A Bilinear Interpolation Based Approach for Optimizing Hematoxylin and Eosin Stained Microscopical Images

Kaya Kuru[1] and Sertan Girgin[2]

[1] Gülhane Military Medical Academy
`kkuru@gata.edu.tr`
[2] INRIA Lille Nord Europe, Villeneuve d'Ascq, France
`sertan.girgin@inria.fr`

Abstract. Hematoxylin & Eosin (H&E) is a widely used staining technique in medical pathology for distinguishing nuclei and cytoplasm in tissues by dying them in different colors; this helps to ease the diagnosis process. However, usually the microscopic digital images obtained using this technique suffer from uneven lighting, i.e. poor Koehler illumination. The existing ad-hoc methods for correcting this problem generally work in RGB color model, and may result in both an unwanted color shift and loosing essential details in terms of the diagnosis. The aim of this study is to present an alternative method that remedies these deficiencies. We first identify the characteristics of uneven lighting in pathological images produced by using the H&E technique, and then show how the quality of these images can be improved by applying an interpolation based approach in the Lab color model without losing any important detail. The effectiveness of the proposed method is demonstrated on sample microscopic images.

1 Introduction

Hematoxylin & Eosin (H&E) technique is one of the most common methods in medical pathology. The method allows to distinguish between nuclei and cytoplasm in tissues by dying them in different colors, namely blue and red. Routinely processed tissue sections are put into hematoxylin (a bluish dye) for 3-5 minutes. After removal of the excess dye, sections are then treated with a bicarbonate solution for about 2 minutes until nuclei stand out sharply blue. This is followed by rinsing and putting the slides into ethanol; sections are stained with eosin (a reddish dye) for 1-2 minutes. The staining procedure is finished by consecutive steps using ethanol, xylene and a mounting medium before coverslipping. This bi-coloring process helps to better depict the microscopic morphology of diseases and eases the diagnosis[1].

---

[1] Coloring process for blue is between yellow and blue and coloring process for red is between red/magenta and green respectively. Thus, sometimes it is not possible to see distinct blue or red colors on images.

With the recent advances in digital imaging, the latest digital cameras coupled with powerful computer software now offer image quality that is comparable with traditional silver halide film photography and greater flexibility for image manipulation and storage; hence, it is increasingly being used for image capture for microscopy – an area that demands high resolution, color fidelity and careful management [1].

Perhaps one of the most misunderstood and often neglected concepts in optical microscopy is the proper configuration of the microscope with regards to illumination, which is a critical parameter that must be fulfilled in order to achieve optimum performance [2]. The intensity and wavelength spectrum of light emitted by the illumination source is of significant importance, but even more essential is that light emitted from various locations on the lamp filament be collected and focused at the plane of the condenser aperture diaphragm [2]. Both the filament and condenser alignment procedures are necessary to achieve good Koehler illumination. Moreover, when working with living cells one must avoid relatively high light intensities and long exposure times that are typically employed in recording images of fixed cells and tissues (where photo bleaching is the major consideration) [3]. The misalignment of the filament and the condenser, low light situations and short exposure time lead to illumination problems; this reduces the quality of the digital images, in particular those obtained by using the H&E technique.

In order to remedy defects on digital H&E images, pathologists rely on various image processing applications to enhance them using the common RGB color model. The RGB color model is an additive color model in which red, green, and blue light is added together in various ways to reproduce other colors; it is optimized for display on computer monitors and peripherals, and is device dependent, meaning that the output may vary from one device to another. Enhancing images is usually a time consuming and mostly ad-hoc process; sometimes even experienced pathologists may not be able to get images in correct appearance. For example, Fig. 1 shows two sample pathological images in
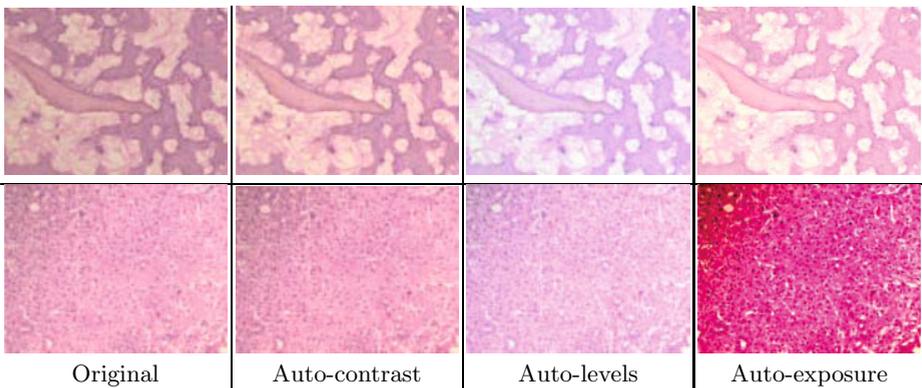

