Multispectral enhancement method to increase the visual differences of tissue structures in stained histopathology images

Pinky A. Bautista* and Yukako Yagi
Department of Pathology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

Received: June 26, 2012
Accepted: July 30, 2012

Abstract. In this paper we proposed a multispectral enhancement scheme in which the spectral colors of the stained tissue-structure of interest and its background can be independently modified by the user to further improve their visualization and color discrimination. The colors of the background objects are modified by transforming their $N$-band spectra through an $N \times N$ transformation matrix, which is derived by mapping the representative samples of their original spectra to the spectra of their target colors using least mean square method. On the other hand, the color of the tissue structure of interest is modified by modulating the transformed spectra with the sum of the pixel's spectral residual-errors at specific bands weighted through an $N \times N$ weighting matrix; the spectral error is derived by taking the difference between the pixel's original spectrum and its reconstructed spectrum using the first $M$ dominant principal component vectors in principal component analysis. Promising results were obtained on the visualization of the collagen fiber and the non-collagen tissue structures, e.g., nuclei, cytoplasm and red blood cells (RBC), in a hematoxylin and eosin (H&E) stained image.

Keywords: Multispectral, enhancement, visualization, multispectral, detection, histopathology

1. Introduction

Color is the most accessible feature of an image for use in image interpretation [1]. In pathology diagnosis, color is a primary feature of stained tissue sections by which diagnostic interpretations are derived. There are many varieties of chemical stains that can be utilized to visualize the tissue structures’ organization. For routine staining, however, hematoxylin and eosin (H&E) staining is the most popular. This stain mainly differentiates the nuclei areas from the cytoplasm areas. When abnormal changes in the ratio between the cytoplasm and nuclei areas are observed other staining techniques, i.e., special stains or immunohistochemical stain, are employed to further probe the cause of the abnormality. Due to the popularity of the H&E staining, a number of investigations have been done to digitally simulate its staining effect and to investigate the discrimination of H&E stained tissue structures which are ordinarily differentiated with the use of other types of stain [2–8]. The success of these research endeavors would eventually allow pathologists to quantify the distribution of tissue structures in the absence of special stains. In effect, the cost of staining and the risk of losing the area of interest in tissue re-cuts are reduced.

Interest on the application of multispectral imaging, which is widely used in remote sensing applications, to pathology diagnosis has grown over the years. Results on several pilot studies on the utility of multispectral imaging to improve the visualization of tissue structures and accuracy of quantitative histopathology image analysis showed promise [2–15]. A multispectral image acquisition system captures the same scene at different spectral wavelength using...
In this paper we introduce a spectral enhancement methodology on the spectral enhancement introduced by Mitsui et al. [15]. If $t_o$ is the original $N$-band spectral transmittance of the multispectral pixel, its enhanced spectral transmittance $t_e$ is given in the following form:

$$t_e = t_o + W\varepsilon$$

and the reference spectrum, $\hat{t}_o$, Eq. 2. The notation $W$, on the other hand, is defined as the weighting matrix whose elements function as weighting factors to the spectral residual-errors at different bands. Hence, through the matrix $W$ the amount of spectral error feed back to the original spectrum can be controlled. Furthermore, the reference spectrum $\hat{t}_o$ in Eq. 2 is the pixel’s original spectrum reconstructed using $M$ dominant principal component vectors in principal component analysis (PCA) and calculated as follows:

$$\hat{t}_o = VV^Tt_o + [I - VV^T]\bar{t}$$

where $\bar{t}$ is the average spectrum of the spectral samples used in the principal component analysis (PCA); $I$ is an $N\times N$ identity matrix; and $V$ is an $N\times M$ matrix whose $j$th column correspond to the $j$th principal component vector, $V_j$:

$$[V]_j = \begin{cases} v_j & j = <M \\ 0 \end{cases}$$

The enhancement formulation in Eq. 1 modifies the color of the object of interest, but not the color of the background objects. That is, in the previous enhancement formulation the users do not have options to modify the colors of the background objects. In addition, the users cannot freely specify the enhancement color for the object of interest as it is directly tied to the bands at which the spectral residual-error, $\varepsilon$, of the object peaks. For example, if the object of interest’s error peaks at 550 nm then its enhancement color is associated to the spectral color of this wavelength. These limitations are addressed in the proposed
2.2. Proposed multispectral enhancement method

