Color Correction in Whole Slide Digital Pathology

Yuri Murakami[1], Hikaru Gunji[1], Fumikazu Kimura[1], Masahiro Yamaguchi[1], Yoshiko Yamashita[2], Akira Saito[2], Tokiya Abe[3], Michiie Sakamoto[3], Pinky A. Bautista[4], and Yukako Yagi[4]
[1]Global Scientific Information and Computing Center, Tokyo Institute of Technology; Tokyo, Japan
[2]Medical Solutions Division, NEC Corporation; Tokyo, Japan
[3]Department of Pathology, School of Medicine, Keio University; Tokyo, Japan
[4]Harvard Medical School, Department of Pathology, Massachusetts General Hospital; Boston, MA

Abstract

Whole slide imaging (WSI) is a technique with both the scanning and viewing of digital images of the entire area of a glass pathology slide. The implementation of WSI has started to accelerate, and the usages of WSI are expected for variety of applications such as image analysis, education, conference, and remote diagnosis. However, for implementation of WSI to clinical applications, color variation in digital pathology images is one of the most important issues. The reasons for color variation of whole slide images are mainly in the two processes: staining tissue samples and scanning glass slides. This paper presents the methods and the results of the color correction of whole slide images, with consideration of these two reasons. For color variation caused by scanning devices, the characterization of the whole slide scanner was performed based on a traditional empirical model using a known color chart, where the color chart is a miniature transparent one arranged on a glass slide. In addition, it was also tested that color correction using Hematoxylin and Eosin stained mouse embryo slide instead of the color chart slide. For the color variation caused by staining conditions, the color distribution of images was corrected so as to fit to that of a reference image. Such correction is relatively easy for pathology image, because the color distribution of a stained pathology slide presents a characteristic color distribution decided by its staining method. The effectiveness of the color correction methods were confirmed for whole slide images of tissue samples stained by some typical staining methods.

Introduction

In pathological diagnosis, pathologists examine tissue slides using optical microscope to determine how cell and tissue morphology has changed, so as to yield the final diagnosis and the subsequent treatment policy. Recently, digital imaging technologies have been introduced into the pathological diagnosis. Especially, technologies in whole slide imaging (WSI) have been widely promoted. WSI systems consist of a whole slide scanner and a viewer system; the scanner produces the image of whole tissue slides at high magnification (e.g. x20), and the viewer system enables users to observe the scanned image at arbitrary location and magnification by an intuitive operation. One of the advantages of digital imaging in pathology is applicability of various techniques of digital image analysis: extraction and quantification of histological components, such as nuclei or cytoplasm, or classification of different tissue types [1]. In addition, digital imaging allows internet information sharing for education, diagnostics, publication and research [2].

However, there are still many issues we must solve before implementation of WSI systems in the clinical environment. One

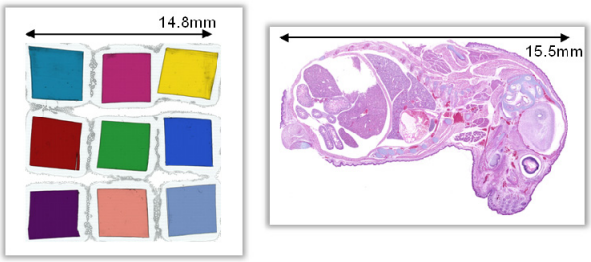

***Figure 1**. A set of calibration slides; MGH color charibration slide (left) and H&E stained mouse embryo slide (right).*

of the most important issues in WSI is the color [3]. By ordinary, tissue sections are stained to enhance its contrast because original tissue sections are almost colorless; Hematoxylin and Eosin (H&E) stain is commonly used for routine diagnostic procedures. Since most staining reactions involve a chemical union between dye and stained substance, the color of stained tissue sections could be determined by staining conditions such as the staining time, the temperature or pH of the solution. However it is difficult to keep the staining conditions constant in practice. In addition, at that stage of image scanning by a WSI scanner, the color of the image is also affected by the characteristics of the scanning system. As a result, the color of pathological images largely varies in the present circumstance. Color variation becomes problem when viewing pathological images on a display devise, compared to when viewing slides directly by microscope. Moreover, image analysis is affected by the color variation; indeed, most software requires the manual instructions to compensate the color variation. For the image analysis to become practical, automated processes based on color correction techniques are strongly required.

