J Pathol Inform

Editor-in-Chief:
Anil V. Parwani , Pittsburgh, PA, USA
Liron Pantanowitz, Pittsburgh, PA, USA

OPEN ACCESS
HTML format

For entire Editorial Board visit : www.jpathinformatics.org/editorialboard.asp

Original Article

Color correction for automatic fibrosis quantification in liver biopsy specimens

Yuri Murakami, Tokiya Abe[1], Akinori Hashiguchi[1], Masahiro Yamaguchi, Akira Saito[2], Michiie Sakamoto[1]

Global Scientific Information and Computing Center, Tokyo Institute of Technology, Meguro-ku, [1]Department of Pathology, School of Medicine, Keio University, Shinjyuku-ku, [2]Medical Solutions Division, BioMedical Imaging and Informatics Group, NEC Corporation, Minato-ku, Tokyo, Japan

E-mail: *Yuri Murakami - murakami.y.ac@m.titech.ac.jp
*Corresponding author

Received: 30 October 13 Accepted: 04 November 13 Published: 31 December 13

This article may be cited as:
Murakami Y, Abe T, Hashiguchi A, Yamaguchi M, Saito A, Sakamoto M. Color correction for automatic fibrosis quantification in liver biopsy specimens. J Pathol Inform 2013;4:36.

Available FREE in open access from: http://www.jpathinformatics.org/text.asp?2013/4/1/36/124009

Abstract

Context: For a precise and objective quantification of liver fibrosis, quantitative evaluations through image analysis have been utilized. However, manual operations are required in most cases for extracting fiber areas because of color variation included in digital pathology images. **Aims:** The purpose of this research is to propose a color correction method for whole slide images (WSIs) of Elastica van Gieson (EVG) stained liver biopsy tissue specimens and to realize automated operation of image analysis for fibrosis quantification. **Materials and Methods:** Our experimental dataset consisted of 38 WSIs of liver biopsy specimens collected from 38 chronic viral hepatitis patients from multiple medical facilities, stained with EVG and scanned at ×20 using a Nano Zoomer 2.0 HT (Hamamatsu Photonics K.K., Hamamatsu, Japan). Color correction was performed by modifying the color distribution of a target WSI so as to fit to the reference, where the color distribution was modeled by a set of two triangle pyramids. Using color corrected WSIs; fibrosis quantification was performed based on tissue classification analysis. **Statistical Analysis Used:** Spearman's rank correlation coefficients were calculated between liver stiffness measured by transient elastography and median area ratio of collagen fibers calculated based on tissue classification results. **Results:** Statistical analysis results showed a significant correlation r = 0.61-0.68 even when tissue classifiers were trained by using a subset of WSIs, while the correlation coefficients were reduced to r = 0.40-0.50 without color correction. **Conclusions:** Fibrosis quantification accompanied with the proposed color correction method could provide an objective evaluation tool for liver fibrosis, which complements semi-quantitative histologic evaluation systems.

Key words: Color correction, fibrosis quantification, liver biopsy, tissue classification, whole slide image

Access this article online

Website: www.jpathinformatics.org

DOI: 10.4103/2153-3539.124009

Quick Response Code:

INTRODUCTION

Whole slide imaging systems enable not only the digital image management and monitor-based observation, but also the application of digital image analysis: Extraction and quantification of histological components or classification of different tissue types.[1] However, for implementation of digital image analysis to practical applications, color variation in digital pathology images is one of the most important issues.[2]

The reasons for color variation of whole slide images (WSIs) are mainly in the two processes: Staining tissue samples and scanning glass slides. The color of stained tissue samples varies depending on the conditions of preparing and staining tissue as well as the type of stain solution. In addition, color of WSIs varies depending on scanner or scanning condition. Since it is difficult to separate these color variations caused by different process, overall color variation is considered to be corrected in most approaches of color correction. Several color correction methods have been proposed based on the characteristic distribution of the colors in pathology images,[3,4] while there are examples where general color correction method is applied.[5] Color correction methods based on spectral information have been reported,[6,7] which focus on the color variation caused by staining process. A challenging study has been started for standardization and validation of the color in digital pathology,[2] which aims to control the color of digital slides displayed at end users. However, there are very few practical systems of pathology image analysis[8] without any manual operation by introducing color correction techniques.

One of the promising applications of image analysis in pathological diagnosis is quantification of fibers for evaluating liver fibrosis. Evaluation of liver fibrosis in patients with chronic liver disease is crucial for understanding the disease state, predicting prognosis and selecting the appropriate treatment.[9,10] At present, histopathological evaluation of fibrosis is performed by identifying the location, degree and pattern of fibrosis and presence of architectural distortion and regenerative nodule formation. However, this staging includes intra - as well as inter-individual variation. To overcome this problem, introduction of any quantified indices has been suggested for evaluating the degree of liver fibrosis.[11-15] In these studies, fiber areas are extracted from the histological image of the liver biopsy specimen by means of digital image analysis and the area occupied by fibers relative to the area of the entire tissue specimen is quantified as a ratio. The area ratio of fiber shows a correlation with Ishak score,[11-14] hepatic venous pressure gradient,[14] and liver stiffness measured by elastography.[15] However, interactive operations were required for the extraction of fibers in most cases,[11,12,14,15] because of the color variation problem as stated above.

This paper presents the color correction method for Elastica van Gieson (EVG)-stained WSIs of liver biopsy. EVG stain is a widely used staining method to visualize elastic and collagen fibers. However, temporal color change as well as color difference among facilities of EVG-stained samples is larger than H and E-stained samples. Therefore, color correction plays an essential role in identifying and quantifying fiber areas from EVG-stained liver samples by image processing. Though the proposed color correction method was presented previously,[4] this paper reveals the method that enables fully-automated processing. In the experimental evaluation, it is shown that robust classification of tissues into collagen fibers, elastic fibers, cytoplasm and nucleus becomes possible by the application of the proposed color correction. In addition, the correlation between the result of fibrosis quantification and the elastography measurements is evaluated with and without the color correction, using the samples prepared in multiple facilities.

MATERIALS AND METHODS

Slide Preparation and Scanning

The WSIs of EVG-stained liver biopsy used in this paper are same as the images used for the quantification of fibers.[15] Here, slide preparation and scanning process to generate the WSIs are reviewed. Liver biopsy specimens were collected from 38 chronic viral hepatitis patients from 4 medical facilities. These specimens were formalin-fixed, paraffin-embedded and sliced to a thickness of 3 μm at each facility and then, stained with EVG at a single facility. The color of stained tissue samples vary when they are prepared at different facilities under different preparing conditions, even if they are stained at the same facility under the same condition. A WSI of each specimen was acquired using the Nano Zoomer 2.0 HT® at ×20 magnification; the resolution of WSIs is equivalent to 0.46 μm/pixel.

