Research Journal of Recent Sciences ISSN 2277-2502
Vol. **2(1),** 17-20, January **(2013)** *Res.J.Recent Sci.*

A Fully Automatic and Haar like Feature Extraction-Based Method for Lip Contour Detection

Zahedi Morteza and Mohamadian Zahra
School of Computer Engineering, Shahrood University of Technology, Shahrood, IRAN

Available online at: www.isca.in
Received 26th September 2012, revised 27th October 2012, accepted 6th November 2012

Abstract

In this paper we propose a fully automatic and efficient method for lip contour detection. At first face region is separated from any background using a variation of AdaBoost classifiers trained with Haarlike features extracted from the face. After that by applying the second classifier trained with mouth Haarlike features, on face region mouth region is extracted, at last sobel edge detection operator applies on mouth region and lip contour is detected. Most of previous methods are based on image color intensity, such methods act very weak on pictures with low contrast and noise, furthermore they are very time consuming. Our method by using Haarlike features is very robust against low contrast and noisy images. It is also very fast, efficient and fully automatic.

Keywords: Lip contour detection, Haarlike features, AdaBoost classifiers, mouth region, face region.

Introduction

Lip contour extraction is beneficial because it forms a preliminary stage of face image analysis, which is essential for application areas including man-machine interaction based on the observed human behavior, video-telephony, face and person identification, bimodal speech recognition, face and visual speech synthesis, facial expressions classification, facial image transformation and etc. All of these applications require an efficient and fully automated mouth feature extraction method that can be achieved using an automatic lip contour detection technique.

Different approaches for the detection or segmentation of lip contours from face Images can be classified into three different categories these are: i. model based approach ii. color/gray level analysis based approach, and iii. approaches based on fuzzy clustering technique.

Model-based techniques like deformable templates[1], active shape models[2] and snakes[3] generally use a set of feature points to approximate the lip contours with spline functions and thus gives a model of the lip contour. In color/gray level analysis based approach[4, 5], lip segmentation is performed directly from the color or intensity space. This kind of algorithm often uses a color transformation or color filter to enlarge the difference between the lip and the skin. The resulting segmentation usually contains patches scattered in the image. For images with weak color contrast, this method cannot satisfactorily outline the boundary of the lip region.

Approaches based on fuzzy[6,7] clustering technique, the classical fuzzy c-means (FCM) algorithm generates a membership distribution using color information to minimize a fuzzy entropy measure. Lip segmentation is then performed using this distribution. In order to enhance the performance of FCM, different modifications of fuzzy clustering methods have been proposed in the past few years. Authors in[8] introduced a new dissimilarity measure that integrates the color dissimilarity and the spatial distance in terms of an elliptic shape function. Due to the presence of this elliptic shape function, their method can differentiate the pixels having similar color information but located in different regions. However, from experimental results, it is found that these approaches cannot perform satisfactorily especially when dealing with images of weak color contrast.

In this work of lip contour detection, first we have applied two different previously trained cascaded AdaBoost classifiers trained with Haarlike features in a hierarchical manner for detecting the face and mouth region respectively from an input image. Detection of lip contour is then performed from this isolated mouth region using sobel edge detection operator.

Experimental results shows that the proposed system is fast and can detect lip contours successfully even from images with noise and weak color contrast.

The rest of paper is organized as follow : a brief summary of proposed system is described in section II, in section III face and mouth region extraction method is presented. Experimental results are shown in section IV, and finally the conclusion of our paper is in section V.

Methodology

Overview of the proposed system: The proposed automated lip contour detection system consists of three parts. These are: i. face detection, ii. mouth region isolation, and iii. lip contour detection from the isolated mouth region. At first by applying a kind of filter we should reduce noises, and then we should convert the color image to a gray one by choosing an appropriate threshold value. For locating the approximate face region from an image, we have adopted the fast and efficient face finding method introduced recently by[9]. The same method, trained using mouth patches, has been used once again for isolating the mouth region from the detected face. Finally, detection of the lip contour has been performed from the isolated mouth region by applying sobel edge detection operator.

