

RESEARCH

Open Access



Lip segmentation using automatic selected initial contours based on localized active contour model

Yuanyao Lu* and Qingqing Liu

Abstract

With the rapid development of artificial intelligence and the increasing popularity of smart devices, human-computer interaction technology has become a multimedia and multimode technology from being computer-focused to people-centered. Among all ways of human-computer interactions, using language to interact with machines is the most convenient and efficient one. However, the performance of audio speech recognition systems is not satisfied in a noisy environment. Thus, more and more researchers focus their works on visual lip reading technology. By extracting lip movement features of speakers rather than audio features, visual lip reading systems can get superior results when noises and interferences exist. Lip segmentation plays an important role in a visual lip reading system, since the segmentation result is crucial to the final recognition accuracy. In this paper, we propose a localized active contour model-based method using two initial contours in a combined color space. We apply illumination equalization to original RGB images to decrease the interference of uneven illumination. A combined color space consists of the U component in CIE-LUV color space and the sum of C2 and C3 components of the image after discrete Hartley transform. We select a rhombus as the initial contour of a closed mouth, because it has a similar shape to a closed lip. For an open mouth, we utilize a combined semi-ellipse as the initial contours of both outer and inner lip boundaries. After attaining the results of each color component separately, we merge them together to obtain the final segmentation result. From the experiment, we can conclude that this method can get better segmentation results compared with the method using a circle as the initial contour to segment gray images and images in combined color space, especially for open mouth. An extremely obvious advantage of this method is the results of open mouth excluding internal information of mouth such as teeth, black holes, and tongue, because of the introduction of the inner initial contour.

Keywords: Lip segmentation, Localized active contour model, Initial contour, Illumination equalization, Color space

1 Introduction

Visual lip reading is a technology which combines machine vision and language perception. Visual lip reading systems identify face region from images or videos by machine vision, then extract the mouth variation features of speakers and determine the pronunciations of these features by recognition model, thereby recognizing the speech contents. This system receives more and more attention in the field of human-computer interaction (HCI), pattern recognition (PR), and artificial intelligence (AI) in recent years. It resolves the problem

that the recognition rate of audio speech recognition system declines dramatically because of interferences or noises. Lip segmentation is fundamental to visual lip reading systems, because the accuracy of segmentation result directly affects the recognition rate.

There are various methods of lip segmentation, such as a method based on MAP-MRF framework [1], a clustering approach without knowing segment number [2], and the active contour model (ACM) involved in this paper. ACM is one of the popular models for image segmentation. There are several advantages of ACM over classical image segmentation methods [3]. First, ACM can obtain sub-pixel accuracy of object boundaries [3, 4]. Second, this model can be developed within the framework of the energy minimization principle. Third, they can get smooth

* Correspondence: luyy2018@163.com; luyy@ncut.edu.cn
School of Electronic and Information Engineering, North China University of Technology, No. 5, Jinyuanzhuang Road, Shijingshan District, Beijing 100144, China

and closed results and can be used for further applications, such as feature extraction and shape analysis [3].

Existing ACM can be classified into two categories: edge-based models and region-based models. Edge-based models adopt image gradient as constraint condition which impels the initial contour to converge to object boundary [5]. Researchers have proposed many methods to improve edge-based models [4, 6–8]. These models have incomplete convergence problem due to fuzzy or weak object boundary and image noise. Region-based models are immune from image noise. They utilize image statistical information as constraint condition, and the performance is superior to edge-based models. These models can segment object area in the case of weak boundary or even without boundary [5]. Papers [9–14] proposed many methods to improve region-based models. In the Mumford-Shah method [9], an image is segmented by minimizing the energy function including boundary and image area [15]. Then, Chan and Vese [9] used Shah’s variational method to build an active contour model, which was developed by Osher’s level set method [16]. Chan-Vese algorithm has a high performance and is easy to implement, so it has been widely used in image segmentation [15]. Global region-based models may not obtain satisfactory results when the texture is heterogeneous or different between object and background. Thus, some researchers attempt to use local regional information as constraint of active contour [5]. Localized active contour model (LACM) is one of widely used methods, proposed by Lankton and Tannenbaum [6]. This model constructs a local region at each point along contours. These regions lead to a set of local energies centered on points.

Li and Cheung and Chin et al. [17–19] proposed several lip segmentation methods based on grayscale images, while Talea and Yaghmaie [20], Kim et al. [21], Hulbert and Poggio [22], Canzlerm and Dziurzyk [23], and Leung et al. [24] adopted methods using color images directly. The pseudo tone is proposed as the ratio of RGB values for lip detection in [22]. Canzlerm and Dziurzyk [23] suggested that the segmentation quality can be improved by suppressing the blue color component, since the blue component is subordinate in the lip region. Leung et al. [24] applied fuzzy color clustering. There are other color spaces that are used to display lip color and background color differences, such as YCbCr [25], NTSC [26], CIE-Lab [27–30], and bi-color space [31].

In this paper, we propose a LACM-based method using two initial contours in a combined color space. At first, illumination equalization is used to the original RGB image to reduce the interference resulting from uneven illumination. Then, we use a combined color space which consists of the U component in CIE-LUV color

space and the sum of C2 and C3 in discrete Hartley transform (DHT). This color space can retain more details of lip information and highlight differences between the lip and other parts such as skin, teeth, and black holes. Ultimately, we utilize a rhombus as the initial contour for closed mouth and combined semi-ellipses as the outer and inner initial contour for open mouth. Thus, we can obtain the segmentation results without internal part of the mouth.

This paper is organized as follows: we review the LACM and Chan-Vese model in Section 2; the proposed method is described in Section 3; experimental results are shown in Section 4; and lastly, the conclusion is provided in Section 5.

2 Overview of LACM and Chan-Vese model

The core of classical ACM is evolving the curves in constrained images so that the target feature is detected in the image by minimizing their energy and length [15]. LACM was proposed by Lankton in 2008 [6]. This model constructs a local region at each point along the contour. These regions lead to a set of local energies centered on these points. In the process of local energy minimization, the curve gradually converges to the object boundary. We use I to represent a pre-specified image defined on the domain Ω , and C represents a closed contour represented as the zero-level set of a signed distance function ϕ , i.e., $C = \{x | \phi(x) = 0\}$ [6]. The interior of C is specified by the approximation of the smoothed Heaviside function:

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

Similarly, the exterior of C is defined as $(1 - H\phi(x))$. The smoothed version Dirac delta function $\delta\phi(x)$ which is the derivative of $H\phi(x)$ is used to specify the area around the curve C :

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

In order to calculate the local energy, LACM introduces a function $\beta(x,y)$ to define the local region in terms of parameter r :

$$\beta(x,y) = \begin{cases} 1, & \|x-y\| < r \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

If the point y is in the local region centered on x , the value of function will be 1; otherwise, it will be 0. The local region is split into interior area and exterior area by the evolving curve, as shown in Fig. 1.

