Chapter 5
Additional Biometric Traits

Every biometric system relies on one or more biometric modalities. The choice of modality is a key driver of how the system is architected, how it is presented to the user, and how match vs. nonmatch decisions are made. Understanding particular modalities and how best to use the modalities is critical to overall system effectiveness.

Whither Biometrics Committee, National Research Council, 2010

Earlier chapters in this book focused exclusively on three specific biometric modalities - fingerprint, face, and iris. These traits have been extensively studied in the literature and have been incorporated in several government, military, and civilian biometric systems around the world. However, apart from these traits, several other biometric attributes have also been studied in the context of applications ranging from border control systems to surveillance to forensic analysis. Examples of such attributes include hand geometry, ear, speech, signature, gait, DNA, and teeth. Further, soft biometric attributes (i.e., attributes that provide some information about the individual, but lack the distinctiveness and permanence to sufficiently differentiate any two individuals) such as scars, marks, and tattoos (SMT), periocular region, and human metrology have also been studied in the biometric literature. This chapter will introduce a few of these traits in order to convey the breadth of work being conducted in the field of biometrics.

5.1 Introduction

As stated in Chapter 1, a wide variety of biometric traits have been proposed and studied in the literature. In some cases, academic curiosity about the uniqueness and permanence of certain biological traits has spurred exploratory research (e.g., iris); in other cases, new application domains have resulted in the exploration of novel biometric traits (e.g., periocular biometrics). Furthermore, certain biometric traits are uniquely suited for some applications and scenarios. For example, voice may be more practical in tele-commerce applications; the ear may be useful in surveillance applications where only the side-profile of the human face is available; gait patterns may be relevant in identification-at-a-distance scenarios; hand geometry may be appropriate for use in systems requiring the verification (as opposed to identification)

A.K. Jain et al., *Introduction to Biometrics*, DOI 10.1007/978-0-387-77326-1_5,

© Springer Science+Business Media, LLC 2011

of a few enrolled identities thereby mitigating some of the concerns associated with using a strong biometric cue such as fingerprint; and the iris or fingerprint may be chosen in applications where the subject is cooperative and in close proximity to the sensor.

Apart from the aforementioned traits, ancillary information such as gender, ethnicity, age, height, and eye color can also be used to improve the matching accuracy of a biometric system. For example, if a female subject (probe) is matched incorrectly against a male subject (in the gallery), then the gender information can be used by the biometric system to reject the match. Ancillary attributes provide additional information about the individual, but lack the distinctiveness and permanence to sufficiently differentiate two individuals. However, they can be used to narrow the search space of potential matches in an identification system (e.g., if the input probe is deemed to be an "Asian Male", then an identification system can constrain its search only to "Asian Male" identities in the database) or when other biometrics traits are not readily available (e.g., using the periocular information when the iris is deemed to be of poor quality). Such traits are commonly referred to as *soft biometrics* in the literature. Unlike some other attributes such as fingerprint and iris, soft biometric traits are not necessarily "unique" to an individual. They can be shared across a significant fraction of the population (e.g., gender) and may lack permanence (e.g., scars, marks, and tattoos, abbreviated as SMT).

Given the diversity of biometric traits discussed in the literature, in the interest of being concise, we restrict our discussion to the following four biometric traits in this chapter: ear, gait, hand geometry, and soft biometrics.

5.2 Ear

The appearance, structure, and morphology of the human ear has been studied as a biometric cue for a number of years. While most face recognition systems extract the attributes of the human face from frontal images, the visibility of the ear in non-frontal poses of the face (e.g., side view) makes it a viable biometric in many scenarios. The human ear is observed to exhibit variations across individuals as assessed by the curves, surfaces, and geometric measurements pertaining to the visible portion of the ear, commonly referred to as the pinna. The structure of the pinna depicting various anatomical features can be seen in Figure 5.1.

As a biometric trait, the ear offers several advantages: (a) the structure of the ear has been observed to be stable despite aging, and ear growth is almost linear after the age of four; (b) the ear, unlike other facial features, is minimally impacted by changes in facial expression; and (c) image acquisition does not involve explicit contact with the sensor.

A typical ear recognition system consists of the following components: (a) an ear detection module (also known as segmentation) that localizes the position and spatial extent of the ear in an image; (b) a feature extraction module that extracts discriminative features from the ear; (c) a matching module that compares the fea-

(1) Helix Rim
(2) Lobule
(3) Antihelix
(4) Concha
(5) Tragus
(6) Antitragus
(7) Crus of Helix
(8) Triangular Fossa
(9) Incisure Intertragica

Fig. 5.1 External anatomy of the ear. The visible flap is often referred to as the *pinna*. The intricate structure of the pinna coupled with its morphology is believed to be unique to an individual, although large-scale evaluation of automated ear recognition systems has not been conducted.

tures extracted from two ear images and generates a match score; and (d) a decision module that processes the match score(s) and establishes the identity of the subject.

5.2.1 Ear detection

A number of techniques have been proposed in order to locate the ear in a given image. These approaches can be categorized into the following groups.

1. Template Matching:
   In a template matching scheme, a template of a typical ear is constructed and is matched with each location in the query image. The location giving the highest score is considered as the region containing the ear. The template may consist of an edge image of the ear or a set of descriptors extracted from the ear such as the response to a set of filters or a histogram of shape curvatures in case a 3D image of the ear is being used for recognition. Detection based on response to a set of pre-selected filters, known as the Viola and Jones technique, is also commonly used for detecting faces in an image.
2. Model-based Detection:
   A model-based detection technique assumes certain characteristics of the shape of the ear and tries to find regions that manifest such characteristics. The shape of the helix, for example, is usually elliptical so a generalized Hough transform tuned for detecting ellipses can be used to locate the ear in an edge image. Features extracted using chain codes[1] can also be used to classify each curve obtained from an image to be a curve associated with the ear, such as a helix or an anti-helix.
3. Morphological-operator-based Detection:
   Since the structure of the ear is usually more intricate than the structure of the remaining region in a profile face image, morphological transformations such

[1] A chain code typically measures the local orientations along the length of the curve.

as the Top-hat transformation can be used. A Top-hat transformation essentially subtracts a morphologically smoothened version of an image from itself, thereby highlighting finer details.

4. Face-geometry-based Detection:
   Since in a profile image the nose can be easily detected as the point with high curvature, it is possible to constrain the search for the ear in an appropriate location relative to the nose.

Ear detection performance can be improved by utilizing a two-stage processing, where the skin region is segmented from the profile image in the first stage and the ear is detected in this reduced search space during the second stage.

5.2.2 Ear recognition

1. Subspace analysis-based techniques:
   Similar to recognizing face images, projecting the ear image onto a set of principal directions is an effective way to obtain a salient and compact representation of an ear. Subspace projection techniques such as PCA, ICA, and LDA have been successfully used in literature for matching ear images. Furthermore, manifold learning based techniques such as Locally Linear Embedding (LLE) and Kernel PCA have also been used to perform ear recognition.
2. Sparse representation-based techniques:
   Optimization techniques that minimize L-1 norm of the distance vector between the transformed query and all the transformed templates in a database have been shown to provide high recognition accuracy in object recognition studies. This technique has also been successfully used for ear recognition.
3. Point set matching-based techniques:
   Elastic bunch graph matching is an effective technique to recognize faces based on responses to a bank of Gabor filters at some fiducial points on the face. This technique has also been successfully used for matching ear images where a number of landmark points can be easily detected, thanks to its complex structure.
   Scale Invariant Feature Transform (SIFT) is a well known technique for matching two images where a set of salient points can be reliably and repeatably extracted from them. In order to match two images using SIFT features, corner points are detected from two images and matched based on image gradient-based features extracted from the neighborhood region of each point. See Figure 5.2 for an example of SIFT points being matched between two ear images.
4. Image filtering-based techniques:
   In certain techniques, the ear image is first enhanced to highlight the discriminative features and suppress the noise. Two common techniques that use this basic procedure are force field transformation and local binary patterns.

   a. Force field transformation:
      A force field transformation essentially obtains the intensity of the forces at

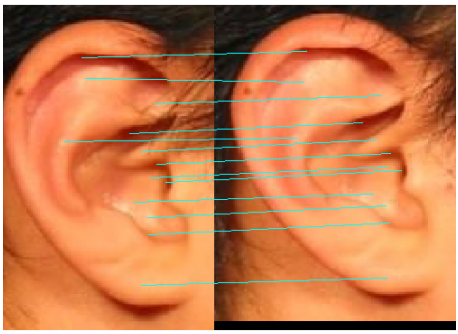


Fig. 5.2 Comparing two ear images by using the SIFT key points matching scheme. Here, the SIFT keypoints are first extracted from each image prior to comparing them.

each location in the image where each pixel is considered as a source of the force with intensity proportional to its value. Force field transformation has been shown to effectively remove the noise in the ear image leading to significant improvement in recognition accuracy. Further, a set of lines indicating the gradients of this force field can be extracted and used for matching. See Figure 5.3 for a depiction of force field extracted from an ear image with the force field lines marked.

b. Local binary patterns:
Local binary patterns essentially characterize each pixel based on variation of intensity of that pixel along a set of directions. The variation along each direction is encoded as a single bit indicating whether the intensity is increasing or decreasing, and the set of bits associated with each direction is used to obtain an integer value for each pixel. Such a transformation effectively reduces the effect of illumination variations and other sources of noise, thereby generating an enhanced image which can be used for robust matching.

