Face and Lip Tracking for Person Identification*

Ying Zhang

Key Laboratory of Information Network Security, Ministry of Public Security,
People's Republic of China (The Third Research Institute of Ministry of Public Security),
339 Bisheng Road, Zhangjiang Hi-Tech Park, Pudong New Area, Shanghai, China
zhangying@stars.org.cn

Abstract. This paper addresses the issue of face and lip tracking via chromatic detector, CCL algorithm and canny edge detector. It aims to track face and lip region from static color images including frames read from videos, which is expected to be an important part of the robust and reliable person identification in the field of computer forensics. We use the M2VTS face database and pictures took from my colleagues as the test resource. This project is based on the concept of image processing and computer version.

Keywords: face recognition, lip tracking, computer forensics.

1 Introduction

Regarding the sustained increase of hi-tech crime, person authentication has aroused a lot of attentions in various fields especially in areas of high security. Thus there is an urgent requirement for robust and reliable identification technology from governments, the military, police, forensic scientists and commercial organizations. Based on the fact that most people are used to identify individuals by their faces, face recognition plays an important role during this process of identification.

Over the past ten years or so, face recognition has developed rapidly and become a popular area of research in computer vision and one of the most successful applications of image analysis and understanding [1]. For example, Chellappa et al. has demonstrated the survey of face detection as well as related psychological research in 1995. They considered static images and clips from videos respectively, generalized algorithms utilized for each one and analyzed their characteristics as well as advantages and disadvantages. [5]

Lip tracking is also an important tool for computer forensics. Sometimes the original evidences are possibly videos with strong noise while it is expected that the investigators could extract information from the voice. In this situation the technology will help forensic scientists make this via tracking the diversification of lip contour in real-time.

* This paper is supported by the Special Basic Research, Ministry of Science and Technology of the People's Republic of China, project number: 2008FY240200.

X. Lai et al. (Eds.): E-Forensics 2010, LNICST 56, pp. 282–286, 2011.

In this paper we will discuss a new way to implement face detection, which includes face detection, expression extraction and tracking of other features. And due to the importance of lip we select it as the representative from all the features and track its motion simultaneously.

2 Algorithms and Implementation

2.1 Face Region Segmentation

There are a lot of algorithms to segment face from the background image (e.g., pattern matching snakes, color localization and neural network). Here we use the chromatic method.

Rough Face Region Detection. Previous work [3] has proved that face region could be approximated via locating pixels in the following range:

$$L_{lim} \leqslant \frac{R}{G} \leqslant U_{lim} . \tag{1}$$

R and B stand for the red and green color component of each pixel respectively. And L_{lim} and U_{lim} are the thresholds which are dependent on the particular light over the facial part in the image [3].

The software ImageJ is utilized to split the color components and get two thresholds as shown in figure 1. After the segmentation, the candidate's points are marked by the color black and then we can get the rough face region.

Fig. 1. The thresholds of face segmentation implemented by ImageJ

Accurate Face Region Segmentation. From the figure 2 we can see that there are some noises in the result image processed by previous step. Thus the elimination is expected to be performed. Via computing the frequency of marked points, if there are some points which are not located in the main block they will be treated as noise and will be removed from the candidate list.

Fig. 2. Noise points

2.2 Lip Tracking

Rough Lip Region Detection. In this step, the two thresholds have been adjusted to locate lip pixels [3]. And then based on the theory that the lip is located in the lower half of face and it is usually symmetric about the vertical middle line of face, we could get rid of the extra points. In addition, we also need to merge broken lip regions which are brought about by the deficiency of lip thresholds.

Accurate Lip Region Detection. CCL (Connected Component Algorithm) is utilized here to find the largest blocks in the rough lip region.

Definition of CCL: The notation of pixel connectivity describes the relationship between two or more points. For two pixels to be connected they have to fulfill certain conditions on the pixel brightness and spatial adjacency [4].

Canny Edge Detector. We use Canny edge detector to describe the lip contour in the accurate lip region. The result of above steps is shown in figure 3.