| Original | Auto-contrast | Auto-levels | Auto-exposure |

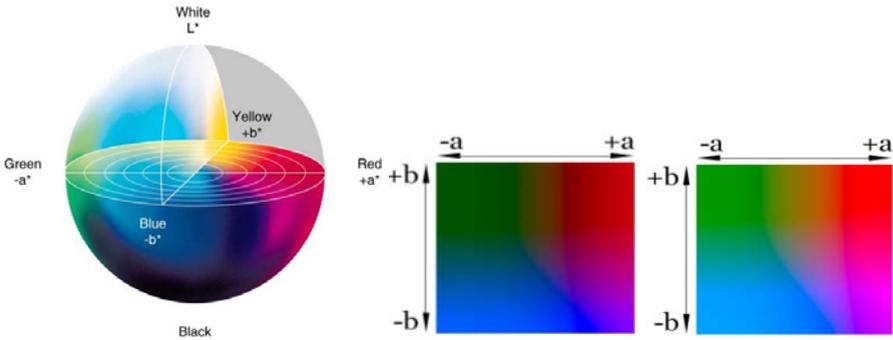**Fig. 1.** Two sample H&E images before and after enhancements in RGB color model

Fig. 2. The CIE LAB color model. The rectangles on the right show the colors in a* and b* dimensions with L* values of 25 and 50.

their original form and after employing three different enhancement procedures. Auto-contrast builds a histogram of the image and spreads out the values so the whole tonal-range is being used, auto-levels adjusts brightness and contrast to produce a balanced image with a good range of color intensities, and auto-exposure measures the darkest and lightest points in the image and adjusts the brightness accordingly. Note that, the effect of auto-contrast is minimal, and a color shift is apparent in the other enhanced images. The appearance of the image with color shift not only disturbs the appearance, but also causes the loss of essential details in the hematoxylin and eosin sections.

The difficulty stems from the fact that the RGB color model does not allow to modify the illumination of an image without altering the information stored in the red and blue channels of the image. Since, the nuclei of cells are colored in blue and the cytoplasms are colored in red in the H&E dying technique, illumination corrections in RGB color model inherently affect the data that the pathologists are interested in. In particular, some delicate details become barely perceptible impair the elaborateness of the images. This brings the need for a solution in which the lightness and the chromaticity of pixels in an image can be changed independent of each other. One possible option, is to use the CIE 1976 (L*, a*, b*), or CIE LAB, color space. The CIE LAB color space aims to conform to human vision and can describe all the colors visible to the human eye. It is a device-independent model established by the International Commission on Illumination [4]. The coordinates of CIE LAB color model, L*, a* and b*, represent the lightness of the color, its position between red/magenta and green, and its position between yellow and blue, respectively (Fig. 2). L* values range from 0 to 100; a value of 0 yields black and a value of 100 indicated diffuse white. Negative a* values indicate green while positive values indicate magenta; similarly, negative b* values indicate blue and positive values indicate yellow. L* component closely matches human perception of lightness. Thus, it is possible to make accurate color balance corrections by modifying output curves in the a* and b* components, or to adjust the lightness contrast using the L* component. As an example, the distribution of colors in a* and b* dimensions for

