Original article

Improving color and shape repeatability of tongue images for diagnosis by using feedback gridlines

Chang Jin Jung a,1, Keun Ho Kim a,*,1, Young Ju Jeon a, Jinsung Kim b

a Medical Engineering R&D Group, Korea Institute of Oriental Medicine, Daejeon, South Korea
b College of Medicine, Kyung Hee University, Seoul, South Korea

Received 28 June 2013; received in revised form 15 January 2014; accepted 15 January 2014

Abstract

Introduction: In traditional East Asian medicine, the color and shape features of a tongue reflect the health of the body. The repeatability of computerized tongue diagnosis is highly dependent on the image acquisition process. In this study, a method for providing frontal and profile feedback gridlines for the tongue image acquisition process was proposed and the resulting color and shape repeatability was assessed.

Methods: Eight subjects were recruited and 120 images were acquired. For each subject, 15 tongue images were acquired with three different gridline types, including no gridline (Group I), frontal gridlines (Group II), and frontal and profile gridlines (Group III) in a pseudo-randomized order. To evaluate repeatability, intra-class correlation coefficients (ICCs) were calculated from the mean values and histogram fitting parameters for Commission Internationale de l’Eclairage L*a*b*, profile angles, and tongue area ratios (TARs) of four edge regions, respectively.

Results: The ICCs for all the color features in Group I were lower as compared with those in Group II and additionally, the ICCs for all the color features in Group III were larger compared with those in Group II. The ICCs of the shape features were lower than those of the color features. The ICCs of TARs in Group I were lower than those in Groups II and III. From the tip to the root of the tongue region, the ICCs of mean luminance tended to decrease in all the groups, but the intensity of the decline in Group III was lower than those in Group I and Group II.

Conclusions: The proposed tongue image acquisition process with frontal and profile gridlines improved the repeatability of color and shape features and potentially improve diagnostic accuracy of computerized tongue diagnosis.

© 2014 Elsevier GmbH. All rights reserved.

Keywords: Color and shape repeatability; Feedback gridlines; Tongue image; Tongue diagnosis; Methodology; Measurement

Introduction

The color and shape features of a tongue reflect the physiological and clinico-pathological condition of the body in traditional East Asian medicine (TEAM) and each part of the tongue is related to corresponding internal organs [1]. Recently, various computerized tongue diagnosis systems (CTDSs) have been developed for accurate diagnosis based on quantitative and objective tongue features. In a CTDS, a patient’s tongue image is captured via a digital camera, and a tongue region is segmented from the image. The quantitative features of the tongue, such as color values in regions of interest, textural parameters, and geometric information, are calculated from the tongue image and are used for the patient diagnosis. For instance, with thickness of a tongue coating, the patient’s condition is classified into presence or absence of pathogenic qi by using the color features of the tongue image [2,3].

Furthermore, previous research has found that the features of the tongue image can be used for an estimation of specific diseases. The color distributions of the tongue in appendicitis and pancreatitis have appeared different from those in normal subjects, and a tongue deviation angle has been shown to be useful for a stroke diagnosis [4,5].

The feature values of color and shape in the tongue image, however, are highly dependent on the image acquisition process. In particular, different tongue positions can have a strong influence on the color and shape information in the tongue imaging owing to differences in the curvature of the tongue and the angle...
of light incidence. A color correction method for the tongue image was developed based on a color checker and used in the various CTDSs to correct the distortion of illuminators and cameras [6], but it is still difficult to correct the color and shape distortion caused from the tongue posture. Establishing the position of the tongue is an important condition for the repeatability of the tongue diagnosis.

Owing to this problem, several CTDSs have provided the contact area of a cheekbone or chin to fix the location of the head during the tongue image acquisition process [7,8]. However, it is still difficult to constantly control the tongue position even if the head is fixed. In this study, a method for providing frontal and profile feedback gridlines for the tongue image acquisition process was proposed and the repeatability of the color and shape in the images of the tongue was assessed.

Materials and methods

Image acquisition device for feedback gridlines

To provide frontal and profile feedback gridlines to subjects, two digital cameras and an LCD monitor were used. A three-charge-coupled device (CCD) digital camera with six-megapixels was located in front of the tongue and a Bayer digital camera was located at the right side of the tongue as shown in Fig. 1(a).

