

Lip Recognition for Authentication and Security

Sunil Sangve¹, Nilakshi Mule²

¹(Dept. of CSE, Zeal Education Society's DCOER Narhe, Pune, India)

²(Dept. of CSE, Zeal Education Society's DCOER Narhe, Pune, India)

Abstract: Lip is used as an identifier from person to person as the size of the upper and lower lips, furrows, grooves, the distances between the lines and the edges vary individualistically. These unique parameters are experimented to distinguish individuals and their uniqueness is also verified. The highlighting advantage in devising the biometric system based on lip recognition is that the data acquisition and handling are simple and hence storing and processing the data becomes an unconstrained task leading to drastic increase in the efficacy of the system. The main focus of this work is to track the lips in the given face image and then extract suitable features from the lips to perform the differentiation task. Accurately and robustly tracking lip motion in image sequences is especially difficult because lips are highly deformable, and they vary in shape, colour, and relation to surrounding features across individuals; in addition, they are subject to both non-rigid (expression, utterance of speech) and rigid motion (head movement). The use of colour space transformation has shown very well for lip image segmentation.

Keywords: Lip Corner Detection, Lip Corner Extraction, Lip Features Classification, Lip Segmentation, Region of Interest

I. Introduction

In the modern society, there is an ever-growing need to determine or verify the identity of a person. Early automated authorization and authentication methods relied on possessions and knowledge. The most important drawbacks of these methods are (i) possessions can be lost, forged or easily duplicated; (ii) knowledge can be forgotten; (iii) both knowledge and possessions can be shared or stolen. The science of biometrics provides an elegant solution to these problems by truly verifying the identity of the individual. For contemporary applications, biometric authentication is automated to eliminate the need for human verification, and a number of new biometrics has been developed, taking advantage of improved understanding of the human body and advanced sensing techniques. Biometric identifiers which are extracted so as to distinguish the individuals from one another should be unique and attainable. Physiological biometric includes Fingerprint recognition, Iris recognition, Face recognition, Hand geometry recognition and Vein recognition. Behavioral biometric is based on measurements and data derived from a physical action. Several work are also being carried out in biometric system like keystroke recognition. Grip recognition, ear recognition, odor recognition, gait recognition and much more to bud in future. Amongst the above mentioned techniques, iris recognition has been observed to have high accuracy and very less scanning time for verification. But on the other hand, it required expensive and highly complex acquisition device. Hence a simple and easily implementable biometric system must be devised to overcome all these constraints.

In this recognition 100 individuals facial images were recorded under standard conditions and algorithms were developed to extracted lip region. The extracted images were subject to sequence of processing techniques. Color plane extraction was first performed to convert the image into gray scale following which multiple filters were applied to improve the SNR. Particle classification technique was employed to the processed image to extract the features. Contour of the lip, surface area of the lower and upper lip and ridges present in the lower lip forms the set of features under consideration. The acquired features were subjected to statistical evaluation to study the systems uniqueness. Though very few researchers have worked on this technique, no effective result has been attained till date. In this work we focus on the lip corner detection. We consider that the face is already recognized then we select the lip contours and segmentation applying the extraction and classification we recognize the corner of lips.

II. LIP MODEL

2.1 Person Identification and Verification

Person recognition tasks can be divided into two classes:

- Person identification
- Person verification

The task of the first class is to determine the person from a closed set, whose features best match the features of the person to identify. Thus, it is assumed that only enrolled person will access the identification system. The verification problem is to validate the claimed identity of a person from an open set. Therefore, verification system has a reject subjects, called by *impostor*.

2.2 Lip Model

2.2.1 Region of interest

Person recognition using lip information is started with obtaining a so-called Region of Interest (ROI). Further the recognition system works with the ROI. There are two types of database which can be used for person recognition based on the labial area. The first kind of databases consists of face images. In this case the recognition system has to be able to localize the lip area in the image [6]. This is commonly executed by an assumption that the distance between the outer points of eyes is equal, approximately, to the distance between the bridge of nose and the middle of the mouth area, Fig. 1.

