A Complete Color Normalization Approach to Histopathology Images Using Color Cues Computed From Saturation-Weighted Statistics

Xingyu Li*, *Student Member, IEEE*, and Konstantinos N. Plataniotis, *Fellow, IEEE*

***Abstract*— *Goal*: In digital histopathology, tasks of segmentation and disease diagnosis are achieved by quantitative analysis of image content. However, color variation in image samples makes it challenging to produce reliable results. This paper introduces a complete normalization scheme to address the problem of color variation in histopathology images jointly caused by inconsistent biopsy staining and nonstandard imaging condition. *Method*: Different from existing normalization methods that either address partial cause of color variation or lump them together, our method identifies causes of color variation based on a microscopic imaging model and addresses inconsistency in biopsy imaging and staining by an illuminant normalization module and a spectral normalization module, respectively. In evaluation, we use two public datasets that are representative of histopathology images commonly received in clinics to examine the proposed method from the aspects of robustness to system settings, performance consistency against achromatic pixels, and normalization effectiveness in terms of histological information preservation. *Results*: As the saturation-weighted statistics proposed in this study generates stable and reliable color cues for stain normalization, our scheme is robust to system parameters and insensitive to image content and achromatic colors. *Conclusion*: Extensive experimentation suggests that our approach outperforms state-of-the-art normalization methods as the proposed method is the only approach that succeeds to preserve histological information after normalization. *Significance*: The proposed color normalization solution would be useful to mitigate effects of color variation in pathology images on subsequent quantitative analysis.**

***Index Terms*—Color variation, histological information preservation, histopathology image, illuminant normalization, saturation-weighted statistics, spectral estimation.**

I. Introduction

DIGITAL histopathology is a research field where color image processing algorithms and pattern recognition methods are exploited to enable computers to understand histopathology images and to make diagnosis decisions. As quantitative analysis on histopathology images is usually achieved through comparing numerical descriptors of a query image to prior knowledge obtained from training data or physicians, deviation of numerical descriptors of a query image from prior knowledge should only reveal the true differences between histological information conveyed by images. However, due to operational inconsistency in histopathology image preparation, images of biopsy samples stained by the same types of chemical dyes usually appear in different colors. Consequently, numerical features extracted directly from images may be distorted by such color variation and deviate from true values, finally resulting in inaccurate segmentation and diagnosis.

To reduce effects of color variation among histopathology images on numerical features and subsequent analysis, some works extracted numerical features from a grayscale version of a query image [1], [2]. However, a large amount of information mainly carried by color is ignored in these approaches. Recent research in digital histopathology has confirmed significance of color information in quantitative analysis on histopathology images with few color variation generated under tightly controlled laboratory conditions [3], [4]. To take advantage of color information for accurate quantitative analysis within large datasets, color variation caused by operational inconsistency in histopathology image preparation should be removed beforehand, so that extracted features represent the real histological characteristics only. In this sense, normalization algorithms to remove color variation in images are crucial in histopathology.

In this paper, we propose an effective normalization scheme to address color variation in histopathology images generated by light-absorbing stains only. Compared to previous works, our method is significant as it addresses three challenges of histopathology image normalization holistically as follows.

1) A histopathology image is a final output of a biopsy processing pipeline including sectioning, staining, and imaging. Any operational inconsistency in this pipeline may cause color variation [5]. Hence, the first challenge of color normalization for histopathology images is to blindly identify different sources of color variation. To this end, this paper introduces a complete color normalization solution based on a microscopic imaging model. Although similar models are exploited in other methods to normalize spectral variation in stains, it should be stressed that our solution solves a more complicated problem usually occurring in clinics, where color variation in images is jointly caused by inconsistent staining and imaging conditions. By identifying the two causes of color variation, an illuminant normalization module and a spectral normalization module are proposed and concatenated to form our complete normalization pipeline.

Manuscript received July 25, 2014; revised December 15, 2014; accepted February 14, 2015. Date of publication February 19, 2015; date of current version June 16, 2015. *Asterisk indicates corresponding author.*

*X. Li is with the Multimedia Lab, Edward S. Rogers Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3H7, Canada (e-mail: xingyu.li@mail.utoronto.ca).

