Applied Intelligence (2019) 49:1823–1840
https://doi.org/10.1007/s10489-018-1352-6

# Improving the performance of the lip identification through the use of shape correction

Carlos M. Travieso[1,2] · Antonio G. Ravelo-García[1,2] · Jesús B. Alonso[1,2] · José M. Canino-Rodríguez[2] · Malay Kishore Dutta[3]

Published online: 13 December 2018


**Abstract**
This work presents automatic identification and verification approaches based on lip biometrics, using a static lip shape and applying a lip correction preprocessing, transforming data from Hidden Markov Model and being classified by Support Vector Machines. The classification system is conclusive for the identification of a person by the shape of the lips, even if the person presents soft facial emotions. Moreover the use of static lips shape has been revealed as a good option for security applications. The experiments have been carried out with two public datasets. One dataset was used to model and validate the approach, and the other dataset has been used to test the model blindly. The accuracy is up to 100% and 99.76% for GDPS-ULPGC and RaFD datasets respectively, using two training samples under a hold-out validation. Based on the results we can conclude that the system is very robust and stable with the highest classification capacity and minimal computation complexity.

**Keywords** Lip-based biometrics · Image processing · Pattern recognition · Automatic identification · Artificial intelligence

## 1 Introduction

Security has become one of the most important areas for states, and biometrics is the safest method to identify a person. As a result, this field is in constant development and has generated a substantial economic growth.

Many kinds of biometric modalities have been developed according to its use, acceptation, technology, and applicability. Nowadays, fingerprint, face and iris are the common sources for security applications, but many other modalities are being developed to improve its accuracy and/or its fusion with other modalities. In particular, in this work, a model for lip identification and verification, that improves the-state-of-the-art [1], is presented.

✉ Antonio G. Ravelo-García
antonio.ravelo@ulpgc.es

[1] Institute for Technological Development and Innovation in Communications, Signals and Communications Department, IDeTIC-ULPGC, University of Las Palmas de Gran Canaria, Las Palmas, Gran Canaria, Spain

[2] Signals and Communications Department, University of Las Palmas de Gran Canaria, 35001 Las Palmas, Gran Canaria, Spain

[3] Centre for Advanced Studies, Dr. A.P.J. Abdul Kalam Technical University, Lucknow, India

In [2], a classification based on lip information is presented where features of shape and texture are extracted by static and dynamic methods for physiological and behavioral applications. Most of the works in the state of the art are based on dynamic information and movement of the lips, among which is that made by Mehra et al. [3], that tries to recognize a speaking person through his lips. They made use of features of the lips drawn by Principal Component Analysis (PCA). This technique produces feature vectors in a reduced dimension, which are the input to neural network classifiers. Thereby, they achieved a success rate of 91.07%.

Another research that tries to recognize a speaking person by his lips was performed by Arsic et al. [4], although in this case with the support of visual and hearing information extracting features by color transformation and a fuzzy clustering technique. Using only visual information, a maximum success rate of 94.44% was reached, whereas when visual combined with auditory information was used, the result reached a 100% success rate in almost perfect conditions.

There are also researches such as Newman et al. [5], who try to recognize the language a person is talking by recognizing the movement of their lips while saying a series of established words. For this goal, they used Active Appearance Models (AAM) that located the face and mouth and produced a vector that represented the shape of the lips for each video frame. Once these vectors are classified, they

obtained results of up to 100% between 75 different languages and 256 established words. Later, Newman and Cox. [6] modified the classification system, to obtain a decision totally independent of the speaking person, getting a success rate of 34.70%. Authors in [7] proposed a novel ordinal contrast measure called Local Ordinal Contrast Pattern (LOCP). It was used in a three orthogonal planes (TOP) configuration to represent video of the mouth region of a talking face as input into a speaker verification system in the form of LOCP histogram. The verification was accomplished using chi-squared histogram distance or LDA classifiers, reaching EER over 0.36.

On the other hand, the use of static information is not commonly used due to its low accuracy, as is observed in [2]. Thus, some works as in Briceño et al. [8] detected the lips contour by a color transformation. Using the contour points to make the feature extraction and classifying with Support Vector Machine (SVM), they achieved a success rate of 100%, with 100 extracted features and 5 training samples, and 99.37%, with 175 extracted features and 2 training samples.

The lips contour extraction is another field of the state-of-the-art. In particular, Raheja et al. studied three facial expressions (happy/sad/neutral) by processing an image containing a human face, extracting the lips contour [9]. They extracted the lips contour by edge detection, generated a binary image, used a filling system and performed a histogram analysis of the binarized image for the identification of the facial expression. Using this system, they achieved a success rate of up to 95%. There exist a number of approaches for lips contour extraction, or lips corner detection, for visual speech/speaker recognition, as in [10].

Recently Shokhan and Khitam [11] presented an approach for biometric identification system based on lips shape recognition in low resolution images of human faces. In other study, Shen et al. [12] applied dynamic lips identification introducing the algorithm of calibration of feature points.

Other recent results have been focused on the use of limb prints. For instance, Kumaran et al. [13] correlated fingerprint with lip print patterns to determine the predominant fingerprint and lip print patterns in Guajarati Population. Sharma et al. [14] studied morphological pattern of lip prints to determine the predominant pattern in Rajasthani population. Thakur et al. [15] studied 40 pairs of twins and their families to evaluate the possibility of variation of lip print patterns. Finally, Alzapur et al. [16] studied the uniqueness, prevalence and gender significance of lip print patterns in human subjects.

In previous contributions [1, 17], we have started to introduce some methods related to this topic. For example, in [1], the lip information is used to build a bimodal biometric approach where lips and faces are fused, and a Support Vector Machine based classifier is proposed. In [17], a new deep learning approach is applied to the lip-based biometric system.

In this new approach an improvement of the color transformation from [1] and the use of lip shape correction made possible a better feature extraction. Moreover, the whole new process presented in this work has been validated with two public datasets obtaining a consistent result (see Table 12).

In this work an improved lip feature extraction and a correction system are proposed enhancing the state of the art and showing that the use of static lips shape is a good option for security applications.

A system capable of identifying a person by the lips characteristics (see Fig. 1) has been designed and implemented in this proposal. For this purpose, the subject should be at rest, face forward, without moving the lips, and with a neutral expression or an expression that does not vary significantly the shape of the lips.

In addition, a feature extraction technique with the highest classification capacity and minimal computation complexity is presented. To face this challenge, a set of experiments allowing the comparison between the performances of the different feature extraction techniques to apply are showed. The goal is to specify which features are the most effective, obtaining the lips characteristics and the identification of users (see Fig. 2). Therefore, the innovation and contribution of this work is to do more efficient previous approach based on shape information, by its correction. The idea is to reduce the possible errors during the preprocessing and processing image stage, reaching more discriminative features and improving the accuracy of this proposal vs. the state-of-the-art.

## 2 Preprocessing

This stage proceeded to carry out a preprocessing of the images previous to the feature extraction process.

### 2.1 Facial image extraction

For the facial image extraction, we have used an algorithm derived from Viola and Jones [18]. All images used in our experiment were successful detected.

### 2.2 Color transformation

The labial zone is usually one of the areas of the face where the red component is more intense. To highlight this component, a color transformation is used, so lips can be quickly and easily identified. At the same time, it was found that the rest of the facial image has higher values in the green component than in the labial area. This is because the mixture of red and green results in yellow, a color similar to the skin's color. Therefore, we have opted for the removal of the green component when

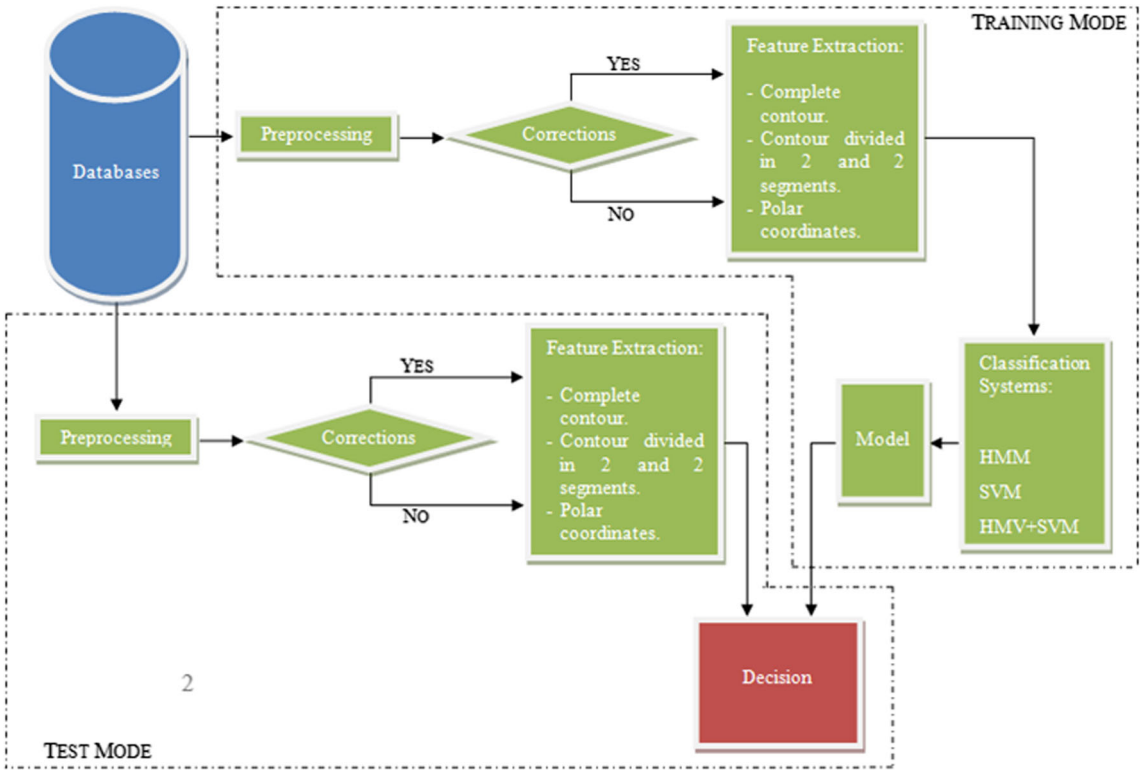


**Fig. 1** Block diagram of the recognition system based on people's lips

performing the color transformation, although, in some cases, the red component was also highlighted.

