# Digital Staining of Pathological Tissue Specimens with PCA-based Feature Extraction and Linear Mapping of Spectral Transmittance

*Pinky A. Bautista[1], Tokiya Abe[1], Masahiro Yamaguchi[1,4], Yukako Yagi[3] and Nagaaki Ohyama[2]*
[1] *Tokyo Institute of Technology, Imaging Science & Engineering Laboratory, Yokohama, Japan;*
[2] *University of Pittsburgh Medical Center, Pittsburgh Pennsylvania, USA;*
[3] *Tokyo Institute of Technology, Frontier Collaborative Research Center, Yokohama, Japan;*
[4] *Akasaka Natural Vision Research Center, NICT, Tokyo Japan*

## Abstract

*Color is an important parameter in pathological diagnosis and it is the very reason why tissue samples are stained, since the morphological structure of the tissue components can only be vividly observed when they are colored differently. Common for routine staining is the Hematoxylin and Eosin(HE) dyes, however to assess diseases related to the condition of the fibrosis, Masson-trichrome (MT) dyes are used instead. The digital transformation of an HE-stained image to its MT-stained (digital staining) equivalent has already been proposed, where the information derived from the 16-band images of the stained specimens were utilized. In this paper we addressed the possible reduction of the spectral dimension requirement to implement the proposed digital staining procedure. To find for the effective spectral dimension, from the classification point of view, principal component analysis (PCA) was applied independently to the five statistical descriptors of the tissue components transmittance spectra, i.e. mean, maximum, minimum, range and standard deviation. In our initial experiments with liver tissue specimens, it was found that ten principal components can be effective to implement the digital staining scheme.*

## Introduction

Dyes are applied to pathological tissue specimens such that pathologists can perform a diagnosis about the particularities of the patient's disease. Once tissue specimens are stained, the different pathological structures composing the tissue can be readily examined. Depending on the pathologist initial findings about the patient's ailment, specific tissue structures are examined to yield a conclusive diagnosis. For diseases related to the condition of the fibrosis, Masson-trichrome dyes are used. These dyes have an effect of transforming the fiber region into blue, the cytoplasm to shades of red and the nucleus to black.

Tissue specimens are observed under a microscope to examine any changes in their morphological structures. With the recent development in filter technology it is now possible to acquire microscopic images at *N* different wavelengths, *N>3,* where the resulting image is referred as multispectral image.[1] A multispectral image carries more information compared to grey-level or RGB images, and reports regarding the capability of a multispectral microscopic imaging system to enhance medical image analysis, and to deal with problems such as dye amount standardization, and digital staining of pathological tissue specimens can be found in literatures.[2-8]

Digital staining is the transformation of an image into an image of desired color. A digital staining methodology for pathological tissue specimens, i.e. the transformation of an HE-stained specimen to its MT-stained equivalent, has already been proposed.[7-8] The method utilizes the information from the 16-band multispectral images of a tissue specimen. In this paper, we incorporate feature extraction scheme, i.e. PCA, in the implementation of the digital staining procedure proposed in Ref. [8] to reduce the transmittance feature dimension. The general framework of the current digital staining methodology is shown in Fig.1. First, tissue components, which are emphasized in the MT-stained tissue slide, are identified from the HE-stained tissue slide, and then classifiers for each of the components' transmittance spectra feature calculated over a 5 × 5 block pixels: mean, maximum, minimum, range and standard deviation, are built. Classification is undertaken on the PCA-extracted features to select the appropriate 16 × 16 transformation matrices, which are devised from the linear mapping of specific sets of spectral transmittance. These matrices convert the transmittance spectra of the classified HE-stained pixels to their MT-stained configurations. The transformed transmittance spectra are finally converted to their corresponding RGB values to provide a perceptible digitally-stained image. Generally, the digital staining procedure can be done in three steps: (i) classification of the various tissue components; (ii) transformation 16-band transmittance spectra of the classified multispectral pixels to their desired transmittance configuration; and (iii) the visualization of the 16-band digitally stained image in the RGB color space.

Incorporating digital processing techniques to pathological diagnosis, such as in the area of staining of tissue slides, would facilitate faster implementation of the diagnosis, and would also provide convenient way for information exchange across the internet, especially that in today's era of information technology, the very ultimate goal is to provide and access information regardless of regional location.