Outline of Color Correction

The proposed color correction method is based on the color distribution of WSIs in RGB color space. The three axes of the RGB color space represent the signal values of three channels of WSIs. In this paper, RGB signals are assumed to be "linear RGB," i.e., the signal value is linear to the power of the transmitted light through a glass slide, though WSIs are often saved as the images consisting of "gamma-corrected RGB" for observation purpose on a monitor. Since "linear RGB" and "gamma-corrected RGB" are convertible with each other outside quantization error, the proposed method can apply to WSIs saved in either color format.

The color distribution of stained tissues forms different shape depending on staining methods. In addition, the shape of the distribution is shifted and/or deformed depending on staining conditions. 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 a reference image. In the proposed method, color correction is supposed to be performed WSI by WSI basis. Below, each part of the color correction is explained.

Color Distribution Model for EVG Stain

Figure 1a shows an example WSI of EVG-stained liver biopsy. Figure 1b shows a magnified part of the WSI

including periportal area. As shown in Figure 1b, EVG staining visualizes collagen fibers by pink-red, elastic fibers by blue-black, cytoplasm by yellow-brown and nuclei by black-brown. Figure 1c shows the distribution of the colors sampled from the periportal area in RGB color space, where the image signals are normalized in the range of 0-1. Since periportal areas contain all the main color components of liver biopsy, the color distribution of this area represents the color distribution of the entire WSIs. Figure 1d shows the distribution of same color samples, which are projected on a plane perpendicular to the direction of the white vector (1, 1, 1). We can see a triangle-shaped distribution of colors on this plane. Each of three corners of the triangle corresponds to the color of single-dye staining; pink-red of collagen fibers, blue-black of elastic fibers and yellow-brown of cytoplasm, whereas black-brown nuclei are projected onto the same corner of cytoplasm because the hue of black-brown and yellow-brown are similar. Based on this distribution, we conceived of making a model to represent the color distribution of EVG-stained tissue samples by two triangular pyramids as shown in Figure 1e. If we make the model represented by such a simple geometrical form, subsequent color correction is easy to be performed. The concrete method of model estimation and color correction are described in the following subsections.

Window Selection for Color Sampling

Color distributions of WSIs are estimated based on the colors sampled from WSIs. Since the number of pixels constituting a WSI is huge (e.g. the WSI in Figure 1a consists of 27,500 × 8,800 = 242,000,000 pixels at ×20), efficient color sampling is required. However, the color sampling from a low-magnification WSI is not appropriate because color mixture among different-color pixels occurs, which makes it difficult to estimate color distribution accurately. Therefore, we select the windows appropriate for color sampling using low-magnification WSIs. Then, color sampling is performed from the window areas of high-magnification WSIs. For this purpose, periportal area is suitable, because periportal areas contain all the main color components of liver biopsy. Other than periportal area, we select the windows from cytoplasm region because cytoplasm occupies over a wide area and it is easy to selectively sample the colors of cytoplasm. Therefore, in the proposed method, two kinds of windows, periportal and cytoplasm, are selected from low-magnification WSIs. Below, the method for selecting respective window is explained.

Ahead of the window selection, the glass slide areas without any tissue components are specified to remove these areas from the analysis. The RGB signal values in glass slide area are higher values compared to the pixels of any tissue components. Therefore, we can specify the glass slide area by simple two-class clustering of the signal values. More specifically, the pixels of WSIs are classified into two classes by K-means algorithm based on RGB image signals and the class with higher RGB signal values are identified by glass slide and labeled by "#Glass."

For window selection of periportal area, the pixels not classified into #Glass are again classified into *K* classes by *K*-means clustering algorithm based on RGB image signals. Finally, small window is scanned over the entire WSI and each window is scored by the function designed for detecting periportal area:

$$\text{Score\_periportal} = \sum_{k=1,..K} \sum_{j=1,..,k} d_{jk} \times \text{Min}\left(N_j, N_k\right) \quad (1)$$

where N_j is the number of the pixels classified into j^{th} class, and d_{jk} is the Euclidian distance between the cluster centers of j^{th} and k^{th} classes in RGB color space. This score is high when widely-separated colors are included at the similar rate in a window, which can be thought as a feature of periportal area. The four highest scored windows are selected as periportal regions.

For the window selection of cytoplasm area, the signal value is inverted for the pixels of #Glass as

$$\begin{aligned} R'(x,y) &= 255 - R(x,y) \text{if} (x,y) \in \#\text{Glass} \\ G'(x,y) &= 255 - G(x,y) \text{if} (x,y) \in \#\text{Glass} \\ B'(x,y) &= 255 - B(x,y) \text{if} (x,y) \in \#\text{Glass} \end{aligned} \quad (2)$$

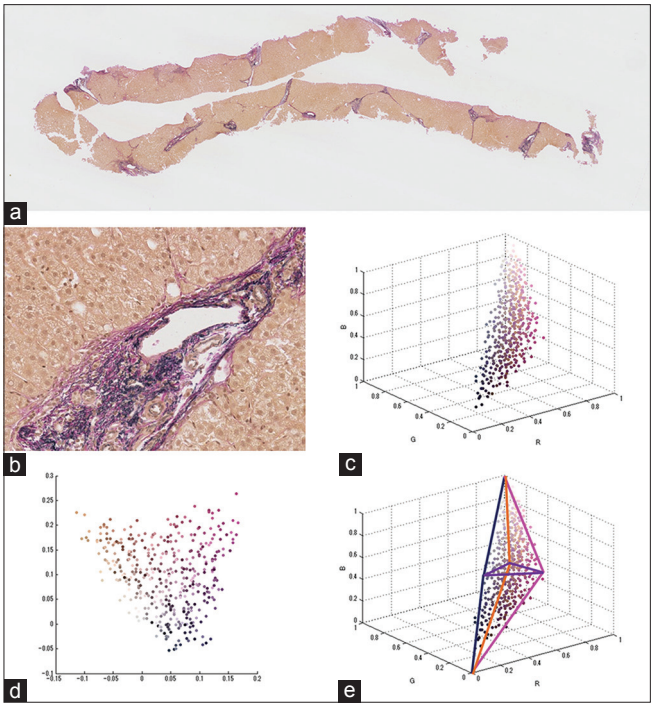

Figure 1: (a) Overview and; (b) magnified view of whole slide image (WSI) of Elastica van Gieson-stained liver biopsy. The color distribution of the WSI (c) in RGB color space; and (d) on a plane perpendicular to white direction; (e) color distribution model represented by two triangular pyramids

After that, small window is scanned over an entire WSI and each window is scored by the function designed for detecting cytoplasm area:

$$\text{Score\_cytoplasm} = \frac{\sigma_R}{m_R} + \frac{\sigma_G}{m_G} + \frac{\sigma_B}{m_B} \quad (3)$$

where m_R, m_G, m_B and σ_R, σ_G, σ_B are the mean and the standard deviation of the signal values of channel image, respectively. This score is low when the average lightness of a window is high with small color variation, which can be thought as a feature of cytoplasm area. In addition, by means of the processing of Equation (2), the score of areas including glass slide becomes high. Thus, the four lowest scored windows are selected as cytoplasm regions, which include less glass slide area.

