

# Exploring Body Shape from mmW Images for Person Recognition

Ester Gonzalez-Sosa, Ruben Vera-Rodriguez,  
Julian Fierrez *Member, IEEE*, and Vishal M. Patel, *Senior Member, IEEE*

**Abstract**—Due to the ability of millimeter waves (mmWs) to penetrate dielectric materials such as plastic, polymer and clothes, the mmW imaging technology has been widely used for the detection of concealed weapons and objects. The use of mmW images has also recently been proposed for biometric person recognition to overcome certain limitations in image acquisition at visible frequencies. This paper proposes a biometric person recognition system based on the shape information extracted from real mmW images. To this aim, we report experimental results using the mmW images with different body shape-based feature approaches such as contour coordinates, shape contexts, Fourier descriptors and row and column profiles. We also study various distance-based and classifier-based matching schemes. Experimental results suggest the potential of performing person recognition through mmW imaging using only shape information, a functionality that could be integrated in the security scanners deployed in airports.

**Index Terms**—mmW imaging, body shape information, border control security, contour coordinates, dynamic time warping, support vector machines, shape contexts, Fourier descriptors, row and column profiles.

## I. INTRODUCTION

**M**ILLIMETER waves (mmWs) are high-frequency electromagnetic waves in the range of 30 – 300 GHz with wavelengths between 10 to 1 mm. These types of waves have been recently found to have interesting properties for various pattern recognition applications. The use of mmW imaging in particular has been gaining interest in the security research community [1]–[3], mainly due to its low intrusiveness and the ability to pass through clothing and atmospheric occlusions. Traditional applications of this technology include concealed weapon detection (CWD) [4]. Radiation at the mmW frequencies is non-ionizing and is therefore considered safe for human exposure. As a result, mmW scanners have been deployed in several international airports, replacing the former X-rays scanners. The suitability of these frequencies for CWD relies on the different response (due to difference of temperatures) between metallic objects and the human body skin.

Research in mmW imaging has focused in increasing the narrow depth of field (distance over which an object is

Ester Gonzalez-Sosa, Ruben Vera-Rodriguez and Julian Fierrez are with the Biometrics and Data Pattern Analytics Lab (BiDA) - ATVS, Escuela Politecnica Superior, Francisco Tomas y Valiente 11, Universidad Autonoma de Madrid, 28049 Madrid. e-mail: {ester.gonzalezs, ruben.vera, julian.fierrez}@uam.es

Vishal M. Patel is with the Department of Electrical and Computer Engineering, Rutgers University, Piscataway, NJ, 08854 USA e-mail: vishal.m.patel@rutgers.edu

Manuscript received ...

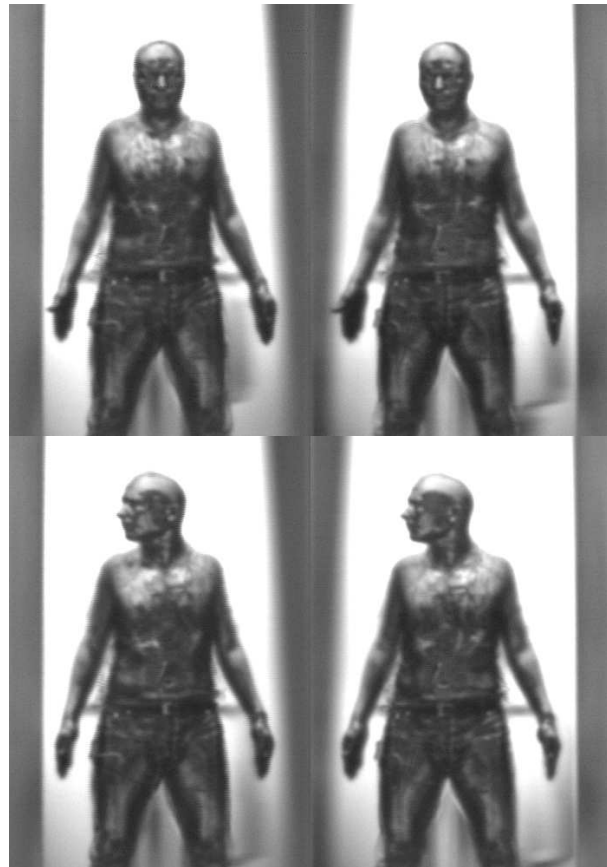


Fig. 1: Two full-scanning screening images from the mmW TNO database acquired with a stereo radiometric scanner. A full-scanning screening image is comprised of two single images with slight variations in pose. Two different head poses are depicted: frontal (above) and lateral (below). Figures are extracted from [9].

considered in focus) and the screening time of mmW imaging systems, in order to perform CWD through a corridor without creating bottlenecks [5], [6]. Automatic detection of concealed weapons, explosives, and contraband through mmW imaging is still an active area of research [4], [7], [8].

In the area of biometric person recognition, some researchers have studied the use of images acquired beyond the visible spectrum (e.g. X-ray [10] and infrared (IR) [11], [12]) with the aim of overcoming the limitations such as illumination variations and body occlusions.

Millimeter wave images present some benefits for biometric



Fig. 2: Benefits of mmWs for imaging: *i*) transparency through body clothes (extracted from [3], and *ii*) transparency through face clutter (images from the mmW TNO database)

person recognition, which make them especially noteworthy for further exploration and study. The key benefits lie in the wave properties enabling it to pass through clothes and other obscurants. As a consequence, shape information retrieved from the mmW images may be more robust to clothing variations than visible images. For the same reason, mmW imaging can also be exploited to retrieve texture information, which could potentially be used as discriminatory information in person recognition applications (see the torsos in Fig. 1). Furthermore, mmWs are also able to pass through accessories such as balaclavas, caps or artificial beards. This property makes mmW images less susceptible to spoofing attacks when compared to images in the visible spectrum [13]. People modifying their body constitution or using sophisticated facial masks may be easily detected using mmW images.

Based on the interest among the security research community in these frequencies, together with the promising recent results for mmW person recognition [9], [14], one may consider the possibility of using mmW images acquired from the screening scanners located at the airports simultaneously for both detecting hidden objects and performing person recognition. This two-fold control procedure could address better the security issues, threads and challenges faced in today's society.

Despite the interesting properties of mmW images, only a few works have used mmW images for person recognition purposes. This shortage of biometric recognition research based on the mmW images is mainly due to the lack of databases of images of people acquired at this frequency band as well as the privacy concerns of these images [15]. Alefs *et al.* developed one of the first reported efforts for person recognition using real mmW passive images acquired in outdoors scenarios [9] (see Fig. 1). They exploited the texture information contained in the torso region of the image through multilinear eigenspaces techniques. On the other hand, the works by Moreno-Moreno *et al.* [16] and by Gonzalez-Sosa *et al.* [14] proposed and analyzed a biometric person recognition system based on shape information extracted from synthetic images from the BioGiga database, exploiting geometrical measures between different silhouette landmarks and features based on contour coordinates, respectively. In all cases, images

were extracted using a frequency of 94 GHz. In the present work, we will form an in depth analysis of the discriminative capability of shape-based features using real mmW images from the mmW TNO database [9]. In the context of mmW scanners placed at airports, subjects are scanned following a cooperative protocol with the arms upwards and legs separated between each other (see Fig. 13). As our idea is to study the possibility of using the same mmW scanned images for both CWD and person recognition applications, we elaborate our study under the assumption of this cooperative protocol.

Body shape information from mmW images may not be substituted as a primary biometric such as face or fingerprint, but it may be useful for narrowing the search of possible suspects with very little effort. Furthermore, since mmW can pass through clothing and other materials, one can perhaps use it to detect spoofing. Besides, for some specific applications such as screening scanners at airports, mmW images have clear benefits which could be used to improve the recognition rates obtained by more traditional biometrics. In the future, it could also be possible to have visible cameras along with mmW scanners and therefore benefit from the properties of the different spectral regions through multimodal fusion schemes.

This paper is structured as follows. A brief review of shape-based biometric recognition applications is described in Section II. The mmW TNO database used in this paper is described in Section III. Section IV describes the different shape-based features used in the biometric system, while Section V addresses the different matching approaches explored in this work. The experimental protocol and results of these methods are presented in Section VI and VII respectively. Finally, Section VIII concludes the paper with a brief summary and discussion.

