Security Technology (ICCST), 2013 47th International Carnahan Conference on. IEEE, 2013

DOI: http://dx.doi.org/10.1109/CCST.2013.6922076

Body Shape-based Biometric Recognition using Millimeter Wave Images

E. Gonzalez-Sosa, R. Vera-Rodriguez, J. Fierrez and J. Ortega-Garcia
Biometric Recognition Group - ATVS, EPS, Universidad Autonoma de Madrid
Avda. Francisco Tomas y Valiente, 11 - Campus de Cantoblanco - 28049 Madrid, Spain
ester.gonzalezs@estudiante.uam.es, {ruben.vera, julian.fierrez, javier.ortega}@uam.es

***Abstract*—The use of MMW images has been proposed recently in the biometric field aiming to overcome certain limitations when using images acquired at visible frequencies. In this paper, several body shape-based techniques are applied to model the silhouette of images of people acquired at 94 GHz. Three main approaches are presented: a baseline system based on the Euclidean distance, a dynamic programming method and a procedure using Shape Contexts descriptors. Results show that the dynamic time warping algorithm achieves the best results regarding the system performance (around 1.3% EER) and the computation cost. Results achieved here are also compared to previous works based on the extraction of geometric measures between several key points of the body contour. An average relative improvement of 33% EER is achieved for the work reported here.**

I. Introduction

Many biometric characteristics are used to identify individuals: fingerprint, signature, iris, voice, face, hand, etc. The majority of these biometric traits are acquired with cameras working at visible frequencies of the electromagnetic spectrum. Such images are affected by, among others factors, lighting conditions and the body occlusion (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, millimeter (MMW) and submillimeter (SMW) waves [1]. The images captured beyond the visible spectrum circumvent, to some extent, some of the mentioned limitations; furthermore, they are more robust to spoofing than other biometric images/traits.

Among the spectral bands out of the visible spectrum, the millimeter 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.

In spite of the previous advantages, to date, there are just a few works on this field. Specifically, just one working with real data [2], and some others based on BIOGIGA database, which is a synthetic database [3], [4]. In [2], Alefs *et al.* proposed a holistic recognition approach based on the texture information of the MMW images. On the other hand, the works by Moreno-Moreno *et al.* proposed a biometric system based on geometric measures between different key points of the contour. This shortage of biometric recognition research based on MMW images is due, in part, to the lack of databases of images of people acquired at GHz. This lack is a consequence of: *i)* the privacy concern 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 this paper, we propose a novel approach for biometric recognition based on the comparison of body contours extracted from images at 94 GHz. This is inspired by previous works, which show that recognition through the shape and boundary of traits such as hand or signature are fairly reliable [5], [6]. In this work, three approaches based on the body shape are considered: *i)* a baseline technique based on the Euclidean distance, *ii)* a programming dynamic technique based on the dynamic time warping algorith (DTW) and *iii)* a shape contexts descriptor. Several experiments are carried out to determine the performance and behavior of these different approaches. Also, three experimental protocols are followed varying the quantity of training data for different contour sizes. Finally, a comparison with the previous works based on geometric measures is also carried out showing an average relative improvement of performance of 33% EER.

This paper is structured as follows. The database, and the procedure carried out to obtain the contours of people is explained in section II. Section III describes the different methods used in this paper to model the contours. The evaluation of these methods is performed in section IV, and conclusions are finally drawn in section V.

II. Database and Contour Extraction

The corpus of the BIOGIGA database consists of synthetic images at 94 GHz of the body of 50 individuals (25 males and 25 females). The images are the result of simulations carried out on corporal models at two types of scenarios (outdoors and indoors) and with two kinds of imaging systems (passive and active). These corporal models were previously generated using the software MakeHuman[1] based on body measurements taken from the subjects. Then, these models were imported to Blender[2], 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 [3].