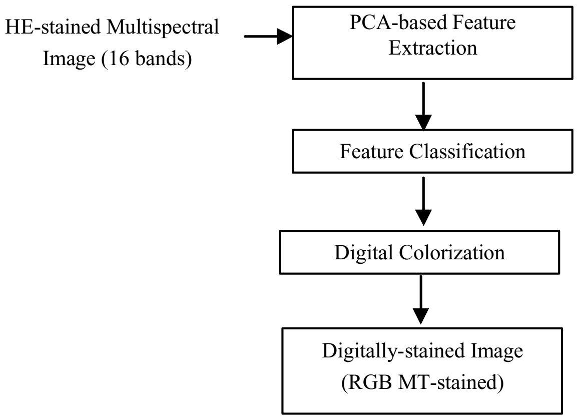

HE-stained Multispectral Image (16 bands)
PCA-based Feature Extraction
Feature Classification
Digital Colorization
Digitally-stained Image (RGB MT-stained)

*Figure 1. General framework of the digital staining procedure.*

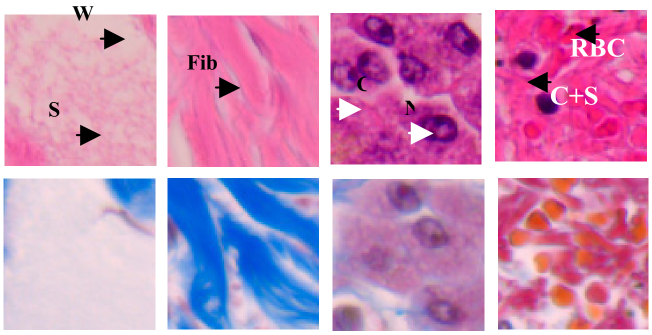


*Figure 2. RGB images of the different tissue components. Top: HE stained; Bottom : MT stained. W–white region; Fib- fibrosis; C- cytoplasm; N-Nucleus; S-serum; C+S –cytoplasm with serum; RBC–red blood cell: Fib-fibrosis*

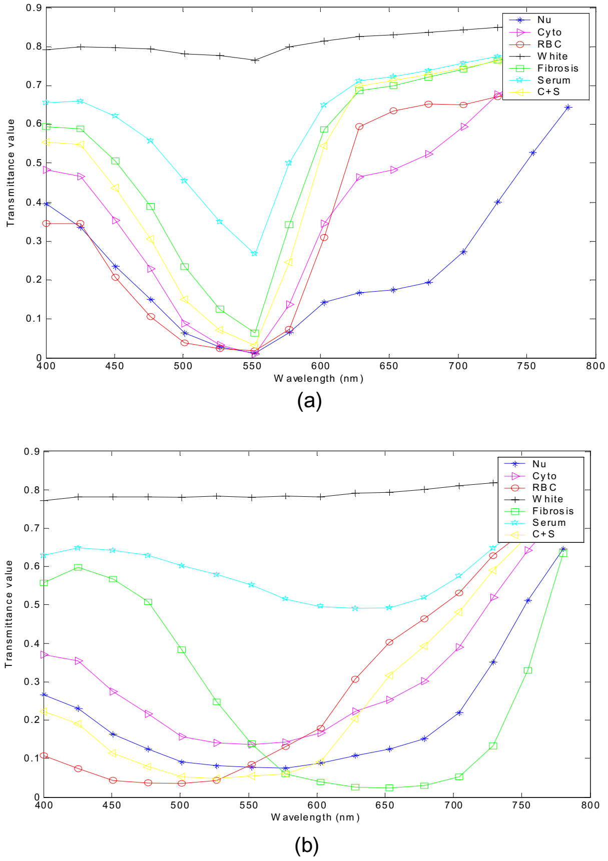
## Image Acquisition

The microscopic multispectral camera developed by Akasaka Natural Vision Research Center, Japan[9] was used to capture the HE and MT-stained images of a liver tissue specimen. Since the HE and MT-stained specimens were extracted from the serial sections of the tissue, the general histological structure of the specimens are similar. Information regarding the tissue specimen within the visible spectrum range, i.e. 400 nm to 700 nm can be obtained from the captured 16-band multispectral images of size 2000 × 2000 pixels.