The LCD monitor was located in front of a subject’s eyes to allow the subject to observe the tongue position and prevent it from changing. In order to minimize the distortion from uncontrolled light, the images were acquired in a semi-closed box. A high-brightness white LED ring array was used as a main illuminator for frontal image acquisition and two white LEDs were used as sub-illuminators for profile image acquisition. The frontal and profile images were acquired within an interval of 20 ms and the brightness of the main illuminator was attenuated during the profile image acquisition to avoid overexposure.

The graphical user interface with gridlines consisted of a frontal preview image and a profile preview image as shown in Fig. 1(b). Because the length and the width of the tongue are different according to each subject, the gridlines consist of five horizontal and five vertical lines, which help the subject judge their tongue position. The gridlines in the frontal preview were used for adjusting the position of the tongue and the red vertical grid in the profile preview was used for adjusting the angle and the posture of the tongue as a subject stuck out their tongue.

Experimental procedures

Eight healthy subjects (mean age: 30.3 ± 3 years) volunteered to participate in this study. Before the experiment, all subjects were allowed to practice the procedures for using the gridlines.

The procedure for guiding the tongue in front consisted of three steps. In the first step, a subject was required to move the locations of the corners of their mouth in line with the top horizontal gridline in the preview image. In the second and the third step, the subject was required to open their mouth while keeping the location of their head steady and sticking out their tongue onto the bottom horizontal gridline in the preview image.

When using both frontal and profile gridlines, the subject was required to change the profile angle of their tongue or head in order to relocate their tongue between the two red vertical gridlines in the profile preview image, following the frontal gridline procedure.

For each subject, fifteen tongue images were acquired with three different gridline types including no gridline (Group I), frontal gridlines (Group II), and frontal and profile gridlines (Group III). Five images were acquired for each gridline type and all the acquisition processes were performed in a pseudo-randomized order.

Image processing

To extract features from images for statistical analysis, image processing was performed for the 120 images which had been acquired. From the frontal image, a tongue region was segmented with the combining polar edge method [9] and the gradient vector flow snake technique [10]. The polar edge points were detected by using the external energy of the snake energy field instead of color information. Following the snake processing, manual validation was performed by using a contour modification tool that provides a graphical user interface and an active contour function. The contour modification tool automatically modified the manual validation result by using the external energy of the snake algorithm, so that the manual operator cannot modify the result. The manual validation played a role in modifications of large misalignment of segmented contours.

Pixel values in the RGB color space were converted to the Commission Internationale de l’Éclairage (CIE) L*a*b* space, and high-luminance regions, which were caused by light reflection from saliva, were removed from the tongue region. The a* value from the CIE L*a*b* color space is widely used for the chroma feature of the tongue diagnosis since the a* value is appropriate for representing the red intensity [11,12].

For evaluating the color, color features were extracted from the entire tongue as well as from separate tongue regions because the color of each part of the tongue represents the state of corresponding internal organs [1]. The mean values of L* and a* were calculated for the entire tongue and for the six separate regions, as shown in Fig. 2(a). The histograms of L* and a* for the entire tongue were computed as features for the color distribution. In order to evaluate the change in the color features in the vertical direction, mean values of L* and a* were calculated from the 10 separate regions as shown in Fig. 2(b).

For evaluating the shape, each tongue area ratio (TAR), which is the ratio of the tongue area to the area of the rectangular grid (EAi), of the four regions including the edge was computed, as shown in Fig. 2(c), where i = A, B, C, and D. The TAR is the ratio of the pixel count in the tongue region to the pixel count of the corresponding region for each of the four regions. A horizontal center distance (HCD) was obtained from the difference between the horizontal center location of the tongue and the center point of the image, as shown in Fig. 2(d).
From the profile image, the profile angle of the tongue was calculated. Points at the root and the tip of the tongue were detected manually from the profile image and three points of the tongue surface contour were detected on the three horizontal lines, which were vertically divided evenly into four regions between the points of the root and tip. The tongue-sticking-out angle (TSA, $\theta$ between the line from $P_{\text{root}}$ to $P_{\text{tip}}$ and the vertical line), profile curvature angle of the root side (PCAr, $\delta$ to be upper curvature from the center point), and profile curvature angle of the tip side (PCAt, $\gamma$ to be lower curvature from the center point) were calculated by using the five points as shown in Fig. 2(e).

**Statistical analysis**

The histograms of the $L^*$ and the $a^*$ were fitted to a Weibull distribution using a maximum likelihood estimation [13] and two parameters of the distribution were computed for the color features using the following Weibull probability density function:

$$y = f(x|\eta, \beta) = \beta \eta^{-\beta} x^{\beta-1} e^{-(x/\eta)^\beta} I_{(0, \infty)}(x),$$  \hspace{1cm} (1)

where $\eta$ (eta) and $\beta$ (beta) represent the shape parameter and the scale parameter, respectively.