In this paper, only passive images at outdoor scenarios are considered. This subset of the database is comprised of 50 subjects, with 6 images per user. Three of them were simulated with clothes, and the other three were simulated without clothes to analyse the effect of clothing and have some variability between the images of the same person. For

[1] http://makehuman.org/
[2] http://blender.org/

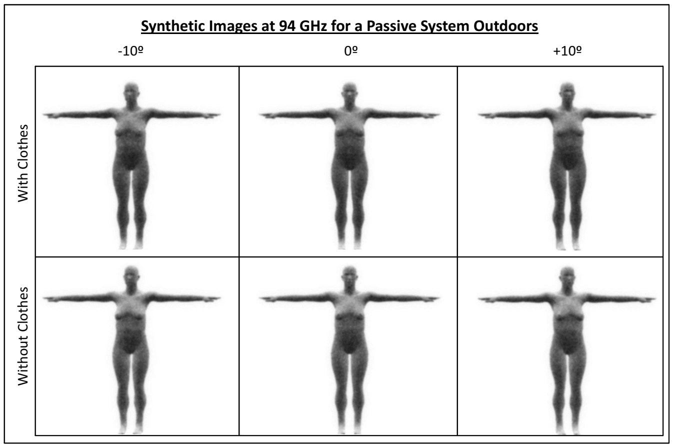


Fig. 1. Synthetic images of one user simulated at 94 GHz with a passive system and outdoors. The figure shows the three different camera angles, and images with clothes and without clothes

this, three angles between the subject and the camera were considered, having images with -10, 0 and +10 degrees. Figure 1 shows an example of the 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 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 aim of this paper is to develop a biometric system based on the contour of the body silhouettes. Therefore, the first step was to binarize the images, separating the background from the body. A characteristic of the images simulated by passive systems is the different grey level they present in different parts of the body. For instance the abdomen is much darker than the feet. This fact hinders the segmentation process. This problem was overcome performing the segmentation in two steps: *i)* border detection and *ii)* morphological detection.

A Canny border detector (whose parameters are previously tuned) is first applied to the image. After that, various morphological operations are conducted on the resulting border image. These morphological operations consist of closing operations with different structural elements in different areas of the image (head, arms, from arms to calf, and feet). Finally, another set of morphological closing removes spurious irregularities, and obtains the final contour of the human body, which is used in the following experimental sections. Figure 2 shows an example of the process of segmentation and contour extraction.

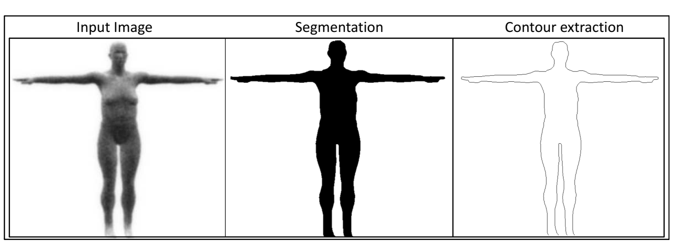


Fig. 2. Main steps followed in our system to extract the contour. From left to right: Original image (of a subject with clothing and a camera angle of +10 degrees), segmented image, contour extraction.

III. Body Shape-based approaches

This section describes the approaches based on the body shape followed in order to build the biometric system. In this case, the biometric information used is based on the x and y coordinates of the body contour. In this work we study and compare three different approaches. The first one is based on computing the euclidean distance (ED) between two silhouettes, the second one is focused on a dynamic programming technique and the third one is based on shape contexts descriptors.

A. Baseline Technique: Euclidean Distance (ED)

This naive approach consists of 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 is applied. Then, the euclidean distance between the two normalized contours is computed following equation 1,

$$Dist\_l2 = \sum_{j=1}^{N} \sum_{i=1}^{2} \sqrt{((a_{ij} - b_{ij}))^2} \tag{1}$$

where a, b represent the sequence of contour coordinates of images a and b respectively; $i = 1, 2$ represents the number of coordinates describing each contour point, in this case: x and y (row and column), and $j = 1, ..., N$ defines every point of a contour, assuming that every contour is characterized by N points.

