A Lip Contour Extraction Method Using Localized Active Contour Model with Automatic Parameter Selection

Xin Liu*, Yiu-ming Cheung*, Meng Li* and Hailin Liu†
*Department of Computer Science, Hong Kong Baptist University, Hong Kong SAR, China
Email: (xliu,ymc,mli)@comp.hkbu.edu.hk
†Faculty of Applied Mathematics, Guangdong University of Technology, Guangzhou, China
Email: lhl@scnu.edu.cn

***Abstract*—Lip contour extraction is crucial to the success of a lipreading system. This paper presents a lip contour extraction algorithm using localized active contour model with the automatic selection of proper parameters. The proposed approach utilizes a minimum-bounding ellipse as the initial evolving curve to split the local neighborhoods into the local interior region and the local exterior region, respectively, and then compute the localized energy for evolving and extracting. This method is robust against the uneven illumination, rotation, deformation, and the effects of teeth and tongue. Experiments show its promising result in comparison with the existing methods.**

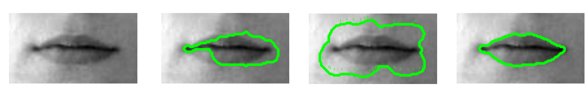
***Keywords*-lip contour extraction; localized active contour model; minimum-bounding ellipse; localized energy.**

I. Introduction

Lip contour extraction has been extensively studied in recent years [1], [2], [3]. It is one of the most important techniques for human-machine interface applications such as lip reading and facial expression analysis. Nevertheless, it is a non-trivial task to find a robust and accurate extraction approach due to large variations caused by different speakers, illumination conditions, appearance of teeth and tongue, low contrast between lip and skin, high deformable level of lips, and so forth.

In the past decade, a number of techniques have been proposed to achieve the lip contour extraction, which can be categorized into two major classes: the edge-based approach and the model-based approach. The edge-based method mainly utilizes the low level spatial cues such as color and edges to achieve lip localization and extraction [2]. Often, the performance of such a method will deteriorate when there is a poor contrast between lip and surrounding skin regions. In contrast, the model-based approach, which builds a lip model with a small set of parameters, generally outperforms the former one. Examples include deformable template (DT) [3] and active shape model (ASM) [4]. The DT algorithm utilizes a parametric model to describe the lip contour, which is, however, sensitive to the deformation and irregularity of the lips. The ASM utilizes a series of landmark points, which are controlled within a few modes derived from a training data, to describe the lip shape. In general, the training process of the ASM is quite time-consuming. Besides the DT and ASM methods, the active contour model (ACM) or snakes [1] is another typical example along this line. The conventional ACM allows an initial contour to deform by minimizing a specific global energy function to produce the desired segmentation. Paper [1] has shown the success of this method in its application domain, but this method is somewhat sensitive to the parameter initialization, uneven illumination and teeth effect.

Further, when objects have heterogeneous statistics, it is found that the localized region-based active contour model (LACM) [5] can generally achieve a satisfactory segmentation result while the conventional ACM fails. In the LACM, the evolving curve splits the local neighborhoods into the local interior region and local exterior region, respectively. Subsequently, the localized energy for evolving and extracting can be computed. However, improper parameters such as large radius or far away evolving curve in LACM can lead to the wrong extracting results as shown in Fig. 1.



(a) (b) (c) (d)

Figure 1. (a) Lip image with some uneven illumination. (b) The extracted result by conventional ACM. (c) The extracted result by the LACM with the improper parameters. (d) The extracted result by the LACM with the proper parameters.

In this paper, we propose an automatic lip contour extraction method based on the localized active contour model, in which, the proper parameters can be automatically selected. We find a minimum-bounding ellipse as the initial evolving curve that can be fitted well in the LACM, meanwhile the local radius in LACM can be automatically selected. Experiments have shown the promising results of the proposed algorithm in comparison with the existing methods.

II. Overview of LACM

This section will overview the primary framework of LACM [5], which assumes that the foreground and background regions would be locally different.

1051-4651/10 $26.00 © 2010 IEEE
DOI 10.1109/ICPR.2010.1053

The statistical analysis of local regions leads to the construction of a group of local energies about each point on the evolving curve. In order to optimize these local energies in its own local region, each point is considered individually. Consequently, the point's component of the local energy is computed by splitting the local neighborhoods into the local interior region and the local-exterior region, respectively, using the evolving curve.