## II. RELATED WORK ON SHAPE-BASED BIOMETRIC PERSON RECOGNITION

By shape information, we mean any type of information that may be retrieved from binary images. Body shape-based recognition techniques are enclosed within the wide area of object shape-based recognition. A general survey about the different techniques of modeling shape can be found in [17], [18]. In the area of person recognition, several biometrics traits exploit the shape information:

*a) Hand Shape:* Different information from hand shape silhouettes has been used to perform authentication: widths and lengths of the hand [19], others also include the palm size [20]; hand contour [21]; or independent component analysis features [22].

*b) Gait:* The appearance-based gait approaches use the lower part of the silhouette of the body normally in a lateral view to model the subject. Features may be acquired from a set of frames belonging to the same cycle, or frame independently. The most common features are based on an average silhouette [23] (Gait Energy Image, Active Energy Image) .

*c) Signature:* Offline and online signature recognition methods explore signature shape information through its area, aspect ratio, contour coordinates, number of edge points, horizontal and vertical histogram, curvature, etc. [24].

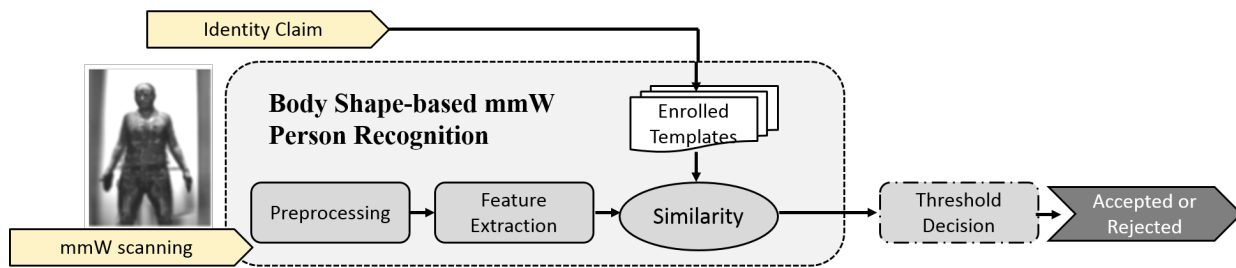


Fig. 3: Body shape-based person recognition using mmW imaging.

*d) Fingerprint and Vein:* There are previous works in the literature that address fingerprint recognition through ridge-based structure description [25]. Likewise, there are some works that utilize the skeleton pattern of the vein [26] or their shape [27] to perform identification. Some knuckle point perimeter distances have been also proved to improve a vein triangulation system [28].

*e) Soft Biometrics:* Body shape information has also been used to extract certain soft biometrics such as gender [29] or even age [30].

### III. THE MMW TNO DATABASE

Further research on mmW imaging has been predominantly blocked by two principal issues: *i)* high cost of acquisition devices, and *ii)* privacy concerns. Certainly, the average cost of commercial mmW scanners surpasses \$150,000<sup>1</sup>, impeding easy access to mmW data. Even more critical are the privacy issues that enclose mmW images. As mmWs can pass through materials and clothes, mmW images from human beings present partially naked figures, worsening the scene regarding the utilization of these data for research purposes.

Images are recorded using a passive stereo radiometer scanner in an outdoor scenario. The passive mmW radiation emitted by the subject arrives to a mirror that provides the mmW passive radiation to two hyperbolic antennas. Therefore, each full scanning is a set of two single images with slightly different points of view of size  $696 \times 499$  (width  $\times$  height). Fig. 1 shows some full-scanning examples. As can be seen, radiation passes completely through the upper part clothes, whereas the lower part clothes remain partially in the mmW image (see for example the belt). This fact has to do with the different attenuation of each specific material.

The database is comprised of images belonging to 50 different male subjects in 4 different scenarios. These 4 different scenarios derive from the combination of 2 different head poses and 2 different facial occlusions. In the first head pose configuration, the subject is first asked to stand in front of the scanner with head and arms position at a fixed rack (*frontal head pose*). In the second pose configuration (*lateral head pose*), the subject is asked to turn his head leftward whereas the torso is asked to remain fixed (it may suffer some small changes due to the head movement). Fig. 1 depicts above a full-scanning with frontal head pose and below a full-scanning with lateral head pose ( $0^\circ$  and  $90^\circ$  in the yaw axis respectively)

<sup>1</sup>Source from <https://www.propublica.org/special/scanning-the-scanners-a-side-by-side-comparison>.

TABLE I: The mmW databases available for person recognition purposes. Here, M stands for male and F for female.

|                | BioGiga [16]         | mmW TNO [9]      |
|----------------|----------------------|------------------|
| Nature         | Synthetic simulation | Real scanner     |
| Architecture   | Active and Passive   | Passive          |
| Scenarios      | Outdoors and Indoors | Outdoors         |
| Subjects       | 25 (M) and 25 (F)    | 50 (M)           |
| Images subject | 24                   | 8                |
| Total Images   | 1200                 | 400              |
| Variations     | Pose and Clothes     | Pose and Clutter |

In order to prove the benefits of mmW imaging above visual imaging, images with different facial configurations were also extracted. In this case, a second round of images with the first and second head pose configurations were extracted but now a large part of the facial region was occluded using an artificial beard or balaclava. We will refer to these two different facial clutter configurations as *clutter* and *non clutter*, respectively.

As mentioned before, each scanning is a set of two images. By dividing this set into single images of  $348 \times 499$ , the TNO database is comprised of 50 subjects  $\times$  2 head pose configurations  $\times$  2 facial clutter configurations  $\times$  2 images per set, making a total of 400 images in the whole mmW TNO database. Table I compares the two mmW datasets available for person recognition.

As shown in Fig. 1, subjects appear with their legs and arms separated from their body (with the arms downwards). In the case of commercial applications, such as in Fig. 13, subjects usually have their arms upwards. Apart from this minor difference, we can say that in both cases subject follows a cooperative protocol. Hence, conclusions drawn from TNO images may be translated to mmW images acquired for border control applications.

### IV. BODY SHAPE FEATURES

The biometric recognition module of this work aims to perform person recognition through body shape-based information extracted from mmW images. Fig. 3 shows the general scheme followed in this work. First we need to preprocess the mmW input image, in order to obtain a binarized image from which to extract shape-based features. This work has been mainly aimed to explore the potential use of mmW images for person recognition purposes. In order to extract conclusions that could directly be related to the discrimination capabilities of body shape feature approaches, we isolate our

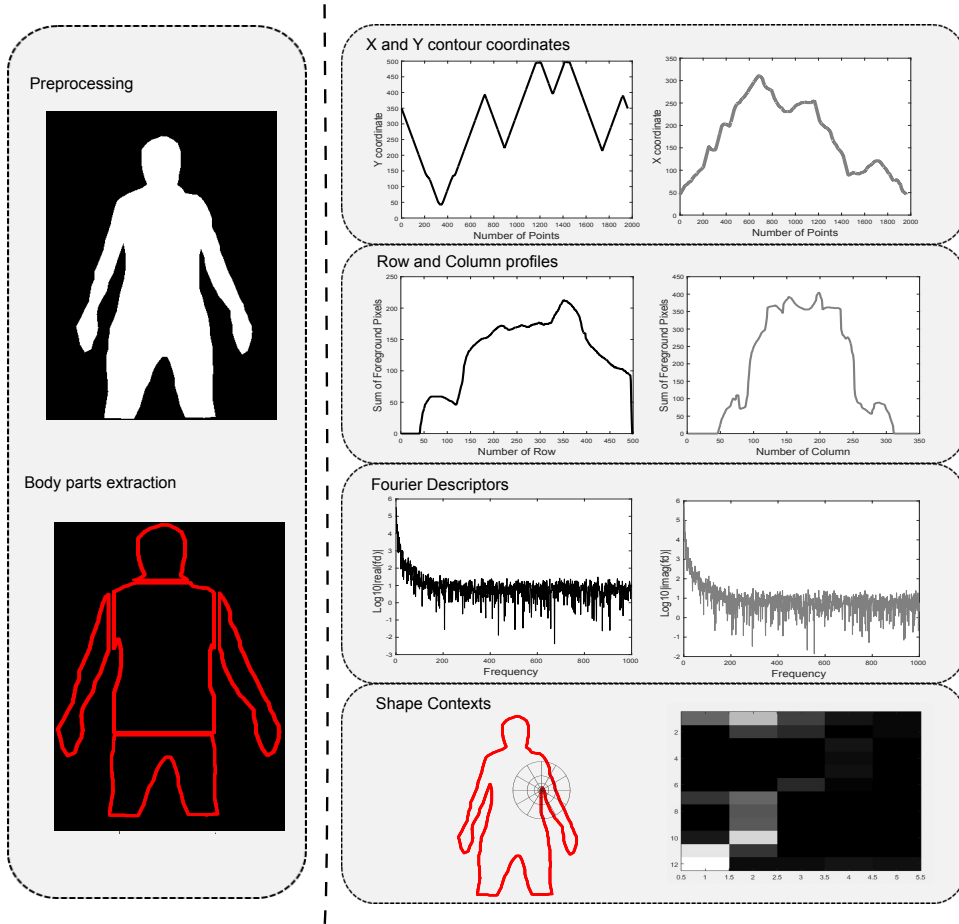


Fig. 4: Body shape features proposed in this work to perform person recognition using mmW images. The left side shows a binarized image (above) and different body part regions extracted after the segmentation (below). The right side shows the different shape-based features employed in this work. From top to bottom: contour coordinates, row and column profiles, Fourier descriptors and shape contexts.

