

Article

Acoustic Biometric System Based on Preprocessing Techniques and Linear Support Vector Machines

Lara del Val ^{1,*}, Alberto Izquierdo-Fuente ^{2,†}, Juan J. Villacorta ^{2,†} and Mariano Raboso ^{3,†}

¹ Departamento de Ciencia de los Materiales e Ingeniería Metalúrgica, Expresión Gráfica de la Ingeniería, Ingeniería Cartográfica, Geodesia y Fotogrametría, Ingeniería Mecánica e Ingeniería de los Procesos de Fabricación, Área de Ingeniería Mecánica, Universidad de Valladolid, Paseo del Cauce 59, 47011 Valladolid, Spain

² Departamento de Teoría de la Señal y Comunicaciones e Ingeniería Telemática, Universidad de Valladolid, Paseo Belén 15, 47011 Valladolid, Spain; E-Mails: alberto.izquierdo@tel.uva.es (A.I.-F.); juavil@tel.uva.es (J.J.V.)

³ E.U. Informática, Universidad Pontificia de Salamanca, Calle Compañía 5, 37002 Salamanca, Spain; E-Mail: mrabosoma@upsa.es

† These authors contributed equally to this work.

* Author to whom correspondence should be addressed; E-Mail: lvalpue@eii.uva.es; Tel.: +34-983-184-443.

Academic Editor: Vittorio M.N. Passaro

Received: 6 May 2015 / Accepted: 10 June 2015 / Published: 17 June 2015

Abstract: Drawing on the results of an acoustic biometric system based on a MSE classifier, a new biometric system has been implemented. This new system preprocesses acoustic images, extracts several parameters and finally classifies them, based on Support Vector Machine (SVM). The preprocessing techniques used are spatial filtering, segmentation—based on a Gaussian Mixture Model (GMM) to separate the person from the background, masking—to reduce the dimensions of images—and binarization—to reduce the size of each image. An analysis of classification error and a study of the sensitivity of the error *versus* the computational burden of each implemented algorithm are presented. This allows the selection of the most relevant algorithms, according to the benefits required by the system. A significant improvement of the biometric system has been achieved by reducing the classification error, the computational burden and the storage requirements.

Keywords: acoustic biometric system; acoustic images; preprocessing techniques; support vector machine

1. Introduction

Biometric systems are based on the subject's characteristics to allow his/her identification [1]. The main biometric systems use elements such as fingerprints, retina, face, voice, *etc.* to characterize people and then classify them for subsequent identification and validation. Each of these systems requires the use of specific sensors to obtain the desired characteristics of the subject. Video cameras are often used as sensors to identify subjects or property although a radar could also be used to obtain the shape of a subject through reflection [2,3]. There are accurate and reliable classification systems based on acoustic radars:

- Animal echolocation, developed by mammals such as bats, whales and dolphins through specific waveforms [4,5], or the identification of different types of flowers by other species [6].
- Acoustic signatures used in passive sonar systems [7,8], which analyze the signal received by a target in the time-frequency domain.

There is little literature on the use of an acoustic radar as a biometric system for human identification and an ultrasonic band rather than an audible frequency band is usually employed [9,10]. In previous works, the authors developed multisensor surveillance and tracking systems based on acoustic arrays and image sensors [11,12]. In another line of work, making the most of the acquired experience in acoustic arrays and image sensors, the authors developed a biometric identification system based on the acoustic images acquired with an electronically scanned array [13]. The system tries to discriminate subjects in terms of their acoustic image, directly related to the subject's shape, height and geometrical characteristics. These characteristics are considered "soft biometrics" and they used to be used along with other "hard biometrics" (e.g., fingerprints) in order to uniquely identify a person.

The system obtained acoustic images by scanning the subjects in four frequencies of the acoustic band and in four different positions, defining an acoustic profile that comprises all of these images. Subsequently, the acoustic profile was compared to previously stored profiles to identify the subject. In this first system, Mean Square Error (MSE) between two images of the same frequency and position is used to compare the acoustic profiles, defining a global error as the sum of the errors associated with each image of the profile. Using the Equal Error Rate (EER) as a quality indicator, this system obtained an EER value of 6.22%, such as other emerging biometric identification systems [14–17].

In a later work [18], the authors analyzed the contribution of each acoustic image—associated with a frequency and position—to the performance of the biometric system, finding that each image provides different degrees of information. Two main conclusions were obtained:

- Each set of images associated to certain frequency provides different information, improving the system performance, thus, the number of frequencies used should be increased.
- Information associated to certain subject positions only provides redundant information and does not improve the quality of the system, thus, the number of positions used should be decreased.

In a second stage of the analysis, a new global error function was proposed by weighting the MSE error of each image proportionately to the information that it provides. In this case, an EER value of 4% was obtained. The use of more efficient classification algorithms would provide an improvement in the classification error, which also represents an EER.

Since Support Vector Machines (SVMs) are algorithms that currently define Machine Learning [19], it was decided to work with them in the classification tasks. Furthermore, SVMs are the unique algorithms used in the classification capable of working with high-dimensional data, such as the case of the acoustic profiles used.

This paper presents an improved biometric system that uses a SVM algorithm for classification and identification of subject. Since high dimensionality of acoustic profiles exponentially increases the computational burden of SVM classifiers, preprocessing and feature extraction techniques have been designed and implemented to improve the classifier performance. This new system is based on the results obtained in previous studies [18].

In Section 2, SVM classification algorithms and associated training techniques are explained. Section 3 describes the biometric system, including acquisition, preprocessing and classification systems. In Section 4, an analysis of the results is done and finally, Section 5 presents the final conclusions.

2. Support Vector Machines

SVMs carry out binary classification by constructing a hyperplane defined by the weight vector w and the bias term b , as shown in Figure 1, so samples of different classes will be divided by a separation, as wide as possible. Thereby, SVM algorithms are called maximum margin classifiers, being γ the margin of separation.

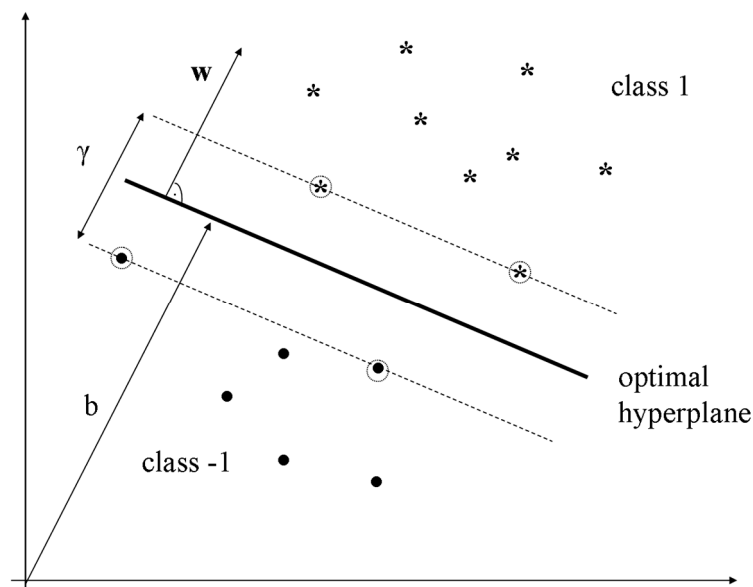


Figure 1. Hyperplane for binary classification.

