Tongue Image Segmentation via Thresholding and Gray Projection

Weixia Liu[1,2,3], Jinmei Hu[2], Zuoyong Li[2,3], Zuchang Zhang[2], Zhongli Ma[4,2], Daoqiang Zhang[5,2]

[1] Straits Institute, Minjiang University, Fuzhou 350121, China

[2] Fujian Provincial Key Laboratory of Information Processing and Intelligent Control, Minjiang University, Fuzhou 350121, China

[3] Internet Innovation Research Center (Minjiang University), The Fujian College's Research Base of Humanities and Social Science, Fuzhou 350121, China

[4] College of Automation, Harbin Engineering University, Harbin 150001, China

[5] Department of Computer Science and Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

[email: Liuweixia0201@126.com, jinmeihu@126.com, fzulzytdq@126.com, youarelucky@yeah.net, mazhongli@hrbeu.edu.cn, dqzhang@nuaa.edu.cn]

*Corresponding author: Zuoyong Li

Abstract

Tongue diagnosis is one of the most important diagnostic methods in Traditional Chinese Medicine (TCM). Tongue image segmentation aims to extract the image object (i.e., tongue body), which plays a key role in the process of manufacturing an automated tongue diagnosis system. It is still challenging, because there exists the personal diversity in tongue appearances such as size, shape, and color. This paper proposes an innovative segmentation method that uses image thresholding, gray projection and active contour model (ACM). Specifically, an initial object region is first extracted by performing image thresholding in HSI (i.e., Hue Saturation Intensity) color space, and subsequent morphological operations. Then, a gray projection technique is used to determine the upper bound of the tongue body root for refining the initial object region. Finally, the contour of the refined object region is smoothed by ACM. Experimental results on a dataset composed of 100 color tongue images

showed that the proposed method obtained more accurate segmentation results than other available state-of-the-art methods.

1. Introduction

Tongue diagnosis [1-3] is popular in Traditional Chinese Medicine (TCM) due to its effectiveness, painlessness, and lack of side-effects. In China, tongue diagnosis has gone through over 3000 years, and its practitioners infer a subject's health status from tongue appearances, such as the color, texture, and coating. TCM has eight famous tongue diagnosis rules [3], which reveal that different tongue body sub-regions reflect different human organs' health statuses. Furthermore, the tongue appearance is an important index to reflect a subject's health condition. For instance, the TCM syndrome, which depicted by the tongue coating's color and texture features, often reflects human health status [4].

Modern Western Medicine gradually regards the human tongue as an extension of the upper gastrointestinal tract indicating the human health status. Accordingly, some researchers [5-6] have also agreed that tongue diagnosis is beneficial for making clinical decision [7]. For instance, some researchers took tongue coating as a risk index for edentate patients' aspiration pneumonia, as it was closely related to many viable salivary bacteria [8]. Moreover, the study in [9] takes tongue amyloidosis as a possible diagnostic indicator for the disease of plasmacytoma. An increasing number of researches endeavour to explore the potential of tongue diagnosis to infer systemic disorders.

However, conventional tongue diagnosis often depends on the experience of practitioners, which is subjective, time-consuming, and instable. Nowadays, several automated computer-aided tongue diagnosis systems [10-12] have been developed by using digital image processing [13-14] and pattern recognition techniques [15]. These automated systems usually first extract the image object and its features, and then feed into a designed classifier for tongue diagnosis. Tongue image segmentation is a prerequisite and also crucial for developing automated tongue diagnosis systems. Several methods [16-19] have been presented for tongue image segmentation in the past few decades. However, it is still challenging due to the personal diversity in tongue appearances.

Among the existing segmentation methods, ACM (i.e., snake)-based methods are popular. ACM [20-22] is a popular deformable shape model for contour extraction, which evolves a given initial contour to the true object contour. The determination of the initial object

contour plays a key role in ACM-based tongue image segmentation. When the initial contour contains strong fake object contour, it is difficult to converge to the true object contour. Motivated by our observation that there is an obvious hue difference between tongue body pixels and their neighboring face pixels, we proposed two tongue image segmentation methods [23-24]. The first method [23] published on the conference is a preliminary version of the second method [24] that is published on the journal. The method in [23] first maps an image from RGB to HSI, and conducts image thresholding on the hue component to extract an initial object region, and then performs image thresholding on the red component to find the gap region between the upper lip and the tongue body root, and finally uses the gap region to remove fake object region and obtain final tongue body region. On the basis of the first method [23], the second method [24] can more accurately find the above gap region by adaptively selecting one of two image thresholding results on the red component. Motivated by both methods and ACM-based methods, we propose this work to improve tongue image segmentation accuracy. Differences between this work and our previous works [23-24] are as follows: (1) this work simplifies the hue component transformation for parameter reducing; (2) this work refines the initial object region by gray projection based determination of the tongue body root's upper bound instead of image thresholding based determination of the above gap region; (3) this work smoothes the tongue body contour via ACM. Specifically, in the proposed method, an initial object region is first extracted via image thresholding on a transformed hue component, and subsequent morphological operations. Then, a gray projection technique is used to determine the upper bound of the tongue body root for refining the initial object region. Finally, the initial object contour is smoothed by ACM. Experiments on a dataset composed of 100 color tongue images with personal diversity in tongue appearances showed that the proposed method improved tongue image segmentation accuracy.

The rest of this article is organized as follows. We first review related works in Section 2, and then introduce the theory and implementation of the proposed method in Section 3. We report experimental results in Section 4, and conclude in Section 5.