proposed system by manually segmenting the mmW images. Concretely, the segmentation operation carried out in this work is done by selecting a set of 200 manual points. With the resulting binarized image, a finer detail contour is extracted. The preprocessing stage of this system also implies the mean normalization of both contour coordinates (column  $x$  and row  $y$ ), in order to achieve robustness against translations. Different feature approaches are considered in this work, classifying them into two main categories: contour-based features and image-based features.

#### A. Contour-based Features

The different feature approaches extracted from the contour coordinates considered in this work are: *i*) contour coordinates, *ii*) Fourier descriptors, and *iii*) shape contexts.

1) *Contour Coordinates (CC)*: are used as the baseline feature approach, being defined as  $\text{contour\_coordinates}(n) = (x_n, y_n), n = 1, \dots, n_{cc}$ , being  $n_{cc}$  the number of pixels that compose the contour that the silhouette edge describes with the background (approximately 2000 points), and  $(x_n, y_n)$  the coordinates of each one of those pixels. The starting point of the sequence

is the middle point of the head. Fig. 4 shows the  $y$  and  $x$  coordinates of the subject silhouette.

2) *Fourier Descriptors (FD)*: are simple to compute and robust against translations and rotations since the effect these transformations cause on the descriptors is completely known [17]. For each pair  $(x_n, y_n)$  we define the complex variable  $u_n = x_n + jy_n$ . The Discrete Fourier Transform of  $u_n$  is obtained as

$$\mathbf{f}(l) = \sum_{k=0}^{n_{cc}} u_k \exp(-j \frac{2\pi}{n_{cc}} lk), l = 0, 1, \dots, n_{cc} - 1. \quad (1)$$

The coefficients  $\mathbf{f}(l)$  are known as the Fourier descriptors of the boundary. Fig. 4 shows real and imaginary part of the Fourier descriptors. As can be seen, the discriminatory information of the Fourier descriptors lies in the lower frequencies, which accounts for the gross shape information.

3) *Shape Contexts (SC)*: were first introduced by Belongie *et al.* [31]. This technique describes a specific point considering the relative distance and angle to the rest of the points within a shape through a logarithmic and polar

histogram. The number of radial bins ( $r\_bins$ ) and theta bins ( $\theta\_bins$ ) are the main parameters of this descriptor. As a result: *i*) the overall number of bins in the logarithmic and polar histogram is  $r\_bins \times \theta\_bins$ , and *ii*) the shape contexts of a shape described by  $N$  points forms a vector of size  $(N \times r\_bins \times \theta\_bins)$ . Given a point  $p_i$  within the shape boundary, their corresponding logarithmic and polar histogram  $\mathbf{h}_i$  is computed as follows:

$$\mathbf{h}_i(k) = \#\{q \neq p_i : (q - p_i) \in bin(k)\}. \quad (2)$$

Equation 2 specifies that component  $k$  of histogram  $\mathbf{h}_i$  contains the number of points different to point  $p_i$  that lies in  $bin(k)$ , where  $bin(k)$  is the bin associated to a particular combination of radial and angular distance and  $k$  ranges from 1 to  $r\_bins \times \theta\_bins$ . The final feature vector of SC is the concatenation of all  $\mathbf{h}_i$  vectors of all points within the boundary. All radial distances from the remainder points of the sequence are computed and then normalized by the average distance. Any point that surpasses a preset relative distance is considered an outlier and it is not taken into account for the log polar histogram. Fig 4 shows an example of the  $12 \times 5$  shape contexts descriptor for a particular point within the mmW body shape. Dark colors imply a high density of points within a bin, while lighter colors imply a lower density.

### B. Image-based Features

Unlike the previous subsection in which features were extracted from contour coordinates, we describe here the shape-based approach that extract features from the binarized image.

1) *Row and Column Profiles (RCP)*: Given the binarized image  $I$ , whose pixels belonging to the foreground ( $fg$ ) are set to 1 ( $I(\mathbf{x}_{fg}, \mathbf{y}_{fg}) = 1$ ) and pixels belonging to the background ( $bg$ ) are set to 0 ( $I(\mathbf{x}_{bg}, \mathbf{y}_{bg}) = 0$ ), we compute row and column profiles as follows:

$$\text{row\_profile}(y) = \sum_{x=1}^{n_c} I(x, y), y = 1, \dots, n_r \quad (3)$$

$$\text{column\_profile}(x) = \sum_{y=1}^{n_r} I(x, y), x = 1, \dots, n_c \quad (4)$$

being  $n_r$  and  $n_c$  the height and width of the normalized body image. Hence, the row profile and column profile dimensionality will be of 499-length and 348-length, respectively. Fig. 4 shows the row and column profiles of the binarized image.

### C. Extraction of Body Regions

Different body parts are extracted using information from the binarized image and the global contour coordinates sequences. Five body regions are extracted: head, torso, right arm, left arm and legs. Concretely, the head is extracted by exploring the first minimum of the row profile; legs are also extracted by looking at the maximum of the row profile (normally coincides with the  $y$  coordinate in which the arms are almost at their border); and the torso is the region between

the head and the legs. Right and left arms are extracted by inspecting the binarized image from the center of the torso. By going leftwards and rightwards respectively, we detect the background region between the torso and the arms, which let us find the vertical line that separates the arms from the body.

It is worth noting that this body parts extraction is feasible due to the constrained position of people the body scanners (they are told to stand in a specific position, which is normally shown previously to the subjects). In this work, body parts regions are described through CC and RCP.

## V. MATCHING

### A. Distance-based Matching

1) *Dynamic Time Warping (DTW)*: In the specific case of row and column profiles, individual distances are computed independently between row profiles and column profiles. These distances are then averaged. With Fourier descriptors, we use their absolute value to compute similarities.

2) *Modified Hausdorff Distance (MHD)*: The Hausdorff distance is an effective method for shape comparisons [32]. Rather than measuring superposition, it estimates proximity. Two sets of points are close in terms of Hausdorff distance if every point of each set is close to some point of the other set. Likewise DTW, this distance may deal with sequences of points with different dimensionality. In this work, we apply the modified Hausdorff distance that aims to be more robust against noise by averaging individual distances rather than selecting the maximum (original version). The MHD between two sets of coordinates  $A = a_1, a_2, \dots, a_{N_a}$  and  $B = b_1, b_2, \dots, b_{N_b}$  is defined as:

$$\text{MHD}(A, B) = \max(h(A, B), h(B, A)). \quad (5)$$

where

$$h(A, B) = \frac{1}{N_a} \sum_{a \in A} \min_{b \in B} \|a - b\|, \quad (6)$$

$$h(B, A) = \frac{1}{N_b} \sum_{b \in B} \min_{a \in A} \|a - b\|. \quad (7)$$

### B. Classification-based Matching

1) *Support Vector Machines (SVM)*: We employ SVMs as classifiers. As SVMs require features with the same length, some previous operations are needed. CC are truncated or interpolated to the mean length of all sequences. FD and SC are recomputed with the new normalized contour coordinates. RCP do not require any operation, as they have already a fixed size. Due to the high dimensionality of those features, a dimensionality reduction operation is also applied using Discrete Continuous Transform (DCT). Only for the case of FD we select the first  $k$  Fourier descriptors instead of computing DCT over Fourier descriptors.

