© 2011,
Scienceline Publication
www.science-line.com
Journal of World's Electrical Engineering and Technology
J. World. Elect. Eng. Tech. 1(1): 12-16, 2012
JWEET

An Efficient Algorithm for Lip Segmentation in Color Face Images Based on Local Information

Hashem Kalbkhani*, Mehdi Chehel Amirani

Department of Electrical Engineering, Urmia University, Urmia, Iran
*Corresponding author's email: st_h.kalbkhani@urmia.ac.ir

Abstract – *Lip detection is used in many applications such as face detection and lips reading. In previous works, researchers have considered whole of face image for lip detection. In this paper we propose a new algorithm. In our algorithm for reducing required calculation and increase accuracy of correct detection, we do not consider whole of the face image. We first remove the upper half part of the face image. Then, for estimate lip area, we divide remained lower half face image to equal parts. For each part we calculate statistical information such as standard deviation, and based on them we detect lip area in face image. For separate lip pixels from skin pixels, we use YCbCr and HSI color spaces at this work. We evaluate our work on CVL face database. Our experiments show that new algorithm gives better results than previous works on this database.*

Keywords*: lip detection, skin, saturation, standard deviation.*

ORIGINAL ARTICLE

INTRODUCTION

Lip area detection and extracting it from face is important in many applications such as face detection, lip reading, and voice detection [1]. In face detection lip is used for verifying detected area. The goal of lip-reading process is to make natural connection between human and computer. This required robust algorithm that is resistant in different light, different skin color and person-independent. Lip detection is a critical preprocessing step in many human-oriented applications, such as speech recognition and dental application [2]. The goal is to exploit the facial visual information of lip movements that contain the image information of speech.

There are several techniques for lip area detection. A large category of techniques are model-based. In these techniques, at the first, a model of the face is built. Then the structure of lip area is described by a set of model parameters [3]. These techniques include snakes, active contour models and several other parametric models. The most advantage of these systems is that the important features are represented in a low-dimensional parameter space and the calculation complexity is decreased. Also, these systems are good performance in conditions such as rotation, scaling and illumination variation. These models have some problems. For example sometimes large training set is needed to cover the high variability range of lips. For lip detection several color spaces and segmentation techniques were used. For example, Gomez et al [4] have used components of RGB color space to create new image by a linear combination of red, green and blue colors. Hsu et al [5] have used *YCbCr* color space to create new image by non-linear combination of *Cb* and *Cr* values. Nasiri et al [2] have used PSO approach to obtain optimized map for extracting lip area in face image.

In this paper, for lip detection we use the enhanced version of *Lip-Map* that proposed in [5]. For better separation between lip and skin pixels, we multiply *Lip-Map* by saturation component of HSI color space. In order to reduce required calculation, we first remove upper half of the face image. After this step, we estimate the lip area. For this purpose, we divide remaining lower half part of into some parts and calculate standard deviation for each part. Based on standard deviation of each part, we determine lip area. Finally, for extract lip pixels from skin pixels, we obtain optimum threshold value to convert the gray scale image into binary image. White pixels in binary image are lip region.

The rest of this paper is organized as follows: section 2, describes our lip detection algorithm. In section 3, we

To cite this paper: Kalbkhani H, Chehel Amirani. M. 2012. An Efficient Algorithm for Lip Segmentation in Color Face Images Based on Local Information. *J. World Elect. Eng. Tech.* 1(1): 12-16.
Journal homepages: http://www.jweet.science-line.com/

present our experimental results on CVL face database. And finally, section 4 concludes this paper.

PROPSED LIP DETECTION ALGORITHEM

At the first step for lip detection, we must choose efficient color spaces. In these color spaces lip must be separable from the skin. According to previous works we choose *YCbCr* color space and for better separation, we propose to use *HSI* color space. From these color spaces, we choose *Cb*, *Cr* and *Saturation* components for our work.

or complexity reduction and better detection of lip area, we remove the upper half of the face image, because we know that the lip is placed at the lower half part of the face area. Then, we propose to divide remaining image into some equal parts and search for lip. Finally by thresholding approach, we separate the lip pixels from the skin pixels. In the following, we introduce these steps in details.

