Int J CARS manuscript No.
(will be inserted by the editor)

Regional Image Analysis of the Tongue Color Spectrum

Satoshi Yamamoto · Norimichi Tsumura · Toshiya Nakaguchi · Takao Namiki · Yuji Kasahara · Katsutoshi Terasawa · Yoichi Miyake

Received: date / Accepted: date

Abstract *Purpose* Kampo medicine (Japanese traditional herbal medicine) can identify "Mibyou" (disease-oriented state) based on visual tongue inspection by a trained physician. Surface colors of uncoated tongue provide essential features. A hyperspectral imaging system which performs regional image analysis of the tongue color spectrum was developed to automatically inspect the uncoated tongue.

Methods Hyperspectral tongue surface images were acquired by a camera equipped with an integrating sphere. The color spectrum from 4 tongue areas — uncoated tongue, coated tongue, lip, and perioral areas — were quantified. The average spectrum of each area was determined and 4 images calculated by focusing on the differences between respective spectra; highlight, shadow, tongue coat, and lip-weighted images were calculated. The uncoated tongue area was extracted by subtraction.

Results Reproducibility was evaluated by applying the camera and extraction algorithm to 44 images from human subjects. 82% were judged as acceptable by a Kampo medicine physician expert.

Conclusions Using a hyperspectral camera and extraction algorithm, the tongue color of the uncoated part was automatically extracted. This technique is suitable for tongue color analysis and may help non-trained users to identify "Mibyou".

Keywords Hyperspectral imaging · Regional image analysis · Japanese traditional herbal medicine (Kampo Medicine) · Medical diagnosis · Tongue image analysis

S. Yamamoto · K. Terasawa
Department of Japanese–Oriental Medicine,
Graduate School of Medicine, Chiba University
1-8-1 Inohana, Chuo-ku, Chiba, 260-8670 Japan
Tel.: +81-43-226-2984
Fax: +81-43-226-2985
E-mail: may-s@umin.net (S. Yamamoto)

N. Tsumura · T. Nakaguchi
Graduate School of Advanced Integration Science,
Chiba University
1-33 Yayoi-cho, Inage-ku, Chiba, 263-8522 Japan

T. Namiki · Y. Kasahara
Department of Frontier Japanese–Oriental Medicine,
Graduate School of Medicine, Chiba University
1-8-1 Inohana, Chuo-ku, Chiba, 260-8670 Japan

Y. Miyake
Research Center for Frontier Medical Engineering,
Chiba University
1-33 Yayoi-cho, Inage-ku, Chiba, 263-8522 Japan

1 Introduction

As an essential factor in visual examination, the appearance of tongue and face, mainly in terms of color, contains a lot of useful information for medical diagnosis [1]. However, inspection of tongue and face is not considered to be important in modern medical diagnosis, since it does not contain much objective information. In Kampo medicine (Japanese traditional herbal medicine), inspection of the face and tongue is one of the principal methods for diagnosis. By assessing the patient's complexion of the tongue, for example, we are able to assess the degree of mental stress, anemia and the Oketsu status (blood stagnation: one of the important pathological criteria in Kampo medicine). However, these visual examinations have not been approved as an objective factor because the skills applied in the examination are difficult to understand. This condition represents an obstacle for Kampo medicine to attain recognition in modern medicine. Development of a useful imaging system for the quantitative analysis of the tongue for removing this obstacle has been eagerly

awaited, as it was difficult to find quantitative information of the tongue by conventional *RGB* image.

Additionally, Kampo medicine contains a number of concepts useful for preventive medicine. For example, "Mibyou" (disease-oriented state: not a disease, but can easily become one if no cure is applied; contains Oketsu status) is one of the most important concepts for preserving health and preventing illnesses from developing by the early recognition of signs of abnormalities and their treatment. Thus, it can also be useful for modern preventive medicine, but these concepts are nevertheless difficult to disseminate widely, since the skills in the diagnosis of Kampo medicine have been thought to require great experience in the traditional examination, such as tongue inspection, pulse diagnosis and/or abdominal palpation, and nothing is thought to be substitutable for experience.

In the methods used up to now, segmentation of the whole tongue was performed with 3-dimensional color spaces [2–8]. In those methods, image contrast was enhanced by the Sobel operator and the edge of the tongue was traced using various algorithms. The whole tongue could be segmented clearly, but elimination of the coating was not the focus, and it was difficult to analyze the color properties of the tongue since the color space is a three-dimensional space such as *RGB* or CIE 1976 *L*a*b**. Moreover, these algorithms have good performance only in certain restricted images containing highlight and shadow, since the efficiency of the Sobel operator is dependent on the image contrast. Although highlight and shadow make the margin of each facial area, they tend to eclipse color information because highlights, mainly composed of specular reflection, show the color property of the light source, and the dynamic range of the color is not high enough in shadow. Tongue color recognition with a tongue color database has also been reported [9, 10]. By these methods, tongue areas are separated clearly, but the database was created on the empirical diagnostic method of traditional Chinese medicine, and thus we need new information free from subjective inspection.

The hyperspectral imaging system has been used for the tongue before [11–14]. In those reports, hyperspectral imaging was focused, but elimination of specular reflection, which eclipses color information, was not considered. Also, tongue area recognition was performed with hyperspectral images using a database created in an empirical manner [15]. In contrast, we have shown the correlation between "Oketsu" and spectral properties of tongue color, which is not dependent on a database [16].

As long as highlight elimination is not considered, whole tongue extraction is not difficult. An example is shown in Fig. 1.

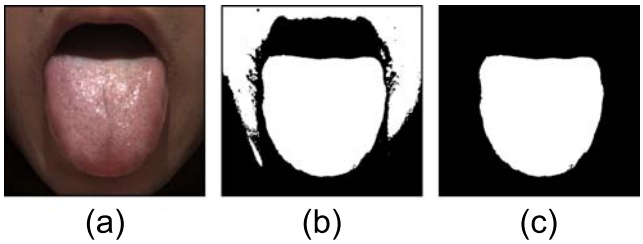


Fig. 1 Example of whole tongue extraction of image with highlight
Whole tongue extraction is performed for the image with highlight. The tongue was illuminated with a light emitting diode from a 45-degree elevated angle to sharpen the tongue edge. (a): Original *RGB* image. (b): Binary image built from *R* channel. *R* channel was selected because luminance of the uncoated tongue region and coated tongue region differs less in *R* channel. (c): The connected area was selected from the central point of the picture.