A challenging study has been started for standardization and validation of the color in digital pathology [3]. To control the color of digital slides displayed at end users, the scope of the study extend to all phases such as preparing specimen, staining, scanning, viewer system, and display. As a first step, a set of calibration slides have been proposed: one is a color chart slide (referred to as MGH color calibration slide in this paper), and the other is an H&E stained mouse embryo slide (Fig. 1). It has been reported that these slides can be used to check the displayed color visually. However, the evaluation has not been done on the basis of colorimetric data. Another work on color correction of pathological images is the method based on multispectral imaging technologies, which aims to correct the colors of tissue slides at different staining conditions [4]. However, it is difficult to apply

multispectral imaging technologies to WSI under present circumstances.

This paper reports two recent works on color correction of pathological images. One is a color correction for the variation caused by whole slide digital scanners. For this purpose, a traditional empirical model (see for example [5]) using a known color chart was applied and the accuracy was evaluated. The other is a color correction for the variation caused by staining. For this purpose, we propose a color correction method based on a characteristic shape of the color distributions of pathology images.

Color Correction for Scanner Variation

Methods

Input-devise characterization and color correction based on it are a kind of mature technologies in the field of color imaging. We applied a devise characterization method based on a traditional empirical model using a known color chart for whole slide scanners.

As a known color chart for whole slide scanners, the MGH color calibration slide (the left panel of Fig. 1) was used in this examination. The MGH color calibration slide was proposed for standardization of WSI system; nine miniature color films are arranged on a glass plate, and the colors were selected to include typical six colors (cyan, magenta, yellow, red, blue, and green) and the three colors often appeared in H&E stained tissue samples. Since the MGH color calibration slide includes only nine colors, we simply adopted linear model which represents the transformation from the RGB image signals of a whole slide image to XYZ tristimulus values:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \mathbf{M} \begin{pmatrix} 1 \\ R \\ G \\ B \end{pmatrix}, \qquad (1)$$

where **M** is a three-by-four calibration matrix. The matrix **M** can be derived by a least squared method on the basis of nine XYZ-RGB pairs of data of the MGH color calibration slide.

To derive an effective calibration matrix, the selection of training samples is important. In general, it is desirable for a set of training samples to have similar spectral features with those of target objects. From this point of view, we also derive a calibration matrix based on the training data sampled from the H&E stained mouse embryo slide (the right panel of Fig. 1).

The model of Eq. (1) cannot represent the non-linearity in the input-output characteristics of scanners. Then, we made a transparent gray chart which consists of eight grey patches with different absorption density, and used this beforehand to examine the input-output characteristics of scanners. Note that if a color chart with more colors is used, more complicated models can be used instead of Eq. (1), and the input-output nonlinearity will be included in the model.

For both the derivation of the color calibration matrix and the evaluation of the color correction results, we need the actual XYZ tristimulus values of slides. For this purpose, a spectral microscope

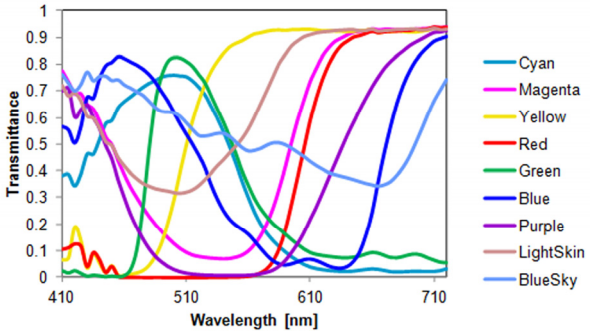

Figure 2. *Spectral transmittance functions of nine color patches of MGH color calibration slide measured by spectral microscopy.*