## VI. EXPERIMENTAL PROTOCOL

The experimental protocol followed in the previous work using the mmW TNO database [9] was very optimistic. It assumed 4 images as input (test images) and 4 images as training. Individual distances were computed comparing pairs



of images under the same head pose and point of view conditions, which is not a realistic situation. Then, the final distance was the minimum over the 4 former individual distances.

The experimental protocol proposed in this work aims to simulate the real situation in which a traveler would enter in the mmW scanner deployed in the security area of an airport. To this aim, enrolment is carried out in the first use of the mmW scanner. At the same time the subject is being scanned to target concealed weapons or dangerous objects, he is also compared with the previously enrolled template associated to the identity claimed in his passport. To simulate this scenario, we report results in verification mode. It would also be possible to compare the mmW image of the person with a watchlist of suspects. This case has been also considered and identification results are reported. Additionally, in order to gain insight from the benefits of using mmW images, we explore different experimental protocols.

*a) Frontal Protocol:* This protocol is set up as the baseline protocol. Among the 4 images with *frontal head pose*, we randomly select 2 images as gallery images and 2 images as probe images.

*b) Cross-Pose Protocol:* With this protocol, we aim to study the robustness of the different shape-based features proposed against pose variations. The mmW TNO database contains images with *frontal head pose* and with *lateral head pose* (see Fig. 1). To this aim, we randomly select 2 images from the *frontal head pose* subset as gallery images, and 2 images from the *lateral head pose* subset as probe images.

*c) Cross-Clutter Protocol:* Millimeter waves are able to pass through materials such as clothes. The mmW TNO database contains images with some clutter artifacts (caps, beards or balaclavas). With this protocol, we aim to gain insight about the influence of wearing artifacts over the task of person recognition. Therefore, we match frontal images without clutter against frontal images with clutter. We report experiments following this protocol for just some approaches, as clutter directly affects texture but not shape information.

In order to compute verification results, we face gallery images against probe images, generating 200 similarity target scores (2 probe images/subject  $\times$  50 subjects  $\times$  2 gallery images/subject) and 9800 similarity non-target scores (2 probe images/subject  $\times$  50 subjects  $\times$  49 impostor subjects  $\times$  2 images/subject). For each particular protocol, we define two random splits of gallery and probe images and then results are averaged. It is also worth noting that for the SVM classifier, the number of target scores is reduced to 100, as probe images are compared to a model trained with 2 gallery images.

In the identification mode, given a probe image, we compare it against each one of the 2 gallery images of a particular subject individually. Then, a final score results as a linear combination of the individual scores.

Verification results are reported in terms of the Equal Error Rate (EER) and Receiver Operating Characteristic curves (ROC). In the ROC curves, True Acceptance Rate (TAR) is depicted against False Acceptance Rate (FAR). EER is the value attained when False Acceptance Rate and False Rejection Rate (FRR=1-TAR) coincide. Identification results

are reported using the Cumulative Match Characteristic curves (CMC) and Rank-1 rates (R1). The CMC curves plot the recognition rate achieved at every rank value. For instance, Rank-1 means the percentage of cases in which the system outputs the genuine subject as the most probable subject, rank-5 means the percentage of cases in which the systems output the genuine subject among the first 5 candidates, and so forth.

## VII. RESULTS

This section describes the experimental results carried out to analyze the discrimination capabilities of different body shape-based feature approaches for the task of person recognition using mmW imaging. First, Section VII-A and Section VII-B study the influence of different configurations of shape contexts and row and column profiles, respectively. Then, the effect of the feature dimensionality is assessed in Section VII-C. In this case, we report results in the verification mode and using the frontal protocol. The performance of the different classifiers with the different feature approaches is addressed in Section VII-D. Then, the feasibility of using body regions separately to perform person recognition is addressed in Section VII-E. Lastly, some insight regarding the challenges ensuing automatic segmentation of mmW images are covered in Section VII-F.

### A. Exploring Shape Contexts

Shape contexts are computed with two main parameters: the number of angular bins ( $\theta\_bins$ ) and the number of radial bins ( $r\_bins$ ) of the log polar histogram. Fig. 5 a) and b) show the performance of shape contexts varying angular and radial bins, respectively. When varying the number of angular bins, the number of radial bins is fixed to 5, while 12 angular bins are used when varying the number of radial bins (5 radial bins and 12 angular bins are the default values proposed by the author [31]). The distance-based approach employed in this experiment is DTW.

If any of these two parameters increase, the log polar histogram of a sequence possesses much more detail on the distribution of all points with respect to a specific one, hence being more able to discriminate between subjects (reduction in EER). For the angular dimension, we obtain better results if we increase the angular dimension up to 48 bins. From these figures we can see that mmW person recognition through shape contexts benefit more significantly from a more detailed radial distribution of points rather than detailed angular distribution. The best configuration is achieved when using 50 radial bins and 12 angular bins, with an EER of 9.5%.

From these experiments we have learnt that shape contexts retain more discriminative information in the radial than in the angular dimension. Indeed, every single point within the human body shape contains points distributed throughout the entire radial dimension. If we think about a particular point  $p$  within the head contour, all points placed in the head would be near to  $p$ , points placed around the neck and upper torso would be middle-distance neighbors, while points belonging to the legs or lower torso would be far-distance neighbors. That will happen in a similar way with any considered point

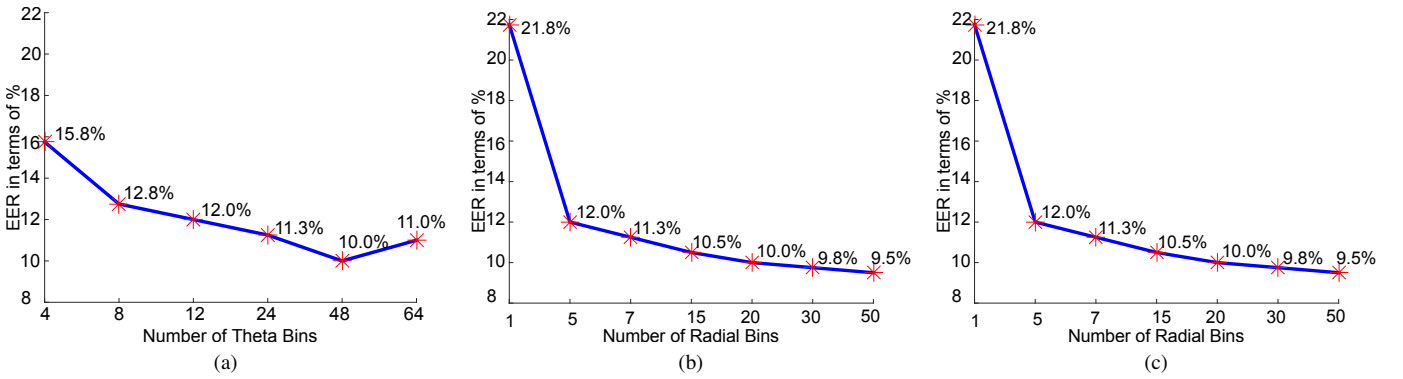


Fig. 5: Performance of shape contexts with: a) number of angular bins (theta bins); b) number of radial bins; c) dimensionality of the descriptor for the best configuration (50 radial bins and 12 angular bins) along with the information about computational time per comparison.

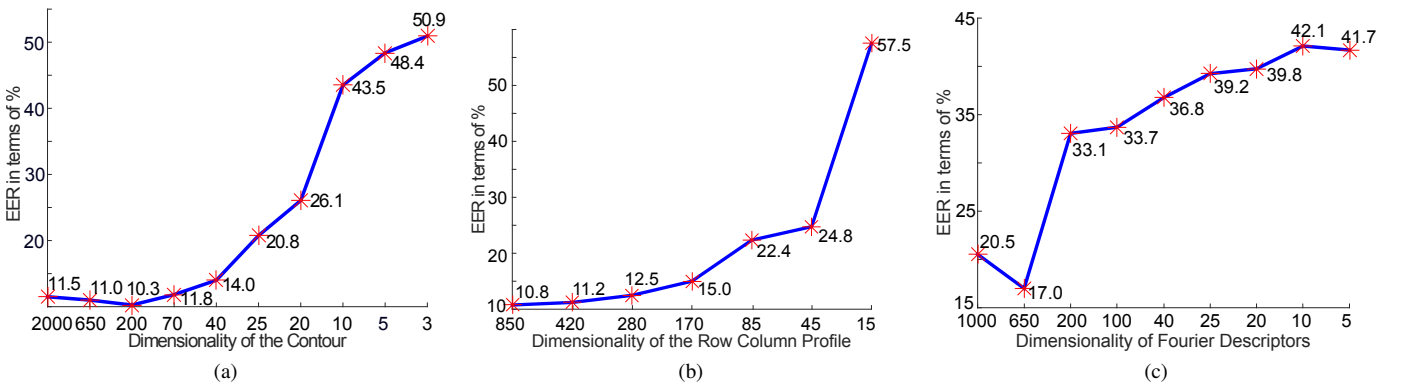


Fig. 6: Influence of the dimensionality of the different shape approaches: a) contour coordinates; b) row and column profiles and c) Fourier descriptors.

within the contour. Unlike the radial dimension, not all points within the contour would have neighbors in all angular bins. That would be the particular case of all outer points of the contour which will only have neighbors from a constrained portion of the entire angular circumference. This evidence is empirically proved with the experiments from Section VII-A.