It was shown that the whole tongue could be extracted easily by using the image with specular reflections. However, these methods were not applicable for images showing the actual color without specular reflections by our system. It was necessary to develop a new algorithm to extract the tongue area since most of the specular reflection and shadow were eliminated by the system.

Therefore, in this article, we propose an automatic hyperspectral image processing method to extract the uncoated tongue area for the clinical inspection of Kampo medicine, using the hyperspectral imaging system to remove most of the specular reflex and shadow. The hyperspectral images were taken by imaging system using an integrating sphere to eliminate specular reflection artifacts. Spectral information on the uncoated tongue area, coated tongue area, lip area and perioral area were sampled from the hyperspectral image, and the sampled information was analyzed to construct the automatic extraction algorithm for the uncoated tongue area. The algorithm thusly constructed was applied to 44 images to confirm its performance.

2 Hyperspectral Imaging System

To reduce specular reflection, we tried 3 illumination systems, two light sources from 45°, two crossed polarizers in front of the camera and the light source, and an integrating sphere (Fig. 2, Table 1). An Integrating sphere is an optical component consisting of a hollow cavity with its interior coated for high diffuse reflectivity. Light rays incident on any point on the inner sur-

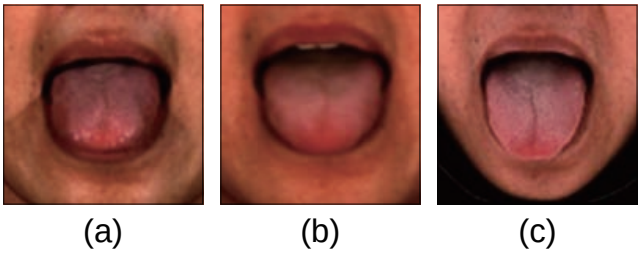


Fig. 2 Comparison of illuminations
(a): two light sources from 45°, (b): two crossed polarizers in front of the light source and the camera, (c): integrating sphere. *RGB* digital camera is used to show images.

	Cost (dollar)	Size (cm)	Light intensity	Light decrease
Two 45° lamps	0	140×15×40	High	–
Crossed polarizers	50	80×15×40	Medium	–
Integrating sphere	200	90×60×80	Medium	+

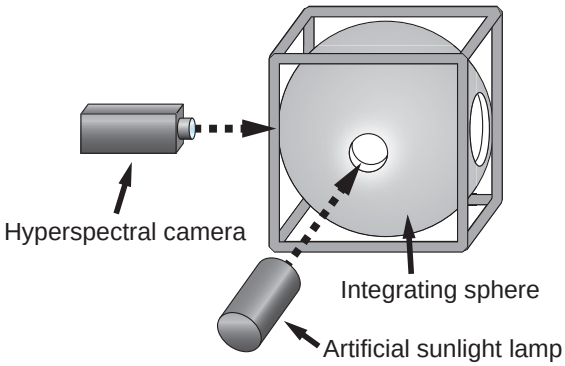
Table 1 Comparison of illuminations
Approximate cost of the system in dollar (camera and lamp excluded), size of the system (width × height × depth, PC and supporting pole excluded, core component size when equipped with the hyperspectral camera), light intensity on tongue, and presence of light decrease between tongue and camera are shown. Note that from 45° requires two lamps, and polarizer and integrating sphere require one lamp; the artificial sunlight lamp is approximately 2,200 dollars.

face are, by multiple scattering reflections, distributed equally to all other such points and effects of the original direction of such light are minimized. Crossed polarizers and an integrating sphere showed good reduction. Cost and size is an advantage of crossed polarizer, but polarizer in front of a camera decreased the light intensity emitted from tongue. Tongue showed temporal color change under dry air and heat, and thus we selected an integrating sphere for illumination.

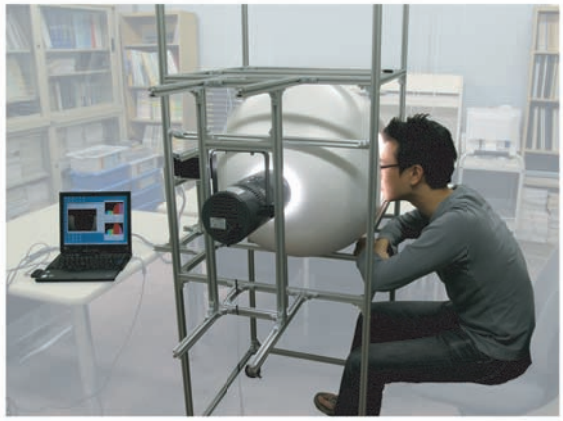
The spectral imaging system is illustrated in Fig. 3.

An artificial sunlight lamp (XC-100A, SERIC LTD., Tokyo, Japan) was utilized as the light source. The integrating sphere (60-cm diameter, made of styrofoam, painted matte white with identical reflectance between 400–800 nm) provided diffuse illumination to eliminate artifacts from specular reflections.

