Simulation of Millimeter Wave Body Images and its Application to Biometric Recognition

Miriam Moreno-Moreno\textsuperscript{a}, Julian Fierrez\textsuperscript{a}, Ruben Vera-Rodriguez\textsuperscript{a} and Josep Parron\textsuperscript{b}

\textsuperscript{a}Biometrics Research Lab.-ATVS, EPS, Universidad Autonoma de Madrid, Campus de Cantoblanco, C/ Francisco Tomas y Valiente 11, 28049 Madrid, Spain
\textsuperscript{b}Antenna and Microwave Systems Group-AMS, Universitat Autonoma de Barcelona, Campus UAB, Bellaterra, 08193 Barcelona, Spain

ABSTRACT

One of the emerging applications of the millimeter-wave imaging technology is its use in biometric recognition. This is mainly due to some properties of the millimeter-waves such as their ability to penetrate through clothing and other occlusions, their low obtrusiveness when collecting the image and the fact that they are harmless to health. In this work we first describe the generation of a database comprising 1200 synthetic images at 94 GHz obtained from the body of 50 people. Then we extract a small set of distance-based features from each image and select the best feature subsets for person recognition using the SFFS feature selection algorithm. Finally these features are used in body geometry authentication obtaining promising results.

Keywords: Millimeter wave images, 94 GHz, image simulation, synthetic database, biometrics, image processing, body geometry, distance-based features, feature extraction, feature selection, SFFS algorithm, verification.

1. INTRODUCTION

Biometric Recognition is the process that allow to associate an identity with an individual automatically, using one or more physical or behavioral characteristics that are inherent in the individual.\textsuperscript{1,2} Many of these biometric characteristics used to identify individuals, such as ear, face, hand, and gait, are extracted from images acquired by 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.). In order to circumvent these limitations, researchers have proposed the use of images acquired at others spectral ranges: X-ray, infrared, millimeter (MMW) and submillimeter (SMW) waves.\textsuperscript{3} The images captured beyond the visible spectrum overcome, 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:\textsuperscript{4} ability to pass through cloth and other occlusions, its health safety, its 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 is only one published research work that performs biometric recognition using that kind of images, using a holistic recognition approach.\textsuperscript{5} 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.

For all above, in this contribution it is first described the generation and content of BIOGIGA, a new database composed of simulated images of people at 94 GHz. Then, we present a new biometric system based on that...
### Table 1. Main parameters and body measures for each subject.

<table>
<thead>
<tr>
<th>Main Parameters</th>
<th>Corporal Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Neck circ.</td>
</tr>
<tr>
<td>Age</td>
<td>Height</td>
</tr>
<tr>
<td>Tone</td>
<td>Upper arm circ.</td>
</tr>
<tr>
<td>Weight</td>
<td>Upper arm length</td>
</tr>
<tr>
<td>Height</td>
<td>Lowerarm length</td>
</tr>
<tr>
<td></td>
<td>Wrist circ.</td>
</tr>
<tr>
<td></td>
<td>Front chest dist.</td>
</tr>
<tr>
<td></td>
<td>Burst circ.</td>
</tr>
<tr>
<td></td>
<td>Underburst circ.</td>
</tr>
<tr>
<td></td>
<td>Waist circ.</td>
</tr>
<tr>
<td></td>
<td>Nape to waist</td>
</tr>
<tr>
<td></td>
<td>Waist to hip</td>
</tr>
<tr>
<td></td>
<td>Shoulder to Neck</td>
</tr>
<tr>
<td></td>
<td>Upperleg height</td>
</tr>
<tr>
<td></td>
<td>Lowerleg height</td>
</tr>
<tr>
<td></td>
<td>Calf circ.</td>
</tr>
<tr>
<td></td>
<td>Angle circ.</td>
</tr>
<tr>
<td></td>
<td>Thigh circ.</td>
</tr>
<tr>
<td></td>
<td>Hips circ.</td>
</tr>
</tbody>
</table>

kind of images totally different and complementary to the only one published so far. Our system is based on the geometry while the other one was based on a holistic approach over the grayscale after various dimensionality reduction techniques.5 Finally, some verification experiments are performed to determine the performance and behavior of the developed system. An analysis of the features extracted and selected from the images is also carried out.