Based on a training set of l known samples formed by data vectors x_i and the corresponding class labels y_i to which they belong:

$$(x_1, y_1), \dots, (x_l, y_l) \in R^N \times \{-1, 1\} \quad (1)$$

Machine Learning algorithms obtain the hyperplane according to an optimization criterion, which must be validated subsequently.

In the validation phase, the class label of a new data vector x can be predicted by projecting x in the weight vector w :

$$f(x) = w \cdot x + b \quad (2)$$

The sign of this projection will reveal the predicted class label. Thus, new samples are mapped into the n -dimensional space and a class will be associated to them, depending on which side of the hyperplane has been mapped.

There are different possible hyperplanes that divide the data space into two subsets. Typically, the maximum margin criterion is used as an appropriate optimization criterion to obtain the hyperplane with the greater margin of separation γ (see Figure 1). Only the vectors (or samples) positioned on the margin—which are called support vectors and that in Figure 1 are surrounded by a circle—are necessary to describe this hyperplane.

For a canonical representation of the hyperplane, the constraints $y_i (w \cdot x_i + b) \geq 1$ must be met to find the margin $\gamma = 2/\|w\|$. The maximization of margin γ is equivalent to the minimization of $(1/2) \|w\|^2$, subject to the same restrictions.

Violation of restrictions involves the introduction of the variable ξ_i , giving rise to the so-called problem of soft-margin SVM optimization:

$$\min \frac{1}{2} \|w\|^2 + C \sum_i \xi_i \quad s.t. \quad y_i (w x_i + b) \geq 1 - \xi_i, \quad \xi_i > 0 \forall i \quad (3)$$

C is the regularization parameter, so that higher values of C correspond to stronger violations penalties.

In order to resolve the problem showed in Equation (3), it is rewritten in terms of positive Lagrange multipliers α_i . In this way, it is required to maximize the following expression:

$$L_D = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \quad (4)$$

subject to restrictions $0 \leq \alpha_i \leq C$ and $\sum_i \alpha_i y_i = 0$, given the relation:

$$w = \sum_i^{N_s} y_i \alpha_i x_i \quad (5)$$

where N_s denotes the number of resulting support vectors. The discriminant function on which the SVM optimization is based is obtained by substituting w in Equation (2):

$$f(x) = \sum_i^{N_s} y_i \alpha_i (x \cdot x_i) + b \quad (6)$$

Essentially, a SVM is a two-class classifier but, in practice, it is very common to find problems associated with $K > 2$ classes. In these cases, a multiclass classifier is needed. There are several methods that combine multiple two-class SVM to obtain a multiclass classifier. The most widespread methods are the one-versus-all and the one-versus-one [19].

Training and Validation

The classifier learns from a training set—samples whose class labels are known—and defines a hyperplane. Then, this hyperplane is used to classify the samples from the validation set—whose classes are unknown. After that, the associated classes are compared with their corresponding classes and the error rate of the classifier is assessed, as shown in Figure 2.

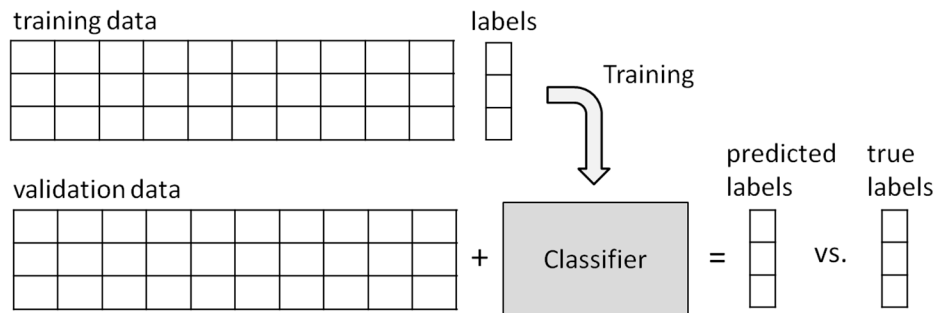


Figure 2. Classifier training.

The number of available samples is finite (N samples) and should be divided among the training set and the validation set. In this work, the classification algorithm is trained, using the two most common training methods:

- Leave-One-Out (LOO)
- Cross Validation (CV)

In LOO [20], training is carried out using $N - 1$ samples, and validation is performed using the sample which has been excluded. Errors are taken into account when the classification is wrong. This process is repeated N times, each time excluding a different sample. The total number of errors gives an estimation of the classification error rate.

On the other hand, the CV method involves taking the available data samples and dividing them into S groups (named folds) [21]. $S-1$ folds are used to train the model, and the remaining fold is used to validate. This procedure is repeated S times, taking a different fold each time to validate the model. Finally, the classification error rate is the average of the errors that have been obtained in each of the S runs. An example of a 5-fold cross-validation ($S = 5$) is shown in Figure 3, where the fold used to validate is highlighted.

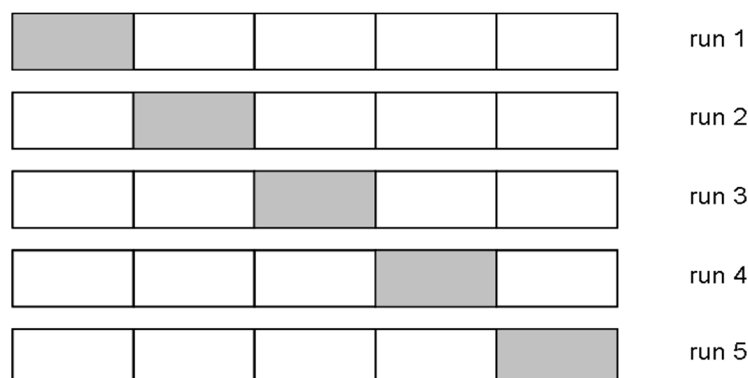


Figure 3. 5-fold Cross Validation.