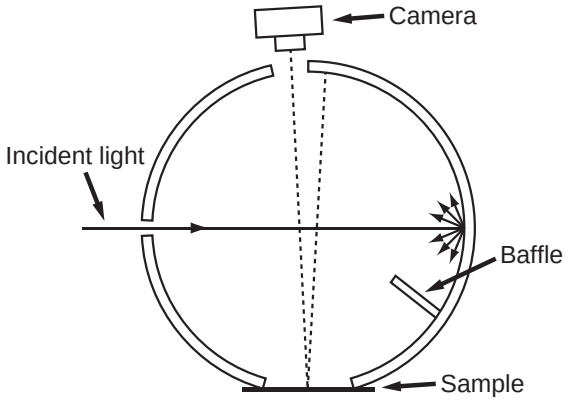
A hyperspectral camera (Hyper Spectrum Camera: HSC1700, Hokkaido Satellite Corp., Hokkaido, Japan) was used for data collection. The camera features a spectral range of 400–800 nm containing 81 bands of 5-nm resolution. It is equipped with transmission grating and an array sensor with an 8-bit monochrome CCD camera with 480×640 pixels. The optical instrument contains both a spectrometer and a scanning mechanism using an internal digital servo-motor. The camera is capable of taking a hyperspectral image every



(a) Schematic diagram of the system



(b) Photograph of the system



(c) Schematic diagram of integrating sphere

Fig. 3 Hyperspectral imaging system for the face
(a): schematic diagram of the system. (b): photograph of the system. PC-controlled hyperspectral camera, artificial sunlight lamp and integrating sphere are shown. (C): schematic diagram of an integrating sphere. Interior surface of the sphere is coated for high diffuse reflectivity. Incident light lays are distributed equally by multiple scattering reflections; baffle is a light barrier used to prevent direct illumination.

16 seconds as full-sized image. Acquired data are then normalized as spectral reflectance; it was calculated at respective pixels as the ratio of the spectral distribution divided by that of the diffuse reflection standard (White Calibration Plate CS-A5, KONICA MINOLTA HOLDINGS, INC., Tokyo, Japan). In comparison with the multi-band camera equipped with color filters, detailed spectral data are acquired, and the resulting spectrum is resistant to motion artifacts because the camera is part of a line-scan system and all the spectra of one pixel are acquired in parallel.

Forty-four hyperspectral images were acquired from 30 healthy subjects, all male Mongoloids aged 27.9±7.6 years (21–51 years). The images were taken with enough intervals (over 1 month) to allow the tongue appearance to change. In case tongue appearance changed with time, images were taken immediately after the tongue was extended; it took only two seconds for scanning the whole tongue data, and most of the specular reflection was eliminated.

3 Analysis of Hyperspectral Image

Nine hyperspectral images were randomly selected from 44 images and directed to this analysis. For each hyperspectral image, 20 small areas were picked manually, 5 small areas each from 4 tongue areas — uncoated tongue area, coated tongue area, lip area and perioral area — as shown in Fig. 4.

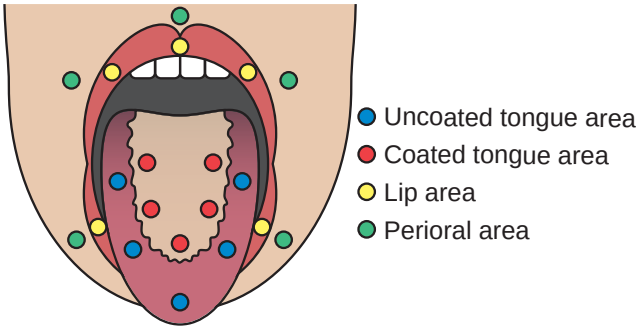
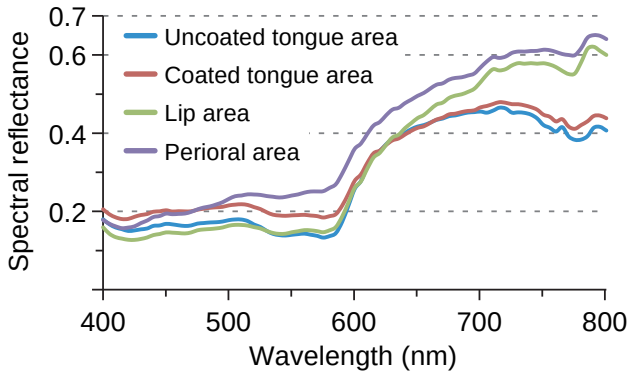


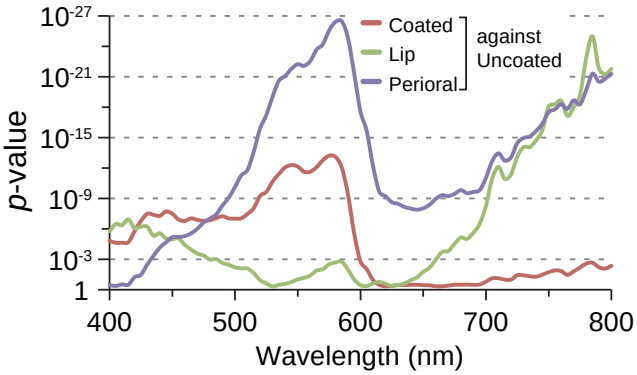
Fig. 4 Schematic diagram of tongue and peripheral area
Small area of each facial area; blue circle: uncoated tongue area, red circle: coated tongue area, yellow circle: lip area, green circle: perioral area. These small areas were selected manually.

Fifty pixels comprised each small area, and the average spectrum was calculated as the representative spectral reflectance of the small area. From 9 persons, a total of 45 small areas were picked from each facial area. The average spectra of 4 facial areas are shown in Fig. 5(a).

The average spectrum of the uncoated tongue area was compared with that of other areas. The spectrum of



(a) Spectral reflectance of 4 facial areas



(b) Significance of difference

Fig. 5 Spectral reflectance of each face area and differences between areas
(a): Spectral reflectance of 4 facial areas; blue line: uncoated tongue area, red line: coated tongue area, green line: lip area, purple line: peripheral area. (b): Significance of difference. Significance of differences between uncoated tongue area and other areas is shown. Each p-value was calculated against the spectral reflectance of uncoated area at the same wavelength with Welch's t-test. Note that p-values are shown on the reverted logarithm axis.

the coated tongue area was similar between 600–800 nm (orange–infrared) and different between 400–600 nm (violet–yellow). The spectrum of the lip area was similar between 400–650 nm (almost whole visible range) and different between 650–800 nm (red–infrared). The spectrum of the perioral area was different between almost all wavelengths, from 400 to 800 nm.