In this paper, I denotes a pre-specified image defined on the domain Ω, C denotes a closed contour represented as the zero level set of a signed distance function ϕ, whose value can be given as: $C = \{u \,|\phi(u) = 0\}$ [6]. The interior of C is specified by the following approximation of the smoothed Heaviside function:

$$\mathcal{H}\phi(u) = \begin{cases} 1, & \phi(u) < -\varepsilon \\ 0, & \phi(u) > \varepsilon \\ \frac{1}{2}\left\{1 + \frac{\phi}{\varepsilon} + \frac{1}{\pi}\sin\left(\frac{\pi\phi(u)}{\varepsilon}\right)\right\}, & otherwise. \end{cases} \tag{1}$$

Similarly, the exterior C can be defined as $(1 - \mathcal{H}\phi(u))$.

The derivative of $\mathcal{H}\phi(u)$, denoted as $\delta\phi(u)$ with

$$\delta\phi(u) = \begin{cases} 1, & \phi(u) = 0 \\ 0, & |\phi(u)| < \varepsilon \\ \frac{1}{2\varepsilon}\left\{1 + \cos\left(\frac{\pi\phi(u)}{\varepsilon}\right)\right\}, & otherwise, \end{cases} \tag{2}$$

is a smoothed version of the Dirac delta used to specify the area adjacent to the curve. Parameters u and v are expressed as independent spatial variables to represent a single point in Ω, respectively. Using this notation, the characteristic function $\mathcal{B}(u,v)$ marking the local regions in terms of a radius parameter r can be described as follows:

$$\mathcal{B}(u,v) = \begin{cases} 1, & \|u - v\| < r \\ 0, & otherwise. \end{cases} \tag{3}$$

Therefore, a localized region-based energy formed from the global energy by substituting local means for global ones is shown below [7]:

$$F = -(\mu_{in}(u) - \mu_{out}(u))^2, \tag{4}$$

$$\mu_{in}(u) = \frac{\int_{\Omega_v} \mathcal{B}(u,v)\cdot\mathcal{H}\phi(v)\cdot I(v)dv}{\int_{\Omega_v} \mathcal{B}(u,v)\cdot\mathcal{H}\phi(v)dv}, \tag{5}$$

$$\mu_{out}(u) = \frac{\int_{\Omega_v} \mathcal{B}(u,v)\cdot(1-\mathcal{H}\phi(v))\cdot I(v)dv}{\int_{\Omega_v} \mathcal{B}(u,v)\cdot(1-\mathcal{H}\phi(v))dv}, \tag{6}$$

where the localized versions of the means $\mu_{in}(u)$ and $\mu_{out}(u)$ represent the intensity mean in local interior and exterior regions around a point u, respectively.

By ignoring the image complexity that may arise outside the local region, we only consider the contributions from the points within the radius r. Furthermore, to keep the curve smooth, a regularization term is added. Also, the arc length of the curve is penalized and weighted by a parameter λ, and the final energy $\mathrm{E}(\phi)$ is given as follows:

$$\mathrm{E}(\phi) = \int_{\Omega_u} \delta\phi(u) \int_{\Omega_v} \mathcal{B}(u,v)\cdot F(I(v),\phi(v))dvdu + \lambda \int_{\Omega_u} \delta\phi(u)\,\|\nabla(u)\|du. \tag{7}$$

By taking the first variation of this energy with respect to ϕ, the following evolution equation is obtained:

$$\frac{\partial\phi}{\partial t}(u) = \delta\phi(u)\int_{\Omega_v} \mathcal{B}(u,v)\cdot\nabla_{\phi(v)}F(I(v),\phi(v))dv + \lambda\delta\phi(u)div\left(\frac{\nabla\phi(u)}{|\nabla\phi(u)|}\right)\|\nabla\phi(u)\|. \tag{8}$$

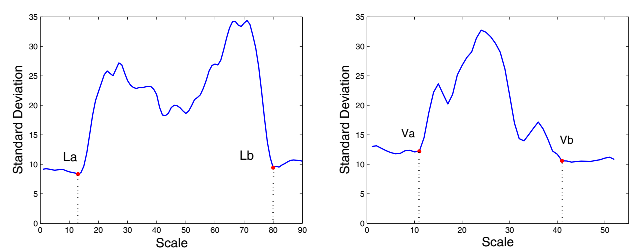
Please note that this ensures that any region-based segmentation energy can be put into this framework.

III. The Proposed Algorithm

Our proposed automatic lip contour extraction system includes an initialization phase and a lip contour extraction phase.

A. *Initialization*

Empirical studies have found that a lip shape is usually close to an elliptical region [8] [3]. Therefore, the lips can be approximatively surrounded by the various elliptical contours according to its special structure. In our method, how to find a minimum-bounding ellipse as the initial evolving curve is of crucial importance to extract the lip contours.



(a) (b)

Figure 2. Standard deviation of columns and rows.

