A PSO Tuning Approach for Lip Detection on Color Images

Jalal A. Nasiri
Communication and Computer Research Center
Department of Computer Engineering
Ferdowsi University of Mashhad, Iran
j.nasiri@wali.um.ac.ir

H. Sadoghi Yazdi
Department of Electrical Engineering
Tarbiat Moallem University of Sabzevar, Iran
sadoghi@sttu.ac.ir

M. Amir Moulavi
Networking Department
Information Technology Service (ITS) Center
Ferdowsi University of Mashhad, Iran
moulavi@acm.org

M. Rouhani
Department of Electrical Engineering
Islamic Azad University, Gonabad Branch, Iran
m.rouhani@ieee.org

A. Eshghi Shargh
Faculty of Engineering
Islamic Azad University of Mashhad, Iran
aref.cs@gmail.com

Abstract

Lip detection is used in many applications such as face detection and lips reading. We have proposed a novel approach for detecting lip using Particle Swarm Optimization (PSO). PSO is used to obtain an optimized map. The image is mapped to YC_bC_r color space. The main idea of the method is based on that lip has the high values of C_r and low values of C_b. The proposed algorithm has been examined on CVL and Iranian databases and we have reached to the %92 correction rate which comparing to the previous approach, there is %11 increase in lip detection. We find out that the proposed algorithm is flexible which it is independent of lightening conditions.

1. Introduction

One of the most critical preprocessing steps in many human-oriented applications is Lip Detection. Many various applications use this technique such as speech recognition, lip reading, audio interaction, dental application [9], [12].

Researcher has used either RGB or Segmentation approaches for lip detection. In RGB approach, Gomez et al [6] has proposed an algorithm in which the image is transformed by a linear combination of red, green and blue chrominance components of the RGB color space. To highlight the details of the lip envelop a high pass filter is then applied to the transformed image. The two generated images are then converted to obtain a binary image. The largest object in this binary image is recognized as the lip. Chang T. C. [1] has used the skin region identification to separate the non-skin holes from skin regions in order to extract facial features from the image. A thresholding box is then created and by which the image is searched line by line for the region to meet the certain criteria. Sadeghi M. et al [11] has proposed a modified version of the predictive validation technique that allows the use of the full covariance matrices used to select the model parameters. Then a Bayesian rule which is based on a subsequent grouping of the mixture components is used to recognize each pixel as lip or non-lip. Another technique is the hybrid edge which is exploited in the work of Eveno, N. et al [5]. In this method, hybrid edges combine pseudo hue and luminance information of the upper, middle and lower section of the lip.

In image segmentation approach, many algorithms have been proposed. For color image segmentation, histogram-based and cluster-based methods have been widely used. Cheng et al [3], [2] proposed a histogram segmentation technique which involves performing a fuzzy partition on a two-dimensional histogram based on the maximum fuzzy entropy principle. In [8], the color lip region is segmented

using a fuzzy thresholding algorithm with connectivity processing. In [4], a hue filter is used to weight the red hue, which is presumed to be the lip region. Then thresholding is used to segment the lip region.

The remainder of this paper is organized as follows; in section 2 an overview of PSO is given. Lips detection with PSO is investigated in section 3. In section 4, we propose our experimental results and we conclude in section 5.

2. Particle Swarm Optimization

In this section we give a brief and short description of what PSO is. Particle Swarm Optimization (PSO) is a population based stochastic optimization technique developed by Dr. Eberhart and Dr. Kennedy in 1995, [7] inspired by social behavior of bird flocking or fish schooling. In PSO, the potential solutions, called particles, fly through the problem space by following the current optimum particles.

