Hyperspectral Imaging of Face for Tongue Region Extraction

Satoshi Yamamoto[1], Norimichi Tsumura[2], Toshiya Nakaguchi[2], Takao Namiki[3], Yuji Kasahara[3]
Katsutoshi Terasawa[1], Yoichi Miyake[4]

[1)] Department of Japanese–Oriental Medicine, Graduate School of Medicine
[2)] Graduate School of Advanced Integration Science
[3)] Department of Frontier Japanese–Oriental Medicine, Graduate School of Medicine
[4)] Research Center for Frontier Medical Engineering
Chiba University, Chiba, Japan

Abstract

In this article, we propose a hyperspectral imaging system and automatic hyperspectral image processing method to extract the uncoated area of the tongue for the clinical inspection of Kampo medicine (Japanese traditional herbal medicine). The automatic processing methods were built based on the analysis of nine taken hyperspectral image. The effectiveness of the method was confirmed by applying to other eight hyperspectral images.

Introduction

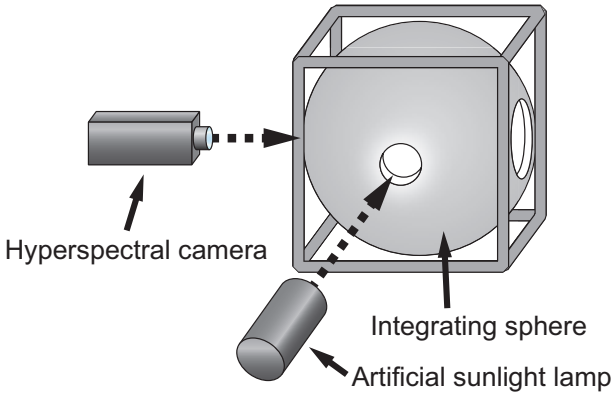
As an essential factor in visual examination, the appearance of tongue and face appearance, mainly in terms of color, contains a lot of useful information for medical diagnosis [1]. However, inspection of tongue and face is not considered to be important in modern medical diagnosis, since it does not contain many 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.

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 with 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, tongue coated 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 8 images to confirm its performance.

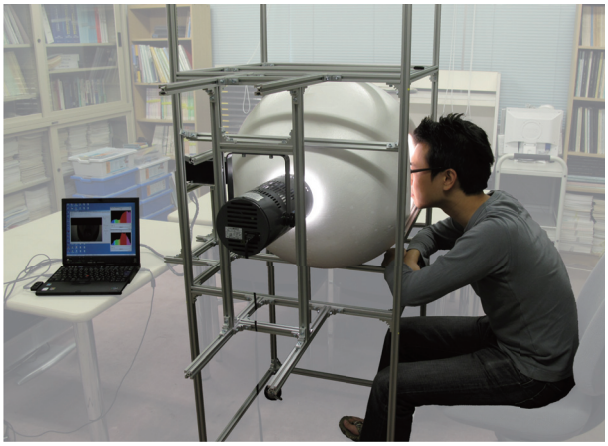
Hyperspectral Imaging System

The hyperspectral imaging system is illustrated in Figure 1.

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



(a) Schematic diagram of the system



(b) Photograph of the system

Figure 1. *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.

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 installs both the spectrometer and a scanning mechanism using an internal digital servo-motor. The camera

is capable of taking a hyperspectral image every 16 seconds as full-sized images. Acquired data are then normalized as spectral reflectance. 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. The spectral reflectance was calculated at each pixel 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).

Nine Hyperspectral images were acquired from 9 healthy subjects, all male Mongoloids aged 22.9±3.0 years (21–30 years). In case tongue appearance changes 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.

Analysis of Hyperspectral Image

Nine hyperspectral images were 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 area, lip area and perioral area — as shown in Figure 2.

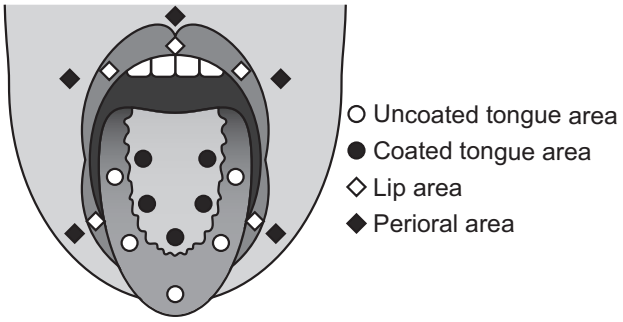


Figure 2. *Schematic diagram of tongue and peripheral area*
Small area of each facial area; open circle: uncoated tongue area, filled circle: coated tongue area, open diamond: lip area, filled diamond: perioral area. These small areas were selected manually.