(a)   (b)
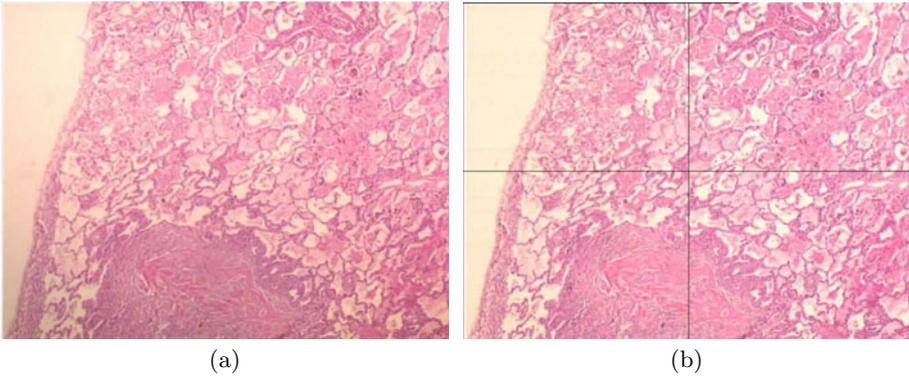
Fig. 3. Two images of the same specimen with light coming from different angles

two different L* values, 25 and 50, are shown in Fig. 2. Despite the change in the lightness, the color information is stable. By taking advantage of this property and the characteristics of the uneven lighting in H&E images, we propose a simple method to correct the illumination of the images and optimize them without losing any important detail which is essential for diagnosis.

The method will be described in detail and its effectiveness will be demonstrated on sample microscopic images in the next section.

2   The Method

As mentioned in the previous section, the CIE LAB color model provides a suitable and convenient way to represent and manipulate H&E images. In order to correct the illumination problems in these images, the remaining missing piece of the puzzle is the characteristics of uneven lighting that is observed in such images. For this purpose, we analyzed a large set of pathological images obtained using the H&E technique. Figure 3 shows two images of a sample specimen with light emitting from different angles. These images have a white area on the left side; because, there is no cell stained by the H&E technique at that area. But, in the first image, one can observe a gradient of increasing intensity (getting darker toward the bottom) which is caused by uneven lighting. The second image, on the other hand, has a relatively uniform illumination and does not suffer from such a problem; however, it still has an illumination problem spreading uniformly all over the image. The L*, a*,and b* and red, green, and blue components of both images in CIE LAB and RGB color models are presented in Fig. 4.

When we examine the lightness component, i.e. L*, of the images more carefully, we see that the L* values of pixels on the left side of the first image decrease linearly along the edge of the image (Fig. 5a); this trend spreads equally all over the image. As expected, the L* values of the pixels in the same region of the second image stay almost constant (Fig. 5b). Each pathological digital image with uneven lighting that we have examined had a similar linear pattern depending
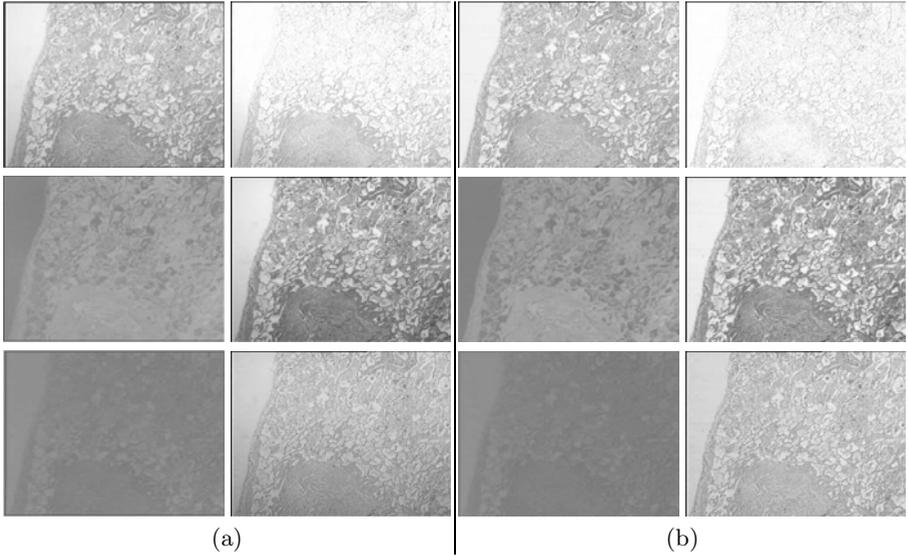
Fig. 4. L*, a*, and b* (first column, from top to bottom) and red, green, and blue (second column, from top to bottom) components of the images in Fig. 3 (a) and (b)