The energy function can be defined as:

$$E(\phi) = \int_{\Omega} \delta(\phi) \int_{\Omega} \beta \cdot F(I, \phi) dy dx + \lambda \int_{\Omega} \delta(\phi) \|\nabla \phi\| dy dx \quad (4)$$

where I is an image and Ω is its domain, λ denotes a parameter of smoothing item, and F is a local force function.

The mean value of local interior region u and local exterior region v are defined as:

$$u = \frac{\int_{\Omega} \beta \cdot I \cdot H(\phi) dy dx}{\int_{\Omega} \beta \cdot H(\phi) dy dx} \quad (5)$$

$$v = \frac{\int_{\Omega} \beta \cdot I \cdot (1-H(\phi)) dy dx}{\int_{\Omega} \beta \cdot (1-H(\phi)) dy dx} \quad (6)$$

The formula of F changes with the change of energy model. In this paper, we choose the Chan-Vese model as the energy model. Thus, F is as following:

$$F = H(\phi)(I-u)^2 - (1-H(\phi))(I-v)^2 \quad (7)$$

and its derivative is denoted as:

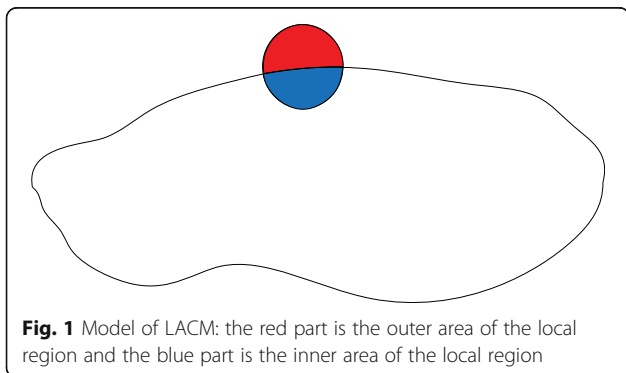


Fig. 1 Model of LACM: the red part is the outer area of the local region and the blue part is the inner area of the local region

$$\nabla_{\phi} F = \delta(\phi) ((I-u)^2 - (I-v)^2) \quad (8)$$

Substituting Eq. (8) into Eq. (4), we can get the local curvature flow:

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \int_{\Omega} \beta \cdot \delta(\phi) ((I-u)^2 - (I-v)^2) dy dx + \lambda \delta(\phi) \operatorname{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) \quad (9)$$

Local force F changes with the variation of mean value u and v based on Eq. (7). From Eq. (9), it can be found that F makes the curvature flow converge to the minimization. The curve will stop moving when the curvature flow approximates to 0 after several times of iterations and only the object boundary satisfies the condition. That means the curve can converge to the object boundary when the local energies do not change any more after iterations.

3 Method

3.1 Mouth localization

The focus of this paper is lip segmentation, so we just take the lip area as region of interest. In order to reduce redundant information, it is necessary to extract the mouth area. In previous studies, researchers have proposed a variety of methods to extract lip area [32, 33]. However, these methods still contain redundant information or have an adverse impact on the subsequent processing. In this paper, we choose a method which can segment the mouth area according to the general structure and proportion of the face [34]. The formula is as follows:

$$\frac{1}{4} W_{\text{face}} < W_{\text{mouth}} < \frac{3}{4} W_{\text{face}} \quad (10)$$

$$\frac{2}{3} H_{\text{face}} < H_{\text{mouth}} < \frac{9}{10} H_{\text{face}} \quad (11)$$

where W_{face} and H_{face} are the width and height of the face and W_{mouth} and H_{mouth} are the width and height of the mouth. From the formula, we can obtain the lip area as shown in Fig. 2. It explains that this method can get satisfactory and effective results in the simple background which is required in this paper.

3.2 Illumination equalization of RGB image

Illumination is a main factor that affects the appearance of an image [35]. The lighting from different directions may cause the uneven illuminations, which often lead to intensity heterogeneity. Therefore, illumination equalization plays a significant role in image analysis and processing [36]. Liew et al. [35] has introduced an effective way to reduce the impacts of vertical direction uneven illumination. This



Fig. 2 Obtained lip areas

method just processes the single point on the boundary and its effect may be influenced by image noise. In this paper, we adopt an improved illumination equalization method proposed in [35] via the analysis of local region along the image boundary, which is more robust against noise and adaptive to the multifarious illumination directions. This method consists of two illumination directions, the horizontal direction and the vertical direction. The uneven illumination can be regarded as the liner along its direction.

We assume the size of a given image is $m \times n$, $L(x, y)$ denotes the luminance values before illumination equalization and $L_h(x, y)$ and $L_v(x, y)$ denote the luminance values after illumination equalization in the horizontal direction and vertical direction, respectively. The size of the local region is $(2p + 1) \times (2q + 1)$; we take the average intensity value of this local area instead of the intensity of a single boundary point. In this paper, we set the size of the local region as 5×7 . The formulas of illumination equalization in the horizontal and vertical direction are as follows:

$$L_h(x, y) = \begin{cases} L(x, y) + \frac{(n-2j+1) \cdot (r(p)-l(p))}{2(n-1)}, & i \in [1, p], \\ L(x, y) + \frac{(n-2j+1) \cdot (r(m-p)-l(m-p))}{2(n-1)}, & i \in (m-p, m], \\ L(x, y) + \frac{(n-2j+1) \cdot (r(i)-l(i))}{2(n-1)}, & i \in [p, m-p]; \end{cases} \quad (12)$$

$$L_v(x, y) = \begin{cases} L(x, y) + \frac{(m-2i+1) \cdot (b(q)-t(q))}{2(m-1)}, & j \in [1, q), \\ L(x, y) + \frac{(m-2i+1) \cdot (b(n-q)-t(n-q))}{2(m-1)}, & j \in (n-q, n], \\ L(x, y) + \frac{(m-2i+1) \cdot (b(j)-t(j))}{2(m-1)}, & j \in [q, n-q]. \end{cases} \quad (13)$$

where $l(i)$ and $r(i)$ denote the mean intensity of left and right borders within a local region of size $(2p + 1) \times (2q + 1)$ at the i th row, respectively. Similarly, $t(j)$ and $b(j)$ denote the mean intensity of top and bottom borders at the j th column individually. The formulas are as follows:

$$l(i) = \frac{1}{(2p+1) \cdot (2q+1)} \sum_{k=-p}^p \sum_{l=1}^{2q+1} L(i+k, l) \quad (14)$$

$$r(i) = \frac{1}{(2p+1) \cdot (2q+1)} \sum_{k=-p}^p \sum_{l=n-2q}^n L(i+k, l) \quad (15)$$

$$t(j) = \frac{1}{(2p+1) \cdot (2q+1)} \sum_{k=1}^{2p+1} \sum_{l=-q}^q L(k, j+l) \quad (16)$$

$$b(j) = \frac{1}{(2p+1) \cdot (2q+1)} \sum_{k=m-2p}^m \sum_{l=-q}^q L(k, j+l) \quad (17)$$

In our works, we apply the illumination equalization separately to the R , G , and B components of the original