Fifty pixels comprise 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 Figure 3(a).

The average spectrum of the uncoated tongue area was compared with that of other areas. The spectrum of the coated tongue area is similar between 600 to 800 nm (orange–infrared) and different between 400 to 600 nm (violet–yellow). The spectrum of the lip area was similar between 400 to 650 nm (almost whole visible range) and different between 650 to 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 [2]. The test statistics of Welch *t* are as follows:

$$t = \frac{\bar{x}_{uta} - \bar{x}_p}{\sqrt{\frac{S_{uta}^2}{n_1} + \frac{S_p^2}{n_2}}} \tag{1}$$

where $\bar{x}_{uta}$ denotes the average of spectral reflectance at a certain wavelength of 45 small areas from the uncoated tongue area, $\bar{x}_p$

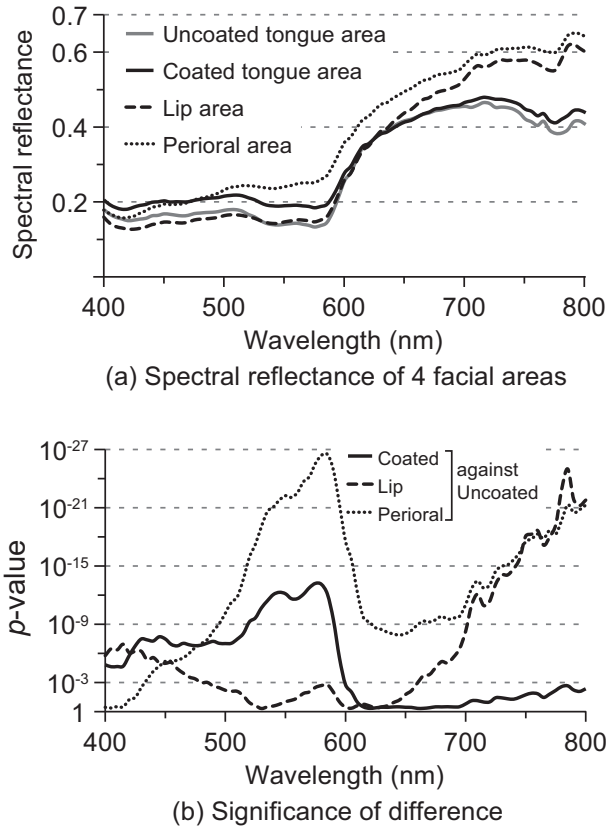


Figure 3. *Spectral reflectance of each face area and differences between areas*
(a): Spectral reflectance of 4 facial areas; gray line: uncoated tongue area, solid black line: coated tongue area, dashed black line: lip area, dotted black line: peripheral area. (b): Significance of difference. Significance of differences between tongue uncoated 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.

denotes that of other areas, S_{uta}^2 denotes the variance of spectral reflectance at a certain wavelength of 45 small areas from the uncoated tongue area, S_p^2 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:

$$df = \frac{\left(\frac{S_{uta}^2}{n_1} + \frac{S_p^2}{n_2}\right)^2}{\frac{\left(\frac{S_{uta}^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_p^2}{n_2}\right)^2}{n_2 - 1}}. \tag{2}$$

In this study, one-sided hypothesis tests were performed. The *p*-value was deduced from the *t* distribution table as probability [2]. The *p*-values, uncoated tongue area against other areas, were calculated for all wavelengths and shown in Figure 3(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 the coated tongue area, significance was highest around 570 nm, in the lip area the highest was around 790 nm, and in the perioral area the highest was both around 570 nm and 790 nm. Based on the above results, we performed extraction of

the uncoated tongue area in 4 steps as shown in Figure 4: 1) highlight elimination, 2) shadow elimination, 3) tongue coat elimination, and 4) lip elimination. The original hyperspectral image is shown in Figure 5(a) as *RGB* image, where *R*=620 nm, *G*=525 nm and *B*=450 nm.