The paper is structured as follows. The generation, baseline corpus and the main characteristics of the used database are presented in Sect. 2. Sect. 3 presents the developed biometric system, describing all its modules: from the preprocessing of the input images to the identity decision. The evaluation of the biometric system is performed in Sect. 4, where the results of some verification and feature selection experiments are shown and analyzed. Conclusions are finally drawn in Sect. 5 together with future work.

## 2. BIOGIGA DATABASE

The corpus of the BIOGIGA database consists of synthetic images at 94 GHz of the body of 50 individuals. The images are the result of simulations carried out on corporal models at two types of scenarios (outdoors, indoors) and with two kinds of imaging systems (passive and active). These corporal models were previously generated based on body measurements taken from the subjects.

### 2.1 Generation of the Database

The generation of the database was carried out jointly at the Universidad Autonoma de Madrid (UAM) and Universidad Autonoma de Barcelona (UAB).6 The process to generate the images that compose the database can be divided into the following stages: 1) Acquisition of body measures of each person, 2) Generation of the 3D body model of each user based on such measures, and 3) Simulation of the images at 94 GHz from each model in the two types of scenarios (indoor/outdoor) and two types of capture systems (passive/active).

### 2.1.1 Corporal Measures

In order to have realistic 3D body models, a set of body measures of each person was taken. These measures are specified in Table 1. Some pictures of each subject were also taken in the visible spectrum with a conventional digital camera).

### 2.1.2 Corporal Model

The body model of each subject was obtained using the free software MakeHuman*, which allows you to generate images of human bodies in 3D. For each person his or her main parameters and body measurements (see Table 1) were introduced through the menus available for it. Subsequently the 3D model was compared with the picture previously taken to verify that the obtained model corresponds to the body the concerned person.

Additionally, MakeHuman generates garments adapted to the body of each model. This allows us to generate, in the following stage, images of people with and without clothes at 94 GHz.

---

*http://www.makehuman.org/
2.1.3 Simulation of Images at 94 GHz

The models generated in MakeHuman are imported to Blender\(^1\), which is a free software devoted to modeling, animation and 3D graphics creation. Blender provides a ray tracing engine that treats the radiation at 94 GHz as if it were light interacting with the scene. For a good simulation the properties of materials (skin, clothes ...) and light sources should be defined according to their characteristics in the range of 94 GHz.\(^7\)

Lastly, the resulting images were postprocessed adding them the typical noise and resolution that the real system to be simulated has.\(^7\)

2.2 Database Description

The database consists of images simulated at 94 GHz, of 50 people (25 females and 25 males), with ages between 15 and 55 years old. For each user the database has four sets of images, each of them simulated by:

- A passive system outdoors (PO).
- A passive system indoors (PI).
- An active system outdoors (AO).
- An active system indoors (AI).

For each user and each of the previous sets the following data was generated:

- Images of the human 3D model with clothes and an angle formed by the camera of -10, 0 and +10 degrees.
- Images of the human 3D model without clothes and an angle formed by the camera of -10, 0 and +10 degrees.

According to what is stated above, for each user the database has \(4 \times 2 \times 3 = 24\) images at 94 GHz. Hence, the total number of images in BIONIGA is \(50 \times 24 = 1200\). Fig. 1 shows some of the images of one subject. In this contribution, only images obtained from passive imaging systems outdoors are considered.

---

\(^1\)http://www.blender.org/
3. SYSTEM DESCRIPTION

3.1 Introduction to the Biometric Systems

A biometric system is the combination of devices, databases and processes that allow the recognition of an individual using one of his/her biometric features. The block diagram of a generic biometric system is shown in Fig. 2

Once the user has been enrolled, biometric systems can operate in two modes: 1) verification mode (the user claims an identity and presents his or her biometric trait, then the system determines if the real identity of the user corresponds to the identity claimed by him or her) and 2) identification mode (the user only present his or her biometric trait without claiming any identity, then the system tries to find his or her identity).

In the acquisition stage the biometric trait is captured. In our case, the biometric trait is the MMW image of the body. After that, the image is processed obtaining a more suitable image to perform the feature extraction, which is the following step. Next, the vector containing the features of the input image is compared with the feature vector of the claimed user (verification mode) or with all the feature vectors in the database (identification mode). Finally, a decision based on the result of that comparison is made.

