

RESEARCH

Open Access

Feature exploration for biometric recognition using millimetre wave body images



Ester Gonzalez-Sosa*, Ruben Vera-Rodriguez, Julian Fierrez, Miriam Moreno-Moreno and Javier Ortega-Garcia

Abstract

The use of millimetre wave images has been proposed recently in the biometric field to overcome certain limitations when using images acquired at visible frequencies. Furthermore, the security community has started using millimetre wave screening scanners in order to detect concealed objects. We believe we can exploit the use of these devices by incorporating biometric functionalities. This paper proposes a biometric recognition system based on the information of the silhouette of the human body, which may be seen as a type of soft biometric trait. To this aim, we report experimental results on the BIOGIGA database with four feature extraction approaches (contour coordinates, shape contexts, Fourier descriptors and landmarks) and three classification methods (Euclidean distance, dynamic time warping and support vector machines). The best configuration of 1.33 % EER is achieved when using contour coordinates with dynamic time warping.

Keywords: Millimetre-waves; Body descriptors; Security applications; Screening scanners; Comparison, Dtw

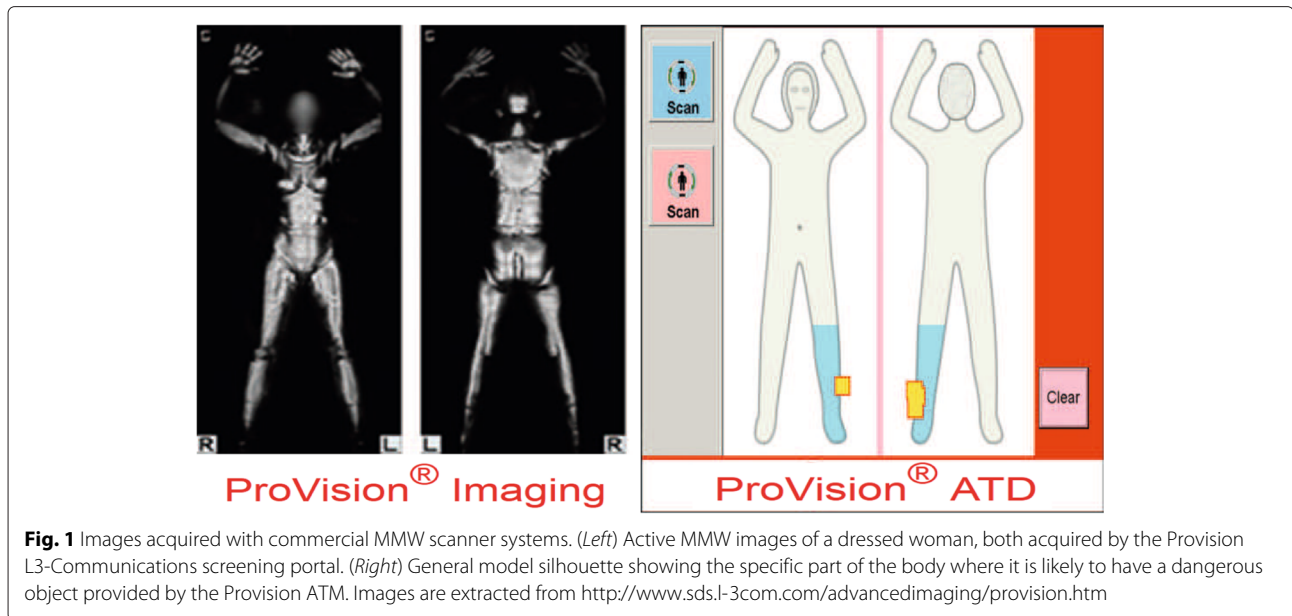
1 Introduction

Many biometric characteristics are used to identify individuals: fingerprint, iris, voice, face, hand, signature etc. The majority of these biometric traits are acquired with cameras working at visible frequencies of the electromagnetic spectrum. Such images are affected, among other factors, by lighting conditions and body occlusions (e.g. clothing, make up, hair etc.). To overcome these limitations, researchers have proposed the use of images acquired at other spectral ranges: X-ray, infrared, millimetre (MMW) and submillimetre (SMW) waves [1]. The images were captured beyond the visible spectrum circumvent and, to some extent, some of the mentioned limitations; furthermore, they are more robust to spoofing than other biometric images/traits [2].

Among the spectral bands out of the visible spectrum, the millimetre waves (with frequency in the band of 30–300 GHz) present interesting properties that can be exploited in biometrics: ability to pass through clothing and other occlusions, innocuous to health, low intrusiveness and the recent deployment and rapid progress of GHz-THz systems in screening applications.

One example of these GHz-THz systems for screening applications are the millimetre wave scanners deployed in several airports such as Los Angeles International Airport, San Francisco International airport in US or Schiphol Airport, Fiumicino Airport, in Europe among others. The MMW scanners have been replacing X-ray scanners throughout the years alleging that since this range of the spectrum is not ionizing, it is therefore less harmful to the health of human beings. These scanners may be implemented in active or passive mode, depending on whether they introduce artificial radiation into the system or not. Another important issue to bear in mind are privacy concerns as these systems can pass through clothes. To minimize privacy issues, operators are usually restricted to use generic silhouettes showing the area of the body where a potential dangerous object may be concealed rather than using real MMW images [3]. Figure 1 shows examples of the output of two MMW active scanners from L3 Communications. The left part of the image shows a view of the output image provided by the ProVision scanner in which it is possible to see the real images from the person. The right part of the image shows the output provided by the ProVision ATD (automatic target detection) scanner in which only the silhouette of the body is depicted. The latter scanner is accepted in both US and European airports.

*Correspondence: ester.gonzalez@siam.es
Department of Electronics and Communications Technology, Madrid, Avenida Francisco Tomas y Valiente 11, Madrid, Spain



Apart from the functionality of detecting dangerous objects, we propose to use also these devices to apply biometric recognition using the body shape as biometric trait. We reckon it may not substitute primary biometrics already deployed but it may be useful for narrowing the search of possible suspects with very little effort.

In this sense, we propose a biometric system that is capable of recognizing people by using the information of shape or appearance contained in MMW images. There have been little works on this field. Specifically, just one working with real data [4], and some others based on BIOGIGA database, which is a synthetic database [5–8].

This shortage of biometric recognition research based on MMW images is mainly due to the lack of databases of images of people acquired at 94 GHz. This lack is a consequence of (i) the privacy concerns these images present and (ii) most of the imaging systems working at the MMW/SMW band are either in prototype form or not easily accessible for research.

In [4], Alefs et al. proposed a holistic recognition approach based on the texture information of the MMW images. Concretely, they exploited the texture information contained in the torso region of the image through multi-linear eigenspace techniques. Likewise, they also analysed the discrimination capability of the face region, evaluating also the fusion between torso and face but it turned out that the best performance of the system was achieved when using only the torso information. On the other hand, the works by Moreno-Moreno et al. [1, 5, 6] put forward a biometric system based on geometric measures between different silhouette landmarks of the contour. It must be acknowledged that since this work used synthetic MMW

images, there was no point in using texture information of any part within the silhouette.

We believe the approach developed by Moreno-Moreno et al. would not be robust enough when applied to real MMW images. This is because the proposed feature extraction technique is highly dependent on the accuracy of a set of landmarks from the body silhouette. The reliability of these landmarks is acceptable with the BIOGIGA images but would drop heavily in real-world images. This hypothesis was also discussed by Alefs et al. in [4], in which they argue that landmarks in millimetre-wave imaging are less robust and have lower location accuracy.

The latter observation motivates us to search for an alternative source of information such as the whole contour of the silhouette, more robust to characterize in noisy images compared to pre-defined landmarks. We may treat the set of contour coordinates as a kind of soft biometrics [9, 10]. Although normally soft biometrics are not discriminative enough to build a biometric system by themselves, they may aid a biometric system by either helping to reach a better decision by fusing hard and soft information or narrowing down the search by allowing the system to only compare with those user models that match with the soft biometrics [11]. Furthermore, Dantcheva et al. [9] bring to the fore that, under certain conditions, a biometric system may be composed of only a vector of different soft biometric traits.