B. Dynamic Programming: Dynamic Warping Technique (DTW)

Dynamic Time Warping (DTW) algorithm is a dynamic programming technique that was first presented by Yashuhara *et al.* [7]. The goal of DTW is to find an elastic match among samples of a pair of sequences that minimize a given distance measure. In the biometric field, it was first used for signature verification [8], [6].

In this work, DTW is used to obtain a cumulative distance between two strings of coordinates, that is known to be minimal. Equation 2 shows the transformation of this minimal distance into a matching score where K is a normalization factor that takes into account the number of aligned points between two sequences.

$$score = e^{-\frac{DIST}{K}} \tag{2}$$

C. Shape Contexts Descriptors (SC)

Shape Contexts descriptors were first introduced by Belongie *et al.* [9]. This technique addresses to describe the feasibility of a specific point by pointing out the relative distance and angle of the rest of the points within a shape. This method considers the set of vectors originating from a point to all other sample points on a shape. The number of radial bins and theta bins are the main parameters of this descriptor.

As a result, the shape contexts of each point is fixed to a $(r\_bins * \theta\_bins)$ vector. Figure 3 shows an example of a shape context descriptor for the point within a diamond in a letter "A" shape. Dark colors mean a high density of points within that bin, while lighter colours implies less density of points. The histogram shows that the majority of points are very near either in the radial and/or angle sense.

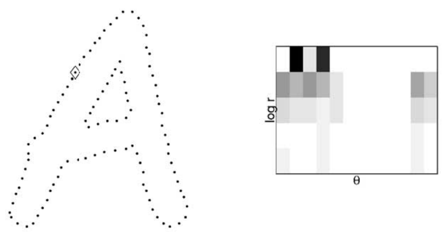


Fig. 3. Example of a sequences of points describing a letter "A" shape and its shape context descriptor for the single point within a diamond, extracted from [9]

This work attempts to study whether this complex descriptor may improve the performance of the system or not. Once the shape context descriptors are computed for all the points describing a contour, DTW algorithm is applied to find the best alignment between them instead of using the contour coordinates.

IV. Experimental Results

This section describes the experimental work carried out to analyze the performance of the different approaches described in Section III. The three aforementioned methods are tested with the contour coordinates of the BIOGIGA database previously described in section II.

In this work, three different experimental protocols are considered: *i)* protocol 1:3, *ii)* protocol 2:3 and *iii)* protocol 3:3, 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. The training images are the ones that the system previously have of each user and are used to enroll the user into the system, while the test images are the ones given by the user when he or she tries to be accepted by the system. In the experiments all the training images are images simulated with clothes, and the test images without clothes (in order to have the most challenging scenario with severe mismatch between enrollment and testing regarding clothing). The configuration for each protocol is as follows:

- P1:3, train (cr_*), test: $(sr_{10}, sr_{-10}, sr_0)$
- P2:3, train (cr_*, cr_*), test: $(sr_{10}, sr_{-10}, sr_0)$
- P3:3, train $(cr_{10}, cr_{-10}, cr_0)$ test: $(sr_{10}, sr_{-10}, sr_0)$

where cr_* stands for images with clothes for the three angles between subject and camera (-10, 0 and 10), and sr stands for images without clothes. It is worth noting that when having 2 or 3 images for training or testing, 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 a sum rule. This is mainly due to the fact that contours do not have the same size in all cases, so it is unfeasible to make the fusion at the feature-level.

A major issue to bear in mind relies on the importance of having contours of the same size or not. Since the database contains six different images per individual, every single contour obtained from its respective image has a number of coordinates that do not necessarily coincide with the number of coordinates of the remainder images from the same subject. Apart from the intra-person variability, there is also a large inter-person difference regarding the size of contours belonging to different subjects. Hence, experiments based on DTW are analyzed with all contours having their original size (MeanS) or normalized to the same size (NormS). ED and SC experiments are only carried out with contours normalized to the same size.