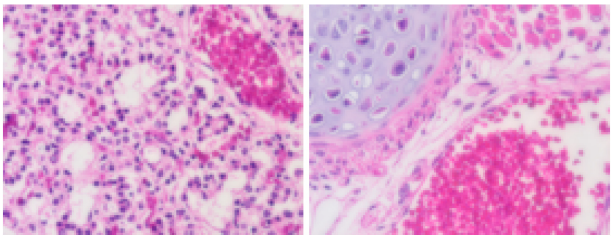

Figure 3. *Two images of H&E stained mouse embryo slide, used for sampling colors for derivation of color calibration matrix.*

system was used, which consists of an optical microscope (BX62; OLYMPUS), a liquid crystal tunable filter (VariSpec™; CRi), and a digital video camera (ORCA-2; Hamamatsu Photonics K.K.). A spectral image acquired by this system has 1344x1024 spatial resolution and 410nm-710nm wavelength range with 10-nm steps. The XYZ values are calculated per pixel from a spectral transmittance image measured by this system using D65 spectrum, which is used as reference images. In the case of mouse embryo slide, the image registration is required between the estimated XYZ image and the RGB image obtained by scanners. Then, the size, location, and rotation were corrected to be aligned each other. However, the resulting images look like different each other because the image by the spectral microscope is slightly-blurred, which is caused by the difference in the modulation transfer functions (MTFs) of the two input devices. To avoid this influence, the color samples are obtained per pixel after resizing the images to one-tenth.

Results

We examined two whole slide digital scanners, NanoZoomer (Hamamatsu Photonics K.K.) and Pannoramic DESK (3DHISTEC Ltd.). First of all, it was confirmed that both scanners have linear input-output characteristics. We derived two color calibration matrices for each scanners based on the MGH color calibration slide and H&E stained mouse embryo slide, respectively.

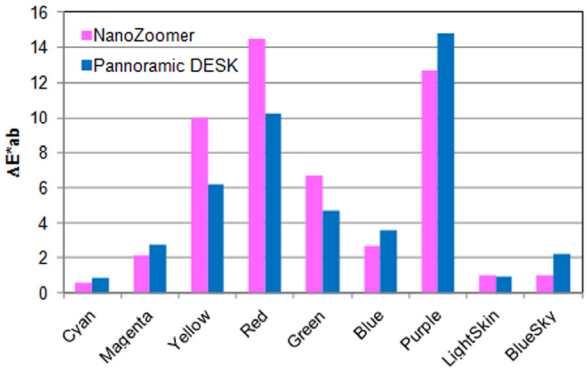

Figure 4. *Color difference between target tristimulus values and estimated tristimulus values from whole slide image of MGH color calibration slide by the calibration matrix derived using MGH color calibration slide itself.*

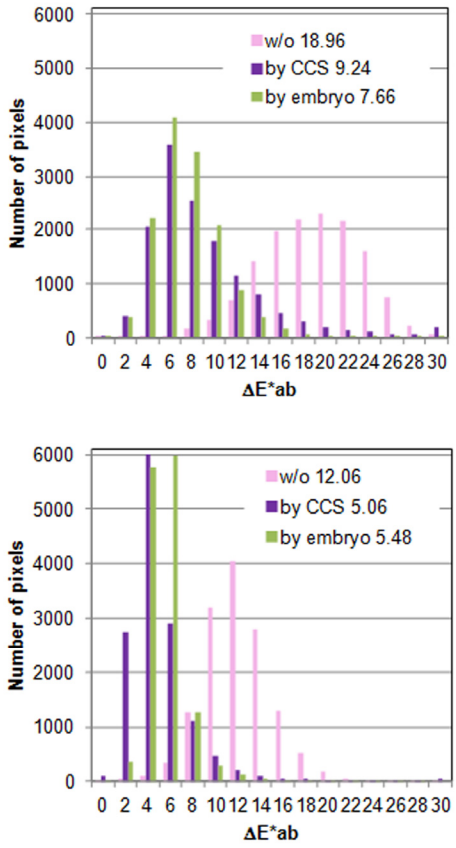

Figure 5. *Comparison of histograms of color differences in images with and without correction for ROI#1 (top) and ROI#2 (bottom). For this comparison, size of the Images are reduced by one tenth from the original. CCS means MGH color calibration slide, and embryo means H&E stained mouse embryo slide. The average of color difference is presented following to respective legend.*