K. N. Plataniotis is with the University of Toronto.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TBME.2015.2405791

2) The second challenge arises from the use of multiple chemical dyes on a single tissue sample. Despite counter colors visually, stains may mix due to colocated histological components, and thus image colors are determined by various combinations of stains. To identify contribution of each stain to color variation, a nonnegative matrix factorization (NMF)-based stain decomposition algorithm is developed. Compared to existing adaptive stain decompositions whose performance are prone to be affected by achromatic pixels, a key characteristic of our algorithm is its noise-resistant property contributed by the proposed saturation-weighted (SW) statistical method. The innovation of SW statistics is advantageous to limit impacts of achromatic pixels on color estimation, and thus to produce consistent stain estimation. This noise-resistance is beneficial especially for stain estimation on images having more blank areas.
3) A third challenge is to avoid histological information loss after image normalization. The capability of preserving histological information (including tissue texture details, spatial structures, and morphology features of histological objects) is a very important criterion to evaluate normalization algorithms in digital histopathology. However, this challenge is never explicitly examined in previous works. In this study, we pay close attention to maintaining histological information when developing algorithms. In evaluation, we particularly design an experiment to examine this capability of the proposed normalization method and existing color normalization algorithms. Our results show that only the proposed method succeeds to maintain tissue features after normalization.

The rest of this paper is organized as follows. The state of the art in histopathology image normalization including problem formulation and previous works is reviewed in Section II. In Section III, we introduce a complete normalization scheme for histopathology images generated by light-absorbing mechanism. By identifying different causes of color variation, we develop an illuminant normalization module and a spectral normalization module. Experimental results and discussions are presented in Section IV, followed by conclusions in Section V.

II. State of the Art

A. Color Variation in Histopathology Images

In histopathology, to enhance visibility of spatial structures of histological components, chemical staining that uses counter-color chemical dyes such as hematoxylin and eosin (H&E) to stain different histological components is usually performed in histopathology routine. Then a digital version of a specimen is generated by microscopic imaging, where a camera captures transmitted light that is not absorbed by stains in different color channels. Accordingly, the ith color component of a histopathology image is formulated by

$$I(p,\lambda_i)=\int_{\lambda_i-\delta}^{\lambda_i+\delta} f_i(\lambda)E(\lambda)e^{-M_i(\lambda)D(p)}d\lambda \qquad (1)$$

where $f_i(\lambda)>0$ for $\lambda\in(\lambda_i-\delta,\lambda_i+\delta)$ represents a sensor response in the ith color channel, $E(\lambda)$ denotes spectral power distribution (SPD) of imaging light, $M_i(\lambda)=[m_{i1}(\lambda),\ldots,m_{iN}(\lambda)]$ represents the absorption spectra of N types of stains around wavelength λ_i, and $D(p)=[d_1(p),\ldots,d_N(p)]'$ is so-called stain depths, or stain proportions, describing amount of stains bounded at location $p=(x,y)$. For completeness, the derivation of the imaging model is specified in the Appendix. It should be noted that $I(p,\lambda_i)$ in (1) represents a pixel value in a color channel without gamma correction, i.e., $\gamma=1$. Corresponding, an RGB-format image in this paper refers to an image represented in the linear RGB color space where $\gamma=1$.

In histopathology images prepared in different pathology laboratories using different equipment, color variation is usually observed due to operational inconsistency in image preparation routine, consisting of biopsy sectioning, staining, and imaging [5]. Correspondingly in the histopathology imaging model formulated in (1):

