Medical Diagnosis Algorithm Based on Tongue Image on Mobile Device

Zibo Zhou¹⁄², Dongliang Peng¹⁄², Fumeng Gao², Lu Leng¹⁄²*¹

Abstract

In traditional Chinese medical (TCM) science, tongue images can be observed for medical diagnosis; however, the tongue diagnosis of TCM is influenced by the subjective factors of doctors, and the diagnosis results vary from person to person. Quantitative TCM tongue diagnosis can improve the accuracy of diagnosis and increase the application value. In this paper, digital image processing and pattern recognition technologies are employed on mobile device to classify tongue images collected in different health states. First, through grayscale integral projection processing, the trough is found to localize the tongue body. Then the tongue body image is transferred from RGB color space to HSV color space, and the average H and S values are considered as the color features. Finally, the diagnosis results are obtained according to the relationship between the color characteristics and physical symptoms.

Key Words: tongue diagnosis, traditional Chinese medicine, image processing, mobile device.

I. INTRODUCTION

Tongue diagnosis is one of the most widely used diagnostic methods in traditional Chinese medicine [1]. The benefit of a tongue diagnosis is that it is simple and straightforward: people who need a health check can quickly determine their pathology through a regular tongue diagnosis.

1.1 Tongue segmentation algorithm

At present, the category of the tongue segmentation algorithm mainly include color model transformation, active contour model, watershed algorithm, neural network and so on. It is the easiest way to think of the separation of the tongue and other parts by color [2]. [3] observed the H channel data distribution characteristics of the image, and used it as a key segmentation factor, combined with the I channel data to segment. The author judges the largest part of the Unicom area in the segmentation result as the tongue, and finally corrects the final segmentation result by the morphological algorithm. Under the technical conditions at that time, it was not possible to accurately distinguish the color of the tongue, so the color of the middle part of the mechanical tongue was taken as the color of the tongue coating, and the color of the tongue took the color of the edge part of the tongue. At the same time, due to the difference in hardware facilities and methods, the results of tongue color recognition are also different. Because the color model conversion method based on color difference is not ideal, researchers have proposed a new active contour model algorithm [4]. Ref. [5] converted the original image from the initial coordinate system to the polar coordinate system, performed boundary enhancement and boundary extraction, and binarizing the result to the initial boundary of the contour. In [6], authors used the double-snake energy function algorithm for tongue segmentation, which improved the accuracy by nearly 10% compared to the ordinary snake algorithm. The authors of [7] used the C2G2FSnake algorithm for tongue segmentation, which increases the curve rate and reduces the complexity of the algorithm. In [8], authors used the maximum inter-class variance method to binarize the original tongue image, and then use the mathematical morphology method to modify...
the boundary of the binary image. The algorithm combines the maximum inter-class variance and mathematical morphology. Segmentation accuracy. It can be seen that the active contour model \cite{9} has attracted the attention of researchers because of the high accuracy of segmentation results and the smooth contour curve. The watershed algorithm usually brings the problem of excessive segmentation. The researchers have proposed some improvement measures for this problem. Marker control is a commonly used improvement method.

The watershed algorithm usually brings the problem of excessive segmentation. The researchers have proposed some improvement measures for this problem. Marker control is a commonly used improvement method. \cite{10} proposed an image segmentation algorithm based on fast two-step mark control. The tongue segmentation technique of the color model conversion method has gradually matured. \cite{11} proposed to first find an initial object region, by transforming and thresholding the morphological components of the image in the HSI color space and morphological operations, and then image clustering the RGB components of the initial object region to find the root of the tongue. The gap region between the upper lip and the upper lip is finally removed by means of the gap region to remove the false tongue region such as the upper lip, and the tongue is extracted from the initial target region. The tongue segmentation algorithm is summarized as shown in Table 1.