### B. Exploring Row and Column Profiles

When assessing the performance of row and column profiles individually (using DTW as the matcher), we observed that column profile obtains better performance than the row profile (14.75% and 22.12% of EER, respectively). Then, the fusion of both profiles outperforms any of the individual cases (10.75% of EER). The superior performance of column profiles over row profiles might be because every subject has a row profile always defined by a neck region, hip, waist and so forth, regardless of their position, making this way the interclass variability small. On the contrary, column profiles depend more on the body position, that is, how the subject places their arms and legs inside the mmW scanner. We assume subjects have the tendency/inertia of positioning their body in a particular way that helps to discriminate better between subjects.

### C. Influence of the Feature Dimensionality

In this section, we analyze the influence of the feature dimensionality with respect to the system performance. Fig. 6 a), b), c) and Fig. 5 c) show the influence of the feature dimensionality for contour coordinates, row and column profiles, Fourier descriptors and shape contexts, respectively. The distance-based approach employed in this assessment is DTW.

For contour coordinates, shape contexts and row and column profiles, the dimensionality reduction is achieved by selecting a reduced number  $N$ , taking 1 sample out of  $k$  from the corresponding feature vector obtained with the highest resolution, being 2000 ( $N \leq 2000$ ) points for CC and SC and  $width + height = 850$  ( $N \leq 850$ ) for RCP. The final dimensionality of the feature vector will depend on the specific feature approach (SC= $N \times r\_bins \times \theta\_bins$ -length vector; CC= $N \times 2$ -length and RCP= $N \times 1$ -length, respectively). Regarding contour coordinates, the performance is moderately improved up to 10.3% if the sequence is down sampled to 200 points. It is also worth noting that similar results are obtained if we use 2000 (11.5%) or 70 points (11.8%). If the sequence is down sampled even more, the performance starts to degrade severely. Concerning row and column profiles, the best performance is achieved with the highest resolution (850 points). Performance is kept almost similar when decreasing down to 200 points, but with less than 200 points there is

not enough discriminatory information to distinguish between subjects.

In the case of the Fourier descriptors, the down sampling operation is different. Rather than taking 1 out of  $k$  samples, we select the first  $N$  components of the sequence ( $N \leq 1000$ ). This is because Fourier descriptor discriminatory information is not distributed homogeneously through the whole sequence, but it is concentrated on the low frequency components. If we pay attention to Fig. 6 c), we note that, if we use less than 650 components, the Fourier descriptors are not able to discriminate reliably. Also, the best result is obtained when we get rid of the highest frequency components (using 650 components). This fact suggests that the most discriminatory information relies only on the low frequency components.

In the case of shape contexts (see Fig. 5 c)), the influence of the feature dimensionality is explored with the best combination of parameters (number of angular and radial bins) discussed in Section VII-A, that turned out to be shape contexts with 50 radial bins and 12 angular bins. The dimensionality of the log polar histogram for a specific point is 600 ( $50 \times 12$ ). The high dimensionality of the whole contour is  $600 \times 2000$ , making it necessary to study the computational time employed per comparison. For that reason, the time per comparison (using DTW), is also plotted in Fig. 5 c). Likewise contour coordinates and row and column profiles approaches, the performance of shape contexts is almost kept if the feature is down sampled to 200 points. As with contour coordinates, performance improvement is also observed here when down sampling the sequence down to 400. Besides, the computational time per comparison between using 2000 and 400 or 200 is greatly reduced. In real applications, one should find a trade-off between performance and computational time involved, that in this particular case is achieved when using shape contexts descriptors with dimensionality between 200 – 400.

By the exploration of the dimensionality of the different feature approaches considered, we have learnt that it is not necessary to work with maximum resolution. If features with lower dimensionalities are used, similar or in some cases better results are obtained.

#### D. Verification and Identification results

1) *Verification Task*: Fig. 7 and 8 present the ROC curves obtained for the different matching and shape feature approaches for the frontal and cross-pose protocols, respectively.

The feature approach that works better with DTW is row and column profiles (10.75% of EER). For MHD, contour coordinates achieved even better results with an EER value of 9.25%. Regarding SVM, row and column profiles are the shape features that performed the best (8.0% of EER). Furthermore, RCP-SVM achieves the best results among all combinations of feature approaches and matchers. Note that this system is also robust to variations in pose, as the performance is maintained while following the cross-pose protocol (see Fig.8 c)).

It is also interesting to notice that MHD is not an appropriate matcher for neither RCP nor FD. This is mainly because neither of these shape approaches are retaining information of shape coordinates, but about their size (RCP) or their

frequency information (FD). Besides, MHD does not take into account any context information when computing the distance. That allows to match points regardless of their specific situation within the shape. Unlike MHD, DTW algorithm takes into account some spatial context information through their global and local constraints. This justifies the better performance of DTW with respect to MHD over RCP and FD approaches. However, the best results for these descriptors are achieved with SVM. Note that contour coordinates keep their verification performance very similar for the different matchers.

Fig 9 shows some examples of false negatives and false positives when using the CC-DTW following the frontal protocol. While false negatives occur when the system wrongly rejects two samples belonging to the same identity, false positives take place when the system wrongly accepts two samples that do not belong to the same identity. As can be seen, arm movements may lead to false negatives. Besides, issues such as similar height (case C) or similar constitution (case D) cause mismatched identities. These examples show the limitations of CC for recognition.

Table II presents the results obtained when following the cross-clutter protocol with respect to the baseline frontal protocol. As can be seen, verification results with the cross-clutter protocol are almost the same, there is only slight degradation. Hence, we deduce that mmW images might be potentially robust to occlusions. However, it is also noticeable that occlusions degrade more significantly texture than shape information. Therefore, it would be more interesting to study the influence of occlusions when assessing texture information from the mmW images.

2) *Identification Task*: Fig. 10 and 11 present the CMC curves obtained for the different matching and shape feature approaches for the frontal and cross-pose protocols, respectively. Some findings from the CMC curves are: *i*) the feature approach that works better with DTW are contour coordinates, with R1 of 82.50% *ii*) contour coordinates is also the best approach when using MHD, this time achieving higher rank-1 rates (90%), and *iii*) the best shape approach with SVM is SC (R1 of 70%). The best system configuration in terms of performance (CC-MHD with 90% of rank-1 in the frontal protocol) is also the most robust against pose variations, as it is able to maintain their performance in the cross-pose protocol.

It is also worth noting the limitations of Fourier descriptors for person recognition applications. In almost all the matching approaches, it is the shape feature approach with worst results in terms of EER and R1. With these results, we foresee that Fourier descriptors may be more suitable to recognize objects that differ much more in their shapes. The interclass variability between body human shapes is quite smaller than other standard object recognition applications.

