Expert Systems with Applications 42 (2015) 8027–8038

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

# Robust tongue segmentation by fusing region-based and edge-based approaches

Kebin Wu [a], David Zhang [b,*]

[a] Electronic Engineering Department/Graduate School at Shenzhen, Tsinghua University, Beijing 100084, China
[b] Biometrics Research Centre and the Department of Computing, The Hong Kong Polytechnic University, Kowloon, Hong Kong

ARTICLE INFO

*Article history:*
Available online 25 June 2015

*Keywords:*
Tongue segmentation
ROI
Region-based
Edge-based
Fusion

ABSTRACT

A tongue diagnosis system can offer significant information for health condition. To ensure the feasibility and reliability of tongue diagnosis, a robust and accurate tongue segmentation method is a prerequisite. However, both of the common segmentation methods (edge-based or region-based) have respective limitations so that satisfactory results especially for medical use are often out of reach. In this paper, we proposed a robust tongue segmentation method by fusing region-based and edge-based approaches. Before segmentation, ROI (region of interest), which will be used as input for the subsequent segmentation, was extracted by a novel way. Next, we merged adjacent regions utilizing the histogram-based color similarity criterion to get a rough tongue contour. It is essentially a region-based method and hence the results are less sensitive to cracks and fissures on surface of the tongue. Then, we adopted a fast marching method to connect four detected reliable points together to get a close curve, which is based on edge features. Contour obtained by region-based approach was utilized to act as a mask during fast marching process (edge-based) and the mask added limits so that the ultimate contour will be more robust. Qualitative and quantitative comparisons show that the proposed method is superior to the other methods for the segmentation of tongue body in terms of robustness and accuracy.



## 1. Introduction

Tongue diagnosis is an indispensable aspect in the Tradition Chinese Medicine (TCM) and has been attracting extensive attention worldwide due to its no-pain attribute and convenience. However, traditional tongue diagnosis depends highly on clinicians' experience and thus different clinicians are likely to reach remarkably different diagnostic results for the same patient. Fortunately, with the development of computer and other corresponding techniques, computerized tongue diagnosis is adopted to this field to overcome this weakness. For instance, Zhang, Wang, Zhang, Pang, and Huang (2005) presented a computer aided tongue diagnosis system (CATDS) to establish the relationship between quantitative features and diseases via Bayesian networks. Chiu (2000) built a computerized tongue examination system (CTES) based on chromatic and textural algorithm. A typical complete tongue diagnosis system is composed of four ordered parts: image acquisition, tongue body segmentation, feature extraction, and diagnosis. When acquiring tongue image, it is inevitable that some non-tongue elements, such as the face, teeth, lips, neck, and others in the background are captured (Pang, Zhang, & Wang, 2005). However, feature extraction and diagnosis are carried out purely on the tongue body. Hence, segmentation is needed in the tongue diagnosis system to segment tongue body from the background. In fact, in medical image analysis, medical image segmentation serves as the same role. Familiar examples include the MRI (Magnetic resonance), CT (Computed topography), and the breast segmentation (Neeraj & Lalit, 2010). From this perspective, tongue segmentation belongs to medical image segmentation and its significance is distinct. It should be noted that precise tongue body segmentation is the premise and basis of correct diagnosis results since accurate feature extraction and diagnosis rely highly on the segmentation result. This close relationship is because pathology information may rest in anywhere of the tongue body. For instance, teeth marks, which lie at the edge of tongue, provide a wealth of disease related information (Maciocia, 1987, chap. 6). Therefore, if the tongue body cannot be segmented with high accuracy, either missing some portion of tongue or including background part, tongue diagnosis using this segmentation result will be misleading. Hence, high accuracy is a great demand in segmentation. Additionally, the demand for robustness is also established due to its background of medical application. A useful

* Corresponding author. Tel.: +852 2766 7271; fax: +852 2774 0842.
*E-mail addresses:* wkb11@mails.tsinghua.edu.cn (K. Wu), csdzhang@comp.polyu.edu.hk (D. Zhang).

http://dx.doi.org/10.1016/j.eswa.2015.06.032
0957-4174/

segmentation approach must be applicable to all subjects, especially to all kinds of patients. Thus, to be applicable in practice, accuracy and robustness are the basic requirements for any tongue segmentation approach. These two criterions are then adopted to evaluate any tongue segmentation approach.

Since an acquired image is a color picture (color abnormality can reveal pathology), many color segmentation methods have been applied to segment tongue body. Generally, these approaches can be divided into two classes: region-based approaches and edge-based ones. For the former, Ning, Zhang, Wu, and Yue (2012) segmented tongue images into many small regions using watershed algorithm and then color-similarity based region merging was executed. Similar attempts could also be found in another literature (Wu, Zhang, & Bai, 2006). For edge-based approaches, studies forwarded recently mainly focus on active contour model (ACM). According to different initialization methods and various curve evolution solutions, there exist many variants. With respect to the initialization, Pang et al. (2005) initialized the evolving contour with the bi-elliptical deformable template. However, Zuo, Wang, Zhang, and Zhang (2004) proposed to obtain initial curve by selecting segments from edge map. Comparing with those focusing on geometrical information, many researchers have attempted to utilize color feature of tongue, especially hue-channel value, to obtain initial contour (Shi, Li, & Li, 2013; Yu, Yang, Wang, & Zhang, 2007; Zhai, Lu, & Zhang, 2009). With regard to the curve evolution, Gradient Vector Flow (GVF) (Yu et al., 2007), Double Snake (Zhai et al., 2009), C2G2FSnake (Shi et al., 2013) and Double Geodesic Flow (Shi, Li, Li, & Xu, 2012) were adopted. However, two key difficulties frustrate these evolutions. Firstly, the initial contours must be close to the true boundaries. The other is that active contours have difficulties progressing into boundary concavities. Li, Yao, Yuan, and Zhou (2010) utilized level set method for curve evolution to address these two problems.