![Table 1. Tongue segmentation threshold table.](image)

<table>
<thead>
<tr>
<th>Method</th>
<th>Category</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>snake-based approach</td>
<td>active contour model</td>
<td>2007</td>
</tr>
<tr>
<td>HIS color</td>
<td>color model transformation</td>
<td>2008</td>
</tr>
<tr>
<td>double snake energy function</td>
<td>active contour model</td>
<td>2009</td>
</tr>
<tr>
<td>two-step marker-controlled</td>
<td>Watershed algorithm</td>
<td>2012</td>
</tr>
<tr>
<td>C2G2FSnake function</td>
<td>active contour model</td>
<td>2013</td>
</tr>
<tr>
<td>Threshold and clustering</td>
<td>color model transformation,</td>
<td>2017</td>
</tr>
</tbody>
</table>

2. Tongue feature extraction algorithm

In recent years, there have been few studies on tongue image feature extraction algorithms, and their research focuses on how to comprehensively and effectively acquire disease-related features.

\cite{12} tried to use wavelet transform to extract the color and texture features of different parts of the tongue image, and statistical analysis of these feature data, and finally to classify the healthy tongue image and the diseased tongue image. The authors of \cite{13} believe that most tongue feature classification methods do not take into account regional information. Therefore, the study uses a color-texture segmentation algorithm to obtain a series of homogeneous regions, and classifies these regions according to the Earth Mover distance, and finally performs corresponding feature analysis. \cite{14} studied different methods to obtain color and texture feature information in the tongue image, and based on this information, the appendicitis disease diagnosis experiment, the experimental results have higher accuracy. \cite{15} classified the shape of the tongue based on the geometric characteristics of the tongue. The study corrected the skew of the tongue with three geometric criteria and classified the shape of the tongue with seven geometric features. In order to obtain better classification results, the research used the analytic hierarchy process to increase the weight of relevant factors and used the fuzzy fusion framework to express the certainty and accuracy between these factors and the tongue category.

At the same time, in order to avoid the subjective and qualitative problems of traditional tongue diagnosis, several computer-aided diagnosis systems have been proposed \cite{16}. For example, \cite{17} proposed a Bayesian network-based system capable of identifying five different diseases with an accuracy of approximately 75%. \cite{18} used the model established by Bayesian decision to map the tongue image features of the sample with the existing classification results and identified the color of the tongue, which played a certain role in the clinical tongue diagnosis. \cite{19} proposed a hybrid image segmentation algorithm combining region-based methods with boundary-based automatic classification methods.

At present, mobile phone penetration rate in some countries is almost 100%. Most people carry mobile phones with them. The built-in sensors of smart phones (such as magnetometers, accelerometers, cameras) enable the development of new sensor systems to measure the status of telephone users and his surrounding environment, so the study of a person who monitors his health through a tongue diagnosis of a smartphone \cite{20} is beginning to rise. The concept of computerized tongue diagnosis is not new, but the research on tongue diagnosis based on mobile phones is relatively rare. Much of the work in this area has been based on the assumption that the tongue image comes from a well-controlled environment and can only be used by Chinese medicine practitioners. \cite{21} developed an Android-based automatic tongue diagnosis application based on Canny algorithm. However, due to various lighting conditions, they do not discuss how to calibrate the image color.
II. TONGUE IMAGE DIAGNOSIS ALGORITHM

1. Tongue image segmentation

1.1. Image graying

In order to diagnose the color of the tongue image, it is very important to extract the image of the tongue position. Most of the tongue segmentation algorithm first converts the image into gray image. The gray-scale integral projection algorithm adopted in this paper also needs to be divided first. The image is converted to a grayscale image. The use of the simple and color singular representation of the gray image can better segment the region. In this paper, the weighted average method is used to obtain the grayscale image, which is used to prepare the segmented tongue image. There are several common grayscale processing methods (R, G, and B represent the three primary colors, namely, red, green, and blue):