To find the minimum-bounding ellipse of lip region, the detection of lip corner dots is needed. Specifically, $I(x, y)$ represents a pixel value at coordinate (x, y), m and n are the maximum values of rows and columns. The left corner, right corner, upper corner, lower corner are denoted as La, Lb, Va, and Vb, respectively.

We project the RGB-based lip image into the gray-level one. From the practical viewpoint, it is common that there exists the noises or uneven illumination effects. Hence, each lip image is performed with a 3×3 mean filter and a contrast stretching adjustment. According to the empirical and statistical methods, the real lip region is usually different from the surrounding regions. Subsequently, the horizontal or vertical lip corner dots, e.g. Fig. 3(b), can be detected by computing the first and last points with the standard deviation, which is changing obviously compared with the adjacent ones. For example, in order to find the horizontal lip corner columns quickly, we can use the following equations:

$$mean_j = \frac{1}{m}\sum_{i=1}^{m}(I(i,j)), j = 1, 2, \cdots n, \tag{9}$$

$$Std_j = \left(\frac{1}{m}\sum_{i=1}^{m}(I(i,j) - mean_j)^2\right)^{\frac{1}{2}}, \tag{10}$$

$$|Std_j - Std_{j+2}| < \Delta S, |Std_j - Std_{j-2}| < \Delta S, \tag{11}$$

where ΔS is a changing threshold. Fig. 2 gives an example. We can easily obtain the coordinate value of La_x, Lb_x. Subsequently, the value of La_y and Lb_y can be computed through the mean coordinate with the minimum gray value of $I(La_x, j), i = 1, \cdots, n$. The other approximative corner dots can also be computed through the above method. Actually, the lip corner dots need not fit the exact position of lips, which is enough within the estimated neighbourhood around geometric positions. Let the (x_c, y_c) be the origin center of the minimum-bounding ellipse, through which the mathematical equations are as follows:

$$x_c = \frac{1}{2}(La_x + Lb_x), y_c = \frac{1}{2}(La_y + Lb_y),$$

$$\theta = \arctan\left(\frac{Lb_y - La_y}{Lb_x - La_x}\right),$$

$$r_a = \frac{1}{2}((Lb_x - La_x)^2 + (Lb_y - La_y)^2)^{\frac{1}{2}},$$

$$r_b = \frac{1}{2}((Vb_x - Va_x)^2 + (Vb_y - Va_y)^2)^{\frac{1}{2}},$$

$$X = (x - x_c)\cdot\cos\theta + (y - y_c)\cdot\sin\theta,$$

$$Y = (y - y_c)\cdot\cos\theta - (x - x_c)\cdot\sin\theta,$$

$$\frac{X^2}{r_a^2} + \frac{Y^2}{r_b^2} = 1, \tag{12}$$

where the inclined angle θ is defined to be positive in the counter-clockwise direction.

B. Contour Extraction

After successfully finding the minimum-bounding ellipse of lip region, we let this ellipse be the evolving curve representing the zero level set C as stated in Section II, which can be fitted well in the LACM. Subsequently, local neighborhoods of the points can be split into the local interior region and local exterior region by the evolving curve.

By computing the local energies at each point along the curve, the evolving curve will deform by minimizing the local energies so that the desired contour extraction is achieved. The steps of lip contour extraction are as follows:

- Locate the lip region, preprocess;
- Obtain the minimum-bounding ellipse;
- Evolve with iteration;
- Extract the lip contours.

Fig. 3 gives an example. Specifically, the local radius r selected by the function $\mathcal{B}(u,v)$ is an important parameter in the LACM. By the rule of thumb, $r = \frac{r_b}{2}$ is an appropriate value in the most cases of lip contour extraction. However, when there exists the lip image with the mouth open widely, the value of local radius selected should be smaller.

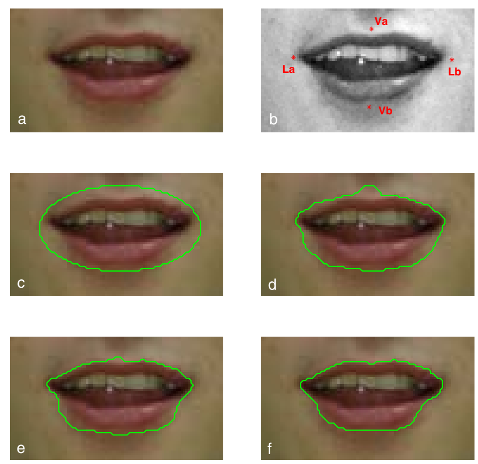


Figure 3. The procedure of lip contour extraction: (a) lip image of size 52×90, (b) lip corner dots, (c) minimum-bounding ellipse, (d)-(f) the extracted results after 20, 40, and 60 iterations, respectively.