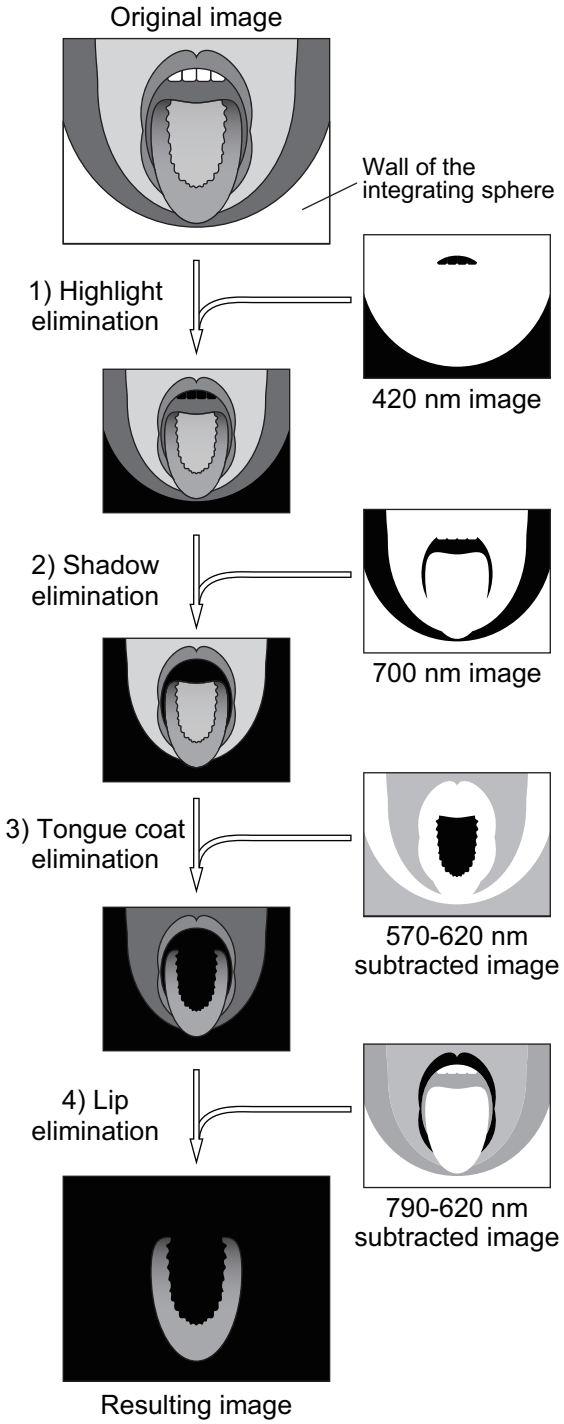


Figure 4. *Schematic diagram of uncoated tongue area extraction*
Process of uncoated tongue area extraction is shown. Highlight, shadow, tongue coat, and lip areas were eliminated sequentially by 4 images, 420 nm image representing highlight, 700 nm image representing shadow, 570−620 subtracted image representing tongue coat mainly, and 790−620 subtracted image representing lip area mainly. Subtracted image 570−620 nm indicates difference between 570 nm and 620 nm and subtracted image 790−620 nm indicates difference between 790 nm and 620 nm, respectively.

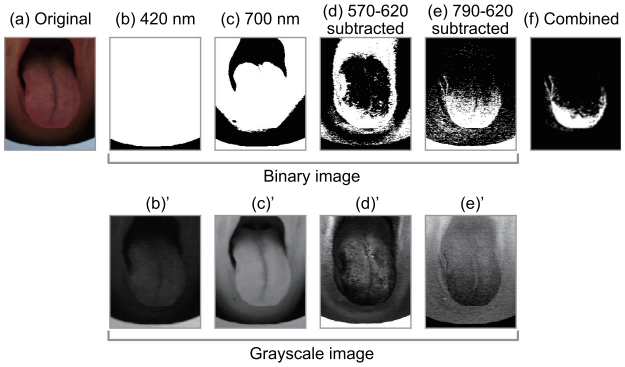


Figure 5. *Example of uncoated tongue area extraction*
Actual images of extraction process are shown. (a): Original image. (b), (b)': 420 nm image; highlight is eliminated. (c), (c)': 700 nm image; shadow is eliminated. (d), (d)': Subtracted image 570−620 nm, indicating difference between 570 nm and 620 nm; note that tongue coating and perioral area are eliminated. (e), (e)': Subtracted image 790−620 nm, indicating difference between 790 nm and 620 nm; note that lip area and perioral area are 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 40% for (e), respectively. (b)'–(e)': grayscale images calculated.

