this soft information can be explicitly fused
with biometric recognition systems to improve the overall recognition when confronting high
variability conditions. The use of soft biometric traits is able to improve the performance of
face recognition based on sparse representation on real and ideal scenarios by adaptive
fusion rules [24]. Depending of the acquisition distance, the discriminative power of the facial
regions can change. This results in some cases in better performance than achieved for the
full face [40]

Soft biometrics introduce a possibility to automatically search databases based on biometric
features obtained from verbal descriptions, resulting in more than 95% identification
accuracy [41].

2.10.4 Liveness detection
Spoofing is the act of masquerading as a valid user by falsifying data to gain an illegitimate
access. Vulnerability of recognition systems to spoofing attacks (presentation attacks) is still
an open security issue in biometrics domain and among all biometric traits. Galbally [42]
propose a technique using 25 general image quality features extracted from one image (i.e.,
the same acquired for authentication purposes) to distinguish between legitimate and
impostor samples. The experimental results, obtained on publicly available data sets of
fingerprint, iris, and 2D face, show that the proposed method is highly competitive compared
with other state-of-the-art approaches and that the analysis of the general image quality of
real biometric samples reveals highly valuable information that may be very efficiently used
to discriminate them from fake traits. Erdogmus et al [43] studied detection problem of more
complex 3D attack types using various texture based countermeasures.

3. Detection of image and video manipulation

Image and video files are changed for numerous reasons with and without a criminal intent. Images are scaled, cropped, rotated and compressed to make them fit for a document or a
website. Contrast or colors are changed to enhance the visibility of details. This processing
is often referred to as manipulation. However, manipulation could also refer to modification
of an image with a criminal intent. One type of modification is a change of the visual content
by hiding or inserting visual information in the original image. The other modification is non
visual addition of information, like a text message in an image that is published on a website
as a means of communication between persons. This modification is referred to as
steganography.

A number of clues can be used for detection of manipulation by visual inspection, like
discrepancies in lighting, brightness levels, color distributions, edges, noise patterns and
compression artifacts in the transitions between the tampered and original parts of the
questioned image. Visual inspection could be not feasible anymore when it is not known
which region in an image is being questioned. A lot of research was focused on automated
detection of regions in an image that might have been tampered with. However, most of the
methods that have been published do only produce indications of regions in an image that
require inspection by an examiner. The SWGIT group has published a document on

A special type of detection is based on the indication of ‘resampling’ [44]–[47]. When a part
of an image is pasted into another image, it is often necessary to apply rotation and resizing
to make them visually fit. This resizing causes a special relationship between color values in
the resized region that could be detected.

Double compression detection in JPEGs [48]–[53] is also widely researched, as well as
using the Photo Response Non Uniformity (PRNU) for detection [54]

574
Another type of image tampering is referred to as ‘copy-paste’ forgery ([46]–[54]). Objects or persons that are visible against a background with a specific texture, like blue air, green grass, trees, etc. are hidden by pasting a copy of a region in the image with the same texture over them. Detection of this type of tampering looks like a simple straightforward process. All regions in an image have to be compared to each other in order to find regions that are copies. However, the challenge is to limit the number of comparisons and to find a computationally efficient method. This is a requirement when a large amount of images has to be checked. A relatively large number of methods have been proposed in the literature for this task. Also a method with support vector machines (SVM) for object based manipulation by the contour in a video file is proposed, with an accuracy from 70 to 95 % claimed by the authors [55].

A related problem is the detection of illegal copies of image and video files. One technique for protection of original image and video files is the use of watermarking. A watermark is in most cases a hidden mark in the image that will get lost in most common copy processes. Although watermarking is already an old technique there is a lot of research going on ([56]–[74]). Furthermore on authenticity a survey is published for video[75].

The number of papers published on these topics show that the problem of reliably detecting image tampering has not been solved yet. There are now a number of software packages available that offer a number of methods that can be tried on a questioned image, however interpretation of the results by an expert remains important.

