Automatic Tongue Diagnosis System

[1]Lun-chien Lo, [1]Mark Chun-cheng Hou, [1]Ying-ling Chen, [2,+]John Y. Chiang, [2]Cheng-chun Hsu

[1]Department of Traditional Chinese Medicine
Changhua Christian Hospital
Changhua, Taiwan

[2]Department of Computer Science Engineering
National Sun Yat-sen University
Kaohsiung, Taiwan
[+]Email:chiang@cse.nsysu.edu.tw

***Abstract*—Tongue diagnosis tops the four diagnoses and hence is crucial to the inspection diagnosis in Chinese Medicine. Clinically, doctors mostly rely on their own knowledge and experience when determining major lesions of a patient by observing the coloration, overall modalities, and volume of sputum of different parts of the tongue. Different doctors may come to drastically different judgments on the same tongue presentation with little overlap. Therefore, it is important to develop scientific methods that can help doctors diagnose based on standardized differentiation procedures and render reliable diagnoses in order to enhance the clinical application value of Chinese Medicine. The computerized automatic capture of characteristics shown on the images of the surface and the back of the tongue comprises breakup of the tongue image and capture of the tongue characteristics. And that, tongue characteristics include the shape, fur, body, and fluid. Sublingual collateral vessels are used to examine traits of veins on the back of the tongue and detect the length, width, number of branches, solidness, and color. By far, 29 images of the surface of the tongue and 69 images of the back of the tongue have been collected.**

Keywords: Automatic tongue diagnosis, inspection diagnosis in Chinese Medicine, sublingual collateral vessels.

I. INTRODUCTION

In Chinese Medicine, it is believed that a person's tongue has much to do with his/her internal organs. Traditionally, there are three ways to relate channels and network vessels to internal organs, namely the stomach channel, triple burner, and bowels and viscera (Fig. 1).

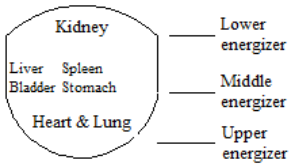


Figure 1. Breakup of the Tongue and Corresponding Body Parts

In tongue diagnosis, the tongue is observed to differentiate the syndrome. Observation of the surface of the tongue is mainly focused on the shape, fur, and body of the tongue. The shape of the tongue can be medium, fat, thin, and tilted. In term of the tongue body, it can be pale, light-colored, light red, red, crimson, or dark-colored with bruise dots, fissures, tooth marks, or red dots. The fur of the tongue can be white, yellow, black, slimy, thick, thin, peeling or none. Body fluid can be normal, insufficient or lack. Observation of the images of the back of the tongue is focused on the sublingual collateral vessels, including the length, width, number of branches, solidness, and color of the back of the tongue. By observing these properties, the doctor can know the patient's qi and blood circulation and whether organs are functioning normally.

Automatic tongue diagnosis system accordingly becomes an inevitable trend. Separation of the tongue section from images of the lower face including the mouth and teeth is the preliminary step in the automation of tongue diagnosis. In the breakup of the tongue section, HSI color space conversion and morphology are combined to break up the image of the surface of the tongue [1,2,3]. Polar coordinates conversion and marginal filtering [4] are combined to detect the six edges and directions and center of the surface of the tongue in the images. The watershed transformation is used together with the active contour model to search the edges of the tongue [5,6]. The use of the gradient of a multi-image search for the active contour [7,8,9]. In this thesis, detect the tongue by color calibration, detection of the control points and active contour model. Based on the image of the surface of the tongue captured, the shape, fur, and body of the tongue, and body fluid among other characteristics of the tongue can be obtained as well. Separation of the image of the back of the tongue, on the other hand, is completed by means of the color spectrum difference between the skin and the back of the tongue, the enhanced gray-scale contrast, and comparison of thresholds. Based on the captured image of the back of the tongue, the length, width, number of branches, solidness, and color of the sublingual collateral vessels are detected with the process. Properties of the surface of the tongue and sublingual collateral vessels can be detected and identified based on the morphological changes and color spectrum differences.