3. System Description

Based on basic radar/sonar principles [22,23], an acoustic detection and ranging system for biometric identification was proposed [24], according to the block diagram in Figure 4.

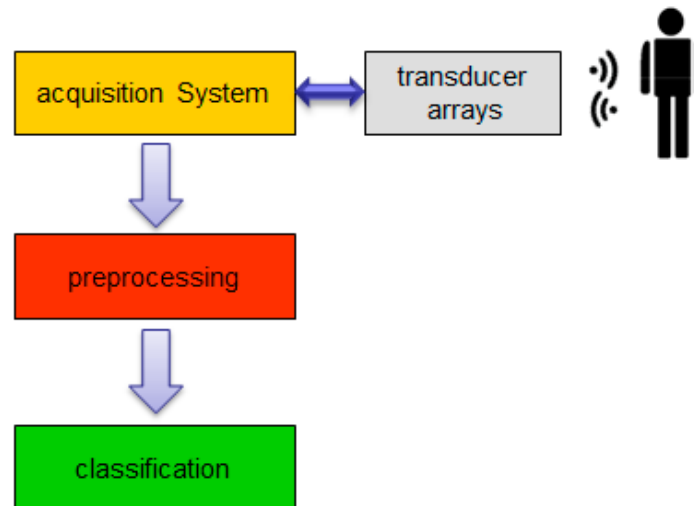


Figure 4. Functional description block diagram.

This system performs four main tasks: (i) subject scanning; (ii) acoustic images acquisition; (iii) images preprocessing and (iv) subject identification, based on classification algorithms.

3.1. Acquisition System

The subject is electronically scanned in the azimuth coordinates using two linear arrays. For each steering angle the system performs: (i) transmission beamforming; (ii) reception beamforming and (iii) match filtering in the range coordinate. After processing all the steering angles, a two-dimensional matrix is formed and stored, representing the acoustic image. Figure 5 shows the block diagram for the acquisition system.

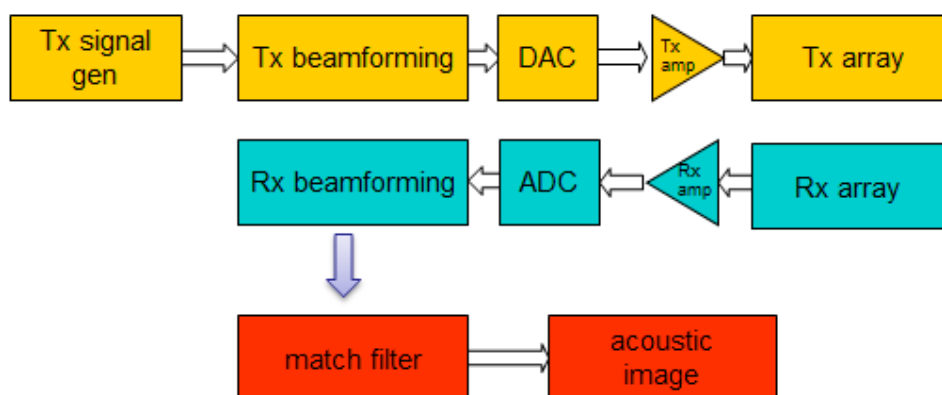


Figure 5. Acquisition system block diagram.

Figure 6 shows an example of an acoustic image, considering that the x axis represents the azimuth angle and the y axis, the range.

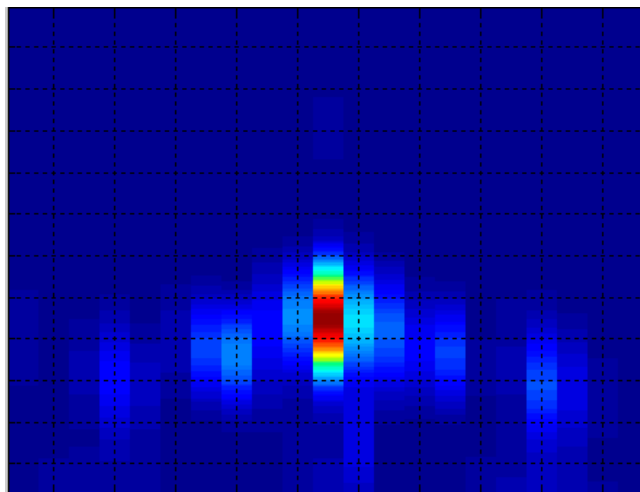


Figure 6. Acoustic image example.

Based on the conclusions of previous works [18], a new system that employs $P = 3$ spatial positions and $F = 9$ frequencies is defined. This system generates P_i acoustic profiles, associated to subject i and formed by $P \cdot F = 27$ images.

The selected positions for the subject under analysis are: front view with arms outstretched (p_1), back view (p_2) and side view (p_3). The nine frequencies are 500 Hz-spaced, from 8 kHz (f_1) to 12 kHz (f_9). The number of beams used for each frequency is shown in Table 1.

Table 1. Number of beams vs. frequency.

| f_1 | f_2 | f_3 | f_4 | f_5 | f_6 | f_7 | f_8 | f_9 |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 13 | 15 | 15 | 17 | 17 | 17 | 19 | 19 | 21 |

3.2. Preprocessing and Parametrization Techniques

With the purpose of reducing the dimension of the acoustic profiles, eliminating redundant or non-significant information and thus reducing the associated computational burden, several preprocessing and parametrization techniques on acoustic images have been evaluated. Among the preprocessing techniques, the following processes are implemented:

- spatial filtering
- segmentation using Gaussian Mixture Models algorithms
- masking
- binarization

On the other hand, a reduced set of parameters was extracted from the acoustic images in order to characterize them. Two families of algorithms were analyzed:

- line-based image coding
- geometric feature extraction

Figure 7 shows the processing scheme.

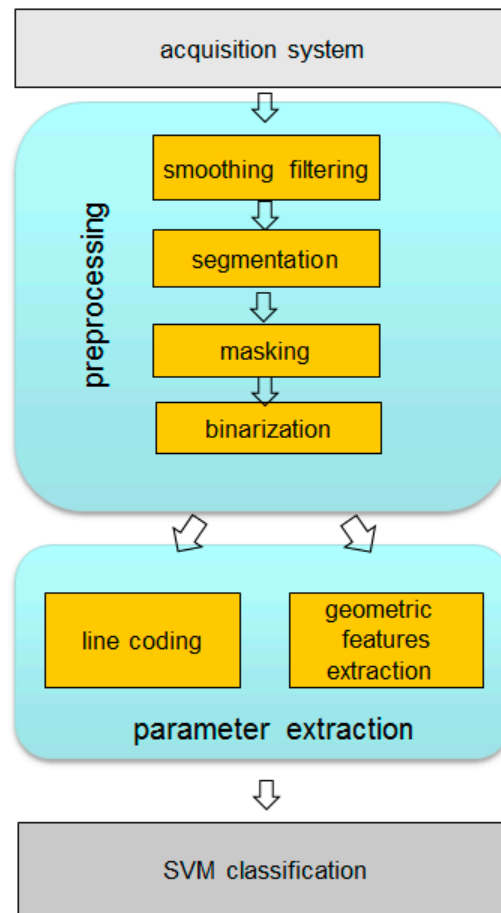


Figure 7. Preprocessing and parametrization techniques.