In the both window selections, it is avoided to select the adjacent windows, in order to select the regions representing the color distribution of the whole area of WSI, but for a part of WSI.

While Balance Adjustment

White balance adjustment is one of the most basic color correction methods. At the first stage of the proposed color correction method, white balance is adjusted, where the mean signal value of glass slide area is used as the signals of white. By means of this process, we can assume the vertex of an upper pyramid representing color distribution is located on the (1, 1, 1) corner.

Estimation of Color Distribution Model

Color distribution model represented by the two triangle pyramids is explained in this section [Figure 2]. If two pyramids share the bottoms and the vertices of the lower and upper pyramids are assumed to be (0, 0, 0) and (1, 1, 1) respectively, we can define the two pyramids by three vectors pointed to the remaining three vertices of the pyramids from the origin. We call these vectors primary vectors; each of primary vectors is labeled by CF (collagen fibers), EF (elastic fibers) and C (cytoplasm); they are represented by three-dimensional column vectors p_{CF}, p_{EF} and p_C, respectively.

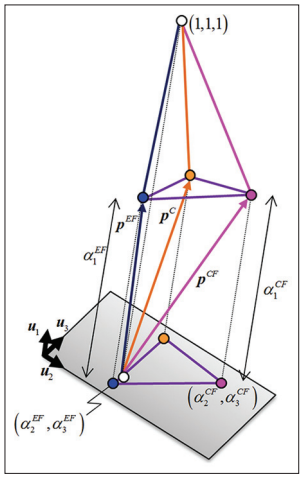

Figure 2: Color distribution model represented by two triangle pyramids. Two pyramids are specified by three primary vectors p^{CF}, p^{EF}, and p^{C}

The primary vector p_C is simply estimated as the average of RGB image signals of all the pixels in the selected windows of cytoplasm. The primary vectors p_{CF} and p_{EF} are estimated from the colors sampled from the selected windows of periportal. To explain how to specify p_{CF} and p_{EF}, let us introduce the following representation:

$$p_{CF} = \alpha_1^{CF} u_1 + \alpha_2^{CF} u_2 + \alpha_3^{CF} u_3 \quad (4)$$

$$p_{EF} = \alpha_1^{EF} u_1 + \alpha_2^{EF} u_2 + \alpha_3^{EF} u_3 \quad (5)$$

where u_1, u_2, u_3 are orthonormal vectors and u_1 is parallel to the direction of white vector (1, 1, 1): $u_1 = 1/\sqrt{3}(1,1,1)^T$. We can say that the plane perpendicular to white [Figure 1c] is spanned by u_2 and u_3. The relationship of these vectors is depicted in [Figure 2]. Since the color distribution forms an apparent triangle on the plane spanned by u_2 and u_3, we can easily specify the two corners (top-right p_{CF} and bottom p_{EF}). That means we can estimate $(\alpha_2^{CF}, \alpha_3^{CF})$ and $(\alpha_2^{EF}, \alpha_3^{EF})$ and the remaining unknowns are α_1^{CF} and α_1^{EF}. One of the unknowns α_1^{CF} is estimated from the color samples c which are projected onto near the point $(\alpha_2^{CF}, \alpha_3^{CF})$ as

$$\alpha_1^{CF} = E\left\{ u_1^T c \,\middle|\, \left(\alpha_2^{CF} - u_2^T c\right)^2 + \left(\alpha_3^{CF} - u_3^T c\right)^2 < Th \right\} \quad (6)$$

where $E\{\}$ represents the averaging operator and Th is a threshold. It is difficult to stably estimate α_1^{EF} by the same approach to α_1^{CF}, because p_{EF} has similar direction with u_1, and most color samples of single-dye staining with various density are projected onto around the point $(\alpha_2^{EF}, \alpha_3^{EF})$. Then, as a practical alternative, we estimate it by the average of the u_1 components of the two already-obtained primaries:

$$\alpha_1^{EF} = \left(\alpha_1^{CF} + \alpha_1^{C}\right)/2 = \left(\alpha_1^{CF} + u_1^T p_C\right)/2 \quad (7)$$

These three primary vectors are used in color correction process explained in the next subsection.

Color Correction of WSIs

The purpose of the color correction is to match the colors of a test WSI with the colors of a reference WSI. For this purpose, the primary vectors of both test and reference images are estimated and the colors of the test image are modified as follows.

Let $\left(p_{CF}^{test}, p_{EF}^{test}, p_{C}^{test}\right)$ be the primary vectors of a test image and $\left(p_{CF}^{ref}, p_{EF}^{ref}, p_{C}^{ref}\right)$ be the primary vectors of a reference image. 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

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \left(p_{CF}^{test}, p_{EF}^{test}, p_{C}^{test}\right) \begin{pmatrix} w_{CF} \\ w_{EF} \\ w_{C} \end{pmatrix} \quad (8)$$

where $\left(w_{CF} \;\; w_{EF} \;\; w_{C}\right)^{T}$ is a set of the weighting coefficients. The weighting coefficients $\left(w_{CF} \;\; w_{EF} \;\; w_{C}\right)^{T}$ are obtained based on Equation (8) and the inverse matrix of $\left(p_{CF}^{test}, p_{EF}^{test}, p_{C}^{test}\right)$. This set of image signals is modified to the corresponding coordinate in the reference pyramid as

$$\begin{pmatrix} R' \\ G' \\ B' \end{pmatrix} = \left(p_{CF}^{ref}, p_{EF}^{ref}, p_{C}^{ref}\right) \begin{pmatrix} w_{CF} \\ w_{EF} \\ w_{C} \end{pmatrix} \quad (9)$$

where $(R', G', B')^{T}$ is the corrected set of the image signals.

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 p_{*}^{ref} and p_{*}^{test} by $p_{*}^{ref} - (1,1,1)^{T}$ and $p_{*}^{test} - (1,1,1)^{T}$ respectively, where * is one of {CF, EF and C}, and by replacing $(R, G, B)^{T}$ and $(R', G', B')^{T}$ by $(R, G, B)^{T}$-$(1, 1, 1)^{T}$ and $(R', G', B')^{T}$-$(1, 1, 1)^{T}$, respectively. By converting all pixels as described above, the color distribution of the test image will be modified to fit to the reference one.