II. CAPTURE OF THE TONGUE SECTION AND ITS TRAITS

A. Color Calibration

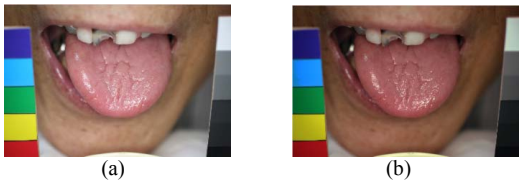
In order for the tongue diagnosis images not to be affected by variants like the environment and equipment when the images of the surface or back of the tongue are being processed and to maintain the persistence in the presentation of color, color correction is first implemented. For the six colors on the right and left sides of tongue diagnosis images (blue, light blue, green, yellow, red, and pink), the standard values measured in the standard environmental setting are being corrected and the compensation formula is as follows (1).

The corresponding standard colors for value R of color k on the color card are $R_{std}[k]$. $G_{std}[k]$, $B_{std}[k]$ color component the same for other colors . The six colors divide the color space into seven subsections. Suppose that R_i is the original pixel

This research is supported by the Committee on Chinese Medicine and Pharmacy, Department of Health, Executive Yuan, Taiwan. Project No. CCMP98-RD-031.

value entered and it is between subsections $R_{origin}[k]$ and $R_{origin}[k+1]$, when corresponding formulas (1) is used in the calculation, R_o is the output result and G_o, B_o so on. Fig. 2a is the original image and b is the corrected result.

$$R_o = (R_i - R_{origin}[k]) \times \left(\frac{R_{std}[k+1] - R_{std}[k]}{R_{origin}[k+1] - R_{origin}[k]}\right) + R_i \quad (1)$$



(a) (b)

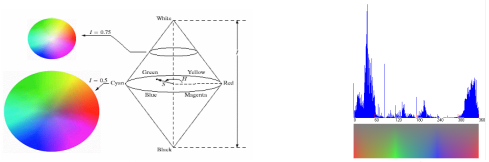
Figure 2. (a) is the original image and (b) is the corrected result.

Fig. 3 shows the variation curve of Values R, G, and B before and after correction. On the left is the standard color value. In the middle is the color card value of the originally captured image. On the right is the corrected result and the correction curve is shown in the small graphics on the right.

Figure 3. Corrected curve of Values R, G, and B.

B. *Skin Hue*

First, the RGB color space is converted into an HSI color space in order to meet human eye perception. HSI color space can be shown as a cone space model in Fig. 4a.



(a) (b)

Figure 4. (a) HSI model and (b) typical tongue diagnosis images hue statistics

After correction, images in Fig. 2b are shown in Fig. 4b: Hue Statistics. The longitudinal axis represents the number of hues while the horizontal axis represents the hue value. Due to the fact that most of the skin hue statistics are between $15°$ and $60°$ and the hue values of the tongue and the lips are close to $0°$ and $360°$, this property is used to separate the tongue section from the skin section. First, in the skin hue, it is assumed that the hue value of the maximum pixel H exists. Let $h1$ and $h2$ be the hue value and $h1$ is initialized as H -1 and $h2$ as H +1, then the range between $h1$ and $h2$ is the removal limit of the skin section. Suppose that $Hue(x)$ is the statistic value of Hue x, $h2$ stays the same and $h1$ is reduced once when $Hue(h1) > Hue(h2)$ and $h1$ stays the same and $h2$ is increased once when $Hue(h1) < Hue(h2)$. Under the conditions mentioned above, the numbers of pixels from Hue $h1$ to Hue $h2$ are accumulated until the skin occupies around half of the area of the whole image as is shown in the following formula.

$$\sum_{x=h1}^{h2} Hue(x) > \frac{1}{2}\sum_{x=0}^{360} Hue(x). \quad (2)$$

Fig. 5a is the result of separating the skin. The color cards on the right and left sides can further be removed through rectangular detection as is shown in Fig. 5b.

(a) (b)

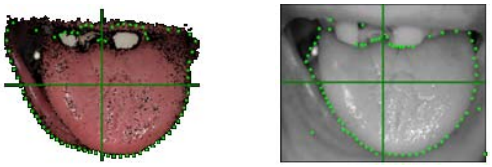
Figure 5. (a) Image of the tongue with the skin removed and (b) image of the tongue with the color card removed.