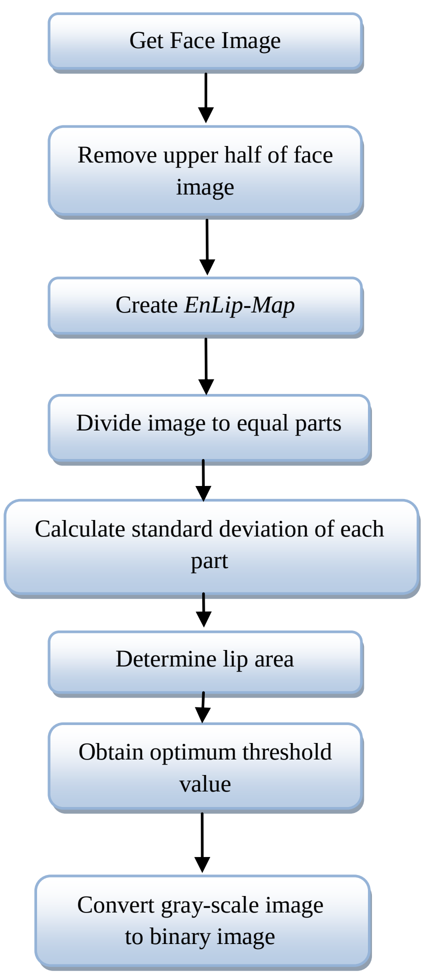


Fig. 1 Block diagram of our algorithm.

A. Choose efficient color space

According to the arrangement of the facial components, we know that the lip area is placed at the lower half part of the face area. So that, in order to reduction calculation rate, and increase the speed of detection, we propose to remove the upper half of the face image. Then we keep the lower half of the face area; and apply the remaining of our algorithm on this part of face image.

Hsu et al [5] have proposed a method for lip detection based on *YCbCr* color space. Pixels of lip area have stronger red component and weaker blue component than other facial regions. Therefore, the chrominance component *Cr* has greater value than the *Cb* in the lip region. Also, lip region has relatively low *Cr/Cb* value, but it has high Cr^2 value. Then, *Lip-Map* is constructed as follows:

$$Lip-Map = C_r^{\,2} \times (C_r^{\,2} - \eta(C_r / C_b))^2 \qquad (1)$$

$$\eta = 0.95 \frac{\sum C_r^{\,2}}{\sum C_r / C_b} \qquad (2)$$

where both Cr^2 and *Cr/Cb* are normalized to the range [0, 1]. We also normalize *Lip-Map* to the range [0, 1]. The parameter η is defined as a ratio of average Cr^2 to the average *Cr/Cb*.

The *Lip-Map* can be emphasizes the lip pixels well. But sometimes in a kind of the people, skin pixels have more common characteristics as lip pixels. This may degrade the performance of lip pixels segmentation. In Fig. 2, two images are presented, that in the right image separation between skin and lip is good, and in the left one separation is degraded. So, for better separation, we tested other color spaces. We searched for spaces that different between lip and face pixels is considerable. Finally, we choose saturation component of *HSI* color space. For better separation, saturation component is multiplied with *Lip-Map*. So, enhanced *Lip-Map* is obtained by bellow equation.

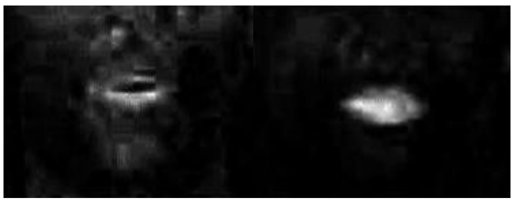
$$EnLip-Map = S \times \{C_r^{\,2} \times (C_r^{\,2} - \eta(C_r / C_b))^2\} \qquad (3)$$

Fig. 3 shows *EnLip-Map* of two images in Fig. 2.

B. Estimate lip area

Before applying thresholding approach, we estimate the lip location in the face image. According to the face and lip size, we propose to segment the lower half part of the face image in two directions.

(a)



(b)

Fig. 2 (a) Original face images of different people; (b) *Lip-Map* of original images, respectively.

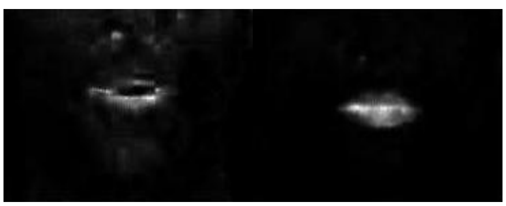


Fig. 3 The *Enhanced Lip-Map* of images in Fig. 2.