2. Related works

Due to the above challenges associated with tongue image segmentation, single conventional image processing techniques, such as edge detection and image thresholding, usually fail to achieve a satisfactory segmentation result. To improve segmentation accuracy, several hybrid methods have been proposed during the past few decades. ACM-based methods are the most popular ones. Studies of ACM-based methods have the following two steps, i.e., the determination of the initial object contour, and the object contour evolution. This paper

focuses on the initial object contour determination. When determining the initial object contour, existing studies have mainly used several low-level image processing techniques, such as prior shape-based ellipse detection, edge detection, region segmentation, and feature point detection.

For example, the bi-elliptical deformable contour extraction method, termed BEDC [15], uses the tongue body shape prior to determine the initial object contour for image segmentation. Specifically, BEDC [15] first defines a specific deformable template, termed BEDT, to roughly describe the tongue body, and then obtains the initial tongue body contour by minimizing the BEDT energy function. Finally, a modified ACM, that replaces conventional internal force with the template force, is used to evolve the initial contour and obtain the final segmentation result. BEDC simply uses two semi-ellipses to model the tongue body shape. However, large personal variation in tongue body shape can result in the initial tongue body contour obtained by BEDC containing undesirable strong edges from neighboring tissues, thus providing a tongue body contour that does not reflect the real tongue body contour.

Zhang et al. [16] combined polar edge detection with ACM to achieve tongue image segmentation. In detail, this method first detects image boundaries in polar coordinate and removes fake object boundaries with the help of an edge mask, and then performs local image thresholding and morphological operations on the edge removal result to obtain a binary polar edge image, and finally adopts a heuristic method to obtain the initial object contour for ACM-based contour smoothing. Unfortunately, there exist two disadvantages in this method: 1) there is no common edge mask for removing fake object boundaries, because tongue body size and shape vary from person to person; 2) the edge filtering scheme, which combines a Sobel edge detector, a Gaussian filter, and image thresholding with morphological operations, fails to remove long fake object boundaries. Further, the Gaussian filter often weakens the true tongue body contour, thus increasing the difficulty of tongue image segmentation.

Ning et al. [17] proposed a method called GVFRM, which combined gradient vector flow, region merging, with ACM to accomplish the task of tongue image segmentation. In detail, GVFRM first suppresses both noise and trivial image details by modifying the conventional GVF as a scalar diffusion equation, then splits the image into many small regions by a watershed algorithm, and finally utilizes region merging to determine the initial object contour for subsequent contour smoothing. However, this method also has two disadvantages: 1) it easily generates segmentation error on the object regions near image borders, because the markers of these object regions are wrongly assigned as background under the invalid assumption that the tongue body and the background should be at the image center and the image borders, respectively; 2) the modified GVF may weaken the true tongue body contour when it suppresses image noise and trivial image details.

Shi et al. presented two ACM-based methods briefly called C2G2F [18] and DGF [19] for tongue image segmentation. In detail, C2G2F [18] first detects four feature points of the tongue body to determine the initial object contour, and divides the initial contour into the upper and the lower parts. Next, the upper and the lower halves are evolved to the true object contour by the parameterized GVF snake model and the geodesic ACM, respectively. Finally, the two half profiles are assembled to obtain the final object contour. However, C2G2F may miss partial feature points, or detect undesirable points. To resolve this issue, Shi et al. [19] developed an upgraded method called DGF. In detail, DGF first roughly localizes the image window of the tongue body via a salient object detector [25], then detects four feature points within the image window and follows the idea of C2G2F [18] to find the initial object contour. Next, DGF uses the geodesic ACM and the geo-GVF ACM to evolve the lower and the upper half contours to the true object contour, respectively. Finally, both half contours are assembled to form the final object contour. However, DGF still has similar limitation to C2G2F.
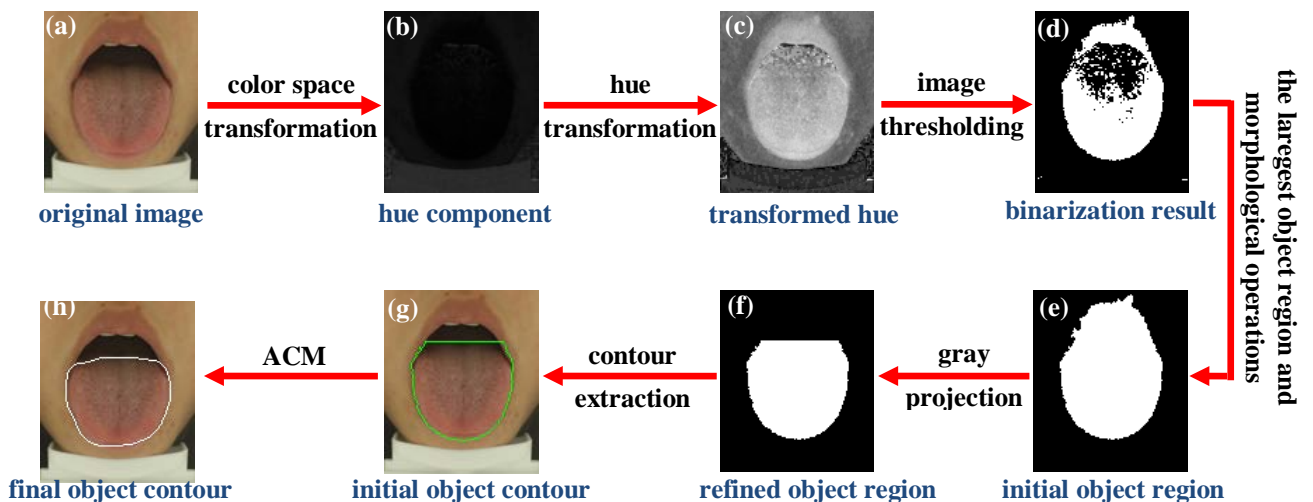

Fig. 1. The flowchart of the proposed method.