1) $D(p)$ is directly related to biopsy sectioning. It conveys histological information contained in an image, and thus should not be modified (more details in Section III-B).
2) $M_i(\lambda)$ quantifies stains' spectra. Inconsistency in biopsy staining, such as different stain manufacturers, may lead to variation in $M_i(\lambda)$ for the same types of stains, finally causing color variation [6], [7].
3) A biopsy digitization condition is characterized by $f_i(\lambda)E(\lambda)$, the product of a camera sensor response and SPD of an imaging illuminant. Variation in biopsy imaging condition, including differences in either $f_i(\lambda)$ or $E(\lambda)$ may produce color variation in images [5], [6]. In this study, instead of a specific $f_i'(\lambda)$ which may be device-dependent and thus unknown, we assume a common $f_i(\lambda)$, then variation in biopsy digitization condition can be quantified as $f_i'(\lambda)E'(\lambda)=f_i(\lambda)E(\lambda)$, where $E'(\lambda)$ is the actually PSD of incident light, and $E(\lambda)$ is the PSD of the resulting illuminant associated with the common $f_i(\lambda)$. Hence, variation in biopsy imaging is transferred to variation in the resulting illuminant $E(\lambda)$.

Therefore, image color $I(p,\lambda_i)$ may vary with respect to variation in biopsy staining and imaging conditions, which are quantified by variation in $M_i(\lambda)$ and $E(\lambda)$, respectively. For clarification, illuminant variation in this paper refers to non-standard $E(\lambda)$, which corresponds to inconsistency in biopsy imaging, and spectral variation in stains, or simply stain variation, refers to inconsistent $M_i(\lambda)$ of the same type of stains, corresponding to disagreement in biopsy staining.

B. Previous Works on Histopathology Color Normalization

We categorize histopathology image centered color normalization solutions in the literature into three distinct groups.

1) Histogram Matching: The first group of normalization algorithms are based on histogram matching in the RGB color space. In [8], after image background removal, histogram matching is performed in the red, green, and blue channels, respectively. In [9], color map quantile matching, a variation of

histogram matching, was proposed for stable color normalization. Since histogram matching on entire images ignores local differences of image content, color associated with one stain may be matched to irrelevant colors. Recently, tissue component segmentation followed by histogram landmark matching was proposed to remove stain variation [14]. As histogram matching-based methods do not distinguish causes of color variation, histological information is hardly preserved after normalization, consequently introducing unwanted bias into subsequent image analysis.

2) Color Transfer: The second group of color normalization solutions relies on the so-called color transfer technique discussed in [15]. After converted to the $l\alpha\beta$ color space [16], the mean and variance in each color channel of a query image are matched to the statistics of a reference image [10]. Since images stained by multiple chemical dyes may have different color distributions, colors associated with different histological components may smear each other after color transfer. To address this problem, before color transfer, an image is divided into regions by segmentation manually [17] or automatically through pixel classification [11], so that each region contains one type of histological objects only. For one thing, manual segmentation of images in large datasets is infeasible. For another, segmentation achieved by hard pixel classification is not reliable because one pixel in a histopathology image may belong to multiple objects due to histological component overlap. Hence, soft pixel classification by stain decomposition was proposed to precede color transfer [11], [18]. An improved normalization method based on study [11] was proposed in [12]. In this method, mean and ranking statistics in each decomposition channel of a test image are mapped to the statistics of a reference image nonlinearly. However, applying color transfer on decomposition results has two limitations. First, sources of color variation in histopathology images are lumped together and not addressed separately. Second, statistics in decomposition channels, which is closely related to histological information in images, is modified and thus tissue features may not be preserved after normalization.

3) Spectral Matching: This class of color normalization solutions aims to remove stain variation only. To that end, algorithms belonging to this group first estimate stain spectra either using adaptive estimators [7], [19] or via dedicated hardware [13], [20], and then match the estimated quantity to a reference stain spectra. Approaches in this category have an advantage over other groups in the sense that if stain variation is the only source of color disagreements in images, with good estimation, spectral matching approaches can preserve histological features. However, if other causes also contribute, as algorithms in this category do not identify their effects, histological features may be modified after normalization.