We first segment it into three equal parts according to x-axis direction. Then for each part, we calculate the standard deviation of the pixels of *EnLip-Map*. Lip in x-axis direction is placed in the part that has the maximum standard deviation. Also, we segment the lower half part of the face image into three equal parts according to y-axis direction. Then for each segment, we calculate the standard deviation of the pixels of *EnLip-Map*. Lip in y-axis direction is placed in the part that has the maximum standard deviation.

According to the segmentation of image into three equal parts in x and y-axis direction, we have nine blocks. According to the two segments (in x-axis and y-axis) that have maximum standard deviation, we choose one of the nine blocks as lip area. But, sometimes whole of the lip is not placed into one block. So that, we propose to consider the half of the blocks that are in the neighboring of the block, which obtained previously. In this manner, we obtain lip area in the face image. Fig. 4 illustrates these steps.

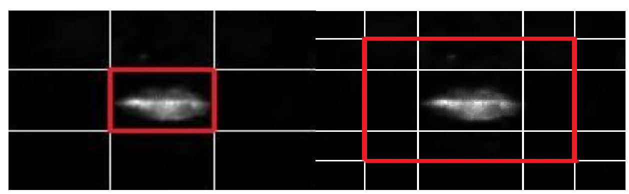
C. Separate lip pixels from skin pixels

After determining the lip area, we must extract lip pixels from the skin pixels. To enhance the gray-scale *EnLip-Map* and improve the separation, we propose to apply the *Top-Hat* morphological operation [6] with ball structuring element. The *Top-Hat* transformation of a gray-scale image f with structuring element b is defined as f minus its opening:

$$T_{hat}(f) = f - (f \circ b) \tag{4}$$

where $\circ$ demonstrates opening operation. After this step, we use threasholding approach. In any threasholding approach, we need to the threshold value for compare gray-scale image with this threshold value. And then convert the gray-scale image into binary one. To obtain the optimum threshold value (*Th*), we use the *Otsu'*s method [6].

(a) (b)



(c) (d)

Fig. 4 (a) Segments in x-axis direction, standard deviations from up to bottom are 0.0092, 0.1368, and 0.0034; (b) segments in y-axis direction, standard deviations from left to right are 0.0068, 0.1355 and 0.0065; (c) selected block according to the standard deviations; (d) lip area.

The method is optimum, because it maximizes the between-class variance, a well-known measure used in statistical discriminant analysis. The important property of this method is that based entirely on computations performed on the histogram of an image, that histogram is an easily obtainable 1-D array.

After obtaining the optimum threshold value, we convert the *EnLip-Map* to binary image (*Bin-Lip*) as follows:

$$Bin-Lip = \begin{cases} 1 & if \quad EnLip-Map \geq Th \\ 0 & if \quad EnLip-Map < Th \end{cases} \tag{5}$$

For noise reduction, after apply threshold, we remove connected components that have very low number of white pixels. Fig. 5 shows the binary image of *EnLip-Map* in Fig. 4.

RESULTS

For evaluation the performance of our new algorithm, we apply it on the images of CVL face database [7]. This database is consists of images from 114 persons. There are 7 images for each person. These images are labeled by side view that includes far left, angle 45˚, angle 135˚ and far right, frontal view that includes serious expression, smile (showing no teeth) and smile (showing teeth). These images are taken under uniform illumination and no flash and projection screen are in the background. All of images are size of 640×480 pixels and have JPEG format. To evaluate the new method, we choose images that have frontal view.

In TABLE 1 we have presented the results of our lip area detection and lip segmentation method. These results indicated that our algorithm can find lip area efficiently. If lip area is found properly, the separation between lip and skin is enhanced. The average rate of correct lip area detection is 98.76%.

The performance of our algorithm is compared with the method of [2] and [5]. Results indicate that our algorithm has better efficiency than previous algorithms. Adding the effect of saturation component and determining lip area before segmentation remove other redundant regions.

These parameters are most effective in increasing the rate of correct detection. After apply threshold value if the number skin pixels are comparable with the lip pixels or the number of lip pixels is very low, we assume that our algorithm is failed in separating lip pixels from skin pixels. In TABLE 2 we compare our results in lip segmentation with the results of [2] and [5]. In Fig. 6 we present some images with extracted lip region. These images indicate the performance of our algorithm.