Furthermore, this approach measured variance before the transformation to determine the most appropriate equation to apply in each case, so as to obtain a more robust system to different forms and variations that can be found in a facial image. To do this, and through a series of heuristic tests, it was concluded that, for the measurement of the variance, the best choice is to use the following equation:

$$I = R + B - 6^{*}G \quad (1)$$

Thus, we will be able to determine whether the input image is one from a subject with a normal face, bearded face, punished face, etc.

The idea is to maximize the contrast between lips and the rest of the face to ease the lip segmentation and detection blocks. But the particularities of each user and images showed irregular answers for the Eq. 1. The solution was to define different equations according to the variability of the intensity of the pixel, based on the variance and under a heuristic experimentation.

Through a new series of heuristics tests, it can be concluded that the best transformation for the different cases is the equation presented in Table 1. Thus, an improvement of the efficiency is achieved by highlighting the lips in a different way according to the case. These values have been obtained using some samples of one of the two datasets of this work.

Once applied such transformation, a normalization of the image is applied, where the different values of the transformed image are placed between 0 and 255 to eliminate variations due to changes, for example, in light intensity (see Fig. 3).

**Fig. 2** Facial extraction block output for three example images

**Table 1** Equations used for the color transformation

| Transformation | Case | Variance |
|---|---|---|
| I = R + B – 1,9* G | Normal faces | $< 2{\cdot}10^9$ |
| I = R + B – 2,2* G | Bearded faces | $< 6{\cdot}10^6$ |
| I = R + B – 4* G | Punished faces | $< 8{\cdot}10^6$ |
| I = R + B – 2,4* G | Punished faces | $< 3{\cdot}10^7$ |
| I = R + B – 2,4* G | Others | $< 6{\cdot}10^7$ |
| *I* = 3*R + B – 6,4* G | Normal faces | The rest |

## 2.3 Labial zone search and image filtering

For a better detection of the lips contour, it has been decided that, since they are always located at the bottom of the image, we can crop the image and thus focus on the region of interest.

Furthermore, since in the image obtained after the crop process, it can be observed that, in some cases, it is obtained within the region of interest the nostrils or some other variation with high red component, we will proceed to eliminate such problem by finding the center of mass.

The facial distribution was used to crop the labial region. It is based on a detected face as input. The location of the mouth is always in the lower half and in the center of that half. That low half of the face can be reduced in its right and left extreme and in the superior and inferior part, although a little more in its ends. That way, the mouth is detected automatically (see Fig. 4).

Once cropped the labial region, we proceeded to perform a two-dimensional digital filter to the image. This filtering is applied so that, especially in cases of faces with beards or stubble, these areas are softened, and the edge of the lip contour becomes more marked. Thus, a faster lighter, but also safer system in detecting lips, is obtained (see Fig. 5).

## 2.4 Black and white transformation

The image binarization is necessary for an easy and accurate extraction of the contour of the lips since this will simplify the processing applied to the image. In order to convert a color image to black and white, a threshold level must be set. This threshold is chosen according to Otsu's method that minimizes intraclass variance of white and black pixels [19]. This Otsu threshold is the starting point. The second point is to decrease a 10 % of the sum between mean and variance from pixel intensity values of mouth cropped image. Then, the object can be increased in order to have a better detection. The obtained values using only the Otsu method are not fully successful. For this reason, the sum of the result obtained by Otsu's method and the mean of the cropped image will be used. Thus, the threshold that is obtained does not only depend on the variance, but also on the average of the image.

Then, the object can be better detected due to the reduction of the noise and other different objects. The last step is the use of morphological operators and filtering for removing the noising point and delimitate better the right object, the mouth. Furthermore, for more credible and less rough detection, a smoothing process is then applied to the contour.

## 2.5 Corrections

After applying the preprocessing described above, we observed the emergence of various irregularities (especially in cases of lip color very similar to the rest of the facial skin and/or with illumination changes). Figure 6 shows the problem and the solution, which will be proposed in this stage. Therefore, authors designed a new block, which is composed by two steps. The first step defines the ellipse-shaped pattern and the second step is the correction of the mouth object.

Step 1: Correction using an ellipse-shaped pattern

The first step of correction applied is based on the use of a pattern for the removal or trimming of the most significant irregularities. This pattern will be ellipse-shaped, as seen in Fig. 7, which will vary in relation to the size of the lips, smaller size (belonging to a child) or larger (belonging to an

**Fig. 3** Color transformation block output

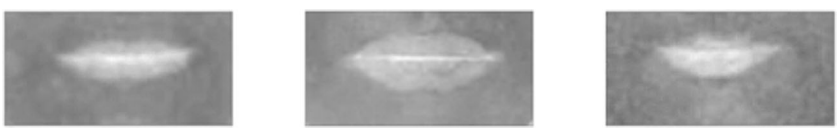

**Fig. 4** Labial zone search and image filtering block output

**Fig. 5** Black and white transformation block output

adult). If the detected mouth object is inside of the ellipse-shaped, the algorithm will be finished without change. But if the mouth object exceeds the ellipse-shaped pattern, the second step is applied.

Step 2: Correction using quadratic regression

As with the corrections applied using an ellipse-shaped pattern, only the removal of the most significant irregularities is achieved. Then, we opted for the use of polynomial regression for the removal of the less significant and more resistant ones.

The goal of the correction stage is the input. That input is some x and y coordinates of the lip shape. The process is as follows: the right and left ends of the mouth is detected. The lower lip contour can be kept and so, the irregularities can be identified. If this is the case, some x and y coordinates before the irregularities will be labelled as the input. The idea is to define a polynomial function which replaces the irregularity.

The polynomial regression, according to Bishop's work [20], calculates the parameters of the polynomial function that better fits the lip contour. As the lip contour follows a polynomial form rather second order, quadratic regression is used, obtained according to Eq. 2.

$$\nabla_w E = \nabla_w \left( \frac{1}{2} \sum_{n=1}^{N} \{ y(x^n; w) - t^n \}^2 \right) \tag{2}$$

where $E$ is the error to minimize, $y$ the quadratic function, $x$ is the position of the samples, $w$ are coefficients of the quadratics function, and $N$ is the number of samples. The minimization of $E$ with respect to $w$ will define the quadratic function.

$$\nabla_w E = 0 \rightarrow \frac{dE}{dw} = 0 \tag{3}$$

Applying the minimization,

$$\frac{2}{2} \sum_{n=1}^{N} \{ y(x^n, w) - t^n \} \cdot \sum_{n=1}^{N} x^n = 0 \tag{4}$$

Extracting the weights of the polynomial function ($w$), the following expression is obtained:

$$w = tn/xn \rightarrow \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ \mathrm{M} \\ w_M \end{bmatrix} = \begin{bmatrix} t^0 \\ t^1 \\ t^2 \\ \mathrm{M} \\ t^N \end{bmatrix} / \begin{bmatrix} x_0^0 & x_1^0 & x_2^0 & \Lambda & x_M^0 \\ x_0^1 & x_1^1 & x_2^0 & \Lambda & x_M^1 \\ x_0^2 & x_1^2 & x_2^2 & \Lambda & x_M^2 \\ \mathrm{M} & \mathrm{M} & \mathrm{M} & \mathrm{O} & \mathrm{M} \\ x_0^N & x_1^N & x_2^N & \Lambda & x_M^N \end{bmatrix} \tag{5}$$

The resolution of expression (5) will give the weights of the polynomic function with minimum error. This process is applied to any irregularity detected by calculating the difference between adjacent values and on irregularities detected by calculating the difference between values which are not adjacent, but with a certain gap between them.

In some cases, and due to the above regression corrections applied, some kind of discontinuity can be found. These errors will be small. To avoid further feature extraction problems due to the possible lack of pixels in the coordinates where we want to take measures, we decided to make a contour processing, through a series of morphological operations where such discontinuities should be solved.