There are already some previous works using the whole contour information of the human silhouette for person recognition [7, 8]. In the preliminary work presented in [7], we developed some baseline techniques using contour coordinates for feature extraction and Euclidean distance

and dynamic time warping for classification. The following work [8] explored some other contour coordinate-based feature extraction such as Fourier descriptors and landmarks.

We are aware that, when using real MMW images fed by commercial scanners, it will be likely to encounter people with different body positions (scanning position) that could worsen the performance of the contour coordinate-based system. However, there are already literature proposing approaches to overcome this type of problems [12, 13]. In [12], a technique is proposed for pose-invariant representation of objects. This is achieved by applying multi-dimensional scaling (MDS) to a set of vertices of the mesh representation of the object. In [13], a preprocessing stage is put forward in order to register all silhouettes of hand images to a fixed pose enabling therefore feature approaches like independent component analysis.

In this paper, we extend the previous works [7, 8] by comparing multiple shape descriptors of the body contour and classifiers. We introduce an additional classification method, support vector machines, completely different from the other matching approaches. Our aim here consists of assessing all these features regarding performance, computational time and robustness.

Figure 2 draws a simple diagram explaining the whole biometric system we develop in this work. As can be seen, there are three principal stages: the contour extraction stage, the feature extraction stage and the comparison stage. Given two millimetre images, first contours are extracted; then a chosen feature approach is computed for each contour and finally a similarity measure between this pair of features is obtained. In the final stage, the output score is thresholded to decide whether this pair of subjects belongs to the same identity or not.

This paper is structured as follows. Section 2 further comment related works. The database and the procedure carried out to obtain the contours of people are explained

in Section 3. Section 4 describes the different feature extraction and classification approaches used to compare the contours. The evaluation of these methods is performed in Section 5, and conclusions are finally drawn in Section 6.

2 Related work on shape-based recognition

There are previous works that have used the shape or appearance of the body to recognize subjects. Please note that body shape-based recognition techniques are enclosed within the wide area of object-shape-based recognition [14, 15].

The lower part of the silhouette of the body is commonly used in gait-based biometric recognition systems where signals are extracted from video sequences of people walking. Such is the case of the work in [16], where they fused information of the gait biometric trait with some shape cues such as body weight, width and some body part proportions. Then in [17, 18], a multimodal system based on footsteps and gait was built. Likewise, in [19], a spatial temporal analysis of the lower part of the human silhouette was used to build a gait recognition system. In most cases, the silhouette of the people is extracted through background subtraction techniques.

There are also examples in which the silhouette of the person is used solely to recognize subjects. In [20], a system detected human silhouettes through background subtraction and modeled the appearance of the individual based on its colour and its spatial distance. They divided the silhouette into three different blobs and incorporated the path length measure, which is the distance from the top of the head to a given point on the path. Likewise in [21], they proposed a method to re-identify a subject seen in the field of one camera who reappears in another camera's field. They extracted a spectral classification of the appearance for each person. Then, they propose a new feature based on colour-position histogram in order to characterize the silhouette in static images.

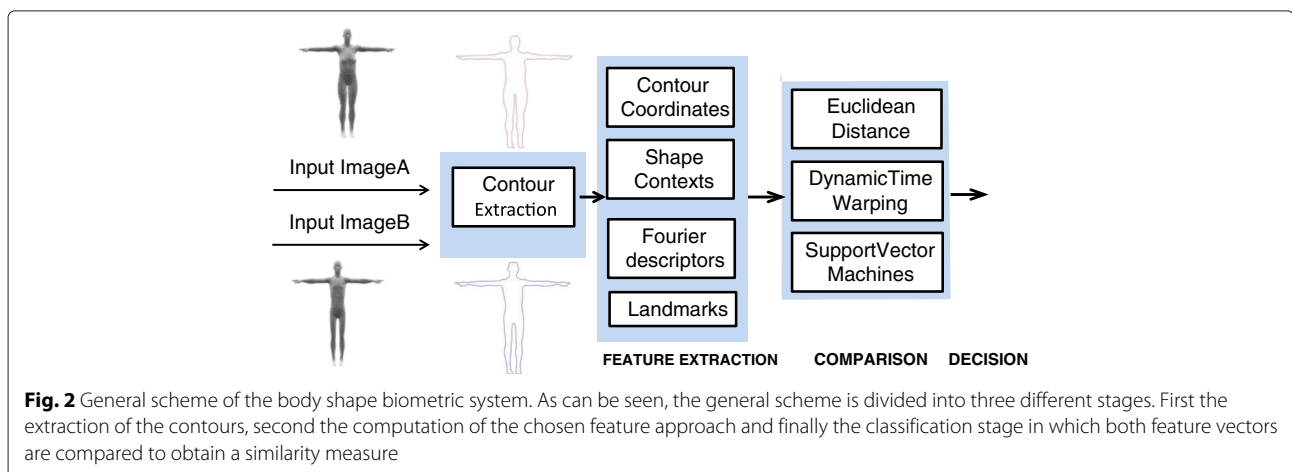


Fig. 2 General scheme of the body shape biometric system. As can be seen, the general scheme is divided into three different stages. First the extraction of the contours, second the computation of the chosen feature approach and finally the classification stage in which both feature vectors are compared to obtain a similarity measure

There are other works in which the silhouette information is not used to recognize subjects but with different motivation. In [22], the silhouette of the body is used to analyse human movements through Fourier descriptors. In [23], they were able to recover 3D human configurations through multiple 2D images in conjunction with the information provided by shape context descriptors [24].

Regarding shape-based recognition applied to other biometrics traits, we may find examples in hand [13, 25–27] and signature biometrics [28, 29].

3 Database and contour extraction

Due to the unavailability of real MMW image databases, in this paper, we use a synthetic database called BIOGIGA. The corpus of the BIOGIGA database consists of synthetic images at 94 GHz of the body of 50 individuals (25 males and 25 females). BIOGIGA images are the result of simulations carried out on corporal models at two types of scenarios (outdoors and indoors) with two kinds of imaging systems (passive and active). These corporal models are previously generated using the software MakeHuman¹ based on body measurements taken from real subjects. Then, these models are imported to Blender², which simulates the effect of the 94-GHz radiation over the human models. A more detailed description of the generation of the BIOGIGA database can be found in [5, 6, 30].

In this paper, only passive images at outdoor scenarios are considered similarly as the previous work using a real database [4]. This subset of the database is comprised of 50 subjects, with 6 images per user. Three of them are simulated with clothes, and the other three are simulated without clothes to analyse the effect of clothing and have some variability between images from the same person.

Pose rotation is also considered in the images, having two images with 10° , two images with 0° and two images with -10° . Figure 3 shows some images from a single subject of the database. As can be seen, images with and without clothes are very similar as the 94-GHz band is almost transparent to clothes; however, the pixel intensity is a bit darker in the images with clothes and small parts of the clothes are still noticeable in the waist and neck. The pose rotation is also observed.

Regarding the contour extraction, the first processing step is to binarize the images, separating the background from the body. A characteristic of the images simulated by passive systems is the different grey level presented in different parts of the body. For instance, the abdomen is much darker than the feet. This fact hinders the segmentation process and hence the binarization. This problem is overcome by performing the segmentation in two steps: (i) border detection, and (ii) morphological operations.