The acquisition stage (which, in this work, corresponds to the simulation of the MMW images) was already explained in Sect. 2. The rest of the modules of the system are described in the next subsections: the image segmentation (Subsect. 3.2) belongs to the preprocessing while the boundary extraction (Subsect. 3.3), landmark generation (Subsect. 3.4), feature generation (Subsect. 3.5) and feature selection (Subsect. 3.6) are part of the feature extraction module. In the Subsect. 3.7 (matching) the similarity measure used to perform the comparison between feature vectors is introduced. The value of that similarity measure determines the final decision given by the system (Subsect. 3.8).

It is worth mentioning that the features extracted from the images are based on distances between points of the silhouette of the body. The process followed to obtain these features is depicted in Fig. 3.

3.2 Image Segmentation

The first step is to binarize the image, 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:

- Border detection.
- Morphological operations.

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 consists 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.

An example image after this segmentation step is shown in Fig. 3(b)
3.3 Boundary Extraction

Once the input image is properly segmented, only the largest connected component is considered, assuming that it is the body. Then, the body boundary is extracted. After that, the middle point of the top of the head is detected. This point is fixed to be the first point of the extracted contour. The rest of the coordinates of the boundary are ordered in a clockwise direction. In addition to the boundary of the body, a basic skeleton of the body is obtained by means of morphological operations.

3.4 Landmark Generation

Six different reference points are first considered: (i) middle point of the top of the head, (ii) the crossing point of the arms line and the vertical line of the torso, (iii) the centroid of the body, (iv) the bifurcation of the skeleton in the abdomen area, (v) the central point of a bounding box including the whole body, and (vi) the pubis. An example image obtained after boundary extraction and the reference points detection is depicted in Fig. 3(c).

For each reference point, the Euclidean distance between the reference point and every point of the boundary is computed. Therefore, a one-dimensional function, showing the radial distance, is obtained for each reference point. An example of the six resulting radial distance functions is shown in Fig. 3(d). Every function is examined to find local maxima and minima. Maxima of the curve correspond to the head, hands, and feet outer points, while the minima correspond to points near the neck, shoulders, axilla, wrist and hips, depending on the considered reference point. Fig. 3(e) shows an example boundary together with the reference points and the maximum and minimum distance points.

In order to have enough geometric measures of the body, several extra points are detected inside the area of the body and in its boundary. To summarize, the following points are detected (see Fig. 3(f)):

- The centroid of some parts of the body: head, arms, torso and legs.
Figure 4. Body boundary, landmarks and distances extracted among them of the subjects and conditions of Fig. 3. These distances form the 21D feature vector. The number next to every distance is the component number in the feature vector. The ones whose number component appears with a subindex \( m \) are calculated as the average of that distance and the analog one of the right extremity. Red triangles represent the reference points and centroids of head, arms, legs and torso. Blue circles represent landmarks on the silhouette of the body (some of them have minimum distance to the reference points). Green circles represent landmarks with maximum distance from the reference points.

- Some points located at the boundary of the above mentioned body parts (for example in case of the torso, the points of the torso boundary located at three different heights are detected: at one quarter, at a half and at three quarters of the torso height.)

### 3.5 Feature Generation

Once all the landmarks are obtained, the Euclidean distance between some of them is calculated. Specifically, 21 distances are obtained, which constitutes the feature vector. Fig. 3(g) and Fig. 4 show the considered distances. In Fig. 4, next to every distance there is a number that represents the component number in the feature vector (e.g. the height is the first component of the feature vector, it is feature number 1).

### 3.6 Feature Selection

In order to find the most discriminative set of distance-based features, and therefore increase the performance of the biometric system, a feature selection is performed. In addition, by reducing the number of features the computational cost decreases too.

Among the different feature selection algorithms,\(^8\) the one employed in this work is the Sequential Floating Forward Selection (SFFS).\(^9\) This suboptimal searching technique is an iterative process in which, in each iteration, a different set of features is used to compute a certain criterion. This is done until the criterion does not improve. For more details see\(^8\) and\(^9\) In our case the criterion is related to the performance of the system.

### 3.7 Matching

Once the feature vector of an image is computed, either containing the 21 components or with only the components selected by SFFS, a similarity measure is computed. This similarity measure between two vectors is called score. In our biometric system the score is the opposite of the Euclidean distance between the two compared vectors. The reason for using the opposite is in order to have higher scores for more similar vectors.
3.8 Decision

The obtained scores are used to (i) decide whether the user is who he or she is claiming or not (verification) or to (ii) provide a list with the most likely identities (identification). In both cases, if the user’s score is higher than a certain threshold, the user is accepted as a genuine user.