Face and mouth region isolation: The Detector in[9] consists of three parts. The first part is an efficient method of encoding the image data known as "integral image" that allows a very quick calculation of the used features. The second element is the application of a boosting algorithm known as AdaBoost[10] in selecting appropriate features that can form a template by modeling an object's variation. The third part is a cascade of classifiers that speeds up the detection process by quickly eliminating unlikely object regions. In our work, classification of images for detecting the face and mouth region has been performed using this method on the basis of the value of simple scalar features. The features are similar to Haar basis functions that operate on the grey level images, and the decision depends on the value of the difference of the sums computed over rectangular regions. Three types of features, introduced in[11], have been used in our work for both face and mouth region detection (figure-1). For reducing the features we can use genetic algorithm[11] and for classification the pictures we can use neural networks[12].

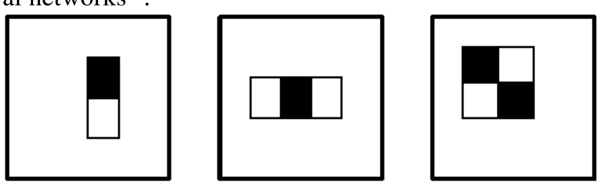

Figure-1
Features used in face and mouth detection[1]

The value of a two-rectangular feature is the difference between the sums of the pixels within two rectangular regions. The regions have the same size and shape and are either horizontally or vertically adjacent. A three rectangular feature computes the sum within two Outside rectangles subtracted from the sum in center rectangle. Finally a four-rectangular feature computes the difference between diagonal pairs of rectangles. Two other types of features can easily be generated by rotating the first two types by 90°. For computing the rectangle features very efficiently, an auxiliary image representation known as the "integral image" has been used here that was originally introduced in[13]. The integral image *ii* at location *(x, y)* is defined as the sum of the pixels of the rectangle ranging from the top left corner at *(0, 0)* to the left of *(x, y)* inclusive:

$$ii(x,y) = \sum_{x' \le x, y' \le y} i(x',y') \qquad (1)$$

Here *i* is the input image in question. The integral image can be computed in one pass over the original image by using the following recurrence relations:

$$\begin{aligned} s(x,y) &= s(x,y-1) + i(x,y) \\ ii(x,y) &= ii(x-1,y) + s(s,y) \end{aligned} \qquad (2)$$

Where *s(x, y)*, is the cumulative row sum and *s(x,-1) =ii (-1, y) =0*. Any rectangular sum can be computed in 4 array references using the integral image. Number of features calculated using the above method for each sub-window is far larger than the number of pixels in the sub-window. Although each feature can be computed very efficiently, computing the complete set is prohibitively expensive. So, a variant of AdaBoost[14] has been used both to select a very small number of significant features, and to train two different cascade of classifiers using the selected features for detecting the face and mouth region. Detailed algorithm for constructing such cascaded classifiers can be found in[13].

The method requires a set of positive and negative image samples of both face and mouth respectively to build such classifier. For the face detector, used positive examples are human faces and the negative examples are regions known not to contain a human face. Similarly, for the mouth region detector, image patches centered on mouth region have been used as positive samples, and the image patches that are randomly displaced from mouth but reside within the face region have been used as negative samples. To achieve robustness of the mouth detector against the variation of mouth region caused by different facial expressions, positive samples of the mouth are created from the images of various facial expressions. Total 2400 positive and 2600 negative samples have been used to train both the face and mouth detector or Small subsets of such positive samples are given in figure-2.

Finally, the trained detectors are scanned across the input image at multiple scales as well as locations in a hierarchical manner for detecting the face first and then to isolate the mouth region from the detected face.

Results and Discussion

Steps: As describe before to implement the proposed method we use AdaBoost classifiers to detect face and mouth regions. So, the classifier is trained using Haarlike features extracted from face and mouth. The only use of color intensity is very time consuming, so Haarlike features instead of using color intensity is introduced in[14].

AdaBoost function is consisting of two sections: the weak classifier and boosting section. The weak classifier tries to find the best threshold value for classification. Boosting section calls the weak classifier several times. After each classification weight value for misclassification samples is change. This method forms a sequence of weak classifiers that act as a strong one.

Extracting such features is describe in face and mouth region isolation section.

At last the sobel edge detector operator is applied on mouth region to extract lip contour.