- Maximum method:
  \[ \text{GRAY} = \max(R,G,B) \] (1)

- Component method:
  \[ \text{GRAY} = R \text{ or } G \text{ or } B \] (2)

- Average method:
  \[ \text{GRAY} = \frac{R + G + B}{3} \] (3)

- Weighted average method
  \[ \text{GRAY} = 0.299 \times R + 0.578 \times G + 0.114 \times B \] (4)

1.2. Grayscale integral projection

In a tongue image, at the junction of two objects, the gray value changes greatly, and the transition of the gray value is prone to occur. According to the distribution characteristics of the color of the tongue image, the tongue and the non-region have relatively obvious boundaries and the body area of tongue is a whole in a block shape. According to these characteristics of the image, we can choose the gray integral projection algorithm to determine the position of the tongue image. The gray integral projection divides the region by the gray integral value in a certain direction and selects the divided region according to the condition. In this paper, the gray level is integrated into the horizontal and vertical directions of the image to determine the position of the tongue.

\[ \text{GS}(x) = \int_{y_1}^{y_2} \text{GRAY}(x,y)dy, \] (5)
\[ \text{GS}(y) = \int_{x_1}^{x_2} \text{GRAY}(x,y)dx. \] (6)

GS(x) represents horizontal gray-scale integration, GS(y) represents vertical integration, and GRAY(x,y) represents gray (x,y) coordinates in \([x_1,x_2],[y_1,y_2]\) images. After obtaining the gray level integral projection of the horizontal and vertical directions of the tongue image, respectively, the distribution curve of the gray level integral in the horizontal and vertical directions will be drawn, and the distribution will show many peaks and troughs, but we know that in the tongue In the image, the tongue and non-tongue regions have different brightness values. Due to the existence of the edge of the tongue, troughs will be generated, and the position of the trough can be found to determine the region of the tongue.

After obtaining the gray integral projection value, the projection function curve is fitted with smooth lines. As shown in Fig 1, the distribution and difference of the gray value of the image can be clearly seen, and then the appropriate interval is set to find the minimum value of the gray integral. (This article limits the distance between two adjacent minimum points to be no less than 20 pixels).

![Grayscale integral projection result](image-url)

Fig. 1. Grayscale integral projection result. (a) source image, (b) Vertical gray level integration, (c) Horizontal gray level integration.
where MIN(i) represents the minimum value of the i-th pixel, and px represents the pixel value of the i-th pixel. The minimum of the three values of R, G, and B is: 

\[
\text{MIN}(i) = \min(R, G, B)
\]

The HSV color space represents the color information of the tongue image more accurately and is more suitable for the classification of the tongue image. In the HSV color space, the color feature of the image is composed of three components: H, S, and V.

\[
H = \frac{60 \ast (G-B) / (SV)}{} \quad \text{if } S \neq 0 \text{ and } \max = R
\]

\[
S = \max(\min(R, G, B) - \min(\min(R, G, B), \min(G, B)), \min(G, B))
\]

\[
V = \max(R, G, B)
\]

where R, G, and B are the red, green, and blue values of the tongue image, respectively. The range of H and S is: [0, 360] for H and [0, 255] for S. The V value is in the range [0, 1].

2.1 Tongue classification

2.1.1 RGB and HSV color space

Color information is an important component in image processing technology. In many image processing algorithms, such as image segmentation algorithm, edge detection algorithm, and algorithm implementation are based on accurate color information. Nowadays, there are many color spaces, such as RGB space, HSI space, HSV space, etc., which can be used to accurately represent a color image. Different color spaces determine their different research roles. Facing an actual image processing problem, how to select the right color space is especially important. This article will introduce the RGB color space and the HSV color space.