on the distinct positions of either the illumination apparatus or the microscope. From a global point of view, this common pattern tells us that the L* values of pixels in the CIE LAB color model and the relationship between them, for example, near the corners, indicate how the uneven lighting spreads through the image, either increasingly or decreasingly with a different angle with the XY plane. Given an image and a set of points on the image that are supposed to be white, i.e. the L* value equal to 100, by first converting the image to the CIE LAB color model, and then interpolating and exterpolating the L* values of these points subtracted from a reference value, in most cases 100, we can create a mask that approximates the distribution of the uneven lighting; subtracting this mask from the L* channel of the image and converting the image back to
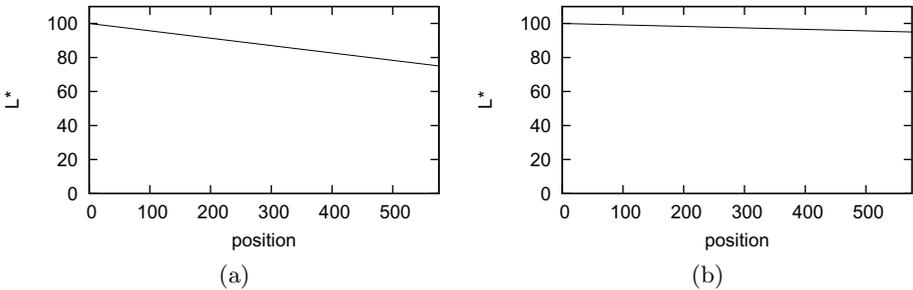

Fig. 5. The L* values of the pixels in the left-most area of the images in Fig. 3
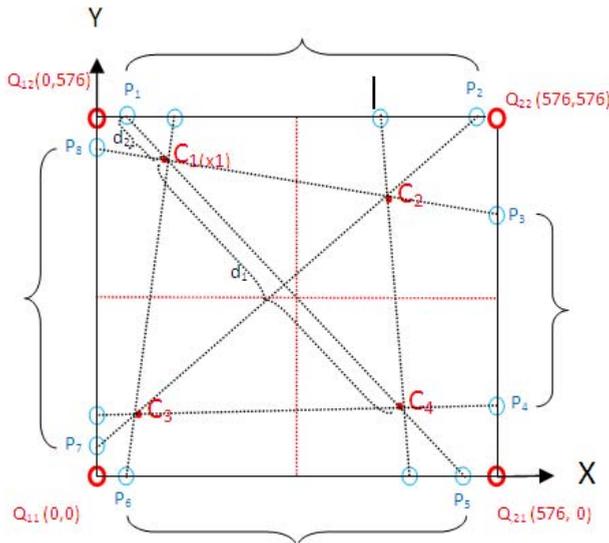
Fig. 6. The four red dots show the data click points and the red circled dots are the points at which we want to interpolate by using the calculated values (L values and xy coordinates values) of blue circled dots

the RGB color model allows us to reduce the uneven illumination, and in most cases eliminate it. The procedures for converting an image from the RGB color model to the CIE LAB color model and vice versa are provided in the appendix. Note that, during this process a* and b* channels of the image, which contain information about nuclei and cytoplasm in tissues, are not altered, and detail enhancement and sharpening can also be applied to this channels before converting the image back to the RGB color model to further improve the quality of the resulting image.

By taking into consideration the observed linear distribution, we opted for a bilinear interpolation scheme using four source points; this scheme allows to process images efficiently and produces good results. Let C_1, C_2, C_3 and C_4 be the four chosen points such that any three of them are not colinear. For each pair of points, C_i and C_j, $i < j$, we first determine the intersection points of the line passing from these two points and the image boundaries; if both points are on the same border of the image, then the intersection points are assumed to be the points themselves. The mask values of the intersection points are then calculated by linear extrapolation of the L* values of the corresponding source points subtracted from the reference value. Once the mask values of intersection points are known, the mask value of any point on a border of the image can be found by calculating a set of projected values and taking their mean; each projected value is calculated by interpolating or extrapolating linearly the mask values of a pair of intersection points on that border (see Fig. 6). Finally, in
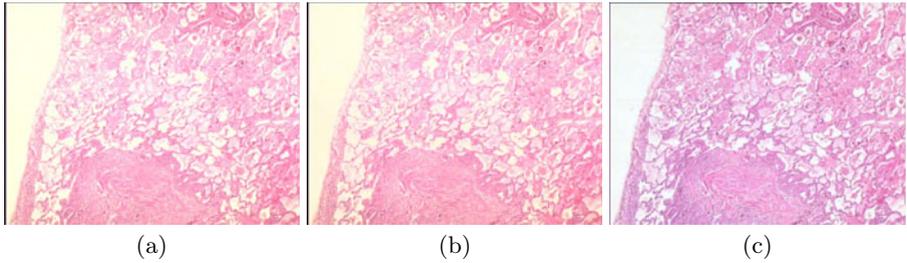
(a)   (b)   (c)