A Canny border detector is first applied to the image. The parameters of this detector have been empirically tuned (0.0005 for the low threshold, 0.10 for the high threshold and 2.5 for the standard deviation of the Gaussian filter). Then, the image is divided into four different bands due to the difference of intensity level between them: head, arms, from arms to calves and feet (see Fig. 4 left). After that, the closing operation is applied to each band to join any part of the silhouette that remains open after the Canny border detector stage. The closing operation is a type of morphological operation which implies the consecutive use of the dilatation and erosion with a defined structural element. Different structural elements have been used for the different band (disk for the upper bands and rectangular for the lower

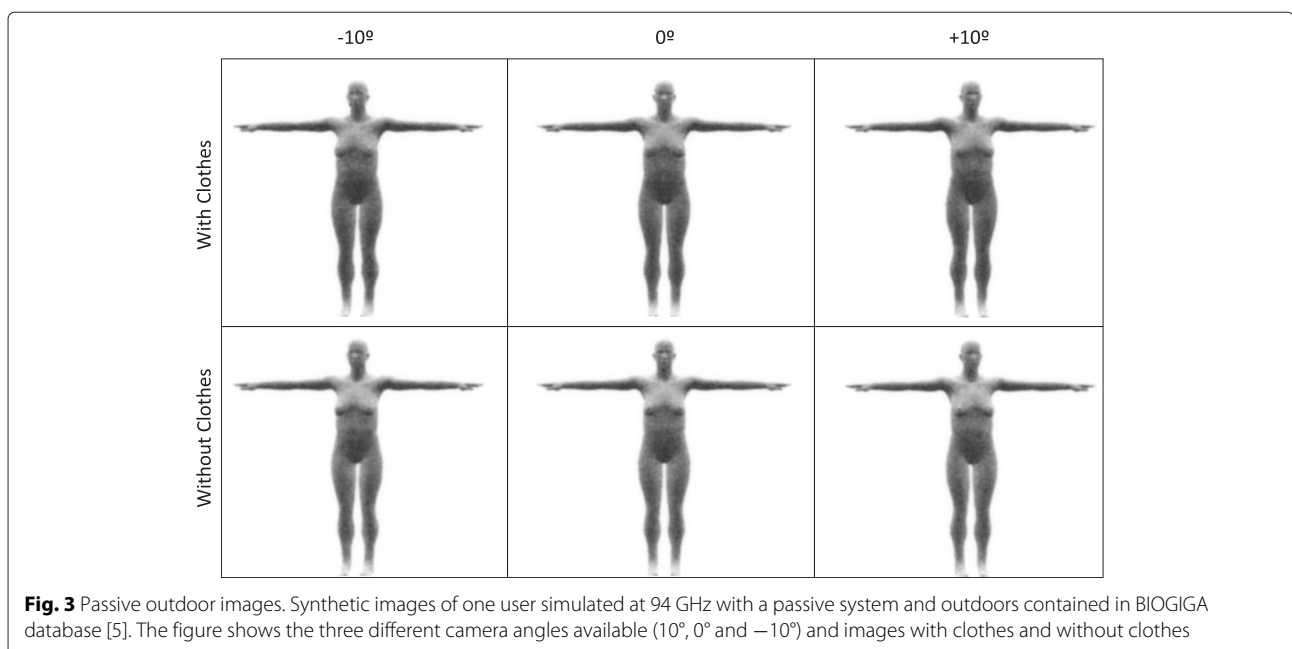


Fig. 3 Passive outdoor images. Synthetic images of one user simulated at 94 GHz with a passive system and outdoors contained in BIOGIGA database [5]. The figure shows the three different camera angles available (10° , 0° and -10°) and images with clothes and without clothes

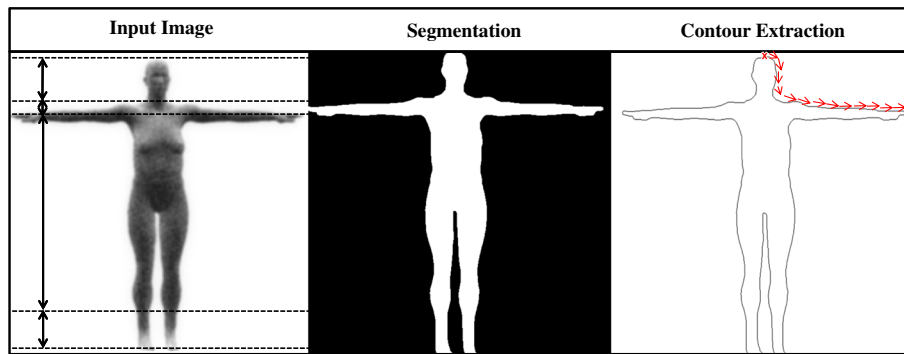


Fig. 4 Contour images. Main steps followed in our system to extract the contour. From left to right: original image, binarized image (through Canny detector and morphological operations applied to the different bands as depicted in the original image) and contour extraction

bands). Once the closing operation is applied to each band, we concatenate the resulting bands to create the improved silhouette. Finally, another set of morphological closing operations removes spurious irregularities to this improved silhouette and leads to the final contour of the human body, which is used in the following experimental sections. Figure 4 shows an example of the process of segmentation and contour extraction for user 1.

4 Shape-based body comparison

In what follows, we aim to describe the last two stages of the scheme drawn in Fig. 2. Within the feature extraction and the classification stage, several techniques may be applied. In this section, we proceed to outline every technique proposed for both the feature extraction and the classification stage.

4.1 Shape descriptors

We have selected four different approaches for the feature extraction stage: (i) contour coordinates themselves, (ii) shape contexts [24], (iii) Fourier descriptors of the coordinates [31] and (iv) silhouette landmarks [6].

4.1.1 Contour coordinates (CC)

Contour coordinates are used as the baseline feature approach. Once the silhouette is computed, the contour

(the external boundary of the silhouette) is extracted starting from the upper middle point of the head in clockwise direction (see Fig. 4 right). By concatenating the x and y coordinate of every single point of the contour following the aforementioned order, we obtain a $2 \times N$ matrix which describes the contour of the subject, being N the number of points of the contour.

The original resolution of the contours (N) extracted from the MMW images is approximately 2800 points. Through subsampling techniques, different contour resolutions ranging from 100 points up to 2800 points are used and analysed in the following sections.

4.1.2 Shape contexts (SC)

Shape context descriptors were first introduced by Belongie et al. [24]. This technique characterizes each point in the shape considering the relative distance and angle to the rest of the points of the shape.

The basic idea of shape contexts is illustrated in Fig. 5, which shows an example of a shape context descriptor for two points in the eight digit shape. Note that the log-polar histogram used in this case with 12×5 dimensions, where 12 accounts for the number of angular bins and 5 accounts for the number of radial bins. Dark colours mean a high density of points within a bin, while lighter colours imply less density of points. In both cases, the majority

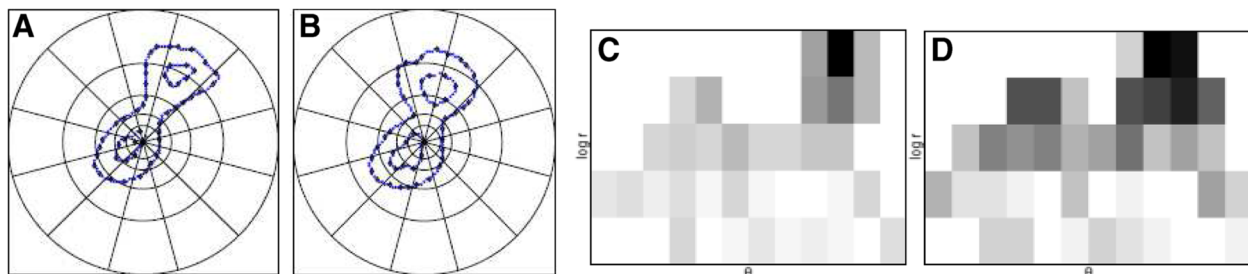


Fig. 5 Shape context example. Example of the computation of a shape context descriptor for two single points within the eight digit shape. Image (a) and (c) represent a point within the eight digit shape and its respective log-polar histogram while image (b) and (d) a different point within the same digit and its associated log-polar histogram. Images extracted from [35]

of the shape points are quite distant from the point being characterized.

To obtain the shape context descriptor of the contour, we need to compute the shape context descriptor for every single point of the contour. This means that every single point of the contour is no longer described by its x and y coordinate but by a $(N \times r_bins \times \theta_bins)$ vector, which gives an account of the distribution of the remaining points of the contour with respect to that particular one. As a result, the shape context descriptor of a shape with N points form a vector of size $N \times r_bins \times \theta_bins$. We decide to use the same configuration of parameters the author originally proposed, this is 5 r_bins and 12 θ_bins .