3. The proposed method

It is well known that ACM-based methods are usually sensitive to the initial object contour. When the initial object contour contains strong fake object contour, it is difficult to converge to the true object contour. To improve the segmentation accuracy, we focused on how to obtain an initial object contour that is close to the true tongue body contour. After exploring the characteristic of the hue component, we present a new segmentation method based on image thresholding, gray projection, and ACM. The flowchart of the proposed method is shown in **Fig. 1**. Our contributions are as follows.

(1) After further exploring our revealed image characteristic that hue values of the tongue body pixels and the upper lip pixels usually are higher or lower than those of their neighboring pixels, we presented a simplified hue component transformation scheme that can reduce the parameter in the proposed algorithm. After performing this scheme, both the tongue body pixels and the upper lip pixels usually have higher hue values than their neighboring pixels. We used the transformed hue values to find the initial object region.

(2) We introduced the gray projection technique to refine the initial object region and obtain the initial object contour, which is smoothed by ACM.

3.1 Extraction of the initial object region

In existing ACM-based methods, the initial object region and its corresponding contour are usually extracted by using image processing techniques, such as prior shape based ellipse detection [15], edge detection [16], region segmentation [17], and feature point detection [18-19]. Differing from these methods, we extracted the initial object region more stably and accurately by image thresholding on the transformed hue values. The detailed process is as follows:

(1) Color space transformation: an image is transformed from RGB to HSI via the following equations:

$$H = \begin{cases} \theta, & G \geq B \\ \theta - 2\pi, & G < B \end{cases} \tag{1}$$

$$S = 1 - \frac{3}{R+G+B} \min\{R, G, B\}, \tag{2}$$

$$I = \frac{1}{3}(R + G + B), \tag{3}$$

where

$$\theta = \arccos\left\{ \frac{[(R-G)+(R-B)]/2}{[(R-G)^2+(R-B)(G-B)]^{1/2}} \right\}. \tag{4}$$

Taking **Fig. 1(a)** as an example, its hue component is shown in **Fig. 1(b)**. This figure explores an image characteristic that the tongue body pixels and the upper lip pixels are usually brighter or darker than their neighboring pixels, where brighter pixels have higher hue values. Based on the image characteristic, we will perform the hue component transformation and image thresholding in the next two steps to extract the initial object region.

(2) Hue component transformation: the image hue component is transformed as

$$H'(i,j) = \max\{H(i,j), H_{max} - H(i,j)\}, \tag{5}$$

where H_{max} indicates the maximum hue value in the image, and (i, j) is the pixel coordinate. In comparison with the hue component transformation in [23-24], here, we simplify the transformation and reduce one parameter. **Fig. 1(c)** shows the transformed hue component of

Fig. 1(a). **Fig. 1(c)** demonstrates that the transformed hue values of most tongue body pixels and upper lip pixels are higher than those of their neighboring pixels. In the next step, we will perform image thresholding to extract the initial object region.

(3) Image thresholding: an image thresholding on the transformed hue component is performed to obtain an image binarization result:

$$B(i,j) = \begin{cases} 1, & \text{if } H'(i,j) > T \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

where

$$T = V_{H'}(\alpha N). \tag{7}$$

In Eq. (7), $V_{H'}$ indicates the sorted vector of H' with descending order, N is the total number of pixels, and α is a parameter controlling the ratio of object pixels in B. The image binarization result of **Fig. 1(c)** is shown in **Fig. 1(d)**.

(4) Extraction of the initial object region: the proposed method finds the maximum white region in B, and successively performs three morphological operations (i.e., "imdilate", "imfill", and "imerode") to refine it as the initial object region shown in $\hat{B}$. The structural elements used in "imdilate" and "imerode" are of shape "disk" and radius "1" as shown in **Fig. 2**. Because the tongue body shape is similar to a disk, the structural element adopts the shape "disk". In addition, Section 4.3 will discuss the impact of the radius "r" of the structural element on segmentation accuracy of our algorithm, and will explain why the radius should be set to 1 in our algorithm. The extracted initial object region is exhibited in **Fig. 1(e)**.

0	1	0
1	1	1
0	1	0

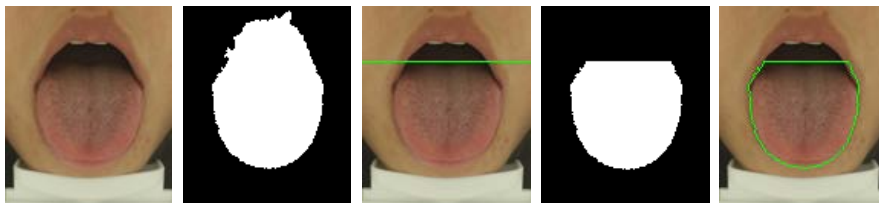**Fig. 2.** Structural element of shape "disk" and radius "1".

Fig. 3. intermediate results of the proposed method when refining the initial object region.