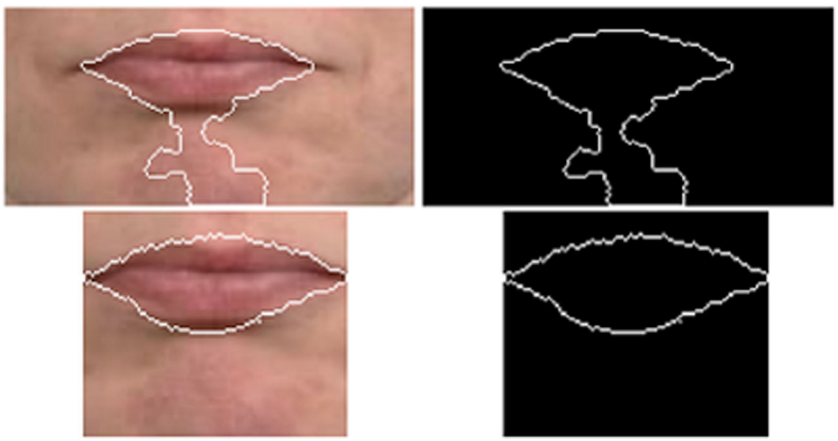

**Fig. 6** Representation of the result obtained by applying the different corrections

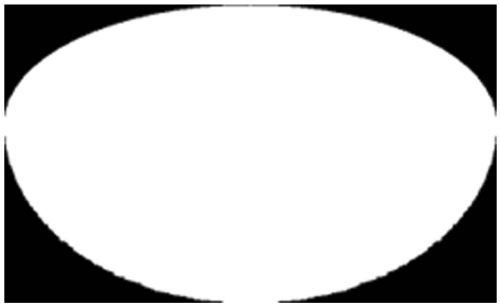

**Fig. 7** Ellipse-shaped pattern

## 3 Feature extraction

It is noted that the lips of a person contain no color distinguishing information, as it may be modified by the use of makeup techniques or changes in light intensity. Furthermore, there are no special features which serve for the distinction of different people except for, the most relevant data, the geometrical measurements that can be obtained from the lip contour.

From this contour, you can draw two important geometric measurements: height and width, and, in polar coordinates, the amplitude for a certain angle from 0 to 360 degrees.

However, before proceeding to the implementation of the feature extraction techniques, the maximum number of measures that may be performed to all images in a database must be checked.

### 3.1 Height and width measures for a complete lips contour

One of the more intuitive geometric measurements of a subject's lips is the height and width measurements of its contour (see Figs. 8 and 9). In Fig. 10, you can see a user's complete lip contour, where the measures to be carried out in this section are indicated.

Thus, for example Fig. 10, we obtain the following graphs of measures:

### 3.2 Height and width measures for contour divided in two and two segments

In Fig. 11, a subject's complete lip contour can be observed. Upper, lower, left and right segments have already been separated through the center of mass located at the intersection between the two shafts making up the segments. Height measures are made along the entire width of the contour of the lips, i.e., not limited by the vertical axis. Likewise, the width measures are not limited by the line of the horizontal axis but are carried along the full height of the contour under study.

Therefore, this part of the feature extraction has been called "contour divided in two and two segments", because, although the contour may appear to be divided into four segments (Fig. 11), actually, it is divided in two and, upon completion of a series of measurements, such division is eliminated, and it is again divided in two different segments. (Figs. 12 and 13).

### 3.3 Polar coordinates measures

Figure 14 shows the complete lips contour, where the intersection of the lines shown in black marks the center of mass, the red arrows are the rho coordinate measurements that are to be made and the red ellipse marks the angles at which to measure, i.e., 0 to 360 degrees.

In this case, the minimum values that can be observed in Fig. 15 correspond to the areas where the amplitude of

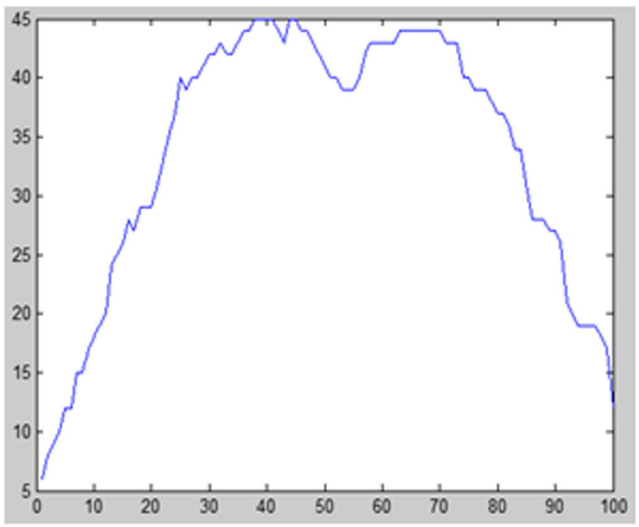


**Fig. 8** Height measures obtained from a complete lip contour

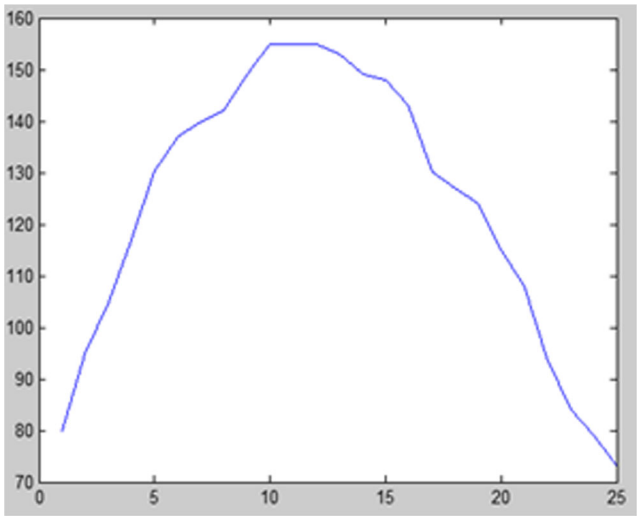


**Fig. 9** Width measures obtained from a complete lip contour

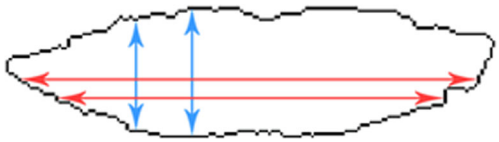

**Fig. 10** Height and width measures for complete lip contour

the lips with respect to the center of mass is minimal, i.e. upper and lower edges of the center of the contour, and the maximum values of this figure correspond to the areas where the amplitude of the contour of the lips is maximum with respect to the center of mass, i.e. the left and right corners. Furthermore, Fig. 15 shows the theta values on the horizontal axis, being observed that these values lie between −180 degrees and 180 degrees.

### 3.4 Comparison between Cartesian and polar features

Some figures have been included in order to show the feature distribution, using boxplot representation. The figures show the number of users on X axis and the value of features on Y axis. The boxplot representation shows statistical information about the feature distribution. In particular, the rectangle shows the second and third quartiles, and the median value is also indicated. The lowest and upper quartiles are represented as extensions of the rectangle with horizontal lines. The feature distribution is clearly shown with this representation. Moreover, the grade of discrimination is observed because the feature is shown one by one. Figures 16, 17, 18 and 19 show as the Cartesian coordinates are more dispersed than Polar Coordinates, which will fit with the experimental results. Also, this level of discrimination has propitiated to include the Fisher kernel improving the accuracy of this approach. The following figures shows the feature distribution (see Figs. 16, 17, 18 and 19).

## 4 Classification systems

In this section, the specification of the classification process is described. We used two classification approaches, Hidden Markov Model (HMM) [21] and Support Vector Machine (SVM) [22]. An HMM-based model has been deployed to

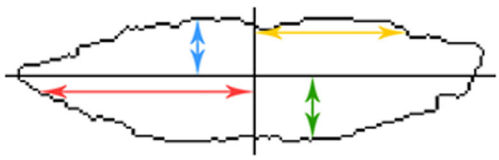

**Fig. 11** Height and width measures for a contour divided in two and two segments

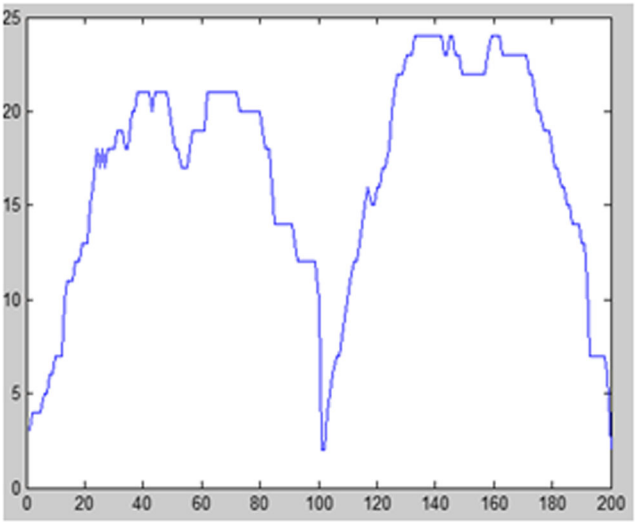


**Fig. 12** Height measures obtained from the upper and lower segments

characterize the sequential information from our features, in a similar fashion to [21]. Finally, these features have been transformed based on a HMM kernel learnt by SVM.

### 4.1 Hidden markov model

HMMs are theoretically sound and traditionally has performed very well in real applications such as speech recognition [23]. We review in this section the theoretical aspects related to our work.

An HMM has two associated stochastic processes; *dynamics* and *observations*. The former is not visible and is usually modeled by the probability of transition between hidden states. The observation process is modeled by the probability of obtaining an observed value given a hidden state (see [21, 24]) for a complete treatment of HMM).