The RGB color space is based on the three primary colors of human vision - red (R), green (G), and blue (B). The color space is considered to be red (R), green (G), and blue (B). Proper mixing can cause perception of all colors in the spectrum. In the RGB color space, each image is composed of one pixel, and the color of each pixel is composed of three components of R, G, and B, which together constitute the color feature of the image. In the tongue image, the R, G, and B values of various tongue colors are not much different, and they do not show certain regularity. In order to better classify the tongue image according to the tongue color, it is necessary to select a more suitable color space, and RGB. The space is converted to other color spaces. This article uses the HSV color space.

The HSV color space is based on human visual perception. In the HSV color space, three components are used to represent the image color, on the H component is used, the S component represents the depth of the color, called saturation. The S component and the H component play an important role in the color classification of the tongue image. The luminance (V) component indicates the degree of lightness and darkness of the color, and the range is [0, 1]. It is these characteristics of the HSV color space that are more suitable for the classification of the color features of the tongue image than the RGB color space. This paper mainly uses the mean value of the H and S components of the HSV color space to determine the color classification threshold.

2.1.2 RGB and HSV space conversion

After experimenting and analyzing the tongue image in different color spaces, it is found that the HSV color space is more suitable for the color classification of the tongue image. After segmenting the image of the tongue position, the space of the segmented tongue is converted and converted. The formula is as in formula (2.8). The range of R, G, and B is: [0-255], max represents the maximum of the three values of R, G, and B, and min represents the minimum of the three values of R, G, and B.

\[
H = \frac{60 \ast (G-B) / (SV)}{} \quad \text{if } S \neq 0 \text{ and } \max = R
\]

\[
S = \max(\min(R, G, B) - \min(\min(R, G, B), \min(G, B)), \min(G, B))
\]

\[
V = \max(R, G, B)
\]

Threshold classification

Common tongue colors are pale white, red, purple, black, etc. The common tongue coating color is divided into yellow, white, gray and so on. For these color types, this paper divides the color detection of tongue image into five categories: light red tongue, white moss, yellow moss, red tongue, and purple black tongue. After the experimental analysis of the collected tongue image, under the HSV color space, the average values of H and S are obtained respectively, \(\bar{H}\) and \(\bar{S}\). Under the HSV space, the \(\bar{H}\) and \(\bar{S}\) values of the tongue images differ greatly and have a certain regularity, that is, from the red tongue, yellow moss, purple black tongue, white moss and red tongue, \(\bar{H}\) value increased in turn, but the white moss and red tongue are basically the same. However, their \(\bar{H}\) values are larger than those of other categories, and their \(\bar{S}\) value is observed. The \(\bar{S}\) value of the reddish tongue is greater than 0.20, while the \(\bar{S}\) value of white moss is less than 0.20. For these two points, we can well classify the tongue image into the five categories mentioned above according to the two values of \(\bar{H}\) and \(\bar{S}\).

The tongue image segmentation is performed by taking the tongue image sample obtained on the Internet and the tongue image sample obtained by taking photos of itself,
and then calculating the $\bar{H}$ and $\bar{S}$ values in the HSV color space to determine the color classification of the tongue image as shown in Table 2. The range of $\bar{H}$ in Table 2 is [0, 255], the unit is degree, and the range of $\bar{S}$ is [0, 1].

### Table 2. Tongue segmentation threshold table

<table>
<thead>
<tr>
<th>Tongue category</th>
<th>$\bar{H}$</th>
<th>$\bar{S}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red Tongue</td>
<td>&lt;17</td>
<td>0.40~0.65</td>
</tr>
<tr>
<td>Yellow Tongue</td>
<td>17~135</td>
<td>0.25~0.45</td>
</tr>
<tr>
<td>Purple Tongue</td>
<td>135~240</td>
<td>0.10~0.35</td>
</tr>
<tr>
<td>White Tongue</td>
<td>$\geq$240</td>
<td>0.07~0.20</td>
</tr>
<tr>
<td>Light Red Tongue</td>
<td>$\geq$240</td>
<td>0.20~0.40</td>
</tr>
</tbody>
</table>

### III. EXPERIMENTAL RESULTS AND ANALYSIS

By downloading the already-collected tongue image and the mobile device to start the tongue diagnosis system to obtain the tongue image for detection, the JPG format picture is used for testing, and the detection result is recorded. Good results have been achieved in both tongue segmentation and tongue color diagnosis.