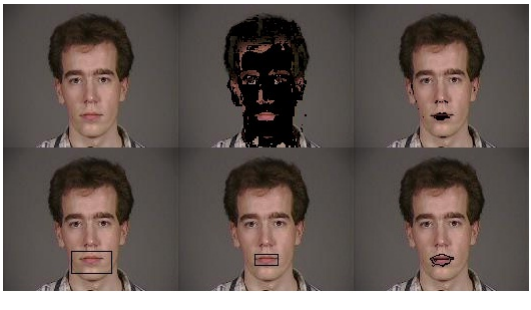

Fig. 3. Result images for face and lip tracking

3 Analysis of Results

3.1 Complexity of Algorithm

The complexity of this algorithm is O(facewidth×faceheight). This could be calculated by the following steps:

(1) Search the possible lip region, complexity here is O(facewidth*faceheight).
(2) Search the possible lip region, complexity here is O(facewidth*faceheight).
(3) Search the rough lip region, complexity here is O(half_faceheight*facewidth)
(4) Search the accurate lip region, complexity here is O(rough_lipwidth*rough_lipheight)
(5) Find the edge of lip, complexity here is O(accurate_lipwidth*accurate_lipheight)

According to the above deduction, the complexity is O(facewidth* faceheight). That means the consumptive time of this algorithm varies in the same manner as size of input image.

3.2 Veracity of Result

Here we evaluate the veracity via comparing the lip contour implemented by my algorithm to the one which is got by hand. The follow histograms show the distributions of lip edge points of the above two situations respectively.

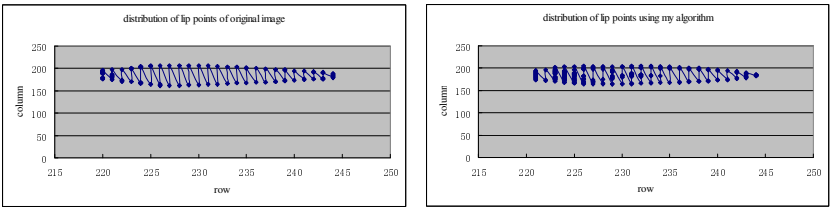

Fig. 4. Distribution of edge points by hand and my algorithm

And then we compare the pixels located in the edge of the two. According to the statistic data, 81.4% edge points have been included in the result.

3.3 Deficiencies

Easy to be Influenced by Other Conditions. The whole program is based on the chromatic algorithm. Point is that the chromatic difference is easy to be influenced by the camera series or the background light. For some images in which the rate of red and green color component didn't vary obviously the algorithm didn't work so well and sometimes even fails.

Only Suitable to Color Images. The basis of this algorithm is that the rate of red and green component is different for each part of the face. Hence it means that only color image is suitable instead of gray level image.

The Deficiency of Canny Edge Detector. Due to the shortage of canny edge detector there are some superfluous edges.

4 Future Application

The previous paper has mentioned that lip tracking system could be used in the security field especially for the field of computer forensics. For the reason that in some places where the speech signal is not so good or in the situation face detection is supposed to be helpful in the person authentication or in the case that lip reading is supposed to help the forensic scientists identify what people talk about in the videos, lip tracking is required to compensate the deficiency.

References

1. Grgic, M., Delac, K.: General Info (2005), http://www.face-rec.org/general-info/
2. Li, S.Z., Jain, A.K.: Handbook of Face Recognition, 10.1007/0-387-27257-7_17
3. Wark, T., Sridharan, S.: A Syntactic Approach to Automatic Lip Feature Extraction for Speaker Identification, Speech Research Laboratory, Signal Processing Research Centre, Queensland University of Technology, Australia (1998)
4. Fisher, R., Perkins, S., Walker, A., Wolfart, E.: Pixel Connectivity (2002), http://www.cee.hw.ac.uk/hipr/html/connect.html
5. Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A.: Face Recognition: A Literature Survey, University of Maryland, Sarnoff Corporation, National Institute of Standards and Technology, USA (2003)
6. Green, B.: Canny edge detector Tutorial (2002), http://www.pages.drexel.edu/~weg22/edge.html
7. Barnard, M., Holden, E.-J., Owens, R.: Lip tracking using pattern matching snakes. In: The 5th Asian Conference on Computer Vision (2002)
8. Mitsukura, Y., Fukumi, M., Akamatsu, N.: A Design of Face Detection System by using Lip Detection Neural Network and Skin Distinction Neural Network. Faculty of Engineering, University of Tokushima (2000)
9. Jiang, X., Wang, Y., Zhang, F.: Visual Speech Analysis and Synthesis with Application to Mandarin Speech Training, Department of Computer Science, Nanjing university, Nanjing Oral School (1999)
10. Gurney, K.: An Introductoin to Neural Networks. T.J. International Ltd., Padstow (1999)