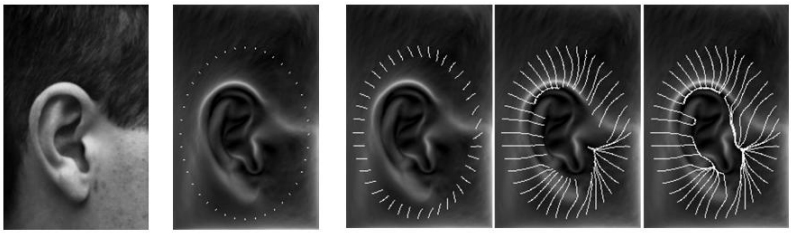


Fig. 5.3 Extraction of force field lines from an ear image using an iterative approach.

5. Geometric measurements-based techniques:
Features obtained by measuring certain geometric characteristics of the ear can also be used as a set of discriminative features. As an example, the centroid of an

ear image obtained from its edge image can be used as a center to draw concentric circles with pre-specified radii. Various measurements, such as number of points on a circle intersecting the edge image or distance between two consecutive intersections, can be used as a feature vector. Characteristics of different curves present in the edge image such as the coordinates of endings and bifurcations can also be used as additional features.

6. Transformation-based techniques:
   Various image transformation techniques such as Fourier transform or wavelet transform can also be applied to extract discriminative features from an ear image. Fourier transform can also be applied in order to obtain a rotation and translation invariant representation of the ear - for example, by using a polar coordinate system and extracting only the magnitude of the Fourier transform.
7. 3D techniques:
   In some scenarios, acquiring a 3D rendition of the ear entity may be possible. 3D images offer depth information that can be used in conjunction with the 2D texture information to improve the recognition accuracy. In the case of 3D ear images, local histograms of shape curvature values can be used to match two ear images. The Iterative Closest Point (ICP) algorithm is commonly used to register and match 3D ears. In the ICP technique, each point in the input ear image is used to obtain a corresponding point in the template image and the input is then rotated and translated in order to minimize the distances between corresponding points. This procedure is iteratively applied until convergence, and the resultant set of distances is used to compute the matching score.

5.2.3 Challenges in ear recognition

Although several algorithms for ear detection and matching have been proposed in the literature, large-scale public evaluation of ear recognition algorithms has not been conducted. Further, there are no commercial biometric systems at this time that explicitly utilize features of the ear for human recognition. But the performances of ear recognition algorithms have been tested on some standard ear datasets. Experiments suggest that ear images obtained under controlled conditions can result in good recognition accuracy. However, the performance of ear recognition methods on non-ideal images obtained under varying illumination and occlusion conditions is yet to be established. Several challenges have to be overcome to make this possible.

1. Ear Occlusion:
   One of the main challenges faced by an ear recognition system is occlusion due to the subject's hair. One way to address such occlusion is by capturing the thermogram along with the visible light image. In a thermogram, the hair can be easily detected (and possibly isolated) as its temperature is usually lower than that of the skin.

2. Earprint Identification:
   Earprints, or earmarks, are marks left by secretions as a result of pressing the ear against a flat surface. These marks mainly consist of impressions of the helix, anti-helix, tragus, and anti-tragus. Other features include the earlobe and the crus of helix, but they are less frequently observed. Earprints can be compared based on details such as the notches and angles in imprinted samples, the positions of moles, folds, and wrinkles, and the position of pressure points. Figure 5.4 provides a set of example earprints lifted from crime scenes.
   Earprints are known to be available in approximately 15% of the crime cases and have also been considered in some court cases as a source of forensic evidence. However, due to significant variations among the multiple impressions of an ear, individualization based on earprint is frequently disputed. The main reasons confounding the individualization of earprint include (a) variable deformations caused by the force applied by the ear to the surface, (b) the duration of the ear's contact with the surface (c) ornamental modifications to the ear, such as piercing, and (d) changes in the shape and size of the ear due to aging.
   Earprint identification is usually manually performed by identifying and matching a set of geometric features from the earprint such as the intersection points of ear curves with a regular grid or locations of certain other landmark points. Fully automated systems utilizing SIFT features have also been designed to match two earprints, but their performances has not been extensively evaluated in operational environments.

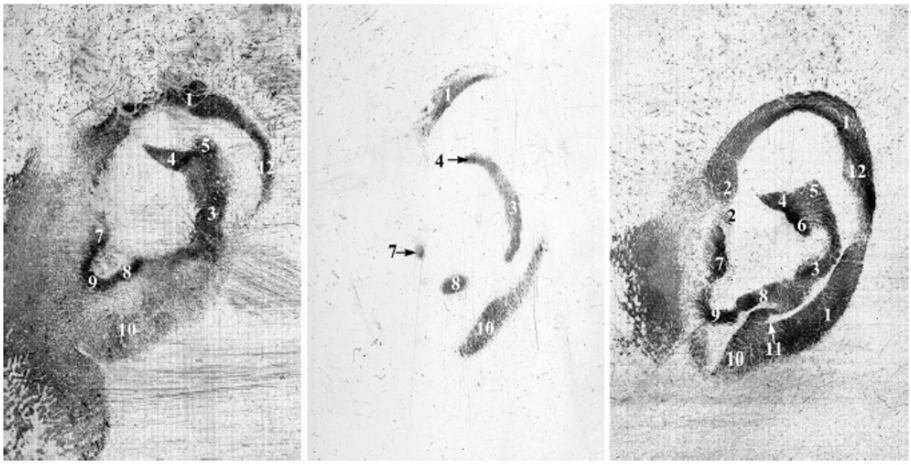


Fig. 5.4 Examples of earprints in which various anatomical features are indicated. This image is reproduced from [21]. The labels in this image are as follows: 1. helix; 2. crus of helix; 3-6. parts of anti-helix; 7. tragus; 8. antitragus; 9. incisure intertragic; 10. lobe.

5.3 Gait

The demand for human identification at a distance has gained considerable traction, particularly due to the need for covertly recognizing individuals in unconstrained environments with uncooperative subjects. In such environments, the person of interest may not be interacting with the biometric system in a concerted manner. Further, the individual might be moving in this environment characterized by variable illumination and a non-uniform background. Biometric modalities such as fingerprint and iris cannot be easily acquired at large stand-off distances.[2] On the contrary, the face and gait modalities can easily be acquired at a distance, although the smaller spatial resolution of the face at long distances can degrade accuracy of face recognition systems. As a result, gait-based human recognition has received some interest for biometric recognition at a distance. *Gait* is defined as the pattern of locomotion in animals. Human gait, therefore, is the manner in which people walk. While the formal definition of gait refers to human *motion*, practical algorithms for gait recognition include both dynamic and static features (such as body shape) of the moving human body. It can be viewed as a behavioral trait that is impacted by the musculo-skeletal structure of the human body.

Gait recognition is perceived as an attractive solution for distance-based identification for a number of reasons. First and most importantly, human gait has been observed to have some person-specific characteristics. Psychological studies by Cutting and Kozlowski showed that humans are capable of deducing gender and recognizing known individuals based on gait. Second, the gait biometric can be acquired passively and, therefore, explicit subject interaction is not required for data acquisition. Passive collection is beneficial in an environment where subjects are being observed covertly. Finally, discriminatory features of human gait can be extracted in low resolution images. This suggests that expensive camera systems may not be required for gait recognition.

Typically, an algorithm for gait recognition begins with a silhouette extraction process. This component aims to isolate (i.e., segment or localize) the contour of the human body from a video sequence. A simple method for accomplishing this is through background subtraction, on a frame-by-frame basis, although more sophisticated methods based on Gaussian Mixture Models and Hidden Markov Measure Field exist as well. Once the silhouette is determined, features can be extracted for further processing. Methods for feature extraction are typically *model-based* or *model-free*.

Model-based approaches incorporate structural information of the human body either based on *a priori* information or through models of the human body deduced from training data. A wide variety of biped models[3] are commonly used, although they vary in terms of complexity and information extracted. The benefit of a model-based approach is that a good model allows for robust and consistent feature extraction. Since features are obtained from structural information, distortion in silhouette

[2] The distance between the subject and the acquisition device is referred to as *stand-off distance*.

[3] Biped models are two-legged models.

shape is less likely to induce error. Model-free approaches, on the other hand, generally aim to extract features based on the movement of the silhouette through time. The primary advantage of a model-free approach is computational simplicity, as many algorithms of this class can be executed rapidly. However, a commonly cited concern of model-free algorithms is their inability to adapt to silhouette distortions arising from variations in camera viewpoint and clothing, or errors in segmentation. In the following section, two popular algorithms for feature extraction will be briefly discussed.

5.3.1 Feature extraction and matching

5.3.1.1 Model-based approach

An example of a model-based approach is a five-link biped model that is used to represent human locomotion. This model is designed for representing gait across the saggital plane (side profile) and is demonstrated in Figure 5.5.

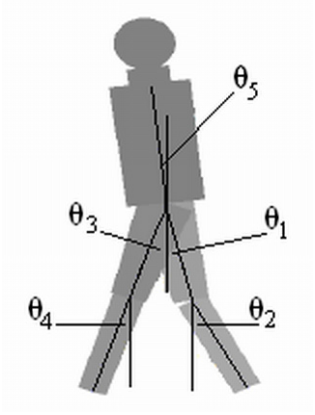


Fig. 5.5 A five-link biped model used for modeling the human body for gait recognition.