C. *Detection of the internal and external control points of the surface of the tongue*

Suppose that the image with the color card removed (Fig. 5b) is I_{erase} and the center coordinate is (x_0, y_0), then the pixel coordinate can be transformed to a polar coordinate(3) .

$$I_{erase}(x,y) = I_{org}(r\cos\theta + x_0, r\sin\theta + y_0). \quad (3)$$

Detection of the tongue surface section is performed through mutual approximation inside and outside the surface of the tongue. The borderline between the external section and the oral cavity can be ascertained by removing the skin section. By means of the edge of the image with the color card removed (Fig. 5b), the center of the surface of the tongue can be detected and through polar coordinate conversion [6], the borderline between the tongue and the skin can be searched from the outside to the inside. Fig. 6a shows the detection results for external control points of the surface of the tongue.



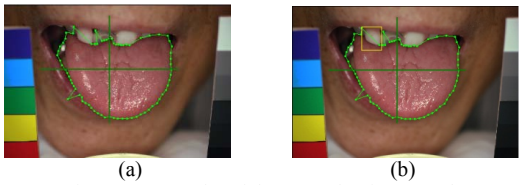
(a) (b)

Figure 6. shows the detection results for the external (a) and internal (b) control points of the surface of the tongue.

The brightness of pixels at the intersection between the internal section of the surface of the tongue and the oral cavity and lower lip and it is used as the basis for separation. Supposed that the gray scale image with the color card removed (Fig. 5b) is I_{gray}, control points are searched for from the inside to the outside. The searched pixels also refer to their adjacent pixels as a small group. When the brightness of the group is lower than 30 on average, it is stopped. If not, the smallest brightness average value is being searched for. Fig. 6b shows the detection results for internal control points of the surface of the tongue.

From the candidate control points inside and outside the surface of the tongue, the closest is chosen as the selected control point based on its distance from the center of the image

of the surface of the tongue as is shown in Fig. 7a. The upper part of the tongue may have been affected by the tongue fur (yellowish, closer to the skin complexion), which results in the control point for skin detection falling on the inner side and wrong capture. The angle of the polar coordinate in the upper section of the surface of the tongue is located between 275 and 315. Control points within this scope are replaced by the candidate points inside the surface of the tongue.



(a) (b)

Figure 7. Results of the control points search

D. Control point smoothness correction

After capturing the control points, influence from foreign signals will cause some protrusion as is shown in Fig. 7a. Control points with a distance from two adjacent control points greater than the average value plus the variant are considered control points with excessive shifts as a result of foreign signals. Let the average distance between control points be M and the variant be V and Search for control points satisfying the conditions of $M+V>d(c_i,c_{i+1})$ and $M+V>d(c_{i-1},c_i),\quad d(c_i,c_{i+1})$, is the distance from Control Point i to Control Point $i+1$. The interpolation method can be applied for smoothness correction.

Corresponding unevenly distributed control points to Fig. 7b represent an overly long distance between two control points as is shown that $d(C_j,C_{j+1})>d(c_k,c_{k+1})+d(C_{k-1},C_k)$, $d(c_j,c_{j+1})$ is the largest distance among all control points from Control Point j to Control Point $j+1$. It is greater than the distance between Control Point k and its two adjacent control points. The correction is made with the interpolation method which generates the center point as the new control point $v(x_n,y_n)$ to replace Control Point k.

Figure 8. Result of the corrected curve of the smooth surface of the tongue.

After smoothness correction with the two methods mentioned in the foregoing, the control points on the edge of the smooth surface of the tongue are shown Fig. 8.

E. Correction of the curve with the active contour model and results

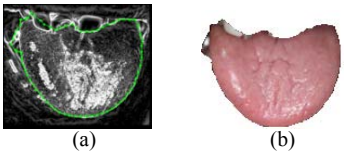
The surface of the tongue appears to be a smooth arc in shape. The borderline between its interior and exterior corresponds to a high-frequency edge. This characteristic is qualified for application of the active contour model. The control points detected in the foregoing are used as the initial points for the active contour and to search for the high-frequency section in order to capture the tongue surface.