(a) Original image, (b) the initial object region, (c) (a) with a green line indicating our determined upper bound of the object, (d) the refined object region, (e) the object contour.

3.2 Refinement of the initial object region

When extracting the initial object region, the proposed method is prone to misclassification of the upper lip and the gap region between the upper lip and the tongue body root. To resolve this issue, we introduce the gray projection technique [26] to find the upper bound of the tongue body root, and employ the bound to remove the upper lip and the gap region. The detailed process is as follows:

(1) Determination of the upper bound of the tongue body root: specifically, we first find the locations of object pixels in the initial object region extraction result, i.e., $\hat{B}$. Then, we take the red component of a tongue image as a gray image, and calculate the average gray value of the object pixels on each image row containing object pixels. Finally, we find the row with the lowest average gray value among the image rows containing object pixels, and this is considered as the upper bound of the tongue body root. If there are two or multiple image rows with the same lowest average gray value, we take the image row with the maximum row number as the upper bound. Taking the subject in **Fig. 1** as an example, **Fig. 3(a)** and **Fig. 3(b)** again show the original tongue image and the extracted initial object region. **Fig. 3(c)** shows the location of the upper bound using a green line on the original image. From **Fig. 3(c)**, it can be observed that the determined upper bound is very close to the root of the true tongue body.

(2) Refinement of the initial object region: we first remove the white pixels on the image rows above the upper bound from the image binarization result $\hat{B}$. This may result in the sole white region in $\hat{B}$ becoming two or multiple white regions. Therefore, we need to choose the largest white region as the refined object region. **Fig. 3(d)** and **Fig. 3(e)** show the refined object region and its corresponding object contour. From **Fig. 3(e)**, it can be observed that the object region is effectively refined.

The underlying principle of the above refinement method is that the above gap region is usually darker than the tongue body and the upper lip. Accordingly, transitional object pixels near the tongue body root are darker than other object pixels. Therefore, among the image rows containing object pixels, the image row with the lowest average gray value can be taken as the upper bound of the tongue body root.

3.3 ACM-based object contour smoothing

After refining the initial object region, the boundary of the refined object region is used as the initial object contour for subsequent ACM-based contour smoothing. The same ACM [22] used in GVFRM [17] is utilized to smooth the initial object contour. To validate the effectiveness of ACM for contour smoothing, **Fig. 4** shows the initial object contour and the

smoothed contour. Obviously, the refined object contour is smoother than the initial object contour.

Fig. 4. Smoothed result of the initial object contour.
(a) Original image, (b) initial contour, (c) smoothed contour.

4. Experimental results

To evaluate segmentation performance of different methods, extensive experiments were conducted on a dataset composed of 100 color tongue images with the sizes of 110×130. First, qualitative comparisons between the proposed method and three state-of-the-art methods (i.e., GVFRM [17], C2G2F [18], and DGF [19]) were performed on eight representative tongue images. Then, quantitative comparisons on the entire tongue image dataset were evaluated by using four common image classification measures, i.e., misclassification error (ME) [27], false positive rate (FPR), false negative rate (FNR) [28], and kappa index (KI) [29]. ME measures the percentage of background pixels erroneously classified into foreground (i.e., object), and conversely, foreground pixels erroneously assigned to background. FPR and FNR measure classification error in detail. FPR measures the rate of the number of background pixels misclassified into foreground to the total number of background pixels in the manual ideal segmentation result (ground truth). FNR measures the rate of the number of foreground pixels misclassified into background to the total number of foreground pixels in the ground truth. FPR and FNR indicate over-segmentation and under-segmentation, respectively. KI measures the ratio of overlapping foreground area between the automatic segmentation result and the ground truth. The definitions of ME, FPR, FNR, and KI are as follows:

$$\mathrm{ME} = 1 - \frac{|B_m \cap B_a| + |F_m \cap F_a|}{|B_m| + |F_m|}, \tag{8}$$

$$\mathrm{FPR} = \frac{|B_m \cap F_a|}{|B_m|}, \tag{9}$$

$$\mathrm{FNR} = \frac{|F_m \cap B_a|}{|F_m|}, \tag{10}$$

$$KI = 2 \frac{|F_m \cap F_a|}{|F_m| + |F_a|}, \tag{11}$$

where B_m and B_a indicate background of the ground truth and a certain method's segmentation result, respectively; F_m and F_a are their respective foreground; and $|.|$ is the cardinality of a set. The four measurements range between 0 and 1. The lower the values of ME, FPR, and FNR, the better the segmentation. And conversely, the higher the value of KI, the better the segmentation.

In our experiments, the parameters α and r in the proposed method were set to 0.3 and 1, respectively. For GVFRM [17], we set the optimal iteration number of GVF-based image diffusion corresponding to the highest average KI value [29]. Other parameters of GVFRM were set according to the literature [17]. Parameters of C2G2F [18] and DGF [19] were in accordance with their own literatures.