Including the coordinates of the centroid, (x,y), the model consists of 7 parameters. Each of the five angles are denoted as sagittal plane elevation angles (SEAs) and defined as the angle between the main axis of the body part and the y axis. Each component of the model is also defined by a height (ℓ), and length component of the top and bottom bases (t and b). Using $\alpha = t/l$ and $\beta = b/l$, each part p_i, $i = 1,\ldots,5$ is represented as $p_i = \{\alpha_i, \beta_i, \ell_i\}$. These part heights are further normalized with respect to the trunk (ℓ_5) to obtain scale invariance. The complete model is defined as follows:

$$H = \{K, R, M\} \tag{5.1}$$

$$K = \{\alpha_1, \beta_1, \alpha_2, \beta_2, \ldots, \alpha_5, \beta_5\} \tag{5.2}$$

$$R = \{r_1, r_2, \ldots, r_5\},\ r_i = \ell_i/\ell_5 \tag{5.3}$$

$$M = \{x, y, \theta_1, \theta_2, \theta_3, \theta_4, \theta_5\} \tag{5.4}$$

Using the above parameters, space domain features are computed for recognition. These include ankle elevation (s_1), knee elevation (s_2), ankle stride width (s_3), and knee stride width (s_4). The Discrete Fourier Transform is computed for each of these features and then used as the primary feature vector for recognition. An illustration of these features is provided in Figure 5.6.

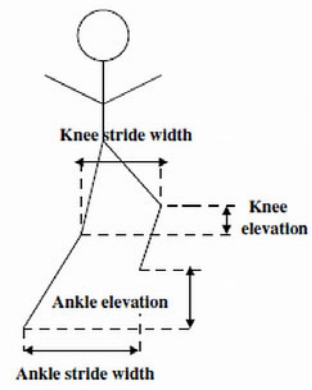


Fig. 5.6 The space domain features extracted from the five-link biped model.

5.3.1.2 Model-free approach

Perhaps the most popular model-free approach is the Gait Energy Image (GEI) algorithm. While it does not result in superior matching performance, this algorithm is often cited as a benchmark for comparison due to its ease of implementation. GEI aims to quantify the gait dynamics of an individual via a single image-based representation. Given N binary silhouette images, $\{S_t(x,y)\}$, at various time instances denoted by t, the gait energy image is defined as:

$$G(x,y) = \frac{1}{N}\sum_{t=1}^{N} S_t(x,y) \tag{5.5}$$

In short, Equation (5.5) represents the averaged silhouette intensity over N frames. Prior to averaging, the images must be normalized such that the height of

each silhouette is the same. In addition, the images must be aligned according to the horizontal centroid. Horizontal alignment allows for the moving shape dynamics to be visualized in the final image. An example is provided in Figure 5.7. Here, the rightmost image in each row represents the gait energy image.

Fig. 5.7 An example showing the normalized silhouette frames along with the gait energy image (GEI). The two rows correspond to two different subjects. In each row, 7 frames are used to derive the GEI which is the rightmost image.

For each gait energy image, gait dynamics is captured in terms of pixel intensity. The energy images are transformed to feature vectors and used for gait recognition.

5.3.1.3 Feature matching

The resulting feature vectors constructed using the methods outlined above often result in large dimensionality, which is difficult to classify, especially when the number of training samples is small. Typically, a combination of Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are used for dimensionality reduction and subspace optimization. The reduced feature vectors are then compared using the Euclidean distance metric.

5.3.2 Challenges in gait recognition

The matching performance of gait recognition algorithms is impacted by factors such as clothing, footwear, walking surface, walking speed, walking direction (with respect to the camera), etc. Further, the gait pattern of an individual can change over time, especially with variations in body mass The impact of these factors is difficult to mitigate and, therefore, evaluation of gait recognition algorithms has been predominantly conducted in controlled environments. This has prevented the incorporation of gait recognition in commercial biometric systems.

5.4 Hand Geometry

Hand geometry, as the name suggests, refers to the geometric structure of the hand. This structure includes width of the fingers at various locations, width of the palm, thickness of the palm, length of the fingers, contour of the palm, etc. Although these metrics do not vary significantly across the population, they can still be used to verify the identity of an individual. Hand geometry measurement is non-intrusive and the verification involves a simple processing of the resulting features. Unlike palmprint, this method does not involve extraction of detailed features of the hand (for example, wrinkles on the skin).

Hand geometry-based verification systems have been commercially available since the early 1970s. The earliest literature on the hand geometry biometric is in the form of patents or application-oriented description. Sidlauskas introduced a 3D hand profile identification apparatus that was successfully used for hand geometry recognition. Hand geometry systems have been deployed in several nuclear power plants across the United States. Also, hand geometry kiosks are present at Ben Gurion airport (Tel Aviv, Israel) for rapid verification of frequent travelers.

A typical hand geometry system consists of four main components: image acquisition, hand segmentation and alignment, feature extraction, and feature matching.

5.4.1 Image capture

Most hand geometry systems acquire an image of the back of the human hand. This image is often referred to as the dorsal aspect of the hand. Thus, most commercial systems require the subject to place their hand on a platen with the palm facing downward. A suitably positioned camera above the hand is then used to acquire an image of the dorsal aspect. Different types of imaging configurations have been discussed in the hand geometry literature as described below.

1. Contact-based vs Contactless: A typical system requires the user to place her hand on a flat surface prior to capturing. Such systems are contact-based and require explicit cooperation of the subject being identified. However, hygiene is an issue of concern to some users in such systems. Furthermore, the large size of the hand (compared to, say, fingers) limits the use of a hand geometry system on smaller devices (e.g., mobile phones). To address these issues, contactless recognition systems have been proposed. However, such systems are required to address the intra-class variability in the captured images due to the articulation of the hand in all three dimensions.
2. Dorsal vs Palmar: Traditionally, an image of the hand is acquired by placing the hand on a flat surface and imaging the back of the hand with a CCD camera. However, there has been interest in capturing the ridge patterns present on the palm and fingers along with the hand shape by imaging the palmar aspect of the hand (that includes the inner palm). See Figure 5.8 (b) for an example. One

drawback of such a system is that it is rather inconvenient for a user to place the hand on a platen with the palm facing upward.

3. Peg-based vs Pegless: In order to guide the positioning of the hand on the platen for imaging purposes (e.g., to prevent the fingers from touching each other), a few pegs are usually placed on the sensor platen. See Figure 5.8 (a). The user is expected to move his hand forward till one of the pegs touches the webbings between a pair of fingers. Although the use of pegs obviates the need for image alignment, it adds to the complexity of using the system and, thus, greater inconvenience to users. See Figure 5.9 for an example where the user has incorrectly placed his fingers around the pegs.

(a) (b)

Fig. 5.8 Different imaging configurations for acquiring a sample of the human hand. a) A constrained hand capture scenario. b) An unconstrained capture scenario.

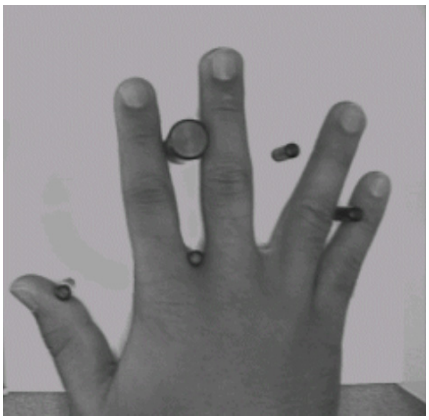


Fig. 5.9 Example showing incorrect placement of the hand in a peg-based system..

5.4.2 Hand segmentation

After the hand image is captured, the hand boundary must be extracted in order to determine the region of interest. To accomplish this, typically, the image is thresholded in order to deduce the region associated with the hand. This is followed by certain morphological operators (e.g., dilation and erosion followed by a connected region analysis) to extract the silhouette of the hand. If the image is very noisy (e.g., due to variable illumination and shadows), then more complex segmentation techniques such as the mean shift algorithm may be needed.

The segmented hand may still contain some artifacts such as the pegs on the platen, rings worn by the user, clothing that covers certain parts of the hand, and discontinuous contours due to non-uniform lighting. These artifacts are removed using specialized image processing techniques that are tailored to specific artifacts. Once a reliable hand shape is obtained, a hand contour is extracted and is used for further processing.

Due to variations in the way in which the users place their hand, the silhouettes extracted from multiple captures of the same hand may not be precisely aligned. It is important to account for these variations before extracting features, especially if the features are not invariant to such geometric transformations. A typical set of transformations include the rotation and translation of the hand, and the movement of individual fingers. It is easier to take into account the global affine transformations such as translation or rotation of the entire hand, while accommodating the movement of a single finger automatically is relatively difficult. To remedy this situation, the segmented hand may be further divided to obtain smaller segments corresponding to individual fingers. Features can then be extracted separately from each of these segments.

5.4.3 Feature Extraction

Typically, two kinds of features are extracted from a hand or a finger silhouette: one-dimensional geometric measurements and two-dimensional shape-based features. The geometric measurements include length and width of fingers, length and width of the palm, and thickness of the fingers. See Figure 5.10 for an example of geometric measurements obtained from a hand image.

It is also possible to use the set of points along the contours of the silhouette (or the segmented image itself) as features. The dimensionality of these features can be reduced in order to obtain a more discriminative and compact representation of the hand. Such shape-based features are expected to be more discriminative than the geometric features due to the fact that they model the structure of the entire hand, rather than portions of the hand, thereby exploiting more information.

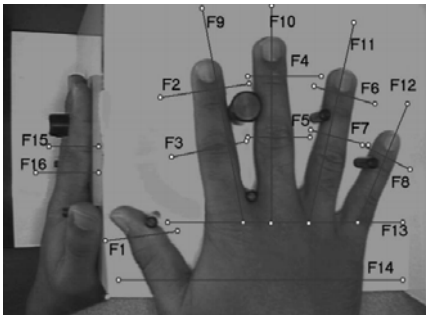


Fig. 5.10 This figure illustrates the axes along which the geometric measurements of the hand image are extracted. As can be seen here, the geometric measurements include the lengths of the fingers, widths of the fingers, width of the palm and the depth of the hand. While individual measurements may not be discriminative enough for biometric recognition, an agglomeration of these measurements results in a feature vector that can effectively be used for biometric verification.