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 (Figure 5(b)'). The grayscale image was converted into a black and white binary image, with a threshold level of 30% (Figure 5(b)).

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 (Figure 5(c)'). The grayscale image was converted into a black and white binary image, with a threshold level of 30% (Figure 5(c)).

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 area from the uncoated tongue area. To calculate the subtracted images, two images, minuend image and subtrahend image, were calculated first. 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 Figure 6. After this process, contribution of the uncoated tongue area could be eliminated. These processes are written as follows:

$$SV = \left(I_m - \frac{U_m}{U_s} \cdot I_s \right) \tag{3}$$

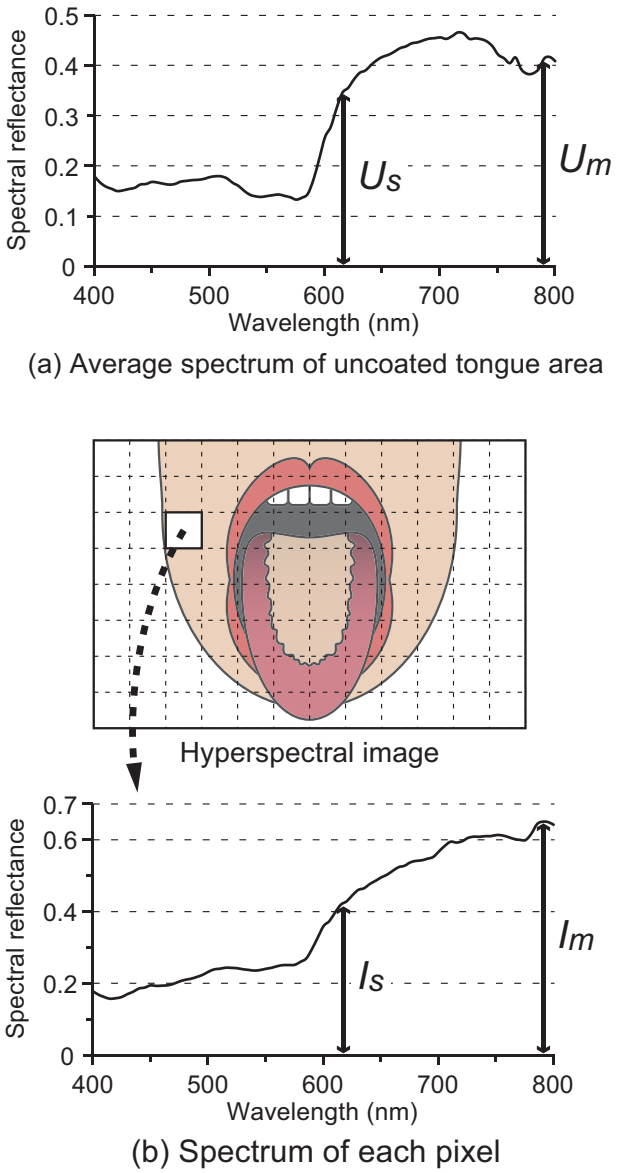


Figure 6. *Schematic representation of subtraction*
This scheme represents the 790−620 nm subtraction. (a): Average spectrum of uncoated tongue area. Us: spectral reflectance at 620 nm, Um: spectral reflectance at 790 nm. (b): Spectrum of each pixel. Is: spectral reflectance at 620 nm, Im: spectral reflectance at 790 nm.

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, I denotes the value of each pixel of the hyperspectral image at a certain wavelength, and subscripts m and s denote the minuend wavelength and the subtrahend wavelength, respectively. 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, Figure 5(d)'). We selected 570 and 620 nm for the following reason. As shown in Figure 3(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% (Figure 5(d)). 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.