#### 3.1. Grayscale image of tongue

Before performing the segmentation of the tongue, the tongue image should be processed to convert the image into a grayscale image. Fig 2 shows the grayscale result of the experimental tongue image.

![Fig. 2. Tongue image grayscale result.](image)

#### 3.2. Tongue segmentation

The gray level integral projection method is used to determine the positional condition of the tongue is slightly harsh, and the photographing light is required to be good, and the collected tongue image basically does not contain other objects. Under such conditions, a better segmentation effect can be obtained. As shown in Fig 3 below, the image inside the white rectangle is the position of the tongue determined by the gray integral projection algorithm, which can accurately locate the position of the tongue, but the lack of part of the non-tongue part in the white rectangle is to a certain extent, it affects the color detection behind.

![Fig. 3. Tongue segmentation result image.](image)

#### 3.3. Tongue image color diagnosis

After the position of the tongue is determined, the average of R, G, and B is separately classified into the RGB space of the segmented tongue image to classify the color of the tongue image, and the same type of tongue image is found, and the R, G, and B values are found. There is no convergence to a certain interval, and there is no regularity, as shown in Table 3. In terms of R-means, only the red tongue value is more prominent than the value of any category, while in other categories, the mean values of G and B do not show any regularity. In the RGB color space, it is basically impossible to find a suitable classification standard, which is seriously affected by the high correlation of its color components.

Then, the cut tongue image is converted into the HSV space, and the collected sample images are color-classified according to the $\bar{H}$ and $\bar{S}$ thresholds of the tongue-segment threshold table in Table 4, which can better detect the red tongue, yellow tongue, white tongue, purple black tongue and red tongue five categories. 48 tongue images were tested, 41 of which were correctly diagnosed, and the
diagnostic rate was over 85.4%.

Table 3. Different RGB image range distribution.

<table>
<thead>
<tr>
<th>Tongue category</th>
<th>R</th>
<th>G</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red Tongue</td>
<td>140-220</td>
<td>60-90</td>
<td>60-85</td>
</tr>
<tr>
<td>Yellow Tongue</td>
<td>110-145</td>
<td>85-130</td>
<td>40-80</td>
</tr>
<tr>
<td>Purple Tongue</td>
<td>40-130</td>
<td>50-110</td>
<td>50-100</td>
</tr>
<tr>
<td>White Tongue</td>
<td>65-135</td>
<td>60-120</td>
<td>60-120</td>
</tr>
<tr>
<td>Light Red Tongue</td>
<td>90-140</td>
<td>40-110</td>
<td>35-130</td>
</tr>
</tbody>
</table>

Table 4. HSV color space recognition.

<table>
<thead>
<tr>
<th>Tongue category</th>
<th>Total number of detections</th>
<th>Correct identification number</th>
<th>Detection rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red Tongue</td>
<td>10</td>
<td>9</td>
<td>90.00%</td>
</tr>
<tr>
<td>Yellow Tongue</td>
<td>9</td>
<td>7</td>
<td>77.78%</td>
</tr>
<tr>
<td>Purple Tongue</td>
<td>11</td>
<td>9</td>
<td>81.82%</td>
</tr>
<tr>
<td>White Tongue</td>
<td>6</td>
<td>4</td>
<td>66.67%</td>
</tr>
<tr>
<td>Light Red Tongue</td>
<td>12</td>
<td>12</td>
<td>100.00%</td>
</tr>
<tr>
<td>Total</td>
<td>48</td>
<td>41</td>
<td>85.42%</td>
</tr>
</tbody>
</table>