The next subsections describe the four experiments carried out in this paper. The first compares the baseline technique based on the Euclidean distance versus the DTW algorithm; the second experiment deals with the effects of normalizing the size of the contours; the third one analyzes the importance of the dimension of the contour sequence while the fourth experiment studies the convenience of using a more complex descriptor of the body shape, such as its shape context. Finally, results achieved here are compared to the results from [4].

A. Comparison between ED with DTW

This experiment analyzes the convenience of using DTW compared to the baseline technique based on the Euclidean distance (ED). Taking into account that the computation cost of DTW is larger than the computation cost of ED, DTW is only worthwhile if its use implies an improvement of the performance.

All the results are depicted in Table I in the form of equal error rates (EER), a popular measure of performance for biometric systems working in a verification mode. For each experiment six EER values are computed according to the protocol used (P1:3, P2:3, P3:3) and the dimension of the contour size. Two different contour sizes are considered: the original size of the contour obtained in the preprocessing stage described in section II, and a reduced contour set of 500 points approximately.

In this case, we are interested in comparing the first two rows of Table I. We conclude that applying DTW turns out in better performance of the system. For instance, applying DTW to the contour coordinates instead of computing their Euclidean distance results in an average relative improvement of 74.22% of EER (going from 5.3% EER to 1.2% EER for P2:3) for the case of using the original size of the contour, and 61.42% (going from 5.3% EER to 2.0% EER for P2:3) for the 500-points contour size. This improvement is due to the power of DTW algorithm of finding a good alignment between strings of points which optimizes the distance between them.

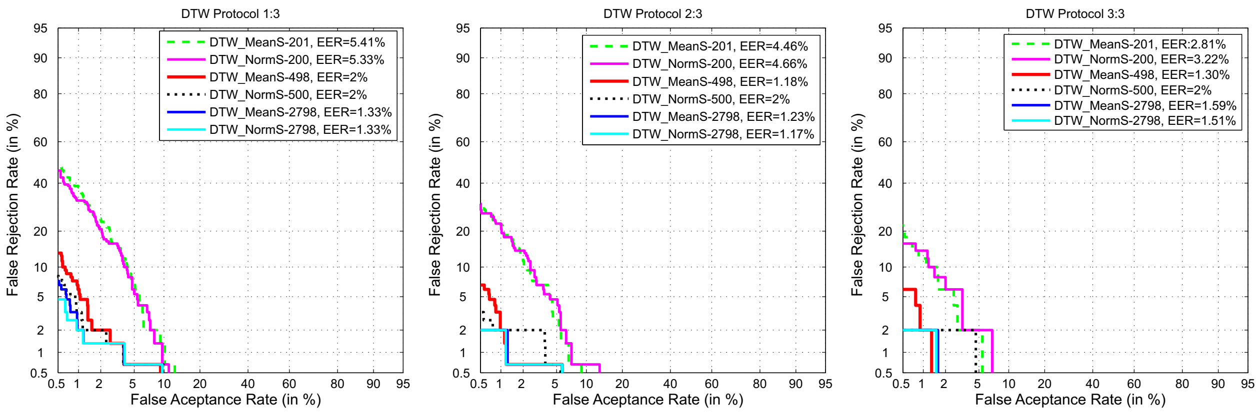
Regarding the three protocols defined, the P2:3 achieves the best results for the DTW approach (1.17% EER), but all three protocols achieve very similar results.

B. Analyzing the effect of contour normalization

In this experiment, the effect of the contour normalization is studied. In this case, we are interested in comparing the

TABLE I. PERFORMANCES OF BODY SHAPE-BASED BIOMETRIC SYSTEMS BASED ON THE EUCLIDEAN DISTANCE, DYNAMIC TIME WARPING AND SHAPE CONTEXTS DESCRIPTORS.