For these methods, each of them has had its fair share of success, but corresponding limitations also accompany with them. Taken as a whole, they fail to satisfy the demands for both accuracy and robustness simultaneously, which are the basic requirements for a successful segmentation as mentioned in the first paragraph. Specifically, region-based image segmentation techniques can provide closed region boundaries and are robust to random noise due to its statistics basis. Nevertheless, the boundaries may be not very accurate, especially when it comes to the edges between lips (face) and tongue due to their similar color (Ning et al., 2012). Similarly, despite of the fact that edge-based methods can locate part of edges exactly, the detected boundaries are likely to be open curves. In addition, some post-procedures, such as edge tracking, should be performed to obtain closed region boundaries. Furthermore, edge-based methods are sensitive to noise and often result in spurious edges. Thus, robustness cannot be guaranteed.

In tongue segmentation application, shortcomings for these two classes of methods will aggravate greatly further owning to the specific characteristics of tongue. Specifically speaking, the difficulties fall broadly under two considerations. Firstly, the difficulties of tongue segmentation arise when the image comes from a patient since tooth marks, evident fissure or abnormal color often company with pathologies (Fig. 1(b) and (d)). Secondly, more common and inherent problems exist. One usual difficult comes from interference by lips and face due to their resemblance with tongue in the aspect of color (Fig. 1(a) and (c)). Abnormal tongue shape in some occasions also challenge many segmentation approaches especially those model-based ones. Besides, image of male with beards may meet difficulty (Fig. 1(e)) and wrinkles in the lower jaw also bring obstacle (Fig. 1(f)).

Based on the abovementioned analysis, it can be summarized that segmentation by a single use of either edge-based or region-based method, can hardly achieve accuracy and robustness simultaneously due to the inherent characteristic of tongue image and the respective intrinsic limitations of these two categories of methods. The aim of this work is to develop a method to segment tongue body accurately and robustly. Motivated by the respective properties of edge-based and region-based methods, we proposed that a combination of these two categories methods often performs better and the adverse influence for each method may decline. Thus, naturally we attempted to combine region-based method with edge-based together to get more accurate and more robust results. We began by a ROI extraction that relies highly on the use of color information and the original image was replaced by the ROI for subsequent segmentation. Then an improvement was made to an existing region-based method MSRM (Maximal Similarity based Region Merging). The tongue contour obtained from the improved MSRM was then combined with an enhanced edge map to generate a weight image. Finally, the edge-based method fast marching worked on this weight map to compute the final tongue contour. Here is how the fusion works: the result in region-based method was utilized to generate a mask, which is necessary when acquiring the weight map in edge-based method. The proposed approach can effectively achieve the requirement of accuracy and robustness, whose performance is validated by qualitative and quantitative comparison.

The remainder of the paper is organized as follows. Section 2 describes techniques used to extract ROI (region of interest) so as to avoid evitable interference and reduce computational complexity simultaneously. In Section 3, we give an introduction about the proposed methods, not only the adopted region-based and edge-based approaches, but also the scheme of combination. To assess the performance, we present the qualitative and quantitative comparison results in Section 4. Finally, Section 5 concludes the paper.

## 2. Preprocessing: extract ROI to enhance robustness

For any tongue segmentation routine, not only tongue body itself but also complex background part will involve in the computation. Generally, the abovementioned background consists of face, beard, lip, teeth, and others in the capture environments, which are collectively called non-tongue components in this paper. Fig. 2 illustrates this observation. Unexpectedly, in most of the acquired images, the non-tongue parts occupy more space than the tongue body, which will possibly bring interference to segmentation accuracy and robustness. Besides, the non-tongue portion will consume much higher computational complexity than the tongue body in most cases. In summary, segmentation directly from the original image will result in a large waste of time and degrading the segmentation performance. Intuitively, if a simple but effective preprocessing can get rid of majority of non-tongue components while maintaining tongue body unchanged, then this preprocessing will benefit the subsequent finer segmentation process in aspect of efficiency and accuracy. In nature, this preprocessing is to extract the region of interest (ROI). To make it more clear, ROI in this paper refers to a rectangular area, whose main content is tongue body. For this purpose, we proposed an efficient ROI extraction approach.

Tongue body differs from the non-tongue portion (with lips as an exception) in the aspect of color. Lips, however, are in the vicinity of tongue body such that lips are allowed to be included in the rough ROI. Thus utilizing color information will be well suited in the extraction. To highlight the difference between tongue and non-tongue, *de-correlation and stretch* algorithms to the original image (denoted as $I$) in RGB format were applied. The *de-correlation* method was to reduce cross-correlation among three channels and *stretch* approach followed to enhance contrast