

An approach to hand biometrics in mobile devices

Alberto de Santos Sierra · Carmen Sánchez-Ávila ·
Aitor Mendaza Ormazá · Javier Guerra Casanova

Abstract This paper focuses on hand biometrics applied to images acquired from a mobile device. The system offers the possibility of identifying individuals based on features extracted from hand pictures obtained with a low-quality camera embedded on a mobile device. Furthermore, the acquisitions have been carried out regardless illumination control, orientation, distance to camera, and similar aspects. In addition, the whole system has been tested with an owned database. Finally, the results obtained ($6.0\% \pm 0.2$) and the algorithm structure are both promising in relation to a posterior mobile implementation.

Keywords Contact-less hand biometrics · Mobile devices · Support vector machines · Security · Segmentation

1 Introduction

Hand recognition is a well-known biometric technique, based on geometric features extracted from the hand of an individual. This technique provides a high degree of acceptance among the users due to its lack of invasivity and the performance rate, very suitable for daily applications like the access to a public library, a university, or a research center [1,2]. Furthermore, the increasing demand of security within mobile devices makes this technique a very plausible

option for embedding biometric applications inside current electronic devices such as PDA, mobile phones, and the like.

However, this article proposes a finger geometry recognition system [3–5], focusing only on geometrical aspects from fingers. Moreover, this method requires a contact-less hand image acquisition [6–8], allowing to take images from the hand with a mobile devices. In fact, only the fingers are required to be captured by the system. The acquisition procedure consists of taking a picture of the hand open, with a mobile device or a low-quality camera.

The acquisition procedure emulate real situations with an unknown background, with no constraints regarding distance to mobile devices, hand orientation, and the like. The proposed features are invariant to these previous changes, being this independency essentially important for an accurate performance of the overall system.

Regarding classifier algorithms, this methods proposes linear support vector machines [9,10] to provide a decision about the identity of a hand image.

Finally, the layout of the paper is as follows: Section 2 presents the data used to validate the algorithms and details the acquisition procedure. Section 3 introduces the proposed approach for hand segmentation, indicating how the features are extracted in Sect. 5. The paper ends presenting the results obtained (Sect. 6) and the conclusions, together with possible future work (Sect. 7).

2 Database acquisition

This biometric method is oriented for mobile applications and therefore, the algorithm must be tested with images acquired from mobile devices. To our knowledge, there exist no public database satisfying our requirements, and therefore an owned database has been collected.

Fig. 1 Different examples of hand acquisitions. It is remarkable that there exist different illuminations, different sizes, orientation, scales, distance to camera, and similar aspects that provides this database with special characteristics and simulate proper standard environments



Such a database contains hand captures of 50 different individuals of an age range from 16 to 60 years old, gathering men and women in the same proportion containing also samples from different races. Therefore, it is a database with a huge variability in terms of size, skin color, and the like.

Furthermore, considering a contact-less approach for biometric hand recognition, every hand image was acquired without placing the hand in any platform neither requiring any removal of rings, bracelets, or watches. Instead, the individual was required to open his/her hand naturally, so that the mobile device (an Apple iPhone) could take a photo of the hand at 10–15 cm of distance.

This acquisition implies no control on illumination, background (behind the hand), hand orientation, and distance to mobile camera, and every acquisition was carried out under natural light. This approach combines several current challenges in hand biometric recognition with the limitation of mobile devices [11].

Both hands were taken, in a total of three sessions: during the first session, 10 acquisitions from both hands are collected; second session is carried out after 10–15 min, collecting this time 5 images per hand. Finally, third session is taken after three months, collecting 5 images per hand. Nowadays, the database is still under construction, and most users have only come through first two sessions.

In fact, the data involved in this document only considers first two sessions and only one hand per user.

The image size provided by the device is $1,600 \times 1,200$ pixels. However, the images were subsampled to obtain images of 400×300 pixels, aiming a realistic simulation of a mobile environment. Notice that the computational cost of using original images would rise extensively, making a

posterior mobile implementation an unfeasible goal. Furthermore, this subsampling is not considered within the process of hand recognition.