For the current experiment, the subjects are the liver tissue slides prepared by the National Cancer Center (NCC) of Japan. Specifically, we have one pair of HE and MT-stained slides, and from which we captured six pairs of HE and MT-stained images. From these images the following tissue components were identified: Nucleus, Cytoplasm, red blood cell (RBC), white region, fibrosis; and components which are associated to the type of protein abundant in them: Serum and Cytoplasm with Serum. The magnified images of these components are shown in Fig. 2. Figure 3, on the other hand, illustrates the components average transmittance over a 5 × 5 block

## Feature Parameters

The spectral transmittance or reflectance configuration of a tissue component reflects, to certain extent, the component's structural and biochemical composition,[10,11] That is, different tissue structures exhibit variability in their transmittance spectra configurations. Owing to this fact, various applications have been implemented utilizing the information contained in the transmittance spectra of the various tissue components.[5,6,10,12] Equation (1) is the expression used to estimate the transmittance of a single pixel, where $t(x,y,\lambda)$ refers to the spectral transmittance of the selected component at location $x,y$; $I_o(\mathrm{x},\mathrm{y},\lambda)$ is the intensity signal of the object image; $I_{ref}$ is the reference signal; $I_{do}(x,y,\lambda)$ and $I_{dref}(x,y,\lambda)$ refer respectively to the dark current signals of the reference image and the object image. The reference image is acquired by imaging a slide with no tissue sample on it, while the dark current image is obtained with no illumination of the CCD camera.



(a)

(b)

*Figure 3. Transmittance-spectra of the different tissue components. (a) HE stained; (b) MT-stained*

$$t(x,y,\lambda) = \frac{I_o(x,y,\lambda)-I_{do}(x,y,\lambda)}{I_{ref}(x,y,\lambda)-I_{dref}(x,y,\lambda)} \qquad (1)$$

Neighboring pixels are said to be correlated. Consideration of the contextual information of an image pixel may therefore enhance the classifier performance as shown in Ref. [3]. The simplest way to exploit the spatial context of an image pixel is to consider its statistical behavior within a defined window. If we let $S_{xy\lambda_k}$ represents the set of coordinates in a rectangular sub-image window of size $uxv$ at wavelength $\lambda_k$ and centered at point $(x,y)$, we have the following statistical descriptors, which are considered in this work.

$$\text{Mean: } t(x,y,\lambda_k)_{\text{mean}} = \frac{1}{uv} \sum_{(s,t,\lambda_k) \in S_{xy\lambda}} t(s,t,\lambda_k). \tag{2}$$

$$\text{Max: } t(x,y,\lambda_k)_{\max} = \max_{(s,t,\lambda_k) \in S_{xy\lambda}} \{t(s,t,\lambda_k)\} \tag{3}$$

$$\text{Min : } t(x,y,\lambda_k)_{\min} = \min_{(s,t,\lambda_k) \in S_{xy\lambda}} \{t(s,t,\lambda_k)\} \tag{4}$$

$$\text{Range: } t(x,y,\lambda_k)_{Range} = t(x,y,\lambda_k)_{\max} - t(x,y,\lambda_k)_{\min} \tag{5}$$

$$\text{Std: } t(x,y,\lambda_k)_{Std} = \sqrt{\frac{1}{uv} \sum_{(s,t,\lambda_k) \in S_{xy\lambda}} \left( t(s,t,\lambda_k) - t(x,y,\lambda_k)_{mean} \right)^2} \tag{6}$$

## PCA-Based Feature Classification

Before classification, feature extraction is done to reduce the dimension of the feature space and the correlation among the feature variables. The benefits of doing so are twofold: it trims down the time for classification, and lessens the design complexity for the classifier. A common feature extraction method is the Principal Component Analysis (PCA). In this method the data are linearly projected onto orthogonal axes, thereby producing new sets of uncorrelated variables.