To evaluate the repeatability of the color and shape features, the repeatability was estimated by using the intra-class correlation coefficient (ICC) that was determined with an analysis of variance for mixed models [14]. For five images from each group corresponding to rats, the ICC was computed from the features of the eight subjects. The sample size for the ICC analysis was determined by using ICC power analysis in PASS 11.0 (NCSS, Kaysville, UT, USA) [15].

**Results**

**ICCs of the color features**

The ICCs of the color and shape features represent the repeatability of the color and the shape of the tongue. No gridline was used during the image acquisition process for Group I. For Group II, only the frontal gridlines were used during the image acquisition process. For Group III, the tongue images were acquired by using both frontal and profile gridlines.

The ICCs of the color and shape features for Group II and Group III were higher than those for Group I. In most cases, the ICCs for Group III were higher than those for Group II in Table 1. The ICC of the mean $L^*$ for the entire tongue was 0.92 for Group III whereas the ICCs of the mean $L^*$ for the entire tongue were lower than 0.9 for Groups I and II.

The ICCs of $\beta$ and $\eta$, which are the scale parameter and the shape parameter of the color distribution, respectively, were greater in Group III than in Groups I and II, as shown in Fig. 3. The ICCs of $\beta$ showed a significant difference between Groups I and III. The ICCs of $\beta$ of $L^*$ tended to be less than those of other features.

All the ICCs of color features from the separate tongue regions were higher than 0.9 in all the groups, as shown in Fig. 4. The ICCs of the color features for all the separate regions, except the AreaAD, were higher than 0.95 in Group III. The variances of $a^*$ ICC for separate regions in Group I were higher than those in Group III. The ICCs of $L^*$ and $a^*$ of AreaAD and AreaAE were lower than those of other regions and the ICCs of the root regions tended to be higher than those of the tip regions. In the tip regions, the ICCs of $L^*$ and $a^*$ for Group III were higher than those for Group II whereas there were little difference between Group II and Group III in the center regions.

The ICCs for all the color features with no gridline were lower as compared with those with only the frontal gridlines and additionally, the ICCs for all the color features with both the frontal and profile gridlines were larger compared with those with only the frontal gridlines.

**ICCs of the shape features**

The ICCs of the shape features were lower than those of the color features in Fig. 5(a). The ICCs of TARs in the EA$_A$ and EA$_D$ regions tended to be lower than those of the EA$_B$ and EA$_C$ regions. The ICCs of TARs in Group I were lower than those in...
Groups II and III. In the EA region, the ICCs of TAR showed large differences between groups.

The PCar and the PCAt signify the profile curvatures of the tip and root sides, respectively, and the TSA represents a tongue-sticking-out angle. The trends of ICCs for most of these features were similar to those of the color features; an exception is PCAt, as shown in Fig. 5(b). The ICCs of TSA, PCar, and PCAt for Group III were higher than those for Groups I and II. The ICC of PCar showed little difference between Groups I and II. On the other hand, the ICC of PCAt in Group II was lower than those for Groups I and III. From the results, the profile gridlines made the ICCs for the angle and curvature features increase. The fontal gridlines seem to be unrelated with PCar and PCAt.

**ICCs of luminance along the vertical direction**

From the tip to the root of the tongue region, the ICCs of mean luminance in the 10 separate areas tended to decrease in all the groups. The ICCs of mean a*, on the other hand, illustrated no change. In the root of the tongue regions corresponding to Area1, Area2, and Area3, there were no large differences between the groups, as shown in Fig. 6. The ICCs of mean luminance in Group I declined steadily from Area1 to Area10. In Group II, the ICCs declined sharply to less than 0.5 in the tip regions corresponding to Area8, Area9, and Area10. The ICCs declined in Group III as well, but the intensity of the decline was lower than it was in Group I and Group II.
Relationship between the color and controllable features

The shape, the horizontal location, and the profile curvature are features that are relatively controlled by the gridlines, because the gridlines serve as references for adjusting the head location or the tongue posture during the feedback process. As a result, they are called controllable features. On the other hand, the color features, such as $a^*$ and $L^*$, are not features that are directly controllable by using the gridlines. According to the ICC analysis, the repeatability of both color and controllable features was improved. These results show that improvements in the repeatability of controllable features are...
The variance of the luminance (VL) was correlated with the variances of the controllable features. The VLS of the subjects with low variances for the controllable features were lower than those of the subjects with high variances for controllable features. On the other hand, the variance of the chroma (VC) was correlated only with the variance of the HCD. The variances of similar to improvements in the repeatability of the color features. To determine the relationships between the color features and the controllable features, a correlation analysis among the variances for each of the features was performed. To estimate the variance of the feature, the standard deviation (SD) was calculated from the 15 feature values, which were extracted from 15 tongue images for each subject. A high SD for a feature corresponds to a low level of repeatability for the corresponding feature. The results of the correlation analysis are illustrated in Table 2.