Since success of spectral matching approaches heavily relies on the accuracy of stain spectral estimation, a very brief review on it is presented here. Spectral estimation, or stain decomposition, is a process that estimates stain spectra and corresponding stain proportions at each pixel in a histopathology image. Early works of stain decomposition can be traced back to [21], where stain spectra are predetermined by experiments. To address spectral variation in stains, spectra were estimated at regions of interest by manual selection [22]. However, such manual operation is time-consuming for large datasets. For adaptive stain decomposition, sparseness analysis followed by relative newton method was proposed to achieve blind source separation [23]. Since this method is designed for hyperspectral images whose channel number is much larger than the number of stain types on biopsies, it is not applicable to RGB images. Based on an imaging model where image colors were linear combinations of stain spectra in the OD domain, a plane fitting method was achieved by singular value decomposition and thresholding for images containing two stains only [7]. For accurate estimation of weak stains, prior knowledge on stains is used in the plane fitting process [19]. However, their performance vary as the prefixed thresholds for stain estimation may be inappropriate for a testing image. For accurate estimation, spectra of chemical stains were detected using dedicated hardware [13], [20]; whereas dependence on devices limited their adoption. Later, NMF was used for spectral estimation [18], [24]. Since both studies do not address an inconsistent convergence issue, stain decomposition may converge to any local minima, leading to wrong estimation. Blind color decomposition (BCD) is achieved by performing expectation-maximization on color distributions in the Maxwell color triangle [25]. Though a heuristic randomization function tries to select stable colors for estimation, BCD method is prone to be affected by achromatic pixels when estimating weak stains' spectra. Recently, spectral estimation via supervised learning on a training set of histopathology images was proposed in [12]. As the learning model relies on statistics of images in the training set, stain estimation may not be accurate, or even fail, when spectral variation in stains occurs between a query image and training images.

Table I provides an overview of color normalization algorithms currently used to process histopathology images. In this summery, spectral matching algorithms, which rely on predetermined spectrum values, or a reduced set of a reference image, offer greater flexibility compared to solutions following the histogram matching or color transfer paradigm. However, it should be noted that none of the existing solution addresses holistically the three challenges of color normalization in histopathology images and thus a new and complete color normalization approach is needed.

III. Proposed Color Normalization Scheme

The block diagram of the proposed normalization approach consisting of two processes is shown in Fig. 1:

1) The offline process defines a standard histopathology image preparation condition including information on $E_s(\lambda)$ and $M_s(\lambda)$, where the subscript s labels a standard condition. For operational flexibility, information on the standard condition can be defined via either predetermined quantities explicitly (for instance, the normalization example in Fig. 5), or a reference image[1] as a reduced set implicitly (Experiment 3 in Fig. 9, for example). In the diagram, a reference image I_s is used, where a standard

[1]Use of multiple references may improve the scheme performance. Such studies will be the subject of a future work.

TABLE I
SUMMARY OF COLOR NORMALIZATION METHODS

Category	Causes of color variation addressed		Histological information preservation	Reference image requirement	Representative reference
	illuminant variation	stain variation			
histogram matching	no	no	no	yes	[8], [9]
color transfer	no	no	no	yes	[10]–[12]
spectral matching	no	yes	maybe	no	[7], [13]