These differences between the uncoated tongue area and other facial areas were evaluated statistically by Welch's t-test [17], since the variance of every two respective samples were assumed not to be equal by F-test. The test statistics of Welch's t are as follows:

$$t = \frac{\overline{x}_{uta} - \overline{x}_{oa}}{\sqrt{\frac{S_{uta}^2}{n_1} + \frac{S_{oa}^2}{n_2}}}, \quad (1)$$

where $\overline{x}_{uta}$ denotes the average of spectral reflectance at a certain wavelength of 45 small areas from the uncoated tongue area, $\overline{x}_{oa}$ denotes that of other areas, S^2_{uta} denotes the variance of spectral reflectance at a certain wavelength of 45 small areas from the uncoated tongue area, S^2_{oa} denotes that of other areas, and n_1 and n_2 denote the number of samples ($n_1 = n_2 = 45$). For statistical analysis, the degree of freedom df is calculated as follows:

$$df = \frac{\left(\frac{S^2_{uta}}{n_1} + \frac{S^2_{oa}}{n_2}\right)^2}{\frac{\left(\frac{S^2_{uta}}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S^2_{oa}}{n_2}\right)^2}{n_2 - 1}}. \quad (2)$$

In this study, one-sided hypothesis tests were performed.

The p-value was deduced from the t distribution table as probability [17]. The p-values, uncoated tongue area against other areas, were calculated for all wavelengths and shown in Fig. 5(b) on the reverted logarithm axis. The p-value is the probability that two groups of samples are subgroups from the same population, and a small p-value means that the result is highly statistically significant. In other words, a smaller p-value is a larger spectral difference. In the coated tongue area, significance was highest around 570 nm, in the lip area the highest was over 700 nm, and in the perioral area the highest was both around 570 nm and over 700 nm.

Based on the above results, we performed extraction of the uncoated tongue area in 4 steps as shown in Fig. 6: 1) highlight elimination, 2) shadow elimination, 3) tongue coat elimination, and 4) lip elimination. The original hyperspectral image is shown in Fig. 7(a) as RGB image, where R=620 nm, G=525 nm and B=450 nm.

3.1 Highlight elimination

As the first step, the highlighted area (wall of the integrating sphere) was eliminated. To extract the grayscale image at a certain wavelength, the spectral reflection at a certain wavelength was extracted from each pixel and normalized as an 8-bit value, with a maximum value of 255 and a minimum value of 0. This first step was performed using a 420-nm image, since the spectral reflectance of the face is lowest at about 420 nm, that is, the difference of the spectra between face and highlight was greatest at 420 nm (Fig. 7(b)'). The grayscale image was converted into a black and white binary image,

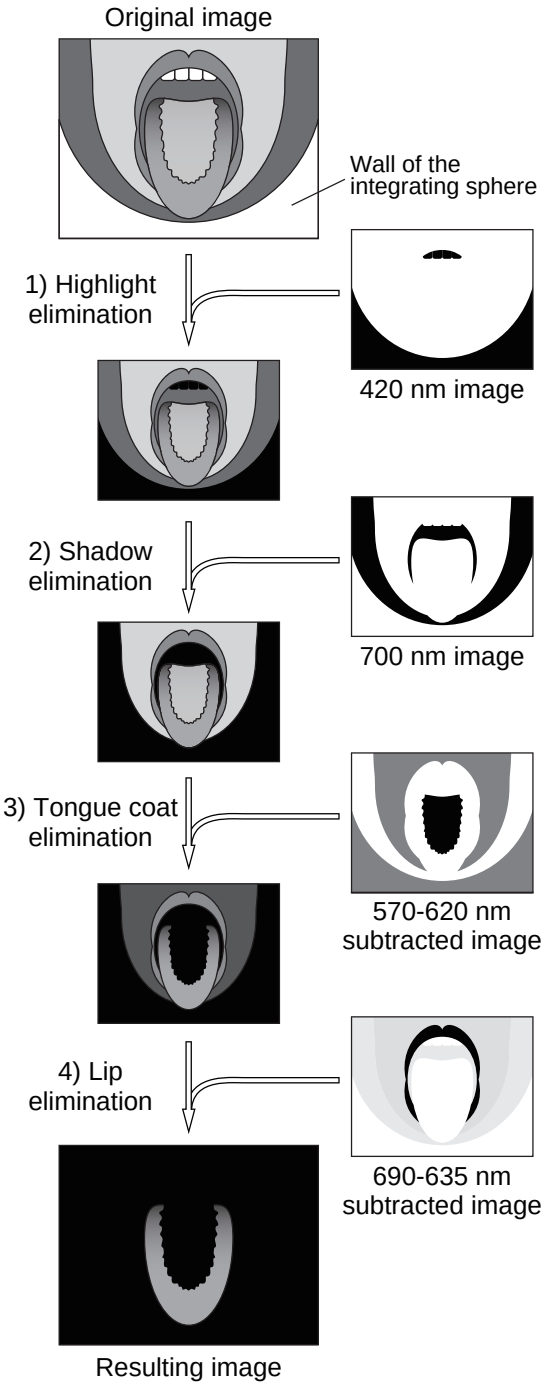


Fig. 6 Schematic diagram of uncoated tongue area extraction Process of uncoated tongue area extraction is shown. Highlight, shadow, tongue coat, and lip areas were eliminated sequentially by 4 images, 420 nm image representing highlight, 700 nm image representing shadow, 570−620 subtracted image representing tongue coat mainly, and 690−635 subtracted image representing lip area mainly. Subtracted image 570−620 nm indicates difference between 570 nm and 620 nm and subtracted image 690−635 nm indicates difference between 690 nm and 635 nm, respectively. In 420 nm, 570−620 nm, and 690−635 nm images, inverted color is shown to indicate the eliminated areas as the dark parts.