#### E. Body Regions

Table III presents the results obtained for different body regions using contour coordinates and row and column profiles for 5 different body regions: left arm, right arm, legs, torso and head. The score level fusion of all parts is also computed. For



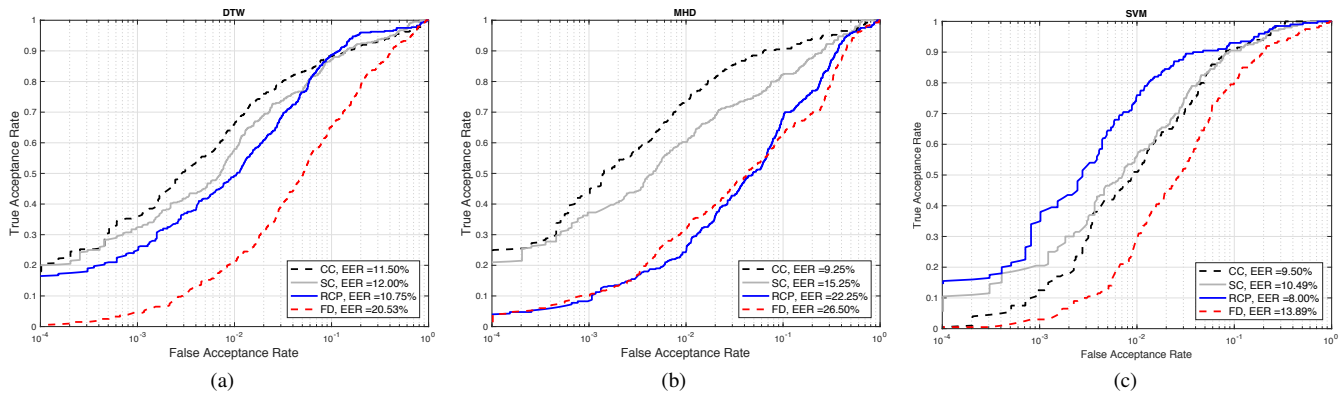


Fig. 7: Verification results of different matching approaches: a) Dynamic time warping, b) Modified Hausdorff distance and c) Support vector machines for the frontal protocol.

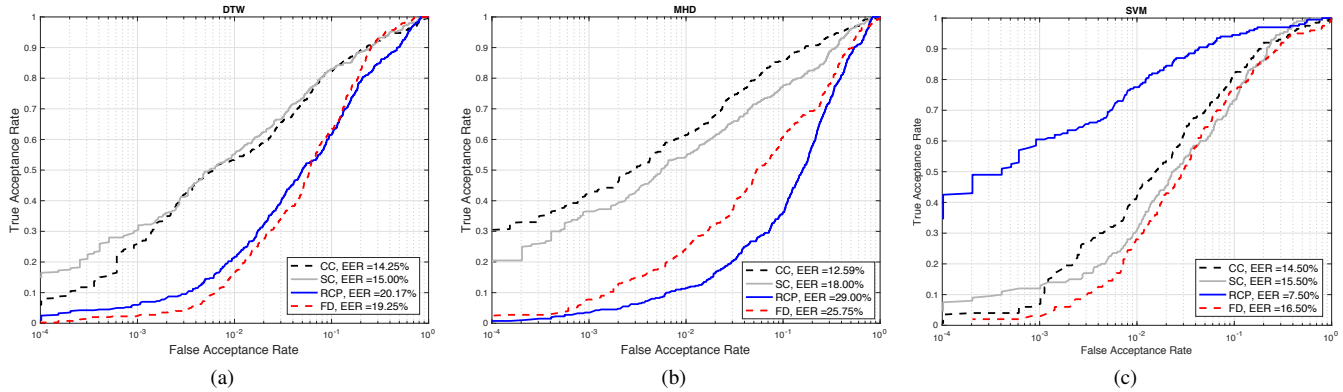


Fig. 8: Verification results of different matching approaches: a) Dynamic time warping, b) Modified Hausdorff distance and c) Support vector machines for the cross-pose protocol.

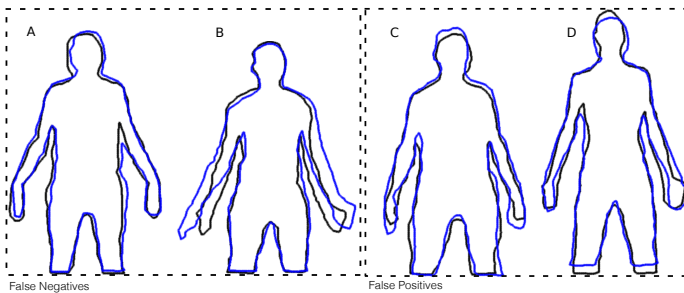


Fig. 9: False negatives and false positives produced by the CC-DTW system following the frontal protocol.

TABLE II: Assessment of the clutter influence over verification experiments following the frontal protocol.

| Body Part | No Clutter | Clutter |
|-----------|------------|---------|
| CC-DTW    | 11.50      | 11.90   |
| RCP-DTW   | 10.75      | 12.5    |

both approaches, the best body regions are left arm, right arm and torso. From last row of Table III, one can deduce that row and column profiles from different body parts contain complementary information that help to improve the overall performance. Nevertheless, the body part approach of row and

TABLE III: EER in % of 5 body regions using contour coordinates and row and column profiles as shape-based features and dynamic time warping as distance-based matcher.

| Body Part   | CC-DTW | RCP-DTW |
|-------------|--------|---------|
| Left Arm    | 22.4   | 23.25   |
| Right Arm   | 21.25  | 23.20   |
| Legs        | 38.33  | 25.75   |
| Torso       | 20.25  | 17.97   |
| Head        | 24.32  | 27.40   |
| All Regions | 28.33  | 14.25   |

column profile does not outperform the corresponding global approach (10.75% of EER for RCP-DTW). This fact leads us to conclude that there is no benefit from addressing mmW human shape through body regions.

### F. Automatic Segmentation

As discussed earlier, we have manually segmented the mmW images to study their discriminative capabilities for person recognition. In this section, we briefly discuss some of the challenges involved when an image segmentation algorithm is used to automatically segment the body shapes from the mmW images. From Fig. 1 one may think that segmenting

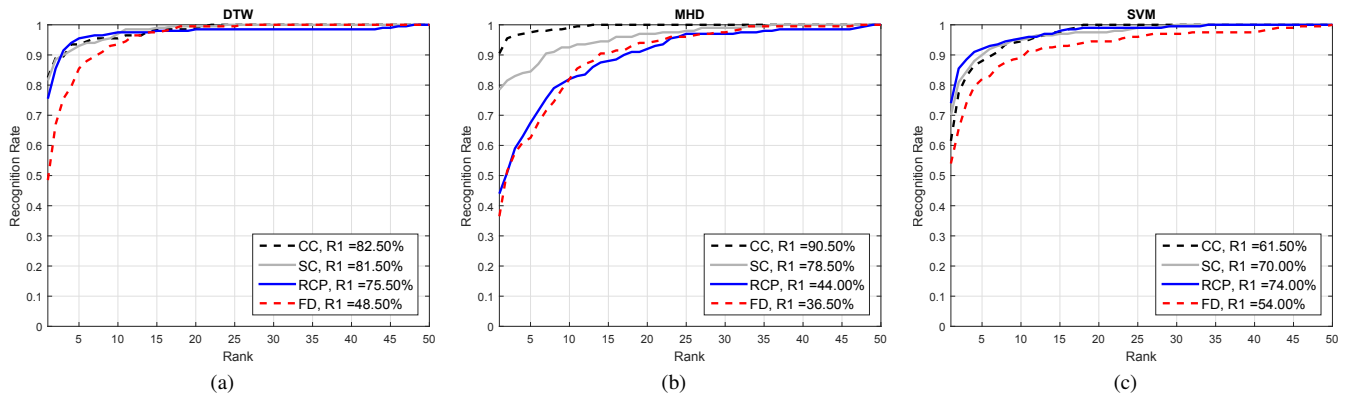


Fig. 10: Identification results of different matching approaches: a) Dynamic time warping, b) Modified Hausdorff distance and c) Support vector machines for the frontal protocol.