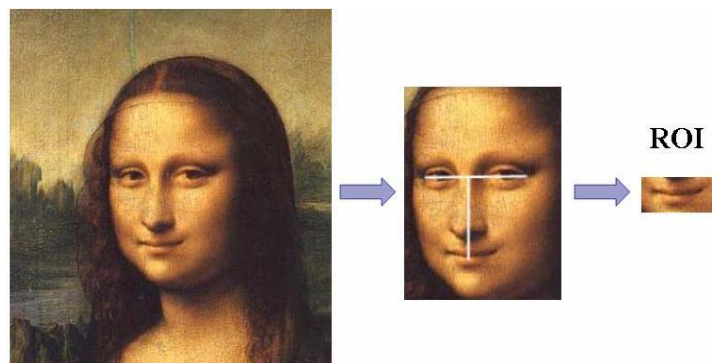


Figure 1. Region of Interest (ROI)

Having defined the ROI, the system has to make an appropriate description of visible motion. Since a physical process is modeled, it could be described in terms of physical movements and positions of the articulators, for instance, the muscle action could be assessed. However, the motion of mouth region musculature is very complex because of this motion is three-dimensional, not directly observed and there are at least thirteen groups of muscles defining lip movements. The technology rests upon construction of Appearance Based Model (ABM) for lip localization, lip tracking and features extraction. This kind of models involves static lips information, namely the inner and outer lip counters. During lip extraction, the Appearance Based Model is expanded to the Active Shape Model (ASM), based on the Point Distribution Model (PDM). The ASM is used to represent shape deformation and consists of two parts - the shape sub model and the intensity sub model.

2.2.2. Shape sub model

The first step of shape modelling is to label training shapes consisting of points around the bound of the training examples. To place this point's collection and to compare equivalent points from different shapes the reference points are needed. Their distance is scale; their orientation with respect to the horizontal is defined as the angle, and center of the scale as the origin[6]. The other points are fixed at equal horizontal distance along the lip contour.

III. Lip Contour Extraction

Empirical studies have found that a lip shape is usually close to an elliptical region. Therefore, the lips can be approximatively surrounded by the various elliptical contours according to its special structure. In our method, how to find a minimum-bounding ellipse as the initial evolving curve is of crucial importance to extract the lip contours. To find the minimum-bounding ellipse of lip region, the detection of lip corner dots is needed. Specifically, $I(x, y)$ represents a pixel value at coordinate (x, y) , m and n are the maximum values of rows and columns. The left corner, right corner, upper corner, lower corner are denoted as L_a , L_b , V_a , and V_b , respectively. We project the RGB-based lip image into the gray-level one. From the practical viewpoint, it is common that there exists the a noise or uneven illumination effects. Hence, each lip image is performed with a 3×3 mean filter and a contrast stretching adjustment. According to the empirical and statistical methods, the real lip region is usually different from the surrounding regions. Subsequently, the horizontal or vertical lip corner dots, e.g. Fig.2 (b), can be detected by computing the first and last points with the standard deviation, which is changing obviously compared with the adjacent ones. For example, in order to find the

horizontal lip corner columns.

3.1 Contour Extraction

After successfully finding the minimum-bounding ellipse of lip region, we let this ellipse be the evolving curve representing the zero level set C , which can be fitted well in the LACM. Subsequently, local neighbourhoods of the points can be split into the local interior region and local exterior region by the evolving curve. By computing the local energies at each point along the curve, the evolving curve will deform by minimizing the local energies so that the desired contour extraction is achieved[2]. The steps of lip contour extraction are as follows:

- Locate the lip region, pre process;
- Obtain the minimum-bounding ellipse;
- Evolve with iteration;
- Extract the lip contours.

Fig. 2 gives an example. Specifically, the local radius r selected by the function $B(u, v)$ is an important parameter in the LACM[2]. By the rule of thumb, $r = rb2$ is an appropriate value in the most cases of lip contour extraction. However, when there exists the lip image with the mouth open widely, the value of local radius selected should be smaller.

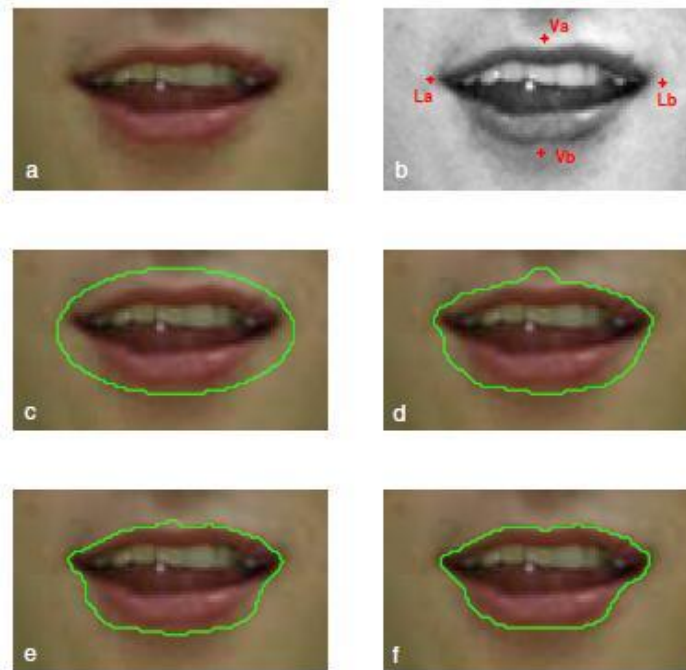


Figure 2. The procedure of lip contour extraction: (a) lip image of size 52×90 , (b) lip corner dots, (c) minimum-bounding ellipse, (d)-(f) the extracted results after 20, 40, and 60 iterations, respectively