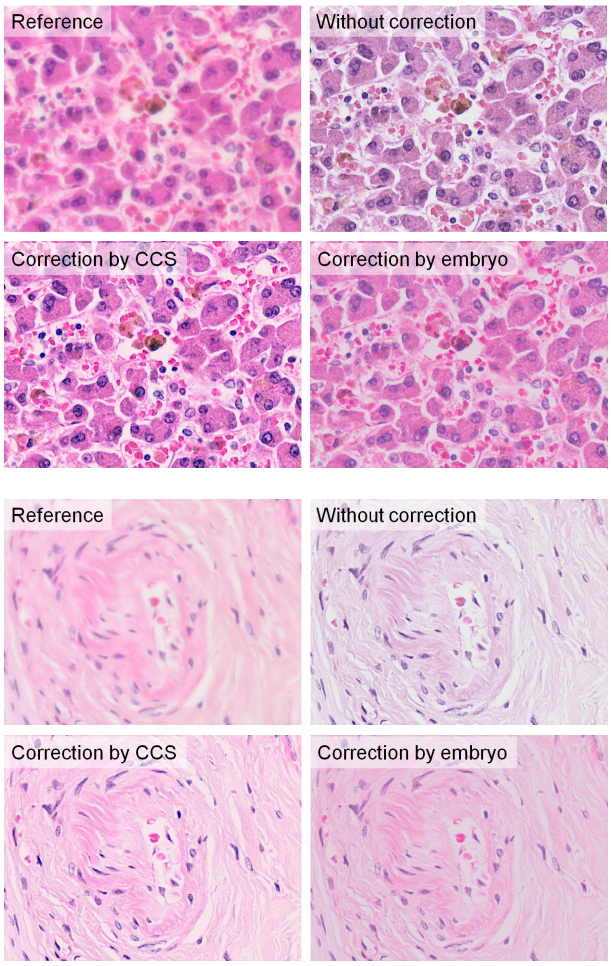

Figure 6. *Color correction results for ROI#1 (top) and ROI#2 (bottom). The images are presented as sRGB-format images. CCS means MGH color calibration slide, and embryo means H&E stained mouse embryo slide.*

Figure 2 shows the spectral transmittance functions, measured by the spectral microscope, of the nine color films of the MGH color calibration slide. Figure 3 shows two image of embryo slide used for acquiring the training data; the presented images are one-tenth resized one for data acquisition. The color calibration matrices were derived based on these data.

For reference, the color of the nine color films themselves were estimated by using the derived matrix based on the MGH color calibration slide, and the color difference in CIE L*a*b color space was calculated (Fig. 4). The average color difference is $\Delta E^*_{ab} = 5.7$ and 5.2 for NanoZommer and Pannoramic DESK, respectively. The tendencies of the error are similar between the two scanners.

Next, an H&E stained liver tissue slide was scanned, and the XYZ image was estimated by applying the color calibration matrices. In order to evaluate the accuracy, the image registration is required between the estimated XYZ image and the reference XYZ image obtained by the spectral microscope. Then, as in the case of embryo slide, the size, location, and rotation were corrected to be aligned each other. Then, to avoid the influence of

the difference of MTFs, the color difference was calculated per pixel after resizing the images to one-tenth.

The results for Pannoramic DESK are presented in this paper. The histograms of the color differences for ROI#1 (mainly cytoplasm) and ROI#2 (mainly fibers) are shown in Fig. 5. In addition, the reference image and the images with and without color correction are shown in Fig. 6. We can confirm that the distribution of the color error becomes smaller and that the color of the corrected image approaches to the reference in both ROIs by either color calibration matrix. In the comparison between the results obtained by the two calibration slides, the embryo slide gives smaller error but it is not significant. This indicates that the color correction based on the color samples from H&E stained tissue samples is promising, but more improvement is required to collecting the sample colors. In addition, we have to say that a simple color correction using the MGH color calibration slide has some efficacy, but not sufficient.