4.1 Results of qualitative evaluation

Fig. 5 exhibits the segmentation results of the four methods on the eight representative tongue images with personal diversity in tongue apperances including shape, size, color, texture, and coating. Among these methods, GVFRM [17] achieves a good segmentation only on the 4th image (i.e., **Fig. 5(d)**), but generates misclassification on the other images. In detail, under-segmentation happens to **Figs. 5(a)-(b), (e)-(f)**, and **(h)** ; and over-segmentation happens to **Figs. 5(a)-(e)** and **(g)**. Both C2G2F [18] and DGF [19] generate misclassification on most images. For C2G2F, under-segmentation happens to **Figs. 5(a)-(b)** and **(g)**, and over-segmentation happens to **Figs. 5(a)-(h)**. Similarly, for DGF, under-segmentation happens to **Figs. 5(a)** and **(g)**, and over-segmentation happens to **Figs. 5(a)-(h)**. In general, DGF alleviates the degree of over-segmentation as compared to C2G2F. Among the four methods, our propposed method obtains the best segmentation result on each representative image, and our object contours are the closest to the true tongue body contours. This group of experiments demonstrates superiority of our proposed method over other three methods. However, our segmentation results shown in **Figs. 5(b)-(c)** and **(g)-(h)** still have some over-segmentation, which will be improved in our future work.
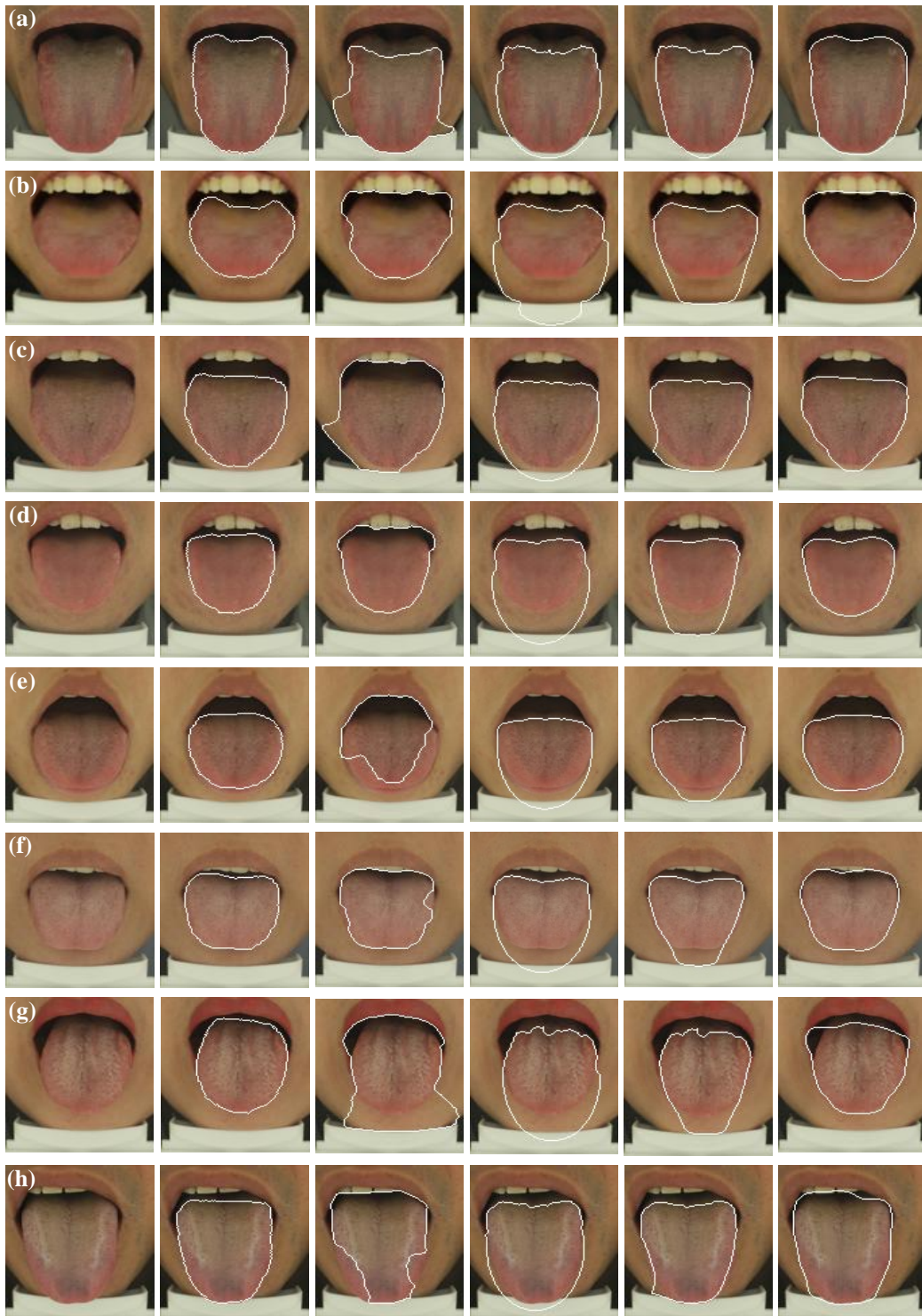
Fig. 5. Segmentation results on the eight representative tongue images, where image columns 1-6 indicate original images, ground truths, GVFRM [17], C2G2F [18], DGF [19], and the proposed method.
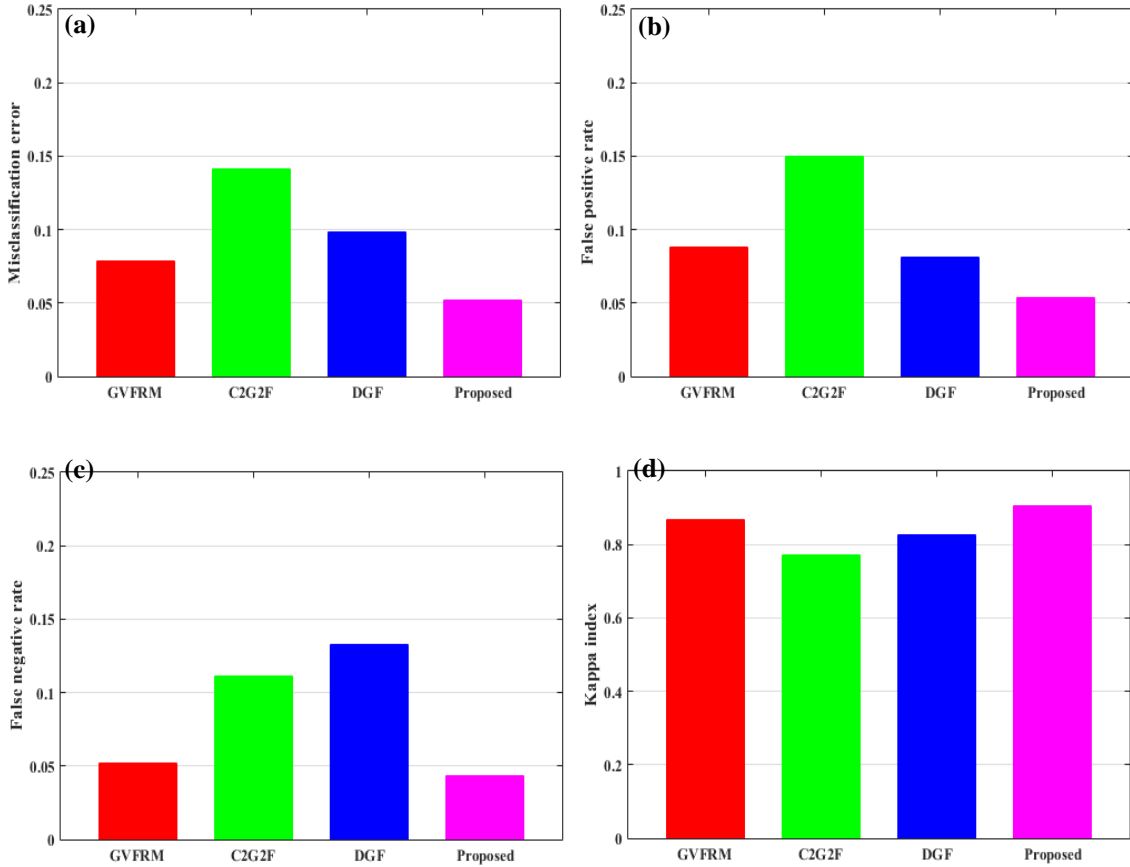
Fig. 6. Bar charts of four average quantitative measurement values on the entire image dataset.

