Automatic Segmentation of Color Lip Images Based on Morphological Filter

Meng Li and Yiu-ming Cheung

Department of Computer Science, Hong Kong Baptist University
mli@comp.hkbu.edu.hk, ymc@comp.hkbu.edu.hk

Abstract. This paper addresses the problem of lip segmentation in color space, which is a crucial issue to the success of a lip-reading system. We present a new segmentation approach to lip contour extraction by taking account of the color difference between lip and skin in color space. Firstly, we obtain a lip segment sample via a color transformation sequence in 1976 CIELAB and LUX color spaces. Secondly, we establish a Gaussian model and make use of the hue and saturation value of each pixel within the lip segment to estimate the model parameters. Subsequently, the memberships of lip and non-lip regions are calculated, respectively. Thirdly, we employ morphological filters to obtain the desirable lip region approximately based on the memberships. Finally, we extract the lip contour via convex hull algorithm with the prior knowledge of the human mouth shape. Experiments show the efficacy of the proposed approach in comparison with the existing lip segmentation methods.

1 Introduction

In the past decade, lip segmentation has received considerable attention from the community because of its widespread applications[1], [2]. In general, lip segmentation is a non-trivial task because the color difference between the lip and the skin regions is not so noticeable sometimes. In the literature, a few image segmentation techniques have been proposed. One class of methods is based on the clustering with color features [3] provided that the number of clusters is known in advance. However, from the practical viewpoint, the number of clusters should be selected adaptively. Consequently, such a method is unable to operate fully automatically. Another class of widely-used methods is model-based ones [1]. Empirical studies have shown its success, but manually work on landmarks is needed for training.

In this paper, we will present a new method for automatic segmentation of lip images provided that the lower part of a face (i.e. the part between nostril and chin) has been available. A color transformation sequence is proposed to enlarge the distinction between the lips and the skin. Different from the existing methods, the proposed one only extracts a segment of lip rather than the whole lip region. Then, based on the lip segment sample, we establish a Gaussian model in a modified HSV color space so that the memberships of lip and non-lip region are calculated, respectively. We further utilize morphological filters to obtain the lip region candidate based on the two memberships. Finally, convex hull algorithm is employed to extract lip contour. Experiments have shown the efficacy of the proposed approach in comparison with the existing methods.

K. Diamantaras, W. Duch, L.S. Iliadis (Eds.): ICANN 2010, Part I, LNCS 6352, pp. 384–387, 2010.

2 Lip Membership Based on Color Space Transformation

We transform the source image into 1976 CIELAB color space and employ the histogram equalization to cover the a^* component to the range of [0, 255], denoted as I_{a^*}. Furthermore, we get the U component via the method proposed in [4] with histogram equalization, denoted as I_U.

Let $I_{sub} = I_{a^*} - I_U$.[1] We establish a Gaussian model for I_{sub} based on the gray-level value for each non-zero pixel with the mean $\hat{\mu}_{sub}$ and the standard deviation $\hat{\sigma}_{sub}$. The candidate lip segment can be obtained by

$$I_{candidate} = \begin{cases} 0 \ \ if \quad I_{sub} \leq \hat{\mu}_{sub} - 2\hat{\sigma}_{sub}, \\ \\ 1 \ \ otherwise. \end{cases} \tag{1}$$

The morphological reconstruction based method proposed in [5] is performed to suppress border connected noisy structures. The output image is denoted as $I^*_{candidate}$. In $I^*_{candidate}$, the nearest connected foreground block to gravity center, as the extracted segment, makes a segment sample that corresponds to the lip segment. Note that it is enough to extract a part of lip rather than the whole region because the extracted segment is used for sample data so as to establish a probability model.

For each pixel in HSV image, we perform the following transformation to get a vector: $C = (H \cdot cos(2\pi \cdot S), H \cdot sin(2\pi \cdot S))^T$. Then, we establish a probability model as follows:

$$P = \frac{1}{2\pi\sqrt{\hat{\Sigma}}} \cdot exp(-\frac{(X - \hat{\mu})\hat{\Sigma}^{-1}(X - \hat{\mu})^T}{2}) \tag{2}$$

where the mean $\hat{\mu}$ and the covariance matrix $\hat{\Sigma}$ can be evaluated by C vectors of the pixels in source image restricted by the lip segment region in the previous steps. As the input of the model of Eq.(2), the C vector for each pixel in the whole source image is obtained. Thus, we can calculate the lip membership denoted as M_{lip}.

Similarly, we can also establish a probability model to calculate the non-lip membership as $M_{non-lip}$. Moreover, considering the convenience of visibility, we project the memberships from $[0, 1]$ to $[0, 255]$.

3 Lip Contour Extraction

We obtain a mask image by letting

$$Mask = 255 - M_{non-lip} - I_U. \tag{3}$$

While the lip membership is labeled as marker, morphological reconstruction operation can be employed.