5.4.4 Feature matching

The features extracted from a segmented hand image can often be denoted as a feature vector in the Euclidean space. Consequently, common distance measures such as Euclidean and Manhattan distances can be effectively used to compare two hand images. In case a sufficient amount of training data is available, more sophisticated distance measures such as the Mahalanobis distance[4] can also be used for robust matching. If the sampling of points on the silhouette are directly used for matching,

[4] Mahalanobis distance essentially weights the different dimensions of the feature vector based on variation of the features along that dimension. For example, if there is a large variation of fea-

then distance measures such as the Hausdorff distance[5] can be used to obtain the matching scores. It is also possible to use machine learning schemes to design classifiers such as multi-class Support Vector Machines (SVM) that can map the input feature set into one of many identities.

5.4.5 Challenges in hand geometry recognition

Hand geometry systems have been successfully deployed in several applications including nuclear power plants, border control systems, recreational centers and time-and-attendance systems. In these applications, the biometric system typically operates in the *verification* mode. Since the hand geometry of subsets of individuals can be similar, the *identification* accuracy due to this biometric modality can be low. Further, the shape of an individual's hand can change with time - a factor that is especially pronounced in young children. More recent research has explored the use of hand geometry in conjunction with fingerprints and low-resolution palmprints in a multibiometric configuration for improved accuracy.

5.5 Soft Biometrics

There are many situations where primary biometric traits (i.e., face, fingerprint and iris) are either corrupted or unavailable, and the soft biometric information is the only available clue to solving a crime. For example, while a surveillance video may not capture the complete face of a suspect, the face image in the video may reveal the suspect's gender and ethnicity, or the presence of a mark or tattoo may provide additional valuable clues. We will discuss some of the soft biometric traits (i.e., periocular, facial marks, and tattoos) below. The periocular biometric is gaining increasing attention since it offers a trade-off between using the entire face image and the iris portion only. Facial marks and tattoos are also gaining widespread attention since they offer complementary information that can be exploited along with primary biometric traits.

---

ture values across one dimension, differences along that dimension are assigned a smaller weight compared to those dimensions along which the variation is small.

[5] Hausdorff distance between two set of points A and B is given by $H(A,B) = \max\{h(A,B),h(B,A)\}$ where $h(A,B) = \max_{a\in A}\{\min_{b\in B} d(a,b)\}$ and d is a measure of distance between two points such as the Euclidean distance. Essentially, Hausdorff distance is the distance between a point in one set that is farthest away from all points in the other set, to its closest point in the other set.

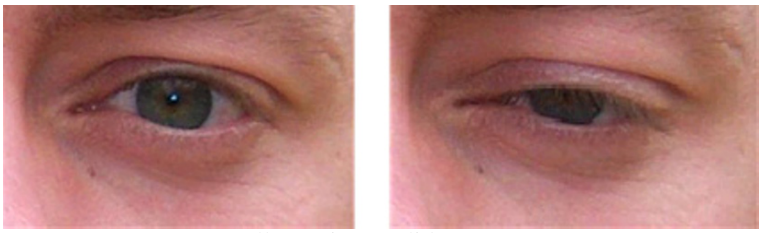
5.5.1 Periocular

The periocular region represents the region around the eyes. It predominantly consists of the skin, eyebrow, and eye. The use of the periocular region as a biometric cue represents a good trade-off between using the entire face region or using only the iris for recognition. When the entire face is imaged from a distance, the iris information is typically of low resolution; this means the matching performance due to the iris modality will be poor. On the other hand, when the iris is imaged at small standoff (typically, 1 meter), the entire face may not be available, thereby forcing the recognition system to rely only on the iris. However, the periocular biometric can be used for a wide range of distances. Figure 5.11 shows sample periocular images collected from two different subjects. Periocular images can also be captured in the NIR spectrum to minimize illumination variation compared to visible spectrum. The key steps in preprocessing, feature extraction, and matching visible band periocular images are described below.

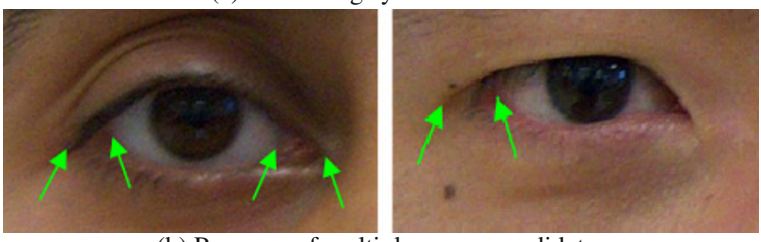
Fig. 5.11 Example of periocular images from two different subjects.

5.5.1.1 Preprocessing

Periocular images contain common components across images (i.e., iris, sclera, and eyelids) that can be represented in a common coordinate system. Once a common area of interest is localized, a global feature representation scheme can be used. A global representation scheme is holistic in that it characterizes the entire periocular region rather than only the local regions. The iris or eyelids are good candidates for the alignment process. While iris detection can be performed fairly well due to the approximately circular geometry of the iris and a good contrast between the iris and the sclera, accurately detecting the eyelids is more difficult. The inner and outer corners of the eye can also be considered as reference points, but there can be multiple candidates as shown in Figure 5.12. Therefore, we will discuss an iris-based image alignment method in this section. The iris can be used for translation and scale normalization of the image, but not for rotation normalization. However, small rotation variations can be overcome by using a rotation tolerant feature representation in the feature extraction stage.

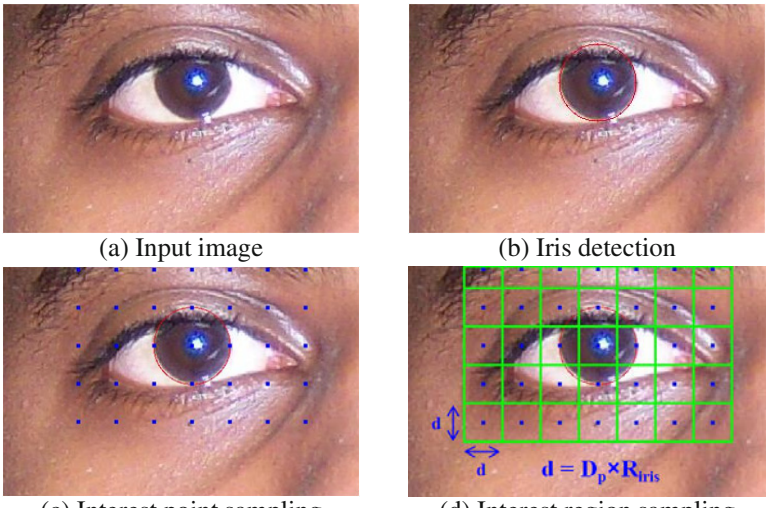


(a) Illustrating eyelid movement



(b) Presence of multiple corner candidates

Fig. 5.12 Example images showing difficulties in periocular image alignment.



(a) Input image (b) Iris detection

(c) Interest point sampling (d) Interest region sampling

Fig. 5.13 Schematic of image alignment and feature extraction processes for periocular images.

5.5.1.2 Feature extraction

The features can be extracted using all the pixel values in the region of interest that is defined with respect to the iris. From the center, C_{iris}, and the radius, R_{iris}, of the iris, multiple ($=n_{pi}$) interest points $p_1, p_2, \ldots, p_{n_{pi}}$ are selected within a rectangular window defined around C_{iris} with a width of $6\times R_{iris}$ and a height of $4\times R_{iris}$ as shown

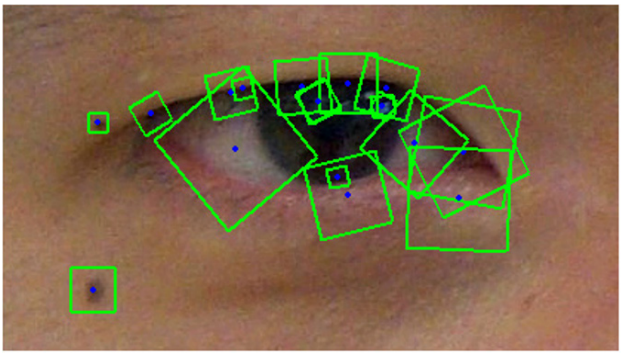


Fig. 5.14 Examples of local features and bounding boxes for descriptor construction using SIFT operator. Each bounding box is rotated with respect to the major orientation or gradient.

in Figure 5.13. The number of interest points is decided based on the sampling frequency ($1/D_p$), which is inversely proportional to the distance between interest points, $D_p \times R_{iris}$. For each interest point p_i, a rectangular region r_i is defined. The dimension of each rectangle (r_i) in the ROI is of size ($D_p \times R_{iris}$) by ($D_p \times R_{iris}$). When D_p=1, the size of the rectangle becomes $R_{iris} \times R_{iris}$ (see Figure 5.13 (d)).

For the construction of the descriptor in each region, r_i, some of the distribution-based descriptors such as gradient orientation (GO) histogram and local binary pattern (LBP) can be used. The responses of GO and LBP are quantized into 8 distinct values to build an eight bin histogram in each sub-region. The eight bin histogram is constructed from a partitioned sub-region and concatenated across the various sub-regions to construct a feature vector. A Gaussian blurring with a standard deviation, σ, can be applied prior to extracting features using the GO and LBP methods in order to smooth out variations across local pixel values.