(a) (b)

Fig. 5 (a) Binary image of the lip area; (b) lip pixels in the original image separated from the skin pixels.

Table 1. Results of our lip area detection and lip segmentation.

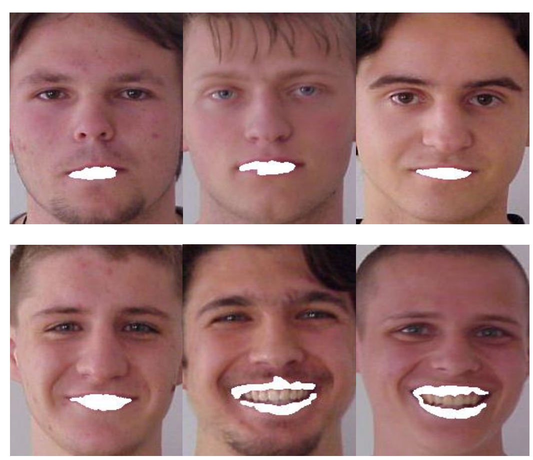
Expression	Accuracy (%)	
	Lip area detection	**Lip segmentation**
Serious	96.16	94.49
Smile (showing no teeth)	100	93.57
Smile (showing teeth)	99.05	92.45
Total	**98.40**	**93.5**

Table 2. Comparison between the results of our algorithm and other methods on CVL database.

Expression	Accuracy (%)		
	Our algorithm	**Method of [2]**	**Method of [5]**
Serious	94.49	92.72	84.54
Smile (showing no teeth)	93.57	90.99	82.72
Smile (showing teeth)	92.45	89.09	80.90
Total	**93.5**	**91.22**	**82.72**

CONCLUSION

In this paper we proposed a novel algorithm for lip area detection and lip segmentation in color face images based on local information. Local information in our algorithm is standard deviation of pixels of different parts in lower half part of face image. To reduce the complexity and achieve better results, we propose to remove the upper half of the face image. For detection of lip area, we used the enhanced version of *Lip-Map* that proposed by Hsu et al. We multiplied this *Lip-Map* by saturation component. After dividing lower half image into some equal parts, in order to finding lip area, we calculate standard deviation of pixels in each area. Then, for separating lip pixels from skin pixels, we obtain optimum threshold value by *Otsu*'s method. Our experimental results show that this algorithm can find lip area precisely. The rate of lip segmentation from skin is also better than other method.



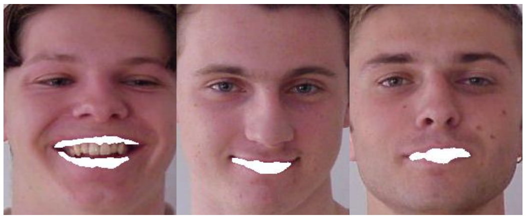


Fig. 6 Some images from the CVL face database with the different expressions and segmented lip pixels.

REFERENCES

[1] J. Shin, B. Jun, D. Kim, "Robust two-stage lip tracker", *IEEE Int. Symposium on Signal Processing and Information Technology (ISSPIT)*, pp. 265-270, Dec. 2009.

[2] J. A. Nasiri, H. S. Yazdi, M. A. Moulavi, M. Rouhani, A. E. Shargh, "A PSO tuning approach for lip detection on color images", *Second UKSIM European Symposium on Computer Modeling and Simulation*, pp. 278-282, Sept. 2008.

[3] E. Skodras and N. Fakotakis, "An unconstrained method for lip detection in color images", *IEEE Int. Conference on Acoustics, Speech and Signal Processing (ICASSP),* pp. 1013-1016, May 2011.

[4] E. Gomez, C. Travieso, J. Briceno, M. Ferrer, "Biometric identification system by lip shape", *Proceedings of the 36th Annual International Carnahan Conference on Security Technology.*, pp. 39-42, 2002.

[5] R. L. Hsu, M. Abdel-Mottaleb, and A. K. Jain, "Face detection in color images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 696-706, May 2002.

[6] R. C. Gonzalez, and R. E. Woods, "*Digital Image Processing*", Prentice Hall, 3rd Edition, 2008.

[7] Computer Vision laboratory, Faculty of Computer and Information Science, University of Ljubljana, Slovenia, *http://www.lrv.fri.unilj.si/facedb.html*