Figure 1 provides some examples of several acquisitions from the database. Notice how the environment is not under control, how the individuals did not remove rings, watches, and the like, and how orientation, hand openness, and distance to mobile camera is not fixed and remains variable for each acquisition.

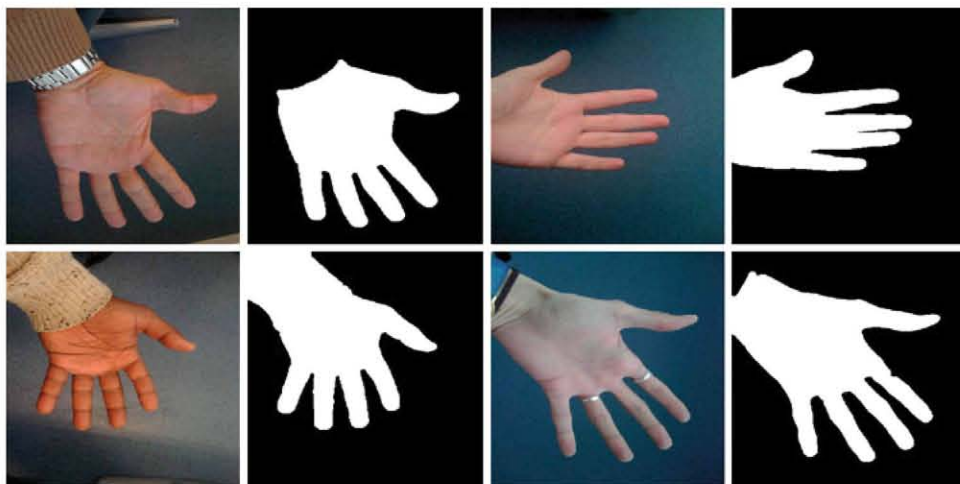
3 Segmentation

Each image is acquired without constraints in terms of illumination, background, and the like (Sect. 2). This characteristic provides the technique with an increment in the acceptance among individuals, based on its lack of invasiveness.

However, this fact represents a more demanding aspect in segmentation, since no previous knowledge is assumed regarding background [8], and as a consequence, isolating hand from background means more effort for the algorithm. In addition, considering mobile restrictions in terms of computational operations, a future implementation on portable devices means less capability in terms of image processing. In other words, the segmentation algorithm must be oriented to achieve good results combining time performance and segmentation accuracy.

The segmentation approach proposed in this document is based on a family of multiscale aggregation algorithms [12]. More in detail, the proposed approach focuses on a sort of algorithms within previous multiscale aggregation procedure, which is known as segmentation by weighted aggregation (SWA [13]). This former approach has provided

Fig. 2 Several examples of the segmentation procedure applied to the database. The result is certainly accurate, but in those cases where hand and some part of the arm is selected, posterior hand contour extraction will only consider the part corresponding to fingers, neglecting the rest



accurate results in segmentation and tested in different public databases.

Our approach models the image, namely I , as a graph $G = (V, E, W)$, where nodes $v_i \in V$ represent pixels in I , edges $e_{ij} \in E$ connect pixels v_i and v_j according to a defined neighborhood system (specifically, 4-connected structure) and weights w_{ij} are associated with each edge e_{ij} , indicating to what extent two nodes v_i and v_j are similar.

Standard mobile cameras provide images in RGB format, but for this application, best results were obtained using the Lab color space (abbreviation for CIE 1976 L^* , a^* , b^*) [14], concretely, considering image I as the layer b from the Lab space.

Furthermore, weights are defined according to Eq. 1,

$$w_{ij} = \exp\left(-\frac{(I_{v_i} - I_{v_j})^2}{2\sigma^2}\right) \quad (1)$$

where I_{v_i} and I_{v_j} represents the intensity values of image I at pixels v_i and v_j ; σ states to what extent pixels v_i and v_j are similar. This parameter has been set to $\sigma = 0.01$ in our database. As future work, it will be considered an automatic decision on this parameter based on features extracted from image I .