First, a spatial filter was implemented to smooth images, in order to reduce multimodalities of torso echoes and improve the segmentation process [10]. Then, a segmentation algorithm was used to differentiate pixels associated to the object from the pixels associated to the background.

The Expectation-Maximization (EM) algorithm is used to adjust a Gaussian Mixture Model (GMM) formed by two Gaussians, associated with foreground and background, respectively. The pixels associated with the background are zeroed.

The dimensions of the images $N \times M$ —where N is the number of rows (dimension in range) and M is the number of columns (dimension in azimuth)—are detailed for each frequency and position in Table 2.

Table 2. Image sizes.

| $N \times M$ | f_1 | f_2 | f_3 | f_4 | f_5 | f_6 | f_7 | f_8 | f_9 |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| p_1, p_2, p_3 | 245×13 | 245×15 | 245×15 | 245×17 | 245×17 | 245×17 | 245×19 | 245×19 | 245×21 |

The profiles formed by the acoustic images are stored to be processed and that is why the size of each pixel of the images has to be defined. The final size of the profiles gives the required storage space, and is related with the computational burden associated to the system. In this case, the value of each pixel is stored in memory using $B = 32$ bits.

Using masking techniques, the size of the images is reduced by adjusting them to the area that the subjects take up on the image. A statistical analysis of the acquired images was performed to determine

the common area for each position and frequency. The sizes of the images obtained by this technique for each frequency and position are detailed in Table 3.

Table 3. Masked image sizes.

| $N \times M$ | f_1 | f_2 | f_3 | f_4 | f_5 | f_6 | f_7 | f_8 | f_9 |
|--------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| p_1 | 145×13 | 145×15 | 145×15 | 145×17 | 145×17 | 145×17 | 145×19 | 145×19 | 145×21 |
| p_2 | 155×11 | 155×11 | 155×11 | 155×11 | 155×11 | 155×11 | 155×11 | 155×11 | 155×11 |
| p_3 | 171×9 | 171×9 | 171×9 | 171×9 | 171×9 | 171×9 | 171×9 | 171×9 | 171×9 |

Finally, the value of the pixels—encoded with 32 bits—is reduced to 1 bit by image binarization. This significantly reduces the storage space required. This operation is equivalent to associate a unit value to the foreground pixels. Figure 8 shows an example of the use of the preprocessing techniques, showing an original Figure 8a, a segmented Figure 8b, a masked Figure 8c and a binarized Figure 8d image.

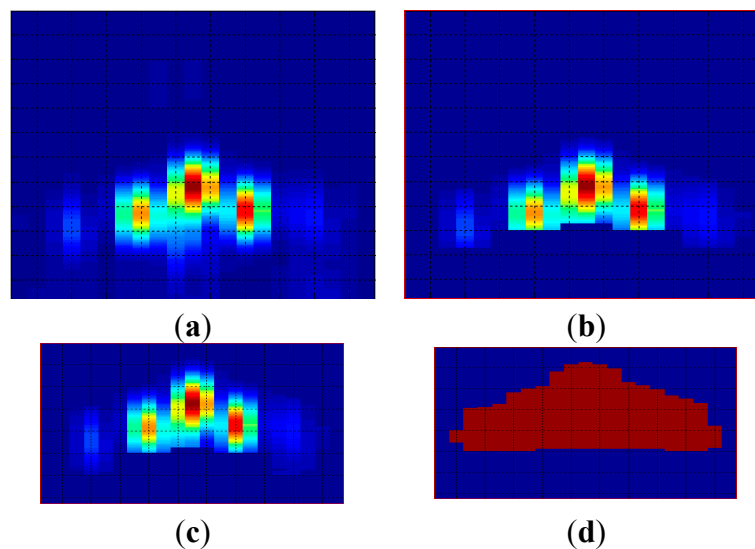


Figure 8. Pre-processed images: (a) original; (b) segmented; (c) masked; (d) binarized.

Table 4. Image sizes using Row-based Image Coding.

| L | f_1 | f_2 | f_3 | f_4 | f_5 | f_6 | f_7 | f_8 | f_9 |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| p_1 | 145 | 145 | 145 | 145 | 145 | 145 | 145 | 145 | 145 |
| p_2 | 155 | 155 | 155 | 155 | 155 | 155 | 155 | 155 | 155 |
| p_3 | 171 | 171 | 171 | 171 | 171 | 171 | 171 | 171 | 171 |

Starting from the binarized images, two feature extraction techniques were applied, significantly reducing the size of the acoustic images. First, Line-based Image Coding algorithms were analyzed. The images are broken down into a set of lines that can be rows or columns. For each line, the number of pixels with unit value is encoded. In this way, the size of each image is significantly reduced, from a $N \cdot M$ size to a L size, where L is N or M , as encoding is performed by rows or columns, respectively. The value of each parameter is stored in the memory using $B = 8$ bits. In row coding, sizes of the images obtained at each position and frequency are shown in Table 4.

In column coding, sizes of the images obtained for each frequency position are detailed in Table 5.

Table 5. Image sizes using Column-based Image Coding.

| L | f ₁ | f ₂ | f ₃ | f ₄ | f ₅ | f ₆ | f ₇ | f ₈ | f ₉ |
|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| p ₁ | 13 | 15 | 15 | 17 | 17 | 17 | 19 | 19 | 121 |
| p ₂ | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 |
| p ₃ | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 |

As an example, Figure 9 represents Line-based Image Coding using rows and columns of an acoustic image of size 6 × 12.

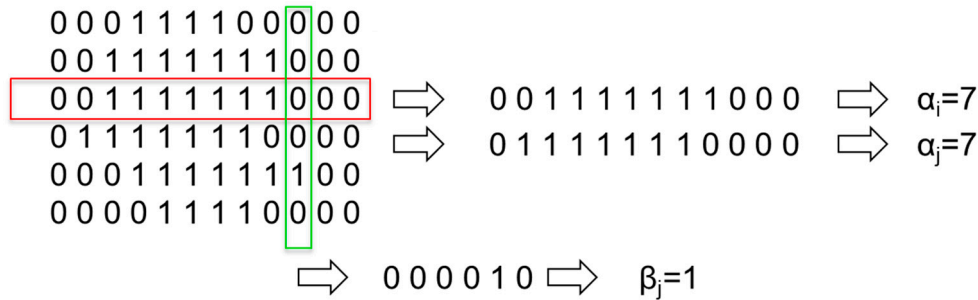


Figure 9. Line-based Image Coding.