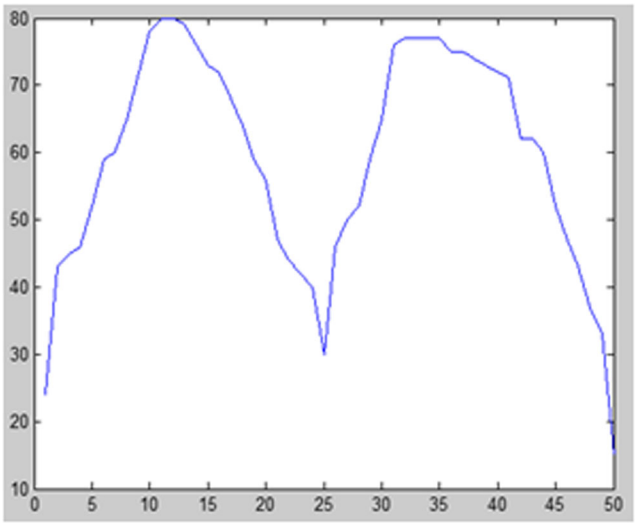


**Fig. 13** Width measures obtained from the left and right segments

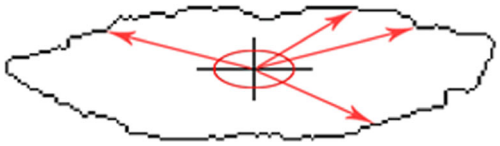

**Fig. 14** Polar coordinates measures

We use in this work a discrete HMM (DHMM) [21] since it forms the basis of our DHMMK detailed in the next section. A DHMM consists of the following parameters:

1) The number of states $N$
2) The number of different observations $M$
3) The transition probabilities matrix $a(N,N)$
4) The initial state probability $\pi(N,1)$
5) The observation probability matrix $b(N,M)$

With DHMM it is possible to extend the possibilities of modelling time sequences extracting dependencies between parts of a shape. In particular $N$ will represent a little sequence or segments of widths and/or heights.

One example of use of DHMM is presented in [25] where the shape of a signature is modeled in an offline signature verification system. The online signature offers both temporal information due to the temporal order of the points and the geometrical dependence of those points. Nevertheless, the offline signature is a shape that only offers the geometrical dependence of those points This has been verified by our previous work using HMM and polar grid features for offline signature verification [25].

Lip contour can be considered as a signature of face, appearing a potential dependence between the points on the lip contour. Moreover, it could also be considered as a one-dimensional manifold, being an ordered sequence in that space. These features suggest us to use the DHMM to model the lip contour descriptors. Therefore, as the transition through the states is produced in a single direction in the lip features, a "left to right" DHMM's turn out to be especially appropriate.

The first step with the DHMM approach is to quantize the feature vector elements. K-Means algorithm is used for clustering to create a set of symbols that are required by the DHMM [26]. These symbols construct a set of states (with a "left to right" structure) that a discrete observation at a step $t$ could be taken from. For our case, $t$ represents the step between the feature vector elements in a similar way as [25]. The number of observed symbols, $M$, is determined by experimentation. In our experiments we used thirty-two symbols. Based on Markov assumptions, given an observation sequence $X = [x_1, x_2, x_3, \ldots, x_t]$ with step $t$, and a trained DHMM $\lambda = [a,b,\pi]$, the probability of $X$ given $\lambda$ can be calculated as follows:

$$P(X/\lambda) = \sum_{S} \prod_{i=1:t} P(x_i/s_i, a) \cdot P(s_i/s_{i-1}, b) \tag{6}$$

Where $S = [s_1, s_2, s_3, \ldots, s_t]$ are the variables for the hidden states. Thus, for a given sequence, the optimal HMM can be selected by maximizing the posterior probability over a set of $C$ trained HMMs:

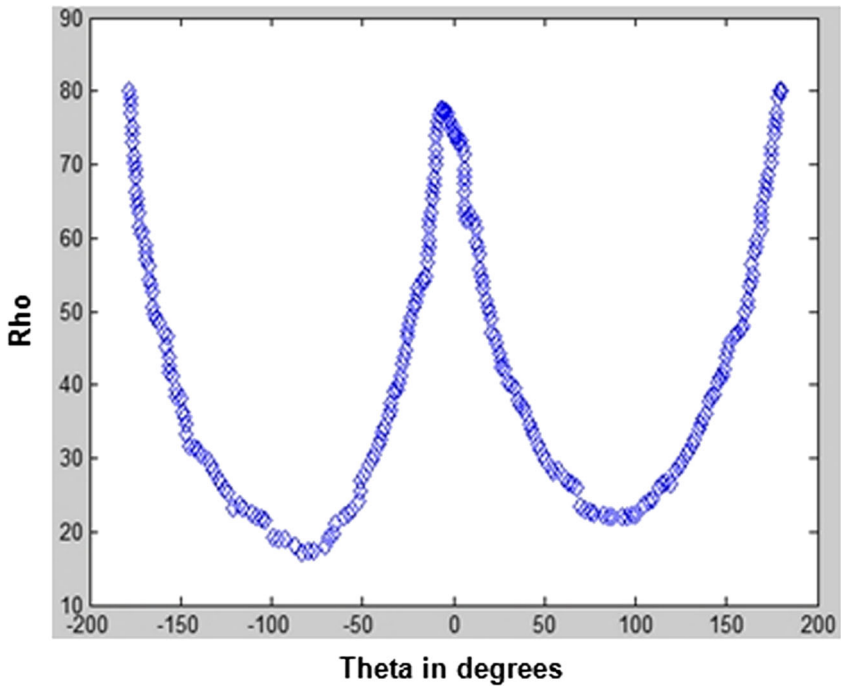


**Fig. 15** Polar coordinates measures obtained from a complete lips contour

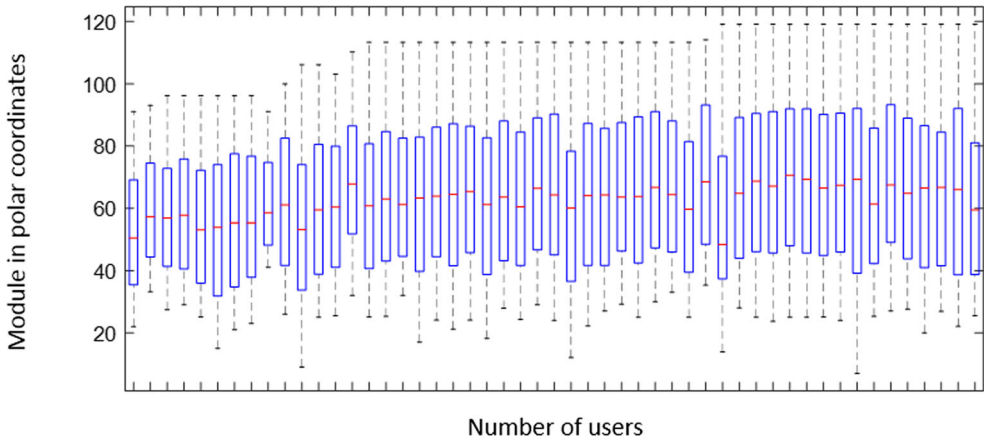


**Fig. 16** Feature Distribution of Polar Coordinates for GDPS-ULPGC dataset

$$\lambda_0 = \max_{j \in C} P(\lambda_j / X) \tag{7}$$

## 4.2 Discrete hidden markov model kernel

The learning of the HMM is based on maximizing the marginal probability of the observations over the hidden states. Thus, it is a generative method and does not fully utilize the inter-class discriminative information presented in the training set. A natural way to address this weakness is to incorporate it into a discriminative learning framework, such as an SVM. This can be achieved by computing the Fisher score of an observation sequence over the learned DHMM parameters, which is calculated by the gradient of the parameter space [27].

From [27], we can calculate that gradient of the logarithm of the probability in Eq. (6) with respect to the HMM parameter $\lambda$, as follows:

$$U(X/\alpha) = \Delta_{\alpha_{i,j}} \log P(X/\lambda) = \frac{\xi(x, s_j)}{\alpha_{i,j}} - \xi(s_j) \tag{8}$$

$$U(X/\beta) = \Delta_{\beta_{i,j}} \log P(X/\lambda) = \frac{\xi(s_i, s_j)}{\beta_{i,j}} - \xi(s_j) \tag{9}$$

We call this a DHMM Kernel where $\xi(x, s_j)$ represents the number of times a certain symbol $x$ is generated by a state $s_j$ in a sequence, while $\xi(s_i, s_j)$ is the frequency of the joint occurrence of two states $s_i$ and $s_j$ at two adjacent time intervals over a sequence, and $\alpha$ and $\beta$ are forward and backward variables, respectively. These are in fact the sufficient statistics for the emission probability $a_{i,j}$ and transition probability $b_{i,j}$ given a sequence. $\xi(s_j)$ represents the frequency of state $s_j$ occurring in a sequence [21, 27]. These values can be directly obtained from the forward-backward algorithm [27].

The DHMM kernel vector $U_X$ for a given sequence $X$ is the concatenation of the two gradient vectors calculated from Eqs. (8) and (9) respectively. The length of the resulting feature vector for a sequence is $N \times (M + N)$.

Therefore, the similarity of two sequences with a learned HMM could be evaluated in a kernel fashion using the corresponding Fisher score vector as follows

$$K(X, Y) = K(U_X, U_Y) \tag{10}$$

where $K(\cdot)$ could be any type of standard kernels for an SVM. We have used the gpdsHMM tool [26].