The fast active contour is a mechanical visual approach to select the edge of an item [10]. The equation is shown as follows:

$$E^*_{snake}=\int_0^1 (E_{int}(v(s))+(E_{image}v(s)))ds. \tag{4}$$

It can be seen as a simplified version of the conventional active contour model[11]. When it selects an item, it only considers pixels in the proximity of the control point and uses the gradient image to search for the high-frequency area when control points are evenly distributed. Changes of the internal and external forces are considered to minimize the total power so that the curve can eventually restrain to the edge of the item.

The image with the color card removed (Fig. 5b) goes through the gray-scale conversion, edge detection, and Gaussian smoothing to obtain the gradient image as is shown in Fig. 9a.Based on the model of the surface of the tongue, the active contour parameters can be modified to obtain a better contour of the surface of the tongue. In light of facts that control points need to be evenly distributed, that the tongue needs to be a smooth arc curve, and that high frequency is the basis for the search of the edge, the internal force parameter α is accordingly set at 0.3 and the external force parameter β is 0.5 while the image force parameter γ is 0.7. It is applicable to a majority of tongue diagnosis images. Fig. 9b is the result of the capture.



(a) (b)

Figure 9. (a) is the gradient image and (b) is the captured image of the surface of the tongue.

F. Characteristics detection

1) Elimination of reflection.

The fluid on the surface of the tongue tends to reflect light and is accordingly the strongly light reflective local area, leading to loss of information on the characteristics of the surface of the tongue below the fluid. To compensate for the loss of authenticity as a result of light reflection, pixels in the surroundings of the light reflective area are referenced when making modifications. Let the brightness value of the already separated HSI color space tongue image be $I_{org}(x_i,y_i)$; the expected value of the image brightness be $I_{\exp}$, and the standard deviation be σ, the light reflective area can then be expressed as $I_{org}(x_i,y_i)>I_{\exp}+\sigma$. Correction of the light-reflective areas is as follows:

$$C_{reflect}(x_i,y_i)=C_{left}(x_a,y_a)\times L_{ratio}+C_{right}(x_b,y_b)\times R_{ratio}+$$
$$C_{up}(x_c,y_c)\times U_{ratio}+C_{down}(x_d,y_d)\times D_{ratio}$$

$$L_{ratio} : R_{ratio} : U_{ratio} : D_{ratio} = \frac{1}{D_{left}(x_a, y_a)} : \frac{1}{D_{right}(x_b, y_b)} : \frac{1}{D_{up}(x_c, y_c)} : \frac{1}{D_{down}(x_d, y_d)}. \quad (5)$$

$C_{reflect}(x_i, y_i)$ represents the light reflective pixels that must be corrected. $C_{left}(x_a, y_a)$ is the non-light-reflective pixel and is closer to the left of $C_{reflect}(x_i, y_i)$. Based on this, $C_{right}(x_b, y_b)$, $C_{up}(x_c, y_c)$,and $C_{down}(x_d, y_d)$, are inferred. $D_{left}(x_a, y_a)$, $D_{right}(x_b, y_b)$, $D_{left}(x_a, y_a)$, and $D_{right}(x_b, y_b)$, are the distances to the four non-light-reflective points adjacent to the light-reflective point $C_{reflect}(x_i, y_i)$. Before searching for each characteristic, it is necessary t perform preliminary removal of the reflection to avoid influence by the light-reflective area and dissatisfactory captures of characteristics.

2) Fissures on the surface of the tongue.

Fissures on the surface of the tongue correspond to areas on the main part of the tongue surface with less bright pixels with irregular lines and shapes. With the reflection on the captured image of the surface of the tongue, it is converted to the gray scale I_{gray} and from top down, the average value M_h and variant V_h for each horizontal line as well as the average value M_v and variant V_v for each vertical line are calculated to meet the pixel value $P(x, y)$ in the $I_{line}(x_i, y_i) < M_h - V_h$ or $I_{line}(x_i, y_i) < M_v - V_v$, and it is the fissure section.

$I_{line}(x_i, y_i)$ is the area of the dark section of the tongue diagnosis image. By means of the connected component analysis, the section with a too small fissure area is removed. The diagonal of the rectangular section with fissures is the fissure length. Fig. 10a shows the fissure search result.

(a) (b)

Figure 10. (a) is fissure and (b) is bruise dot result.

3) Bruise dots on the surface of the tongue.