In order to compute the similarity between two shape contexts, different distance methods may be applied.

4.1.3 Fourier descriptors (FD)

Although Fourier descriptors [31] are a 40-year-old technique, they are still considered as a good description tool [15].

These descriptors are simple to compute and robust against translations and rotations. To apply this technique to our system, first contour coordinates are converted into complex numbers. Secondly, we apply the discrete Fourier transform to these complex numbers to obtain the Fourier descriptors.

With the resulting DFT, we may recover the original trajectory with varying precision depending on the number of Fourier coefficients used. If, instead of using all Fourier descriptors, we use only the first P coefficients, we will have an approximation of the contour. Bearing in mind that high-frequency components account for fine detail and low-frequency components determine the global shape, keeping only the first P coefficients will smooth the shape.

4.1.4 Landmarks (LM)

The last feature approach implies the use of landmark points along the contour. These landmarks consist of a reduced set of key points, obtained automatically as in [6]. Figure 6 shows an example of the situation of these 14 points. In particular, they mark the most singular parts of the human silhouette, among them: head, neck, hands, underarms, waist, hip, pubis and feet. Each landmark is characterized by its position coordinates (x and y).

In this work, we evaluate the results obtained with landmarks as features and compare them with the results achieved with the other approaches. Note that the dimensionality of these features is much smaller compared to the other approaches.

4.1.5 Landmarks with shape contexts (LM-SC)

Although this last approach is not a new descriptor, we believe the combination of these two approaches may

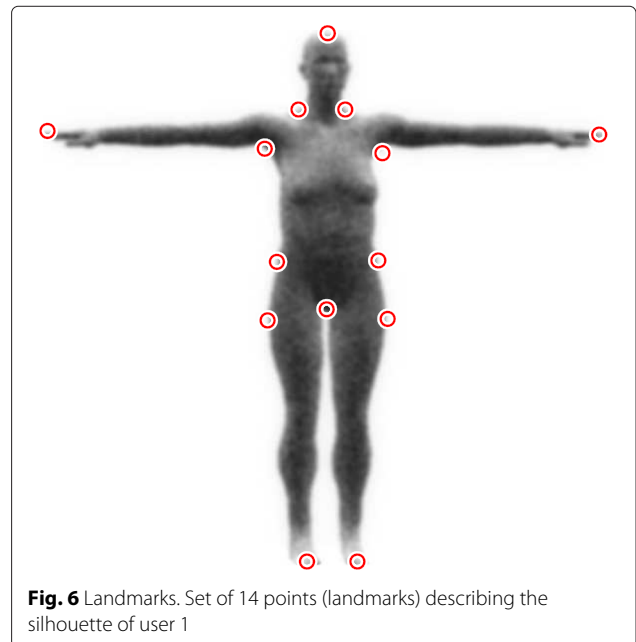


Fig. 6 Landmarks. Set of 14 points (landmarks) describing the silhouette of user 1

lead to a better understanding of the silhouette while keeping a reasonable dimensionality of the feature vector. For this aim, we compute the shape context descriptor to each of the 14 landmarks. In this way, we describe the silhouette not through the set of 14 landmarks but through a histogram describing the context of each point of the contour.

4.2 Similarity computation

Regarding the classification stage, the Euclidean distance (ED) and the dynamic time warping (DTW) algorithm are proposed as similarity distances and, as a proper classifier, support vector machines (SVM) are used.

4.2.1 Baseline technique: Euclidean distance (ED)

This simple approach consists in computing a dissimilarity measure between the contour coordinates of two silhouette images. The only restriction of this method is the fact that distances need to be computed between sequences of the same length. Therefore, a normalization of the length of the sequences must be applied. This normalization is achieved by interpolating or truncating each sequence to the average length of all sequences. Then, the Euclidean distance between the two normalized contours is computed.

4.2.2 Dynamic programming: dynamic time warping (DTW)

The goal of DTW is to find an elastic match among samples of a pair of sequences.

In this work, DTW is used to obtain the optimal alignment between two sequences of points that min-

imizes the cumulative distance between them. This is achieved through non-linear mapping and certain constraints between these two sequences.

These points may be defined (i) by the coordinates x and y in the case of contour coordinates or (ii) by 60-dimensional vectors of shape contexts. The DTW algorithm is the same in both cases. In each iteration, the algorithm computes the Euclidean distance between a point from the first sequence and a point from the second sequence. The resulting DTW distance d is finally transformed to a similarity measure with an exponential normalization: $score = \exp(-\frac{d}{K})$, where K is a normalization factor that takes into account the number of aligned points between the sequences.

4.2.3 Support vector machines (SVM)

Regarding the SVM configurations, it must be noted that experiments are developed only for the protocol 3:1, in which three images for training are used.

Besides, as the feature vector for the support vector machines must be of fixed length, some operations are

first applied to the features in order to meet this requirement. Concretely, for contour coordinates, we first apply PCA independently to each dimension (x and y). Then, we select the first 30 components of each projected vector and finally we concatenate them obtaining a vector of 60 components. In the case of shape contexts, we subsampled the contour down to 50 points and then obtained the shape contexts features for this subset, obtaining a vector of 300 components. In what concerns Fourier descriptors, contours are previously normalized to zero mean and then, once computed the Fourier descriptors, the performance of the system using different number of frequency components was assessed and we chose the 20 low frequency components. Lastly, for landmarks and shape contexts over landmarks, we only concatenate all points in a single vector (14-vector for landmarks and 840-vector for landmarks over shape contexts).

We empirically prove that a polynomial kernel of grade 3 improves with respect to other kernels for all feature approaches.