Quantification of Fibrosis

A fibrosis quantification method[15] is applied to the color-corrected WSIs. In this method, every pixel is classified into five classes; collagen fibers, elastic fibers, cytoplasm, nucleus and glass slide. The classification is performed by a quadratic discriminate function based on the color distribution. After classifying all pixels of WSIs, the area ratio of collagen fibers relative to the area of the entire tissue specimen is calculated per 1 mm × 1 mm tiles and the median area ratios of collagen fibers are derived. The reason why median area ratio of small tiles is used in spite of directly calculated area ratio is to avoid the necessity of trimming the pre-existing fibrous tissue.

For tissue classification, it is required to provide training data for five classes; the training data consist of a set of RGB image signals sampled from each tissue component. If there are color variations among the samples used for sampling training data, the classifier would suffer from a reduced performance. In addition, we have to provide training data for every set of tissue samples, if there are color variations among them. Such implementation is not practical. Color correction proposed in this paper can solve these problems and to realize practical implementation.

Experimental Framework

The purpose of the experiments is to examine whether tissue classification followed by fibrosis quantification is succeeded by classifier trained by using WSIs of tissue samples prepared at different facilities by introducing color correction method. For this purpose, we divided WSIs into four data groups, each of which includes the WSIs of the tissue samples prepared at the same facility. Since the samples were prepared at four facilities, there are four data groups named A, B, C and D. The number of WSIs of each data group is 13, 7, 6 and 12 for A, B, C and D, respectively. Color correction was applied to all 38 WSIs, where one of the WSI of data group A was used as the reference WSI. For training and testing purpose, the areas of four tissue components and glass slide area were manually selected from the WSIs, which constituted a labeled dataset. A tissue classifier was trained per the labeled dataset of each data group. As a result, we prepared four kinds of classifiers corresponding to the data groups, referred to as classifier A, B, C and D. After that, tissue classification followed by fiber quantification was performed for every WSI by using four classifiers.

RESULTS

Color Correction

Figure 3 shows example results of window selection for a WSI. Window selection was performed on the WSI downscaled to ×0.8 magnification from the native magnification ×20. The window sizes are 48 × 48 and 12 × 12 pixel for periportal and cytoplasm areas respectively. The selected windows are depicted on the WSI in black and blue for periportal and cytoplasm areas respectively. In addition, the selected regions extracted from WSI at ×20 magnification are shown. We can see that the appropriate area is selected as both periportal and cytoplasm in Figure 3. As shown by the examples, the windows were selected properly for all WSIs used in the experiments.

The results of color correction are shown in Figures 4-6. Figure 4 shows (a) overview, (b) magnified view and (c) the color distribution in RGB color space of the reference WSI. Figures 5 and 6 show two examples of color correction results from data groups C and D in the same manner as Figure 4. The sets of two pyramids in black and red represent the estimated models of the reference and test WSIs respectively. We can see that the corrected color distribution of test WSI is modified to fit to the reference one. As a result, the color tone of the color-corrected WSI is similar to that of the reference WSI.

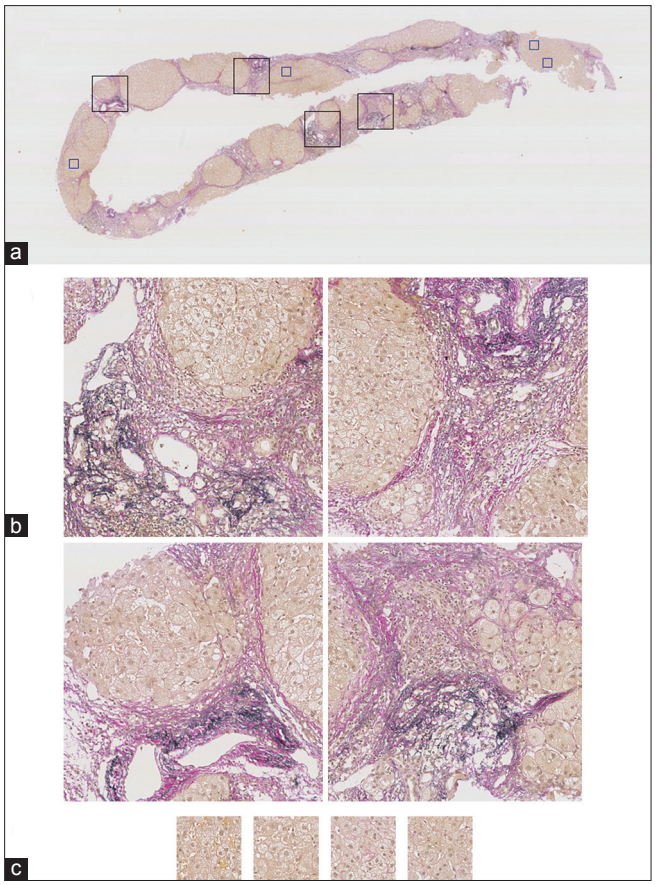

Figure 3: Example results of automatic window selection: (a) Selected windows shown by rectangles on whole slide image in black (periportal) and blue (cytoplasm); (b) four window areas extracted from high-magnification whole slide image for periportal; and (c) cytoplasm

In addition to the visual inspection, we evaluated the color correction results quantitatively. For this purpose, the color gamut of each WSI was calculated and the similarity was measured with the color gamut of the reference WSI.

The color gamut of a WSI is calculated as follows. First, totally 160,000 color samples are randomly selected from the 4 periportal-window images, and a 3-dimensional histogram in RGB color space $h(i_R,i_G,i_B)$ is created from the color samples, where the number of histogram bins is 16 × 16 × 16; i.e. $i_R,i_G,i_B = 1,\cdots,16$. The count of each histogram bin represents the number of color samples that has a color corresponding to the bin. Then, a color gamut $cg(i_R,i_G,i_B)$ is defined by

$$\begin{cases} cg(i_R,i_G,i_B)=1 \; if \; h(i_R,i_G,i_B)\geq Th \\ cg(i_R,i_G,i_B)=0 \; if \; h(i_R,i_G,i_B)<Th \end{cases} \quad (10)$$

where Th = 160000/10000 = 16. The objective of the thresholding is to eliminate the influence of noise and to obtain a gamut stably. The color gamut $cg(i_R,i_G,i_B)$ consists of binary variables which represents inside (1) or outside (0) of the color gamut in a three dimensional RGB color space.