Bruise dots in contrast to the main part of the surface of the tongue appear to be purple black and gray black pumps or small lots. In HSI color space, the hue, saturation, and brightness of bruise dots all go through obvious changes when compared with other parts of the surface of the tongue. The hue of bruised dots is different from that of the tongue body with higher saturation and lower brightness. This is used as the basis for filtration. Preliminary filtered results are then processed by means of mathematical morphology to eliminate foreign signals. After that, the connected component analysis is used to filter those with smaller areas to obtain the bruise dots on the surface of the tongue as is shown in Fig. 10b.

4) Tooth marks on the surface of the tongue. .

Tooth marks on the surface of the tongue result from a overly big tongue body squeezed by adjacent teeth. They appear like dark wavy lines on the tongue body. By RGB color composition, tooth marks are dark red or crimson. Their G color spectrum is lower than that of ordinary tongue body and tongue fur. This is accordingly used as a tool to detect location of tooth marks. Meanwhile, in light of the fact that tooth marks will not appear inside the oral cavity, the upper part of the tongue can be removed first. The areas with possible tooth are kept and areas that are not edges of the tongue are removed. Fig. 11a shows the result of tooth mark detection.

(a) (b)

Figure 11. (a) is the tooth mark and (b) is red dot detection result.

5) Red dots on the surface of the tongue.

Red dots in contrast to the main part of the surface of the tongue appear to be bright red pumps or small lots. In RGB color space, the G value of red dots is smaller than that of their adjacent tongue body with higher saturation than that of the tongue body and lower brightness. After preliminary locations of red dots are detected, those with an overly big area are removed. Fig. 11b shows the detection result.

6) Separation from tongue fur from tongue body.

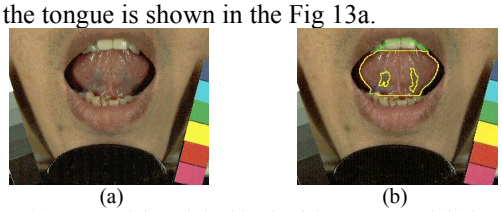
Tongue fur is the furry substance that covers the tongue body. It can be white, yellow, black or slimy, thick, thin, peeling, and none. Based on the HSI color model, the hue difference of tongue fur among white, yellow, and black, and the fact that the saturation and brightness of tongue fur are higher than that of the tongue body are used as the basis to separate tongue fur from the tongue body. On the obtained tongue section, the distance between tongue fur pixel points is measured and it is determined whether the tongue fur is slimy, thick, thin, peeling or there is no fur. The following figure shows the tongue fur detection result.

(a) (b)

Figure 12. (a) is the tongue fur and (b) is the separated tongue body.

G. Detection of sublingual collateral vessels

A typical image of the sublingual collateral vessels on the back of the tongue is shown in the Fig 13a.



(a) (b)

Figure 13. (a) Image of the original back of the tongue and (b) image of the curve of the back of the tongue.

The difference between the image of the back of the tongue and image of the skin lies in the distribution of pixel component G. Therefore, Formula (6) can be used to enhance the contrast of the image of the back of the tongue.

$$x = \frac{R-G}{|G-B|+1} \quad if\ R > G, R > B \tag{6}$$

By means of adjustment of the coefficients, the contrast is further extended.

$$\frac{V}{m-0} = \frac{x-VLow}{X-0} \quad if\ x \le M \ , \ \frac{V}{255-m} = \frac{x-VHigh}{X-m} \quad if\ x > M \tag{7}$$

V is the variant of pixel value X and M is the average of pixel value X. *VLow* is the lower limit of M - V while *Vhigh* is the upper limit of $M+V$. m is the adjusted coefficient. The different average M is used to adjust Coefficient m to obtain proper enhancement effect as is shown in the following figure. Based on the adjusted coefficient *m*, the bi-level threshold is set and the connected section with a maximum area is searched for. After that, the image of the back of the tongue is obtained as is shown in Fig. 13b.

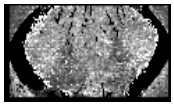


Figure 14. Images of enhanced contrasts.

1) Detection of sublingual collateral vessels.

Sublingual collateral vessels are areas that appear to be purple black and blue back closer to the center of the tongue on the back of the tongue. In RGB color space, the R value of sublingual collateral vessels is lower than that of the tongue body but the G and B values are higher. This can serve as the basis in the detection and filtration. Detected results are binary images. The connected component analysis is used to detect those with the largest and secondary areas. Results are shown in 16b.