Color Correction for Staining Variation

Color distribution of stained tissue samples

In pathological diagnosis, tissue samples are usually stained to visualize their structures. Certain stains are often combined to reveal more details and features than a single stain alone. For instance, H&E staining uses two kinds of dye, hematoxylin (H) and eosin (E); H stains cell nuclei blue, while E stains cytoplasm, connective tissue and other extracellular substances pink or red. As a result, the colors in the microscopic images of H&E stained tissues present a characteristic distribution. Figure 7 shows the color distribution in RGB color space, where the plots are sampled from a whole slide image of an H&E stained liver tissue sample. This RGB color space is the color space of the whole slide scanner, and the image signals are normalized into 0-1 range. We can see that the colors distribute around the gently-curved surface. Therefore, this distribution can be approximately explained by the H-only color locus, E-only color locus, and their mixtures, as shown in Fig. 8.

Another example is shown in Fig. 9: Elastica van Gieson (EVG) staining. EVG staining uses four kinds of solutions and result in the visualization of black-brown nuclei, black-stained elastic fibers, red collagen and yellow-stained cytoplasm and muscles. We can see a triangle distribution of colors on the bottom panel in Fig.9. Each of three corners of the triangle corresponds to the color of single-dye staining regions. Since the hue of black-brown and yellow are similar, these colors are projected onto the same corner of the triangle. As a result, for EVG staining sample, we can assume the model consists of two triangular pyramids as shown in Fig.10. The sides of the triangular pyramids correspond to the color locus of single-dye stained regions.

Color correction method

The color distribution of stained tissues with several kinds of dyes can be approximately expressed by a relatively simple model as shown in Figs. 8 and 10. In addition, the basic shape of this color distribution is expected to be unchanged, although the colors of an image changes depending on the staining condition. Therefore it is reasonable to think that the color correction is realized by modifying the color distribution of a test image so as to fit to that of reference image.

On the basis of this idea, in this paper, we propose a color correction method for EVG staining tissue samples. As shown in Fig. 10, the color distribution of EVG stained samples is approximately represented by two triangular pyramids. Then, in the proposed method, the color distributions of both test and reference images are estimated by using the pyramid model, and the color distribution of the test image is modified to fit to the reference model. Below, this color correction method is firstly presented. Then, the estimation of the pyramid model is presented later.

If one of the corners of the lower and upper pyramid is assumed to be (0, 0, 0) and (1, 1, 1) respectively, we can define the two pyramids by the three vectors pointed to the remaining three corners of the pyramids from the origin. We call these vectors primary vectors, and represent the primary vectors of the reference image by $\mathbf{p}_1^{ref}$, $\mathbf{p}_2^{ref}$, and $\mathbf{p}_3^{ref}$ and the primary vectors of a test image by $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$.

When a set of RGB image signals of the test image, $(r, g, b)^T$, is located in the lower pyramid, it is represented by the additive mixture of the primary vectors as

$$\begin{pmatrix} r \\ g \\ b \end{pmatrix} = \left(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\right) \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}, \tag{2}$$

where $(w_1, w_2, w_3)^T$, is the weighting coefficients. Based on the relationship of Eq. (2), this set of image signals is warped to the corresponding coordinate in the reference pyramid as

$$\begin{pmatrix} r' \\ g' \\ b' \end{pmatrix} = \left(\mathbf{p}_1^{ref}, \mathbf{p}_2^{ref}, \mathbf{p}_3^{ref}\right) \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}, \tag{3}$$

where $(r', g', b')^T$ is the corrected set of the image signals, and the weighting coefficients $(w_1, w_2, w_3)^T$ are obtained based on Eq. (2) and the inverse matrix of $(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3)$.

When a set of RGB image signals is located in the upper pyramid, the color correction can be done in the same manner by replacing $\mathbf{p}_i^{ref}$ and $\mathbf{p}_i$ by $\mathbf{p}_i^{ref} - (1, 1, 1)$ and $\mathbf{p}_i - (1, 1, 1)$, respectively, and replacing $(r, g, b)^T$ and $(r', g', b')^T$ by $(r, g, b)^T - (1, 1, 1)$ and $(r', g', b')^T - (1, 1, 1)$, respectively. By warping all pixels as described above, the color distribution of the test image will be modified to fit to the reference one.