In Figure 9, it can be observed that rows 3 and 4 gave an identical encoding, although they are different. To improve the information of each line and avoid ambiguous encodings, a second parameter that stores the starting position of the first nonzero pixel per line is added. With this improvement, the image dimension is doubled. As an example, Figure 10 represents the new encoding methods for the previous image.

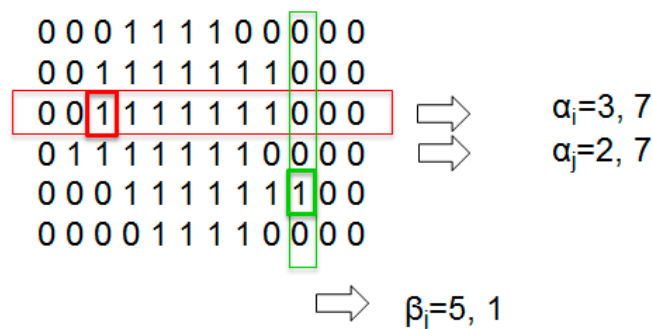


Figure 10. Line-based Image Coding with position.

Secondly, geometric feature extraction algorithms were analyzed. They show the following properties of images:

- Area : A
- Centroid: (c_x, c_y)
- Perimeter: P

In this case, one parameter for area and perimeter, and two parameters for centroid are extracted from each image. The value of each parameter is stored in the memory using B = 32 bits. Figure 11 shows the geometric features extracted from Figure 8d.



Figure 11. Geometric feature extraction.

3.3. Classification

In this work, tests based on linear SVM algorithms were performed. A linear SVM was used because, if the number of features is large, one may not need to map data of a higher dimensional space and using the linear kernel is good enough [25]. Besides, although the dimension of the acoustic images is reduced with the preprocessing techniques, it is still too high in order to be used in SVMs based on a Gaussian kernel, involving an increment of the processing time and the computational burden, without an improvement in the classification error rate.

It was implemented with Matlab, specifically using the LIBSVM library, which allows multiclass SVM classification according to the *one-versus-one* algorithm [26]. The methods LOO and CV, with 10, 5 and 4 folds were used for algorithm training.

In the linear SVM, the regularization parameter C was set to 5000, since a usual practice is to assign C to the range of output values of the SVM algorithm [27], *i.e.*, the maximum number of possible errors, which coincides with the total number of samples.

4. Analysis of Results

4.1. Scenario Definition

This study assumes that the system is used as an access control to enter in a laboratory, where only five subjects have authorized access. The SVM algorithm must be able to classify the subjects who try to access the laboratory in six different classes:

- a class for each of the five authorized subjects
- a class associated to all other people, considered as intruders.

To evaluate the performance of the SVM classification algorithm in acoustic biometric system, 5000 profiles were used. They are divided into six classes according to the following distribution:

- 500 acoustic profiles for each of the five authorized people.
- 2500 acoustic profiles for 25 intruders.

In order to have a population sample as general as possible, the subjects whose acoustic images were used have different morphological characteristics, as shown in Table 6. In this case, unlike previous tests [13,18], acoustic images of each subject were obtained on different days and with them wearing different clothes. In this way, the system will be able to classify the subjects according to who they actually are, without clothes being a distinctive factor.

Table 6. Morphological features.

| Id | # Signatures | Gender | Constitution | Height |
|-------------------|---------------------|---------------|---------------------|---------------|
| Authorized | | | | |
| 00 | 500 | male | thin | average |
| 01 | 500 | female | normal | average |
| 02 | 500 | female | normal | small |
| 03 | 500 | male | strong | tall |
| 04 | 500 | male | very strong | tall |
| Intruders | | | | |
| | 125 | male | strong | average |
| | 150 | female | thin | average |
| | 150 | male | thin | average |
| | 300 | female | normal | average |
| | 100 | male | very strong | tall |
| | 125 | male | normal | average |
| 05-29 | 125 | male | strong | tall |
| | 100 | female | normal | small |
| | 125 | female | strong | small |
| | 125 | female | strong | tall |
| | 325 | male | strong | average |
| | 350 | male | thin | small |
| | 400 | female | thin | small |

Based on this scenario, a set of experiments were conducted to analyze the performance of the proposed classification algorithm using different acoustic profiles. Figure 12 shows experiments that were carried out in this study. The system was tested with raw, preprocessed, binarized, line-based encoded and geometric features extracted profiles.

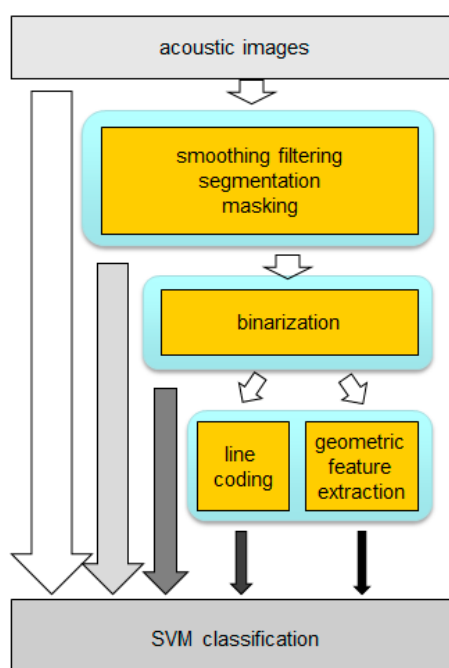


Figure 12. Experiments.

4.2. Raw Profiles

These acoustic profiles use raw images, without preprocessing. Each of the profiles has a size in the order of 3.5×10^6 ; this value is obtained from the following Equation (7):

$$\text{size}_{\text{profile}} = B \cdot \sum_{i=1}^P N_i \cdot \sum_{j=1}^F M_{i,j} \quad (7)$$

where B is the number of bits used to store the value of each pixel of the acoustic images, P is the number of positions used in the system, N_i is the number of rows for each position, F is the number of frequencies and $M_{i,j}$ is the number of columns for each position and frequency. The specific values of these variables are shown in Section 3.2.