Fig. 7. (a, b) The resulting images after applying the proposed method to the sample images in Fig. 3; (c) the resulting image when correction is applied in the RGB color model

Original   CIE LAB   RGB

Fig. 8. The histograms of the b* channel of the sample image in CIE LAB color model (a) before and after correction in (b) CIE LAB and (c) RGB color models

order to find the mask value of any point (x, y) in the image, we apply linear interpolation in x and y directions separately using the mask values of the border points (i.e. interpolate the mask values of the two border points on column x and similarly on row y). Although each step is linear in the sampled values and in the position, the estimation as a whole is not linear but quadratic in the sample location.

Figure 7 shows the resulting images after applying the proposed method to the sample images in Fig. 3 with illumination problems. One can observe that the uneven lighting (poor Koehler illumination) is removed in both images. Correcting them in RGB color model with similar algorithm on all channels causes a color shift (Fig. 7c), demonstrating the advantage of using the CIE LAB color model. The histograms of the b* channel of the sample image with uneven illumination (Fig. 3a) before and after correction in CIE LAB and RGB color models are presented in Fig. 8. There is a nonsignificant difference between the histograms of the original image and the one processed in the LAB color model; the mean of the b* values is 133.32 in the first one (standard deviation 5.53) and 130.65 in the second (standard deviation 5.48). On the other hand, there is a significant difference between the histograms of the original image and the one processed in the RGB color mode with a shift to the left (the mean of the b* values is 118.75 with standard deviation 5.62); that is to say, there is a significant loss of essential data in terms of diagnosis.
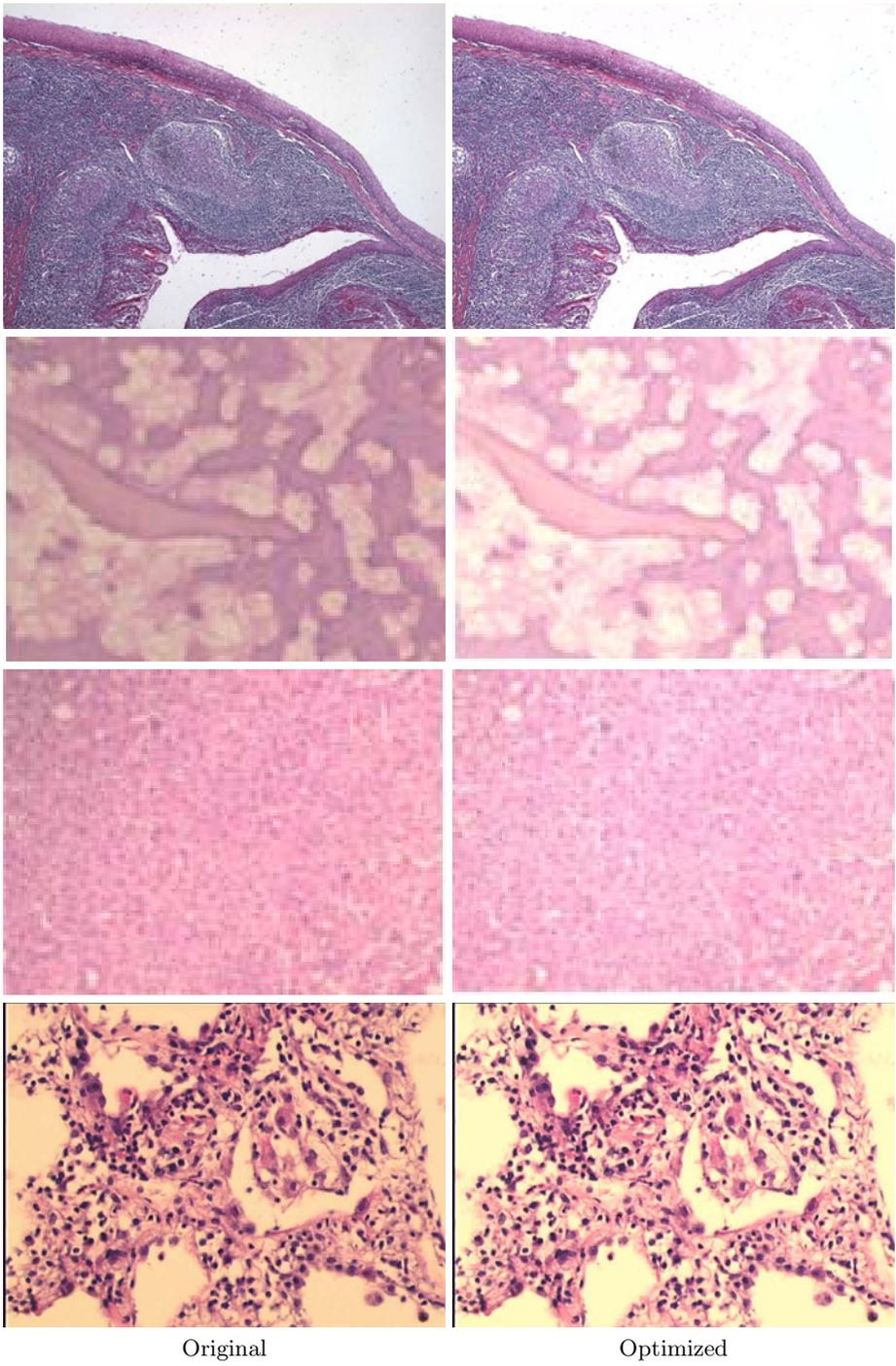
Original   Optimized