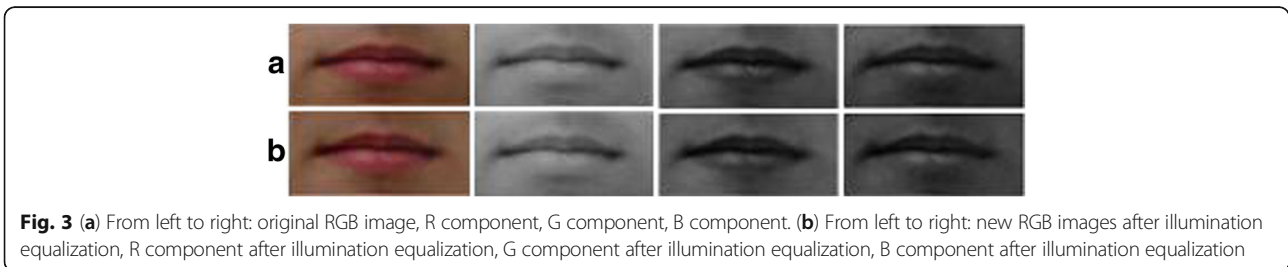
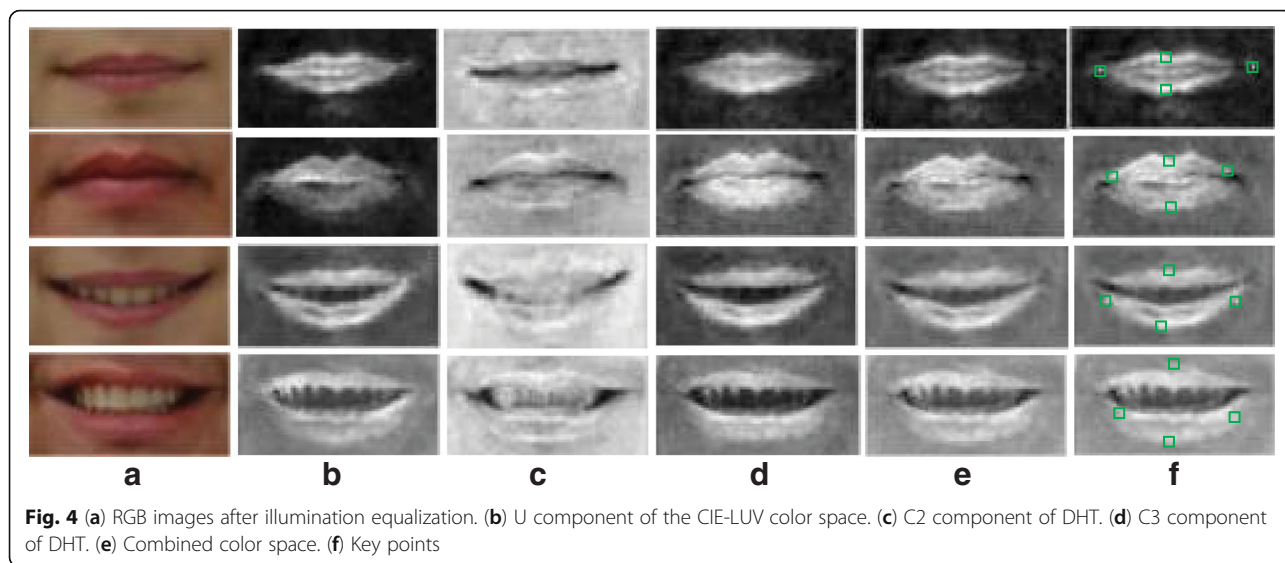


Fig. 3 (a) From left to right: original RGB image, R component, G component, B component. (b) From left to right: new RGB images after illumination equalization, R component after illumination equalization, G component after illumination equalization, B component after illumination equalization



RGB image. By merging the components after illumination equalization, we can get a new RGB image that has uniform illumination. The method is utilized in horizontal direction firstly and then is used in the vertical direction based on the obtained results. Figure 3 shows an example of lip image and its components before and after illumination equalization. The intensity value of each image is changed even though we cannot detect it via our eyes. By this way, we obtain an image which can reduce the interference of uneven illumination and it is favorable for the subsequent segmentation.

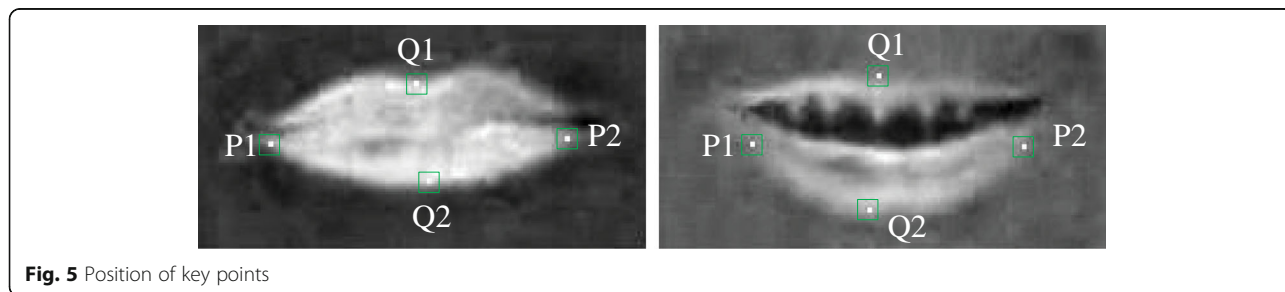
3.3 Color space and key points

3.3.1 Color space

In previous researches, gray images are used to segment lip area most frequently. Some researchers utilize the three components of RGB image respectively and then add the results together. Others may use some new color space like bi-color space [31] and method based on the HIS and RGB color model [37]. In our study, we choose a combined color space consisting of CIE-LUV color space and DHT.

3.3.1.1 CIE-LUV color space CIE-LUV color space is obtained from CIE-XYZ space. In this space, L denotes the luminance, which is separated from the other components. U and V represent chrominance components. CIE-LUV color space has better robustness when the luminance of image has some changes compared with other color space. The experiment shows that the U component has a high performance of brightness difference between the lip region and the background region. This difference makes the lip segmentation from the background more convenient. Therefore, we deem that the color characteristics of the lip largely rely on the U component. But the U component has some drawbacks, for instance, the lip edge is a bit fuzzy. In order to compensate for this shortcoming, we adopt the DHT.

3.3.1.2 Discrete Hartley transform DHT has higher computational efficiency, which is one of its advantages. When the input signal is a real number, its transform formula only contains real number. DHT also has better symmetry. DHT is used in the image processing field increasingly because of these advantages. In this paper, we choose DHT to compensate for the deficiency of the U component of CIE-LUV.