4.2 Results of quantitative evaluation

The segmentation results for the entire image dataset achieved with GVFRM [17], C2G2F [18], DGF [19], and the proposed method were evaluated by ME, FPR, FNR, and KI. Quantitative comparisons are shown in **Figs. 6(a)-(d)**. In detail, the means and standard deviations of ME values obtained by the four methods were 0.079 ± 0.042, 0.141 ± 0.049, 0.098 ± 0.044, and 0.052 ± 0.026, respectively. The means and standard deviations of FPR values were 0.088 ± 0.060, 0.150 ± 0.061, 0.081 ± 0.050, and 0.054 ± 0.032, respectively. The means and standard deviations of FNR values were 0.052 ± 0.083, 0.111 ± 0.079, 0.133 ± 0.091, and 0.043 ± 0.056, respectively. The means and standard deviations of KI values were 0.869 ± 0.067, 0.772 ± 0.083, 0.826 ± 0.080, and 0.906 ± 0.047, respectively. These quantitative results demonstrate the higher segmentation accuracy of the proposed method than other three methods.

Table 1. Average ME values obtained by the proposed method under different combinations of α and r

α \ r	1	2	3	4	5
0.1	0.2086	0.2074	0.2075	0.2057	0.2041
0.2	0.0982	0.0930	0.0929	0.0904	0.0877
0.3	0.0520	0.0523	0.0524	0.0525	0.0525
0.4	0.0767	0.0782	0.0784	0.0790	0.0807
0.5	0.1196	0.1221	0.1243	0.1253	0.1282

Table 2. Average KI values obtained by the proposed method under different combinations of α and r

α \ r	1	2	3	4	5
0.1	0.3886	0.3953	0.3945	0.4040	0.4122
0.2	0.7896	0.8040	0.8048	0.8115	0.8188
0.3	0.9056	0.9052	0.9047	0.9046	0.9046
0.4	0.8675	0.8654	0.8648	0.8642	0.8616
0.5	0.8105	0.8080	0.8052	0.8039	0.8001

4.3 Parameter selection

The proposed method has two parameters, i.e., α and r. The proposed method uses α to control the ratio of exracted object pixels from the transformed hue component, and uses r as the structural element radius of porhoplogical operations when extracting the initial object region. We disscussed the impact of α and r on the segmentation accuracy for the entire dataset composed of 100 tongue images, where α and r were selected from {0.1, 0.2, 0.3, 0.4, 0.5} and {1, 2, 3, 4, 5}, respectively. The average ME values and the average KI values are listed in **Table 1** and **Table 2**, respectively. Lower ME values indicate better segmentation, while higher KI values indicate better segmentation. Both tables show that the segmentation accuracy first increases and then decreases with increases in α under each r. The best segmentation accuracy with the lowest ME value and the highest KI value was obtained with α=0.3 under each r. When α=0.3, both tables show that the segmentation accuracy decreases with increases in r. The best segmentation accuracy with the lowest ME value and the highest KI value was obtained with α=0.3 and r=1. Therefore, we set α and r to 0.3 and 1 in our experiments, respectively.