IV. Lip Corner Detection

Accurately and robustly tracking lip in image sequences is especially difficult because lips are highly deformable, and they vary in shape, colour, and relation to surrounding features across individuals; in addition, they are subject to both non-rigid(expression, utterance of speech) and rigid motion (head movement). Reliable detection and localization of the two lip corners is a very important step in any lip tracking system. In this the basic concept of visual object modelling and detection is presented [8]. Thereafter, the concept is applied for locating the two lip corners in the face image. In some cases, a better approximation of the shape of an object is to use the Gaussian distribution as a base function and the distribution of the vector x can be approximated by a weighted mixture of M Gaussian distributions, that is

$$P(x) = \sum_{m=1}^M c_m N(x; \mu_m, C_m) \quad (4.1)$$

Where M is the number of mixture components, c_m is the weight of the m^{th} component. This can be

used if the distribution originally comes from different sources, which are normally distributed. For example, consider the height of men and women in a certain country; the height of men can be described by $N(\mu_1, \sigma_1)$, and for women $N(\mu_2, \sigma_2)$. For the whole population, if the women are 55% and the men are 45% of that population, a mixture of two Gaussian distributions could approximate the distribution of heights for the whole population

$$P(x) = 0.45N(x; \mu_1, \sigma_1) + 0.55N(x; \mu_2, \sigma_2) \quad (4.2)$$

For robust lip and face detection, a precise estimation of the position of the feature points in the face is needed. The surrounding area of these points characterizes their shape and appearance, which differs from person to person during the utterance of speech in different illumination conditions, and due to perspective distortion. The Gaussian distribution as mentioned above can represent these variations.

The mean template of an object can be considered as a representation of that object. However, lip shapes have high inter-person variation. Therefore the lip boundaries in the mean lip templates are indistinguishable from the surrounding face area as shown in Fig.3. This loss in shape information results in poor lip detection performance when mean lip templates are used [8]. If the variations of the appearance are significant, supervised appearance modelling is used. Given m different appearance classes, each different appearance category or class is treated separately, where each class is represented by a single Gaussian distribution with an average template T_i and $i = 1, \dots, m$. The whole object is represented by a Gaussian mixture using Eq. 4.1.

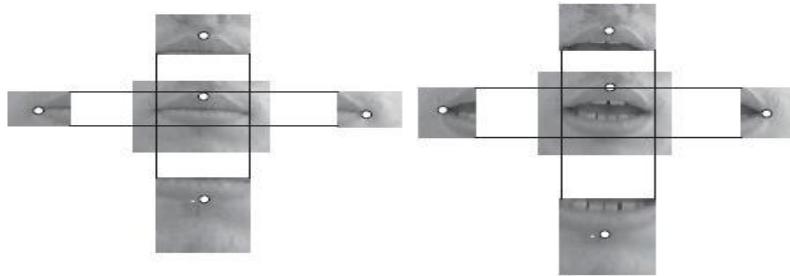


Figure 3. A selection of the area around the upper lip, the lower lip, and the corners

V. Lip Corner Extraction

5.1.1 PCA Training

After face detection, we need to extract lip corners to initialize the level set algorithm. Lip corners are extracted from the face image using PCA template matching. First we train a PCA model for patches of fixed dimension centered around left and right lip corners, separately, using a large training database of face images. We experimented with different patch window sizes to get the best result. We downsized (by 2) extracted face image for faster processing. We obtained a PCA lip-corner patch space of reduced dimension for each corner. Mean images and the primary PCA Eigen-patch images are shown in Figure 4. Finally, a Gaussian mixture model (GMM) is built using the PCA coefficients and the coordinates of each lip corner [7]. We expect that this model yields high likelihood for feature vectors extracted from correct lip corners. We obtained the best result for corner detection with a 41×61 window. Mean error of detection as measured in Euclidean distance between the predicted lip corner and the ground truth is 4.4 and 5.2 pixels for left and right lip corners respectively.



