A Mouth Detection Approach Based on PSO Rule Mining on Color Images

Jalal A. Nasiri [α*], H. Sadoghi Yazdi [β], Mahmoud Naghibzadeh [α]

Department of Computer Engineering, Ferdowsi University of Mashhad, Mashhad, Iran [α]
Department of Electrical Engineering, Tarbiat Moallem University of Sabzevar, Sabzevar, Iran [β]
Commutation and Computer Research Center *
j.nasiri@wali.um.ac.ir, sadoghi@sttu.ac.ir,naghibzadeh@um.ac.ir

Abstract

Finding mouth in face is a bottleneck of many applications such as face detection and lip reading. In this paper explain a new approach for mouth detection using Particle Swarm Optimization (PSO). PSO is used to mining the rule of between pixels of mouth and other pixels an optimized map. The image is mapped to YCbCr color space. The main idea of the method is based on that Mouth has the high values of Cr and low values of Cb. 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 Mouth detection. We find out that the proposed algorithm is flexible which it is independent of lightening conditions.

Keywords: Mouth detection, PSO, Color images YCbCr Color Space.

1. Introduction

One of the most critical preprocessing steps in many human-oriented applications is Mouth Detection. Many various applications use this technique such as speech recognition, Lip reading, audio interaction, dental application [1,2].

Researcher has used either RGB or Segmentation approaches for Mouth detection. In RGB approach, Gomez et al [3] 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 Mouth 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 Mouth. Chang T. C. [4] 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 [5] 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 Mouth or non-Mouth. Another technique is the hybrid edge which is exploited in the work of Eveno, N. et al [6]. In this method, hybrid edges combine pseudo hue and luminance information of the upper, middle and lower section of the Mouth.

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 [7], [8] proposed a histogram segmentation technique which involves performing a fuzzy partition on a two-dimensional histogram based on the maximum fuzzy entropy principle. In [9], the color Mouth region is segmented using a fuzzy thresholding algorithm with connectivity processing. In [10], a hue filter is used to weight the red hue, which is presumed to be the Mouth region. Then thresholding is used to segment the Mouth region.

The remainder of this paper is organized as follows; in section 2 an overview of PSO is given. Mouth 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,[11] 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 [11]. 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}) [11].

$$\begin{aligned} v_{id} &= wv_{id} + c_1 * rand1 * (p_{id} - x_{id}) \\ &\quad + c_2 * rand2 * (p_{gd} - x_{id}), \\ if\ |v_{id}| &> v_{\max} \qquad v_{id} = sign(v_{id} v_{\max}) \end{aligned} \tag{1}$$

$$x_{id} = x_{id} + v_{id} \tag{2}$$

Where w is the inertia weight and c1 and c2 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. Mouth Detection with PSO

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

3.1 Mouth Detection Algorithm

The main idea of Mouth detection method is based on characteristics of Mouth in YCbCr color space which demonstrates that lib region have high Cr and low Cb values [12]. In order to detect the Mouth, the following formula should be computed on every pixel of the image to separate the Mouth region from non-Mouth region. The formula is as follows:

$$LipMap = Cr^2\left(Cr^2 - \eta\frac{Cr}{Cb}\right)^2 \qquad \textbf{(3)}$$

$$\eta = .95\frac{\frac{1}{m}\sum Cr(x,y)^2}{\frac{1}{m}\sum Cr(x,y)/Cb(x,y)} \qquad \textbf{(4)}$$

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 Cr and low Cb values. $(C_r)^2$ emphasizes pixels with higher Cr value and also (C_r/C_b) component completes our idea that Mouth regions have high Cr and low Cb values.

Our simulation results show that this formula does not work properly for various kinds 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 Mouth from non-Mouth pixels. In figure 1 some of the Mouth detected with this formula are demonstrated.