The system is initialized with a population of random solutions and searches for optima by updating generations. Each particle keeps track of its coordinates in the problem space which are associated with the best solution (fitness) it has find currently. It can be taken as a particle that flying in the n-dimensional space. In order to find the optimal solution, each particle adjusts its flying velocity according to its own flying experience and its companions' flying experience at each iteration. Hence each particle has a position (x_{id}) and a speed (v_{id}) in the multidimensional space [7]. The particle position indicates the possible solution in the multidimensional space and the speed indicates the amount of change between the current and next positions. The algorithm stores the previous best position of each particle (p_{id}) and the best global position of particles (p_{gd}) [7].

$$\begin{aligned} v_{id} &= wv_{id} + c_1 * rand_1 * (p_{id} - x_{id}) \\ &\quad + c_2 * rand_2 * (p_{gd} - x_{id}) \\ if\ |v_{id}| > v_{max} \quad v_{id} &= sign(v_{id} v_{max}) \qquad (1) \\ x_{id} &= x_{id} + v_{id} \qquad (2) \end{aligned}$$

Where w is the inertia weight and c_1 and c_2 are the acceleration constants.

PSO has been successfully applied in many research and application areas. The main reason for choosing PSO as a basis for our work is that there are few parameters to adjust. One version, with slight variations, works well in a variety of applications.

3. Lips Detection with PSO

In this section we discuss the details of our proposed algorithm for detecting lip by using particle swarm optimization. The facial color images are converted to YC_bC_r color space.

3.1. Lips Detection Algorithm

The main idea of lip detection method is based on characteristics of lip in YC_bC_r color space which demonstrates that lib region have high C_r and low C_b values [10]. In order to detect the lip, the following formula should be computed on every pixel of the image to separate the lip region from non-lip region. The formula is as follows:

$$LipMap = C_r^2 \left(C_r^2 - \eta \frac{C_r}{C_b} \right)^2 \qquad (3)$$

$$\eta = 0.95 \frac{\frac{1}{m} \sum C_r(x,y)^2}{\frac{1}{m} \sum C_r(x,y)/C_b(x,y)} \qquad (4)$$

Where $(C_r)^2$,(C_r/C_b) all are normalized to the range [0 1]. This formula is designed to brighten pixels with high C_r and low C_b values. $(Cr)^2$ emphasizes pixels with higher C_r value and also (C_r/C_b) component completes our idea that lip regions have high C_r and low C_b values. Our simulation results show that this formula does not work properly for various kind of images including the color of skin. Moreover it is not flexible under various lightening conditions. We need an optimization algorithm that clusters the pixels of lip from non-lips' pixels. In Fig. 1 some of the lips detected with this formula are demonstrated.

3.2. PSO Tuning Approach

As discussed earlier, we need an approach that can cluster lip from other face region efficiently. We have proposed a novel approach based on particle swarm optimization (PSO) and YC_bC_r color characteristic of lip.

Our goal in the proposed algorithm is to obtain the values of A and B in the following formula. The initial values of A and B are random numbers [0 1]. This should be accomplished in a way to separate the lip's pixels from non-lip ones:

$$PSO - LMap = A * C_r^2 \left(A * C_r^2 - B * \frac{C_r}{C_b} \right)^2 \qquad (5)$$

In order to achieve the optimal values of A and B, two sets which include the pixels of lip and pixels of skin are created. These two sets are used to tune the A and B values and are concatenated to generate a new set $train$. For every pixels of set $train$, formula 5 is calculated to obtain the

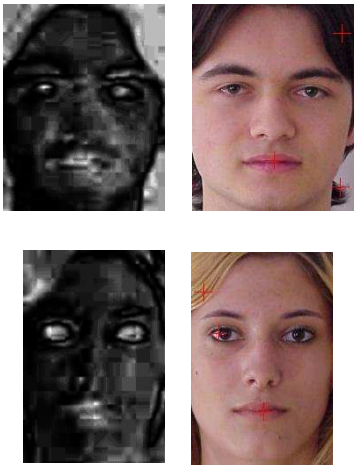

Figure 1. Sample of Lip Detection with formula 3 and 4

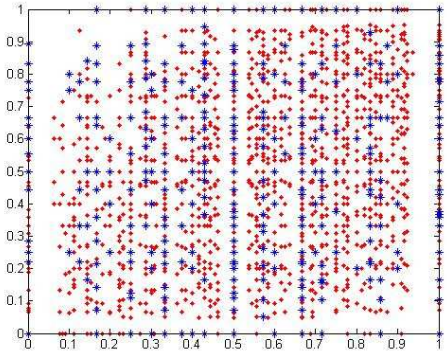

Figure 2. Lips and non-lips pixel before PSO tuning

$lip - map$ set. A clustering algorithm such as FCM is used to cluster the $lip - map$ set.