5. Conclusions

To improve tongue image segmentation accuracy, we developed a method that integrates image thresholding, gray projection, and ACM. The main feature of the proposed method is that it allows more accurate determination of the initial tongue body contour with the use of image thresholding and gray projection. Experimental results on the dataset composed of 100 tongue images demonstrate that the proposed method achieves higher segmentation accuracy

than other state-of-the-art methods. However, the proposed method is prone to generation of over-segmentation, which needs to be resolved in future work.

Acknowledgments

References

[1] G. Maciocia, Tongue Diagnosis in Chinese Medicine, Seattle, WA: Eastland Press, 1995.
    Article (CrossRef Link)

[2] Z. Xie, Practical traditional Chinese medicine, Beijing: Foreign Language Press, 2000.
    Article (CrossRef Link)

[3] B. Kirschbaum, Altas of Chinese tongue diagnosis, Seattle, WA: Eastland, 2000.
    Article (CrossRef Link)

[4] C. H. Hsu, M. C. Yu, C. H. Lee, T. C. Lee and S. Y. Yang, "High eosinophil cationic protein level in asthmatic patients with Heat Zheng," *American Journal of Chinese Medicine*, vol. 31, no. 2, pp. 277–283, February 2003. Article (CrossRef Link).

[5] J. A. Ship, J. Phelan and A. Kerr, Biology and pathology of the oral mucosa, Fitzpatrick's Dermatology in General Practice, McGraw-Hill, New York, July 2003. Article (CrossRef Link)

[6] Y. Zadik, S. Drucker and S. Pallmon, "Migratory stomatitis (ectopic geographic tongue) on the floor of the mouth," *Jounal ofthe American Academy of Dermatology*, vol. 65, no. 2, pp. 459–460, August 2011. Article (CrossRef Link)

[7] J. K. Anastasi, L. M. Currie and G. H. Kim, "Understanding diagnostic reasoning in TCM practice: tongue diagnosis," *Alternative Therapies in Health & Medicine*, vol. 15, no. 3, pp. 18–28, May/June 2009. Article (CrossRef Link)

[8] S. Abe, K. Ishihara, M. Adachi and K. Okuda, "Tongue-coating as risk indicator for aspiration pneumonia in edentate elderly," *Archivesof Gerontology & Geriatrics*, vol. 47, no. 2, pp. 267–275, September–October 2008. Article (CrossRef Link).

[9] S. Hoefert, E. Schilling, S. Philippou and H. Eufinger, "Amyloidosis of the tongue as a possible diagnostic manifestation of plasmacytoma," *Mund Kiefer Gesichtschir*, vol. 3, no. 1, pp. 46–49, January 1999. Article (CrossRef Link).

[10] B. Pang, D. Zhang, N. Li and K. Wang, "Computerized tongue diagnosis based on Bayesian networks," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 10, pp. 1803–1810, September 2004. Article (CrossRef Link).

[11] B. Pang, K. W. Wang, D. Zhang and F. Zhang, "On automated tongue image segmentation in Chinese Medicine," in *Proc. of International Conference on Pattern Recognition*, pp. 616–619, August 11-15, 2002. Article (CrossRef Link).

[12] B. Pang, D. Zhang and K. Wang, "The bi-elliptical deformable contour and its application to automated tongue segmentation in Chinese medicine," *IEEE Transactions on Medical Imaging*, vol. 24, no. 8, pp. 946–956, August 2005. Article (CrossRef Link).

[13] J. Zhai, J. Zhou, Y. Ren and Z. Wang, "Salient object detection via multiple random walks," *KSII Transactions on Internet and Information Systems*, vol. 10, no. 4, pp. 1712–1731, April 2016. Article (CrossRef Link).

[14] S. M. Li, J. G. Zhu, C. Gao and C. W. Li, "A two-stage cascade foreground seeds generation for parametric min-cuts," *KSII Transactions on Internet and Information Systems*, vol. 10, no. 11, pp. 5563–5582, November 2016. Article (CrossRef Link)

[15] S. Lim and D. Lee, "Real-time eye tracking using ir stereo camera for indoor and outdoor environments," *KSII Transactions on Internet and Information Systems*, vol. 11, no. 8, pp. 3965–3983, August 2017. Article (CrossRef Link)

[16] H. Zhang, W. Zuo, K. Wang and D. Zhang, "A snake-based approach to automated segmentation of tongue image using polar edge detector," *International Journal of Imaging Systems & Technology*, vol. 16, no. 4, pp. 103–112, February 2006. Article (CrossRef Link).

[17] J. Ning, D. Zhang, C. Wu and F. Yue, "Automatic tongue image segmentation based on gradient vector flow and region merging," *Neural Computing & Applications*, vol. 21, no. 8, pp. 1819–1826, November 2012. Article (CrossRef Link).

[18] M. Shi, G. Li and F. Li, "C2G2FSnake: automatic tongue image segmentation utilizing prior knowledge," *Science China Information Sciences*, vol. 56, no. 9, pp. 1–14, September 2013. Article (CrossRef Link).

[19] M. Shi, G. Li, F. Li and C. Xu, "Computerized tongue image segmentation via the double geo-vector flow," *Chinese Medicine*, vol. 9, no. 1, pp. 7–16, February 2014. Article (CrossRef Link).