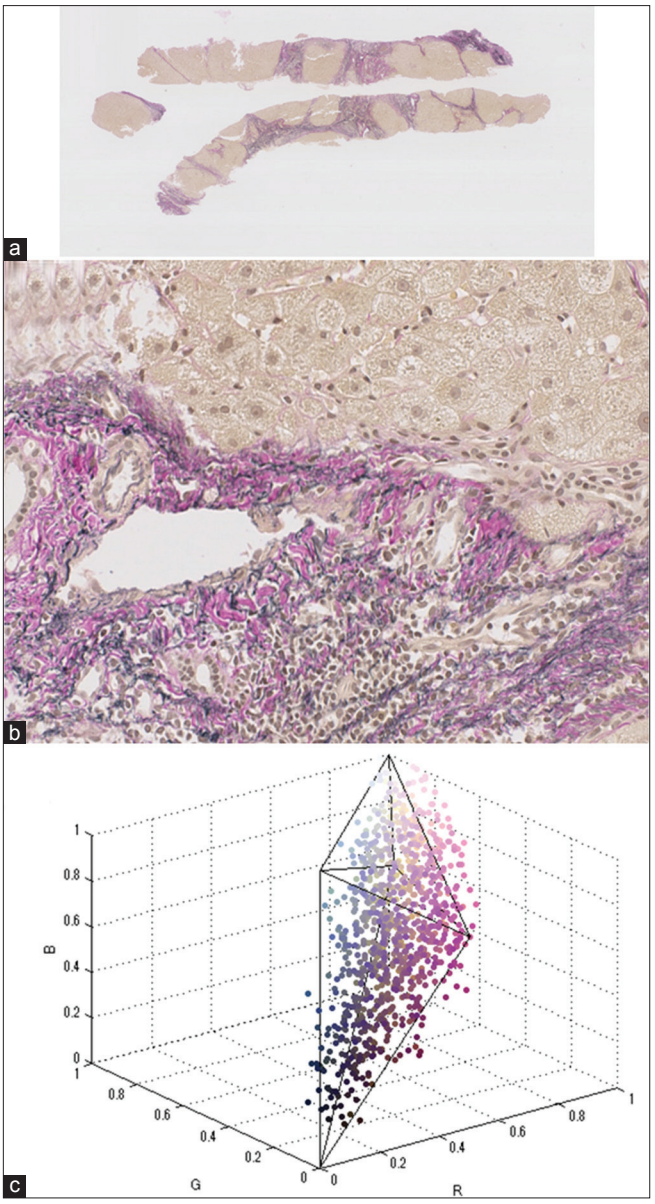

Figure 4: (a) Overview; (b) magnified view; and (c) color distribution of reference whole slide image. Estimated color distribution model is presented by two triangle pyramids in black

Once the color gamut of WSIs are obtained, we can calculate the index to measure the similarity between two color gamuts $cg_I(i_R,i_G,i_B)$ and $cg_{II}(i_R,i_G,i_B)$ as

$$\text{Similarity index} = \frac{\sum_{i_R,i_G,i_B}\begin{bmatrix} cg_I(i_R,i_G,i_B)\ \text{AND} \\ cg_{II}(i_R,i_G,i_B) \end{bmatrix}}{\sum_{i_R,i_G,i_B}\begin{bmatrix} cg_I(i_R,i_G,i_B)\ \text{OR} \\ cg_{II}(i_R,i_G,i_B) \end{bmatrix}} \quad (11)$$

The meaning of the similarity index is explained

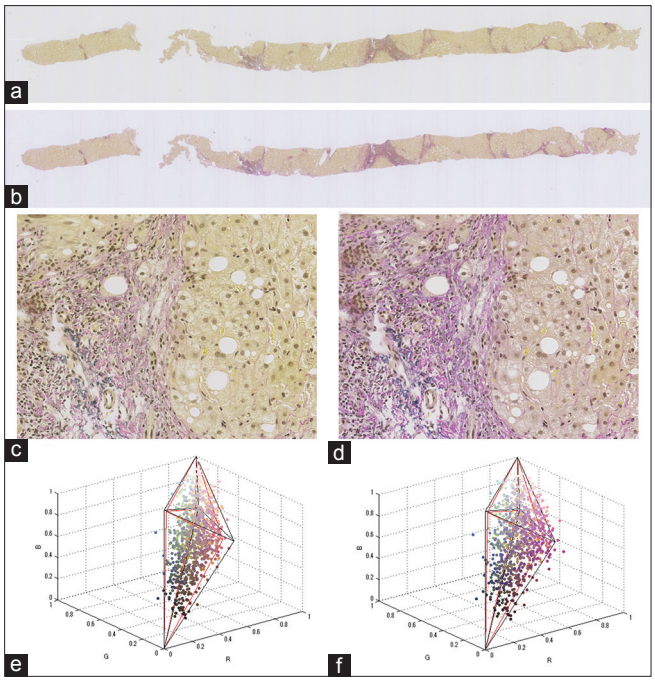

Figure 5: Example color correction result of a test whole slide image of data group C. (a) Overview; (c) magnified view; and (e) color distribution of original WSI; (b) Overview; (d) magnified view; and (f) color distribution of color-corrected WSI. Estimated color distribution models of test and reference WSIs are presented by two triangle pyramids in red and black respectively

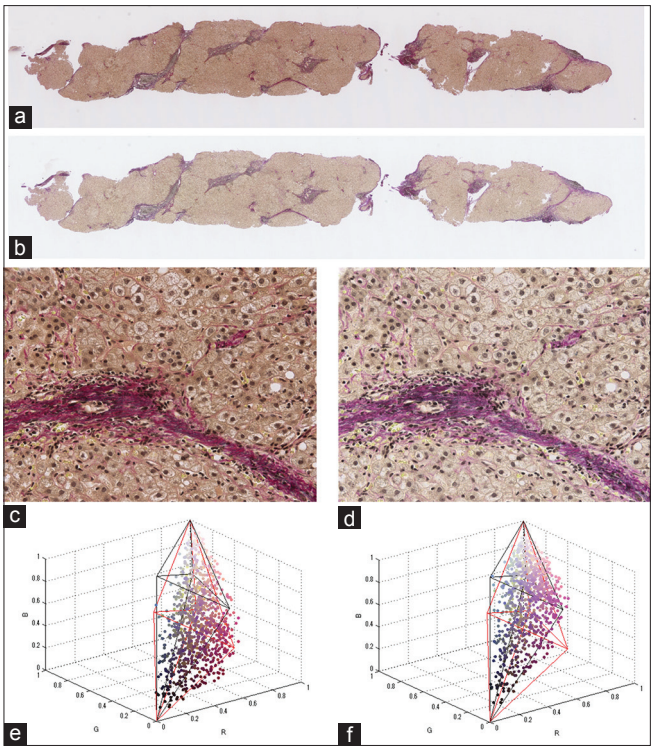

Figure 6: Example color correction result of a test whole slide image of data group D. (a) Overview; (c) magnified view; and (e) color distribution of original WSI; (b) Overview; (d) magnified view; and (f) color distribution of color-corrected WSI. Estimated color distribution models of test and reference WSIs are presented by two triangle pyramids in red and black respectively

using [Figure 7], where a two dimensional color space at $i_G = 8$ is depicted instead of three dimensional space for simplicity. Figures 7a and b show the colors inside of two compared color gamuts $cg_I(i_R, i_G, i_B)$ and $cg_{II}(i_R, i_G, i_B)$, respectively. Figure 7c shows the colors inside of both of the two gamuts by filled circles and the colors inside of either of two by filled and open circles. According to Equation (11), we can find that the similarity index represent the ratio between the number of filled circles to the number of filled and open circles. The similarity index is increasing as overlapping area is increasing; the index ranges from 0 to 1.