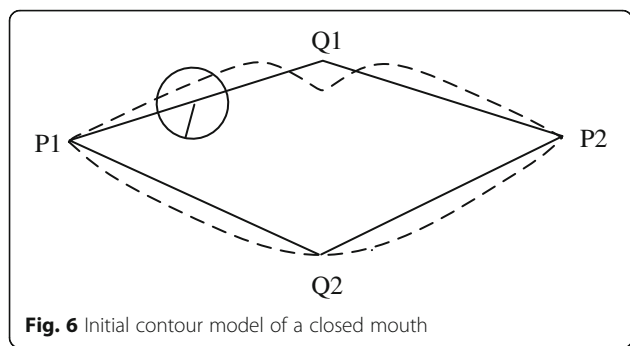


Fig. 6 Initial contour model of a closed mouth

From the experiment, the sum of the C2 and C3 components shows the difference distinctly between the lip region and the background region. Thus, we choose the two components to retain more lip details.

We combine the *U* component in CIE-LUV color space with the sum of the C2 and C3 components of the DHT to extract the color feature. This combined color space can highlight the object region from background region more obvious. It retains more useful lip edge information and provides a more favorable condition for the subsequent lip segmentation study. We can see the components and the final results in Fig. 4.

3.3.2 Key point localization

Determining key points is the critical stage of lip segmentation. In this paper, we define key points in the combined color space. Four points are considered as key points. They are two lip corners, one point on the upper lip and one on the lower lip. We coordinate the horizontal ordinates of the other two points by the median of the horizontal ordinates of the corners.

Two lip corners are detected firstly and then the other two points are defined based on the location of corners. Algorithms of determining lip corners are

quite mature. Researchers have proposed various algorithms of lip corner determination. Rao and Mersereau used the gray level information of lip area pixels to define the location of corners, while some researchers utilized the method of lip edge detection to find corners [34, 38, 39]. In order to reduce the inaccuracy, we take a small area as the lip corner search area instead of a single point. Here, we briefly explain this method through the example of the corner of the mouth. First, we find the row with the maximum pixel value as the baseline. Then, columns of each point which have maximum pixel value in the upper and lower 10 lines of the baseline are detected. Finally, we take the pixels with minimum column and maximum column as left and right corner, respectively. For the other two points, we get their coordinates by the same method. Figure 5 shows the position of these points on the lip.

3.4 Initial contour

After transforming color space and finding key points, the initial lip contour is determined by the coordinates of key points in the new color space. In our research, we divide lip shape into two classes according to the ordinate of Q1 and Q2. If the absolute value of ordinate difference between Q1 and Q2 is greater than half the height of lip image, we deem the mouth is open. On the contrary, we consider the mouth is closed. We will explain each case in the rest of this part.

3.4.1 Closed mouth

We suppose the mouth is closed if the absolute value of ordinate difference between Q1 and Q2 is less than half the height of lip image. In this case, we choose rhombus as the initial lip contour, since rhombus is similar to lip shape. Vertexes of the rhombus are key points we have gained. Figure 6 shows the model.

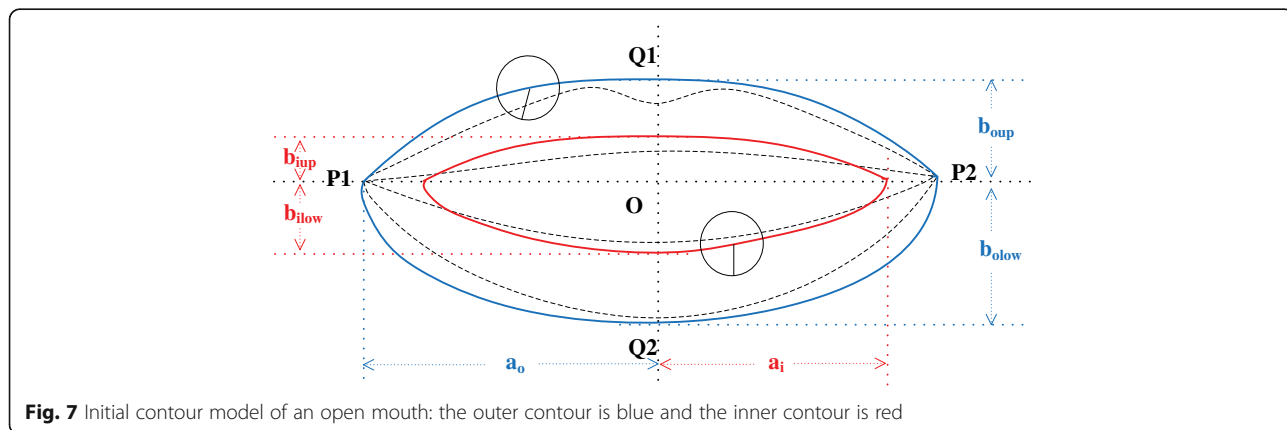


Fig. 7 Initial contour model of an open mouth: the outer contour is blue and the inner contour is red