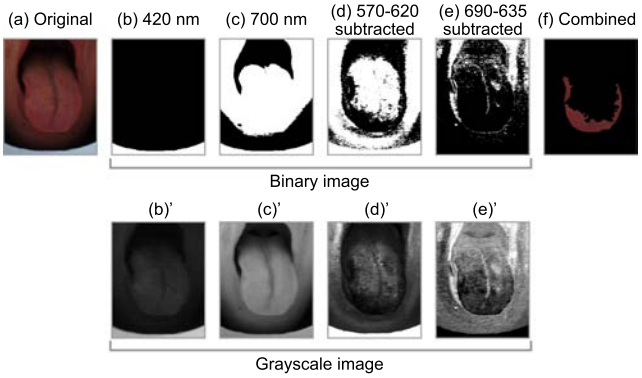


Fig. 7 Example of uncoated tongue area extraction
Extraction process is shown. (a): Original image. (b), (b)': 420 nm image; highlight is eliminated. (c), (c)': 700 nm image; shadow is eliminated. (d), (d)': Subtracted image 570–620 nm, indicating difference between 570 nm and 620 nm; note that tongue coating and perioral area are eliminated. (e), (e)': Subtracted image 690–635 nm, indicating difference between 690 nm and 635 nm; note that lip area is eliminated. (f): combined image; extracted uncoated area. (b)–(e): black and white binary images from grayscale images, threshold levels were set at 30% for (b), (c), (d) and 50% for (e), respectively. In (b), (d), (e), white areas show excluded areas, and in (c), dark area shows excluded area. (b)'–(e)': grayscale images calculated.

with a threshold level of 30% (Fig. 7(b)), and the white area was defined as the highlight.

3.2 Shadow elimination

As the second step, the shadow area was eliminated. This second step was performed with a 700-nm image, since the spectral reflectance of the face is highest at about 700 nm, that is, the difference of the spectra between face and shadow was greatest at 700 nm (Fig. 7(c)'). The grayscale image was converted into a black and white binary image, with a threshold level of 30% (Fig. 7(c)), and the dark area was defined as the shadow. Highlight image and shadow image were combined and linear noise between highlight and shadow was removed by one pixel of opening and closing.

3.3 Tongue coat elimination

As the third step, the coated tongue area was eliminated together with the perioral area. A subtracted image was calculated to eliminate the other areas from the uncoated tongue area. To calculate a subtracted image, two images, minuend image and subtrahend image, were calculated first. Each image was calculated as the summation of three images, the wavelength image, and ± 5-nm image to decrease pixel noise. Then the subtrahend image was multiplied following the ratio of spectral reflectance of the uncoated tongue area at two wavelengths, for minuend image and subtrahend image, in order to equalize the contribution of the uncoated tongue area in these two images as shown in Fig. 8.

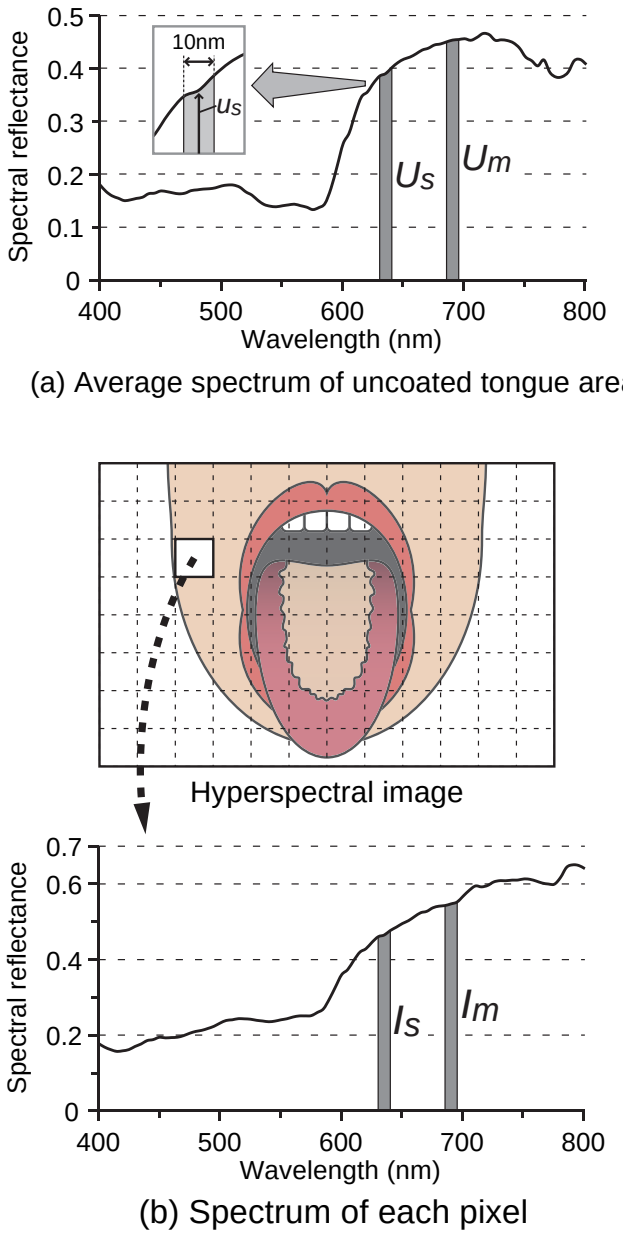


Fig. 8 Schematic representation of subtraction
This scheme represents the 690–635 nm subtraction. The process is represented with a continuous curve for clarity. (a): average spectrum of uncoated tongue area. Us: summation of spectral reflectance at 635 nm and ± 5 nm, Um: summation of spectral reflectance at 690 nm and ± 5 nm. (b): spectrum of each pixel. Is: summation of spectral reflectance at 635 nm and ± 5 nm, Im: summation of spectral reflectance at 690 nm and ± 5 nm.