The similarity indices were calculated for the 37 test WSIs without any correction (WO), with only white balance adjustment (WB) and with color correction (CC), where one of the two compared gamuts was fixed to the color gamut of the reference WSI. Figure 8 shows the box plots of the similarity indices. The median of each series are 0.54, 0.57, and 0.68 for WO, WB and CC, respectively. The differences of the median were evaluated by the Wilcoxon Rank Sum test, a non-parametric test. As a result, there were significant differences between WO versus CC ($P < 0.01$) and WB versus CC ($P < 0.01$), while there was no significant difference between WO versus WB. From these results, it was confirmed that the color correction has a significant effect to modify the color distribution of test WSIs to match to the reference WSI. However, what matter is that color-corrected WSIs are applicable to automated image analysis; the examination from this viewpoint is presented in the next subsection.

Tissue Classification

In the classification experiments, other than the images corrected by the proposed color correction method, we prepared the images corrected by only white balance adjustment for comparison; i.e. CC and WB. The WSIs without any correction were not used for comparison because the color variation of glass slide area becomes a crucial problem. In the case of WB, white-balanced WSIs were used for both training and testing of tissue classification.

Tissue classification was performed for all 38 WSIs by using four kinds of classifiers. Figure 9 shows an example result of tissue classification for a WSI in data group D by classifier A. In the classification results, collagen fibers, elastic fibers, cytoplasm, nucleus and glass slide area are shown by red, dark blue, green, yellow and light blue respectively. Some collagen fibers are misclassified as cytoplasm or nucleus in the case of WB because the color of collagen fibers is very faint in this tissue sample. As a result, in this case, it is presumed that quantified fiber area from the WSI without color correction becomes incorrect. On the other hand, we do not see

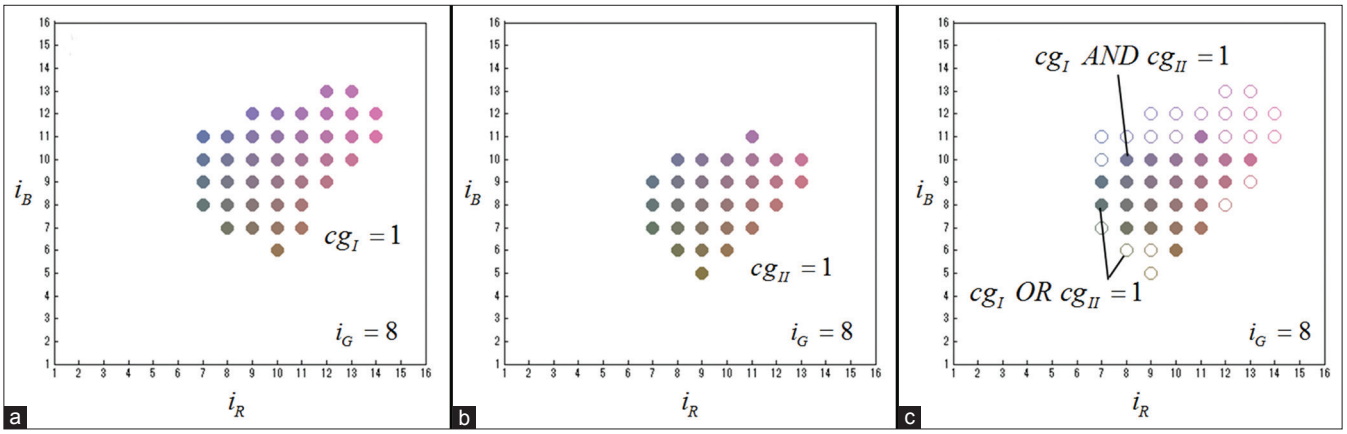

Figure 7: Conceptual diagram of similarity index calculation: (a and b) Respective color gamut of compared two whole slide images; and (c) overlapping gamut (filled circles) compared to conjunction of two gamuts (filled and open circles). Every gamut is presented in two-dimensional space for explanation purpose

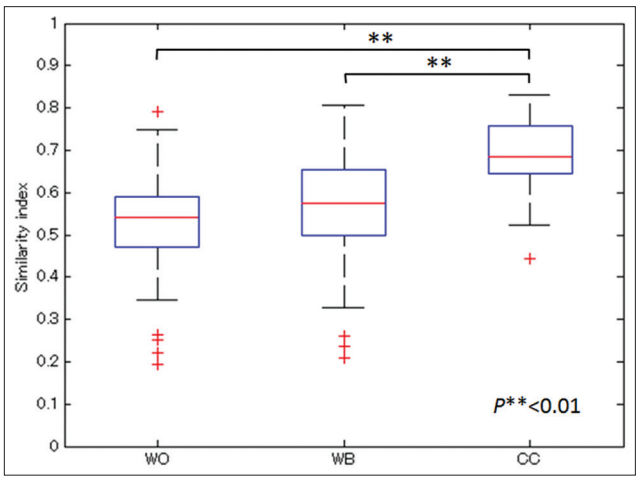

Figure 8: Box plots of similarity indices of color gamuts for whole slide image without any color correction, with only white balance adjustment, and with color correction where one of the two compared gamuts was fixed to the color gamut of the reference WSI

such apparent misclassification in the results with color correction.

To evaluate the classification results quantitatively, classification accuracy of each of four tissue component per each WSI for each classifier were calculated based on the labeled dataset. The results were divided into two groups: in the case that classified WSIs belong to the same data group as the classifier (intra-facility test) and the case that classified WSIs belong to the different data group as the classifier (inter-facility test). The box plots of the accuracy are shown in Figure 10, where each panel shows the results related to each tissue class: collagen fibers (CF), elastic fibers (EF), cytoplasm (C), and nucleus (N). We evaluated the difference of median between the accuracy of CC and WB by the Wilcoxon Rank Sum test. As a result, first of all, there is no significant difference in every pair in the intra-facility case. As for inter-facility case, though there is no significant difference in the results of the class CF and EF, we can see significant differences between CC and WB in the results of the class C ($P < 0.01$) and N ($P < 0.01$); the median of the accuracy of CC (0.93 and 0.85 for C and N, respectively) is significant higher than that of WB (0.91 and 0.80 for C and N, respectively). These results indicate that CC improves the classification accuracy compared to WB when the WSIs from different facilities are treated.