The algorithm works by gathering those pixels v_i and v_j with higher values in terms of weights w_{ij} . This operation leads to a number of segments, where each segment represents a group of pixels with similar properties in terms of texture, collecting the overall information contained within the segment. This texture information is gathered in terms of average values and intensity deviation, so that texture information is collected, as pixels are gathered in subsequent scales. Therefore, texture information is conformed along multiscale aggregation, as a consequence of gathering subgraphs with similar properties in terms of color.

Posterior steps consider this extracted information in order to group more segments, building a pyramid, where segments

are gathered under super-segments according to their metrics in terms of weights w_{ij} . This procedure is repeated until only two segments remain: one segment corresponding to the hand, the other corresponding to background.

Finally, Fig. 2 provides an idea on the performance of the segmentation algorithm.

4 Hand contour processing

This section describes in detail how to extract the required points from the hand (extremities and valleys) considering the contour extracted in previous Sect. 3 and providing an accurate result, since a posterior feature extraction step (Sect. 5) will strongly depend on this procedure and therefore the performance of the whole biometric system. Thus, a proper points extraction requires a special attention.

First of all, the aimed points must be defined more precisely. This procedure seeks to find nine points: four extremities (those corresponding to finger extrema) and five valleys (corresponding to fingers base). Figure 3 provides a visual example of the aimed points.

Many authors in literature [6–8] considered the idea of detecting extremities and valleys based on hand contour. However, this paper focuses on a reliable method that provides accurate results without knowledge on the orientation, scale, and changes in general due to the acquisition step.

This method first isolates fingers from palm by means of mathematical morphology [15], obtaining five regions corresponding to each finger. Opening morphological operator is proposed with a disk structural element of size 40, which conserves rounded palm shape, but erases those blobs corresponding to fingers. Therefore, fingers are obtained by deleting those pixels in common between original segmented image and the palm blob (obtained by the opening operator). In fact, some spurious blobs could appear after this operation, being erased by the application of another opening

Fig. 3 Visual results of the extrema detection procedure. Nine points are detected independently from position, orientation, and image size, provided the segmented hand region. In addition, points have been exaggerated in size for the sake of visibility, and they are represented by a small square