In this test, an average classification error rate of 0.46% with a standard deviation of 0.120 was obtained. Comparing the error rate obtained using SVMs, which represents an Error Equal Rate, with the EER obtained with the classifier based on mean squared error (MSE), the classification error rate was reduced significantly, *i.e.*, from 4% [18] to 0.46%.

4.3. Preprocessed Profiles

In this case, raw profiles were first filtered, then segmented via GMM and, finally, masked, as explained in Section 3.2. Now, each preprocessed profile has a storage size in the order of 1.6×10^6 . This value was obtained using Equation (7).

A mean error classification rate of 0.46% with a standard deviation of 0.121 was obtained. Comparing these results with those obtained using raw profiles, it can be observed that when the profile size is reduced-equivalent to reducing computational burden-the error rate does not change. This shows that eliminated data does not provide relevant information to the classification task.

After that, the preprocessed profiles were binarized. Each profile has now a size in the order of 4×10^5 . In this case, an error rate of 0.75% with a standard deviation of 0.255 was obtained. If these results are compared with those obtained using preprocessed profiles without binarization, it can be observed that size reduction slightly increased the classification error rate, however, at a lower ratio than the reduction in size.

This shows that the most relevant data of the profiles is related to the shape of the subjects, not to the specific value of their pixels. This is why working with binarized profiles is the next step to achieve size reduction. Table 7 summarizes the results obtained in these tests, corresponding to the classification based on raw, preprocessed and binarized acoustic profiles.

Table 7. Classification error rates for raw, preprocessed and binarized acoustic profiles.

| Acoustic Profile | Error Rate | σ (Error Rate) |
|------------------|------------|-----------------------|
| Raw | 0.46% | 0.120 |
| Preprocessed | 0.46% | 0.121 |
| Binarized | 0.75% | 0.255 |

4.4. Parameter Extraction

Line-Based Image Coding

First, line-based image coding was done using the length of rows or columns of the acoustic images. Then, the algorithm was improved, including the initial position of rows and columns. The size of the acoustic profiles obtained through this line coding is calculated using Equation (8):

$$\text{size}_{\text{profile}} = B \cdot K \cdot \sum_{i=1}^P \sum_{j=1}^F L_{i,j} \quad (8)$$

where B is the number of bits used to store the value of each parameter, K is the number of parameters used to encode each line—whose possible values are shown in Table 8—P and F are the number of positions and frequencies, respectively, and $L_{i,j}$ is the number of lines, *i.e.*, rows, columns or the sum of both. The specific values of these variables are shown in Section 3.2.

Table 8. Number of parameters per line.

| Line-Based Image Coding | K |
|-----------------------------------|---|
| Based on Line Length | 1 |
| Based on Line Length and Position | 2 |

In the first case, profiles have a size in the order of 3×10^4 , 2.6×10^3 and 3.6×10^4 , when row, columns and both rows and columns are encoded. The results obtained in these tests are shown in Table 9. Line coding based on length reduces considerably the size of the acoustic profile, although classification error rate increases. On the other hand, despite error rate getting worse, its values are still acceptably low. The lowest error rate value is obtained using both row and column coding, with a mean value of 1.43% and a standard deviation of 0.390.

Table 9. Classification error rates for line coding based on line length.

| Line Coding | Error Rate | σ (Error Rate) |
|--------------|------------|-----------------------|
| Row | 1.93% | 0.498 |
| Column | 1.97% | 0.546 |
| Row + Column | 1.43% | 0.390 |

In the second case, profiles are twice as large as the ones in the first case, since the initial position of the line is also encoded. These sizes are in the order of 6.8×10^4 , 5.3×10^3 and 7.3×10^4 , if rows, columns or both rows and columns are encoded. In this case, although the profile dimension increases, the classification error rate decreases. If both row and columns are encoded according to length and position, an error rate of 0.47% is obtained. This error rate value is similar to the one obtained when using raw profiles. The obtained results are shown in Table 10.

Table 10. Classification error rates for line coding based on length and position.

| Line Coding | Error Rate | σ(Error Rate) |
|--------------------------------|-------------------|--|
| Row + Position | 1.47% | 0.407 |
| Column + Position | 1.86% | 0.391 |
| Row + Column + Position | 0.46% | 0.061 |

4.5. Geometric Feature Extraction

For this set of tests, the size of the profiles obtained by extracting geometric features of the acoustic images is calculated using Equation (9):

$$\text{size}_{\text{profile}} = B \cdot K \cdot P \cdot F \quad (9)$$

Where B is the number of bits needed to store the value of each parameter, K is the number of extracted features; P and F are the number of positions and frequencies, respectively.

The sizes of these acoustic profiles are in the order of 8.6×10^2 , if area or perimeter are used as geometric feature, and 1.7×10^3 , if centroid is employed. It can be observed that using geometric features reduces profile size, but the classification error rate increases excessively, as it is shown in Table 11. The obtained error rates are between 11% and 15%. These values are considerably higher than the reference error rate of 4% [18].

Table 11. Classification error rates using geometric features.

| Geometric Parameters | Error Rate | σ(Error Rate) |
|------------------------------------|-------------------|--|
| Area | 11.07% | 0.297 |
| Centroid | 12.04% | 0.310 |
| Perimeter | 15.04% | 0.325 |
| Area + Centroid + Perimeter | 6.86% | 0.430 |

Aiming to reduce error rate, the employed features were joined to form new profiles. In this case, although the profile size increases, the obtained error rate is reduced to 6.86%, as Table 11 shows. This value of the error rate is above the reference one.

4.6. Results Discussion

Figure 13 shows the classification error rate obtained for each test, and Figure 14 shows the corresponding computational burden. This computational burden is calculated as the product of the number of support vectors, employed by the SVM for the classification, and their size.

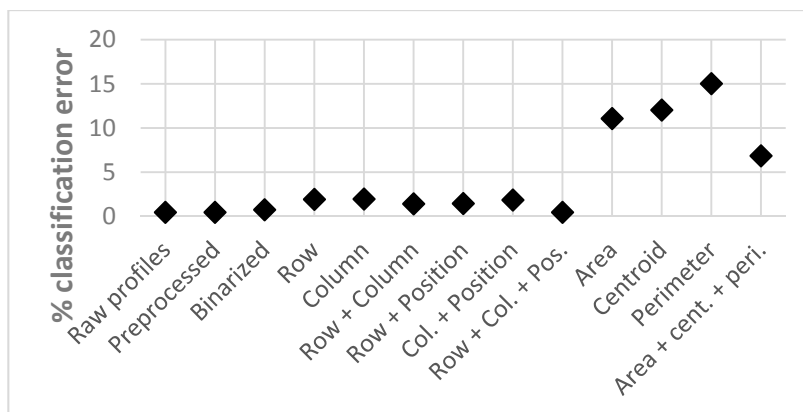


Figure 13. Classification error rates.