Fig. 9. Sample images before and after optimization in the CIE LAB color model

We have implemented the proposed method both in Java and in DOTNET[2]. The program allows the user to load an image obtained using the H&E staining technique, choose the source points and specify the reference L* value. In order to facilitate the process of choosing suitable points, the program can highlight the probable white areas of the image (i.e. with L* values in CIE LAB model close to 100 and not smaller than 90) and automatically determine the source points by dividing the image in four quadrants and picking in each quadrant the point among possible white points which is closest to the centroid of those points; if such a point does not exist in a quadrant, then the centroid of the points that are symmetric to the found points in other quadrants with respect to the XY axis is chosen as the source point and its mask value is set to the average of the mask values of the source points in other quadrants[3].

Some resulting corrected images that are produced by the application are presented in Fig. 9. Distinctive features of H&E technique in the images are made noticeably evident after the optimization of the illumination problem. No essential detail of nuclei and cytoplasm, in terms of diagnosis is lost, thanks to the LAB color model.

3 Conclusion

The digital camera systems have immediate advantages in the world of medicine - images can be distributed easily and without delay for second opinions or for educational purposes. However, no matter how good the camera is, it cannot produce outstanding images from a poorly configured microscope. Both the filament and condenser alignment procedures are necessary to achieve good Koehler illumination in microscopic images; furthermore, low light situations arise frequently and short exposure times may be compulsory. These affect the quality of the resulting images, and often they suffer from illumination problems.

In this paper, we proposed a simple but effective method for correcting uneven lighting in pathological images obtained using the Hematoxylin & Eosin staining technique. The method exploits the characteristics of such images by first converting them into the CIE LAB color model, in which the lightness and the chromaticy of pixels can be modified independent of each other, and then employing bilinear interpolation to mask out the uneven lighting that spreads through an image. The empirical studies on a diverse set of sample images show that the proposed method is effective in improving the quality of H&E images; unlike other ad-hoc methods in RGB color space, a color shift is not observed in the resulting images and essential details are kept. The developed application provides pathologists with little or no image processing experience an effective means of optimizing the H&E images for printing, image analysis, or

---

[2] The program can be downloaded from `http://goo.gl/UGQR3`

[3] Possible white areas can be seen easily in four quadrants of the images after they are highlighted, although they are not evident by naked eye. Choosing points other than white areas may cause unexpected results.

telepathology. We may explore some other enhancement for H&E images to improve diagnosis by adopting some basic image processing algorithms to other channels, "a" and "b", of LAB color mode as well as "L" channel as a future work.