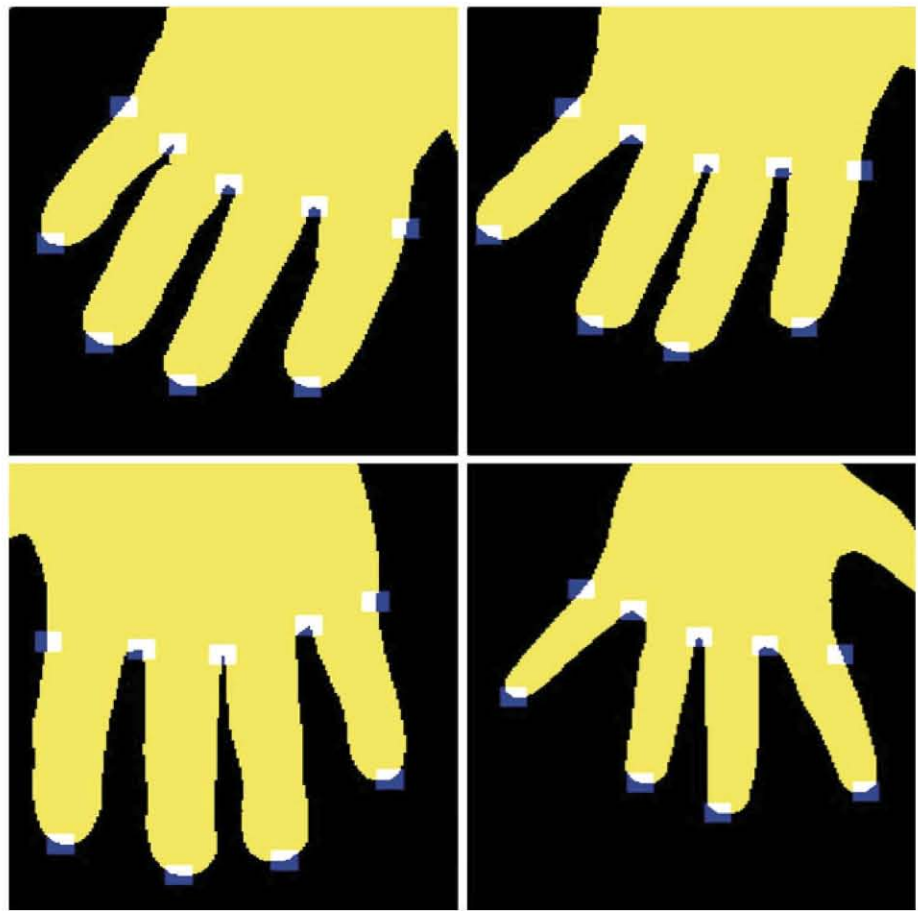
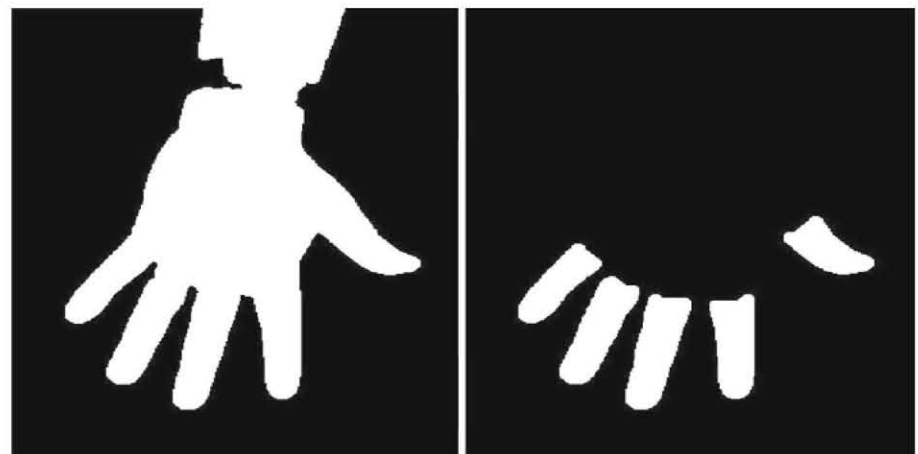


Fig. 4 Finger regions detection: image on the *left* represents the original image segmented. Image on the *right*, the detected regions. Notice how centroids associated with regions can be divided into two groups: thumb and rest of fingers, and the fact that little detection can be obtained based on thumb finger, since it is the most opposite finger to thumb



morphological operator with a small disk structural element (size 5). This set of operations would lead to the result presented in Fig. 4.

The centroid of each region, ξ_i , is obtained in order to cluster previous five regions into two groups: one group gathering four regions (index, heart, ring, and little); another group containing only the region corresponding to thumb. The motivation behind this operations relies on the fact that thumb is usually the most distant finger in relation to each other

fingers. Thus, thumb is detected and also little, since little centroid is opposite to thumb centroid in terms of distance.

Notice that a reasoning based on regions size would lead to false classifications, since not always is little the most little finger in the hand, but it is ensured that little is the most distant finger to thumb.

Calculating tips and valleys is straightforward. Regarding tips, the furthest point of each finger blob in relation to hand centroid is selected as the tip. This value selection is

invariant to rotation, since tip point is the furthest point of each finger to hand center. In addition, valleys are detected by considering the boundary hand contour region between two finger pairs and selecting the closest point to hand centroid.

This procedure obtains three valleys: index-heart, heart-ring, and ring-little valleys. In order to calculate the remaining two points of interest, the same idea is applied but considering the hand boundary between index and thumb and calculating the closest point to the index-heart valley. Similarly, the other point is calculated by considering the hand boundary between little and wrist and by obtaining the closest point to the ring-little valley.

These set of calculations provide the nine points of interest. These points are essential for a posterior feature extraction step (Sect. 5), and therefore identifying an individual precisely rests on an accurate points detection.