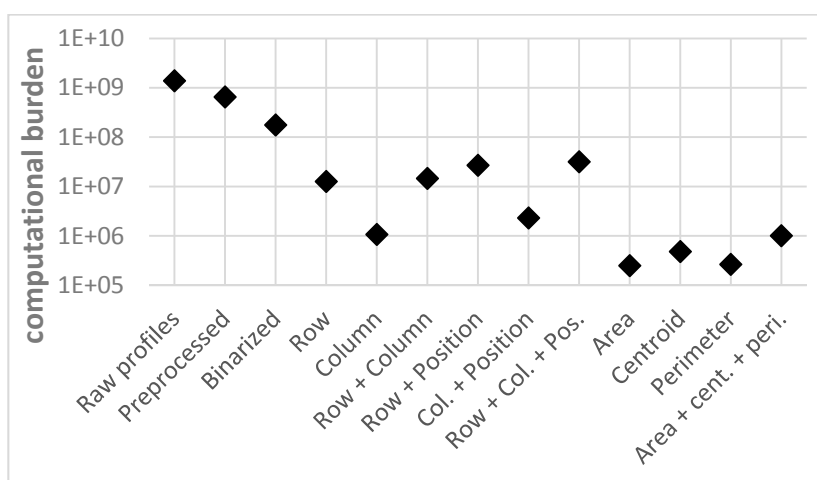


Figure 14. Computational burden.

In order to analyze these parameters and their relationship, classification error and computational burden sensitivities were defined as follows:

- Classification error sensitivity

$$S_e = \frac{(pre)processed_error}{raw_error} \tag{10}$$

with *raw_error* as the error using raw profiles.

- Computational burden sensitivity

$$S_b = \frac{(pre)processed_burden}{raw_burden} \tag{11}$$

with *raw_burden* as the burden using raw profiles.

Error sensitivity shows how the error rate increases due to profile size reduction. In the same way, burden sensitivity shows how burden decreases due to profile size reduction. Given that S_b values are always lower than 1, $1/S_b$ has been analyzed in order to compare both sensitivities in a similar way. Sensitivity values are shown in Table 12.

Table 12. Classification error and computational burden sensitivities.

| | S_e | $1/S_b$ |
|-----------------------------|-------|---------|
| Raw Profiles | 1.00 | 1.00 |
| Preprocessed | 1.00 | 2.13 |
| Binarized | 1.62 | 7.85 |
| Row | 4.17 | 109.72 |
| Column | 4.25 | 1297.28 |
| Row + Column | 3.09 | 95.02 |
| Row + Position | 3.17 | 51.43 |
| Col. + Position | 4.02 | 599.31 |
| Row + Col. + Pos. | 1.00 | 43.90 |
| Area | 23.94 | 5581.96 |
| Centroid | 26.02 | 2883.10 |
| Perimeter | 32.51 | 5242.42 |
| Area + Cent. + Peri. | 14.83 | 1382.36 |

Figure 15 shows $1/S_b$ versus S_e in order to evaluate the relationship between error increment and burden reduction.

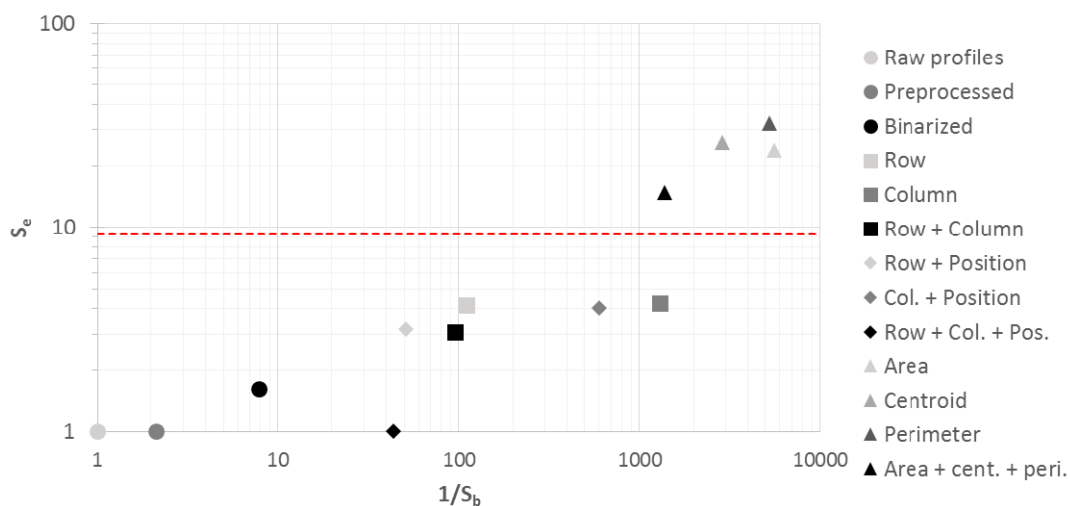


Figure 15. Error increment vs. burden reduction.

The dashed line in Figure 15 represents a reference S_e value of 9.30, which corresponds to the reference error rate, 4%, obtained in previous works. Thus, those algorithms—whose error variation sensitivity is higher than this reference value—should not be taken into account. The use of geometric features of the acoustic images reduces computational burden by a factor close to 1400, but the classification error increases 14%, so they must be discarded.

The optimal working area, placed in the lower right part of Figure 15, corresponds to a high reduction of the computational burden—high $1/S_b$ values—and a low increment of the error classification rate—low S_e values. However, computational burden reduction involves a reduction of the amount of data that represents the acoustic profile, which brings about the possible elimination of relevant information for the biometric classification and the increment of the classification error.

On the other hand, those algorithms which eliminate non-relevant information of the acoustic profiles reduce the computational burden without an increment of the error classification rate. The preprocessed algorithm and the line coding algorithm based on length and position of rows and columns present the same error classification rate as the algorithm which uses raw profiles and a reduction of the computational burden of 2.13 and 43.90, respectively. Therefore, the line coding algorithm based on length and position of rows and columns shows the best performance among the assessed algorithms.

5. Conclusions

An innovative biometric system, which significantly improves the performance of previous systems developed by the research group, is presented in this paper. Its improvement has been achieved through an increment of the number of frequencies analyzed, a reduction of the number of scanning positions, the use of preprocessing techniques on acoustic images and the use of SVM algorithms for the classification task. Reliability and robustness of the system were improved by employing a large set of subjects called intruders, by increasing the number of acoustic profiles captured for each subject and by this regarding the clothes they were wearing during the test so as not to affect the classification.