Quantification of Fibrosis and Relationship with Liver Stiffness

Based on the tissue classification results of WSIs, we can quantify the degree of fibrosis as a median area ratio of collagen fibers.[15] Since there are four classification results for a WSI by different classifiers of A, B, C, and D, four quantified values were derived for every WSI. As Abe *et al.* presented,[15] there is a correlation between the median area ratio of collagen fibers and the stiffness of the liver measured by transient elastography (FibroScan®; EchoSens, France). To evaluate the results of the proposed color correction followed by tissue classification, the correlation with the liver stiffness was derived based on our classification results. Figure 11 shows the scatter plot of liver stiffness versus median area ratio of collagen fibers. The Spearman's rank correlation coefficients for respective classifiers A, B, C and D are shown in Table 1. The results show that there is a significant correlation (r = 0.61-0.68) when color correction was introduced, while slightly lower correlation (r = 0.40-0.50) with only WB adjustment. From this result, we can say that the proposed method can ameliorate the color variations from the WSIs among the tissue samples prepared at different facilities, which results in the applicability of tissue classification analysis over the samples prepared by different facilities without any manual operation. The correlation coefficients were slightly reduced compared with the previous report (r = 0.73 [$P < 0.01$]).[15] We should examine whether there is a significant difference between the results by using a larger collection of data.

CONCLUSION

This paper proposes a color correction method for WSIs of EVG-staining liver biopsy. In this method, a color distribution represented by a set of two pyramids is estimated from the sampled colors from a WSI and color correction is performed by modifying the color distribution of a target WSI so as to fit to that of a reference WSI. For practical and accurate color sampling, we propose a method to select the windows of periportal and cytoplasm

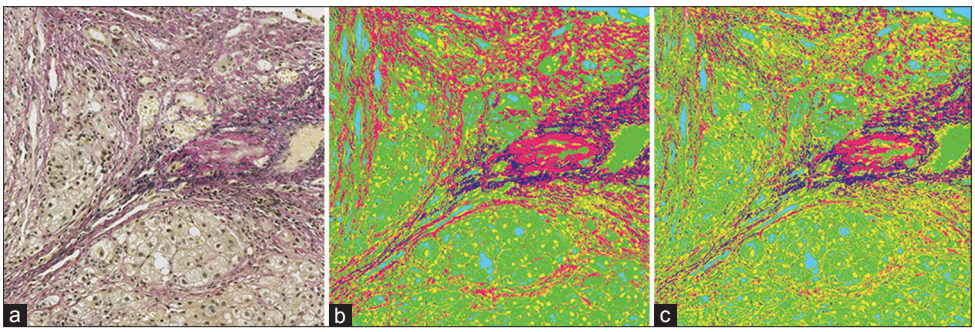

Figure 9: (a) Color image; (b) tissue classification result by classifier A with color correction; and (c) with only white balance adjustment of part of the whole slide image in data group D. In tissue classification results, collagen fibers, elastic fibers, cytoplasm, nucleus and glass slide were in red, dark blue, green, yellow and light blue

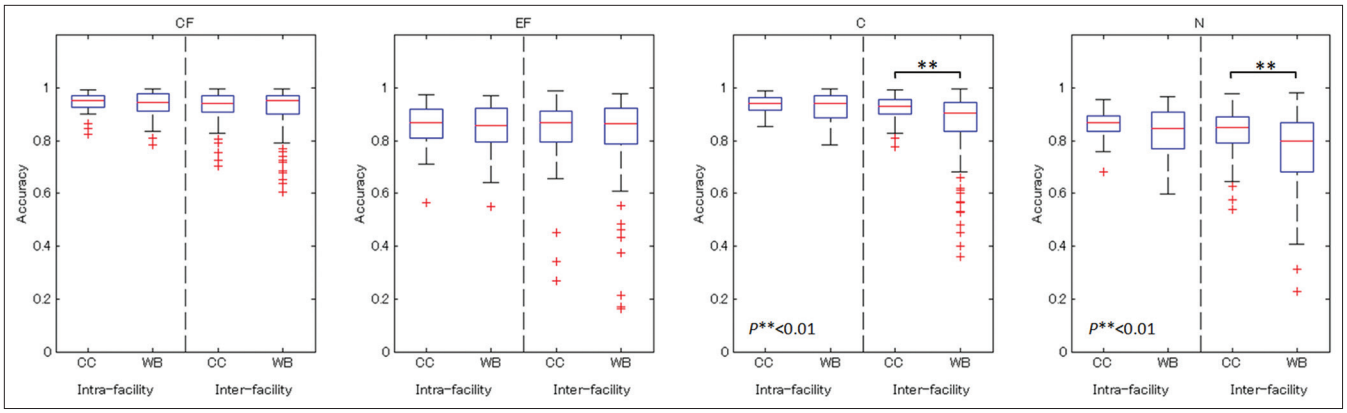

Figure 10: Box plots of classification accuracy of collagen fibers (CF), elastic fibers (EF), cytoplasm (C) and nucleus (N), calculated based on labeled dataset extracted from color-corrected and white-balanced whole slide images in inter-and intra-facility analysis