The pyramid model should be estimated from the color distribution of images. We present here an idea for that; other estimation methods can be used for this process. First, we manually select the area of cytoplasm (light brown area) roughly, and the average vector of the RGB image signals is set to one of the primary vectors. Because the cytoplasm area spread widely, the area selection is easy, which will be automated in the near future. Second, we specify the remaining two primary vectors on the plane perpendicular to the direction of the white vector (1, 1, 1) (the same plane in the bottom panel in Fig. 10. Since the color distribution forms an apparent triangle on this plane, we can easily specify the two corners (top-right and bottom), where the primary vector derived by the cytoplasm area corresponds to the top-left corner of the triangle. Third, the (1, 1, 1) components of the two primary vectors have not decided, which should be determined.

For the primary vector corresponding to the top-right corner, several color samples projected onto near the corner are selected, and the average vector of these color samples is set to the primary

vectors. For the primary vector corresponding to the bottom corner, the average of the (1, 1, 1) components of the two already-obtained primaries is used as the (1, 1, 1) components of this primary.

Results

The proposed color correction method was applied to whole slide images of liver biopsy samples scanned by Nanozoomer. Since all the images were scanned by the same scanner, the color correction for the scanner has not been applied in this experiment.

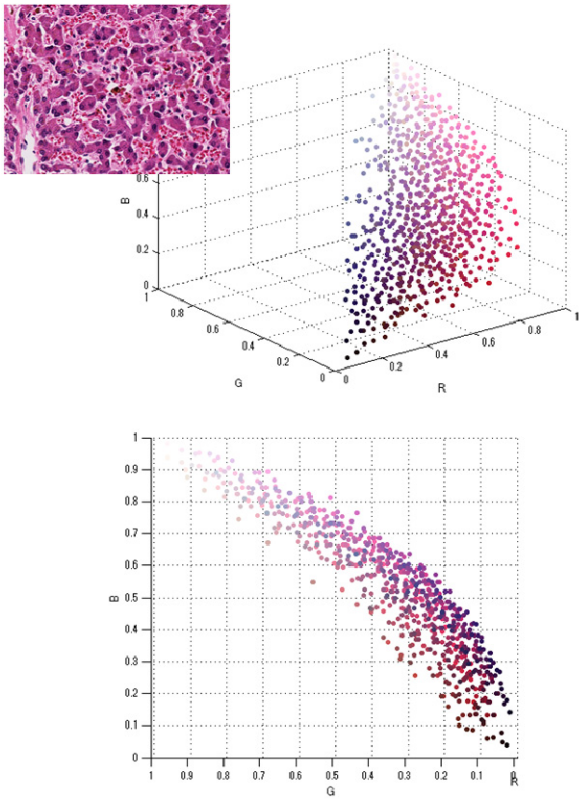

***Figure 7**. Color distribution in RGB color space of H&E stained tissue sample, presented at left-top, from different two perspectives.*

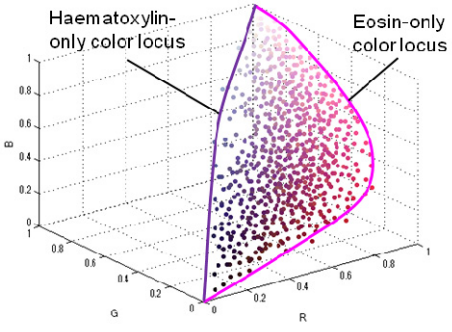

***Figure 8**. An example model of color distribution of H&E stained sample.*

Figure 11 shows the reference image used for the experiment and its color distribution in RGB color space with the estimated pyramid model. We can see that the model well fit to the distribution. Figures 12 and 13 show the test images and their color distributions, before and after color correction, respectively. We can confirm that the color distributions are corrected to fit to the reference model, and the colors in the images are corrected to approach to the reference image also.