As an overview, the steps of the hand contour processing are stated as follows:

1. Palm Blob: Opening operation (disk, 40) to avoid fingers.
2. Fingers: Subtract palm blob result to original segmented image.
3. Avoid small regions in fingers by opening operation (disk, 5).
4. Calculate and identify fingers:
 - (a) Thumb calculation: Furthest centroid in comparison to fingers centroids.
 - (b) Little calculation: The most opposite to thumb.
 - (c) Calculate the rest fingers.
5. Calculate important points:
 - (a) Tip: Blob point furthest from hand centroid
 - (b) Valley: Closest point to hand centroid of the boundary between the corresponding tips.
 - (c) Lateral: Closest point to the corresponding valley point in relation to the opposite boundary.

5 Feature extraction

The biometric information contained within the hand must be reduced to few features so that a feasible comparison between two acquisition is possible. This features represent the biometric template. Considering the characteristics of the database (hands with different scales, sizes, orientations, and the like), feature extraction must be performed regardless any of the aforementioned characteristics.

Three different groups of distances are defined for each hand:

- First distance represents the length of the finger, from the tip of the finger (extrema) to the center of the line formed by the two corresponding inter-finger points (valleys)

lying at both sides of the given finger. The set of all finger lengths is represented by $\Lambda = \{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$, where each length λ_i corresponds to one finger (index, heart, ring, and little, respectively).

- Second distance consists of dividing the finger in m parts, measuring perpendicularly finger widths at each position m , with a separation of $\frac{\lambda_i}{m}$. This thickness measure is gathered under $\Theta_i = \{\theta_1, \dots, \theta_m\}$ for each finger.
- Finally, third distance depends on the finger tip curvature. Extracting curvature information from the extremity of the finger involves calculating the distance from the center of the line lying at a distance $\frac{\lambda_i}{m}$ of the tip of the finger to different points of the tip of the finger in a radial manner. Considering the tip of the finger as semicircular, these lengths are taken dividing the corresponding space into q parts, so that the variation of the angle will be given by $180^\circ/q$. Similarl to previous distances, $\Omega_i = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ gathers all third distances for each corresponding finger.

Reader may notice that the sets of features Λ , Θ , and Ω , despite being independent from orientation, are not invariant to scale or size. This invariance is easily obtained by dividing each component from Θ and Ω by the corresponding length in Λ .

Finally, each finger contains $m + q$ features, and thereby, each individual is represented by $4 \times (m + q)$ features, being this the unique template of a user.

In order to compare templates among individuals, this paper proposes SVM with linear kernel functions as a proper and accurate classifier, which has provided the best results when compared to other classifiers or kernel functions [8].

6 Results

Evaluating a biometric system has typically involved the use of the well-known parameter equal error rate (EER) [7], which provides information on the accuracy of a biometric system in identifying or verifying a given individual.

In order to validate the proposed approach, the acquisitions within the database were used according to a k -fold cross-validation scheme (Leave-one-out, concretely), since the samples in the database correspond to 50 individuals.

A modification of this validation technique was considered to assess to what extent the number of samples in training affects the final performance of the whole system. This study is of great importance since it indicates precisely how many samples are required to create a proper template for each individual. A compromise between samples during training and system performance must be achieved.

The data involved only consider acquisitions from first two sessions, remaining as future work how the system behaves

Table 1 Equal error rate (%) performance in relation to parameters m and q

	$m = 5$	$m = 10$	$m = 15$	$m = 20$
$q = 5$	7.3 ± 0.2	7.4 ± 0.1	8.2 ± 0.3	8.7 ± 0.2
$q = 10$	7.5 ± 0.3	6.0 ± 0.2	6.9 ± 0.2	7.3 ± 0.1
$q = 15$	7.3 ± 0.1	6.4 ± 0.3	6.3 ± 0.2	7.4 ± 0.3
$q = 20$	7.9 ± 0.2	7.2 ± 0.2	7.2 ± 0.2	7.5 ± 0.2

Notice that a greater number of features does not imply better results in terms of identification and verification

with data acquired several months (third session) after the training data.

Therefore, the system evaluation will consider two different aspects: first, the relation between accuracy in identifying individuals and feature parameters (m and q). Secondly, to what extent previous accuracy remains constant by decreasing the number of training samples.