The transmittance-feature data set can be represented by a matrix $\Omega$ of size $mxn$; the $m$ rows correspond to the number of measurements for the $p$ classes of tissue components, and the $n$ columns to the number of variables, i.e. $n$=16 bands. The $n$ principal components ($n$ orthogonal axes) are based on the eigenvectors computed from the covariance matrix of the original data set. To reduce the dimensionality of the input feature space to $r < n$, the eigenvectors are ordered based on the magnitude of their associated eigen values, and only those eigenvectors whose eigen values are the largest are chosen. The projected measurements of a transmittance feature vector $\boldsymbol{\alpha}$ can be expressed in the following form[11]

$$\boldsymbol{y} = \boldsymbol{A}^T \boldsymbol{\alpha} \tag{7}$$

where $\boldsymbol{A}$ is a projection matrix having a dimension of $nxr$ .On the assumption that the projected data has a gaussian distribution, the following condition can be applied to label an input vector $\boldsymbol{\alpha}_i$ :

$$\boldsymbol{\alpha}_i \in \omega_i \text{ if } p(\boldsymbol{y}_{ir}|\omega_i)p(\omega_i) > p(\boldsymbol{y}_{ir}|\omega_j)p(\omega_j) \quad \text{for all } i \neq j, \tag{8}$$

And for a quadratic classifier, the discriminant function is of the following form:

$$g_i(\boldsymbol{\alpha}) = \log\left(P(\omega_i)\right) - \frac{1}{2}\log\left|\sum{}_{ir}\right| - \frac{1}{2}\left(\boldsymbol{y}_{ir} - \boldsymbol{\mu}_{ir}\right)^T \sum{}_{ir}^{-1}\left(\boldsymbol{y}_{ir} - \boldsymbol{\mu}_{ir}\right) \tag{9}$$

$P(\omega_i)$ refers to the a priori probability for class $\omega_i$, while $\boldsymbol{\mu}_{ir}$ and $\sum_{ir}$ is the mean and covariance of the projected transmittance feature, for a particular transmittance class, $\omega_i$.

$$\boldsymbol{\mu}_{ir} = \boldsymbol{E}\left\lfloor \boldsymbol{y}_{ir} \right\rfloor \tag{10}$$

$$\sum{}_{ir} = \boldsymbol{E}\left\lfloor \left(\boldsymbol{y}_{ir} - \boldsymbol{\mu}_{ir}\right)^2 \right\rfloor \tag{11}$$

A pixel is associated to class $\omega_i$ when the condition $g_i(\alpha) > g_j(\alpha)$ for all $i \neq j$ Ref. [13] is satisfied

## Digital Colorization

To reflect the MT-stained color of the classified HE-stained pixel, the 16-band spectral-transmittance configuration of such pixel is first converted to its MT-stained configuration. The transformation process is done through a 16 × 16 transformation matrix, which was obtained through a linear mapping process between sets of HE and MT-stained transmittance spectra. The visualization of the digitally-stained image is realized by projecting the transformed 16-band spectral transmittance onto the RGB color space.

### *Linear Mapping of Spectral Transmittance*

To obtain a transformation matrix that would convert the transmittance of a classified pixel to its desired transmittance configuration, linear mapping of spectral transmittance data sets was introduced in Ref. [7]. The procedure was implemented on the basic assumption that a linear relationship exists between sets of transmittance spectra of tissue components stained with different dyes. Putting the linear assumption in equation form, we have:

$$\boldsymbol{T}_{MT} = \boldsymbol{T}_{HE} w \tag{12}$$

where $\boldsymbol{T}_{MT}$ and $\boldsymbol{T}_{HE}$ are $N x 16 (N = q\, x\, p)$ matrices consisting of the transmittance samples of MT-stained components and HE-stained components, respectively; $q$ indicates the number of transmittance spectra samples per tissue component; $p$ denotes the number of transmittance classes, i.e. number of tissue components, included in the transformation, e.g. nucleus, cytoplasm etc. The parameter $\boldsymbol{w}$ represents the 16 × 16 transformation matrix whose optimal solution, $\boldsymbol{w}$ , is calculated as follows:

$$\hat{w} = \boldsymbol{T}_{HE}{}^{+} \boldsymbol{T}_{MT} \tag{13}$$

where $\boldsymbol{T}_{HE}{}^{+}$ refers to the pseudo-inverse of matrix $\boldsymbol{T}_{HE}$.[14] For a given HE-stained transmittance $t_{HE}(\boldsymbol{\lambda})$, the corresponding estimate of its MT-stained transmittance is given by:

$$\boldsymbol{t}_{MT} = \boldsymbol{t}_{HE} w^{+} \tag{14}$$

### Weight Factor Calculation

The transformation matrices are associated with weight factors. Supposing we have *M* transformation matrices derived from the linear mapping of specific sets of transmittance spectra, and each is assigned to transform a group of transmittance spectra, then we are going to have *M* weight factors to be calculated. In the current experiment, these factors are calculated from the outputs of the *C* classifiers, which were implemented for the *C* different transmittance feature vectors. The decision of the *cth* classifier can be expressed in the following form:

$$e_c(\alpha) = \omega_i \tag{15}$$

where $\alpha$ is the input feature vector, and $\omega_i, \forall_i \in \{1,2,....C\}$ is the class favored for by the classifier. To calculate for the weight factors, the *C* classes of transmittance spectra are assigned to the *M* different transformation matrices:[8]

$$\Delta_{T_m} = \frac{V_{T_m}}{\sum_{m=1}^{M} V_{T_m}} \tag{16}$$

$V_{Tm}$ refers to the accumulated votes cast for class $\omega_i$ that is assigned to transformation matrix $T_m$. As an example, a transformation matrix $T_m$ assigned to transform $n_t$ classes of transmittance spectra would have a possible accumulated votes of $n_t$ , i.e. $V_{Tm} = n_t$ ,as a vote is assigned a binary value of 1 when the pre-assigned transmittance class is favored for by one of the classifiers, otherwise it is 0. With the weight factors calculated from Eq. 16, the transformed transmittance at $kth$ band, is given by the subsequent equation:

$$t_{MT}(\lambda_k) = \sum_{m=1}^{M}\sum_{j=1}^{16} \Delta_{T_m} * \left(t_{HE}(\lambda_j) w^m_{j,k}\right) \tag{17}$$

where $w^m_{j,k}$ represents an entry at $j,k$ in the $m^{th}$ transformation matrix. Note that the notation used for the transmittance spectra in Eq. 17 disregards the spatial location of the pixel, since the emphasis at this point is on the transformation of the spectral transmittance.

### RGB Visualization

Classification coupled with linear mapping of spectral transmittance results to a 16-band virtual MT-stained image. In order for us to visualize this image, the estimated MT-stained transmittance, $t_{MT}$ , of the classified HE-stained pixel is converted to its corresponding RGB values：

$$\begin{bmatrix}\alpha_R\\ \alpha_G\\ \alpha_B\end{bmatrix} = \begin{bmatrix}X_R & X_G & X_B\\ Y_R & Y_G & Y_B\\ Z_R & Z_G & Z_B\end{bmatrix}^{-1}\left(\begin{bmatrix}X\\ Y\\ Z\end{bmatrix} - \begin{bmatrix}X_{black}\\ Y_{black}\\ Z_{black}\end{bmatrix}\right) \tag{18}$$

where X,Y,Z correspond to the CIE 1931/1964 chromacity coordinates：

$$\begin{aligned} X &= \int \bar{x}(\lambda)E(\lambda)t_{MT}(\lambda)d\lambda\\ Y &= \int \bar{y}(\lambda)E(\lambda)t_{MT}(\lambda)d\lambda\\ Z &= \int \bar{z}(\lambda)E(\lambda)t_{MT}(\lambda)d\lambda \end{aligned} \tag{19}$$

The parameters $\bar{x}(\lambda), \bar{y}(\lambda), \bar{z}(\lambda)$ correspond to the CIE XYZ color matching functions and $E(\lambda)$ to the illumination spectrum. Furthermore, $X_i, Y_i, Z_i$ , $i = R, G, B$ , are the XYZ chromaticity coordinates of the RGB primary colors, and $X_{black}, Y_{black}, Z_{black}$ are the XYZ chromaticity coordinates of the monitor background.

## Experiment Results

Table 1 shows the number of training samples for the different tissue components.