Figure-2
A subset of positive face and mouth samples

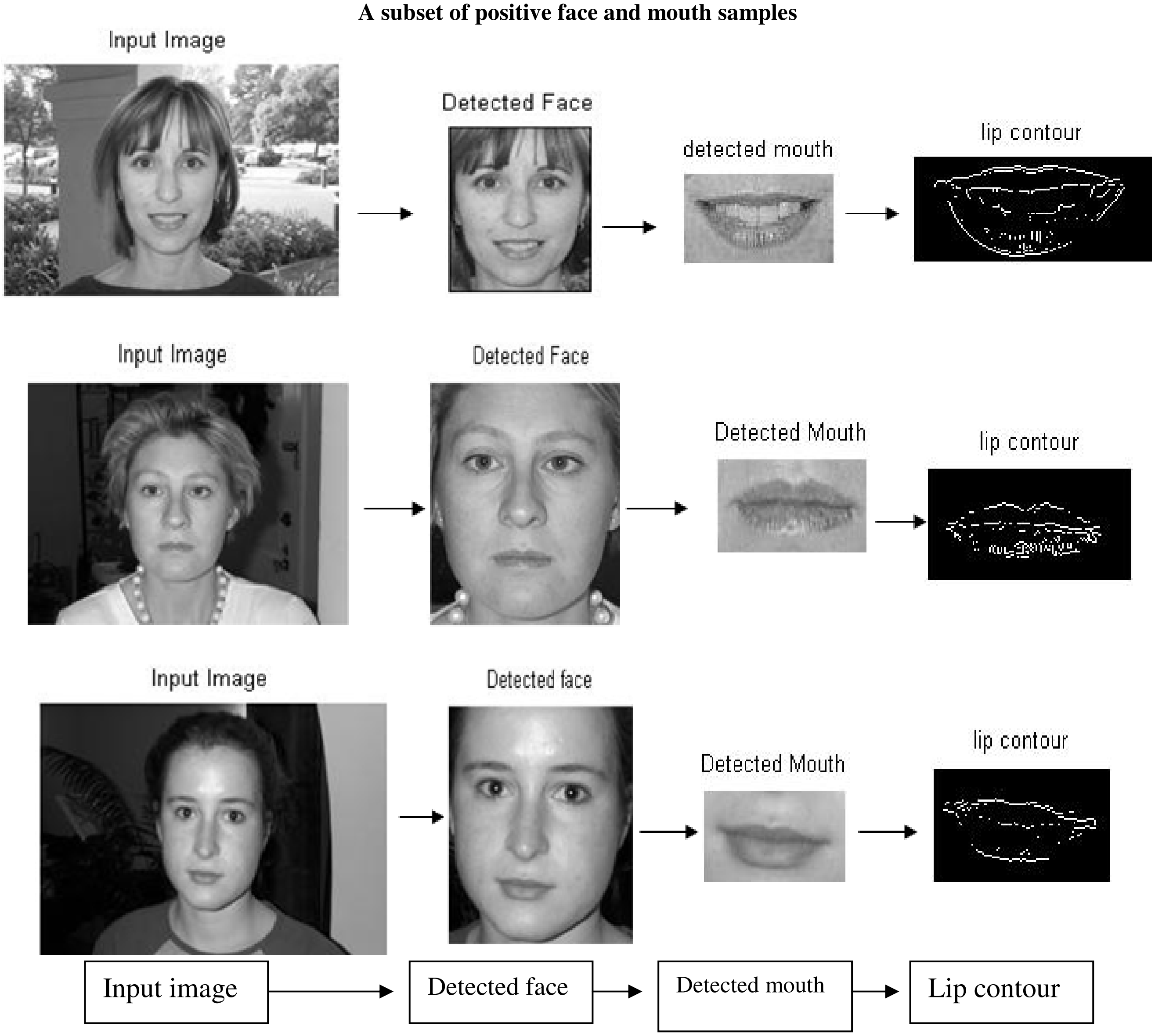

Figure-3
The results of proposed system

Experimental results: Haarlike features extracted from face are given to AdaBoost classifier as input and face region is extracted. Then by giving features extracted from mouth as input to second AdaBoost classifier the mouth region is extracted.

After that by applying the sobel edge detector on mouth region, lip contour is detected. For evaluating the performance of the developed automated lip contour detection system, it has been applied over *Cohn-Kanade AU-Coded Facial Expression* Database[15].