Lip Elimination

As the fourth step, the lip area was eliminated together with the perioral area. The image at 790 nm was selected as a minuend image and at 620 nm as a subtrahend image (subtracted image 790−620, Figure 5(e)'). We selected 790 nm because, as shown in Figure 3(b), lip area and perioral area were significantly different from the uncoated tongue area at 790 nm. We selected 620 nm for the same reason as the third step. The grayscale image was converted into a black and white binary image, with the threshold level at 40% (Figure 5(e)). 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 a Gaussian blur filter was applied to eliminate the solitary pixels and line noise between areas (Figure 5(f)). The performance of this algorithm was validated by applying it to 8 other hyperspectral facial images (Figure 7). This algorithm was applied with the same threshold level as in Figure 5 to show the effect of difference between persons. Although error of extraction still exists on the face, the uncoated tongue area and the coated area were accurately separated. In these persons, #3 had no coating, #8 had less coating, #4 had more coating, and the others had moderate coating. It is noted that the difference in these quantities for coating can be successfully extracted in Figure 7(f). It was shown that the uncoated tongue area was extracted automatically using the hyperspectral camera.

Discussion

Although the color of the tongue presents a lot of useful information for medical diagnosis, there has not been much research focused on tongue color such as the spectral properties of the tongue, as the skills in the diagnosis of Kampo medicine have been thought to require great experience in the traditional examination and are not to be substituted. However, those visual examinations need to be objectively evaluated in order to globalize Kampo medicine. It has always been considered that the skills of Kampo medicine must be based on extraordinary levels of experience and education, and those extremely high expectations have restricted the encouragement of the dissemination of Kampo medicine. The hyperspectral imaging system, one of the objective measurement systems available, will help Kampo medicine to become appreciated as a proper medical diagnostic system.

Tongue area extraction and segmentation have been previously performed using *RGB* images [3, 4, 5, 6]. 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 specular reflection and shadow were eliminated by the system.

In this algorithm, the uncoated tongue area was extracted precisely by using the hyperspectral camera, as the spectral prop-