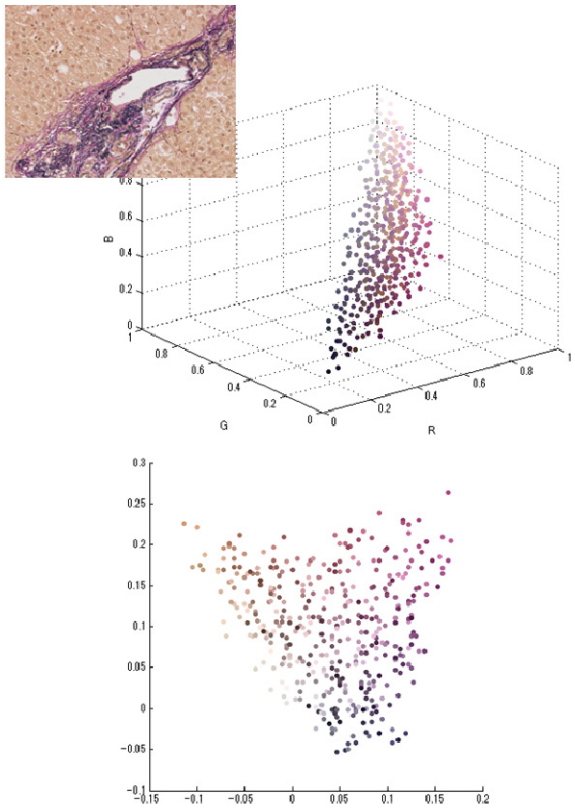

***Figure 9**. Color distribution in RGB color space of EVG stained tissue sample, presented at left-top, from different two perspectives. The bottom panel presents color distribution on a plane perpendicular to the direction of the white vector.*

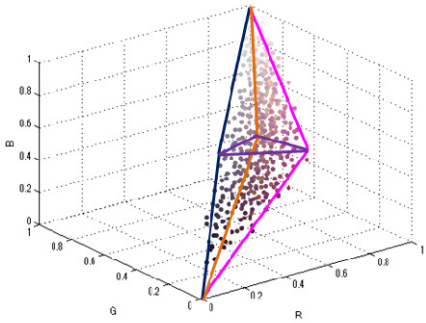

***Figure 10**. Model of color distribution of EVG stained tissue sample, consisting of two triangular pyramids*

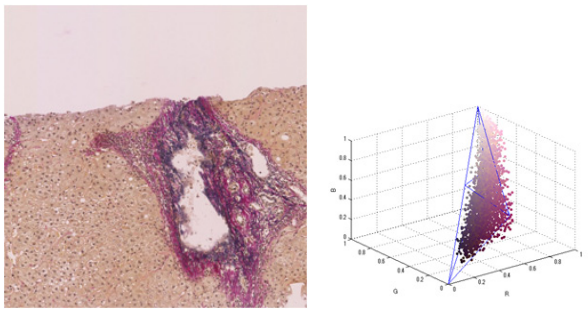

***Figure 11**. Target image (left) and its color distribution with the estimated two triangular pyramids (right).*

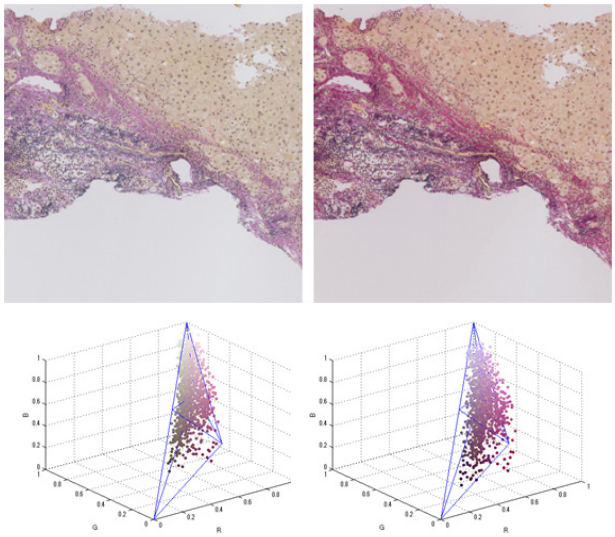

***Figure 12**. Top row shows color images before (left) and after (right) color correction. Bottom row shows color distributions before (left) and after(right) color correction with color distribution model of target image.*