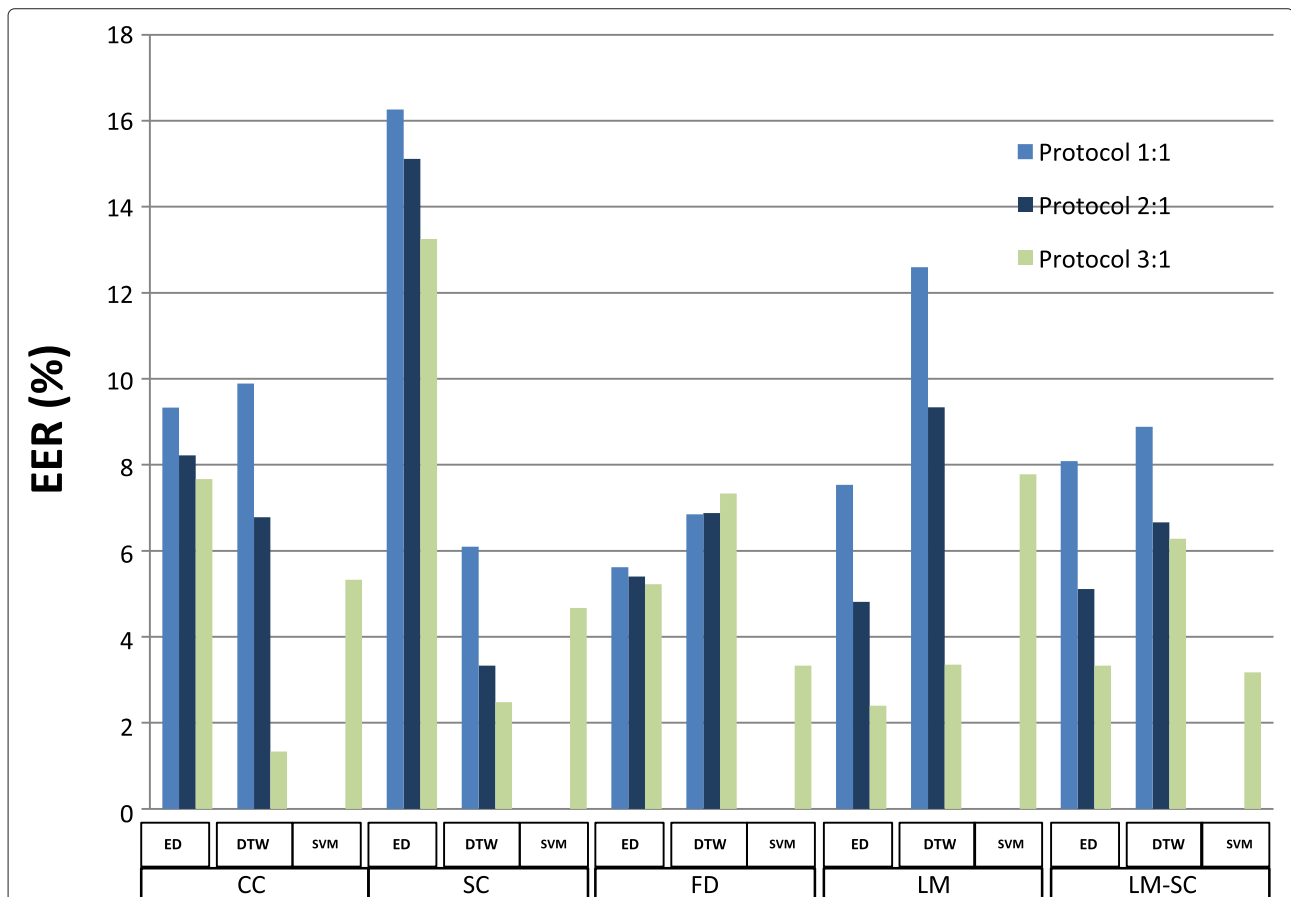


Fig. 7 Results. Performance in terms of % EER of the 15 different approaches for 2800 point contours and for protocols 1:1, 2:1 and 3:1. CC contour coordinates, SC shape contexts, FD Fourier descriptors, LM landmarks, LM-SC shape contexts over landmarks, ED Euclidean distance, DTW dynamic time warping and SVM support vector machine. SVM approaches are only tested with protocol 3:1

5 Experiments

This section describes the experimental work carried out to analyse the performance of the different feature and classification approaches described in Section 4. The aforementioned methods are tested with the contour coordinates of the BIOGIGA database previously described in Section 3.

In this work, three different experimental protocols are considered: (i) protocol 1:1 (P1:1), (ii) protocol 2:1 (P2:1) and (iii) protocol 3:1 (P3:1), where the first number refers to the number of training images considered per user, and the second number to the number of test images per user (one in all cases). In order to have the most challenging scenario with severe mismatch between enrolment and testing regarding clothes, the database is divided into two sets, where the images with clothes are used for the training set and the images without clothes are used for the test set. Both training and test images present variability in pose rotation; each set has images with 10° , 0° and -10° .

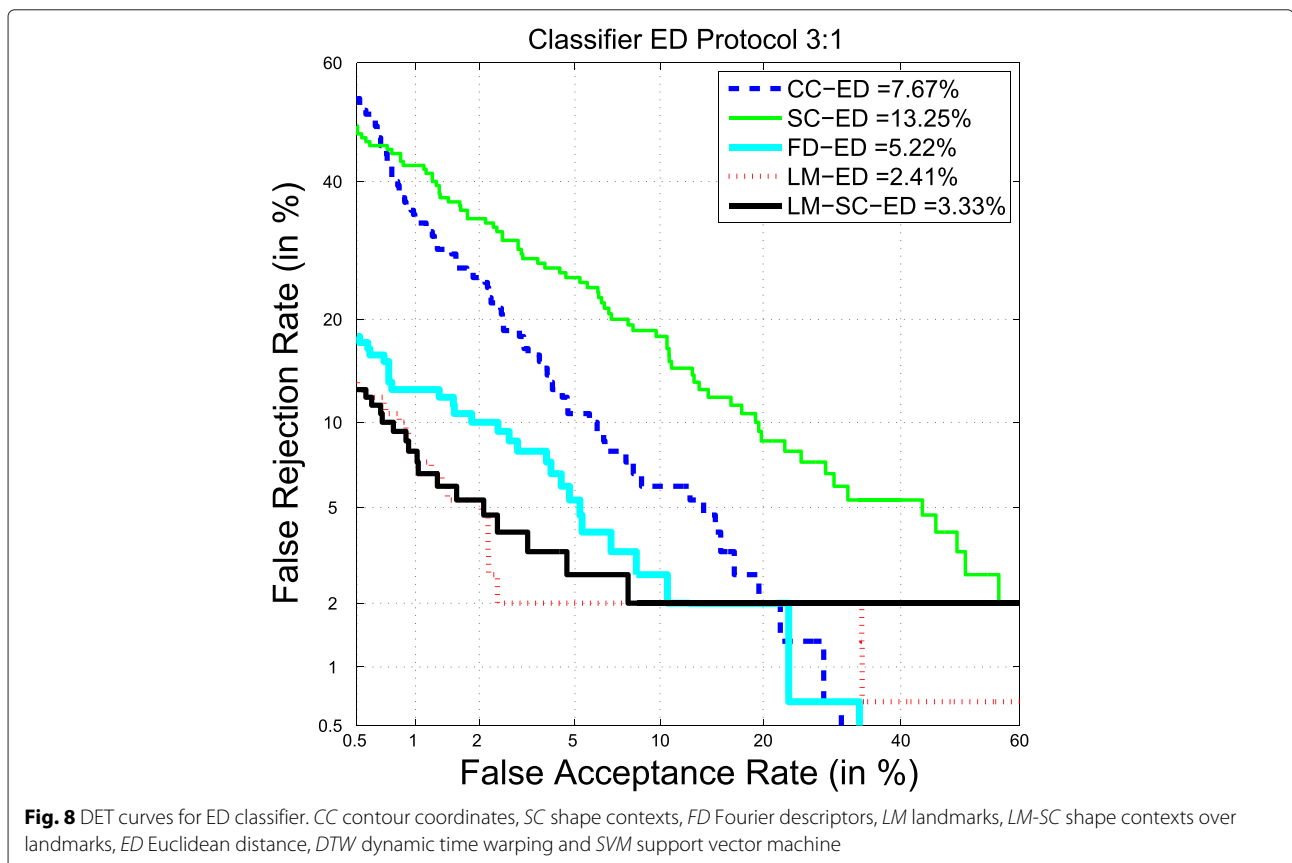
When having 2 or 3 images for training, the fusion of the information contained in the images is carried out at the score level, i.e. all single comparisons between training and test are done image by image, and then the scores are fused using the sum rule [32].

As mentioned in Section 4, experiments based on DTW are analysed with all contours having their original size. ED experiments are carried out with contours normalized to the same size.

5.1 Results

The first experiment carried out compares the performance of the different approaches. Bearing in mind that there are five different feature extraction approaches and three different classification approaches, we have 15 possible system configurations. Figure 7a shows the performance in terms of equal error rate (EER) for each of these approaches and the three protocols considered (P1:1, P2:1 and P3:1).

First, from Fig. 7 we observe that the EER of the system decreases as the number of training images increases. It is also worth noting the outstanding improvement of performance when applying DTW algorithm to the contour coordinates and shape context approaches (especially in P3:1) compared to the baseline Euclidean distance. Applying DTW to the Fourier descriptors does not result in better performance compared to ED since these transformed features are already resampled to have the same dimension. In the case of landmarks or shape contexts over landmark feature approaches, it is not worth applying



DTW to a feature vector of such a small dimensionality and with a fixed number of points. Regarding the average performance between all protocols, the best approaches are CC-DTW (6 %), LM-SC-ED (5.50 %), FD-ED (5.41 %), LM-ED (4.91 %) and SC-DTW (3.96 %). The best individual result is achieved for CC-DTW and for P3:1, obtaining a 1.33 % of EER. However, considering the other two protocols, it can be seen that the performance of the CC-DTW approach is worse compared to other cases. For example, the SC-DTW produces lower EER rates for protocols 1:1 and 2:1. For the case of FD-ED approach, we observe that we obtain similar EER regardless of the number of training images so it can be deduced that although FD-ED does not achieve the best results, as expected, this approach is robust against variations in pose. Regarding the LM-ED approach, we see that using a vector with a dimensionality quite smaller than the vectors used in any of the previous approaches, produces comparable results to the best approaches. However, as was already mentioned in Section 1, in a real database, we believe the localization of these keypoints would not be as robust as in this synthetic database.