IV. Experimental Result

We have applied our approach to the 500 frontal face images with the different mouth shapes obtained in a uniform illuminance environment. The database consists of 200 face images from the CVL face database [9], 100 face images from the GTAV face database [10] and 200 lip images from our laboratorial database. In our experiments, we set the parameter λ at 0.3.

Figure 4. The extracted results of lip images from our laboratorial database by the proposed approach.

Figure 5. The extracted results of lip images from the CVL database by our proposed approach.

Fig. 4 and Fig. 5 show the experimental samples of lip contour extraction by the proposed approach. Table I presents the comparison result between the proposed approach and the other two existing methods (Conventional ACM and DT), in which the performance of lip contour extraction was measured by $\frac{n_c}{n}$, where n is the total number of the test images, and n_c is the number of correctly extracted ones.

Table I
THE COMPARISON OF THE PROPOSED APPROACH WITH THE EXISTING CONVENTIONAL ACM AND DT METHODS.

Method	Performance of the test database			
	CVL	GYAV	Our lab	Average
Conventional ACM	79.5%	81%	76.5%	78.6%
DT method	92.5%	94%	92%	92.6%
Our approach	97%	95%	96.5%	96.4%

It can be seen that the average performance of the conventional ACM method was just 78.6%, which often failed in the appearance of the teeth and tongue. The average performance of DT algorithm reached 92.6%, which was much better than the former. Nevertheless, its performance became deteriorate when the lips are irregular. In contrast, the average performance of the proposed algorithm reached 96.4%. It has extracted the lip contours quite well, even for the deformable or irregular lips. Actually, the proposed method is more tolerant to the uneven illumination, rotation, deformation, and the effects of teeth and tongue. Further, we have also examined those unsatisfactory results (i.e. the remaining cases of 3.6%), and found that the images in these cases all have either the very poor contrast between the lip and the surrounding skin regions, or the obvious beard effects around the lips.

V. Conclusion

This paper has proposed a robust approach with automatic selection of the proper parameters in the LACM for lip contour extraction. We obtain a minimum-bounding ellipse as the initial evolving curve, through which an optimal extraction of the lip image into lip and non-lip regions has been achieved. This algorithm is robust against the uneven illumination, rotation, deformation, and the effects of teeth and tongue. Experiments have shown the promising results.

Acknowledgment

The work described in this paper was supported by the Research Grant Council of Hong Kong SAR with Project No: HKBU 210309, a Faculty Research Grant of Hong Kong Baptist University with the Project No: FRG/07-08/II-88, and the National Science Foundation of China (60974077).

References

[1] G. Chiou and J. N. Hwang, "Lipreading from color video," *IEEE Transactions on Image Processing*, vol. 6, no. 8, pp. 1192–1195, 1997.

[2] X. Zhang, R. M. Mersereau, M. Clements, and C. C. Broun, "Visual speech feature extraction for improved speech recognition," in *Proc. International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, 2002, pp. 1993–1996.

[3] A. W. C. Liew, S. H. Leung, and W. H. Lau, "Lip contour extraction from color images using a deformable model," *Pattern Recognition*, vol. 35, no. 12, pp. 2949–2962, 2002.

[4] J. Luettin, N. Thacker, and S. Beet, "Visual speech recognition using active shape models and hidden markov models," in *Proc. International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, 1996, pp. 817–820.

[5] S. Lankton and A. Tannenbaum, "Localizing region-based active contours," *IEEE Transactions on Image Processing*, vol. 17, no. 11, pp. 2029–2039, 2008.

[6] R. Tsai and S. Osher, "Level set methods and their applications in image science," *Communications in Mathematics and Science*, vol. 1, no. 4, pp. 623–656, 2003.

[7] A. Yezzi, A. Tsai, and A. Willsky, "A fully global approach to image segmentation via coupled curve evolution equations," *Journal of Visual Communication and Image Representation*, vol. 13, pp. 195–216, 2002.

[8] M. Li and Y. M. Cheung, "Automatic lip localization under face illumination with shadow consideration," *Signal Processing*, vol. 89, no. 12, pp. 2425–2434, 2009.

[9] F. Solina, P. Peer, B. Batagelj, S. Juvan, and J. Kovac, "Color-based face detection in the "15 seconds of fame" art installation," in *Proc. Mirage, Conf. Computer Vision/Computer Graphics Collaboration for Model-Based Imaging, Rendering, Image Analysis and Graphical Special Effects,Rocquencourt, France, Mar.10-11*, 2003, pp. 38–47.

[10] "http://gps-tsc.upc.es/gtav/researchareas/upcfacedatabase/gtavfacedatabase.htm."