There are three phases in the proposed multispectral enhancement method: (i) modification of the pixels’ spectral colors; (ii) detection of the object of interest; and (iii) multispectral enhancement. In the spectral modification phase the user first manually selects the background objects whose colors he/she desires to modify in the enhancement. After which the user specifies the envisioned colors of these objects in the enhanced image. The N-band spectra of the objects’ original and target colors, which could either be calculated or measured, are also provided by the user at this stage: (i) the spectra of the objects’ original colors; and (ii) the spectra of the objects’ target colors, which are the colors envisioned by user. The error of the object varies across the different spectral bands. The spectral residual-error vector in Eq. 7 to a constant. Let \( \mathbf{g}_r \) be an \( N \times 1 \) vector representing the modified spectral-color whose \( i \)th element, \( i = 1, 2, \ldots, N \), is computed as follows:
\[ \gamma_i = \sum_{j \in b_o} \epsilon_j \] (8)

By incorporating Eq. 8 to the original spectral error expression in Eq. 7 we will have:

\[ \gamma = A [I - VV^T] t_o - A [I - VV^T] \tau \] (9)

\[ [A]_{ij} = \begin{cases} 1 & j \in b_o, \forall i \\ 0 & \text{otherwise} \end{cases} \] (10)

where \( A \) is an \( N \times N \) matrix, and the variable \( b_o \) is a set of band numbers which correspond to the spectral bands at which the spectral errors of the object of interest are large, i.e., bands for enhancement. Since the magnitude of the spectral residual-error, Eq. 9, no longer varies between spectral bands, the bands for enhancement can be defined by the user by assigning appropriate weighting factors to the spectral errors at the selected spectral bands. Let us define the \( N \times N \) weighting matrix \( W \) in Eq. 1 as follows:

\[ [W]_{ij} = \begin{cases} k_i & i = j \in b_d \\ 0 & \text{otherwise} \end{cases} \] (11)

where \( k_i \in \mathbb{R} \), and is specified by the user. Moreover, \( b_d \) is the set containing the user-defined bands for enhancement. Introducing \( W \) in Eq. 11 to the spectral residual-error in Eq. 9 results to the following form:

\[ o_d = W \gamma \] (12)

where \( o_d \) is an \( N \times 1 \) vector representing the weighted spectral residual-error of a multispectral pixel. The desired color for enhancement is obtained by boosting the spectral errors at the bands specified by the user in \( b_d \), and suppressing the errors at other bands. Hence, through the weighting matrix \( W \) the color for enhancement can be controlled by the user. In should be noted that while the spectral bands in \( b_o \) are identified from the spectral residual-error configuration of the object of interest, those in \( b_d \) are freely specified by the user.

### 2.2.3. Multispectral enhancement

The proposed multispectral enhancement scheme is derived by introducing modifications to Eq. 1. The \( N \times N \) matrix \( Q_g \) determined from Eq. 6 is introduced and the second term of the equation is replaced with the expression given in Eq. 12. The proposed enhancement scheme can then be expressed as follows:

\[ t_e = [W] [A] [I - VV^T] + Q_g \] (13)

The present modifications broaden the functionality of the enhancement scheme which was originally introduced in [15]. These modifications allow the user to not only control the enhancement color of the object of interest but also to re-define the colors of the background objects. Equation 13 generates the same results as Eq. 1 when and if the \( N \times N \) matrix \( Q_g \) is an identity matrix and \( \gamma = \epsilon \). In order for the \( N \)-band enhanced spectrum \( t_e \) to be observed on ordinary color display, the spectrum is transformed to its equivalent RGB color values [21].

The actual implementation of the proposed multispectral enhancement method requires that the \( N \times N \) matrices \( Q_g, V, W, \) and \( A \) must all be determined offline. Spectral samples of the various objects in the image, i.e., object of interest and its background objects, are needed to calculate the entries of matrices \( Q_g \) and \( V \). Information on the enhancement band set, \( b_o \), is also needed to effectively calculate the modified spectral residual-error, Eq. 9, and determine the spatial location of the non-zero entries for the matrix \( A \), Eq. 11. The user-defined bands for enhancement \( b_d \) should also be known to define the weighting matrix \( W \), Eq. 12. The process flow of enhancing an \( N \)-band spectral transmittance \( t_o \) is illustrated in Fig. 1.
3. Experimental results

3.1. Multispectral images used

We used the microscopic multispectral imaging system developed by Olympus, Japan to capture 13 sets of stained liver tissue images. Each set is composed of H&E and Masson’s trichrome stained multispectral images of the tissue. The H&E and Masson’s trichrome stained sections are contiguous so that the pair of images in a set shares similar structural makeup. The spatial dimension of each multispectral image...
is 1434 × 1050 pixels and its spectral dimension is \( N = 63 \). The 63 spectral band images have spectral resolutions of 5 nm and span the visible spectrum from 410 to 720 nm. The H&E stained multispectral images were used as inputs to the proposed multispectral enhancement process while the Masson’s trichrome stained images were used as reference for the visual evaluation of the enhanced images.