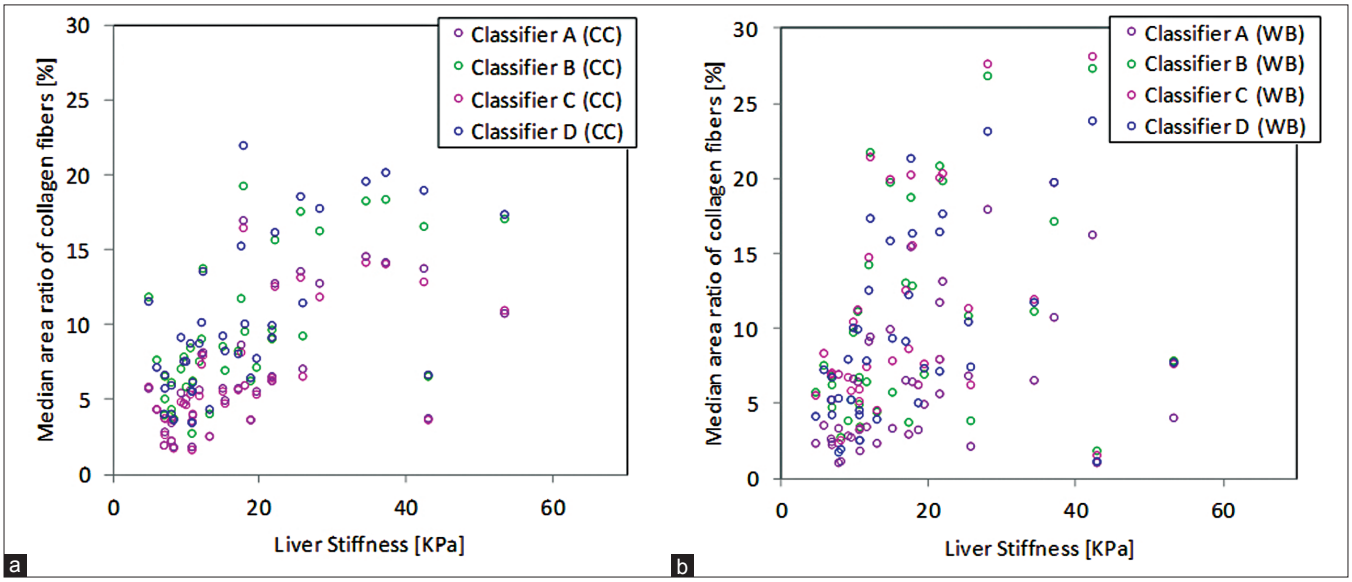

Figure 11: Scatter plots of liver stiffness measured by transient elastography versus median area ratios of collagen fibers (a) with color correction; and (b) with only white balance adjustment

Table 1: Spearman's rank correlation coefficients between median area ratio of collagen fiber and liver stiffness measured by transient elastography

	Classifier A	Classifier B	Classifier C	Classifier D
Color correction	0.680**	0.633**	0.667**	0.636**
White balance	0.486**	0.399*	0.460**	0.502**

**P<0.01, *P<0.05

areas. As a result of the quantitative evaluation of color correction results, it was confirmed that the similarity of the color distribution between test and reference WSIs significantly increased by applying color correction. In addition, we performed tissue classification based on the color information followed by fibrosis quantification using the color-corrected WSIs. Tissue classification results were evaluated based on the labeled dataset. As a result, it was confirmed that the classification accuracy of color-corrected WSIs were significantly higher compared to white-balanced WSIs for the two classes of cytoplasm and nucleus in the case of inter-facility analysis. In addition, evaluation of fibrosis quantification showed that there was a significant correlation between the quantified value calculated from color-corrected WSIs and liver stiffness (r = 0.61-0.68 [$P < 0.01$]). The correlation coefficient was increased from the coefficient calculated from the WSIs corrected by white balance only (r = 0.40-50 [$P < 0.01$]).

Although the evaluation using a larger set of WSIs is required in the future, there is a possibility that fibrosis quantification accompanied with the proposed color correction method provides an objective evaluation tool for liver fibrosis, which complements semi-quantitative histologic evaluation systems. The 38 samples used in the experiments were prepared in four facilities, but stained at the same facility. Therefore, the color of WSIs might largely vary when they are stained at different facilities. The extent to which the proposed color correction method has effect on such data set should be examined in the future. In addition, the quantification of the degree of fibrosis is required not only for liver, but also other organs. Therefore, our future works also include the investigation of the applicability of the proposed color correction methods to EVG-stained tissue samples of other organs, such as kidney and lung.

REFERENCES

1. Gurcan MN, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B. Histopathological image analysis: A review. IEEE Rev Biomed Eng 2009;2:147-71.
2. Yagi Y. Color standardization and optimization in whole slide imaging. Diagn Pathol 2011;6 Suppl 1:S15.
3. Macenko M, Niethammer M, Marron JS, Borland D, Woosley JT, Guan X, *et al*. A method for normalizing histology slides for quantitative analysis. Proceedings of the 6th IEEE International Conference on Symposium on Biomedical Imaging: From Nano to Macro; 2009 Jun 28-1; Boston, Massachusetts. NJ: IEEE Press Piscataway; 2009.
4. Murakami Y, Gunji H, Kimura F, Yamaguchi M, Yamashita Y, Saito A, *et al*. Color correction in whole slide digital pathology. Proceedings of the 20th Color and Imaging Conference; 2012 Nov 12-16. Los Angeles, CAL: IS and T, 2012.
5. Magee D, Treanor D, Crellin D, Shires M, Smith K, Mohee K, *et al*. Colour normalisation in digital histopathology images. Proceedings of Optical Tissue Image Analysis in Microscopy, Histopathology and Endoscopy (MICCAI Workshop); 2009 Sep 20-24. London, UK: Daniel Elson, 2009.
6. Tani S, Fukunaga Y, Shimizu S, Fukunishi M, Ishii K, Tamiya K. Color standardization method and system for whole slide imaging based on spectral sensing. Anal Cell Pathol (Amst) 2012;35:107-15.
7. Abe T, Yamaguchi M, Murakami Y, Ohyama N, Yagi Y. Color correction of pathological images based on dye amount quantification. Opt Rev 2005;12:293-300.
8. Saito A, Cosatto E, Kiyuna T, Sakamoto M. Dawn of the digital diagnosis assisting system, can it open a new age for pathology? Proceedings of SPIE8676 Medical imaging 2013: Digital Pathology; 2013 Feb 9. Lake Buena Vista. FL: SPIE, 2013.
9. Intraobserver and interobserver variations in liver biopsy interpretation in patients with chronic hepatitis C. The French METAVIR Cooperative Study Group. Hepatology 1994;20:15-20.
10. Desmet VJ. Knodell RG, Ishak KG, Black WC, Chen TS, Craig R, Kaplowitz N, Kiernan TW, Wollman J. Formulation and application of a numerical scoring system for assessing histological activity in asymptomatic chronic active hepatitis [Hepatology 1981;1:431-435]. J Hepatol 2003;38:382-6.
11. Hui AY, Liew CT, Go MY, Chim AM, Chan HL, Leung NW, *et al*. Quantitative assessment of fibrosis in liver biopsies from patients with chronic hepatitis B. Liver Int 2004;24:611-8.
12. Lazzarini AL, Levine RA, Ploutz-Snyder RJ, Sanderson SO. Advances in digital quantification technique enhance discrimination between mild and advanced liver fibrosis in chronic hepatitis C. Liver Int 2005;25:1142-9.
13. Goodman ZD, Stoddard AM, Bonkovsky HL, Fontana RJ, Ghany MG, Morgan TR, *et al*. Fibrosis progression in chronic hepatitis C: Morphometric image analysis in the HALT-C trial. Hepatology 2009;50:1738-49.
14. Calvaruso V, Burroughs AK, Standish R, Manousou P, Grillo F, Leandro G, *et al*. Computer-assisted image analysis of liver collagen: Relationship to Ishak scoring and hepatic venous pressure gradient. Hepatology 2009;49:1236-44.
15. Abe T, Hashiguchi A, Yamazaki K, Ebinuma H, Saito H, Kumada H, *et al*. Quantification of collagen and elastic fibers using whole-slide images of liver biopsy specimens. Pathol Int 2013;63:305-10.