Finally, this section presents a discussion on limitations regarding SVM when only one user is enrolled in the system (a mobile phone, for instance).

6.1 Performance evaluation in relation to feature parameters m and q

Considering feature extraction section, there exist two parameters of interest: m and q . The performance of the system will depend strongly on these parameters, and Table 1 will come up with the conclusion that high values of m and q do not improve performance. Furthermore, Table 1 provides that best performance of this system is achieved with $m = 10$ and $q = 10$, obtaining an EER of $6.0\% \pm 0.2$, which is a very promising result considering both the acquisition procedure and the biometric technique involved.

Furthermore, EER variation between first and second session is insignificant. In other words, the system is able to identify individuals independently of the time between enrollment and subsequent accesses.

6.2 Relation between EER, false acceptance rate, false rejection rate, and number of samples in training

Previous experiments were carried out considering $N - 1$ samples as training and 1 sample as testing. The database contains $N = 15$ samples per users, considering both first two sessions. Obviously, it is extensively uncomfortable for

individuals to make 14 acquisitions of their hands. The purpose of this section is to study whether that number of acquisitions can be reduced without losing significant accuracy.

Table 2 illustrates how a higher value of EER is obtained as the number of training samples decreases. Together with EER, the false acceptance rate (FAR) and false rejection rate (FRR) are provided as information on the false positives and false negative detection rates, respectively. Additional efforts must be considered as future work to deal with reducing the number of samples in training.

6.3 SVM limitations

Features are classified to their corresponding classes based on SVM, as stated before, and SVM requires in this scheme at least two users to make a feasible comparison. But, what would happen if there exist only one unique user on the system?

In order to tackle with this problem, a possible solution could provide synthetic/distorted data as real biometric data stored on the device. This solution will allow an individual to use the system normally, since when compared to synthetic/distorted data, the subject must be correctly identified (according to EER). In other words, the synthetic/distorted data will be considered as data extracted from a non-existing user.

Evidently, if a second user is enrolled in the system, the synthetic/distorted data can be ignored, and the system may continue working correctly.

In the case of a user not enrolled in the system, which is trying to verify himself/herself, the SVM identification system will indicate the hand in the database more similar to the user, with the corresponding matching percentage. It is easy to add a minimum threshold from which the system can conclude that the hand belongs to the user in the database. This threshold study and its implementation remains as future work.

Finally, all these experiments have been carried out on a PC Computer @ 1.86 GHz, implemented in MATLAB 7 R14. A future implementation on a mobile device remains as future work.

7 Conclusions and future work

This paper proposes a biometric system based on finger geometry, oriented to a mobile device environment, where

Table 2 Relation between EER, FAR, FRR, and training samples

Training samples	1	3	5	7	10	12	14
EER	12.5 ± 0.3	10.3 ± 0.1	9.5 ± 0.3	8.1 ± 0.2	7.6 ± 0.2	6.4 ± 0.1	6.0 ± 0.2
FAR	0.002	0.001	0.003	0.001	0.002	0.002	0.003
FRR	24.3 ± 0.2	20.7 ± 0.1	18.2 ± 0.1	16.4 ± 0.3	15 ± 0.1	13 ± 0.2	12 ± 0.1

Notice how the EER increases in relation to a reduction in number of training samples

the acquisition procedure is uncontrolled, and hand images can present differences in scale, size, and orientation. A very important characteristic of this proposed technique relies on not requiring any platform as to where to locate the hand, being completely invasive. Furthermore, the images are required to contain only the fingers of the individual, making easier the acquisition of the biometric data.

The segmentation procedure offers the possibility of good accuracy combined with low computational cost, which means that this algorithm is suitable for mobile implementation. Moreover, the feature extraction procedure provides with accurate results to build the template, gathering invariant features to these previous factors with the benefit of being easily extracted and computed.