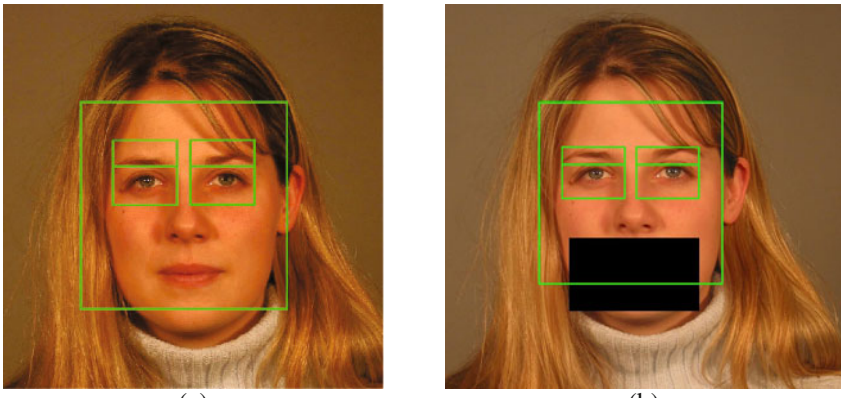
Alternatively, a set of salient key points can be detected in scale space following the Scale Invariant Feature Transformation (SIFT) method. Features are extracted from the bounding boxes for each key point based on the gradient magnitude and orientation. The size of the bounding box is proportional to the scale (i.e., the standard deviation of the Gaussian kernel in scale space construction). Figure 5.14 shows the detected key points and the surrounding boxes on a periocular image. While the GO and LBP features are extracted only around the eye, the SIFT features are extracted from all salient regions. Therefore, the SIFT approach is expected to provide more distinctiveness across subjects.

5.5.1.3 Feature matching

For the GO and LBP features, the Euclidean distance is used to calculate the match scores. The distance ratio-based matching scheme is used for the SIFT as described below.

Given an image I_i, a set of SIFT key points $K_i = \{k_{i1}, k_{i2}, \cdots, k_{in}\}$ is detected. In matching a pair of images I_i and I_j, all the keypoints K_i from I_i and K_j from I_j are compared to determine how many keypoints are successfully matched. The Euclidean distances from k_{ia} to all the key points in K_j are calculated to obtain the closest distance d_1 and the second closest distance d_2. When the ratio d_1/d_2 is sufficiently small (less than a threshold th_{ratio}), k_{ia} is considered to have a matching key point in K_j. By using the ratio of d_1 and d_2, both the similarity and the uniqueness of point pairs are considered.

Recent studies on person identification using periocular biometric traits, both in visible and NIR spectra, show high identification accuracies (over 80%). However, such an accuracy is possible only when the images are of good quality and exhibit low intra-class variations. It has also been shown that the periocular biometric can assist person identification when the face is occluded. Figure 5.15 shows examples of automated face and periocular region detection for (a) a full face and (b) an occluded face. Figure 5.16 shows example matching results of periocular images.



(a) (b)

Fig. 5.15 Examples of automated face and periocular region detection on (a) a full face and (b) an occluded face. Face recognition performance degrades with the occluded face, but the periocular biometric still shows good identification performance. Images are from the FRGC 2.0 database.

5.5.2 Face marks

Advances in sensing technology have made it easy to capture high resolution face images. From these high resolution face images, it is possible to extract details of skin irregularities, also known as facial marks. This has opened new possibilities in face representation and matching schemes. These skin details are mostly ignored and considered as noise in a typical face recognition system. However, facial marks can be used to (a) supplement existing facial matchers to improve the identification

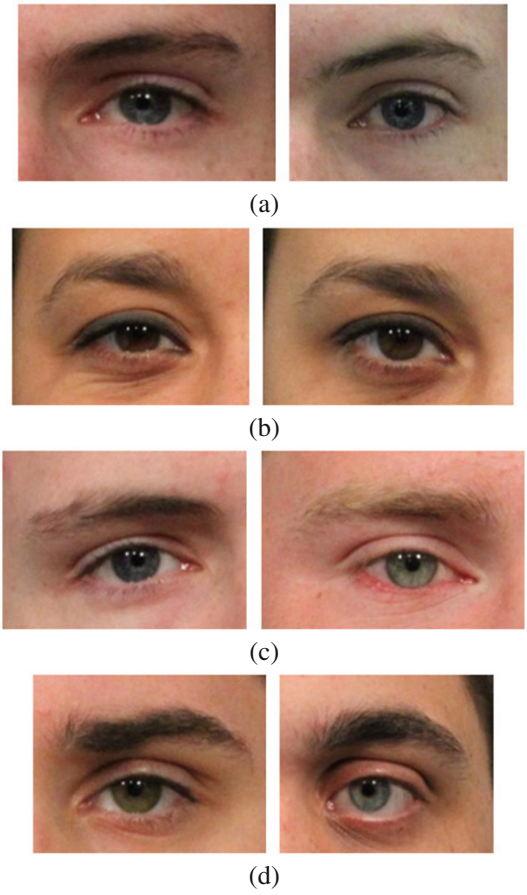
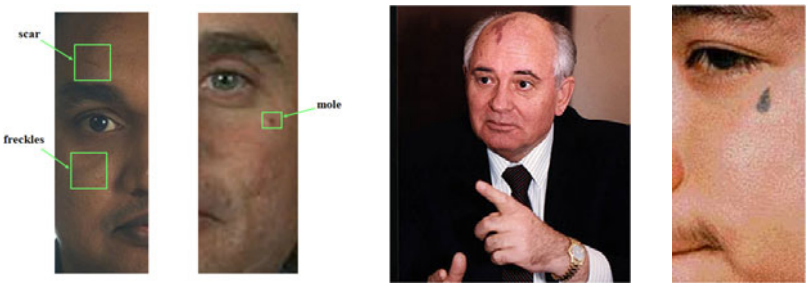


Fig. 5.16 Example matching results with four different pairs of periocular images from four different subjects. Image pairs in (a) and (b) are successfully matched whereas image pairs in (c) and (d) failed to match. The reasons for failure are due to intra-class variations caused by pose and illumination changes and movements of eyebrow and eyelids.

accuracy, (b) facilitate fast face image retrieval, (c) enable matching or retrieval with partial or off-frontal face images, and (d) provide more descriptive evidence about the similarity or dissimilarity between face images, which can be used as evidence in legal proceedings. Figure 5.17 shows some representative facial mark types.

5.5.2.1 Preprocessing

To represent the face marks in a common coordinate system, primary facial features such as eyes, eye brows, nose, mouth, and face boundary (Figure 5.19) are detected by using an Active Appearance Model (AAM) or Active Shape Model



(a) Scar, mole, and freckles[1] (b) Large birthmark[2] (c) Tattoo[3]

[1] FERET database.
[2] http://www.wnd.com/index.php?fa=PAGE.view&pageId=63558.
[3] http://blog.case.edu/colin.mulholland/2007/09/20/a_way_of_expression.

Fig. 5.17 Examples of facial marks.

(ASM). These primary facial features will be disregarded in the subsequent facial mark detection process. Example landmarks detected in face images are shown in Figure 5.19.

Using the detected landmarks, the face image is mapped to the mean shape to simplify the mark detection, matching, and retrieval. Let S_i, i = 1, ..., N represent the shape of each of the N face images in the database (gallery) based on the set of landmarks. Then, the mean shape is simply defined as $S_\mu = \sum_{i=1}^{N} S_i$. Each face image, S_i, is mapped to the mean shape, S_μ, by using the Barycentric coordinate based texture mapping process. First, both S_i and S_μ are subdivided into a set of triangles. Given a triangle T in S_i, its corresponding triangle T' is found in S_μ. Let r_1, r_2, and r_3 (r'_1, r'_2, and r'_3) be the three vertices of T (T'). Then, any point, p, inside T is expressed as $p = \alpha r_1 + \beta r_2 + \gamma r_3$ and the corresponding point p' in T' is similarly expressed as $p' = \alpha r'_1 + \beta r'_2 + \gamma r'_3$, where $\alpha + \beta + \gamma = 1$. This way, the pixel value at p is mapped to p'. Figure 5.18 shows the schematic of the Barycentric mapping process. By repeating this mapping process for all the points inside all triangles, the texture in S_i is mapped to S_μ.

After this mapping process, all face images are normalized in terms of scale and rotation and this allows for the representation of each facial mark in a face-centered common coordinate system. Figure 5.19 shows the schematic of mean face construction.

A blob detection operator is applied to the face image mapped into the mean shape. To suppress false positives in the blob detection process caused by the presence of primary facial features, a generic mask, denoted by M_g, is constructed from the mean shape S_μ. However, the generic mask does not cover user-specific facial features such as a beard or small wrinkles around the eyes or mouth that are likely to increase the false positives. Therefore, a user specific mask, M_s, is also built using the edge image. The edge image is obtained by using the conventional Sobel operator. The user specific mask M_s, constructed as a sum of M_g and edges that are

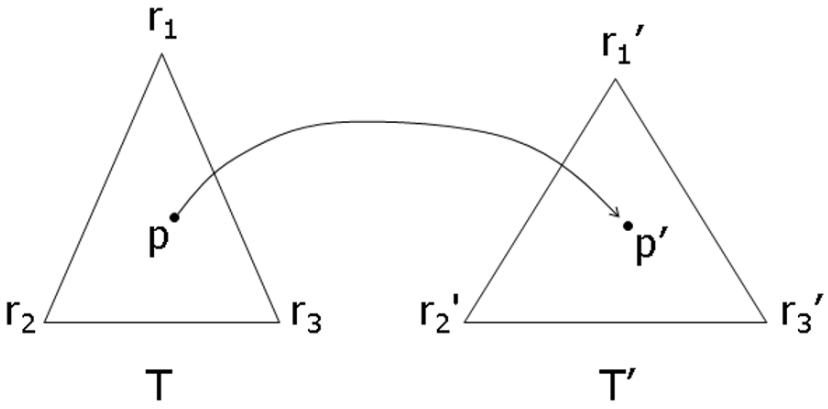


Fig. 5.18 Schematic of texture mapping process using the triangular Barycentric coordinate system.