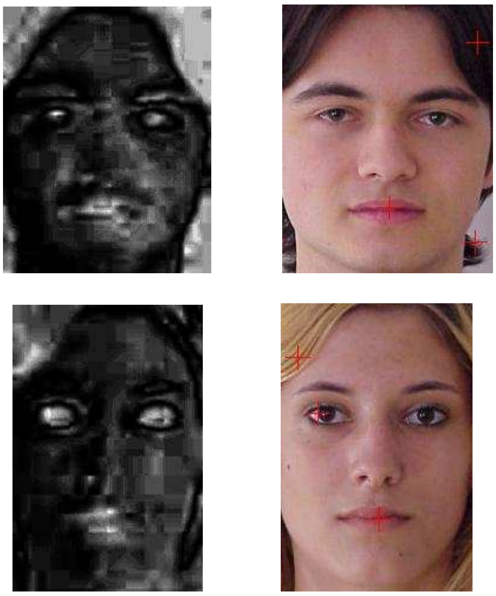


Figure 1: Sample of Mouth Detection with formula 3, 4.

3.2 PSO Tuning Approach

As discussed earlier**,** we need an approach that can cluster Mouth from other face region efficiently. We have proposed a novel approach based on particle swarm optimization (PSO) and YCbCr color characteristic of Mouth.
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 Mouth's pixels from non-Mouth ones:

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

In order to achieve the optimal values of A and B, two sets which include the pixels of Mouth 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 ***Mouth-map*** set. A clustering algorithm such as FCM is used to cluster the *Mouth-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 Mouth'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.

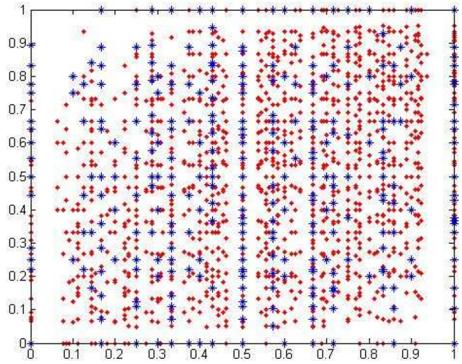


Figure 2: Mouth and non-Mouth pixel before PSO tuning.

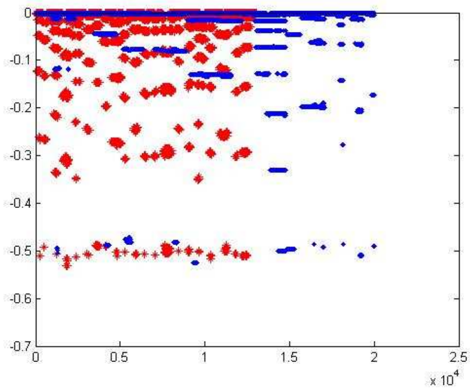


Figure 3: Mouth and non-Mouth pixel after PSO tuning

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. 3 and Table. 4, 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.

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.

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

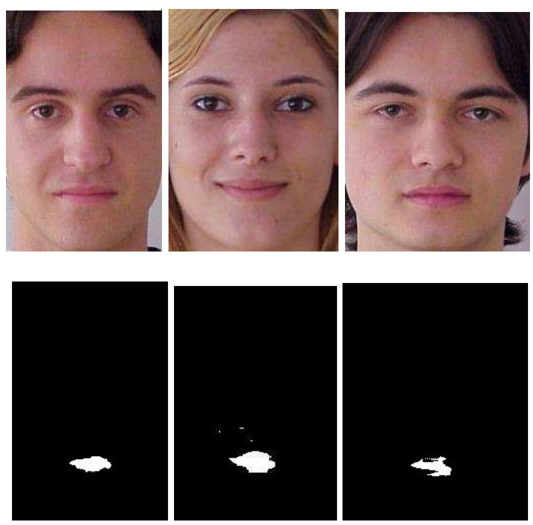
Gender	Female	Male	Total
No. of image	28	22	50
Data Rate (%)	78.56	81.18	80