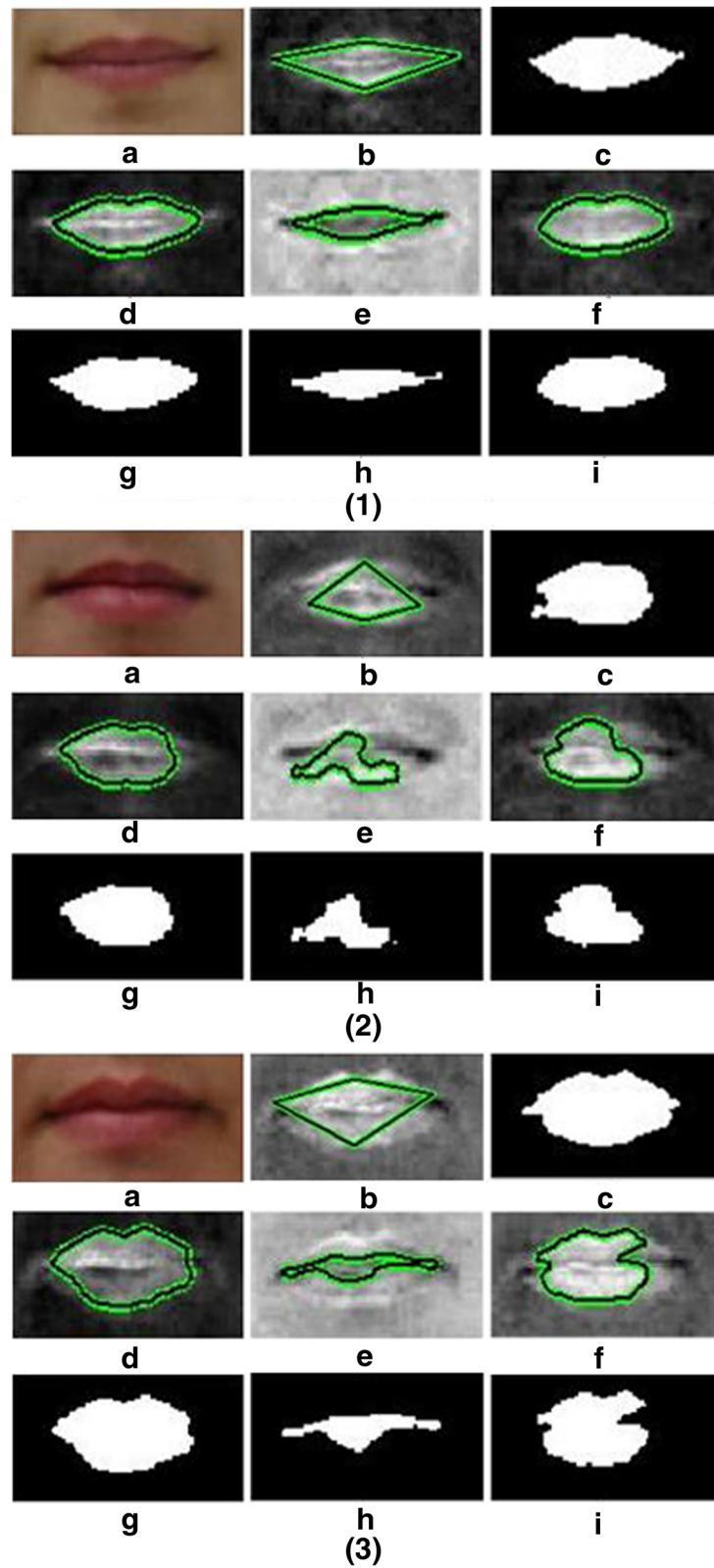
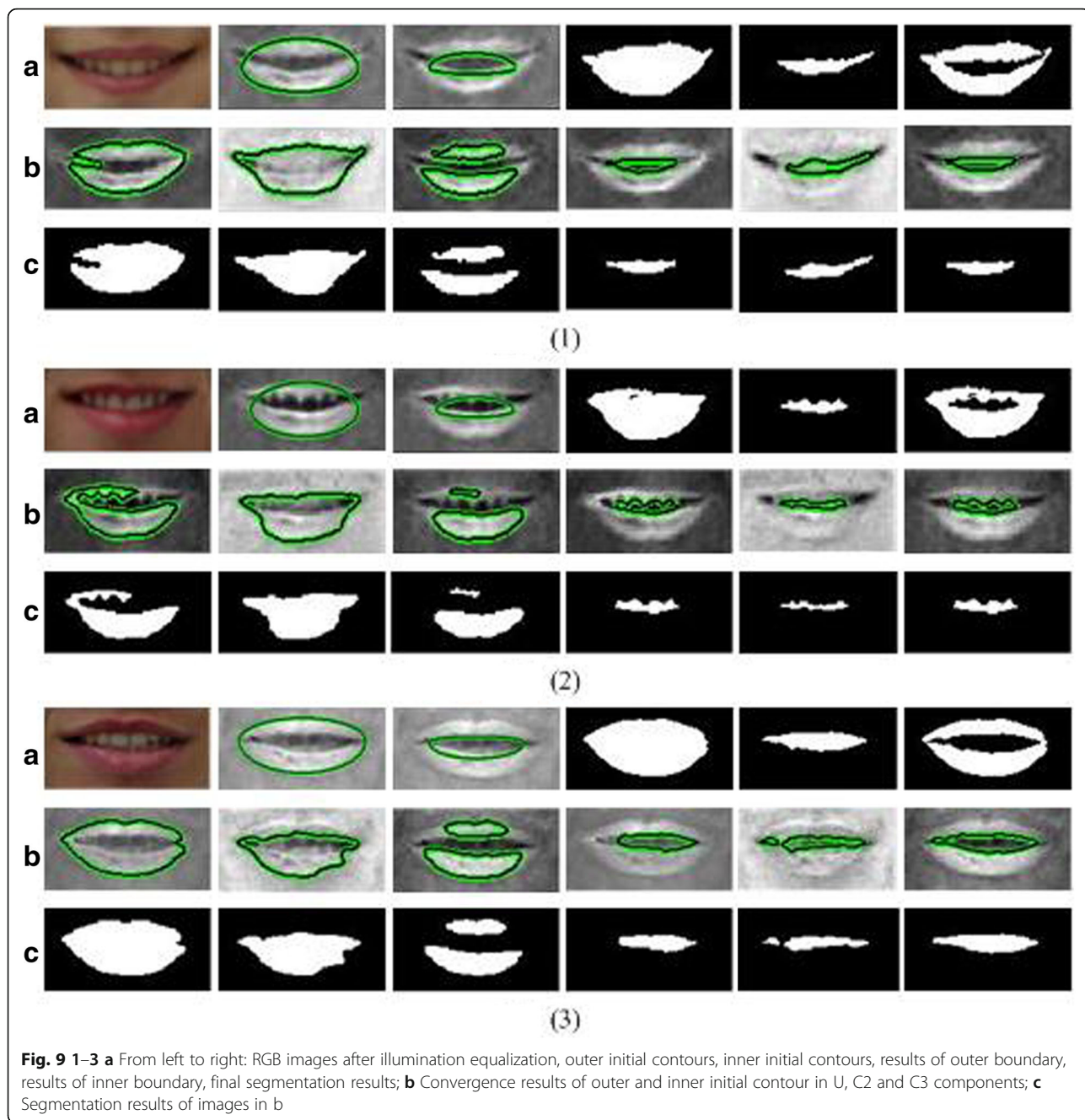


Fig. 8 1–3 **a** RGB image after illumination equalization. **b** Initial contour in combined color space. **c** Final result. **d** Convergence result of U component. **e** Convergence result of C2 component. **f** Convergence result of C3 component. **g** Segmentation result of U component. **h** Segmentation result of C2 component. **i** Segmentation result of C3 component



3.4.2 Open mouth

When the absolute value of ordinate difference between Q1 and Q2 is greater than half the height of lip image, we consider the mouth is open. We find the shape of an open mouth is like an ellipse rather than a rhombus. However, when the mouth opens widely, a single ellipse shape may be inaccurate. Mark et al. proposed that the upper and lower lip shapes are different so that we can use two semi-ellipses as the initial contour [35, 40], as shown in Fig. 7.

We utilize key points to build the combined semi-ellipse. In previous studies, the segmentation result of LACM only has outer contour and the inner contour is unknown. In order to improve and perfect the lip segmentation, we propose the inner initial contour according to the shape of outer contour. The equations are formulated as follows:

$$O_x = \frac{1}{2}(P1_x + P2_x), O_y = \frac{1}{2}(P1_y + P2_y) \quad (18)$$

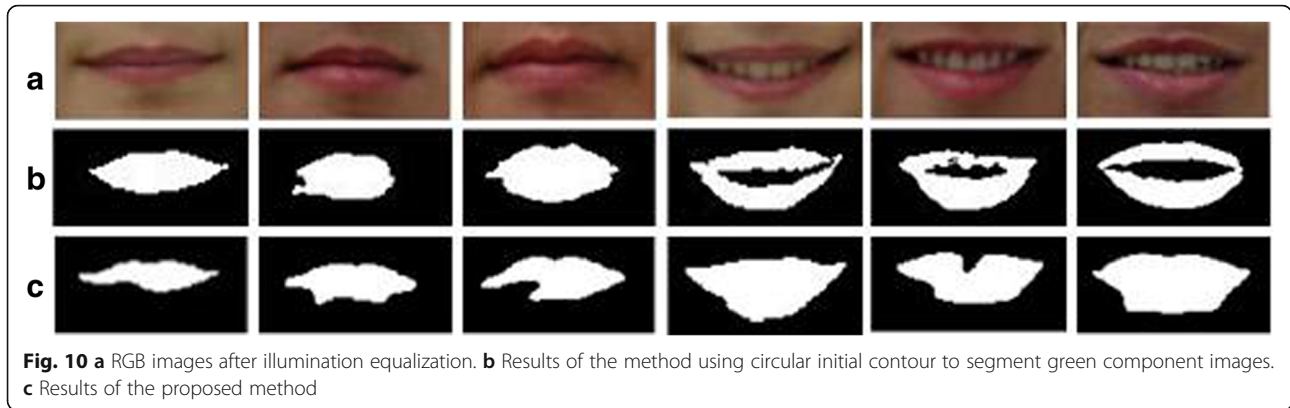


Fig. 10 **a** RGB images after illumination equalization. **b** Results of the method using circular initial contour to segment green component images. **c** Results of the proposed method

$$a_o = \frac{1}{2} \sqrt{(P1_x - P2_x)^2 + (P1_y - P2_y)^2}, \quad a_i = \frac{3}{4} a_o \quad (19)$$

$$b_{oup} = \sqrt{(Q1_x - O_x)^2 + (Q1_y - O_y)^2}, \quad b_{iup} = \frac{1}{2} b_{oup} \quad (20)$$

$$b_{olow} = \sqrt{(Q2_x - O_x)^2 + (Q2_y - O_y)^2}, \quad b_{ilow} = \frac{1}{2} b_{olow} \quad (21)$$

$$\theta = \arctan\left(\frac{P2_y - P1_y}{P2_x - P1_x}\right) \quad (22)$$

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} x - O_x \\ y - O_y \end{bmatrix} \quad (23)$$

$$\frac{X^2_{oup}}{a^2_o} + \frac{Y^2_{oup}}{b^2_{oup}} = 1, \quad \frac{X^2_{olow}}{a^2_o} + \frac{Y^2_{olow}}{b^2_{olow}} = 1 \quad (24)$$

$$\frac{X^2_{iup}}{a^2_i} + \frac{Y^2_{iup}}{b^2_{iup}} = 1, \quad \frac{X^2_{ilow}}{a^2_i} + \frac{Y^2_{ilow}}{b^2_{ilow}} = 1 \quad (25)$$

where O is the original center of the model, O_x and O_y are its abscissa and ordinate, and a_o and a_i are semi-

major axes of outer and inner contour, respectively. b_{oup} and b_{olow} are upper and lower semi-minor axes of outer contour. Similarly, b_{iup} and b_{ilow} are upper and lower semi-minor axes of inner contour. θ is the inclined angle, and it is positive along the counterclockwise direction. The local region radius r of both contours is the same, $r = b_{oup}/4$. The value of r ensures that the local region contains lip area and cannot result in over-convergence. According to the experiment, these parameters are applicable in most lip images.

4 Results and discussion

4.1 Results

The experiment is implemented in MATLAB R2013a, and we have tested 500 face images from AR face database. We apply the initial contours to each component of combined color space and then merge the results of each component to gain the final convergence result.

For an open mouth, the outer contour converges to the outer lip boundary firstly. We use R_o to denote the result, and it contains the internal contents of the mouth. Then, the inner contour converges to the inner lip boundary; the result is labeled as R_i , and it is the supernumerary part of the mouth except the lip. Finally, the ultimate result is $R = R_o - R_i$. It is relatively conspicuous that the teeth and black hole are removed perfectly by the inner contour. Figure 8 shows the initial contours and the results for

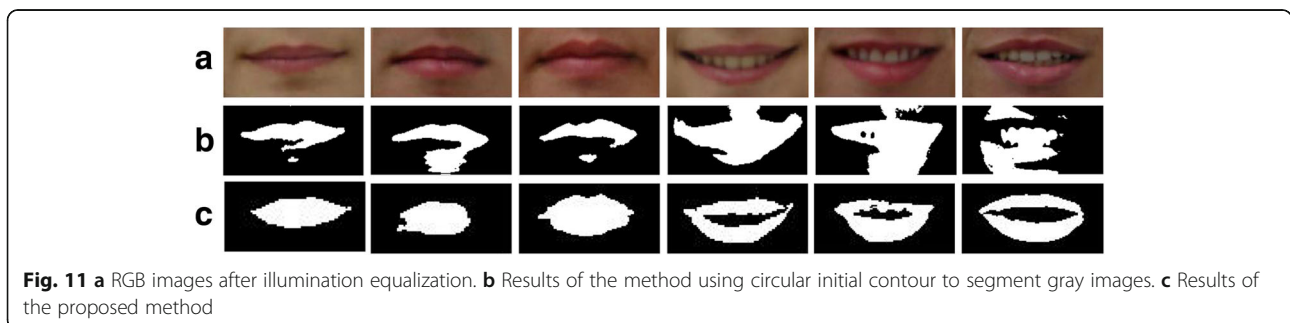
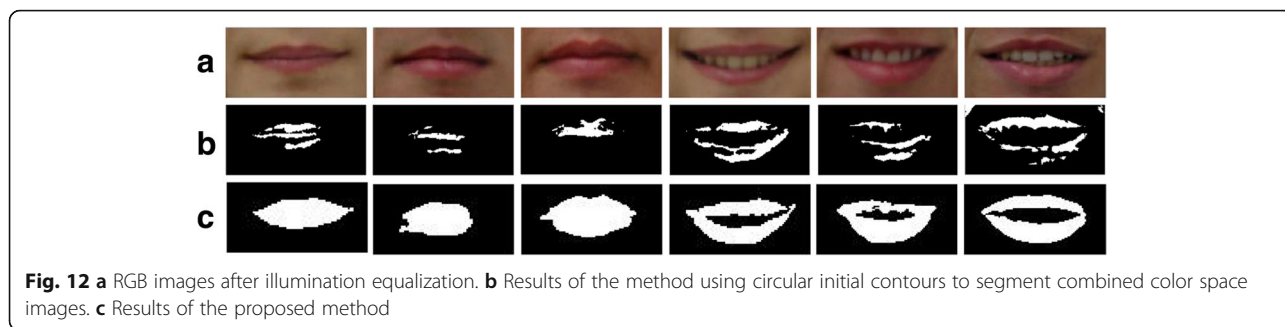


Fig. 11 **a** RGB images after illumination equalization. **b** Results of the method using circular initial contour to segment gray images. **c** Results of the proposed method



closed mouth. Figure 9 shows the initial contours and the results for open mouth.