It has been verified that, as the size of the acoustic profiles decreases by using the preprocessing techniques, the classification error increases, because relevant information is removed. However, the line coding algorithm based on length and position of rows and columns allows reducing computational burden in several orders of magnitude without increasing the classification error rate. In this case, the information of the profiles eliminated by the algorithm is not relevant to the classifier. Thus, this preprocessing algorithm has been selected to be used in the improved biometric system.

On the other hand, the fact that this line coding algorithm was based on binarized images shows that the relevant information for the classifier is associated to the contour of the image. Finally, it was observed that the geometric features extracted from the acoustic images do not provide enough information for the classifier.

Our research group is currently working on improving the biometric system by using bidimensional arrays, employing new algorithms based on Gaussian Mixture Models and creating a large database of acoustic profiles.

Acknowledgments

The authors thank Anibal Figueiras for his technical support with Support Vector Machines.

Author Contributions

All authors contributed equally to the report research and writing of this paper.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Jain, A.; Bolle, R.; Pankanti, S. *Introduction to Biometrics*, 1st ed.; Springer: New York, NY, USA, 1996; pp. 1–41.
2. Crispin, J.; Maffett, A. Radar cross-section estimation for complex shapes. *IEEE Proc.* **1965**, *53*, 972–982.
3. Neubauer, W. A summation formula for use in determining the reflection from irregular bodies. *J. Acoust. Soc. Am.* **1963**, *35*, 279–285.
4. Baker, C.; Vespe, M.; Jones, G. 50 million years of waveform design. In Proceedings of Forum on Engineering and Technology, London, UK, 22 November 2006; pp. 7–21.
5. Balleri, A.; Woodbridge, K.; Baker, C.J.; Holderied, M.W. Flower Classification by bats: Radar comparisons. *IEEE Aerosp. Electron. Syst. Mag.* **2009**, *5*, 4–7.
6. Helversen, D.; Holderied M.W.; Helversen, O. Echoes of bat-pollinated bell-shaped flowers: Conspicuous for nectar-feeding bats. *J. Exp. Biol.* **2003**, *6*, 1025–1034.
7. Chevalier, L.F. *Principles of Radar and Sonar Signal Processing*, 1st ed.; Artech House: Boston, MA, USA, 2002.
8. Ricker, D.W. *Echo Signal Processing*, 1st ed.; Kluwer: Dordrecht, The Netherlands, 2003.
9. Moebus, M.; Zoubir, A.M. Three-dimensional ultrasound imaging in air using a 2D array on a fixed platform. In Proceedings of IEEE International Conference on Acoustic, Speech and Signal Processing, Honolulu, HI, USA, 15–20 April 2007; pp. 961–964.
10. Moebus, M.; Zoubir, A.M. Parameterization of acoustic images for the detection of human presence by mobile platforms. In Proceedings of IEEE International Conference on Acoustic, Speech and Signal Processing, Dallas, TX, USA, 14–19 March 2010; pp. 3538–3541.
11. Duran, J.D.; Fuente, A.I.; Calvo, J.J.V. Multisensorial modular system of monitoring and tracking with information fusion techniques and neural networks. In Proceedings of IEEE International Carnahan Conference on Security Technology, Madrid, Spain, 5–7 October 1999, pp. 59–66.
12. Izquierdo-Fuente, A.; Villacorta-Calvo, J.; Raboso-Mateos, M.; Martinez-Arribas, A.; Rodriguez-Merino, D.; del Val-Puente, L. A human classification system for a video-acoustic detection platform. In Proceedings of International Carnahan Conference on Security Technology, Albuquerque, NM, USA, 12–15 October 2004; pp. 145–152.
13. Izquierdo-Fuente, A.; del Val-Puente, L.; Jiménez-Gómez, M.I.; Villacorta-Calvo, J. Performance evaluation of a biometric system based on acoustic images. *Sensors* **2011**, *11*, 9499–9519.
14. Jain, A.K.; Nandakumar, K.; Ross, A. Score normalization in multimodal biometric systems. *Pattern Recogn.* **2005**, *38*, 2270–2285.
15. Lee, E.C.; Jung, H.; Kim, D. New finger biometric method using near infrared imaging. *Sensors*. **2011**, *11*, 2319–2333.
16. Lee, E.C.; Park, K.R. Image restoration of skin scattering and optical blurring for finger vein recognition. *Opt. Laser Eng.* **2011**, *49*, 816–828.
17. Hamdy, O.; Traoré, I. Cognitive-based biometrics system for static user authentication. In Proceedings of 4th International Conference on Internet Monitoring and Protection, Venice, Italy, 24–28 May 2009; pp. 90–97.

18. Izquierdo-Fuente, A.; del Val-Puente, L.; Villacorta-Calvo, J.; Raboso-Mateos, M. Optimization of a biometric system based on acoustic images. *Sci. World J.* **2014**, *2014*, doi:10.1155/2014/780835.
19. Cristianini, N.; Shawe-Taylor, J. *Support Vector Machines and Other Kernel-Based Learning Methods*, 1st ed.; Cambridge University Press: Cambridge, UK, 2000.
20. Theodoridis, S.; Koutroumbas, K. *Pattern Recognition*, 1st ed.; Academic Press: Melbourne, Australia, 2008.
21. Bishop, C.M. *Pattern Recognition and Machine Learning*, 1st ed.; Springer: New York, NY, USA, 2006.
22. Skolnik, M.I. *Introduction to Radar Systems*, 3rd ed.; McGraw Hill: New York, NY, USA, 2001.
23. Izquierdo-Fuente, A.; Villacorta-Calvo, J.J.; Val-Puente, L.; Jiménez-Gomez, M.I. A simple methodology of calibration for sensor arrays for acoustical radar system. In Proceedings of 118th Convention Audio Engineering Society, Barcelona, Spain, 28–31 May 2005.
24. Wirth, W.D. Radar Techniques Using Array Antennas. *IEE Radar, Sonar, Navigation and Avionics Series 10*, 1st ed.; The Institution of Electrical Engineers: London, UK, 2001.
25. Hsu, C.W.; Chang, C.C.; Lin, C.J. *A Practical Guide to Support Vector Classification. Technical Report*; Department of Computer Science, National Taiwan University: Taipei, Taiwan, 2003.
26. Chang, C.C.; Lin, C.J. LIBSVM: A library for support vector machines. *ACM Trans. Intellig. Syst. Technol.* **2011**, *2*, 1–27.
27. Cherkassky, V.; Ma, Y. Practical selection of SVM parameters and noise estimation for SVM regression. *Neural Netw.* **2004**, *17*, 113–126.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).