### Classification with PCA Features

In section 3 we have identified five differing transmittance features, and in order for us to determine the appropriate number of principal components, which capture most of the variance of these features, we have to look into the eigen value energy, which can be derived from the covariance matrix of the input transmittance feature data set. Given the eigen values $\lambda_i \quad i = 1,2...n$ , to find for the first *k* principal components that account for most of the input data variance, *k* is set such that *Th,* satisfies a pre-defined threshold requirement:

$$Th = \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{m} \lambda_i} \tag{20}$$

We applied PCA independently to the different feature sets, and the resulting cumulative variances are shown in Fig. 4. The plots demonstrate that two principal components are enough, for a threshold of 90%, for the three feature sets: mean, minimum and maximum, while we need at least three components for the range and standard deviation feature sets.

**Table 1: Number of Training Samples for Each Tissue Component**

| Tissue components<br>( $\omega_i$ ) | Nu<br>( $\omega_1$ ) | Cyto<br>( $\omega_2$ ) | RBC<br>( $\omega_3$ ) | White<br>( $\omega_4$ ) | Fiber<br>( $\omega_5$ ) | Serum<br>( $\omega_6$ ) | C+S<br>( $\omega_7$ ) |
|---|---|---|---|---|---|---|---|
| # of samples | 450 | 450 | 300 | 450 | 450 | 300 | 150 |

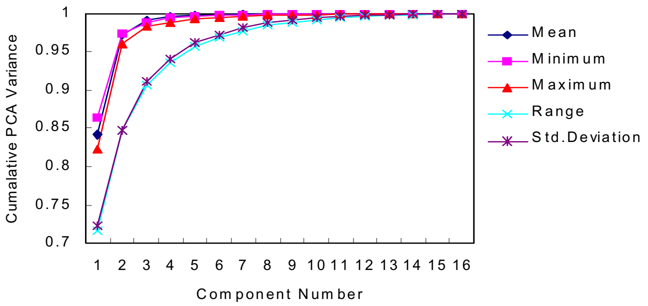


*Figure 4. Plot of the PCA cumulative variance.*

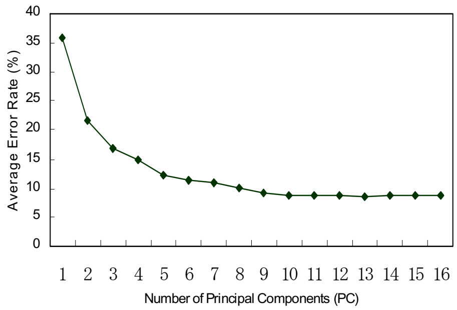


*Figure 5. Integrated average error, for the different feature vectors, with respect to number of PCs.*

However, since the covariance matrix, from which the PCA vectors were derived, was generated without giving emphasis on class details, an accumulated variance of, say, 90%, may not necessarily give an optimum classification result. To determine the optimum number of PC, for the purpose of classification, we randomly divided the data samples shown in table 1, into training and data sets. We performed the classification procedure ten times, with different sets of training and testing data for each time. The integrated average error rate, for the different feature vectors, with respect to the number of components is shown in Fig. 5; the average error dips at a point corresponding to ten PCs and does not significantly change after this point. The classification results with ten PCs for a particular area in an HE-stained specimen is shown in Fig. 6. Apparently, there are areas in the image that the classifiers don't have a common agreement about the class of the pixels found in the area. The corresponding HE-stained and MT-stained of the color-coded images in Fig. 6 is shown in Fig. 7

### Colorization

Once the HE-stained multispectral pixels are classified, the next step is to impart the right MT-stained color impression on them, and this can be done by transforming their transmittance spectra to their corresponding MT-stained configuration.