3.2. Spectral transmittance data

From the captured images we identified six different tissue components namely: nucleus, cytoplasm, red blood cells (RBC), collagen fiber, smooth muscle and duct. We obtained the H&E stained spectral samples of these tissue components and the spectral samples of the white areas (areas which do not contain any tissue structures) to calculate for the principal component (PC) vectors \( v \), Eq. 4, and to investigate the spectral-error enhancement bands, \( b \), Eq. 10. Figure 2 shows the spectral samples projected onto the first three dominant principal component (PC) axes. The cluster plots provide an overview on the differentiation of these tissue components based on their H&E staining color. Clearly, discrimination between tissue components which are stained with the same type of dye, such as the cytoplasm, smooth muscle and collagen fiber which are mainly stained with the eosin dye, is not very direct. We can observe from Fig. 2 that there are significant overlaps in the spectral clusters of these tissue components.

3.3. Generation of the \( N \times N \) transformation matrix

Linear mapping procedure is most effective when the spectral transmittance classes are linearly separable. Hence, to solve for the \( N \times N \) matrix \( Q \) in Eq. 6 we considered the spectral samples of nuclei, cytoplasm and RBC whose spectral clusters display clear-cut separations in the linear PCA space. In order for the white areas in the image, the areas which do not contain any tissue structures, to be preserved in the enhancement result, we also considered the spectral samples of the white areas in the mapping. The \( N \times C, \ N = 63, \ C = 4 \), matrix, \( F \), in Eq. 6 contains the average H&E stained spectral samples of the nuclei, cytoplasm, and RBC, and that of the white area, and matrix \( G \) contains the spectra of their target colors. The color patches in Fig. 4 demonstrate the original and target colors of these tissue components. The color patches at the leftmost panel demonstrate their original colors, while the two groups of color patches displayed at the right panel illustrate the tissue components’ target colors. The first group of target colors was specified based on the Masson’s trichrome staining colors of the tissue components. The second group of target colors, on the other hand, was freely specified by the user. The corresponding transmittance spectra of the first group of target colors were obtained from the Masson’s trichrome multispectral images that were captured, while the transmittance spectra for the second group were calculated from the specified color values [7].

The transformation matrices using these sets of target colors were calculated using the first and second group of target colors as \( Q_{ga} \) and \( Q_{gb} \), respectively.

3.4. Determination of the spectral residual-error

The spectral data set that we used to derive the principal component (PC) vectors \( v \), Eq. 4, to determine the spectral residual-error of a multispectral pixel consisted of the spectral samples of the non-collagen fiber
tissue structures, which include the nuclei, cytoplasm, RBC, smooth muscle, and duct. The spectral samples of the white areas were also included. This implies that the derived principal component vectors could well estimate the spectra of the non-collagen fiber tissue structures but not the spectrum of the collagen fiber itself. We used five principal component (PC) vectors, i.e., $v_j, j \leq 5$, to re-construct the original spectrum.
Fig. 5. RGB images of the hematoxylin and eosin (H&E) stained images along with their Masson’s trichrome stained counterparts. Images on the first column correspond to the H&E stained images and images on the second column illustrate how the different tissue structures in the H&E stained images appear in a Masson’s trichrome stained tissue section.
of the pixel using Eq. 3. This number was based on the experimental results reported in [8]. To determine the modified spectral-residual error using Eq. 8, we first determined the sets of bands for enhancement, \( b_\text{d} \), from the spectral error configuration of the collagen fiber. Figure 4a displays the average spectral residual-errors of the different tissue components in their original configurations. We selected the bands at which there is a significant difference between the spectral error’s magnitude of the collagen fiber and the rest of the tissue components, and these correspond to \( b_\text{d} = \{550, 555, 560\, \text{nm}\} \). Figure 4b demonstrates the effect of re-assigning the spectral error values to the result of Eq. 8. We can observe that the magnitude of the modified spectral residual-error, as shown in Fig. 4b, is constant across the different spectral bands, which is in contrast to the original spectral residual-error displayed in Fig. 4a where there are obvious variations in the error values between different spectral bands.