Table 2

<table>
<thead>
<tr>
<th>Feature</th>
<th>sTAR</th>
<th>sPCAr</th>
<th>sPCAt</th>
<th>sHCD</th>
</tr>
</thead>
<tbody>
<tr>
<td>sL</td>
<td>0.588</td>
<td>0.705</td>
<td>0.434</td>
<td>0.708</td>
</tr>
<tr>
<td>sC</td>
<td>-0.048</td>
<td>0.080</td>
<td>0.028</td>
<td>0.676</td>
</tr>
</tbody>
</table>

sTAR, mean of the SDs for four tongue area ratios; sPCAr, SD of the profile curvature angles of the root side; sPCAt, SD of the profile curvature angle of the tip side; sHCD, SD of the horizontal center distances; sL, SD of \( L^* \) corresponding to the luminance values of the entire tongue; sC, SD of \( a^* \) corresponding to chroma values of the entire tongue.

Discussion

In this study, ring-arrayed LEDs that illuminate the tongue directly were used. Despite the use of a diffusing film in front of the LEDs, the center region of the image, which is close to the LEDs, is brighter than its edge region. These differences in light intensity cause the luminance to change according to location. The lack of homogeneity of the illumination also produces differences in the chroma properties. The white light from LEDs is generated by combining blue, green, and red diodes or by combining a single LED chip with phosphors [16]. The combination of two or more light sources with different wavelengths causes the emission spectra to change according to emission angles. Christian et al. conducted an experiment on optical ray tracing to optimize the angular homogeneity from white LEDs and showed that the chromaticity coordinates differ with pixel location in the light detector [17]. Owing to the lack of homogeneity of the illumination, the variance of HCD is closely related with the repeatability of the luminance and chroma features.

The curvature of the tongue, in particular PCAr, was related to the VL. The subjects who had low variances for PCAr tended to have low VL, as shown in Fig. 7(a). The sizes of the PCAr and
PCAt are highly dependent on the distance between the light source and the tongue. During the image acquisition process, subjects were required to attach their cheekbone to the facial contact area of the device and to keep the location of the root region of the tongue relatively constant. Consequently, the distance from the light source was largely determined by the profile angle of the tongue from the root region to the tip region. When the size of the PCAr is small, the center of the tongue is close to the light source, and likewise, when the size of PCAt is small, the tip of the tongue is distant from the light source. The difference in the distance results in the difference in the light intensity on the surface of the tongue. According to analysis of the ICCs, the repeatability of PCAt displayed large differences between the cases where both gridlines were used and where only the frontal gridlines were used. This trend happens to the luminance of the tip region in the same way. In the tip of tongue region, the profile curvature seems to play a crucial role in the repeatability of the luminance.

The tongue diagnosis is one of the most important diagnostic and clinical procedures in TEAM, why its color and shape show the change of the patient’s condition. The color of the normal tongue is pale red, but its color becomes paler, blue or purplish according to the patient’s pathological conditions [18]. In clinical research, Kanawong et al. showed the relationship between the color features of a digital tongue image and ZHENg, which is the theoretical abstraction of the symptom profiles of individual patients in TEAM theory [2,19]. The color features made various ZHENgS classified with high accuracy, whose results presented the possibility of quantitative and objective diagnosis.

Pixel values of the tongue image represent the color and shape of the tongue. In particular, the chroma values of pixels sensitively reflect the change of the color. Therefore, the patient’s pathological condition is determined dependent on the chroma of pixels in the tongue region in CTDS. In this study, the repeatability of CIE $a^*$ values that represent the chroma of red and green was evaluated, and it was improved by using the frontal and profile gridlines. This result provides the evidence that the diagnostic repeatability can be increased by using the gridlines. TEAM says that the tongue region is divided into four areas, such as its tip, its two sides, its center and its root, and each area is known to reflect the state of the corresponding organ [20]. According to the result of this study, the tip of the tongue showed the low repeatability in luminance and chroma values without gridlines. With the frontal and profile gridlines, on the other hand, the repeatability of the tip of the tongue was improved and the repeatability showed little difference between the tip and the root areas. Consequently, it is suggested that the frontal and profile gridlines are a good solution for making an accurate diagnosis in the tip of the tongue area, whereas it is possible to make a wrong diagnosis based on mistaken information without gridlines.