From Fig. 7, we also conclude that support vector machines do not improve the best EER obtained for each

feature approach with P3:1 except the FD and the LM-SC approaches. Concretely, we obtain a very promising performance of 3.33 % for the first 20 low frequency components when using SVM (using the complete set of Fourier descriptors does not help to obtain low error rates). Likewise, it seems that with the LM-SC SVM configuration, we can obtain reasonable results (3.17 % of EER). This latter approach may overcome the downsides of the landmark features when using real MMW images. Bearing in mind that neighbour points should have similar shape context descriptors, we may relax the landmark accuracy and use the shape contexts over these landmarks instead.

Figures 8, 9 and 10 plot the different DET curves, analysing each classifier separately: Fig. 8 for the Euclidean distance classifier; Fig. 9 for the DTW algorithm and Fig. 10 for the support vector machines. Comparing Figs. 8 and 9, we can clearly notice the improvement of CC and SC approaches when applying the DTW algorithm. In summary, we conclude that the usage of ED is suitable for features such as LM. DTW achieves the best performance for CC and SC descriptors, whereas FD and LM-SC performs well in conjunction with the SVM classifier.

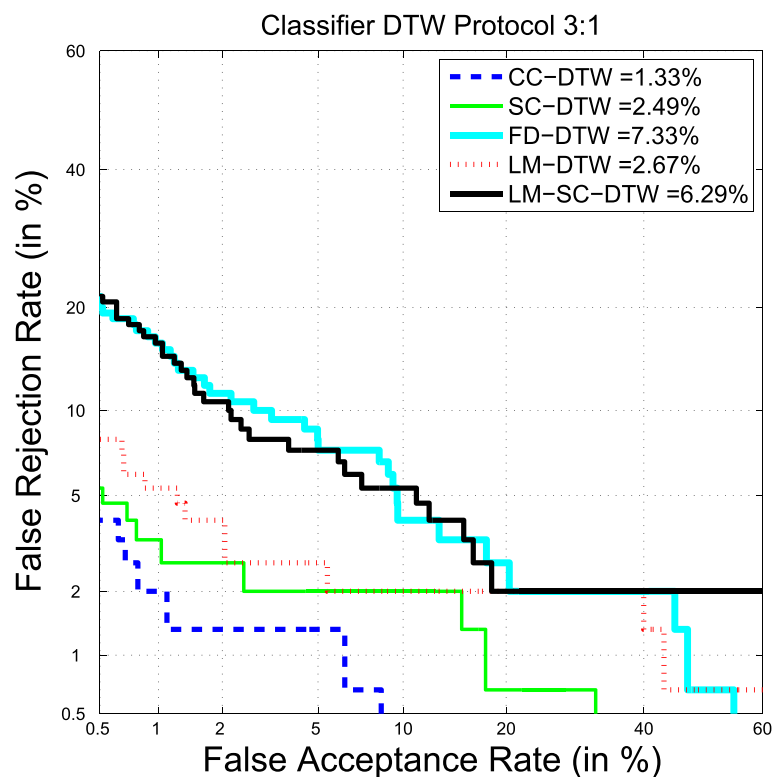


Fig. 9 DET curves for DTW classifier. CC contour coordinates, SC shape contexts, FD Fourier descriptors, LM landmarks, LM-SC shape contexts over landmarks, ED Euclidean distance, DTW dynamic time warping and SVM support vector machine

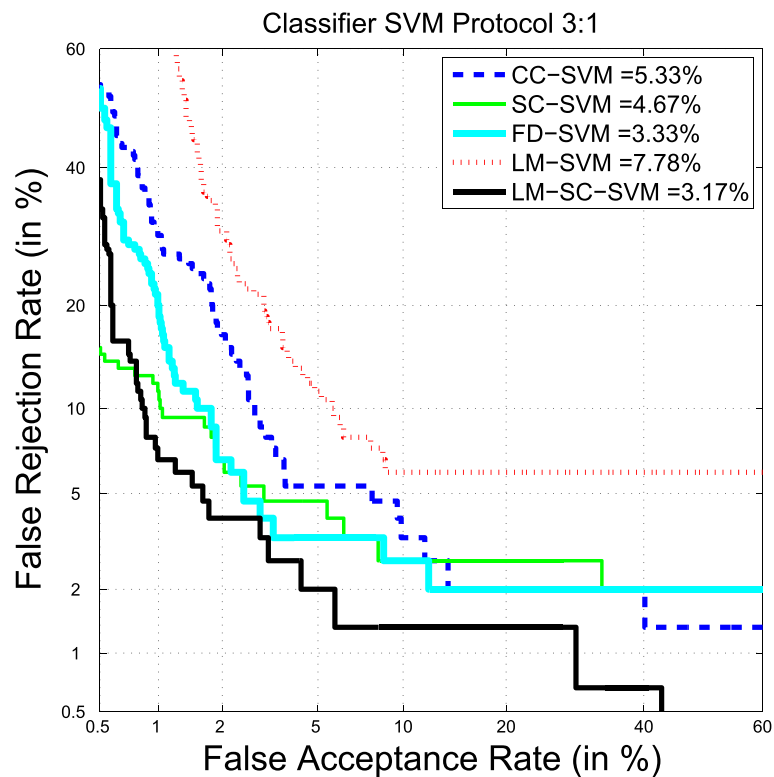


Fig. 10 DET curves for SVM classifier. *CC* contour coordinates, *SC* shape contexts, *FD* Fourier descriptors, *LM* landmarks, *LM-SC* shape contexts over landmarks, *ED* Euclidean distance, *DTW* dynamic time warping and *SVM* support vector machine

We have also assessed the computation cost of each approach (Intel i7-3770 CPU @ 3.4 GHz RAM 8 GB in Matlab R2012b). Figure 11 depicts the time in seconds spent during all the process, taking into account the time needed to compute the features and the time to compare a pair of feature vectors. The main conclusion we can extract from this time comparison is the fact that DTW algorithm implies an increment of the computational time mainly for high dimensional feature vectors like CC, SC and FD. This fact is magnified when dealing with shape context features, mainly caused by the larger dimensionality of this vector.

The computational time of the best approaches (CC-DTW and LM-SC-SVM) concerning robustness and performance is in both cases under 1.5 s, amount of time that would be feasible within the screening scanner system.