The simplest way to implement the digital transformation of an unstained or stained specimen to its desired stained image impression is to perform a linear mapping of the available transmittance data sets at once. However, this notion would not work, especially when the transmittance clusters of the different tissue components, in linear space, are not compactly defined. It is in Ref. [7] that the idea of providing several transformation matrices, instead of having a single transformation matrix only, was introduced. In this paper, we also utilized several transformation matrices to effect the digital transformation of an HE-stained image to its MT-stained equivalent

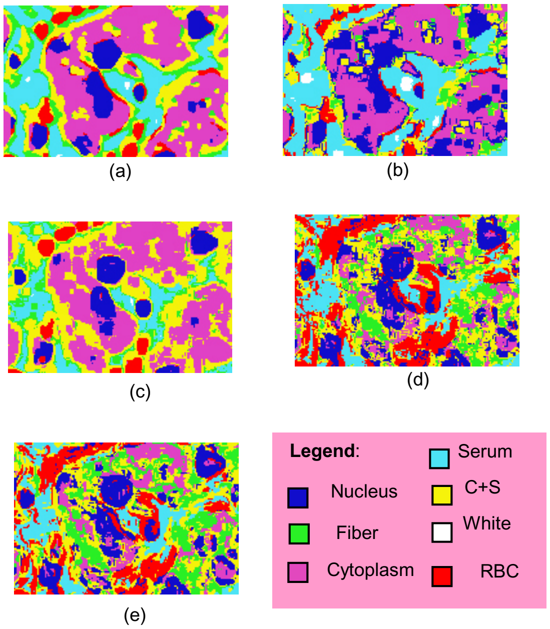


*Figure 6. Classification results of the different feature vectors: (a) Mean; (b) Maximum; (c) Minimum; (d) Range; (e) Std. dev.*

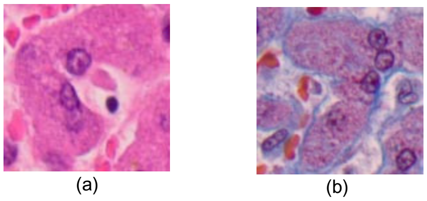


*Figure 7. Corresponding HE and MT-stained of the classified area in Fig. 6: (a) HE-stained; (b) MT-stained*

In Ref. [8], the authors utilized three transformation matrices that initially satisfy the digital staining objective. Namely, Matrix 1 ($T_1$), which was generated using the transmittances of cytoplasm, red blood cell (RBC)and the white region; Matrix 2 ($T_2$), which was produced with the transmittance spectra of Nucleus, Cytoplasm, Fibrosis and the white region; and Matrix 3 ($T_3$), which was derived using the same transmittance combination as with Matrix 2, but with the addition of the Serum transmittance spectra. Transmittance spectra of HE-stained pixels classified as nucleus, cytoplasm or RBC were transformed to their MT-stained configurations using $T_1$. On the other hand, $T_2$ was used to transform the HE-stained transmittance spectra of the white region and the fibrosis, while $T_3$ was assigned to transform the transmittance of the region associated to the presence of Serum. Lastly, the transmittance of the C+S component was assigned to the three transformation matrices. These matrices as well as the weighting factor calculations implemented in Ref. [8] were adopted in the current work

Figure 8 displays the resulting digitally–stained images. It can be observed that digitally-stained images acquire the same general impression as their real counterparts. At a closer look on the components that make up these images however, we can find that there are still areas which are not replicated at their best. The misrepresentations of these components in the virtual MT-stained images are due to the limitations of the classifier design and the design of the transformation matrices.

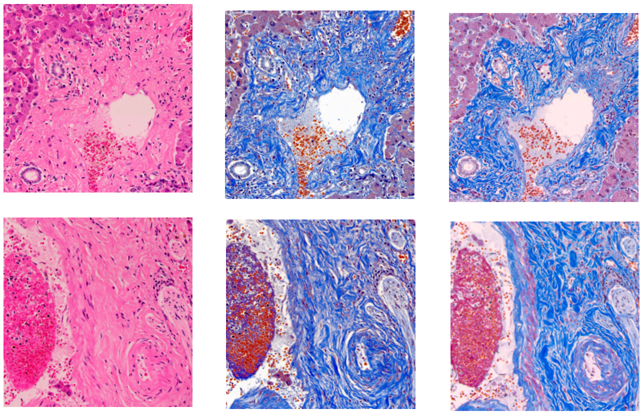


*Figure 8. Right: Liver specimens which are actually stained with Hematoxylin and Eosin (HE); Middle: Result of the digital staining procedure; Left: the specimen when actually stained with Masson Trichrome (MT) The contrast of the digitally stained images was adjusted using Photoshop.*

## Conclusion

The initial implementation of digital staining in the context of pathological images has already been introduced, and in this paper we have addressed the effective spectral dimension which would yield results (digitally-stained images) comparable to the images which were produced utilizing the original spectral dimension. It has been shown that a spectral dimension of ten is effective to produce fair results; these results prove the significance of multispectral imaging to pathological diagnosis. The effectivity of this dimension, however, has to be evaluated further with other samples of liver specimens and ideally with other specimens of different tissue type.