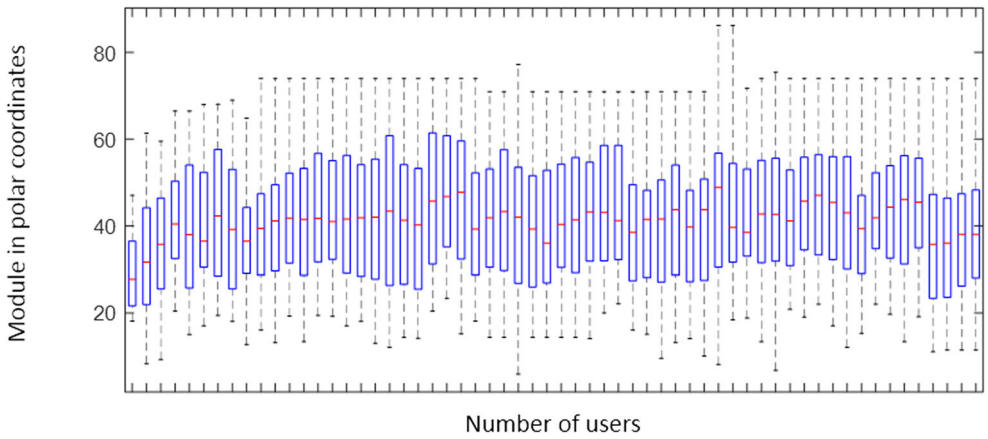


**Fig. 17** Feature Distribution of Polar Coordinates for RaFD dataset

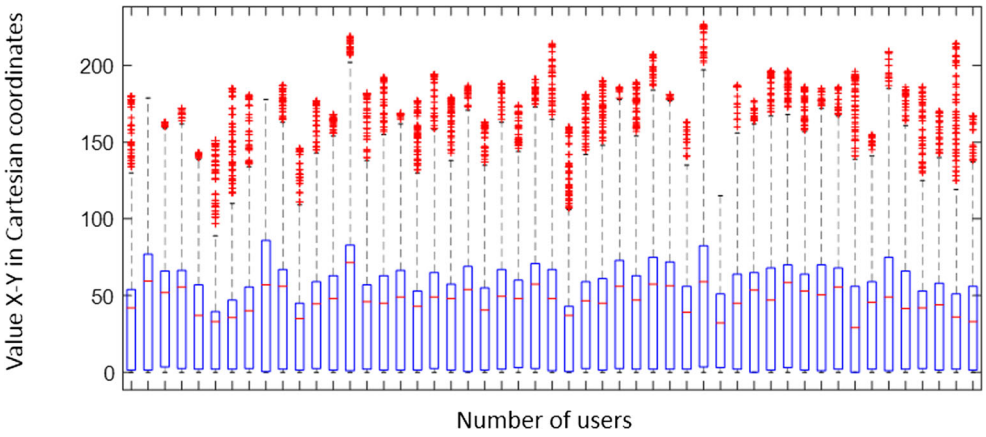


**Fig. 18** Feature Distribution of X-Y Cartesian Coordinates for GPDS-ULPGC dataset

## 5 Experimental methodology

### 5.1 Databases

#### 5.1.1 GPDS-ULPGC

The GPDS-ULPGC database was created by the Group of Digital Signal Processing at the University of Las Palmas de Gran Canaria. It is a public database and available on http://gpds.ulpgc.es/ and used in [1]. It will be used for the development and implementation of the system and for checking its correct functioning.

All users are in frontal position and with no facial expression (or neutral facial expression). Furthermore, it is observed that the light conditions are normal, subjects are Caucasian or Asian and different variations can be found (beard, glasses, acne, etc.) ( see Table 2).

#### 5.1.2 RaFD

The RaFD (Radboud Faces Database) database was created by the Behavioral Science Institute of the Radboud University Nijmegen, located at Nijmegen (The Netherlands). It is a public dataset and available in [28]. This database will not be used for the design of the system but, for the necessary experiments to verify the correct operation of the system.

Because of the necessary conditions for a proper detection and extraction of the lips contour, we have only been able to use the images of three of the eight expressions present in this database (neutral, sadness and indifference). Furthermore, images presenting changes in the camera angle have been eliminated, as well as the multiple users showing the mouth open in the selected expressions (see Table 3).

### 5.2 Feature extraction

As already mentioned, prior to conducting experiments, a test has been done to estimate the maximum measurements achievable on both databases. Table 4 shows the measures obtained from those tests, after checking different kinds of widths and heights.

Furthermore, a number of tests to check whether it is possible to use a smaller number of samples than the maximum possible already established were run. We appreciated a decrease in success rates, so we opted to use

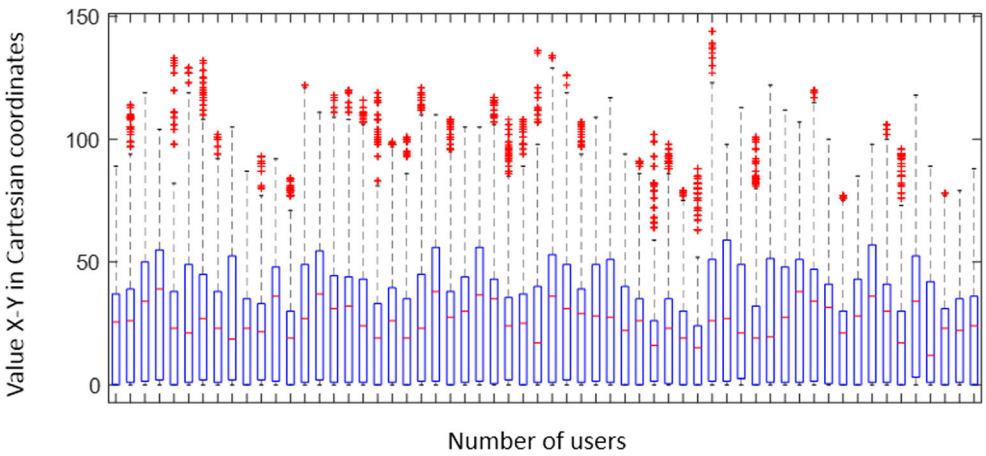


**Fig. 19** Feature Distribution of X-Y Cartesian Coordinates for RaFD dataset

**Table 2** Characteristics of the GPDS-ULPGC database

| Total Number of Samples | 510 |
|---|---|
| Number of Classes | 51 |
| Number of Samples per Class | 10 |
| Format | JPEG |
| Size | 1024 × 768 |
| Quantification | 24 bits |
| Resolution | 96 ppp |

the maximum possible measurements for the configuration of the feature extraction from the lips of a person.

The feature size for each dataset is different because the final lip size was different per each dataset. We have similar image resolution but different distance between camera and user. Therefore, this size must be adjusted.

### 5.3 Experimental methodology

It must be specified that a sum of three experiments with the same parameters with each classifier will be run, i.e., for the GPDS-ULPGC database and feature extraction of the contour divided in two and two segments, three experiments will be performed where the measures will be the same.

The recognition obtained from each experiment represents the identification of a person by the measures obtained from his lips contour. It is possible to obtain two values of interest, the mean and variance of the recognition rates that represent the mean recognition and its respective variance for the three experiments already mentioned. This work has followed a hold-out validation according to [29], repeating each experiment three times.

Other interesting point is the use of two datasets. One of them is used to establish the different variables and another dataset is used to validate this methodology allowing this approach to be independent of the dataset. If the accuracy is kept, then this approach will show its robustness.

Likewise, the number of samples used for the training of the system was established as a percentage in relation to the total number of samples in a database. We reduced it from a 50% hold-out to a 20% hold-out because, through a series of heuristic tests, we found that to be able to observe a significant reduction in success rates obtained by the HMM + SVM classifier, it is necessary to reduce the training percentage in an equally significant way.

**Table 3** Characteristics of the RaFD database

| Total Number of Samples | 540 |
|---|---|
| Number of Classes | 60 |
| Number of Samples per Class | 9 |
| Format | JPEG |
| Size | 1024 × 681 |
| Quantification | 24 bits |
| Resolution | 96 ppp |

**Table 4** Maximum measures achievable on both databases

| Database | Measures in complete contour | | Measures in contour divided in two and two segments | |
|---|---|---|---|---|
| | Heights | Widths | Heights | Widths |
| GPDS-ULPGC | 100 | 25 | 200 | 50 |
| RaFD | 60 | 20 | 120 | 40 |

At the same time, to reduce the quantity of experiments to be performed, we will proceed to the pre-establishment, through a new series of heuristic tests, of a fixed number of states used in the HMM and the HMM + SVM classifiers. In case of HMM + SVM classifier, we obtained better results using polar coordinates measures than with Cartesian coordinates. We also proceeded to differentiate the number of states used for each of these cases, searching the best accuracy, using GPDS-ULPGC dataset. Thus, we obtained Table 5.

## 6 Results

### 6.1 HMM classifier

Results obtained from the experiments made with a classifier based on Hidden Markov Models are presented (see Table 6).

Lower accuracy is obtained for a minor training percentage with the same feature extraction technique. The use of lip corrections improves the accuracy for this classifier, with independency of the feature extraction technique. But in general, it is observed that the use of the static lip with shape information is a bad approach for identifying, matching up the conclusions in [2].

### 6.2 SVM classifier

Table 7 presents the results obtained from the experiments made with a classifier based on Support Vector Machines.

This approach has better response than HMM classifier as observed in Table 7. Even the accuracy has good results with major training percentages. Moreover, the effect of lip corrections shows a better response for this classifier.