In order to demonstrate the proposed method, we use LACM with circular initial contour and the proposed initial contours to segment green component of RGB images, gray images, and images in combined color space, respectively. The iteration numbers of all the methods are set to 300, except that inner contour is 100. Figure 10 displays the results of the method using circular initial contour to segment green component images and the results of our method. Figure 11 shows the results of the method using circular initial contour to segment gray images and the results of the proposed method. In Fig. 12, we explain the results of the method using circular initial contours to segment combined color space images and the results of using the proposed method. Figure 13 shows the results of the method using the proposed initial contours to segment gray images and the proposed method.

It can be seen that the results of the proposed method are more accurate, more complete, and smoother than compared methods. Our method contains more details because we utilize a combined color space which consists of the *U* component in CIE-LUV color space and the sum of C2 and C3 color space of DHT. For open mouth, our method can get the segmentation results only involving the lip area instead of containing teeth or holes. On account of utilizing illumination equalization,

the proposed method avoids the problem of partial convergence or over-convergence owing to the non-uniform illumination. Furthermore, we can get appropriate initial contour according to key points which are defined by our method. This avoids the issue that local region excludes the lip area or contains more lip region even involving the information inside of mouth. Hence, it can converge to the real lip boundary precisely using our method.

4.2 Discussion

To evaluate the accuracy of our algorithm, we use two measures [41]. The first measure is defined as

$$OL = \frac{2(A_1 \cap A_2)}{A_1 + A_2} \times 100\%$$

It determines the percentage of overlap (OL) between the segmented lip region A_1 and the ground truth A_2 . The second measure is the segmentation error (SE), defined as

$$SE = \frac{OLE + ILE}{2 \times TL} \times 100\%$$

where OLE is the number of non-lip pixels being classified as lip pixels (outer lip error) and ILE is the number of lip pixels classified as non-lip ones (inner lip error); TL denotes the number of lip pixels in the ground truth.

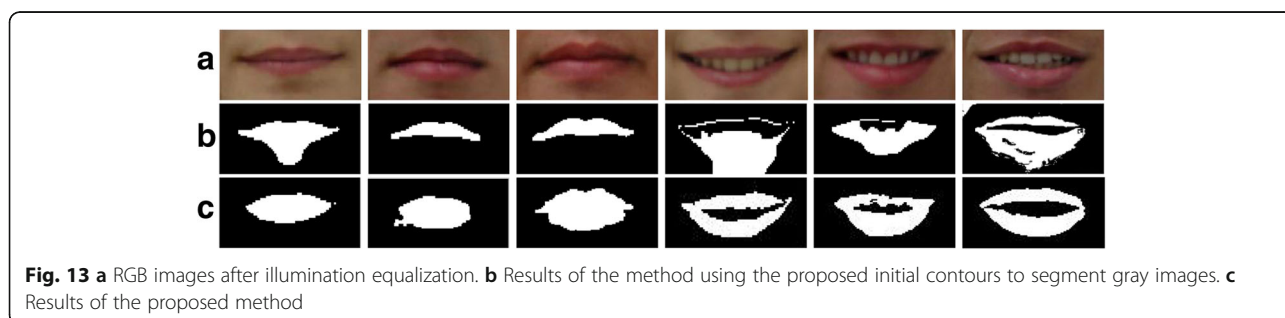


Table 1 Overlap and segmentation error values of two methods

		1 (%)	2 (%)	3 (%)	4 (%)
The proposed method	OL	83.1	88.5	80.7	75.4
	SE	17.7	13.1	21.1	29.2
Green component with circular initial contour	OL	80.3	85.1	69.9	61.8
	SE	16.6	13.1	36.4	48.0
Gray images with circular initial contour	OL	82.6	85.7	75.4	70.9
	SE	17.1	15.8	23.6	26.2
Combined color space images with circular initial contour	OL	81.1	84.6	76.3	74.1
	SE	16.3	16.8	10.2	11.4
Gray images with the proposed initial contours	OL	82.7	83.4	71.6	70.3
	SE	19.6	16.0	23.7	22.5

The ground truth of each image is obtained by manual segmentation.

We divide the images into two groups according to open mouth and closed mouth. Each group is randomly divided into two equal parts. Part 1 and 2 are closed mouth images and part 3 and 4 are open mouth images. The average values of each part are shown in Table 1. It can be found the OL values of the proposed method are higher than compared methods in the closed mouth group. Meanwhile, in the open mouth group, the OL values of the proposed method are higher than compared methods obviously, and the SE values of the former are lower than the latter evidently.

5 Conclusions

In this paper, we have proposed a LACM-based method using two initial contours to segment lip area. For open mouth, we utilized a combined semi-ellipse as the initial contours. And for closed mouth, a rhombus is used as the initial contour. Before the initial contour is determined, we did a set of preparatory works. First, we applied illumination equalization to RGB images to reduce interference of uneven illumination. Then, we adopted a combined color space which involves the U component in CIE-LUV color space and the sum of $C2$ and $C3$ components of DHT. Finally, we determined shape of initial contours due to the positions of four key points in the combined color space. It can be seen from the experiment that this method can obtain accurate, smooth, and robust results. Meanwhile, our results are more similar to the true lip boundary comparing with the results obtained by using a circle as initial contour to segment gray images and combined color space images. Nevertheless, as many other methods, this method cannot get particularly satisfactory results in the case of images containing gums, because gums are similar to lips in the color and texture. We will resolve this problem in future research.

Funding

The research was partially supported by the National Natural Science Foundation of China (no. 61571013), the Beijing Natural Science Foundation of China (no. 4143061), the Science and Technology Development Program of Beijing Municipal Education Commission (KM201710009003), and the Great Wall Scholar Reserved Talent Program of North China University of Technology (NCUT2017XN018013).

About the authors

Yuanyao Lu received his PH.D. degree from Chinese Academy of Sciences, Beijing, China, in 2006. He is now an associate professor with the Department of Electronic and Information Engineering, North China University of Technology, Beijing, China. His current research is mainly on image processing, visual perception, pattern recognition, and artificial intelligence. He has published more than 20 papers in international journals and conferences. Qingqing Liu received her B.S. degree from Shenyang University of Chemical Technology in 2014 and is now a master in the Department of Electronic Information Engineering, School of Electronic Information Engineering, North China University of Technology (NCUT).

Availability of data and materials

The datasets analyzed in current study is available from the corresponding author on reasonable request.

Authors' contributions

YL is the first and corresponding author. QL made the algorithm and experiments. Both authors read and approved the final manuscript.