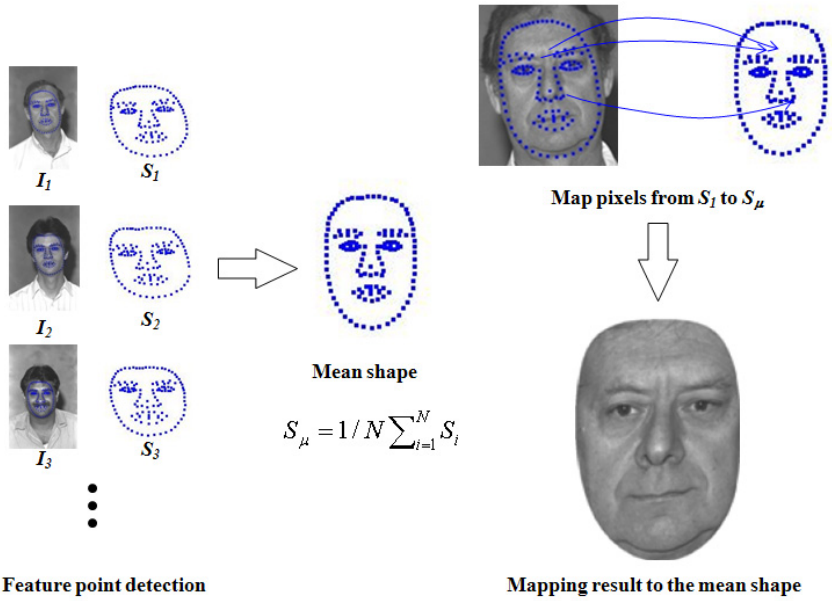


Fig. 5.19 Schematic showing the construction of the mean face image.

connected to M_g, helps in removing most of the false positives appearing around the beard or small wrinkles around the eyes or mouth.

5.5.2.2 Mark detection and classification

Lindeberg proposed that local maxima over multiple image scales of normalized Gaussian derivatives (i.e., $\sigma^2\nabla^2 G$) reflects the characteristic size of a local structure . This allows for a blob detection with automatic scale selection, which is invariant with image scale. Motivated by this, we detect facial marks via scale space analysis. The scale space extrema detection starts by constructing a normalized multi-scale representation of the face image by convolving the input images, $I(x,y)$, with a Laplacian of Gaussian (LoG) filter with a sequence of σ_k as

$$D(x,y,\sigma_k) = \sigma_k^2\nabla^2 G(x,y,\sigma_k) * I(x,y), \quad k = 1,2,\cdots,n, \tag{5.6}$$

where $\sigma_k^2\nabla^2 G$ is the scale-normalized Laplacian of Gaussian operator, and $\sigma_k = k\sigma_0, k = 1,2,\cdots,n$, with σ_0 being a constant value (=$\sqrt{2}$) for the initial scale.

Next, the local extrema over both spatial and scale space in every $3 \times 3 \times 3$ image block is detected . The detected candidate mark locations have the following characteristics:

- The detected location contains candidate facial marks.
- The detected scale (σ_k) indicates the size of the corresponding facial mark.
- The absolute value of $D(x,y,\sigma)$ reflects the strength of the response. This strength may be used as the confidence value to select stable marks.
- The sign of $D(x,y,\sigma)$ helps in assessing the pixel intensity of the facial mark. A positive (negative) sign represents a dark (bright) facial mark with brighter (darker) surrounding skin.

The overall mark detection process is shown in Figure 5.20.

For each detected local extrema, a local bounding box, whose size is proportional to the associated scale, is determined. Pixels in the bounding box are binarized with a threshold value selected as the mean of the surrounding pixels. A local extrema is darker or brighter than its surrounding region, so the average value of the surrounding area can serve to effectively segment the blob in the bounding box. Next, the blobs are classified in a hierarchical fashion: 'linear' versus 'all' followed by 'circular' versus 'irregular'. For determining linearity, two eigenvalues λ_1 and λ_2 are obtained through the eigen decomposition of the spatial coordinates of the blob pixels. When λ_1 is significantly larger than λ_2, the mark is classified as being a linear blob. For circularity detection, the following observation is used: a circle, with radius M_2 will enclose most of the blob pixels if they are circularly distributed. Therefore, a decision of 'circular' or 'irregular' can be made based on the ratio of the number of pixels within and outside of this circle. The pixel intensity of the blob can be deduced based on the sign of $D(x,y,\sigma)$, as stated earlier. The schematic for blob classification is shown in Figure 5.21.

Figure 5.22 illustrates five examples of facial detection results using the proposed mark detection and classification method. It can be observed that the proposed method is robust to noise and provides good estimates of the size and class of the marks.

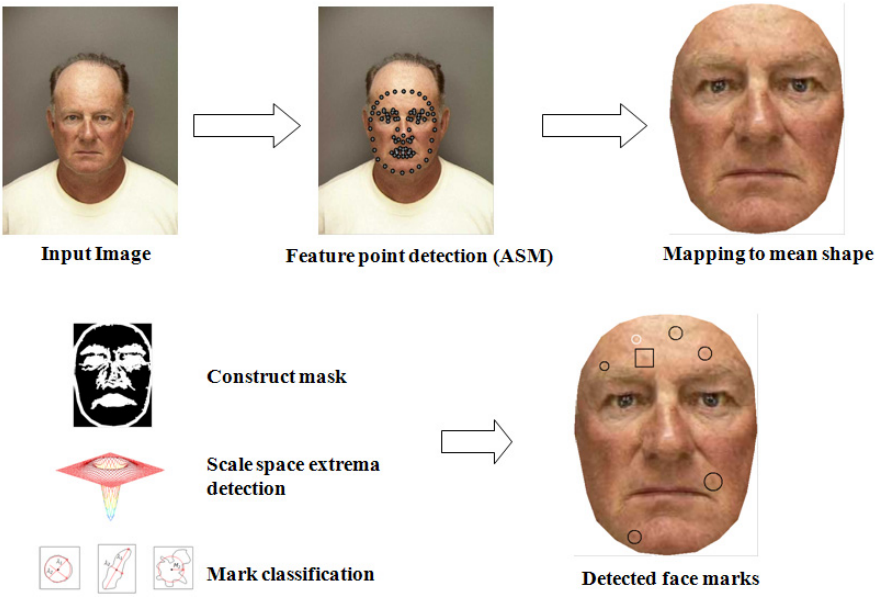


Fig. 5.20 Schematic of the mark detection process.

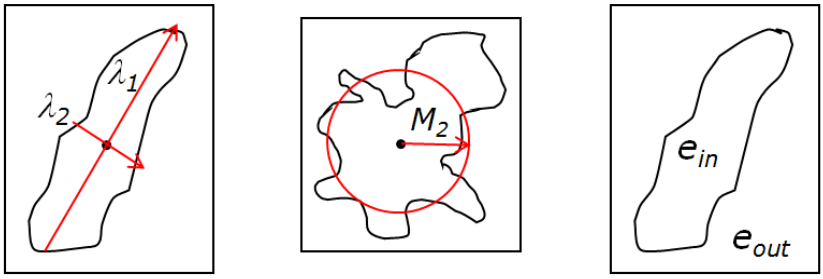


Fig. 5.21 Schematic of the mark classification scheme. Left: The eigenvalues λ_1 and λ_2 are computed based on the the spatial coordinates of the blob pixels. In this example, the largest eigenvalue, λ_1, is significantly larger than the smallest eigenvalue, λ_2, thereby suggesting linearity. Middle: The circular nature of the mark can be deduced after fitting a circle to the blob pixels. Here, a circle of radius M_2 is used to enclose the pixels inside the blob. Since the ratio of the number of blob pixels within the circle to the number of blob pixels outside the circle is large, this mark may be labeled as being 'circular'. Right: The sign of the difference in pixel intensities between the interior (e_{in}) and exterior (e_{out}) of the blob can help assess whether the mark is bright or dark.

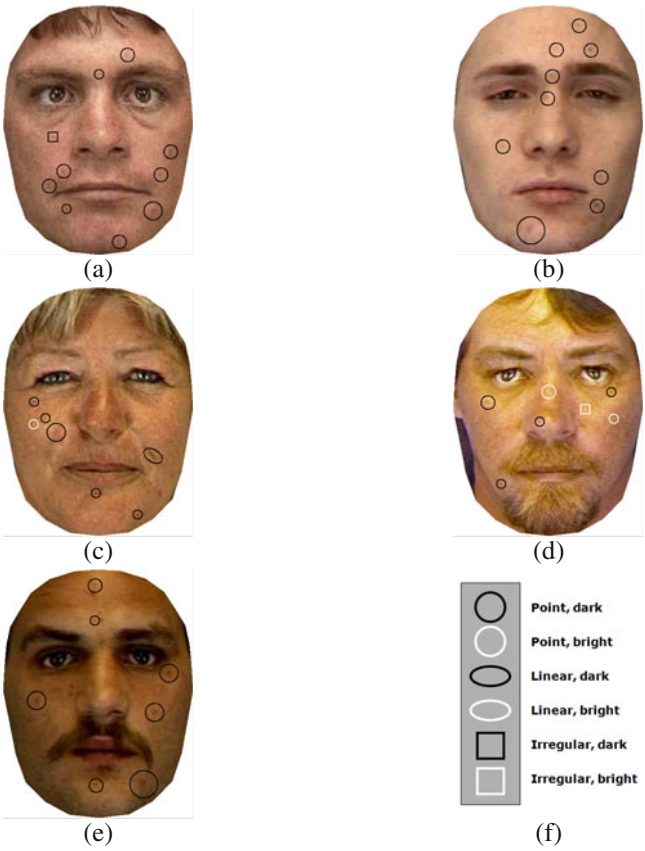


Fig. 5.22 Example mark detection and classification results. (a), (b), (c), (d), (e) are example images with detected marks. (f) Symbols used to denote six different mark classes.