After this process, contribution of the uncoated tongue area could be eliminated. These processes are written as follows:

$$SV = \left(I_m - \frac{U_m}{U_s} \cdot I_s \right) \left(U_n = \sum_{n-1}^{n+1} u_n, \quad I_n = \sum_{n-1}^{n+1} i_n, \quad n = m, s \right), \tag{3}$$

where SV denotes the calculated subtracted value for the subtracted image, u denotes the spectral reflectance of the uncoated tongue area at a certain wavelength band, i denotes the value of each pixel of the hyperspectral image at a certain wavelength band, and subscripts m and s denote the minuend wavelength and the subtrahend wavelength, n represents minuend/subtrahend wavelength band, $n-1$ and $n+1$ represent the adjacent bands at $n-5$ nm and $n+5$ nm, respectively. U and I denote the summation of u and i from $n-1$ to $n+1$, respectively. U was calculated from the average spectrum of uncoated tongue area, which is shown in Fig. 5. After this subtraction, SV was normalized as an 8-bit value, where the maximum value was 255 and the minimum value was 0. This third step was performed by using the image at 570 nm as a minuend image and at 620 nm as a subtrahend image (subtracted image 570–620 nm, Fig. 7(d)'). We selected 570 and 620 nm for the following reason. As shown in Fig. 5(b), the coated tongue area and perioral area were significantly different from the uncoated tongue area at 570 nm. The coated tongue area and lip area showed no significant differences from the uncoated tongue area at 620 nm, and thus 620 nm was selected as the control wavelength. The grayscale image was converted into the black and white binary image, with a threshold level of 30% (Fig. 7(d)), and the white area was defined as the tongue coat area. The coated tongue area was properly eliminated, but in this step the lip area could not be distinguished from the uncoated area. Most of the perioral area was eliminated.

3.4 Lip elimination

As the fourth step, the lip area was eliminated. The image at 690 nm was selected as a minuend image and at 635 nm as a subtrahend image (subtracted image 690–635, Fig. 7(e)'). Although the lip area and perioral area were significantly different from the uncoated tongue area around 790 nm as shown in Fig. 5(b), we selected 690 nm because the noise increases much more when the wavelength exceeds 700 nm. We selected 635 nm because it showed better contrast than 620 nm in this subtraction. The grayscale image was converted into a black and white binary image, with a threshold level of 50% (Fig. 7(e)), and the white area was defined as the lip area. The lip area was properly eliminated. By the latter 2 steps, almost all of the perioral area was eliminated.

Finally, four binary images were processed and effectively combined to obtain the binary uncoated area, and one pixel of opening and closing was performed to remove the dot noise due to thresholding (Fig. 7(f)). The performance of this algorithm was validated by applying it to 44 hyperspectral facial images (Fig. 9). This algorithm was applied with the same threshold level as in Fig. 7. The resulting image was then evaluated by experienced physicians of Kampo medicine. Among the 44 images, 36 (82%) were evaluated as "acceptable" (Fig. 9(a)) and 8 (18%) were evaluated as "not acceptable" (Fig. 9(b)), because lip was not properly removed (Fig. 9(b)-1, 4), and tongue was over-removed (Fig. 9(b)-1, 2).

Note that the difference of coating can be extracted in acceptable cases. It was shown that the spectral property is suitable for extracting the uncoated tongue area automatically in most cases.

4 Discussion

Tongue area extraction and segmentation have been previously performed using 3-dimensional color spaces, such as *RGB* images [3–8]. In these methods, the tongue could be extracted clearly, although elimination of the coating was incomplete, and it was difficult to analyze the color properties of the tongue since the color space was three-dimensional, such as *RGB* or CIE 1976 *L*a*b**. Moreover, these algorithms have good performance only in restricted images that contain highlight and shadow. Although highlight and shadow create the margin of each facial area, they eclipse the color information. It was necessary to develop a new algorithm to extract the tongue area since most of the highlight and shadow were eliminated by the system.

The hyperspectral imaging system has been used for tongue segmentation free from a database [11, 13, 14]. Their method contains spectral angle mapper cube or support vector machines for segmentation, which are effective methods for whole tongue extraction. In contrast, our algorithm allowed us to extract the uncoated tongue area. We suppose that segmentation free from an empirical database would become more accurate with a combination of these methods.

In this algorithm, the uncoated tongue area was extracted exactly in most cases by using the hyperspectral camera, as the spectral property differs between