The number of pixels which are clustered as a wrong category is computed. This number is the $fitness$ function which should be minimized.

Particles fly in two dimensions A and B until the optimized point which is the minimized value of $fitness$ function, is reached.

In the following figures 2 and 3, values of lip's pixels and face's pixels are demonstrated before and after PSO tuning. As indicated in the figures, the first image is a high density composition of two set of pixels, however in the second one the pixels are separated properly.

4. Experimental Results

This section provides simulation results to evaluate our algorithm with and without PSO tuning. We apply our algorithm on CVL [13] and Iranian databases. Summary of the detection results (including the number of images, detection rates) on the CVL and Iranian databases are presented in Table. 1 and Table. 2, respectively. The detection rate is computed by the ratio of the number of correct detection to that of all the images tested. Sample of detections on CVL and Iranian databases are demonstrated in Fig. 4 and Fig. 5.

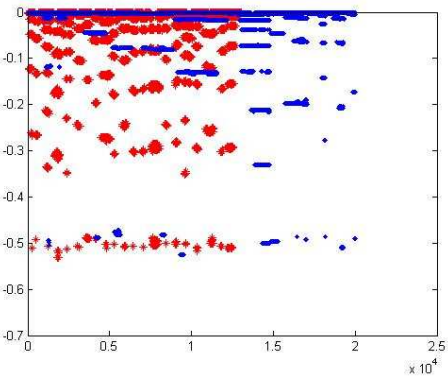

Figure 3. Lips and non-lips pixel after PSO tuning

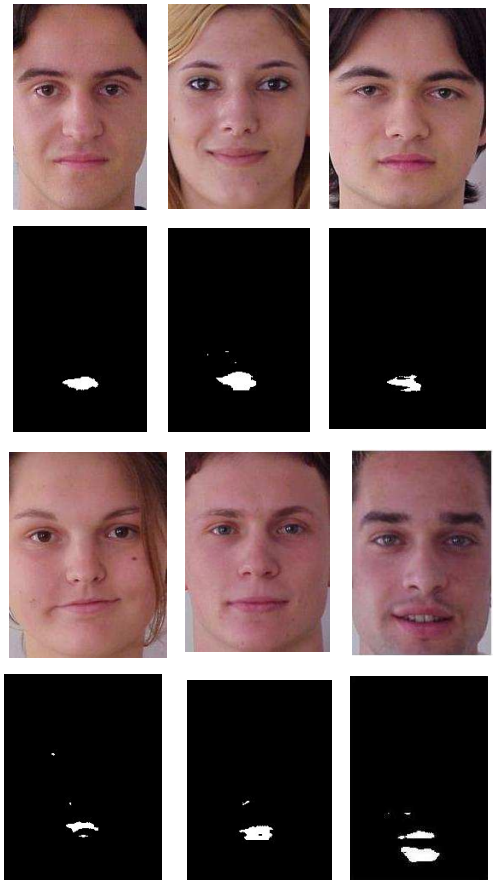

Figure 4. Sample of CVL Detection

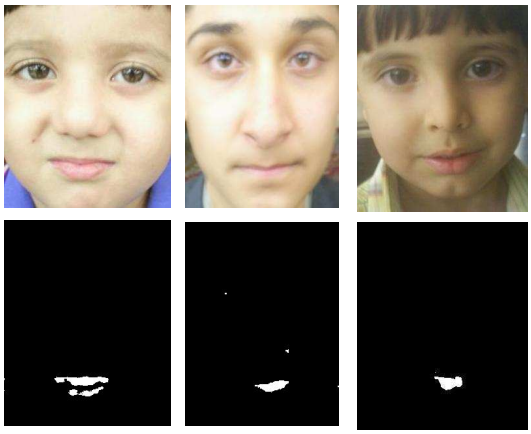

Figure 5. Sample of Iranian Detection

Table 1. Results on CVL database without PSO

Expression	*Serious*	*Smile*	*Grin*	*Total*
No. of image	110	110	110	330
Data Rate (%)	84.54	82.72	80.90	82.72