We further utilize a gray-level threshold selection method proposed in [6] to transform the reconstruction result into a binary image denoted as B_{RT} with boundary

[1] In this paper, all equations are employed in positive area. That is, as long as a result is negative, it will be set at 0 automatically.

connected structures suppressed, and mark the biggest continued foreground block by B_{lip_1}.

According to the following equation

$$I_{TTM} = I_U - I_{a^*}, \tag{4}$$

the region we have obtained covers the teeth, tongue and some parts of oral cavity approximately.

We further transform I_{TTM} into a binary image as B_{TTM} by the threshold selection method. Then, morphological closing is employed to $B_{RT} \cup B_{TTM}$ by performing a 5×5 structuring element operation. We select the biggest foreground block denoted as B_{lip_2} in the closing operation result. Hence, the binary image $B_{lip_1} \cup B_{lip_2}$ can represent the whole lip region. Furthermore, we utilize morphological opening with 3×3 structuring element to smooth the edge, resulting as B_{lip}. Finally, the quickhull algorithm proposed in [7] is employed to draw the contour of lip.

4 Experimental Results

Comparison is made to demonstrate the performance of the proposed approach with Liew03 proposed in [3], and Guan08 in [8]. Four databases: (1) AR face database [9], (2) CVL face database [10], (3) GTAV face database, and (4) a database established by ourselves, are used to test the accuracy and robustness in different capture environments. We randomly selected 800 images in total and manually segmented the lip to serve as the ground truth. Some segmentation results can be found in Figure 1.

Two measures (OL and SE) defined in [3] are used to evaluate the performance of the algorithms. Table 1 shows the segmentation results on the four different databases.

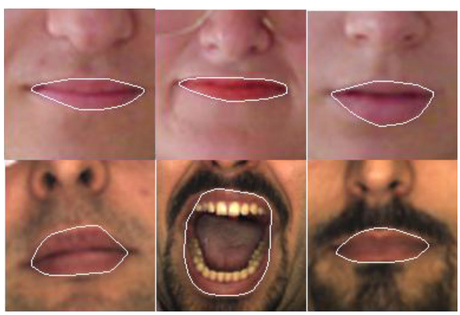

Fig. 1. Some samples of lip contour extraction in different databases

Table 1. The segmentation results across the four databases

Algorithm	Liew03	Guan08	Proposed Approach
average OL, %	80.73	45.10	**90.12**
average SE, %	20.15	55.21	**9.33**

5 Conclusion

In this paper, we have proposed a new approach to automatic lip segmentation via the probability model in color space and morphological filter. This approach features the high stability of lip segmentation and robust performance against the disparate capture environment and different skin color. Experiments have shown the promising result of the proposed approach in comparison with the existing methods.

Acknowledgment

The work described in this paper was supported by the Faculty Research Grant of Hong Kong Baptist University with the project code: FRG2/08-09/122 and FRG2/09-10/098, and by the Research Grant Council of Hong Kong SAR under Project HKBU 210309.

References

1. Matthews, I., Cootes, T., Bangham, J.: Extraction of visual features for lipreading. IEEE Transactions on Pattern Analysis and Machine Intelligence 24, 198–213 (2002)
2. Cetingul, H., Yemez, Y., Erzin, E., Tekalp, A.: Discriminative analysis of lip motion features for speaker identification and speech-reading. IEEE Transaction on Image Processing 15(10), 2879–2891 (2006)
3. Liew, A.W., Leung, S., Lau, W.: Segmentation of color lip images by spatial fuzzy clustering. IEEE Transactions on Fuzzy Systems 11(4), 542–549 (2003)
4. Lievin, M., Luthon, F.: Nonlinear color space and spatiotemporal mrf for hierarchical segmentation of face features in video. IEEE Transactions on Image Processing 13(1), 63–71 (2004)
5. Soille, P.: Morphological Image Analysis: Principles and Applications. Springer, Heidelberg (1999)
6. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics 9(1), 62–66 (1979)
7. Barber, C., Dobkin, D., Huhdanpaa, H.: The quickhull algorithm for convex hulls. ACM Transactions on Mathematical Software 22(4), 469–483 (1996)
8. Guan, Y.: Automatic extraction of lips based on multi-scale wavelet edge detection. Computer Vision, IET 2(1), 23–33 (2008)
9. Martinez, A., Benavente, R.: The ar face database. CVC Technical Report No.24 (June 1998)
10. Solina, F., Peer, P., Batagelj, B., Juvan, S., Kovac, J.: Color-based face detection in the '15 seconds of fame' art installation. In: Proceedings of Conference on Computer Vision / Computer Graphics Collaboration for Model-based Imaging, Rendering, Image Analysis and Graphical Special Effects, Versailles, France, pp. 38–47 (2003)