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



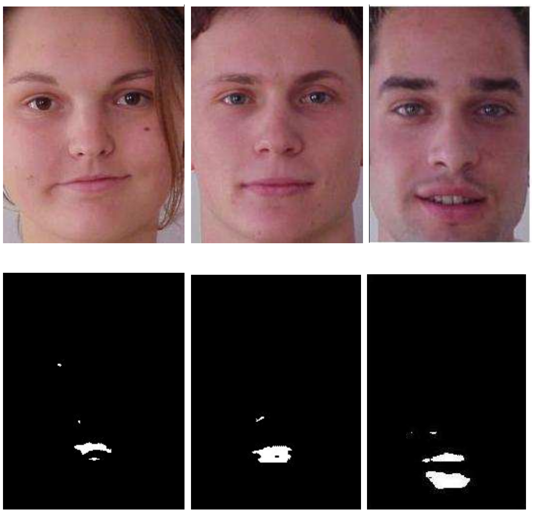


Figure 4: Sample of CVL Detection

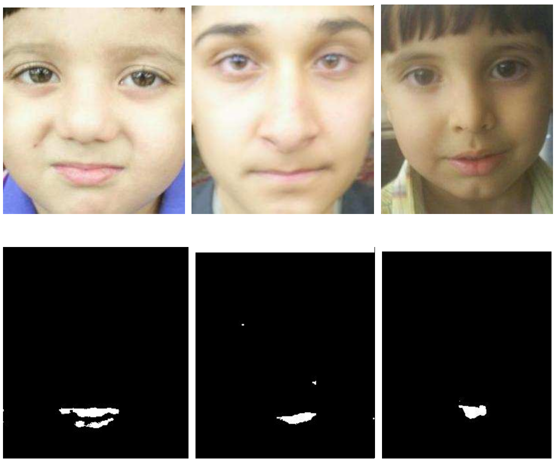


Figure 5: Sample of Iranian Detection

5. Conclusions

In this paper we have presented a Mouth detection algorithm based on PSO for color images. Our method detects Mouth in face image which is extracted over the entire image. We have achieved a method for Mouth 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, Ferdowsi University of Mashhad, Iran.

7. References

[1] Guoyun Lv, Rongchun Zhao, Dongmei Jiang, Yan Li, Sahli, H., "A Robust Visual Feature Extraction based BTSM-LDA for Audio-Visual Speech Recognition", Communications and Networking in China, 2007. CHINACOM '07. Second International Conference on 22-24 Aug. 2007 Page(s):637 - 641

[2] Saitoh, T., Konishi, R., "Mouth Reading using Video and Thermal Images",SICE-ICASE, 2006. International Joint Conference,Oct. 2006 Page(s):5011 – 5015

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

[4] Chang, T.C., Huang, T.S., Novak, C. "Facial feature extraction from color images". Proceedings of the 12th IAPR International. Conference on Pattern Recognition, Computer Vision & Image Processing. 1994. Vol. 2. pp. 39 - 43

[5] Sadeghi, M. Kittler, J. Messer, K. "Modelling and segmentation of lip area in face images" Proceedings of IEEE Conference on Vision, Image and Signal Processing. 2002. Vol. 149, pp. 179 - 184

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

[7] H. D. Cheng, Y. H. Chen, and X. H. Jiang, "Thresholding using twodimensional histogram and fuzzy entropy principle," IEEE Trans. Image Processing, vol. 9, pp. 732–735, Apr. 2000.

[8] H. D. Cheng, J. R. Chen, and J. Li, "Threshold selection based on fuzzy C-partition entropy approach," Pattern Recogn., vol. 31, no. 7, pp. 857–870, 1998.

[9] S. Lucey, S. Sridharan, and V. Chandran, "Chromatic lip tracking usinga connectivity based fuzzy thresholding technique," presented at the 5th Int. Symp. Signal Processing Applications ISSPA'99, Brisbane, Australia, Aug. 22–25, 1999.

[10] T. Coianiz, L. Torresani, and B. Caprile, “2D deformable models for visual speech analysis,” in Speechreading by Humans and Machines, D. G. Stork and M. E. Hennecke, Eds. New York: Springer-Verlag, 1996

[11]Kennedy, J., Eberhart, R. C., "Particle swarm optimization " Proceedings of IEEE International Conference Piscataway, NJ, pp. 1942-1948, 1995 .

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

[13] Franc Solina, Peter Peer, Samo Juvan, and Jure Kovac. Color-based face detection in the ”15 seconds of fame” art installation. In Mirage 2003, Conference on Computer Vision / Computer Graphics Collaboration for Model-based Imaging, Rendering, image Analysis and Graphical special Effects, pages 38–47, INRIA Rocquencourt, France, March 2003.