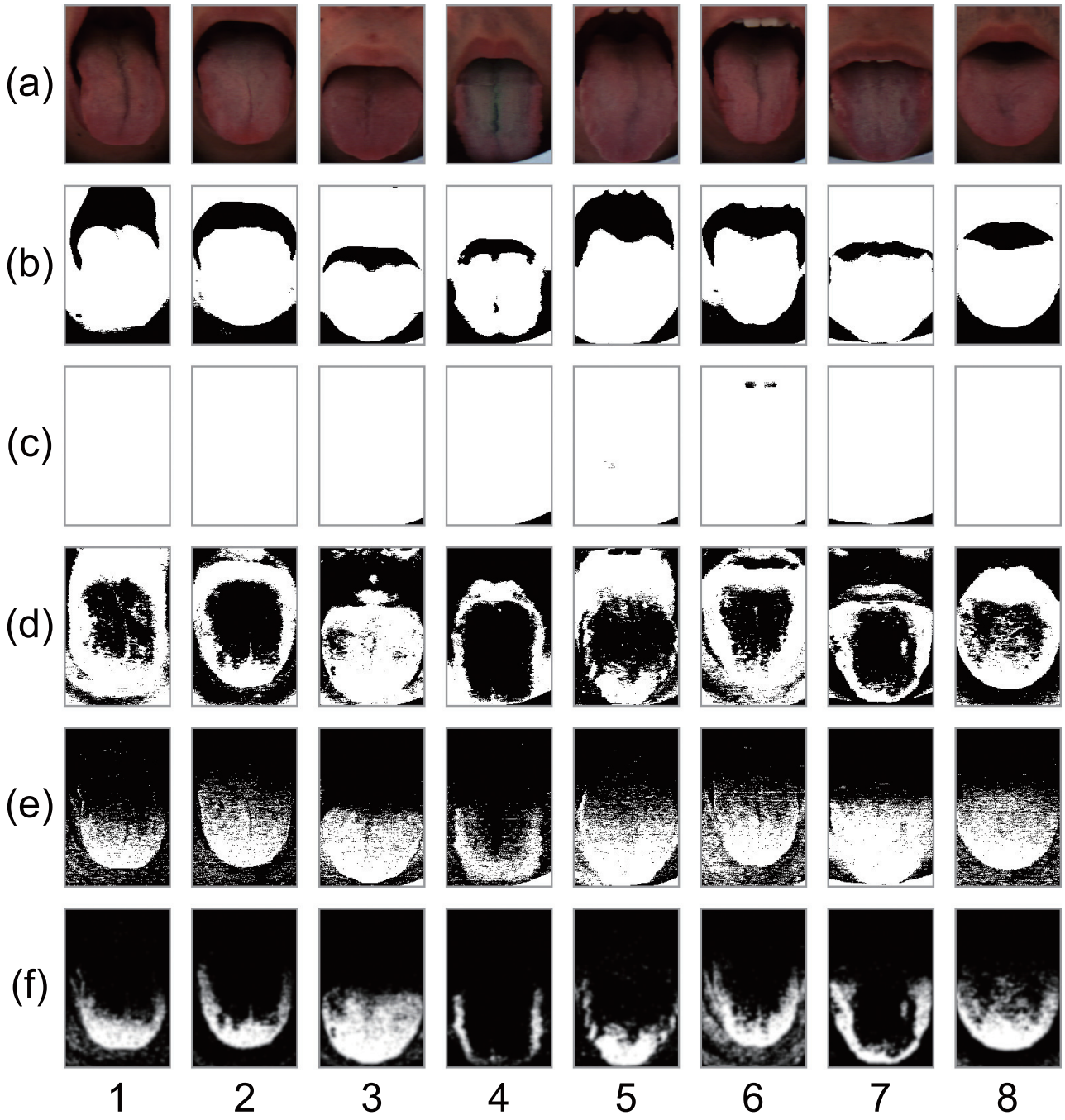


Figure 7. *Extraction of tongue uncoated region*

The algorithm to extract tongue uncoated region was applied to the other 8 images. Numbers: Personal identification digit. (a): RGB image (R=620 nm, G=525 nm, B=450, (b): 420-nm binary image; highlight is eliminated, (c): 700-nm binary image; shadow is eliminated, (d): subtracted binary image 570−620 nm, which indicates difference between 570 nm and 620 nm; tongue coating and perioral area are eliminated, (e): subtracted binary image 790−620 nm, which indicates difference between 790 nm and 620 nm; lip area and perioral area are eliminated. (f): combined image; extracted uncoated area. Threshold levels are set at 30% for (b), (c), (d) and 40% for (e), respectively. Note that the quantity of tongue coating varies between individuals and was properly eliminated in (d) and (f). #3: no coating, #8: less coating, #4: more coating and #1,2,5,6,7: moderate coating within normal limits.

erties differs among respective facial parts: uncoated tongue area, coated 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, since specular reflection and shadow were eliminated by the imaging system with integrating sphere without any image processing. However, error of extraction still exists because the threshold levels to obtain the binary image are optimized to the hyperspectral image of one person. Automated optimization for each person together with pattern recognition is the next task.

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 quantify the coating by area and quantity, not by color.

It has been shown that face color could be estimated by three principal components, and these three principal components could be estimated by three elemental colors such as *RGB* [7]. Contents of pigmentation of the face, such as melanin or hemoglobin, can also be estimated by independent component analysis [8, 9]. Tongue color and lip color are also thought to be estimable by the three components.

For the next experiment, we are collecting more samples of variable appearance and plan to analyze the color information of the extracted uncoated tongue area in detail to find some component(s) that fits as a diagnostic factor. In the next stage, we are going to perform principal component analyses of the uncoated tongue area and principal vector rotation to fit the clinical symptoms.

Acknowledgement

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

References

[1] The Japan Society for Oriental Medicine, Introduction to Kampo: Japanese traditional medicine (Elsevier Japan, Tokyo, Japan, 2005) pg. 53.

[2] Richard A. Johnson, Gouri K. Bhattacharyya, Statistics: Principles and Methods (WILEY, Hoboken, NJ, 1996) pg. 334, 403.

[3] J. Wu, Y. Zhang, J. Bai, Tongue Area Extraction in Tongue Diagnosis of Traditional Chinese Medicine, Proc. IEEE-EMBS, pg. 4955. (2005).