IV. CONCLUSIONS

Through the experimental results, it is found that the accurate recognition of the color of the tongue image is based on the accurate determination of the position of the tongue. If the tongue image is not correctly segmented, the $\bar{H}$ and $\bar{S}$ under the HSV color space are used to judge the color. Significant. The gray-scale integral projection algorithm used in this paper does not accurately determine the position of the tongue. The four corners of the segmented rectangular tongue image still contain non-tongue images. Although it is a small amount, it still affects the back. Color detection. In order to improve the accuracy of the segmentation of the tongue, it is necessary to achieve accurate segmentation of the tongue after the approximate determination of the position of the tongue. At the same time, in the system design, the image should be roughly judged before the diagnosis of the tongue image. The image of the tongue object is not detected.

The detection of $\bar{H}$ and $\bar{S}$ values in the HSV color space is very good in the five categories of red tongue, yellow moss, white moss, purple black tongue and red tongue. However, this paper only divides the color into five categories, and the later research It can be considered that the color category is added based on this classification, and the category is subdivided. Different types of diseases in different positions of the tongue are a very good research direction for image feature extraction of different parts of the tongue. When the classification threshold is determined, many sample images are downloaded from the website, and their collection environments are not the same, which has a greater impact on the classification threshold determination. The same time, the sample base is too small when the threshold is determined. The sample collected in the same environment should be used to correct the color classification threshold of the tongue image. At the same time, the sample size required is as large as possible. The accurate classification threshold is the guarantee of the reliability of tongue diagnosis.

Acknowledgement

This work was supported partially by the National Natural Science Foundation of China (Grants No. 61866028, 61763033, 61662049, 61741312, 61881340421, 61663031, and 61866025), the Key Program Project of Research and Development (Jiangxi Provincial Department of Science and Technology) (20171ACE50024, 20161BBE50085), the Construction Project of Advantageous Science and Technology Innovation Team in Jiangxi Province (20165BCB19007), the Application Innovation Plan (Ministry of Public Security of P. R. China) (2017YYCXJXST048), and the Open Foundation of Key Laboratory of Jiangxi Province for Image Processing and Pattern Recognition (ET201680245, TX201604002), Innovation Foundation for Postgraduate (YC2018094, YC2017067), and “Triple-little” Extracurricular Academic Projects (2018ZD071, 2017YBRJ034).

REFERENCES


Authors

Zibo Zhou received his BS in Tongda College of Nanjing University of Posts and Telecommunications, Yangzhou, P. R. China in 2017. Currently he is pursuing his MS degree in School of Software, Nanchang Hangkong University, Nanchang, P. R. China. His research interests include palmprint recognition, image processing and deep learning.

Dongliang Peng received his BS in Nanchang Hangkong University, Nanchang, P. R. China in 2017. His research interests include palmprint recognition and image processing.

Fumeng Gao received his BS in Xinyu University, Xinyu, P. R. China in 2016. Currently he is pursuing his MS degree in School of Software, Nanchang Hangkong University, Nanchang, P. R. China. His research interests include palmprint recognition, image processing, and computer vision.
**Lu Leng** received his Ph.D. degree from Southwest Jiaotong University, Chengdu, P. R. China, in 2012. He performed his post-doctoral research at Yonsei University, Seoul, South Korea, and Nanjing University of Aeronautics and Astronautics, Nanjing, P. R. China. He was a visiting scholar at West Virginia University, USA. Currently, he is an associate professor at Nanchang Hangkong University.

He has published more than 60 international journal and conference papers and been granted several scholarships and funding projects for his academic research. He is the reviewer of several international journals and conferences. His research interests include image processing, biometric template protection, and biometric recognition.

Dr. Leng is a member of the Institute of Electrical and Electronics Engineers (IEEE), the Association for Computing Machinery (ACM), the China Society of Image and Graphics (CSIG), and the China Computer Federation (CCF).