**Table 5** Pre-establishment of a fixed number of HMM states

| Classifier | Cartesian coordinates | Polar coordinates |
|---|---|---|
| HMM | 70 | 60 |
| HMM + SVM | 90 | 20 |

**Table 6** Results obtained with the HMM classifier

| Feature extraction technique | Preprocessing | Database | Training percentage | Success rates |
|---|---|---|---|---|
| Contour divided in two and two segments (70 states) | With corrections | GPDS-ULPGC | 50% | 79.22% ± 7.54 |
| | | | 20% | 60.21% ± 3.02 |
| | | RaFD | 50% | 26.94% ± 0.23 |
| | | | 20% | 16.27% ± 4.10 |
| | Without corrections | GPDS-ULPGC | 50% | 71.76% ± 29.53 |
| | | | 20% | 71.24% ± 14.20 |
| | | RaFD | 50% | 30.69% ± 0.41 |
| | | | 20% | 19.52% ± 25.53 |
| Complete contour (70 states) | With corrections | GPDS-ULPGC | 50% | 51.76% ± 1.85 |
| | | | 20% | 45.34% ± 18.08 |
| | | RaFD | 50% | 21.53% ± 0.93 |
| | | | 20% | 13.97% ± 5.06 |
| | Without corrections | GPDS-ULPGC | 50% | 58.56% ± 20.35 |
| | | | 20% | 46.32% ± 40.79 |
| | | RaFD | 50% | 24.44% ± 45.89 |
| | | | 20% | 17.46% ± 15.67 |
| Polar coordinates (60 states) | With corrections | GPDS-ULPGC | 50% | 39.08% ± 8.05 |
| | | | 20% | 29.49% ± 68.02 |
| | | RaFD | 50% | 13.89% ± 5.27 |
| | | | 20% | 10.16% ± 0.98 |
| | Without corrections | GPDS-ULPGC | 50% | 30.33% ± 57.57 |
| | | | 20% | 28.76% ± 30.06 |
| | | RaFD | 50% | 16.81% ± 12.90 |
| | | | 20% | 12.46% ± 14.42 |

### 6.3 HMM + SVM classifier

From the results, it is observed that classifiers used previously have not fully resulted in a conclusive system. For that reason, we have opted for the use of HMM and Fisher kernel for a data transformation and introduce this transformation as input of an SVM classifier. Thus, an improvement in the results is expected. The results obtained with this classification system are presented below.

As it can be observed in Table 8, this classification system obtains better results and the accuracy has been improved when the lip corrections were applied. This result refutes the

**Table 7** Results obtained with the SVM classifier

| Feature extraction technique | Preprocessing | Database | Training percentage | Success rates | $\gamma$ |
|---|---|---|---|---|---|
| Contour divided in two and two segments | With corrections | GPDS-ULPGC | 50% | 96.60% ± 0.05 | $4{\cdot}10^{-5}$ |
| | | | 20% | 90.20% ± 8.35 | $5{\cdot}10^{-5}$ |
| | | RaFD | 50% | 49.58% ± 21.53 | $1{\cdot}10^{-4}$ |
| | | | 20% | 36.75% ± 1.49 | $9{\cdot}10^{-5}$ |
| | Without corrections | GPDS-ULPGC | 50% | 94.25% ± 8.05 | $9{\cdot}10^{-5}$ |
| | | | 20% | 89.79% ± 5.67 | $6{\cdot}10^{-5}$ |
| | | RaFD | 50% | 59.72% ± 85.13 | $1{\cdot}10^{-4}$ |
| | | | 20% | 43.17% ± 0.47 | $1{\cdot}10^{-4}$ |
| Complete contour | With corrections | GPDS-ULPGC | 50% | 95.29% ± 4.31 | $7{\cdot}10^{-5}$ |
| | | | 20% | 86.76% ± 0.42 | $1{\cdot}10^{-4}$ |
| | | RaFD | 50% | 47.22% ± 11.34 | $1{\cdot}10^{-4}$ |
| | | | 20% | 34.13% ± 0.02 | $9{\cdot}10^{-5}$ |
| | Without corrections | GPDS-ULPGC | 50% | 93.20% ± 2.51 | $6{\cdot}10^{-5}$ |
| | | | 20% | 87.91% ± 1.34 | $6{\cdot}10^{-5}$ |
| | | RaFD | 50% | 48.06% ± 187.56 | $1{\cdot}10^{-4}$ |
| | | | 20% | 41.67% ± 8.33 | $9{\cdot}10^{-5}$ |
| Polar coordinates | With corrections | GPDS-ULPGC | 50% | 45.23% ± 3.43 | $5{\cdot}10^{-6}$ |
| | | | 20% | 28.92% ± 11.77 | $3{\cdot}10^{-6}$ |
| | | RaFD | 50% | 22.36% ± 24.36 | $9{\cdot}10^{-6}$ |
| | | | 20% | 15% ± 5.50 | $6{\cdot}10^{-5}$ |
| | Without corrections | GPDS-ULPGC | 50% | 51.37% ± 4.77 | $3{\cdot}10^{-6}$ |
| | | | 20% | 30.23% ± 3.92 | $5{\cdot}10^{-6}$ |
| | | RaFD | 50% | 24.58% ± 0 | $5{\cdot}10^{-5}$ |
| | | | 20% | 14.21% ± 3.08 | $8{\cdot}10^{-6}$ |

**Table 8** Results obtained with the HMM + SVM classifier

| Feature extraction technique | Preprocessing | Database | Training percentage | Success rates | γ |
|---|---|---|---|---|---|
| Contour divided in two and two segments (90 states) | With corrections | GPDS-ULPGC | 50% | 100% ± 0 | $3{\cdot}10^{-6}$ |
| | | | 20% | 98.77% ± 0.06 | $9{\cdot}10^{-7}$ |
| | | RaFD | 50% | 99.17% ± 0.52 | $4{\cdot}10^{-6}$ |
| | | | 20% | 93.25% ± 1.15 | $2{\cdot}10^{-6}$ |
| | Without corrections | GPDS-ULPGC | 50% | 99.22% ± 0.15 | $4{\cdot}10^{-6}$ |
| | | | 20% | 99.43% ± 0.14 | $4{\cdot}10^{-6}$ |
| | | RaFD | 50% | 92.64% ± 1.52 | $6{\cdot}10^{-6}$ |
| | | | 20% | 93.65% ± 1.44 | $2{\cdot}10^{-6}$ |
| Complete contour (90 states) | With corrections | GPDS-ULPGC | 50% | 99.87% ± 0.05 | $1{\cdot}10^{-4}$ |
| | | | 20% | 100% ± 0 | $8{\cdot}10^{-7}$ |
| | | RaFD | 50% | 99.72% ± 0.23 | $1{\cdot}10^{-4}$ |
| | | | 20% | 98.02% ± 0.59 | $1{\cdot}10^{-5}$ |
| | Without corrections | GPDS-ULPGC | 50% | 100% ± 0 | $4{\cdot}10^{-7}$ |
| | | | 20% | 99.59% ± 0.26 | $1{\cdot}10^{-4}$ |
| | | RaFD | 50% | 99.72% ± 0.23 | $2{\cdot}10^{-5}$ |
| | | | 20% | 97.30% ± 0.81 | $8{\cdot}10^{-5}$ |
| Polar coordinates (20 states) | With corrections | GPDS-ULPGC | 50% | 100% ± 0 | $1{\cdot}10^{-4}$ |
| | | | 20% | 100% ± 0 | $7{\cdot}10^{-5}$ |
| | | RaFD | 50% | 100% ± 0 | $1{\cdot}10^{-4}$ |
| | | | 20% | 99.76% ± 0.06 | $4{\cdot}10^{-5}$ |
| | Without corrections | GPDS-ULPGC | 50% | 100% ± 0 | $5{\cdot}10^{-5}$ |
| | | | 20% | 100% ± 0 | $1{\cdot}10^{-5}$ |
| | | RaFD | 50% | 99.72% ± 0.06 | $5{\cdot}10^{-6}$ |
| | | | 20% | 99.29% ± 1.08 | $1{\cdot}10^{-5}$ |

conclusion in [2]. Now, we can show the shape from static lip reaches a high level of success, useful for security applications. An important detail is the invariance with the lip sizes. In particular, for features based on polar coordinates, with corrections and using HMM + SVM approach, this system is working with good performance and keeps the accuracy with independence of the lip size and number of training.

## 6.4 ROC

Finally, the system has also been applied on an authentication problem using both databases and following a similar experimental protocol as in [8], obtaining the receiver operating characteristics (ROC) curves shown in Fig. 20. The performance can now be verified from the ROC curves (see Fig. 20) confirming that the proposed approach tends to be much more robust for lip size changes and variation of number of training samples.

## 6.5 System times

Here, the different temporary costs of the system are presented. It should be borne in mind that all exposed times have been measured when experiments were performed on AMD Phenom (TM) II ×4 3.20 GHz, 4 GB RAM and 500 GB HDD and using MATLAB 2012b. In Tables 9, 10 and 11 the preprocessing and features extraction times, the training computational time and test computational time, respectively, are shown.

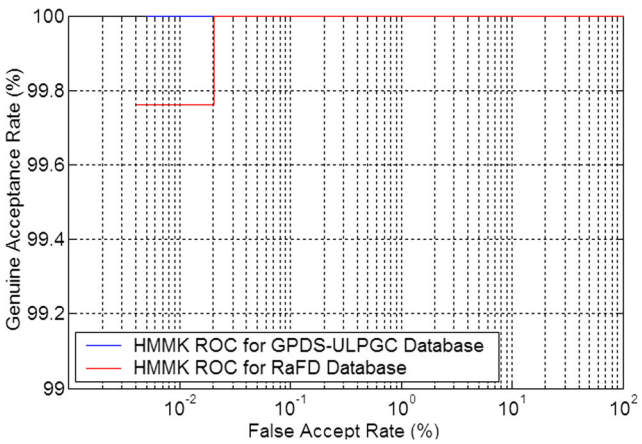


**Fig. 20** ROC curve for GPDS-ULPGC and RaFD databases under our best model based on DHMMK using 2 training samples (better viewed in color)

**Table 9** Preprocessing and feature extraction temporary costs

| Database | Technique | Temporal Cost per Sample (ms) |
|---|---|---|
| GPDS-ULPGC | Preprocessing without corrections | 100.5 |
| | Preprocessing with corrections | 133.1 |
| RaFD | Preprocessing without corrections | 60.1 |
| | Preprocessing with corrections | 82.9 |
| GPDS-ULPGC | Contour divided in two and two segments | 42 |
| | Complete contour | 28 |
| | Polar coordinates | 30 |
| RaFD | Contour divided in two and two segments | 23 |
| | Complete contour | 15 |
| | Polar coordinates | 16 |

The computational time for preprocessing and feature extraction can be considered as real time, because it is less than 180 milliseconds. Adding the test computational cost, for HMM and SVM, the times are inferior to 250 milliseconds. For HMM + SVM approach, up to 2 s are necessary per sample with polar coordinates but that time can be reduced by GPU's or C language programming and it is planned for future works.