4. EXPERIMENTS

4.1 Performance Evaluation of a Biometric System

Once the system is implemented, an evaluation of its performance is desirable. This is done through quantitative measures, obtained from experiments, that reveal how well the system works and allow us to compare its performance with the performance of other biometric systems.

Although the developed system works in both kind of operational modes (verification and identification), the experiments and results shown below are obtained only when working in verification mode. In that case two main metrics are used to quantify the performance of the biometric system: 1) the False Acceptance Rate (FAR), which is the probability that an impostor is accepted by the system as a genuine one, and 2) the False Reject Rate (FRR), that reflects the probability that a genuine user is rejected by the system.

As the threshold value increases, the FRR increases too but the FAR decreases (the system becomes more secure by rejecting more genuine users and accepting less impostors). There is a value of the threshold for which the FAR and the FRR coincides. The value of both rates at that threshold is known as the Equal Error Rate (EER) and is the most widely used parameter to evaluate the performance of a biometric system working in verification mode. Plotting the FRR versus the FAR for all the different threshold values, gives the so-called DET curve (Detection Error Tradeoff curve). This curve is an common way to graphically represent the behaviour of a biometric system, where the EER corresponds to the point where the DET curve and the bisector intersect each other (See Fig. 5(a)). The closer the curve is to the origin (and therefore the lower the EER is), the better the system is.

4.2 Experimental Protocol

Although the BIOGIGA database contains 6 PO, 6 PI, 6 AO and 6 AI images per user, in the following analysis only PO and PI images are considered. The graphical results will be shown exclusively for PO images, for PI images the results are quite similar. The same analysis for images simulated by active systems (AO and AI images) will be part of future work, in which a different image segmentation should be followed due to clearly visible differences between the images simulated by passive and active systems (e.g. the last ones present higher contrast).

We consider three different experimental protocols: 1) protocol 1:3, 2) protocol 2:3, and 3) 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 our 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).

A mathematical model for each user is generated in the training and in the test set. Each model is simply the arithmetic mean feature vector of the considered images. Finally, the user test model is compared to the training model of each user in the database obtaining the corresponding score.

The experiments are carried out following the previous protocols in two ways: (i) considering all the components in the feature vector of the images, and (ii) after applying the SFFS algorithm, considering only the features selected by it. In short, there will be 6 different kind of experiments: each of then corresponds to a protocol working with all features or only with the selected ones.
Table 2. Verification and SFFS feature selection results.

<table>
<thead>
<tr>
<th>Protocol</th>
<th>With all features</th>
<th>With features selected by SFFS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EER (%)</td>
<td>Number of selected features</td>
</tr>
<tr>
<td>1:3</td>
<td>2.0000</td>
<td>6</td>
</tr>
<tr>
<td>2:3</td>
<td>2.0000</td>
<td>4</td>
</tr>
<tr>
<td>3:3</td>
<td>2.0000</td>
<td>3</td>
</tr>
</tbody>
</table>

4.3 Results

The Table 2 includes the values of the EER for each experiment, together with the features selected by the SFFS algorithm, as DET curves, for the 6 experiments performed.

The EER is quite low (2%) when all the 21 features of the feature vector are considered. The SFFS algorithm selects different features in each experiment obtaining improved values of the EER (less than 0.1%). In fact, in protocols 2:3 and 3:3 the EER becomes zero, therefore, the corresponding DETs are only a dot in the origin in those cases.

Regarding the feature selection, first, it is worth mentioning that, according to the Table 2, as the number of training images increases, the number of selected features that optimizes the EER decreases. Bearing in mind The Curse of Dimensionality\textsuperscript{11} this behavior is somehow surprising. According to this principle, a higher amount of training data allow to generate a more complex model (a higher dimension model, or a vector with more features), obtaining a better performance. However, our results show that, when there are more training data available, we obtain better results by considering less features. This could be due to the fact that, when there are more training data, the most discriminative features selected by the SFFS algorithm (mainly features #1, #16 and #2) become more relevant over the rest, while the rest of features are discarded in the feature selection process.