Indeed, the introduction of digital staining technique to the medical arena is of great practical value, especially for tele-pathology where exchange of medical information is done through the internet. However, before this technique can be fully utilized for practical purposes, it should be: (i) robust to varying staining intensity of the specimens; (ii) able to capture the differing characteristics of the tissue structures comprising the specimens; in addition (iii) the automation of the procedure should also be dealt with. We would look into these issues in our future work.

## Acknowledgment

This work is partly supported by the Special Coordination Funds for Promoting Science and Technology of the Ministry of Education, Culture, Sports, Science and Technology. The authors would also like to thank Mr. Yoshifume Kanno of NTT Data Corp. and Mr. Hiroyuki Fukuda of the National Institute of Telecommunication and Communication Technology, Akasaka Natural Vision Research Center, Japan for their technical support.

## References

1. D. Farkas, C. Du G. Fisher, C. Lau. et al. “Non-invasive image acquisition and advance processing in optical bioimaging,” Computerized Med. Imag. Graphic, 22, 89(1998)
2. P. R. Barber, B. Vojnovic, G. Atkin, et al. “Applications of cost-effective spectral imaging microscopy in cancer research,” J. Phys.D: Appl.Phys. 36,1729 (2003)
3. Y. Liu, T. Zhao and J. Zhang, “Learning Multispectral texture Features for Cervical Cancer Detection”, Proc. International Symposium on Biomedical Imaging: Macro to nano, pg.169 (2002)
4. M. A. Roula, A. Bouridane, F. Kurugollu and A. Amira, “A Quadratic Classifier Based on Multipsectral Texture Features for Prostate cancer Diagnosis” Signal Processing and Its Applications, Proc. Seventh International Symposium on, 2, pg.37 ( 2003)
5. F. Keiko M. Yamaguchi, N. Ohyama and K. Mukai, “Development of support system for pathology using spectral transmittance- the quantification method of stain conditions,” Proc. SPIE Medical Imaging, 4684,pg.1516(2002).
6. T. Abe, M. Yamaguchi, Y. Murakami, et al. ”Color correction of pathological images for different staining-condition slides” Proc. 6th International Workshop on Enterprise Networking and Computing in Healthcare Industry, IEEE, pg. 218 ( 2004).
7. P. A. Bautista, T. Abe, M. Yamaguchi, et al., ”Digital staining of unstained pathological tissue samples through spectral transmittance classification”, Optical Review, 12, 1 ( 2005)
8. P. A. Bautista, T. Abe, M. Yamaguchi, et al., ”Digital staining of pathological tissue specimens using spectral transmittance” Proc. SPIE Medical Imaging , 5747, pg. 1892 (2005)
9. H. Fukuda, N. Ohyama, M. Yamaguchi and T. Wada, Advancing Practice and Innovation through Informatics, Pittsburgh, PA, 2002.
10. B.C. Wilson and S.L. Jacques, “Optical Reflectance and Transmittance of Tissues: Principles and Applications” IEEE Journal of Quantum Electronics ,26, 2186 ( 1990) .
11. R.S. Gurjar, B. Backman, L. T. Perelman, et al., ”Imaging human epithelial properties with polarized light scattering spectroscopy”, Nature Medicine, 7, 1245 ( 2001).
12. M. Takeya, N. Tsuruma, H. Haneishi and Y.Miyake, ”Estimation of transmittance spectra from multiband micrographs of fungi and its application to segmentation of conidia and hyphae,” Applied Optics, 38,3644 (1999).
13. A.Webb, Statistical Pattern Recognition, 2nd Edition, John Wiley & Sons Ltd., 2002.
14. Matlab Technical Computing Version 6.1

## Author Biography

*Pinky A. Bautista is currently in her second year as a PhD student at the Imaging Science and Engineering Laboratory of Tokyo Institute of Technology, Japan*