### 6.6 Comparison with other authors

As mentioned above, a previous work, made by Briceño et al. [8] deals with the recognition of a static person by the characteristics of his/her lips and using the same dataset. In that work, they used the GPDS-ULPGC database performing the experiments through a color transformation, feature extraction based on contour points and classification using SVM. Although some similarities between that work and the present study can be observed, important differences exist, and this must be taken into consideration. For example, Briceño et al. [8] used a unique database. However, this study used the same database to be able to create the system along with a second database to verify the correct operation and robustness. In addition, the RaFD database contains images with soft facial expressions, showing that this system can be used in different conditions, which, together with any emotion detection system, provides an improvement in safety, as you can identify a person and verify that he is not under stress or fear conditions.

Furthermore, in the work made by Briceño et al. [8], the authors used a single equation for the color transformation, as well as one unique feature extraction technique to obtain lip characteristics. In the present study, however, different equations are used in the preprocessing of the color transformation for the different face cases, so as an improvement in the efficiency of the system is achieved by highlighting better the lips for each case. In addition, some corrections are applied to correct different irregularities.

Likewise, several feature extraction techniques have been used, so that we can determine which of them gets a better success rate, i.e., which one is the most suitable for the extraction of the characteristics of the lips. In addition, before the extraction of the characteristics, a test to estimate the maximum measurements achievable on each database is made, so that we can extract the characteristics in a more reliable way and without repeating measurements.

Finally, while in the work of Briceño et al. [8] the authors used the classification by HMM and HMM + SVM, they did not prove the efficiency of the SVM classifier by itself. Moreover, the results may be promising as they obtained a 100% success rate using 40 states, 150 measures and 5 training samples and a hit rate of 99.37% using 175 measures, 40 states and 2 training samples. However, the system presented in this study gets better results with the same database, the use of less states and the same number of samples used for the training of the system.

In 2017 it was published the most recent document based on lip biometric identification where landmarks

**Table 10** Training temporary costs of the different classifiers

| Database | Feature extraction technique | HMM (s) | SVM (s) | HMM + SVM (s) |
|---|---|---|---|---|
| GPDS-ULPGC | Contour divided in two and two segments | 953.7 | 1361.7 | 43,339.8 |
| | Complete contour | 550.8 | 1137.3 | 41,687.4 |
| | Polar coordinates | 1014.9 | 3677.1 | 6645.3 |
| RaFD | Contour divided in two and two segments | 1009.8 | 1441.8 | 42,751.8 |
| | Complete contour | 583.2 | 1204.2 | 43,016.4 |
| | Polar coordinates | 1074.6 | 3893.4 | 13,300.2 |

**Table 11** Test temporary costs of the different classifiers

| Database | Feature extraction technique | HMM (ms) | SVM (ms) | HMM + SVM (ms) |
|---|---|---|---|---|
| GPDS-ULPGC | Contour divided in two and two segments | 398.8 | 83.2 | 3130.3 |
| | Complete contour | 200.5 | 83.7 | 3316.6 |
| | Polar coordinates | 307.8 | 81.2 | 1725.6 |
| RaFD | Contour divided in two and two segments | 398.8 | 83.2 | 4185.6 |
| | Complete contour | 200.5 | 83.7 | 4272.3 |
| | Polar coordinates | 307.8 | 81.2 | 1289.4 |

with a probabilistic neural network was applied reaching an accuracy of 87.14% [30]. As can be observed in Table 12, our system outperforms that work and other reference papers. Therefore it is justified the robustness of this proposal and the innovation of using a lip correction preprocessing block could be used as a methodology to be applied in future works.

In an attempt to compare this study with the work presented by other authors, the next table is presented:

## 7 Discussion

In this work we have demonstrated that the use of shape data on static lips can be a competitive approach, versus the rest of categories given in Wang and Liew [2] (the use of dynamic lips).

The preprocessing technique with the use of corrections applied in a rather appropriate way, achieved better results with the HMM and HMM + SVM classifiers, without significantly increasing the computational cost of the system.

The classification using only Hidden Markov Models resulted in an inefficient system under any configuration of preprocessing or feature extraction techniques, for use of lip contours, due to the low maximum success rates obtained (79.22%).

The SVM classifier, however, obtained better results than the classifier stated before. We cannot conclude that this classifier is fully effective in identifying a person, despite it achieves a maximum success rate of 96.60% for the GPDS-ULPGC database, it gets a maximum success rate of 59.72% for the RaFD database. This result only demonstrates that the system is better prepared for the GPDS-ULPGC database than for the RaFD database, since it contains images with a bigger size of the lips.

Furthermore, through a classification with HMM, lower success rates are obtained with the feature extraction under polar coordinates, where it reaches a maximum success rate of 39.08%. In case of using SVM classifier, this rate is improved to a maximum of 51.37%, while with the HMM + SVM classifier, a 100% success rate is obtained for both databases. For the last classifier, the results also hold if we reduce dramatically the training percentage of the classification system.

Although, in Cartesian coordinates we also got a high success rate of 100% for the GPDS-ULPGC database and the HMM + SVM classifier, for the RaFD database we did not obtain the same result. In addition, for both databases we need a larger number of states than for the feature extraction by polar coordinates. This leads to an increase in the computational cost of the classifying block.

Likewise, the feature extraction technique called contour divided in two and two segments, with respect to the configuration using the complete contour, obtained an improvement

**Table 12** Comparison with other authors

| Bibliographic reference | Type of information | Success rates |
|---|---|---|
| Mehra et al. [3] | Dynamic Lip (movement) | 91.07% |
| Arsic et al. [4] | Dynamic Lip (movement) | 94.44% (visual information) and 100% (visual and auditory information) |
| Chan et al. [7] | Dynamic Lip (movement) | EER = 0.36% |
| Briceño et al. [8] | Static Lip and Shape | 100% (5 training samples) and 99.37% (2 training samples) |
| Travieso et al. [1] | Static Lip and SVM | EER = 12.75% |
| Travieso et al. [17] | Static Lip and Shape | 99,83% (2 GPDS training samples) and 98.10% (2 RaFD training samples) |
| Wrobel et al. [30] | Static Lip and Shape | 87,14% and EER = 12,57% (60% training Samples) |
| Wang and Liew [31] | Static Lip and Shape | EER = 18.50% |
| Our work | Static Lip and Shape | 100% (2 GPDS training samples) and EER = 0%<br>99.76% (2 RaFD training samples) and EER = 0.23% |

in the success rate with the HMM classifier. However, for the cases of SVM and HMM + SVM classifiers there was a minimal improvement in the success rates obtained by the feature extraction technique called complete contour, even though, in the first technique, the number of measures used are half those used in the second technique.

Thus, from all discussions drawn, we can deduce that the classification system by combining Hidden Markov Models, Fisher transformation and Support Vector Machines is conclusive for the identification of a person by the shape of the lip, even if the person presents soft facial emotions. In particular, it is a very robust system for the preprocessing with corrections and polar coordinates feature extraction technique, where, as already mentioned before, it achieved a 100% success rate in most experiments, with a more stable system than the one obtained if the preprocessing had been applied without corrections.

## 8 Conclusions

This work presents identification and verification approaches, using geometrical information, with lip corrections, and classifying by HMM + SVM. It has reached up to 100% and 99.76% for GDPS-ULPGC and RaFD datasets, respectively, where both are public datasets.

The innovation of this proposal demonstrates its robustness, and it is observed in the comparison versus the state of the art. The previous accuracies improve all results on lip-based identification and verification approaches.