Secondly, the features selected most frequently in the three experiments are the height (feat.#1), the waist width (feat.#17), the wingspan (feat.#2) and the chest width (feat.#16) (they are represented with its number in Fig. 4 together with the rest of the features). The SFFS selects the most discriminative set of features for each experiment. That is, a group of uncorrelated features that optimizes the verification performance for an specific protocol. However, when considering independently each feature, we obtain that the most discriminative

![Figure 5. (a) DET curves for the three protocols, using all the features or the ones selected by the SFFS algorithm. (b) Mean normalized value of each feature for 7 randomly selected users, the features have been ranked in descending order of their discrimination power.](image)
ones are: the height (feat.#1), the waist width (feat.#17), the waist-pubis distance (feat.#5) and the leg length (feat.#6).

The computation of the most discriminative individual features was done ranking them in descending order of their discrimination power, which can be considered as the quotient \(\text{average separation}/\text{overlap}\). The numerator is the average separation among the mean values of the feature for all subjects, while the denominator is the overlap between the errors bars (standard deviation) of the same feature among all the subjects. This is graphically plotted in Fig. 5(b) where the mean value and standard deviation (as error bar) of every feature is plotted for 7 users, having the features ordered in descending order of \(\text{average separation}/\text{overlap}\). The feature number (the same one as the one in Fig. 4) is written below the \(x\) axis with bold numbers. It can be seen the decreasing separation between mean values and the increasing overlap.

The four most discriminative individual features are the first four ones in the Fig. 5(b), that is features 1, 17, 5 and 6, as we have previously stated. The height (feat.#1) and the waist width (feat.#17) are selected by SFFS algorithm in the three protocols, but the waist-pubis distance (feat.#5) and the leg length (feat.#6) are not. This is due to the fact that, although they are quite discriminative as isolated features, when forming a discriminative set of features, they are discarded because they provide info already given by other features that were first selected. For example, in protocol 2:3 the selected features are the height (feat.#1), the waist width (feat.#17), the chest width (feat.#16) and the wingspan (feat.#2). The height and the wingspan are highly uncorrelated, in fact only with those two features it would be possible to distinguish different users. The chest and the waist width seem to be more correlated (both are widths of the torso), however, the images of the database reveal that they are not so correlated having people with the same chest width but different waist width (for instance women with a chest width similar to a man, tend to have less waist that men).

5. CONCLUSIONS

A new method to simulate images of people at 94 GHz has been proposed. This method first generates a 3D corporal model of the person from his or her body measures, afterwards the model is imported to Blender software, where the radiation at 94 GHz is treated as if it were light interacting with the scene, obtaining images of the person 3D model at that frequency. Images of 50 users were simulated at two types of scenarios (outdoors, indoors) and with two kinds of imaging systems (passive and active). These images comprise the BIOGIGA database, the first database of synthetic images of people at 94 GHz.

Next, a complete biometric system has been developed making use of the images of BIOGIGA. The use of MMW images instead of images acquired at other spectral bands present some advantages (mainly the transparency of clothing at that frequency range). In the development of the system we have proposed a feature extraction method based on the detection of relevant points of the body followed by the measure of 21 distances among them. These distances form the feature vector of every image. These vectors are compared in the matching stage obtaining a final decision: the user is accepted by the system or not (in the verification operational mode).

Finally some verification experiments have been conducted using three different protocols (with one, two or three train images per user). Firstly, all the components in the feature vector of the images are considered in the experiments and, secondly, only the features selected by the SFFS algorithm are taken into account. Also a simple analysis of the selected features was carried out. The obtained verification performance is quite high (EER around 2% or lower) and the use SFFS algorithm considerably improves the performance of the system (EER around 0.1% or lower). In addition SFFS reveals the most discriminative set of features, that does not necessarily contain the most discriminative individual features.

The limitations of this work are related to the special characteristics of the database used. The BIOGIGA images are limited when compared to the real images acquired in practice. However, the synthetic images are very similar to the ones really captured at 94 GHz. Also, the synthetic images used are based on real measures from people. Therefore, the proposed features can be directly applied and are useful for practical MMW images.

Finally, a similar system should be developed when using images simulated with active systems. Furthermore, these systems (based on simulated images) should be improved and tuned to operate with real images acquired with MMW imaging systems. In that case the segmentation module should be adapted while the rest would
remain the same. These systems will allow us to perform identification and/or verification experiments with real images and could be applied in the near future at airport security checkpoints and another screening scenarios.

ACKNOWLEDGMENTS

This work has been partially supported by projects Bio-Challenge (TEC2009-11186), Contexts (S2009/TIC-1485), TeraSense (CSD2008-00068) and “Catedra UAM-Telefonica”.

REFERENCES