The method proposed based on SVM with linear kernel functions allows a performance in EER of $6.0\% \pm 0.2$, which is a promising result considering the involved constraints: images acquired, features extracted, and the fact that only fingers are considered. Furthermore, the image size affects also negatively to the results (400×300 pixels), but a deeper understanding of the relation between size and EER deserves more future work. An interesting goal is to decrease also the number of training sample without having effects on the overall accuracy. In order to compare performance rates with those approaches in literature, authors would like to apply these algorithms to other public databases, so that a fair comparison with other approaches could be carried out.

Many different scenarios may benefit from this biometric technique like mobile phones, laptops, and any device with a camera embedded. Concretely, considering mobile phones, notice that a given individual can make a picture of his/her hand by him or herself and that will be enough to access the mobile device.

On the other hand, an implementation oriented to mobile devices, finishing third session from database and increasing the number of individuals remains as a desirable future work, together with increasing the accuracy of the system in identifying individuals. Furthermore, authors would like to research the independence of the technique regarding mobile devices, in other words enrolling and accessing the system with different mobile devices. In addition, it is of interest to study the effect of geometry modifications due to changes in temperature or any external factor, provoking for instance, thinning or thickening. Their effect will be studied according to standard normalization SC37 Biometrics committee.

Finally, since the hand images have been acquired with the palm up, a possible fusion scheme with palmprint biometrics could be of interest as future research lines.

References

1. Bolle, R., Pankanti, S.: *Biometrics, Personal Identification in Networked Society: Personal Identification in Networked Society*. Kluwer Academic Publishers, Norewell, MA (1998)
2. Arif, M., Brouard, T., Vincent, N.: Personal identification and verification by hand recognition. In: *IEEE International Conference on Engineering of Intelligent Systems*, pp. 1–6 (2006)
3. Su, C.L.: Overlapped finger geometry signal processing and finger shape comparisons for person identification. *Informatica* **18**, 447–456 (2007)
4. Kang, B.J., Park, K.R.: Multimodal biometric authentication based on the fusion of finger vein and finger geometry. *Opt. Eng.* **48**, 090501 (2009)
5. Malassiotis, S., Aifanti, N., Strintzis, M.: Personal authentication using 3-d finger geometry. *IEEE Trans. Inf. Forensics Secur.* **1**, 12–21 (2006)
6. Yoruk, E., Konukoglu, E., Sankur, B., Darbon, J.: Shape-based hand recognition. *IEEE Trans. Image Process.* **15**, 1803–1815 (2006)
7. de Santos Sierra, A., Casanova, J., Avila, C., Vera, V.: Silhouette-based hand recognition on mobile devices. In: *43rd Annual 2009 International Carnahan Conference on Security Technology*, pp. 160–166 (2009)
8. Morales, A., Ferrer, M., Alonso, J., Travieso, C.: Comparing infrared and visible illumination for contactless hand based biometric scheme. In: *42nd Annual IEEE International Carnahan Conference on Security Technology (ICCST)*, pp. 191–197 (2008)
9. Cortes, C., Vapnik, V.: Support-vector networks. In: *Machine Learning*, Springer, New York, pp. 273–297 (1995)
10. Cristianini, N., Shawe-Taylor, J.: *An introduction to support vector machines: and other kernel-based learning methods*. 1st edn. Cambridge University Press, Cambridge (2000)
11. Shahin, M., Badawi, A., Rasmy, M.: A multimodal hand vein, hand geometry, and fingerprint prototype design for high security biometrics. In: *Cairo International Biomedical Engineering Conference (CIBEC)*, pp. 1–6 (2008)
12. Alpert, S., Galun, M., Basri, R., Brandt, A.: Image segmentation by probabilistic bottom-up aggregation and cue integration. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–8 (2007)
13. Son, T., Mita, S., Takeuchi, A.: Road detection using segmentation by weighted aggregation based on visual information and a posteriori probability of road regions. In: *IEEE International Conference on Systems, Man and Cybernetics (SMC)*, pp. 3018–3025 (2008)
14. Hunter, R.S.: Photoelectric color difference meter. *J. Opt. Soc. Am.* **48**, 985–993 (1958)
15. Gonzalez, R.C., Woods, R.E.: *Digital Image Processing*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA (2001)