Also anti forensics methods are described based on methods that are published, to prevent tampering from being detected. [76]–[83]

4. Camera identification of images and video

In criminal investigations of child porn production and distribution, identification of the source of a digital image has become very important, because a specific camera, (or a cell phone camera, a webcam, or a flatbed scanner) could be linked to a suspect using other types of evidence. Identification of images that might have a common source can also be helpful in these investigations. The developments that have been started in the period of the previous review have not been stopped and have lead to a number of new methods and software packages. The most used method is based on the estimation of a specific type of fixed pattern noise in an image that is caused by PRNU - Photo Response Non Uniformity[84]. The method is also useful in other cases such as murder and fraud to find a links between a camera and images that have been taken [85], [86].

For identification of a specific camera as the source of a specific image, the PRNU patterns have to be estimated from reference images from the camera and the noise that can be filtered out from this specific image. These patterns have to be compared and a similarity measure is used as a measure for the strength of the evidence that the camera is the source. Common practice is to compare the PRNU pattern of a specific image with the PRNU patterns from a large number of camera’s. The quality of the estimation of the PRNU pattern from an image depends heavily on the image content and this can be taken into account. However, if there are more images available from the same, unknown source, e.g. the frames in a video file, much better estimations of the PRNU pattern can be obtained by averaging techniques. However several methods are presented to improve the calculation speed as well as clustering images if the camera is not available. Also the use of GPUs is discussed within these methods ([86]–[89]).
Other sources of fixed pattern noise [90]–[92] that have been investigated are based on detection of image artifacts from differences in image processing in the camera chips. If the camera does not work also switching the camera module appears to work for finding a link between the camera module and the images taken[93].

In the forensic practice of a case in which a specific camera has to be identified, a collection of similar cameras from the same brand and type are needed for validation of the results. For using PRNU as evidence, the analyst has to interpret the comparison results. The ENFSI working group for Digital Imaging has conducted two proficiency tests to find out what different experts might report to the court about camera identification. In the practice of investigation of large amounts of images, PRNU is also useful to get indications of possibly common sources. A number of studies have been found on the implementation of this application.

5. Video Image processing and Image Search

The ENFSI worked on a proficiency test of image processing, S-Five https://s-five.eu/ . The project was for Standardization of Forensic Image and Video Enhancement, and resulted in a best practice manual for processing these. Also a test set with a collaborative exercise and the use of different software products and algorithms have shown that different results are obtained in image enhancement depending on the blur function ([94]–[96]).

Several approaches have been published [97], where also attention is given on logging and archiving processes.

A methodology to event reconstruction has been published by Quentin [98]. They proposed a systematic approach to using trace images in the framework of an investigation. The method is cyclic and iterative. In other publications [99], [100], Quentin provided solutions for image classification, by adding systematic description of image content to metadata as well as timeline chronological reconstruction. For 3D reconstruction using image from fixed and mobile cameras, photogrammetry is proposed.

Several software products have been developed for searching in large amounts of images and video and using similarity detection of images. The software appears to be useful in practice for searching large amounts of image databases, and big multimedia analysis [101]–[103]. The development of deep learning methods.

The development and implementation of deep learning methods

The LASIE project of FP7 funded by the European Commission has published an overview of methods for image search in relation to big data http://www.lasie-project.eu/wp-content/uploads/2015/05/LASIE_D6.1.pdf

Forensic video analysis is also assisted by algorithms, which is published for example by Geronimo [104] on retrieving certain actions, such as walking or fighting and the search for specific persons. Also relevance feedback is important for semi supervised search [105]

The recent deployment of very large scale camera networks with fixed and moving surveillance cameras has led to a novel field of object tracking. Hsu [106] proposes vehicle matching and tracking cross camera by affine and scale invariant feature transform.

5.1 Video games forensics
Another field which is approaching is video games forensics[107]. An investigator might not expect video games and their data files are used as a crime or hide data.

5.2 Video carving
Within the field of video carving several approaches have been made. The software defraser which is partly open source as published by the Netherlands Forensic Institute at https://sourceforge.net/projects/defraser/ . Since video file formats define only a limited number of mandatory features, they leave room for interpretation. Differences in file structures appear to be available [108].

6. References


[107] M. Ebrahimi, Lei Chen, “Emerging cyberworld attack vectors: Modification, customization, secretive communications, and digital forensics in PC video games,” in *2014*