Table 2. Results on the Iranian database without PSO

Gender	*Female*	*Male*	*Total*
No. of image	28	22	50
Data Rate (%)	78.56	81.18	80

Table 3. Results on CVL database with PSO

Expression	*Serious*	*Smile*	*Grin*	*Total*
No. of image	110	110	110	330
Data Rate (%)	92.72	90.99	89.09	91.22

Table 4. Results on the Iranian database with PSO

Gender	*Female*	*Male*	*Total*
No. of image	28	22	50
Data Rate (%)	92.85	91	92

CVL Database CVL database consists of head and shoulder images taken from 114 people in 7 kinds of expressions. Among 7 images taken from a person, 3 of them are suitable for our purpose. These three photos are frontal view and with different expressions: serious, smile and grin.

Iranian Database Iranian database consists of head and shoulder images taken from 50 people. Images in Iranian database are taken under various lighting conditions.

5. Conclusions

In this paper we have presented a lip detection algorithm based on PSO for color images. Our method detects lip in face image which is extracted over the entire image. We have achieved a method for lip detection regardless of color of skins and lighting conditions. Our simulation results demonstrate the effectiveness of improvement phases have increased the correct detections by about %11 and we reach the detection rate about %92 in total.

6. Acknowledgment

We would like to thank *Communication and Computer Research Center* of Ferdowsi University of Mashhad, Iran.

References

[1] T. Chang, T. Huang, and C. Novak. Facial feature extraction from color images. In *Proceedings of the 12th IAPR International. Conference on Pattern Recognition, Computer Vision & Image Processing*, pages 39–43, 1994.

[2] H. D. Cheng, J. R. Chen, and J. Li. Threshold selection based on fuzzy c-partition entropy approach. *Pattern Recognition*, 31(7):857–870, 1998.

[3] H. D. Cheng, Y. H. Chen, and X. H. Jiang. Thresholding using twodimensional histogram and fuzzy entropy principle. *IEEE Transaction on Image Processing*, 9:732–735, 2000.

[4] T. Coianiz, L. Torresani, and B. Caprile. 2d deformable models for visual speech analysis. In *Speechreading by Humans and Machines*. 1996.

[5] N. Eveno, A. Caplier, and P.-Y. Coulon. A parametric model for realistic lip segmentation. In *Proceedings of the 7th International Conference on Control, Automation, Robotics and Vision*, volume 3, pages 1426–1431, 2002.

[6] E. Gomez, C. Travieso, J. Briceno, and M. Ferrer. Biometric identification system by lip shape. In *Proceedings of the 36th Annual International Carnahan Conference on Security Technology.*, pages 39–42.

[7] J. Kennedy and R. C. Eberhart. Particle swarm optimization. In *Proceedings of IEEE International Conference*, pages 1942–1948, 1995.

[8] S. Lucey, S. Sridharan, and V. Chandran. Chromatic lip tracking using a connectivity based fuzzy thresholding technique. In *5th International Symposium on Signal Processing Applications ISSPA99*, 1999.

[9] G. Lv, R. Zhao, D. Jiang, Y. Li, and H. Sahli. A robust visual feature extraction based btsm-lda for audio-visual speech recognition. In *Second International Conference on Communications and Networking*, pages 637–641, 2007.

[10] H. Rein-Lien, A. M. Mohamed, and K. J. Anil. Face detection in color images. *IEEE Trans. Pattern Analysis and Mechine Intelligence*, 24(5), 2002.

[11] M. Sadeghi, J. Kittler, and K. Messer. Modelling and segmentation of lip area in face images. In *Proceedings of IEE Conference on Vision, Image and Signal Processing*, volume 149, pages 179–184, 2002.

[12] T. Saitoh and R. Konishi. Lip reading using video and thermal images. In *International Joint Conference SICE-ICASE*, pages 5011–5015, 2006.

[13] F. Solina, P. Peer, S. Juvan, and J. Kovac. Color-based face detection in the 15 seconds of fame art installation. In *Conference on Computer Vision / Computer Graphics Collaboration for Model-based Imaging, Rendering, image Analysis and Graphical special Effects*.