**Acknowledgements** This work has been suported by Ministerio de Economía y Competitividad (TEC2016-77791-C4-1-R).

## References

1. Travieso CM, Zhang J, Miller P, Alonso JB, Ferrer MA (2011) Bimodal biometric verification based on face and lips. Neurocomputing 74(14):2407–2410
2. Wang SL, Liew AWC (2012) Physiological and behavioral lip biometrics: A comprehensive study of their discriminative power. Pattern Recogn 45(9):3328–3335
3. Mehra A, Kumawat M, Ranjan R, Pandey B, Ranjan S, Shukla A, Tiwari R (2010) Expert system for speaker identification using lip features with PCA. In Intelligent Systems and Applications (ISA), 2010 2nd International Workshop on (pp. 1–4)
4. Arsic I, Vilagut R, Thiran JP (2006) Automatic extraction of geometric lip features with application to multi-modal speaker identification. In Multimedia and Expo, 2006 IEEE International Conference on (pp. 161–164)
5. Newman JL, Cox SJ (2009). Automatic visual-only language identification: A preliminary study. In Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on (pp. 4345–4348). IEEE
6. Newman JL, Cox SJ (2010) Speaker independent visual-only language identification. In Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on (pp. 5026–5029). IEEE
7. Chan CH, Goswami B, Kittler J, Christmas W (2012) Local ordinal contrast pattern histograms for spatiotemporal, lip-based speaker authentication. IEEE Transactions on Information Forensics and Security 7(2):602–612
8. Briceño JC, Travieso CM, Alonso JB, Ferrer MA (2010) Robust identification of persons by lips contour using shape transformation. In Intelligent Engineering Systems (INES), 2010 14th International Conference on (pp. 203–207). IEEE
9. Raheja JL, Shyam R, Gupta J, Kumar U, Prasad PB (2010) Facial gesture identification using lip contours. In Machine Learning and Computing (ICMLC), 2010 Second International Conference on (pp. 3–7). IEEE
10. Rohani R, Alizadeh S, Sobhanmanesh F, Boostani R (2008) Lip segmentation in color images. In Innovations in Information Technology, 2008. IIT 2008. International Conference on (pp. 747–750). IEEE
11. Shokhan MH, Khitam AM (2015) Biometric Identification System by Lip Shape. International Journal of Advanced Computer Research 5(18):19
12. Shen X, Wang Z, Wang Y, Chen H (2015) Dynamic Lip Identification Algorithm Based on Automatic Calibration of Feature Points. Journal of Information & Computational Science 12(15):5659–5665
13. Kumaran M, Bastia BK, Kumar L, Patel SH (2017) Correlation between Fingerprint and Lip Print Pattern in Gujarati Population. Medico-Legal Update 17(1)
14. Sharma R, Sharma K, Preethi N, Degra H, Rajmani H (2017) Cheiloscopy: A Study of Morphological patterns of Lip Prints in Rajasthani population. Journal of Research in Medical and Dental Science 3(1):35–38
15. Thakur B, Ghosh B, Puri N, Bansal R, Yadav S, Sharma RK (2017) A comparative study of lip print patterns in monozygotic and dizygotic twins. International Journal of Research in Medical Sciences 5(5):2144–2149
16. Alzapur A, Nagothu RS, Nalluri HB (2017) Lip prints-A study of its uniqueness among students of MediCiti Medical College. Indian Journal of Clinical Anatomy and Physiology 4(1):68
17. Travieso CM, Zhang J, Miller P, Alonso JB (2014) Using a Discrete Hidden Markov Model Kernel for lip-based biometric identification. Image Vis Comput 32(12):1080–1089
18. Viola P, Jones MJ (2004) Robust real-time face detection. Int J Comput Vis 57(2):137–154
19. Otsu N (1979) A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics 9(1): 62–66
20. Bishop CM (1995) Neural networks for pattern recognition. Oxford University Press, Oxford
21. Rabiner LR (1989) A tutorial on hidden Markov models and selected applications in speech recognition. Proc IEEE 77(2):257–286
22. Cervantes J, Li X, Yu W, Li K (2008) Support vector machine classification for large data sets via minimum enclosing ball clustering. Neurocomputing 71(4):611–619
23. Hernando J, Nadeu C, Mariño J (1997) Speech recognition in a noisy car environment based on LP of the one-sided autocorrelation sequence and robust similarity measuring techniques. Speech Comm 21(1–2):17–31

24. Rabiner L, Juang B (1986) An introduction to hidden Markov models. IEEE ASSP Mag 3(1):4–16
25. Ferrer MA, Alonso JB, Travieso CM (2005) Offline geometric parameters for automatic signature verification using fixed-point arithmetic. IEEE Trans Pattern Anal Mach Intell 27(6):993–997
26. David S, Ferrer MA, Travieso CM, Alonso JB (2004) GPDSHMM: A hidden Markov model toolbox in the MATLAB environment. CSIMTA, Complex Systems Intelligence and Modern Technological Applications, 476–479
27. Jaakkola T, Diekhans M, Haussler D (2000) A discriminative framework for detecting remote protein homologies. J Comput Biol 7(1–2):95–114
28. Langner O, Dotsch R, Bijlstra G, Wigboldus DH, Hawk ST, Van Knippenberg AD (2010) Presentation and validation of the Radboud Faces Database. Cognit Emot 24(8):1377–1388
29. Braga-Neto UM, Dougherty ER (2004) Is cross-validation valid for small-sample microarray classification? Bioinformatics 20(3):374–380
30. Wrobel K, Doroz R, Porwik P, Naruniec J, Kowalsk M (2017) Using a probabilistic neural network for lip-based biometric verification. Eng Appl Artif Intell 64:112–127
31. Wang S-L, Liew AW-C (2012) Physiological and behavioral lip biometrics: A comprehensive study of their discriminative power. Pattern Recogn 45:3328–3335


**Carlos M. Travieso-González** received the M.Sc. degree in 1997 in Telecommunication Engineering at Polytechnic University of Catalonia (UPC), Spain; and Ph.D. degree in 2002 at University of Las Palmas de Gran Canaria (ULPGC-Spain). He is Head of Signals and Communications Department at ULPGC, Associate Professor from 2001 in ULPGC and accredited as Full Professor by the Spanish Government; teaching subjects on signal processing and learning theory. His research lines are biometrics, biomedical signals and images, data mining, classification system, signal and image processing, machine learning, and environmental intelligence. He has researched in more than 48 International and Spanish Research Projects, some of them as head researcher. He is co-author of 4 books, co-editor of 21 Proceedings Book, Guest Editor for six JCR-ISI international journals and up to 20 book chapters. He has over 400 papers published in international journals and conferences (65 of them indexed on JCR – ISI - Web of Science). He has published six patents and one more is under revision on Spanish Patent and Trademark Office. He has been supervisor on 8 PhD Thesis (9 more are under supervision), and 130 Master Thesis. He is founder of The IEEE IWOBI conference series and President of its Steering Committee, of The InnoEducaTIC conference series; and of The APPIS conference series. He is evaluator of project proposals for European Union (H2020), Spanish Government (ANECA), Medical Research Council (MRC, UK), DAAD (Germany), Argentinian Government and Colombian Institutions. He has been reviewer in different indexed international journals (<65) and conferences (<200) since 2001. He is member of IASTED Technical Committee on Image Processing from 2007 and member of IASTED Technical Committee on Artificial Intelligence and Expert Systems from 2011. He will be APPIS 2019 General Chair, and was IEEE-IWOBI 2018 General Chair, APPIS 2018 General Chair, InnoEducaTIC 2017 General Chair, IEEE-IWOBI 2017 General Chair, IEEE-IWOBI 2015 General Chair, InnoEducaTIC 2014 General Chair, IEEE-IWOBI 2014 General Chair, IEEE-INES 2013 General Chair, NoLISP 2011 General Chair, JRBP 2012 General Chair and IEEE-ICCST 2005 Co-Chair. He is Associate Editor on Computational Intelligence and Neuroscience journal (Hindawi – Q2 JCR-ISI). He was Vice-Dean from 2004 to 2010 in Higher Technical School of Telecommunication Engineers in ULPGC; and Vice-Dean of Graduate and Postgraduate Studies from March 2013 to November 2017.

**Antonio G. Ravelo García** is a docent and researcher in the Department of Signal and Communications (DSC) and Institute for Technological Development and Innovation in Communications (IDeTIC) at University of Las Palmas de Gran Canaria (ULPGC), Spain. He received the M.Sc. in Telecommunication Engineer and Ph.D. degrees from the Universidad de Las Palmas de Gran Canaria. He has participated in multiple research projects and publications. His current research interest includes biomedical signal processing, nonlinear signal analysis, data mining and intelligent systems in multimedia, communications, environment, health and wellbeing.

**Jesús B. Alonso-Hernández** received the Telecommunication Engineer degree in 2001 and the Ph.D. degree in 2006 from University of Las Palmas de Gran Canaria (Spain) where he is an Associate Professor in the Department of Signal and Communications from 2002. He has numerous papers published in international journals and international conferences, and different patents. His research interests include signal processing in biocomputing, biometrics, nonlinear signal processing, recognition systems, audio characterization and data mining. He has been guest editor of different special issues of Journals as Cognitive Computation, Neural Network and Applicacions or Neurocomputing.

**José Miguel Canino-Rodríguez** received the Technical Telecommunication Engineering (Radio-communication) diploma (University of Las Palmas de Gran Canaria, -ULPGC-, Spain, 1982), the Science Teacher Diploma (University of La Laguna, Spain, 1984), Physical Science degree (National University of Distance Education –UNED-, Spain, 1988) and a Telecommunications Technologies PhD, (ULPGC, 2010).

He has worked as a technical staff at Spanish Radio & Television Corporation (RTVE, 1982-1997), Associate Professor at University of Las Palmas de Gran Canaria (ULPGC, 1991-1997) and he is currently a Professor at ULPGC since 1997, where he has taught several subjects related to electrometric propagation and navigation systems.

Currently, his research work is focused on Intelligent Systems and Machine Learning Applications: agent-oriented methodologies for intelligent systems, machine learning for agent decision making, agent-oriented coordination and negotiation protocols, etc.

**Malay Kishore Dutta** is associated with the Centre for Advanced Studies, Dr. A.P.J.Abdul Kalam Technical University, Lucknow, India and working as a Professor of Computer Science and Engineering. Currently he is also holding the position of Dean, Research & Development and Industrial Consultancy. Earlier he was associated with Department of Electronics and Communication Engineering, Amity School of Engineering and Technology, Noida, Uttar Pradesh, India. He completed his M.Tech in Electronics Engineering from Central University, Tezpur, Assam, India, securing the first position (Gold Medalist). He completed his PhD in the area of multimedia data security and signal processing from Gautam Budh Technical University, India. He has published 230 peer-reviewed research articles in international journals and conferences. He has filed three patents in the area of signal processing and is the principal investigator of various DST-funded projects. He is also in many reviewer panels and editorial boards of many Journals.