5.2 Effect of contour resolution

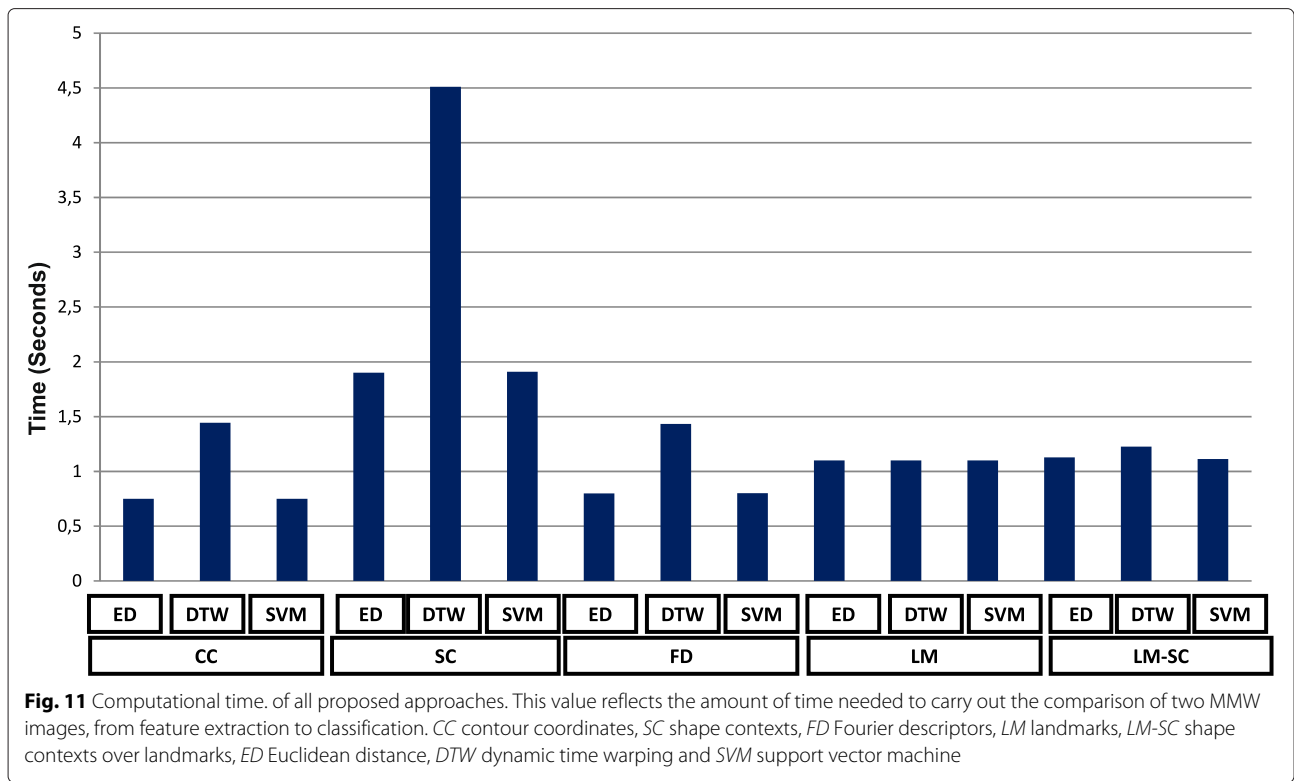
A second experiment is carried out to analyse the effect of the contour size in the recognition performance, for the cases of contour coordinates and shape contexts using DTW. Figure 12a, b represents the performance of the system against the resolution of the contour. Considering the case of CC-DTW, it is very interesting to note a notable

drop of EER when dealing with a contour resolution of more than 500 points. Even though the EER obtained with the largest resolution (2800 points) is slightly better than the EER obtained with 500 points, there is a big difference in terms of computational time between using a 2800-CC-DTW approach rather than a 500-CC-DTW one. Concretely, the computational time drops from 1.48 to 0.81 s when reducing the resolution of the contour from 2800 points down to 500 points. This is an important issue to bear in mind for real-time applications.

Conversely, when using shape context descriptors, the EER drops as the resolution of the contour increases but there is no clear knee point as in the previous case. In this case, we need to find a tradeoff between a suitable EER and a reasonable amount of time.

6 Conclusions

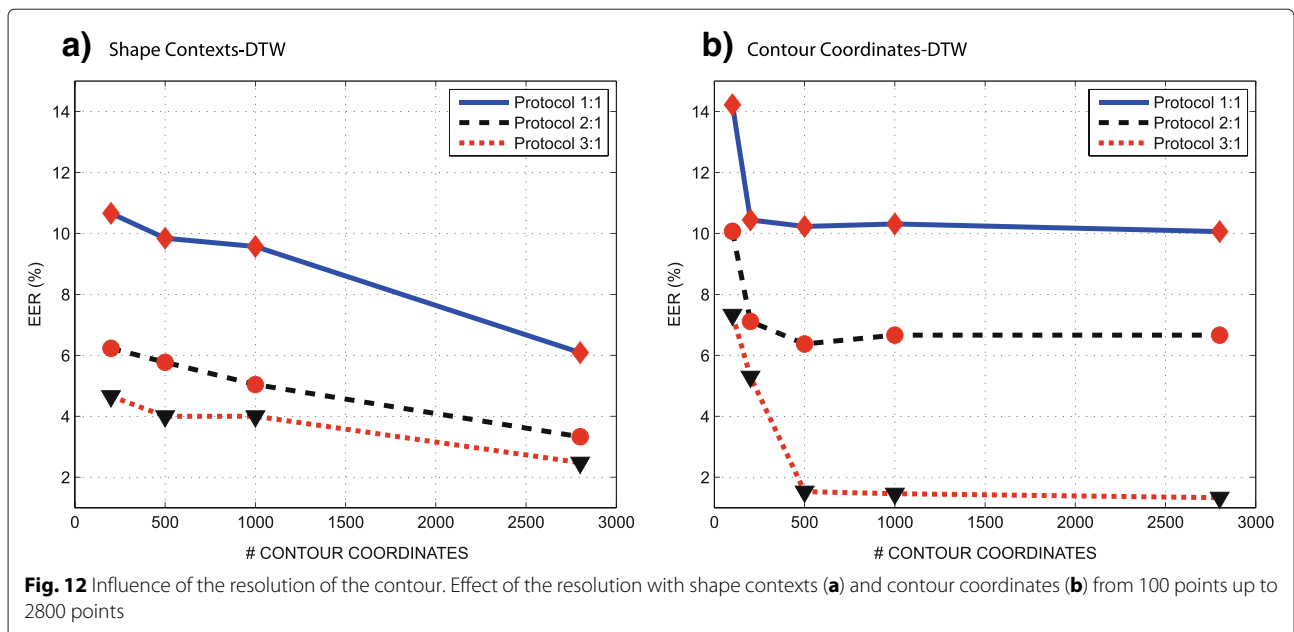
In this paper, a complete body shape biometric system has been developed for MMW body images using the BIOGIGA database. The use of MMW images instead of images acquired at other spectral bands presents some advantages, mainly the transparency of clothing at that frequency, allowing extracting easily the contours from the images. Different approaches have been



analysed ranging from naive approaches such as contour coordinates for the feature extraction stage or the Euclidean distance for the classification stage to complex schemes such as shape contexts or Fourier descriptors for the feature extraction stage or the dynamic time warping algorithm and support vector machines for the classification stage.

The best result is obtained when using the DTW algorithm directly to the coordinates for the contours (CC-DTW) with the highest resolution for protocol 3:1 (1.33 % EER).

However, when working with images extracted from real MMW sensors in which the contour extraction stage may be more difficult due to the presence of noise, illumination



variation and so forth, we believe approaches such as SC-LM may be more robust than the CC as SC-LM approach uses richer information.

It is known that fusing complementary information would lead to a better performance of the system. To this aim, we propose for future work the fusion of some of the feature descriptors explained in this work or even fusing information with the previous system [5] based on geometrical distances between landmarks. Also, future research will explore this system using real MMW images.

Besides, for future work, it would be interesting to assess the influence of using an image-to-class distance instead of the image-to-image distance that we have used in the experiments. There exists an interesting approach named image to class dynamic time warping (I2C-DTW) [33, 34] that manages to find an optimal warping path between an image and a class considering both the time dimension and the within-class dimension.

Endnotes

¹<http://makehuman.org/>

²<http://blender.org/>

Abbreviations

DFT: discrete Fourier transform; sDTW: dynamic time warping; ED: Euclidean distance; EER: equal error rate; FD: Fourier descriptors; GHz: gigahertz; LM: landmarks; MMW: millimetre waves; PCA: principal component analysis; SC: shape contexts; SVM: support vector machines.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

In this paper, an experimental comparison of different feature extraction and classification methods applied to the task of biometric recognition of persons using millimetre wave images is carried out. We assess the variety of approaches concerning performance, robustness and computational time. Furthermore, an update literature revision on both body shape recognition and biometric recognition using millimetre wave images, including commercial developments, is also presented. All authors read and approved the final manuscript.

Acknowledgements

This work has been partially supported by projects TeraSense (CSD2008-00068), Bio-Shield (TEC2012-34881) and BEAT (FP7-SEC-284989) from EU. E. Gonzalez-Sosa is supported by a PhD scholarship from Universidad Autonoma de Madrid.