III. EXPERIMENT RESULTS

By far, ideal images of the surface and the back of the tongue have been collected, based on which fissures, tooth marks, bruise dots, red dots, tongue fur, tongue body, length of sublingual collateral vessels and fullness are determined to facilitate subsequent interpretations. Characteristics of the tongue captured by the computer are similar to those obtained by a group of tongue diagnosis experts. By far, images with fissures that are needed to be improved account for around 20%; bruise dots 20%; tooth marks 20%; red dots 10% and sublingual collateral vessels 15%. Unsatisfactory results of tongue diagnosis images interpretation are mostly attributable to light reflection, insufficient lighting, among other environmental factor that result in loss of authenticity. Currently, efforts have been made to improve light variants in order to enhance the accuracy of the tongue diagnosis identification system. Others like captures of the tongue shape and tongue fluid, among others, are ongoing. Hopefully, the differentiation function of the tongue diagnosis system can be greatly enhanced in the future.

IV. CONCLUSION

In this thesis, HSI color space conversion, polar coordinate conversion, the active contour model, among other techniques are used to capture the contour of the surface of the tongue. Based on the color difference between the back of the tongue and the skin, contrast enhancement, threshold separation, and connected component analysis are used in the separation effort. The captured images of the surface and back of the tongue are analyzed based on the color composition of different characteristics in order to detect the properties of the surface of the tongue and sublingual collateral vessels. The results basically meet the interpretation by experts. In the future, efforts will be focused on enhanced identification of other characteristics of the surface of the tongue in order to offer steady interpretation results.

V. REFERENCES

[1] Zhao Zhong-xu, Wang Ai-min, Shen Lan-sun. "Color Tongue Image Segmentation Based on Mathematical Morphology and HSI Model," *Journal of Beijing Industry University*, 25(2), pp. 67-71, 1999.

[2] Du Jian-qiang, Lu Yan-sheng, Zhu Ming-feng, Zhang Kang and Ding Cheng-hua, "A Novel Algorithm of Color Tongue Image Segmentation Based on HSI," *Proc. of BMEI*, VOL. 1, pp.733-737, May 2008.

[3] Soo-Chang Pei, "Image sampling structure conversion by morphological filters," *Signal Processing: Image Communication*, NO. 1, pp. 13-24, March 1994.

[4] Wangmeng Zuo, Kuanquan Wang, Zhang, D. and Hongshi Zhang, "Combination of polar edge detection and active contour model for automated tongue segmentation," *The 3rd Inter. Conf. on Image and Graphics*, pp.270-273, 18-20 Dec. 2004.

[5] Jia Wu, Yonghong Zhang and Jing Bai, "Tongue Area Extraction in Tongue Diagnosis of Traditional Chinese Medicine," *The 27th Annual Inter. Conf. the Engineering in Medicine and Biology Society*, pp.4955-4957, 17-18 Jan. 2006.

[6] Vincent, Luc, and Pierre Soille, "Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations," *IEEE Tran. of PAMI*, VOL. 13, NO. 6, pp. 583-598, June 1991.

[7] Shengyang Yu, Jie Yang, Yonggang Wang and Yan Zhang, "Color Active Contour Models Based Tongue Segmentation in Traditional Chinese Medicine," *The 1st Inter. Conf. on ICBBE*, pp.1065-1068, 6-8 July 2007.

[8] G. Sapiro, Dario L. Ringach. "Anisotropic diffusion on multivalued images with application to color filtering," *IEEE Trans. on Image Proc.*, VOL. 5, NO. 11, pp. 1582-1586, March 1996.

[9] Chenyang Xu and Prince J.L, "Snakes, shapes, and gradient vector flow," *IEEE Trans. on Image Proc.*, VOL. 7, NO. 3, pp.359-369, March 1998.

[10] Williams, D. J. and Shah, M. "A fast algorithm for active contours," *Proc. 3rd Inter. Conf. on Computer Vision 1990*, pp. 592-595, 4-7 Dec. 1990.

[11] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Inter. Journal of Computer Vision*, VOL. 1, NO. 4, pp. 321-331, 1987.