[20] L. Wang, Y. Chang, H. Wang, Z. Wu, J. Pu and X. Yang, "An active contour model based on local fitted images for image segmentation," *Information Sciences*, vol. s418-419, pp. 61–73, December 2017. Article (CrossRef Link).

[21] B. Han and Y. Wu, "A novel active contour model based onmodified symmetric cross entropy for remote sensing river image segmentation," *Pattern Recognition*, vol. 67, pp. 396–409, July 2017. Article (CrossRef Link).

[22] K. Zhang, H. Song and L. Zhang, "Active contours driven by local image fitting energy," *Pattern Recognition*, vol. 43, no. 4, pp.1199–1206, April 2010. Article (CrossRef Link).

[23] Z. Li, Z. Yu, W. Liu and Z. Zhang, "Tongue image segmentation via color decomposition and thresholding," in *Proc. of 4th International Conference on Information Science and Control Engineering*, pp. 752–755, July 21-23, 2017. Article (CrossRef Link).

[24] Z. Li, Z. Yu, W. Liu, Y. Xu, D. Zhang and Y. Cheng, "Tongue image segmentation via color decomposition and thresholding," *Concurrency and Computation: Practice and Experience*, vol. e4662, pp. 1–9, September 2018. Article (CrossRef Link).

[25] J. Feng, Y. Wei, L. Tao, C. Zhang and J. Sun, "Salient object detection by composition," in *Proc. of IEEE International Conference on Computer Vision*, pp. 1028–1035, November 6-13, 2011. Article (CrossRef Link).

[26] L. Zhang and J. Qin, "Tongue-image segmentation based on gray projection and threshold-adaptive method," *Journal of Clinical Rehabilitative Tissue Engineering Research*, vol. 14, no. 9, pp. 1638–1641, February 2010. Article (CrossRef Link).

[27] W. A. Yasnoff, J. K. Mui and J. W. Bacus, "Error measures for scene segmentation," *Pattern Recognition*, vol. 9, no. 4, pp. 217–231, April 1977. Article (CrossRef Link).

[28] T. Fawcelt, "An Introduction to ROC Analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, June 2006. Article (CrossRef Link).

[29] J. L. Fleiss, J. Cohen and B. S. Everitt, "Large sample standard errors of kappa and weighted kappa," *Psychological Bulletin*, vol. 72, no. 5, pp. 323–327, May 1969. Article (CrossRef Link).

Weixia Liu received the B.S. degree in Applied Mathematics from Zhengzhou University, Zheng-zhou, China, in 2007. She received the M.S. degree in Applied Mathematics from Xi'an Jiaotong University, Xi-an, China, in 2010. She is currently a lecturer in Stratis Institute of Minjiang University, Fuzhou, China. She has published several papers in international/national journals. Her current research interests inclue image processing, pattern recognition and statistical analysis.

Jinmei Hu received the B.S degree In Industrial Analysis from East China University of Science and Technology (ECUST), Shanghai, China, in 1997. She received the M.S. degree in Computer Science and Technology from Fuzhou University, Fuzhou, China, in 2004. She is currently an associate professor in Department of Computer Science of Minjiang University, Fuzhou, China. She has published several papers in international/national journals. Her current research interests include image processing and data mining.

Zuoyong Li received the B.S. and M.S. degrees in Computer Science and Technology from Fuzhou University, Fuzhou, China, in 2002 and 2006, respectively. He received the Ph.D. degree from the School of Computer Science and Technology at Nanjing University of Science and Technology, Nanjing (NUST), China, in 2010. He is currently a professor in Department of Computer Science of Minjiang University, Fuzhou, China. He has published over 50 papers in international/national journals. His current research interests include image processing, pattern recognition and machine learning.

Zuchang Zhang received the B.S. degree in Automatic Control from Harbin University of Science and Technology, Harbin, China, in 2002. He is currently a lecturer in Department of Computer Science of Minjiang University, Fuzhou, China. He has published several papers in international/national journals. His current research interests include image processing, control theory and control engineering.

Zhongli Ma received the B.S., M.S. and Ph.D. degrees in College of Automation from Harbin Engineering University, Harbin, China, in 1998, 2003 and 2006, respectively. She received the post-doctoral from the College of Shipbuilding Engineering of Harbin Engineering University, Harbin, China, in 2015. She is currently a Professor in College of Automation from Harbin Engineering University, Harbin, China. She has published over 40 papers in international/national journals. Her current research interests include image processing, pattern recognition and machine learning.
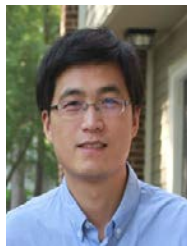
Daoqiang Zhang received the B.Sc. and Ph.D. degrees in Computer Science from Nanjing University of Aeronautics and Astronautics, China, in 1999 and 2004, respectively. He joined the Department of Computer Science and Engineering of Nanjing University of Aeronautics and Astronautics in 2004, and is a professor at present. His research interests include machine learning, pattern recognition, data mining, and medical image analysis. In these areas he has published over 150 technical papers in referred journals or conference proceedings, including reputable international journals such as IEEE Trans. Pattern Analysis and Machine Intelligence, IEEE Trans. Medical Imaging, Neuroimage, Human Brain Mapping, Medical Image Analysis, etc., and top-tier international conferences such as NIPS, IJCAI, AAAI, MICCAI, CVPR, etc., with 7000+ citations by Google Scholar. He served as a program committee member for several international and native conferences.