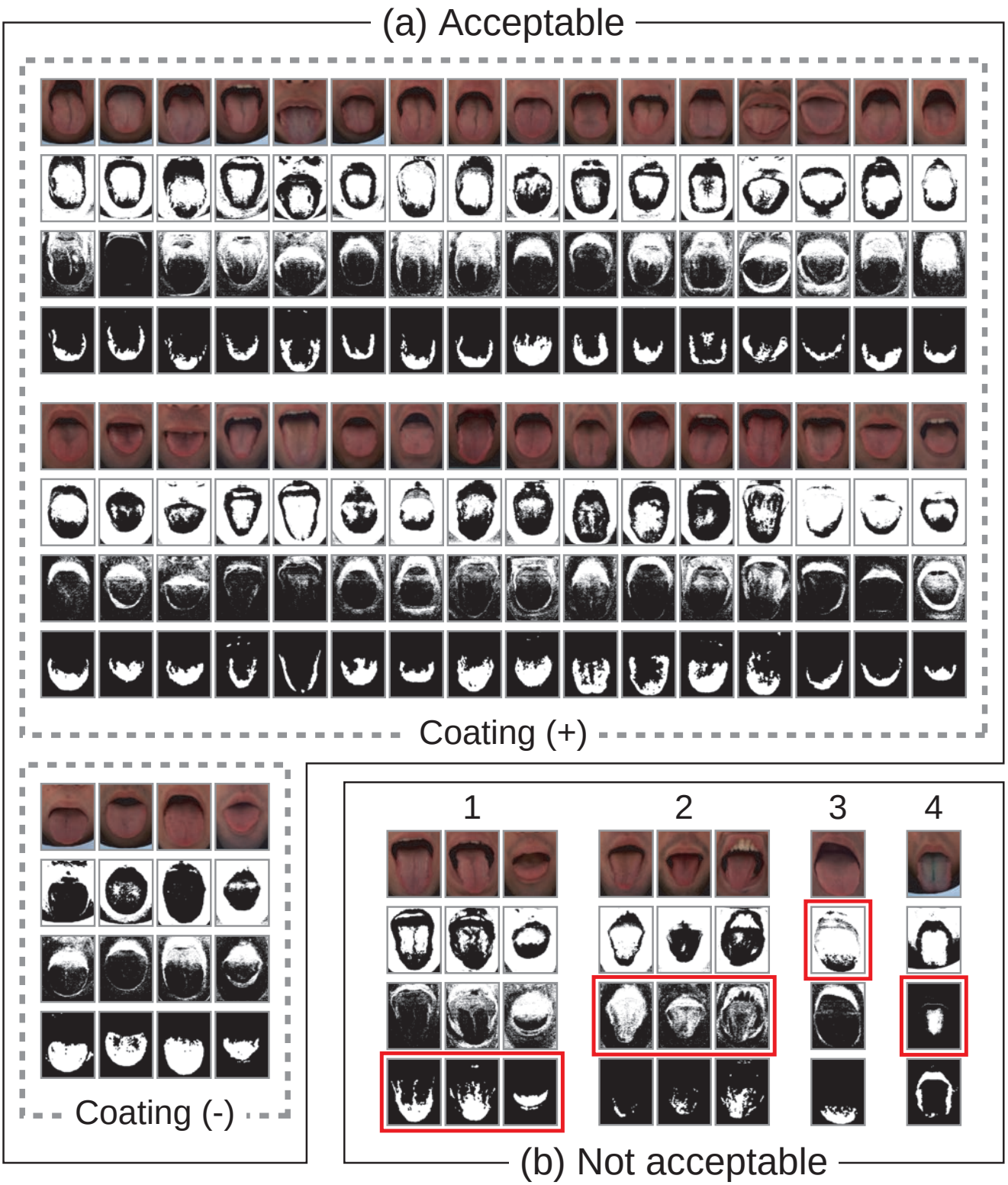


Fig. 9 Extraction of uncoated tongue region
The algorithm to extract uncoated tongue region was applied to 44 images. In each column, *RGB* image (R=620 nm, G=525 nm, B=450), subtracted image 570–620 nm (tongue coat elimination), subtracted image 690–635 nm (lip elimination) and resulting image are shown from the top. (a): Images classified into "Acceptable", separated with or without coating. Note that tongue and quantity of coating are properly extracted. (b): Images classified into "Not acceptable". They are classified into 4 categories; 1: incomplete lip elimination, 2: excess tongue in lip elimination, 3: excess tongue in tongue coat elimination, 4: unexpected tongue coat in lip elimination.

respective facial parts: uncoated tongue area, coated tongue area, lip area and perioral area. Employing this algorithm, the color of the tongue area can be analyzed without the effect of coating. Furthermore, we can analyze actual spectral properties of the tongue without specular reflection by the system, as highlight and shadow were eliminated by the imaging system with the integrating sphere without any image processing. However, in some cases, elimination errors still remain. We classified them into four categories: 1) incomplete lip elimination (Fig. 9(b)-1), 2) excess tongue in lip elimination (Fig. 9(b)-2), 3) excess tongue in tongue coat elimination (Fig. 9(b)-3), and 4) unexpected tongue coat in lip elimination (Fig. 9(b)-4). The first error, incomplete lip elimination, occurs when tongue color is intact and lip color is scarlet with erosive change. This error is expected to be corrected by taking shape into account, such as by the Snake method or gradient vector flow method [4, 18]. The second error, excess tongue in lip elimination, was observed when tongue color was deep red. For elimination of this error, collecting hyperspectral tongue images of similar color and developing another algorithm with other wavelength(s) is required. The third error, excess tongue in tongue coat elimination, seems to be of edematous change of the tongue. To solve this problem, more similar images are required. The fourth error, unexpected tongue coat in lip elimination, occurs when tongue coating has a gray color. In this case, lip was properly eliminated when the threshold level was set at 30%. However, in order to obtain acceptable results for all the images, if all the variations of the tongue are taken into consideration, the uncoated tongue area could be properly extracted in most cases. More improvement is needed, as tongue color variety should not be so large as to change the performance of the algorithm; difference among the tongue areas was much greater than that in one area between individuals. We are currently in the process of collecting and analyzing a wide variety of tongue images.

In contrast, variations of tongue coating in terms of quantities, colors and textures also contain a large amount of clinical information in Kampo medicine, although the tongue coating color can be easily influenced by some foreign elements such as coffee, colored foods and smoking, and thus can interfere with the analysis. In this algorithm, coating could be discriminated even after a cup of coffee, and thus we are planning to subjectively measure the coating by area and quantity, not by color.

For diagnosis, a tongue color database is frequently employed [9, 15, 19]. These databases are based on the criteria of traditional Chinese medicine based on subjective inspections by physicians. However, recognition performance is quite good with their criteria, such as Bayesian networks. Additionally, it has been shown that tongue appearance is correlated with some diseases [19, 20]. It has been suggested that pancreatitis and appendicitis patients can be recognized by tongue image. We suppose it would be clearer if detailed color information would be additionally employed. Now we need to develop a new algorithm free from empirical subjective inspection, using actual color information.

The hyperspectral imaging system has a problem in terms of the color depth of the camera. The hyperspectral camera has an 8-bit depth per band, which limits the dynamic range of the image. One method to solve this problem is to build a high-dynamic range image from 3 shots with different exposure times, but this is not possible for tongue image because of the unavoidable movements. It would of course be useful for skin. Meanwhile, an integrating sphere was an unexpected solution. It provides diffuse illumination, i.e., it reduces the contrast (dynamic range) of the image.

The system has another big problem, that of cost; the commercially available hyperspectral camera costs approximately 50,000 dollars. We used the hyperspectral camera to obtain detailed color information of the tongue because hyperspectral images provide for greater sensitivity and robustness to the algorithm, since the color spectrum can be reconstructed from a 3-band imaging system when detailed color components are determined in the color spectrum. Detailed color information is already examined in respect to the skin [21–23]. It has been shown that face color could be estimated by three principal components, and these could be estimated by three elemental colors such as RGB [21]. Contents of pigmentation of the face, such as melanin or hemoglobin, can also be estimated by independent component analysis [22, 23]. Tongue color and lip color are also thought to be estimable by the three components. In the next stage, we will analyze the color information of the extracted uncoated tongue area in detail to find some component that fits as a diagnostic factor, and reconstruct these components from 3-band images. They also showed that spectral absorbance, the negative logarithm of spectral reflectance, can be calculated as the linear combination of respective pigments after modified Beer–Lambert law. In this study, the properties of spectral difference among the areas were not focused upon, since it was not necessary for segmentation, although, for future work, this needs to be determined to find changes inside the tongue.