5.5.2.3 Mark matching

The detected facial marks are encoded into a 48-bin histogram representing the morphology, color, and location of facial marks. To encode the location information of facial marks, the face image in the mean shape space is subdivided into eight different regions as shown in Figure 5.23. Each mark is encoded by a six digit binary number representing its morphology and color. When there is more than one mark in the same region, a bit by bit summation is performed. The six bin values are concatenated for the eight different regions in the order as shown in Figure 5.23 to generate the 48-bin histogram. If a mark is observed on the borderline of the face segments, it is included into both the regions. Given the indices obtained from face images, the histogram intersection method is used to calculate the matching scores. Let $H^1(i)$ and $H^2(j)$ be the two histograms representing the mark indices, then the histogram

intersection is calculated as $\sum_{k=1}^{48} \left(H^1(k) \& H^2(k)\right)$, where & represents the logical *and* operation. The score range of the mark index based matching is [0,48].

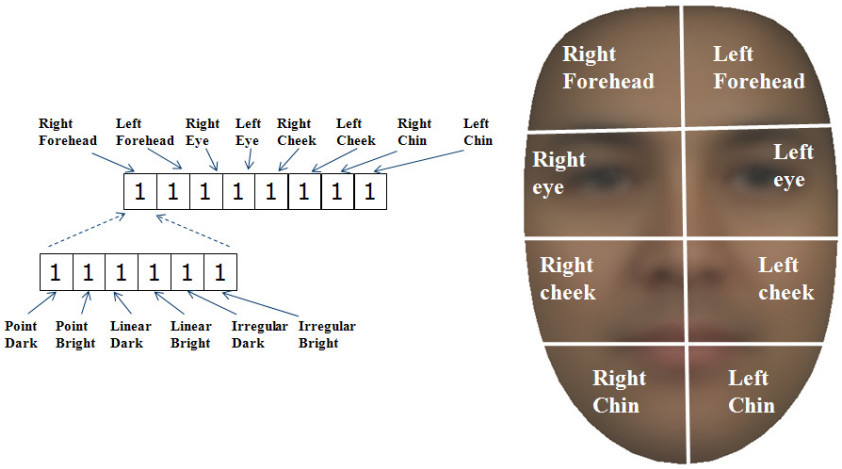


Fig. 5.23 Schematic of the mark-based indexing scheme.

Figure 5.24 shows example mark detection and matching results with an identical twin database collected by the University of Notre Dame. In all the four pairs of identical twins shown in this figure, facial marks help in differentiating identical twins when combined with a leading commercial face recognition engine.

5.5.3 Tattoo

The use of tattoos imprinted on human body in suspect identification started with the Bertillon system. Since then, images of tattoos on the human body have been routinely collected and used by law enforcement agencies to assist in suspect and victim identification. When the primary biometric traits are unavailable or corrupted, tattoos can be used to identify victims or suspects. Figure 5.25 shows examples of tattoo images used in victim and suspect identifications.

Tattoos provide more discriminative information than the traditional demographic indicators such as age, height, race, and gender for person identification. Many individuals acquire tattoos in order to display their personality, or to indicate their their membership in a group. Therefore, recognition of tattoos can provide a better understanding of an individuals' background and membership in various organizations, especially criminal gangs. Figure 5.26 shows example tattoo images showing membership in Mexikanemi Mafia gang.

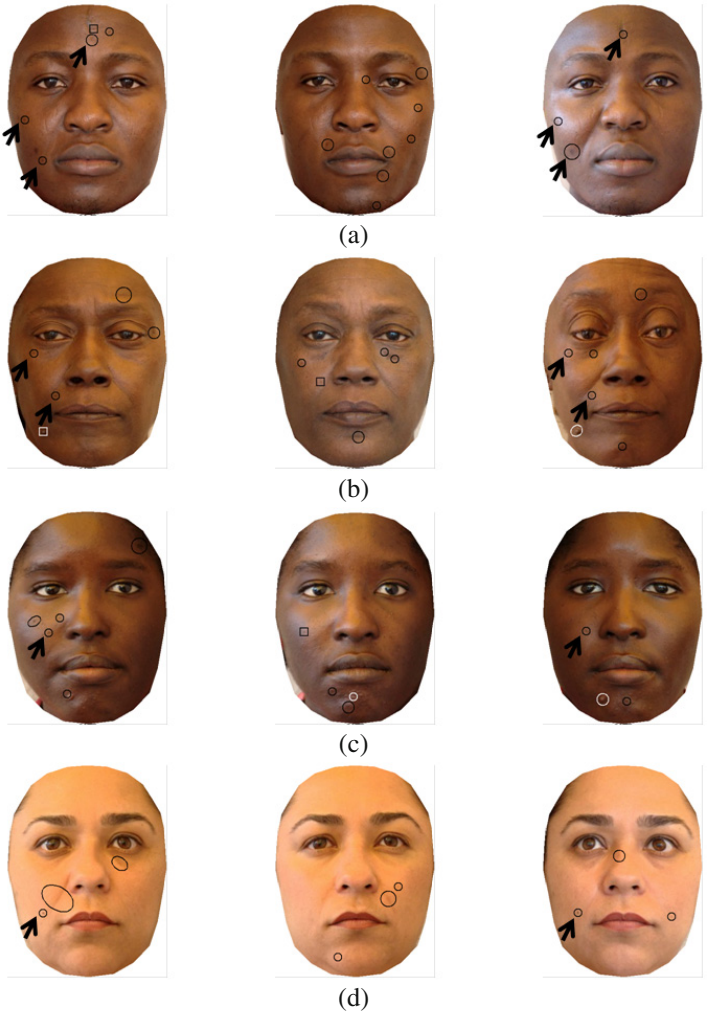


Fig. 5.24 Example matching results with four different identical twins. Images in the first and third columns belong to the same subject and images in the second column are the twins. The first, second, and third columns correspond to the probe, incorrect match using FaceVACS only, and correctly matched true mates using FaceVACS and mark index, respectively. Black arrows in each row show the facial marks correctly detected and classified that contribute to individualize the identical twins.

The current practice of matching tattoos is based on a set of predefined keywords. Assigning keywords to individual tattoo images is both tedious and subjective. A content-based image retrieval system for tattoo image retrieval can overcome many of the limitations of the keyword-based tattoo retrieval. We will briefly describe a system that extracts local image features based on the Scale Invariant Feature

Transform (SIFT). Contextual or side information, i.e., location of tattoo on the body and keyword (class) assigned to the tattoo, is utilized to improve both the retrieval time and retrieval accuracy. Geometrical constraints are also introduced in SIFT keypoint matching to reduce false retrievals.

(a) (b)

Fig. 5.25 Example tattoo images used in identifying (a) an Asian Tsunami (2004) victim in Indonesia and (b) a suspect.

(a) (b) (c) (d)

Fig. 5.26 Example gang tattoo images indicating the membership in the Mexikanemi Mafia gang in Texas.

5.5.3.1 Tattoo classes and body location

The ANSI/NIST-ITL1-2011 standard defines eight major classes (i.e. human, animal, plant, flag, object, abstract, symbol, and other) and a total of 70 subclasses (e.g. male face, cat, narcotics, American flag, fire, figure, national symbols, and wording) for categorizing tattoos. A search of a tattoo image database involves matching the class label of a query tattoo with the labels of the tattoos in the database. This tattoo matching procedure based on manually assigned ANSI/NIST class labels has

the following limitations: (a) class label does not capture the semantic information contained in tattoo images; (b) labeling millions of tattoo images maintained by law enforcement agencies is both subjective and time consuming; (c) tattoos often contain multiple objects and cannot be adequately classified into a single ANSI/NIST classes; (d) tattoo images have large intra-class variability; and (e) ANSI/NIST classes are not complete for describing new tattoo designs.

Tattoo location on the body is a useful piece of information because it can be tagged precisely and objectively. Hence, searching for similar images at the same body location can significantly reduce the matching time without any loss of matching accuracy. The National Crime Information Center (NCIC) has defined 31 major categories (e.g., arm, calf, and finger), and 71 sub categories (e.g., left upper arm, right calf, and left hand finger) for denoting body location of tattoos.

5.5.3.2 Feature extraction

Scale Invariant Feature Transform (SIFT) is a well known and robust local feature based approach used for object recognition. It has been shown that SIFT features provide better performance than low-level image attributes, (e.g., color, texture, and shape) for tattoo image matching and retrieval. SIFT extracts repeatable characteristic feature points from an image and generates descriptors representing the texture around the feature points. These feature points are invariant to image scale and rotation, and are shown to provide robust matching across a range of affine distortion, change in 3D viewpoints, additive noise, and change in illumination.