Received: 1 November 2014 Accepted: 12 August 2015

Published online: 02 September 2015

References

1. M Moreno-Moreno, J Fierrez, J Ortega-Garcia, in *Biometric ID Management and Multimodal Communication*, LNCS. Biometrics beyond the visible spectrum: imaging technologies and applications, vol. 5707 (Springer, 2009), pp. 154–161
2. A Hadid, N Evans, S Marcel, J Fierrez, in *IEEE Signal Processing Magazine, Special Issue on Biometric Security and Privacy*. Biometrics systems under spoofing attack: an evaluation methodology and lessons learned (IEEE Xplore, September 2015)
3. DM Sheen, DL McMakin, TE Hall, Three-dimensional millimeter-wave approaches for concealed weapon detection. *IEEE Trans. Microwave Theory Tech.* **49**(9), 1581–1592 (2001)
4. B Alefs, R den Hollander, F Nennie, E van der Houwen, M Bruijn, W van der Mark, J Noordam, Thorax biometrics from millimetre-wave images. *Pattern Recogn. Lett.* **31**(15), 2357–2363 (2010)
5. M Moreno-Moreno, J Fierrez, R Vera-Rodriguez, J Parron, in *Biometric Technologies for Human Identification, Proc. of SPIE*. Simulation of millimeter wave body images and its application to biometric recognition, vol. 8362 (SPIE, 2012)
6. M Moreno-Moreno, J Fierrez, R Vera-Rodriguez, J Parron, in *Proc. IEEE Intl. Carnahan Conf. on Security Technology, ICCST*. Distance-based feature extraction for biometric recognition of millimeter wave body images (IEEE Xplore, 2011), pp. 1–6
7. E Gonzalez-Sosa, R Vera-Rodriguez, J Fierrez, J Ortega-Garcia, in *Proc. IEEE Intl. Carnahan Conf. on Security Technology, ICCST*. Body shape-based biometric recognition using millimeter wave images (IEEE Xplore, 2013)
8. E Gonzalez-Sosa, R Vera-Rodriguez, J Fierrez, J Ortega-Garcia, in *Proc. International Conference on Pattern Recognition, ICPR*. Comparison of body shape descriptors for biometric recognition using MMW images (IEEE Xplore, 2014)
9. A Dantcheva, C Velardo, A D'angelo, J-L Dugelay, Bag of soft biometrics for person identification. *Multimedia Tools Appl.* **51**(2), 739–777 (2011)
10. D Reid, S Samangoeei, C Chen, M Nixon, A Ross, Soft biometrics for surveillance: an overview. *Machine learning: theory and applications*. Elsevier, 327–352 (2013)
11. P Tome, J Fierrez, R Vera-Rodriguez, M Nixon, Soft biometrics and their application in person recognition at a distance. *IEEE Trans. Inf. Forensic Secur.* **9**(3), 464–475 (2014)
12. S Katz, G Leifman, A Tal, Mesh segmentation using feature point and core extraction. *Vis. Comput.* **21**(8-10), 649–658 (2005)
13. E Yoruk, E Konukoglu, B Sankur, J Darbon, Shape-based hand recognition. *IEEE Trans. Image Process.* **15**(7), 1803–1815 (2006)
14. D Zhang, G Lu, Review of shape representation and description techniques. *Pattern Recognit.* **37**(1), 1–19 (2004)
15. M Yang, K Kpalma, J Ronsin, et al., *A survey of shape feature extraction techniques*. (Peng-Yeng Yin, ed.) (Pattern Recognit. IN-TECH, 2008), pp. 43–90. <https://hal.archives-ouvertes.fr/hal-00446037/document>
16. RT Collins, R Gross, J Shi, in *Proc. Fifth IEEE International Conference on Automatic Face and Gesture Recognition*. Silhouette-based human identification from body shape and gait (IEEE Xplore, 2002), pp. 366–371
17. R Vera-Rodriguez, J Fierrez, JSD Mason, J Ortega-Garcia, in *Biometrics (ICB), 2013 International Conference on*. A novel approach of gait recognition through fusion with footstep information (IEEE Xplore, June 2013), pp. 1–6. doi:10.1109/ICB.2013.6613014
18. R Vera-Rodriguez, JSD Mason, J Fierrez, J Ortega-Garcia, Comparative analysis and fusion of spatio-temporal information for footstep recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(4), 823–834 (2012)
19. L Wang, T Tan, H Ning, W Hu, Silhouette analysis-based gait recognition for human identification. *IEEE Trans. Pattern Anal. Mach. Intell.* **25**(12), 1505–1518 (2003)
20. K Yoon, D Harwood, L Davis, Appearance-based person recognition using color/path-length profile. *J. Vis. Commun. Image Represent.* **17**(3), 605–622 (2006)
21. D-N Truong Cong, L Khoudour, C Achard, C Meurie, O Lezoray, People re-identification by spectral classification of silhouettes. *J. Signal Process.* **90**(8), 2362–2374 (2010)
22. Z Ling, C Zhao, Q Pan, Y Wang, Y Cheng, in *Proc. IEEE International Conference on Automation and Logistics*. Analyzing human movements from silhouettes via fourier descriptor, (2007), pp. 231–236
23. G Mori, J Malik, Recovering 3D human body configurations using shape contexts. *IEEE Trans. Pattern Anal. Mach. Intell.* **28**(7), 1052–1062 (2006)
24. S Belongie, J Malik, J Puzicha, Shape matching and object recognition using shape contexts. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(4), 509–522 (2002)
25. J Burgues, J Fierrez, D Ramos, J Ortega-Garcia, in *Biometric ID Management and Multimodal Communication*. LNCS. Comparison of distance-based features for hand geometry authentication, vol. 5707 (Springer, 2009), pp. 325–332
26. A Morales, E González, MA Ferrer, On the feasibility of interoperable schemes in hand biometrics. *Sensors.* **12**(2), 1352–1382 (2012)
27. H Dutağacı, E Yörük, B Sankur, Comparative analysis of global hand appearance-based person recognition. *J. Electronic Imaging.* **17**(1) (2008).

<http://electronicimaging.spiedigitallibrary.org/article.aspx?articleid=1099857>

28. J Fierrez, J Ortega-García, *On-line signature verification*. (AK Jain, A Ross, P Flynn, eds.) (*Handbook of Biometrics*, Springer, 2008), pp. 189–209
29. M Martínez-Díaz, J Fierrez, RP Krish, J Galbally, Mobile signature verification: feature robustness and performance comparison. *IET Biometrics*. **3**(4), 267–277 (2014)
30. M Moreno-Moreno, J Fierrez, P Tome, R Vera-Rodríguez, J Parron, J Ortega-García, in *Proc. of XXVI Simposium Nacional de Union Científica Internacional de Radio, URSI 2011*. Biogiga: Base de datos de imágenes sintéticas de personas a 94 GHz con fines biométricos (Madrid, Spain, September 2011). <http://atvs.ii.uam.es/files/ursi2011.pdf>
31. E Persoon, K-S Fu, Shape discrimination using fourier descriptors. *IEEE Trans. Syst. Man Cybernet.* **7**(3), 170–179 (1977)
32. J Fierrez-Aguilar, *Adapted fusion schemes for multimodal biometric authentication*. (PhD thesis, Universidad Politécnica de Madrid, 2006)
33. H Cheng, Z Dai, Z Liu, in *IEEE International Conference on Multimedia and Expo (ICME)*. Image-to-class dynamic time warping for 3d hand gesture recognition (IEEE Xplore, 2013), pp. 1–6
34. X Wei, C-T Li, Z Lei, D Yi, SZ Li, Dynamic image-to-class warping for occluded face recognition. *IEEE Trans. Inf. Forensic. Secur.* **9**(12), 2035–2050 (2014)
35. H Zhang, J Malik, in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Learning a discriminative classifier using shape context distances, vol. 1 (IEEE Xplore, 2003), p. 242

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