Figure 4. Mean lip-corner-patch (top row) and primary PCA eigen-patch (bottom row) images for left and right lip corners.

5.2.2.. PCA Template Matching

Once we train a PCA lip corner model, we find the lip corners in a given image using PCA template matching. After resizing the test image, we look in a patch around each pixel and reduce dimension by PCA. We get a feature vector which consists of the PCA coefficients in that neighbourhood concatenated by x and y coordinates of the pixel. GMM likelihood (or score) of each pixel is calculated using the model trained before, and the pixel with the greatest GMM score is decided to be the lip corner[7]. Our testing data is distinct from the training data. In addition, the subjects in training and testing images are distinct as well. Sample GMM score images and the found corners are shown in Figure 5.

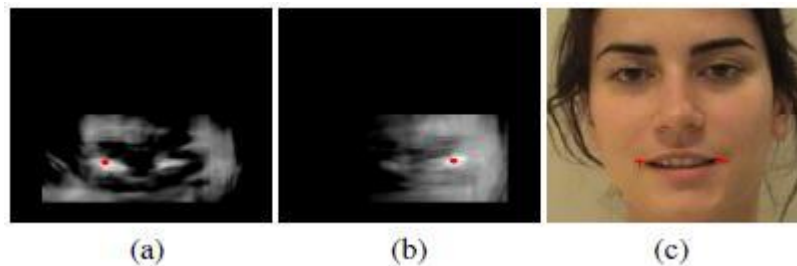


Figure 5. GMM score images for (a) left lip-corner, (b) right lip-corner. (c) Found lip corners shown Super imposed on the face image.

VI. Feature Classification

The most general approach to classifying these examples is to use the posterior probabilities of class membership $P(\omega_i|x)$, where x is the observed feature vector. One way to build a classifier is to develop a model that estimates the posterior probabilities directly; this can be done using non-linear output nodes with a neural network. An alternative is to use a set of density estimators and Bays' theorem. In order to build a classifier system based on density estimators, a density model Feature Classification for each class is required. The probability density estimate $p(x|\omega_i)$ is obtained for each class by training the model only on data from that class. Using Bays' theorem [8]. Theoretically, a Bayesian classifier is optimal as it is the best classifier which minimizes classification error. However, it is very difficult to create such a classifier as it requires the true conditional density functions and the true a priori class probabilities. Practically this is infeasible as we can only get estimates due to sample size limitation and analytical complexity. However, parametric classifiers can be used to approximate such density functions. The model used in this work for density estimation is Gaussian mixture model (GMM). This model is classed as a semi-parametric estimation model since it defines a very general class of functional forms for the density model where the number of adaptive parameters can be increased in a systematic way (by adding more components to the model) so that the model can be made arbitrarily flexible.

VII. Conclusion

The study has clearly depicted that the features in lip are unique characteristics of an individual. Lip images from 100 subjects were acquired. The distance between the vertical lines in the lower lips were compared between the subjects. This indicates that the distance between the lines, shape of lip and corner parameter varies from person to person. Thus from the study it is strongly evident that lip has a unique parameters in the form of grooves, furrows, distance between the lines of the lower lip and many more to be explored.

Acknowledgment

I would like to thanks Prof. Sunil Sangve & Prof T.T.Waghmare for his valuable suggestions.

References

Journal Papers:

- [1]. "Automatic lip contour extraction from color images" S.L. Wanga, W.H. Laua,*, S.H. Leungb
- [2]. "A Lip Contour Extraction Method Using Localized Active Contour Model with Automatic Parameter Selection" Xin Liu, Yiu-ming Cheung, Meng Li
- [3]. "Automatic Outer Lip Contour Extraction in Facial Images" Juan-Bernardo Gómez Mendoza, Flavio Prieto
- [4]. G. Chiou and J. N. Hwang, "Lip reading from color video," IEEE Transactions on Image Processing", vol. 6, no. 8,
- [5]. X. Zhang, R. M. Mersereau, M. Clements, and C. C. Broun, "Visual speech feature extraction for improved speech recognition," in Proc. International Conference on Acoustics.
- [6]. "Person Recognition Based On Lip Movements" Olga Shipilova
- [7]. "Lip Segmentation Using Adaptive Color Space Training" Erol Ozgur, Berkay Yilmaz, Harun Karabalkan, Hakan Erdogan, Mustafa Unel

Theses:

- [8]. "Real Time Lip Motion Analysis for a Person Authentication System using Near Infrared Illumination" Master Thesis Faisal Shafait 20th April 2005.