Consent for publication

Written informed consent was obtained from the patient for the publication of this report and any accompanying images.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 7 November 2017 Accepted: 29 December 2017

Published online: 01 February 2018

References

1. Y Cheung, M Li, X Cao, X You, Lip segmentation under MAP-MRF framework with automatic selection of local observation scale and number of segments. *IEEE Trans. Image Process.* **23**(8), 3397–3411 (2014)
2. Y Cheung, M Li, Q Peng, CLP Chen, A cooperative learning-based clustering approach to lip segmentation without knowing segment number. *IEEE Trans. Neural Netw. Learn. Syst.* **28**(1), 80–93 (2017)

3. S Gao, J Yang, in *21st International Conference on Pattern Recognition (ICPR 2012)*. Saliency-seeded localizing region-based active contour for automatic natural object segmentation (2012), pp. 3644–3647
4. V Caselles, R Kimmel, G Sapiro, Geodesic active contours. *Int. J. Comput. Vis.* **22**, 61–79 (1997)
5. M Srikham, in *21st International Conference on Pattern Recognition (ICPR 2012)*. Active contours segmentation with edge based and local region based (2012), pp. 1989–1992
6. S Lankton, A Tannenbaum, Localizing region-based active contours. *IEEE Trans. Image Process.* **17**(11), 2029–2039 (2008)
7. M Kass, A Witkin, D Terzopoulos, Snakes: active contour models. *Int. J. Comput. Vis.* **1**, 321–331 (1988)
8. CY Xu, JL Prince, Snakes, shapes and gradient vector flow. *IEEE Trans. Image Process.* **7**(3), 359–369 (1998)
9. T Chan, L Vese, Active contours without edges. *IEEE Trans. Image Process.* **10**(2), 266–277 (2001)
10. D Mumford, J Shah, Optimal approximation by piecewise smooth function and associated variational problem. *Commun Pure Appl Math.* **42**(5), 577–685 (1989)
11. JA Yezzi, A Tsai, A Willsky, A fully global approach to image segmentation via coupled curve evolution equations. *J Vis Commun Image Represent.* **13**(1–2), 195–216 (2002)
12. R Ronfard, Region-based strategies for active contour models. *Int. J. Comput. Vis.* **13**(2), 229–251 (1994)
13. N Paragios, R Deriche, Geodesic active regions and level set methods for supervised texture segmentation. *Int. J. Comput. Vis.* **46**, 223–247 (2002)
14. S Osher, R Fedkiw, *Level Set Method and Dynamic Implicit Surfaces* (Springer-Verlag, New York, 2002)
15. T Hoang Ngan Len, Marios Savvides, *A Novel Shape Constrained Feature-Based Active Contour Model for Lips/Mouth Segmentation in the Wild*. *Pattern Recognition.* **54**, 23–33 (2016)
16. S Osher, JA Sethian, Fronts propagating with curvature dependent speed: algorithms based on Hamilton-Jacobi formulations. *J. Comput. Phys.* **79**(1), 12–49 (1988)
17. M Li, Y-M Cheung, *Automatic Segmentation of Color Lip Images Based on Morphological Filter, AMT'10 Proceedings of the 6th International Conference on Active Media Technology* (Springer-Verlag Berlin, Heidelberg, 2010), pp. 296–305
18. SW Chin, KP Seng, L-M Ang, KH Lim, in *Proceedings of the International Multi Conference of Engineers and Computer Scientists 2009*. New lips detection and tracking system, vol 1 (2009), pp. 18–20
19. M Li, Y-m Cheung, Automatic lip localization under face illumination with shadow consideration. *Signal Process.* **89**(12), 2425–2434 (2009)
20. H Talea, K Yaghmaie, in *Communication Software and Networks (ICCSN), 2011 IEEE 3rd International Conference on*. Automatic combined lip segmentation in color images (2011), pp. 109–112
21. Y-U Kim, S-K Kang, S-T Jung, in *Information Reuse & Integration, 2009. IRI '09. IEEE International Conference on*. Design and implementation of a lip reading system in smart phone environment (2009), pp. 101–104
22. A Hulbert, T Poggio, Synthesizing a color algorithm from examples. *Science* **239**, 482–485 (1998)
23. U Canzlerm, T Dziurzyk, in *Proceedings of IAPR Workshop*. Extraction of non manual features for video based sign language recognition (2002), pp. 318–321
24. SH Leung, SL Wang, WH Lau, Lip image segmentation using fuzzy clustering incorporating an ellipse shape function. *IEEE Trans. Image Process.* **13**(1), 51–62 (2004)
25. X Wang, Y Hao, F Degang, C Yuan, in *Neural Networks and Signal Processing, 2008 International Conference on*. ROI processing for visual features extraction in lip-reading (2008), pp. 178–181
26. SL Wang, WH Lau, SH Leung, Automatic lip contour extraction from color image. *Pattern Recogn.* **37**, 2375–2387 (2004)
27. E Skodras, N Fakotakis, in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*. An unconstrained method for lip detection in color images (2011), pp. 1013–1016
28. SL Wang, WH Lau, SH Leung, AWC Liew, in *Acoustics, Speech, and Signal Processing, 2004. Proceedings (ICASSP '04). IEEE International Conference on*. Lip segmentation with the presence of beards, vol 3 (2004), p. iii-529
29. MM Hosseini, S Ghoirani, in *Intelligent Systems, 2009. GCIS'09. WRI Global Congress on*. Automatic lip extraction based on wavelet transform, vol 4 (2009), pp. 393–396
30. Aris Nasuha, *Lip Segmentation and Tracking Based Chan-Vese Model*. *Intl Conf. on Information Technology and Electrical Engineering (ICITEE 2013)*, Oct 7-8 (2013)
31. X Ma, Z Hongqiao, in *Control Conference (CCC), 34th Chinese*. Lip segmentation algorithm based on bi-color space (2015)
32. Y Gong, Speech recognition in noisy environments: a survey [J]. *Speech Comm.* **16**, 261–291 (1995)
33. LI Jian, CHENG Changkui, JIANG Tianyan. Wavelet de-noising of partial discharge signals based on genetic adaptive threshold estimation [J]. *IEEE Trans. Dielectr. Electr. Insul.*, 2012, 19 (20):543-549
34. LI Ruwei, BAO Changchun, XIA Bingyin, et al. Speech enhancement using the combination of adaptive wavelet threshold and spectral subtraction based on wavelet packet decomposition [C]. *ICSP 2012 Proceedings*, 2012: 481-484
35. AWC Liew, SH Leung, WH Lau, Lip contour extraction from color images using a deformable model. *Pattern Recogn.* **35**(12), 2949–2962 (2002)
36. Y-m Cheung, X Liu, X You, A local region based approach to lip tracking. *Pattern Recogn.* **45**, 3336–3347 (2012)
37. Z Li, Z Yu, W Liu, Z Zhang, in *Information Science and Control Engineering (ICISCE), 4th International Conference*. Tongue image segmentation via color decomposition and thresholding (2017)
38. Lewis T W., Powers D M. Lip feature extraction using red exclusion [C]. *Pan Sydney Workshop on Visual Information Processing*. Sydney: Australian Computer Society, 2000, 61-67
39. Rao R. and Mersereau R. Lip modeling for visual speech recognition [C]. *28th Annual Asimolar Conference on Signals, Systems, and Computers*, Pacific Grove, 2. IEEE Computer Society, 1994, 1:587-590
40. B Mark, H Eun Jung, O Robyn, in *Proceedings of the Asian Conference on Computer Vision*. Lip tracking using pattern matching snakes (2002), pp. 23–25
41. AWC Liew, SH Leung, WH Lau, Segmentation of color lip images by spatial fuzzy clustering. *IEEE Trans. Fuzzy Syst.* **11**(4), 542–549 (2003)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com