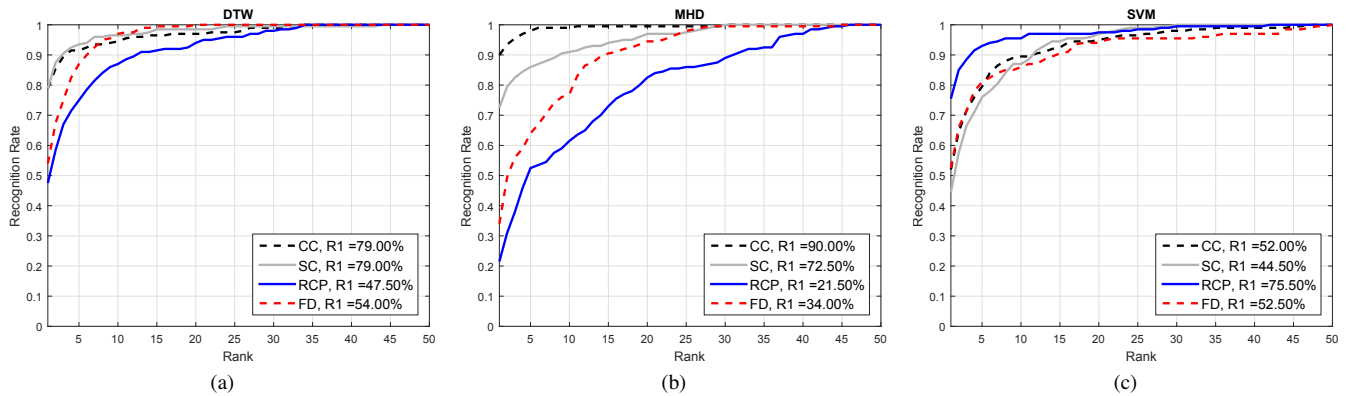


Fig. 11: Identification results of different matching approaches: a) Dynamic time warping, b) Modified Hausdorff distance and c) Support vector machines for the cross-pose protocol.

body shapes is a straightforward task. However, this is not the case for all the images in the database. In fact, images from different subjects vary greatly in terms of their appearance, making it more challenging to find an automatic segmentation algorithm that works well in all cases. Fig. 12 shows a few segmented images using the Active Contours algorithm [33] from the mmW TNO database. As can be seen from this figure, there are still some issues to overcome before introducing an automatic segmentation module into the pipeline of the proposed system: intra leg region not segmented, missing upper torso, missing head, etc. Table IV presents some verification results in terms of EER following the frontal protocol for CC and RCP extracted from manual and automatic segmented images, using DTW as the matcher. Though the performance is poor for both shape-based approaches, RCP seems to be less affected by the poorly segmented images. The inaccurate results of automatic contours prevent us to clearly see the suitability of different body shape approaches for the task of person recognition.

### VIII. CONCLUSIONS

The use of mmW images has been recently introduced in computer vision applications such as weapon detection and biometric person recognition applications. This is the first work addressing the problem of person recognition through



Fig. 12: Examples of automatic segmentation using the active contours algorithm [33] from different subjects in the mmW TNO database.

TABLE IV: Verification results in terms of EER for person recognition following the frontal protocol using body shape features extracted from manually and automatically segmented images.

| Segmentation | CC-DTW | RCP-DTW |
|--------------|--------|---------|
| Manual       | 10.75  | 11.90   |
| Automatic    | 45.75  | 29.25   |

body-shape information using real mmW images. Although there is still room for improvement, the experiments carried out show that person recognition through shape information

contained in mmW images is feasible. More specifically, we have learnt that: *i*) the performance is degraded severely when  $N$  is reduced to 600 – 200 points, depending on the specific feature approach, *ii*) more discriminative information is retained in the radial rather than the angular dimension of shape contexts, *iii*) Fourier descriptors are not good at discriminating between instances of shapes which are very similar among them, which is the case of human body silhouettes, *iv*) the best mmW person recognition in terms of performance and pose robustness is row and column profiles along with support vector machines (RCP-SMV with an EER of 8%) and contour coordinates and modified Hausdorff distance (CC-MHD with the R1 of 90%) for the verification and identification tasks, respectively, and *v*) addressing mmW person recognition using body regions is less convenient than using the global shape.

Comparing our conclusions with those extracted when using BioGiga database in [14], we see that contour coordinates with DTW have performed in a similar way in both mmW databases, but shape contexts or Fourier descriptors have worsened their performance considerably, showing us that these descriptors are less suitable for real mmW images. An additional factor to bear in mind apart from the real/synthetic difference between the mmW TNO and BioGiga databases, is the gender difference. The mmW TNO database only contains images from male subjects, while BioGiga is a gender-balanced database. This difference in gender may also be affecting the performance of shape features used in this work.

The reader should notice that conclusions from our work are constrained by some limitations. First, to prevent additional segmentation mistakes that could distort us to extract conclusions pertaining exclusively to the potential use of mmW for person recognition applications, we have reported all our results using manually-segmented images. Future works should report results using automatically segmented images. Second, some of our conclusions can be dependent on the particular characteristics of the mmW TNO database. For instance, conclusions related to the minimum amount of components needed to achieve optimal results could change if we deal with a database with a larger number of subjects.

To obtain more reliable conclusions about the benefits and limitations of mmW images for biometric purposes, we should compare our approach with the corresponding images acquired in other regions of the spectrum such as the visible range. Currently, there are no available databases of visible and mmW images for research purposes. However, this is certainly an interesting issue which is worth studying in the future.

Real applications of CWD typically use active images like the one shown in Fig 13. Although in this work we are using passive images due to the lack of active images for research, one may notice that the passive images employed in this work are somewhat similar to active images. It is worth noting though that active images have a better resolution and contrast than passive images, factors that could be beneficial when applying shape approaches to these images.

In our future work, we will analyze the texture information retained in the mmW images for person recognition [35]. Furthermore, we will study the discriminative capability of the mmW signature based on both shape and texture features.

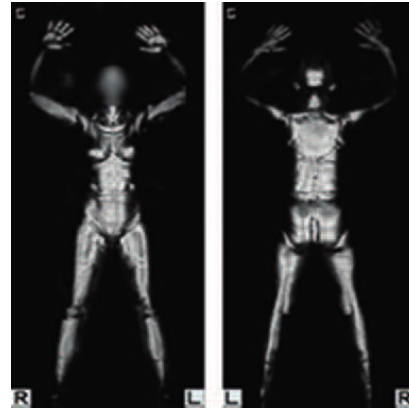


Fig. 13: Example of a mmW image acquired with an active mmW L3 ProVision Imaging scanner (extracted from [34]).

#### ACKNOWLEDGMENT

This work has been partially supported by project CogniMetrics TEC2015-70627-R (MINECO/FEDER), and the SPATEK network (TEC2015-68766-REDC). E. Gonzalez-Sosa is supported by a PhD scholarship from Universidad Autonoma de Madrid. Vishal M. Patel was partially supported by US Office of Naval Research (ONR) Grant YIP N00014-16-1-3134. Authors wish to thank also TNO for providing access to the database.



**Ester Gonzalez-Sosa** received the B.S. in computer science and M.Sc in electrical engineering from Universidad de Las Palmas de Gran Canaria in 2012 and 2014, respectively. In September 2012, she joined the Biometric Recognition Group - ATVS at the Universidad Autonoma de Madrid, where she is currently a PhD candidate. She has carried out several research internships in worldwide leading groups in biometric recognition such as TNO, EU-RECOM, or Rutgers University. Her research interests include pattern recognition, signal processing, and biometrics, with emphasis on face, body, soft biometrics and millimeter imaging. Gonzalez-Sosa has been the recipient of UNITECO AWARD from the Spanish Association of Electrical Engineers and the competitive Obra Social La CAIXA Scholarship.



**Ruben Vera-Rodriguez** received the M.Sc. degree in telecommunications engineering from Universidad de Sevilla, Spain, in 2006, and the Ph.D. degree in electrical and electronic engineering from Swansea University, U.K., in 2010. Since 2010, he has been with the Biometric Recognition Group, Universidad Autonoma de Madrid, Spain, first as the recipient of a Juan de la Cierva Post-Doctoral Fellowship from the Spanish Ministry of Innovation and Sciences, and as an Assistant Professor since 2013. His research interests include signal and image processing, pattern recognition, and biometrics, with emphasis on signature, face and gait verification and forensic applications of biometrics. Dr. Vera-Rodriguez is actively involved in several National and European projects focused on biometrics.



**Julian Fierrez** received the M.Sc. and the Ph.D. degrees in telecommunications engineering from Universidad Politecnica de Madrid, Spain, in 2001 and 2006, respectively. Since 2002 he has been affiliated with the Biometric Recognition Group (ATVS), first at Universidad Politecnica de Madrid, and since 2004 at Universidad Autonoma de Madrid, where he is currently an Associate Professor. From 2007 to 2009 he was a visiting researcher at Michigan State University in USA under a Marie Curie fellowship.

His research interests include general signal and image processing, pattern recognition, and biometrics, with emphasis on signature and fingerprint verification, multi-biometrics, biometric databases, system security, and forensic applications of biometrics. Dr. Fierrez has been actively involved in multiple EU projects focused on biometrics (e.g. TABULA RASA and BEAT), has attracted notable impact for his research, and is the recipient of a number of distinctions, including: EBF European Biometric Industry Award 2006, EURASIP Best PhD Award 2012, Medal in the Young Researcher Awards 2015 by the Spanish Royal Academy of Engineering, and the Miguel Catalan Award to the Best Researcher under 40 in the Community of Madrid in the general area of Science and Technology.