### 3.5. Enhancement of multispectral images

Figure 5 displays the original H&E stained images. The Masson’s trichrome stained images on their right serve as reference for the visualization of the collagen fibers, which are denoted by arrows, in the enhanced image. In the enhancement, the user-defined bands for enhancement \( b_\text{d} \) were set to 480 to 520 nm at a step of 5 nm. The spectral residual-errors at these bands were weighted by a factor of \( k = 10 \), and the weighting factors for the rest of the errors were set to \( k = 0 \). The enhanced H&E stained images are presented in Fig. 6. The enhanced images at each column represent the enhancement results for a particular transformation matrix \( Q_g \): (i) \( Q_g = I \), where \( I \) is an identity matrix; (ii) \( Q_g = Q_{gb} \); and (iii) \( Q_g = Q_{gb} \). The derivation of the \( Q_{gb} \) and \( Q_{gb} \) matrices are discussed in Section 3.3. The different sets of enhanced images show improved visualization of the collagen fiber when compared to their original H&E stained images in Fig. 5. The results further illustrate the effect of the transformation matrices to the enhancement results. For instance, the color of the enhanced images on the second column share similarity to the Masson’s trichrome stained images in Fig. 5, especially the color impression of the background, i.e., non-collagen fiber areas; and the enhanced images on the third column show better discrimination between the nuclei, cytoplasm, and RBC.

The spectral bands at which the spectral residual-error were weighted with non-zero weighting factors are listed in Table 1. Here, we can see that the user can freely change the color of the object of interest, which is the collagen fiber. The enhancement color the collagen fiber is associated to the spectral color of the bands specified in \( b_\text{d} \).

### 3.6. Color discrimination evaluation

We used the CIELAB color difference as metric to evaluate the color discrimination between the different tissue components [7]. We randomly extracted spectral samples for the different tissue components from the 13 multispectral H&E images that we captured and applied spectral enhancement using Eq. 13. Table 2 shows the average Euclidian distance between the centers of the Lab color vectors of the different tissue components for the three different cases of transformation matrices. It is noted that the original color difference between the collagen fiber and smooth muscle is less than 5. This implies that from an H&E stained tissue section these tissue structures are perceptually similar. It is shown that application of spectral enhancement helps improves the color discrimination between these two structures. Furthermore, the color difference values presented in Table 1 illustrate the effectiveness of introducing the transformation matrix \( Q_g \) to the enhancement formulation in the overall multispectral enhancement result. By appropriately designing the transformation matrix we can further improve the visualization and color discrimination of both the collagen fiber and its background structures.
Fig. 6. Enhancement results illustrating the effect of using different designs of transformation matrices. The bands for enhancement were set to 520-530 nm and the weighting factors were set to $k = 20$ at these bands. Enhanced images at the first column demonstrate the enhancement results when $Q_g$ is set to an identity matrix; enhanced images at the second and third columns illustrate the effect of the transformation matrices derived using the target colors illustrated in Fig. 4b, c.

4. Discussion

In this work we have introduced a concept of multispectral enhancement which allows the user to vary the enhancement color of the object of interest as well as correct or modify the color of the background structures to the colors envisioned by the user. The option to modify the color of the background tissue structures is beneficial especially when the tissue sections from which the images were captured are poorly stained. Results of our experiments show that by appropriately designing the transformation matrix, the matrix
Fig. 7. Enhancement results illustrating the effect of choosing different bands for enhancement. The transformation matrix used to obtain these images was $Q_{gb}$, i.e., using target colors illustrated in Fig. 4b. The bands for enhancement used to produce these images are reported in Table 1. By choosing different bands for enhancement the enhancement color of the object of interest, which is at present the collagen fiber, can be varied.

Table 2

<table>
<thead>
<tr>
<th>Tissue components being compared</th>
<th>Original</th>
<th>Transformation matrices used ($\lambda_p = [520-530 \text{ nm}], \Delta = 20$)</th>
<th>$Q_{gb}$</th>
<th>$Q_{gb}$</th>
<th>$Q_{gb}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuclei &amp; Cytoplasm</td>
<td>34.64</td>
<td>35.35</td>
<td>33.28</td>
<td>57.02</td>
<td></td>
</tr>
<tr>
<td>Nuclei &amp; RBC</td>
<td>47.92</td>
<td>61.13</td>
<td>52.56</td>
<td>125.19</td>
<td></td>
</tr>
<tr>
<td>Cytoplasm &amp; RBC</td>
<td>27.10</td>
<td>34.73</td>
<td>43.09</td>
<td>84.97</td>
<td></td>
</tr>
<tr>
<td>Fiber &amp; Cytoplasm</td>
<td>7.52</td>
<td>36.15</td>
<td>39.62</td>
<td>55.84</td>
<td></td>
</tr>
<tr>
<td>Fiber &amp; Smooth muscle</td>
<td>4.77</td>
<td>51.67</td>
<td>62.01</td>
<td>66.86</td>
<td></td>
</tr>
</tbody>
</table>