The *Cohn-Kanade AU-Coded Facial Expression* Database includes 2105 digitized image sequences from 182 adult subjects of varying ethnicity (69% female, 31% male, 81% Euro-American, 13% Afro-American and 6% other groups), performing multiple tokens of most primary FACS (Facial Action Coding System) action units. In figure-3 you can see the obtained results.

Conclusion

A fully automated technique of lip contour detection using Haarlike features is presented in this paper. Experimental result shows that the developed automated lip contour detection system is quite robust against the variability of lip contours emerged from situations like individual appearance (spatial variability), locutions, facial expressions (temporal variability) as well as lighting (spatiotemporal variability), and performs satisfactorily while dealing with the images of weak color contrast, on the other hand time processing of our system is optimized.

We believe that the proposed lip contour detection system can play a significant role in applications where fully automated detection of mouth feature or lip contour is necessary.

References

1. Yuille A.L., Kohen D.S. and P.W. Hallinan, Feature Extraction from Faces Using Deformable Templates, In *Proc. IEEE Intl. Conf. Computer Vision and Pattern Recog.,*San Diego, CA, 104–109 **(1989)**
2. Luettin J., Thacker N.A. and S.W. Beet., Visual SpeechRecognition Using Active Shape Models and Hidden Markov Models, In *Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Proc.,* Atlanta, GA, **2**, 817–820 **(1996)**
3. Kass M., Witkin A. and D. Terzopulos, Snakes: Active Contour Models, *Intl. J. Computer Vision,* **1(4)**, 321–331 **(1987)**
4. Eveno N., Caplier A. and Coulon P.Y. New Color Transformation for Lips Segmentation, In *Proc. IEEE 4 Workshop Multimedia Signal Proc.,* France, 3–8 **(2001)**
5. Wark T., Sridharan S. and Chandran V., An Approach to Statistical Lip Modeling for Speaker Identification via Chromatic Feature Extraction, In *Proc. 4th Intl. Conf. Pattern Recog.,* Brisbane, Australia, 123–125 **(1998)**
6. Shamshirband Shahaboddin and Za'fari Ali., Evaluation of the Performance of Intelligent Spray Networks Based On Fuzzy Logic, *Res. J. Recent Sci*., **1(8)**, 77-81, August (**2012**)
7. Patil Pallavi and Ingle Vikal, Obtaining a high Accurate Fault Classification of Power Transformer based on Dissolved Gas Analysis using ANFIS, *Res.J.Recent Sci*., **1**(**2**), 97- 99 **(2012)**
8. Leung S., Wang S. and Lau W., Lip Image Segmentation Using Fuzzy Clustering Incorporating an Elliptic Shape Function, *IEEE Trans. Image Proc.,* **13(1)**, 51–62 **(2004)**
9. Viola P. and Jones M.J., Robust Real-Time Face Detection, *Intl. J. Computer Vision.,* **57(2)**, 137–154 **(2004)**
10. Freund Y. and Schapire R.E., A Decision-theoretic Generalization of On-line Learning and an Application to Boosting, In *Proc. 2nd European Conf. Computational Learning Theory,* Barcelona, Spain, 23–37 **(1995)**
11. Sharifi M. and Shahriari B., Pareto Optimization of Vehicle Suspension Vibration for a Nonlinear Half-car Model Using a Multi-objective Genetic Algorithm, *Res.J.Recent Sci*., **1(8)**, 17-22 **(2012)**
12. Agbo G.A., Ibeh G.F. and Ekpe J.E., Estimation of Global Solar Radiation at Onitsha with Regression Analysis and Artificial Neural Network Models, *Res.J.Recent Sci*., **1(6)**, 27-31 **(2012)**
13. Viola P. and Jones M.J., Robust Real-Time Face Detection, *Intl. J. Computer Vision.,* **57(2)**, 137–154 **(2004)**
14. Papageorgiou Oren, and Poggio, A general framework for object detection, *International Conference on Computer Vision*, **(1998)**
15. Kanade T., Cohn J.F. and Tian Y., Comprehensive Database for Facial Expression Analysis, In *Proc. 4th IEEE Intl. Conf. Automatic Face and Gesture Recog.,* France 46–53 **(2000)**