[4] H. Zhang, W. Zuo, K. Wang, D. Zhang "A Snake-Based Approach to Automated Segmentation of Tongue Image Using Polar Edge Detector" Int. J. Imaging Syst. Technol., 16(4), 103 (2007).

[5] S. Yu, J. Yang, Y. Wang, Y. Zhang, Color Active Contour Models Based Tongue Segmentation in Traditional Chinese Medicine, Proc. iCBBE, pg. 1065. (2007).

[6] Z. Fu, W. Li, X. Li, F. Li, Y. Wang, Automatic Tongue Location and Segmentation, Proc. ICALIP, pg. 1050. (2008).

[7] F. H. Imai, N. Tsumura, H. Haneishi, Y. Miyake "Principal component analysis of skin color and its application to colorimetric color reproduction on CRT display and hardcopy" JIST, 40(5), 422 (1996).

[8] N. Tsumura, H. Haneishi, Y. Miyake "Independent component analysis of spectral absorbance image in human skin" Opt. Rev., 7(6), 479 (2000).

[9] N. Tsumura, N. Ojima, K. Sato, M. Shiraishi, H. Shimizu, H. Nabeshima, S. Akazaki, K. Hori, Y. Miyake "Image-based skin color and texture analysis/synthesis by extracting hemoglobin and melanin information in the skin" ACM Trans. Graph., 22(3), 770 (2003).

Author Biography

Satoshi Yamamoto *received his B.S. in biology from National Institution for Academic Degrees and University Evaluation (2003), his M.S. in medical science from Nagoya University (2003) and his B.M. and M.D. from Chiba University (2007). Currently he is a Ph.D. student at Chiba University Graduate School of Medicine and also working in Chiba University Hospital as a member of Japanese–Oriental (Kampo) Medicine. His work has focused on new findings between traditional medicine and optical engineering.*

Norimichi Tsumura *received the B.E., M.E. and Dr. Eng degrees in Applied Physics from Osaka University in 1990, 1992 and 1995, respectively. He moved to the, Chiba University in April 1995, as an Assistant Professor. He is currently Associate Professor in Chiba University. He got Charles E. Ives Award (Journal Award: IS&T) in 2002, 2005. He is interested in the color image processing, computer vision & graphics and biomedical optics.*

Toshiya Nakaguchi *received the B.E., M.E., and Ph.D. degrees from Sophia University, Tokyo, Japan in 1998, 2000, and 2003, respectively. From 2006 to 2007, he was a research fellow in Center of Excellence in Visceral Biomechanics and Pain, in Aalborg Denmark from 2006 to 2007. Currently, he is an Assistant Professor at the Graduate School of Advanced Integration Science, Chiba University, Japan. His current research interests include the computer assisted surgery and medical VR training.*

Takao Namiki *received his M.D. from Chiba University, School of Medicine (1985) and his Ph.D. from Chiba University (1993). He is the specialist of cardiology and has also specialized in traditional Japanese–Oriental (Kampo) Medicine. Currently, he is an Associate Professor at the Graduate School of Medicine, Chiba University.*

Yuji Kasahara *received his B.M. and M.D. from Toyama Medical and Pharmaceutical University (1991), his Ph.D. in medicine from Toyama Medical and Pharmaceutical University (2002). Currently he is Associate Professor in Chiba University and also working in Chiba University Hospital as a member of Japanese–Oriental (Kampo) Medicine. His work has focused on new findings of the new cure that used traditional medicine and education of traditional medicine.*

Katsutoshi Terasawa *received his M.D. and Ph.D. from Chiba University, School of Medicine in 1970 and 1979, respectively. He served as the Dean at Toyama Medical and Pharmaceutical University Hospital, and Professor of Department of Japanese Oriental Medicine and Vice-president at Toyama Medical and Pharmaceutical University. Currently he is served as a professor of Department of Japanese–Oriental (Kampo) Medicine, Graduate School of Medicine, Chiba University and President of The Japan Society for Oriental Medicine.*

Yoichi Miyake *was retired from Chiba University in March, 2009 and he became emeritus professor and research professor. He serves now as a chair of Joint Industry-Academia-Government cooperative project on the medical image processing. He is a fellow and honoree member of the IS&T. IS&T and SPIE, named him as the Electronic Imaging Honoree of the Year, 2000. He was served as president of the SPSTJ and vice president of the IS&T.*