	Contour Size 2800 approx.			**Contour Size 500 approx.**		
Experiment	**EER 1:3**	**EER 2:3**	**EER 3:3**	**EER 1:3**	**EER 2:3**	**EER 3:3**
ED NormS	5.33 %	5.33%	4.89%	5.33%	5.33%	4.89%
DTW NormS	**1.33 %**	**1.17%**	**1.51%**	2.00%	2.00%	2.00%
DTW MeanS	1.33 %	1.23%	1.59%	**2.00%**	**1.18%**	**1.30%**
DTW SC NormS	2.00 %	2.00%	2.00%	2.38%	2.00%	2.00%



(a) Protocol 1:3 (b) Protocol 2:3 (c) Protocol 3:3

Fig. 4. DET curves for the case of DTW approach for contours with sizes 2800, 500 and 200, being normalized (NormS) or with their original sizes (MeanS), and for Protocols 1:3, 2:3, and 3:3.

second and third rows of Table I, which reports the results obtained for the cases of normalizing the contour sequences to the same size (NormS) and using the original size or a sampling of them (MeanS), respectively for the DTW approach.

As can be seen from the second and third row, when using the 2800 contour-size, normalizing the contours yields a very small improvement of performance which is not very significant. However, when using the sequences of contours of 500 points, there is an average relative improvement of 25.34% of EER when using the subsampling of the original size (1.2% EER for P2:3) compared to normalising the size to 500 (2% EER for P2:3). Therefore, we can conclude that using the original size or a subsampling of it, its a better option than normalizing it to the same size. A reason for this is that contours of the same person are more or less similar in size, while more different to other persons, making this a discriminative factor.

C. Studying the effect of the contour size using DTW

In this experiment we analyze the effect of the size of the contour sequences for the case of DTW (second and third rows in Table I).

We studied three different dimensions of the contours: 2800 and 500 as given in the table, and also a smaller size of 200 points. Also, we show the results for both normalized contours (NormS) and contours having different sizes (MeanS), for the three protocols defined. These results are given with DET curves in Figure 4.

As can be seen from Figures 4(a), 4(b) and 4(c), the best results in all cases are obtained when using the original size of the contour, i.e., 2800 in average. As the dimension of the contour is reduced, the performance of the system worsens. The relative improvement between the worst and the best scenario is about 75.42% for the protocol 1:3, 74.9% for the protocol 2:3 and 53.11% for the protocol 3:3.

Comparing the results from the different protocols, we can see how all the DET curves approach the origin when more data is used for the training stage, i.e., better results are obtained for the Protocol 3:3. It is worth noting that in this case the number of scores is smaller compared to the other two protocols, and this makes the EER values not to be smaller compared to the other cases.

D. Using DTW with shape contexts

The last experiment carried out in this work analyzes the effect of using more information about the contour than just the two x, y contour coordinates. In this case we apply a shape contexts descriptor to each contour coordinate before DTW and analyze the results obtained.

In this specific case, the shape context approach was set with a log-polar histogram of 12 bins of radial distance and 5 bins of angle distance, resulting a vector of 60 components per each point of the contour. This is applied to the contour with size normalized to 2800 points and 500 points.

It is worth noting that in this case the score is computed as following:

$$score = -\frac{DIST}{K} \quad (3)$$

The exponential is not applied as in Equation 2 because the results were much better this way. The results are depicted in the fourth row of Table I. With these results, we conclude that complex schemes such as shape contexts do not result in a significant improvement of the system. A further study analyzing different sizes of the radial and angle distances are also proposed for further work.