*Number of spectral samples: Nuclei ~30700; Cytoplasm ~20400; RBC ~7500; Fiber ~31200; muscle ~7200.

used to convert the original spectrum of the multispectral pixels to the target spectral configuration, better visual perception of the different tissue structures can be achieved. The multispectral enhancement method proposed in [7] which was also extended in [23] to enhance the collagen fiber areas also allows the user to define the color for enhancement. However, the method is most effective only when the hues of the background pixels are similar, which is not true for the multispectral enhancement method proposed in this paper.

Figure 8a shows magnified areas cropped from the original H&E and Masson’s trichrome images. The
Fig. 8. Areas cropped from the original H&E stained and enhanced images. (a) These images show the effectiveness of the present enhancement scheme to differentiate collagen fibers and smooth muscles which are originally not clearly discriminated from an H&E stained tissue section. This set of images also show the variation of the response of the collagen fiber areas to the enhancement, i.e., some areas are more saturated than others. (b) These images show areas which are mislabeled in the present enhancement. Cytoplasm pixels were mislabeled to Collagen fiber.
leftmost and rightmost images respectively illustrate the staining patterns of H&E and Masson’s trichrome stains. We can observe that while the smooth muscle and collagen fiber areas display indistinct H&E staining patterns, they are prominently discriminated in the Masson’s trichrome stained image. The digitally enhanced H&E stained images at the center panels show similar color discrimination as the Masson’s trichrome stained image wherein the collagen fiber and smooth muscle areas also display distinct colors. It has been shown in [8] that the spectral difference between the smooth muscle and collagen fiber lies in their differing reaction to eosin stain. The spectral bands at which the spectral residual-error of the collagen fiber has large values are correlated to the eosin stain absorption bands, Fig. 4a. One of the issues pointed by authors in [8] in the simulation of the Masson’s trichrome staining is the color of the nuclei which appear reddish in the digitally stained images. The present experimental results show that designing the transformation matrix $Q_g$ using the spectral samples that are linearly separable improves the enhanced color of the nuclei.

An important aspect in any enhancement algorithms is consistency in delivering the desired results regardless of the image condition. For the same weighting factor value $k$ we can observe differing responses of the collagen fiber areas to the enhancement as noted by the variation in their color saturation, Fig. 8a. We can also observe that some pixels belonging to the cytoplasm were mis-detected to be part of the collagen fiber in that they display the enhancement color assigned to collagen fiber, Fig. 8b. This could be the underlying reason why the color difference between the cytoplasm and collagen fiber is smaller in comparison to the color difference between the collagen fiber and smooth muscle after enhancement, Table 2. This result also shows that there is a greater variance in the spectra of the pixels belonging to cytoplasm areas. It is thought that the current results can be improved further by: (i) incorporating information on the staining condition of the multispectral pixel in the definition of the weighting matrix; (ii) integrating statistical properties of the multispectral pixels to the enhancement formulation; and (iii) increasing the number of spectral samples to encompass the staining variations of the different tissue sections. Moreover to address variations in H&E staining protocols, staining correction method such as the one proposed in [24] could be employed to pre-process the images.

5. Conclusion

We have presented a multispectral enhancement methodology whereby we could effectively visualize the color differences between tissue structures which display similar H&E staining patterns. Moreover, the options in the current multispectral enhancement approach allow us to digitally simulate the effect of chemical staining, i.e., Masson’s trichrome staining, by designing appropriately the transformation and weighting matrices. The viability of the method to provide clearer discrimination between the different tissue structures has been illustrated by the enhancement results of the H&E stained images. The present enhancement methodology can be very useful to visualize tissue structures which are not emphasized by the original stain thereby reducing the cost of staining and the time to deliver a diagnosis. The technology of whole slide scanners is evolving. Integrating multispectral acquisition capability in the scanner designs are being looked into. In this case, the proposed multispectral enhancement can be incorporated into the scanners’ image acquisition or image analysis software to deliver digitally enhanced H&E images which possess improved visualization of the tissue structure of interest.

Acknowledgment

This work is supported in part by Olympus, Inc. Japan.

References

P.A. Bautista and Y. Yagi / Multispectral enhancement method to increase the visual differences