5.5.3.3 Feature Matching

In addition to the conventional SIFT matching process as described in section 5.5.1.3, a set of local geometric constraints can be applied to reduce the number of false matching points. Let M_{ij} represent the set of matching keypoints between two images I_i and I_j. Then, M_{ij} can be expressed in terms of two different subsets $M_{ij} = M_{ij,T} \cup M_{ij,F}$, where $M_{ij,T}$ represents the set of true matching points and $M_{ij,F}$ represents the set of false matching points. It is expected that removing the false matching points will increase the retrieval accuracy. The number of false matchings in the presence of viewpoint variations or blurring in the image is likely to be large. When a key point belongs to $M_{ij,F}$, it is likely to match to many other key points. On the other hand, a keypoint in $M_{ij,T}$ is likely to match to either one or a very small number of other key points. Given a query image I, it is matched with all the images in the gallery database D and the number of matching points is obtained for each gallery image. Let L_m, $m = 1, 2, 3, \ldots$, represent the set of key points in the query image that are matched into the same key point in D. Let the size of the area covered by L_m be A_m. Then, L_m is regarded as belonging to $M_{ij,F}$ if A_m is larger than a threshold t ($t = 0.2$). All the matching keypoints not in $M_{ij,F}$ are regarded as true matching points. Finally, the number of key points that belong to $M_{ij,T}$ is used to retrieve the

top-N candidate images from the tattoo image database. Figure 5.27 shows example matching results on a duplicate and non-duplicate tattoo image pairs. The number of matching keypoints for a duplicate pair is observed to be much larger than that for a non-duplicate pair.

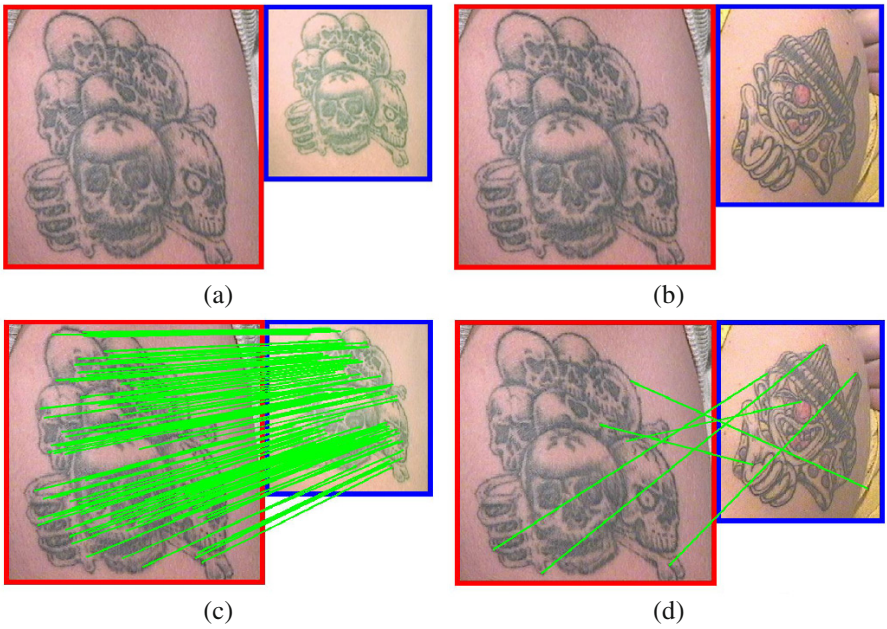


Fig. 5.27 Tattoo matching examples. A pair of (a) duplicate and (b) non-duplicate tattoo images and their SIFT matching results. There are 129 and 12 matching keypoints in the (c) duplicate and (d) non-duplicate pairs, respectively. Number of octaves and scales are selected as 3 and 4, respectively. The feature vector associated with each keypoint has 128 dimensions.

Figure 5.28 shows example tattoo images retrieval results based on the SIFT matcher, body location, and ANSI/NIST class label for three different queries. For query 1, three different duplicate images from the database were successfully retrieved at rank 1, 2 and 6. The corresponding match scores for these three retrieved images are: 163, 157 and 26. The low score for the third retrieved image at rank 6 is possibly due to low contrast (fading) of the tattoo. Two duplicate tattoos were retrieved, at ranks 1 and 2, for query 2 and one duplicate image was retrieved for query 3 at rank 6. Note that due to the small size of the tattoo in query 3, the number of SIFT keypoints extracted is small and hence all the match scores for the retrieved images are low as well.

Fig. 5.28 Retrieval examples. Each row shows a query and its top-7 retrieved images along with the associated match scores.

5.6 Summary

A wide variety of biometric traits have been studied in the literature. Given the growing application domain for biometrics, it is likely that the use of soft biometric[6] traits such as ear, gait, hand geometry, and soft biometrics will become prevalent as well as necessary in some contexts. This would allow for the design of effective biometric systems that utilize the most appropriate set of biometric traits based on factors such as the application scenario, nature of the target population, availability of biometric traits, computational resources available, etc. Thus, investigating new biometric modalities for potential use especially in unconstrained and uncooperative user scenarios has its benefits.

At the same time, one needs to establish the distinctiveness of a biometric trait, its permanence and susceptibility to spoofing. But this can be a rather tedious and long-winding exercise. For example, although fingerprint matching has been used for over 100 years in forensics, the uniqueness or individuality of the fingerprint is still an ongoing area of research. Similarly, modeling of the aging process of the face and its effect on face matching accuracy is an active area of research. In the case of iris, there is a lack of longitudinal data to even begin to investigate the aging phenomena. Thus, as new biometric traits are studied, the onus is on researchers to analyze the advantages and disadvantages of each trait and determine the value added.

[6] Scars, marks, and tattoos (SMT) are an integral part of the FBI's Next Generation Identification (NGI) system.

References

1. ANSI/NIST-ITL 1-2011, Data Format for the Interchange of Fingerprint, Facial, & Scar, Mark & Tattoo (SMT).
2. A. Bertillon (R.W. McClaughry Translation). *Signaletic Instructions including the theory and practice of Anthropometrical Identification*. The Werner Company, 1896.
3. G. Amayeh, G. Bebis, A. Erol, and M. Nicolescu. Peg-free hand shape verification using high order zernike moments. In *Proceedings of the IEEE Workshop on Biometrics at CVPR*, New York, 2006.
4. B. Bhanu and H. Chen. *Human Ear Recognition by Computer*. Springer, 1st edition edition, 2008.
5. M. Burge and W. Burger. Ear biometrics in computer vision. In *Proc. of the* 15^{th} *International Conference on Pattern Recognition ICPR*, pages 826–830, Barcelona, Spain, 2000.
6. J. Bustard and M. Nixon. Robust 2D ear registration and recognition based on SIFT point matching. In *Proc. of the Biometrics: Theory, Applications, and Systems BTAS*, Washington, DC, USA, 2008.
7. K. Chang, K. Bowyer, S. Sarkar, and B. Victor. Comparison and combination of ear and face images in appearance-based biometrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25:1160–1165, 2003.
8. J. Cutting and L. Kozlowski. Recognizing friends by their walk. *Bulletin of the Psychonomic Society*, 9(5):353–356, 1977.
9. N. Duta. A survey of biometric technology based on hand shape. *Pattern Recognition*, 42:2797–2806, 2009.
10. H. Dutagaci, B. Sankur, and E. Yoruk. A comparative analysis of global hand appearance-based person recognition. *Journal of Electronic Imaging*, 17(1), 2008.
11. R. H. Ernst. Hand ID system. United States Patent Number US 3576537, 1971.
12. J. Han and B. Bhanu. Individual recognition using gait energy image. *IEEE Transactions on Patter*, 28(2):316–322, 2006.
13. D. Hurley, M. Nixon, and J. Carter. Force field feature extraction for ear biometrics. *Computer Vision and Image Understanding*, 98(3):491–512, 2005.
14. I. H. Jacoby, A. J. Giordano, and W. H. Fioretti. Personnel identification apparatus. United States Patent Number US 3648240, 1972.
15. A. K. Jain, A. Ross, and S. Pankanti. A prototype hand geometry-based verification system. In *2nd International Conference on Audio- and Video-based Biometric Person Authentication (AVBPA)*, pages 166–171, Washington D.C., March 1999.
16. L. Kozlowski and J. Cutting. Recognizing the sex of a walker from a dynamic point light display. *Perception and Psycophysics*, 21:575–580, 1977.
17. A. Kumar, D. C. M. Wong, H. C. Shen, and A. K. Jain. Personal authentication using hand images. *Pattern Recognition Letters*, 27:1478–1486, 2007.
18. J-E. Lee, A. K. Jain, and R. Jin. Scars, Marks and Tattoos (SMT): Soft Biometric for Suspect and Victim Identification. In *Proceedings of Biometrics Symposium at the Biometric Consortium Conference*, Florida, USA, September 2008.
19. T. Lindeberg. Feature detection with automatic scale selection. *International Journal of Computer Vision*, 30(2):79–116, 1998.
20. D. G. Lowe. Distinctive image features from scale invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.
21. L. Meijerman, A. Thean, C. van der Lugt, R. van Munster, G. van Antwerpen, and G. Maat. Individualization of earprints: Variation in prints of monozygotic twins. *Forensic Science, Medicine, and Pathology*, 2(1):39–49, 2006.
22. R. P. Miller. Finger dimension comparison identification system. United States Patent Number US 3576538, 1971.
23. M. Nixon, T. Tan, and R. Chellappa. *Human Identification Based on Gait*. Springer, 2006.
24. U. Park and A. K. Jain. Face matching and retrieval using soft biometrics. *IEEE Transactions on Information Forensics and Security*, 5(3):406–415, 2010.

25. R. K. Rowe, U. Uludag, M. Demirkus, S. Parthasaradhi, and A. K. Jain. A Multispectral Whole-hand Biometric Authentication System. In *Proceedings of Biometric Symposium, Biometric Consortium Conference*, Baltimore, USA, September 2007.
26. D. P. Sidlauskas. 3d hand profile identification apparatus. United States Patent Number US 4736203, 1988.
27. R. Zhang, C. Vogler, and D. Metaxas. Human gait recognition. *Proceedings of the IEEE International Workshop on Computer Vision and Pattern Recognition*, 2004.