E. Discussion

Finally, the results achieved in this work can be compared to previous experiments using the same database and protocols [4]. In that work, the biometric system was based on 21 geometric distances between different key points of the contour. In that case, results of 2% EER were achieved for the three protocols. In this work we obtain an average relative improvement of 33.17% EER for the best case (DTW NormS with 2800 contour size), having in average 1.33% EER. This previous work by Moreno-Moreno *et al.* [4] also applied a SFFS feature selection algorithm improving their EER results very significantly. In this case, this SFFS algorithm could also be employed to further improve the system performance.

V. Conclusions

In this paper, a complete body-shape based biometric system has been developed for MMW body images (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 to extract easily the contours from the images. Different approaches have been analyzed ranging from naive approaches such as the Euclidean distance to complex schemes such as the dynamic programming techniques and shape context descriptors. The best results are obtained when using the DTW algorithm directly to the contour coordinates for the contours with the best resolution. In this case, EER results in the range of 1.2%-1.5% are achieved, which correspond to a relative improvement of 33% EER compared to previous works based on geometric measures between key points of the contour carried out over the same database.

The limitations of this work are related to the special characteristics of the database used. Images from BIOGIGA database are limited when compared to real images acquired in practice, but there are not publicly available databases for the moment. However, the synthetic images are based on real measures from people and are similar to real data captured at 94 GHz.

For future work, we propose to keep investigating other feature descriptors for the body shape such as Fourier descriptors [10] and also based on the texture of the images [2]. Also, the fusion of both shape-based and geometric-based proposed biometric systems is of interest.

Acknowledgments

This work has been partially supported by projects TeraSense (CSD2008-00068), Bio-Shield (TEC2012-34881), Contexts (S2009/TIC-1485) and "Cátedra UAM-Telefónica". E. Gonzalez-Sosa is supported by Obra Social La Caixa-Convocatoria 2012, and R. Vera-Rodriguez by a Juan de la Cierva Fellowship from the Spanish MINECO.

References

[1] M. Moreno-Moreno, J. Fierrez, and J. Ortega-Garcia, "Biometrics beyond the visible spectrum: Imaging technologies and applications," in *Proc. Biometric ID Management and Multimodal Communication*. Springer, 2009, pp. 154–161.

[2] B. Alefs, R. den Hollander, F. Nennie, E. van der Houwen, M. Bruijn, W. van der Mark, and J. Noordam, "Thorax biometrics from millimetre-wave images," *Pattern Recognition Letters*, vol. 31, no. 15, pp. 2357–2363, 2010.

[3] M. Moreno-Moreno, J. Fierrez, R. Vera-Rodriguez, and J. Parron, "Distance-based feature extraction for biometric recognition of millimeterwave body images," in *Proc. IEEE Intl. Carnahan Conf. on Security Technology, ICCST*, 2011, pp. 326–331.

[4] ——, "Simulation of millimeter wave body images and its application to biometric recognition," in *Proc. International Conferencie SPIE Defense Security and Sensing, SPIE*, vol. 8362, 2012, pp. 83 620E–1.

[5] E. Yoruk, E. Konukoglu, B. Sankur, and J. Darbon, "Shape-based hand recognition," *IEEE Transactions on Image Processing*, vol. 15, no. 7, pp. 1803–1815, 2006.

[6] A. K. Jain, F. D. Griess, and S. D. Connell, "On-line signature verification," *Pattern Recognition*, vol. 35, no. 12, pp. 2963 – 2972, 2002.

[7] M. Yasuhara and M. Oka, "Signature verification experiment based on nonlinear time alignment: a feasibility study," *IEEE Trans. Systems Man Cybernet*, vol. 17, pp. 212–216, 1977.

[8] A. Kholmatov and B. Yanikoglu, "Identity authentication using improved online signature verification method," *Pattern recognition letters*, vol. 26, no. 15, pp. 2400–2408, 2005.

[9] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 4, pp. 509–522, 2002.

[10] E. Persoon and K.-S. Fu, "Shape discrimination using fourier descriptors," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 7, no. 3, pp. 170–179, 1977.