5 Conclusion

An hyperspectral imaging system and tongue region extraction method were proposed in this article. This system is equipped with an integrating sphere, an artificial sunlight lamp, and a hyperspectral camera, and a hyperspectral image was taken without specular reflection in a short period. Tongue area without coating was properly extracted by subtracting other areas, focusing on the spectral differences among respective facial areas. Experimental results revealed that spectral properties differ among facial areas, and tongue areas are effectively distinguishable with these differences. For the future work, we are planning to analyze the spectral property of the extracted region to find correlations with clinical symptoms to find the "Mibyou" factor as early as possible, recovering spectral properties from 3-band images, such as *RGB* images, and finding component changes in the tongue that cause spectral differences of tongue color.

Acknowledgements This work is supported by Joint Industry–Academia–Government Cooperative Project.

References

1. Sato Y, Hanawa T, Arai M, Cyong J, Fukuzawa M, Mitani K, Ogihara Y, Sakiyama T, Shimada Y, Toriizuka K, et al (2005) Introduction to Kampo — Japanese Traditional Medicine. The Japan Society for Oriental Medicine. Tokyo: Elsevier Japan
2. Wu J, Zhang Y, Bai J (2005) Tongue area extraction in tongue diagnosis of traditional Chinese medicine. In: Int Conf Eng Med Bio Soc, Shanghai, China, pp 4955–4957
3. Pang B, Zhang D, Wang K (2005) The bi-elliptical deformable contour and its application to automated tongue segmentation in Chinese medicine. IEEE Trans Med Imag 24(8):946–956
4. Zhang H, Zuo W, Wang K, Zhang D (2006) A snake-based approach to automated segmentation of tongue image using polar edge detector. Int J Imaging Syst Technol 16(4):103–112
5. Yu S, Yang J, Wang Y, Zhang Y (2007) Color active contour models based tongue segmentation in traditional Chinese medicine. In: Int Conf Bioinformatics Biomed Eng, Wuhan, China, pp 1065–1068
6. Fu Z, Li W, Li X, Li F, Wang Y (2008) Automatic tongue location and segmentation. In: IEEE Int Conf Audio Lang Image Process, Shanghai, China, pp 1050–1055
7. Jian-qiang D, Yan-sheng L, Ming-feng Z, Kang Z, Cheng-hua D (2008) A novel algorithm of color tongue image segmentation based on HSI. In: Int Conf Biomed Eng Informat, Hainan, China, vol 1, pp 733–737
8. Gao Z, Cui M, Lu G (2008) A Novel Computerized System for Tongue Diagnosis. In: Int Seminar Future Inform Technol Manage Eng, pp 364–367
9. Pang B, Zhang D, Li N, Wang K (2004) Computerized tongue diagnosis based on Bayesian networks. IEEE Trans Biomed Eng 51(10):1803–1810
10. Wang Y, Yang J, Zhou Y, Wang Y (2007) Region partition and feature matching based color recognition of tongue image. Pattern Recognition Lett 28(1):11–19
11. Liu Z, Yan J, Zhang D, Li Q (2007) Automated tongue segmentation in hyperspectral images for medicine. Appl Opt 46(34):8328–8334
12. Liu Z, Li Q, Yan J, Tang Q (2007) A novel hyperspectral medical sensor for tongue diagnosis. Sensor Review 27(1):57–60
13. Qing-Li L, Yong-Qi X, Jian-Yu W, Xiao-Qiang Y (2007) Automated tongue segmentation algorithm based on hyperspectral image. J Infrared Millim Waves (Hongwai yu Haomibo Xuebao) 26:77–80, (in Chinese)
14. Zhi L, Zhang D, Yan J, Li Q, Tang Q (2007) Classification of hyperspectral medical tongue images for tongue diagnosis. Computerized Med Imaging Graph 31(8):672–678
15. Li Q, Liu Z (2009) Tongue color analysis and discrimination based on hyperspectral images. Compt Med Imaging Graph 33(3):217–221
16. Yamamoto S, Ogawa-Ochiai K, Tsumura N, Nakaguchi T, Namiki T, Kasahara Y, Terasawa K, Miyake Y (2010) Analysis of multi-spectral tongue image and correlation with disease-oriented state. In: Computer Assisted Radiology and Surgery, 24th International Congress and Exhibition (CARS 2010), Geneva, Switzerland, (accepted)
17. Johnson R, Bhattacharyya G (1996) Statistics: Principles and Methods, 3rd edn, WILEY, Hoboken, NJ, USA, chap DRAWING INFERENCES FROM LARGE SAMPLES, COMPARING TWO TREATMENTS, pp 334–335, 403–417
18. Xu C, Prince J (1998) Snakes, shapes, and gradient vector flow. IEEE Trans imag process 7(3):359–369
19. Pang B, Zhang D, Wang K (2005) Tongue image analysis for appendicitis diagnosis. Inform Sci 175(3):160–176
20. Zhang D, Pang B, Li N, Wang K, Zhang H (2005) Computerized diagnosis from tongue appearance using quantitative feature classification. Am J Chinese Med 33(6):859–866

21. Imai F, Tsumura N, Haneishi H, Miyake Y (1996) Principal component analysis of skin color and its application to colorimetric color reproduction on CRT display and hardcopy. J Imag Sci Technol 40(5):422–430
22. Tsumura N, Haneishi H, Miyake Y (2000) Independent component analysis of spectral absorbance image in human skin. Opt Rev 7(6):479–482
23. Tsumura N, Ojima N, Sato K, Shiraishi M, Shimizu H, Nabeshima H, Akazaki S, Hori K, Miyake Y (2003) Image-based skin color and texture analysis/synthesis by extracting hemoglobin and melanin information in the skin. ACM Trans Graph 22(3):770–779