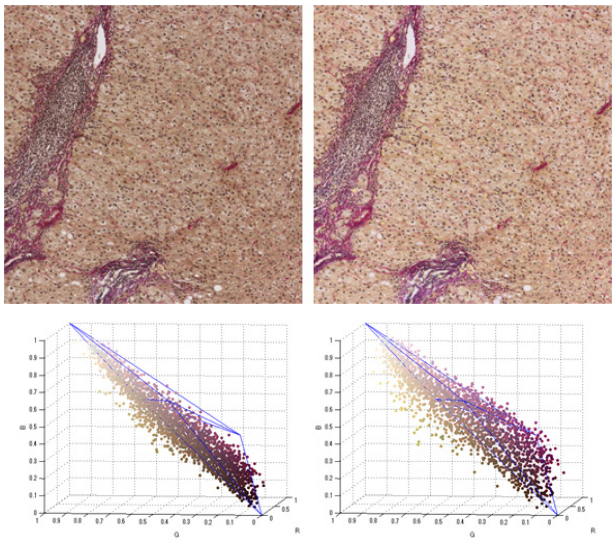

***Figure 13**. Top row shows color images before (left) and after (right) color correction. Bottom row shows color distributions before (left) and after(right) color correction with color distribution model of target image.*

Conclusions

Colorimetric accuracy of the color correction based on the calibration for whole slide digital scanners was examined using the recent-proposed MGH color calibration slide and H&E stained mouse embryo slide. For the quantitative evaluation of the images of H&E stained tissue samples, corresponding reference images with the accurate color information were prepared by using spectral microscope, and the image registration was performed. As a result, the color accuracy of the corrected image of H&E stained liver tissue is not small even by using the embryo slide; ΔE^*_{ab} = 5-8. One of the reasons is that a simple linear model was applied for the color correction. By improvement in the way to correct the sample colors and introducing more complicated/nonlinear model, the accuracy will be improved.

In addition, a color correction method was proposed for the images of the tissue samples stained by several dyes, which is used for compensate the color variation caused by the staining conditions. The method is based on the fact that the color distribution of stained samples is approximately represented by simple models. By taking EVG staining as an example, a color correction method was proposed. The proposed method includes the modeling of the color distribution by two triangular pyramids. Then, the efficacy of the proposed method was tested for EVG stained liver biopsy samples. As a result, it was confirmed that the proposed method appropriately correct the color distributions to fit to the reference one. We expect that color correction methods based on the same concept can be applied to the images of other staining method.

Acknowledgement

This research is supported by New Energy and Industrial Technology Development Organization (NEDO) under the Research and Development Project for pathological image recognition.

References

[1] M. N. Gurcan, L. E. Boucheron, A Can, A. Madabhushi, N. M. Rajpoot, B. Yener, “Histopathological Image Analysis: A Reivew,” IEEE Review in Biomedical Engineering, 2, 147 (2009).

[2] R. S. Weinstein, A. R. Graham, L. C. Richter, G. P. Barker, E. A. Krupinski, A. M. Lopez, Y. Yagi, and J. R. Gilbertson, “Overview of Telepathology, Virtual Microscopy, and Whole Slide Imaging: Prospects for the Future,” Human Pathology, 40 1057 (2009).

[3] Y. Yagi, “Color Standardization and Optimization in Whole Slide Imaging,” Diagn. Pathol., 6 Suppl 1, S15 (2011).

[4] T. Abe, M. Yamaguchi, Y. Murakami, N. Ohyama and Y. Yagi, “Color Correction of Pathological Images Based on Dye Amount Quantification,” Opt. Rev., 12, 293 (2005).

[5] V. Heikkinen, T. Jetsu, J. Parkkinen, M. Hauta-Kasari, T. aaskelainen, and S. D. Lee, “Regularized Learning Framework in the Estimation of Reflectance Spectra from Camera Responses,” J.Opt. Soc. Am. A 24, 2673 (2007).