**Vishal Patel** received the B.S. degrees in electrical engineering and applied mathematics (Hons.) and the M.S. degree in applied mathematics from North Carolina State University, Raleigh, NC, USA, in 2004 and 2005, respectively, and the Ph.D. degree in electrical engineering from the University of Maryland College Park, MD, USA, in 2010. He is currently an A. Walter Tyson Assistant Professor in the Department of Electrical and Computer Engineering (ECE) at Rutgers University. Prior to joining Rutgers University, he was a member of

the research faculty at the University of Maryland Institute for Advanced Computer Studies (UMIACS). His current research interests include signal processing, computer vision, and pattern recognition with applications in biometrics and imaging. He has received a number of awards including the 2016 ONR Young Investigator Award, the 2016 Jimmy Lin Award for Invention, A. Walter Tyson Assistant Professorship Award, the Best Paper Award at IEEE BTAS 2015, and Best Poster Awards at BTAS 2015 and 2016. He is an Associate Editor of the IEEE Signal Processing Magazine, IEEE Biometrics Compendium, and serves on the Information Forensics and Security Technical Committee of the IEEE Signal Processing Society. He is a member of Eta Kappa Nu, Pi Mu Epsilon, and Phi Beta Kappa.

## REFERENCES

- [1] R. Appleby and R. N. Anderton, "Millimeter-wave and submillimeter-wave imaging for security and surveillance," *Proc. of IEEE*, vol. 95, no. 8, pp. 1683–1690, 2007.
- [2] L. Yujiri, "Passive millimeter wave imaging," in *IEEE MTT-S International Microwave Symposium Digest*, 2006, pp. 98–101.
- [3] V. M. Patel, J. N. Mait, D. W. Prather, and A. S. Hedden, "Computational millimeter wave imaging: problems, progress and prospects," *IEEE Signal Processing Magazine*, vol. 33, no. 5, pp. 109–118, 2016.
- [4] C. D. Haworth, Y. R. Petillot, and E. Trucco, "Image processing techniques for metallic object detection with millimetre-wave images," *Pattern Recognition Letters*, vol. 27, no. 15, pp. 1843–1851, 2006.
- [5] J. N. Mait, D. A. Wikner, M. S. Mirotznik, J. van der Gracht, G. P. Behrmann, B. L. Good, and S. A. Mathews, "94-ghz imager with extended depth of field," *IEEE Transactions on Antennas and Propagation*, vol. 57, no. 6, pp. 1713–1719, 2009.
- [6] V. M. Patel and J. N. Mait, "Compressive passive millimeter wave imaging with extended depth of field," *Optical Engineering*, vol. 51, no. 9, pp. 091 610–1, 2012.
- [7] G. N. Sinclair, R. N. Anderton, and R. Appleby, "Outdoor passive millimetre wave security screening," in *Proc. of IEEE Int. Carnahan Conference on Security Technology*, 2001, pp. 172–179.
- [8] X. Shen, C. R. Dietlein, E. Grossman, Z. Popovic, and F. G. Meyer, "Detection and segmentation of concealed objects in terahertz images," *IEEE Transactions on Image Processing*, vol. 17, no. 12, pp. 2465–2475, 2008.
- [9] 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.
- [10] H. Chen and A. K. Jain, "Dental biometrics," *Encyclopedia of Biometrics*, pp. 343–351, 2015.
- [11] S. Z. Li, R. Chu, S. Liao, and L. Zhang, "Illumination invariant face recognition using near-infrared images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, pp. 627–639, 2007.
- [12] A. Morales, E. Gonzalez-Sosa, and M. A. Ferrer, "On the feasibility of interoperable schemes in hand biometrics," *Sensors*, vol. 12, no. 2, pp. 1352–1382, 2012.
- [13] A. Hadid, N. Evans, S. Marcel, and J. Fierrez, "Biometrics systems under spoofing attack: an evaluation methodology and lessons learned," *IEEE Signal Processing Magazine*, vol. 32, no. 5, pp. 20–30, 2015.
- [14] E. Gonzalez-Sosa, R. Vera-Rodriguez, J. Fierrez, M. Moreno-Moreno, and J. Ortega-Garcia, "Feature exploration for biometric recognition using millimetre wave body images," *EURASIP Journal on Image and Video Processing*, no. 1, pp. 1–13, 2015.
- [15] J. F. Tugan, "Privacy and body scanners at EU airports," *Novatica Special English Edition*, pp. 49–54, 2013.
- [16] M. Moreno-Moreno, J. Fierrez, R. Vera-Rodriguez, and J. Parron, "Simulation of millimeter wave body images and its application to biometric recognition," in *Proc. of SPIE, Biometric Technologies for Human Identification*, vol. 8362, 2012.
- [17] M. Yang, K. Kpalma, and J. Ronsin, "A survey of shape feature extraction techniques," *Pattern Recognition*, pp. 43–90, 2008.
- [18] D. Zhang and G. Lu, "Review of shape representation and description techniques," *Pattern Recognition*, vol. 37, no. 1, pp. 1–19, 2004.
- [19] J. Burgues, J. Fierrez, D. Ramos, M. Puertas, and J. Ortega-Garcia, "Detecting invalid samples in hand geometry verification through geometric measurements," in *IEEE Proc. of Int. Conf. on Pattern Recognition*, 2010, pp. 113–120.
- [20] C. Oden, A. Ercil, and B. Buke, "Combining implicit polynomials and geometric features for hand recognition," *Pattern Recognition Letters*, vol. 24, no. 13, pp. 2145–2152, 2003.
- [21] A. K. Jain and N. Duta, "Deformable matching of hand shapes for user verification," in *Proc. of IEEE Int. Conf. on Image Processing*, vol. 2, 1999, pp. 857–861.
- [22] E. Yorum, E. Konukoglu, B. Sankur, and J. Darbon, "Shape-based hand recognition," *IEEE Transactions on Image Processing*, vol. 15, no. 7, pp. 1803–1815, 2006.
- [23] S. Gabriel-Sanz, R. Vera-Rodriguez, P. Tome, and J. Fierrez, "Assessment of gait recognition based on the lower part of the human body," in *IEEE Proc. of International Workshop on Biometrics and Forensics*, 2013, pp. 1–4.
- [24] M. Martinez-Diaz, J. Fierrez, and S. Hangai, *Signature Matching*. Springer, 2015, pp. 1382–1387.

- [25] A. K. Jain and K. Cao, "Fingerprint image analysis: role of orientation patch and ridge structure dictionaries," *Geometry Driven Statistics*, vol. 121, p. 288, 2015.
- [26] A. Yuksel, L. Akarun, and B. Sankur, "Hand vein biometry based on geometry and appearance methods," *IET Computer Vision*, vol. 5, no. 6, pp. 398–406, 2011.
- [27] L. Wang and G. Leedham, "Near-and far-infrared imaging for vein pattern biometrics," in *Proc. of IEEE Int. Conf. on Video and Signal Based Surveillance*, 2016, pp. 52–52.
- [28] A. Kumar and K. V. Prathyusha, "Personal authentication using hand vein triangulation and knuckle shape," *IEEE Transactions on Image Processing*, vol. 18, no. 9, pp. 2127–2136, 2009.
- [29] E. Gonzalez-Sosa, A. Dantcheva, R. Vera-Rodriguez, J.-L. Dugelay, F. Bremond, and J. Fierrez, "Image-based gender estimation from body and face across distances," in *Proc. of IEEE Int. Conf. on Pattern Recognition*, 2016.
- [30] J. Lu and Y.-P. Tan, "Gait-based human age estimation," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 4, pp. 761–770, 2010.
- [31] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 509–522, 2002.
- [32] M.-P. Dubuisson and A. K. Jain, "A modified Hausdorff distance for object matching," in *Proc. of the IEEE Int. Conf. on Pattern Recognition*, vol. 1, 1994, pp. 566–568.
- [33] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, Feb 2001.
- [34] <https://ep.probeinternational.org/2010/01/22/whole-body-active-millimeter-wave-imager/>.
- [35] E. Gonzalez-Sosa, R. Vera-Rodriguez, J. Fierrez, and V. M. Patel, "Millimetre wave person recognition: hand-crafted vs. learned features," in *IEEE Int. Conf. on Identity, Security and Behavior Analysis*, 2017, pp. 1–9.