The tongue shape represents the states of qi or blood in TEAM. A patient with qi deficiency has a thick and large tongue, while a patient suffering from blood deficiency has a thin and triangle tongue. To estimate the tongue shape from an image, Bo et al. proposed various quantitative geometric features [21]. The features, calculated from the contour of the tongue, were used to classify round, obtuse triangular, and square tongues, which are associated with gastritis, hyperthyroidism, and heart disease, respectively. In this study, the TARs were calculated to evaluate the repeatability of tongue shape, and the frontal and profile guidelines improved the repeatability of TARs. Since the tongue shape derived from the image can change according to the angle between a camera and a tongue, it is necessary to control the location and the angle of the tongue. The frontal and profile guidelines improved the accuracy of the location and the angle, and are useful for deriving the tongue shape from the image reliably.

The color pixel values of the tongue image are also used for estimating the thickness of the tongue coating (TTC) in the CTDS. The tongue coating indicates the state of the stomach function [20]. Furthermore, the relationship between the tongue coating indicators and the oral malodor was observed [22]. Recently, Sun et al. and Zhao et al. showed that the tongue coating is related with metabolites in chronic gastritis and chronic hepatitis B patients [23,24]. In order to estimate TTC in the CTDS, Kim et al. calculated the pixel ratios of tongue coating (PRTC) corresponding to six divided areas from the entire region by using the CIE $a^*$ values [3]. TTC was categorized into no coating, thin coating, and thick coating based on PRTC, which accords with a TEAM diagnostic method. One of the crucial criteria for classifying the pixels into tongue coating and tongue body is an intensity of red color since the red intensity of the tongue coating is lower than those of the tongue body. The CIE $a^*$ values represent the red intensity of pixels, and the distribution of the CIE $a^*$ in the tongue region is highly related with
TTC. For evaluating the relationship between the guidelines and repeatability of the PRTC, ICCs of the PRTC were calculated in the same way as the color repeatability analysis. The PRTC was obtained from image binarization with a threshold of CIE a* value, in a similar fashion to Kim et al.’s method, and small clusters were removed. The threshold of CIE a* was determined by analyzing differences between the results of the manual tongue coating segmentation and the image binarization from 20 samples. The repeatability of the PRTC was improved by applying the frontal and profile guidelines (see Table 3), which seems to be caused by the improvement of the CIE a* repeatability. The PRTC algorithm is one of various TTC estimation methods, but the frontal and profile guidelines seem to be useful for improving the repeatability of various TTC estimation methods because its color features are highly related to the CIE a* value.

In this study, the evaluation of repeatability was performed by using healthy volunteers, and as such it was difficult to evaluate clearly the improvement of diagnostic repeatability with various clinical symptoms. However, the evaluated color and shape features are the essential values for the computerized tongue diagnosis, and are closely related to the TEAM doctor’s diagnostic parameters in a clinic. Furthermore, the improvement of TTC repeatability in the healthy volunteers, whose PRTC were variously distributed from 0.09 to 0.55, suggests that the frontal and profile guidelines can be useful in TTC diagnosis for patients. Based on the results of this study, if a clinical experiment is planned with the frontal and profile guidelines, it is expected that it will be possible to acquire clinical data with diagnostic accuracy (Fig. 8).

**Conclusions**

For accurate tongue diagnosis from quantitative and objective tongue features, frontal and profile feedback gridlines have been proposed. The repeatability of the color and the shape features was improved with this method. Moreover, the method using both frontal and profile gridlines showed better performance than the method with only the frontal gridlines. The angle, the horizontal location, and the shape of the tongue affected the repeatability of the chroma and luminance, but these were improved by controlling these factors with gridline feedback. It is expected that this will improve the accuracy of computerized tongue diagnosis and become a standard for medical devices in integrative medicine. Further research should test this approach on people with different TEAM pathologies.

**Conflict of interest**

No conflict of interest declared.

**Acknowledgements**

This work was supported by the grant (K13260) from the Korea Institute of Oriental Medicine (KIOM) funded by the Korea Government. The author wishes to thank Boncho Ku, KIOM, for his professional advice on statistics.

**References**