References

1. Drent, P.: Digital imaging - new opportunities for microscopy. Microscopy, Nikon (2004),
   `http://www.microscopyu.com/articles/digitalimaging/drentdigital.html`
2. Parry-Hill, M., Sutter, R.T., Davidson, M.W.: Microscope alignment for kohler illumination. Microscopy, Nikon (2004),
   `http://www.microscopyu.com/tutorials/java/kohler/`
3. Hazelwood, K.L., Olenych, S.G., Griffin, J.D., Murphy, C.S., Cathcart, J.A., Davidson, M.W.: Introduction to digital imaging in microscopy. Hamamatsu (2004),
   `http://learn.hamamatsu.com/articles/microscopyimaging.html`
4. Fairchild, M.: Color appearance models, 2nd edn. Wiley - IS & T series in imaging science and technology. J. Wiley (2005); ISBN: 978-0-470-01216-1
5. Ford, A., Roberts, A.: Colour space conversions. Mendeley (1998),
   `http://www.poynton.com/PDFs/coloureq.pdf`

Appendix

Conversion between RGB and CIE LAB Color Models

In the study of color perception, one of the first mathematically defined color spaces was the CIE XYZ color space, created by the International Commission on Illumination (CIE) in 1931. CIE XYZ may be thought of as derived parameters from CIE RGB color space, the red, green, blue colors. CIE LAB color space is based directly on the CIE XYZ color space as an attempt to linearize the perceptibility of color differences. The non-linear relations for L*, a*, and b* are intended to mimic the logarithmic response of the eye. In order to convert an image from RGB color space to CIE LAB color space (or vice versa), the CIE XYZ color space is used as an intermediate color space at transformation phases from one color space into other [5]. The conversion formulas between CIE XYZ and RGB color spaces are as follows:

$$X = 0.412453 * R + 0.357580 * G + 0.180423 * B$$
$$Y = 0.212671 * R + 0.715160 * G + 0.072169 * B$$
$$Z = 0.019334 * R + 0.119193 * G + 0.950227 * B$$
$$R = 3.240479 * X - 1.537150 * Y - 0.498535 * Z$$
$$G = -0.969256 * X + 1.875992 * Y + 0.041556 * Z$$
$$B = 0.055648 * X - 0.204043 * Y + 1.057311 * Z$$

The conversion formula from the CIE XYZ color space to the CIE LAB color space is defined as

$$L^* = 116f(\frac{Y}{Y_n}) - 16, a^* = 500[f(\frac{X}{X_n}) - f(\frac{Y}{Y_n})], b^* = 200[f(\frac{Y}{Y_n}) - f(\frac{Z}{Z_n})]$$

where $f(t) = t^{\frac{1}{3}}$ for $t > 0.008856$, otherwise $f(t) = 7.787t + \frac{16}{116}$ and X_n, Y_n and Z_n denote the CIE XYZ tristimulus values of the reference white point. The division of the $f(t)$ function into two domains is done to prevent an infinite slope at $t = 0$. $f(t)$ is assumed to be linear below some $t = t_0$, and is assumed to match the $t^{1/3}$ part of the function at t_0 in both value and slope.

$$t_0^{1/3} = t_0 + b \text{ (match in value) and } 1/(3t_0^{2/3}) = a \text{ (match in slope)}$$

In other words, the value of b was chosen to be 16/116. The above two equations can be solved for a and t_0 to obtain

$$a = \frac{1}{3\delta} = 7.787037 \text{ and } t_0 = \delta^3 = 0.008856$$

where $\delta = \frac{6}{29}$. The reverse transformation can then be calculated by applying the following rules:

1. Define $f_y = (L^* + 16)/116$
2. Define $f_x = f_y + a^*/500$
3. Define $f_z = f_x - b^*/200$
4. If $f_y > \delta$ then $Y = Y_n f_y^3$, else $Y = (f_y - 16/116)3\delta^2 Y_n$
5. If $f_x > \delta$ then $X = X_n f_x^3$, else $X = (f_x - 16/116)3\delta^2 X_n$
6. If $f_z > \delta$ then $Z = Z_n f_z^3$, else $Z = (f_z - 16/116)3\delta^2 Z_n$