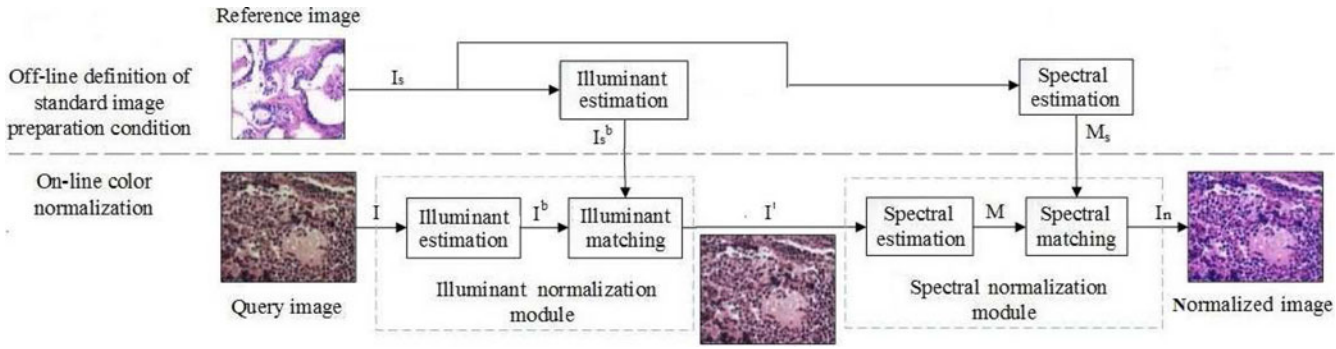

Fig. 1. Diagram of the proposed normalization pipeline, where the online normalization processing includes an illuminant normalization module and a stain spectral normalization module. The subscript s labels a standard condition for histopathology image preparation.

imaging illuminant and stain spectra are obtained by illuminant estimation and spectral estimation, respectively. Note, to qualify an image to be a reference, it must a) contain the same types of stains occurring in a query image, and b) have image background and histological components stained by different chemical dyes clearly presented. Otherwise, inaccurate, or even irrelevant, quantities are estimated in this offline process, finally affecting overall normalization performance.

2) The online process normalizes colors in a query image I. For each query image, the imaging illuminant and stain spectra are estimated and matched to reference quantities generated by the offline process. Since side information on query images' preparation, such as information about imaging device, is hardly guaranteed for large image sets, our study focuses on an operational scenario where no side information on a query image, other than knowledge about the stain type, is available to the system.

As shown in Fig. 1, the complicated normalization pipeline is formed by two modules, one to normalize imaging illuminant and the other to remove spectral variation in stains. It should be noted that following our analysis in Sections III-A and B, PSD of imaging light $E(\lambda)$ affects image color $I(p,\lambda_i)$ linearly in the linear RGB color space, while $I(p,\lambda_i)$ varies following an exponential function with respect to stain spectra $M(\lambda)$. Thus, it is necessary to identify the two independent factors, and propose linear and nonlinear algorithms to normalize color variation due to the two factors individually. Besides, as illuminant variation and stain variation are independent and can be addressed separately, if color variation is known to be introduced by one cause, an image can be processed by the corresponding module only.

A. Illuminant Normalization Module

Illuminant variation, corresponding to inconsistency in biopsy imaging in this study, introduces color bias in images. To remove the color bias, $E(\lambda)$ of a query image should be estimated and matched to a standard $E_s(\lambda)$, which is either estimated from a reference image or defined by a predetermined quantity such as the CIE illuminant D_{65} [26]. However, estimation of $E(\lambda)$ directly from an image is relatively difficult due to the integral effect of a camera. Based on the theory of metamer [27], we deduce an equivalent intensity matching algorithm to achieve illuminant normalization.

1) Illuminant Matching: Following (1), intensities of an image generated under illuminant with standard SPD are

$$I_s(p,\lambda_i) = \int_{\lambda_i-\delta}^{\lambda_i+\delta} f_i(\lambda)E_s(\lambda)e^{-M_i(\lambda)D(p)}d\lambda. \tag{2}$$

Applying the first mean value theorem for integration [28] to (1) and (2), we get

$$I(p,\lambda_i) = E(\epsilon_1)\int_{\lambda_i-\delta}^{\lambda_i+\delta} f_i(\lambda)e^{-M_i(\lambda)D(p)}d\lambda \tag{3}$$

$$I_s(p,\lambda_i) = E_s(\epsilon_2)\int_{\lambda_i-\delta}^{\lambda_i+\delta} f_i(\lambda)e^{-M_i(\lambda)D(p)}d\lambda \tag{4}$$

where $\epsilon_j \in (\lambda_i-\delta,\lambda_i+\delta)$ for $j=1,2$. Therefore

$$I(p,\lambda_i)/I_s(p,\lambda_i) = E(\epsilon_1)/E_s(\epsilon_2). \tag{5}$$

That is, we need to estimate a ratio of SPDs, $E(\epsilon_1)/E_s(\epsilon_2)$, to normalize image intensity $I(p,\lambda_i)$ to $I_s(p,\lambda_i)$. Since backgrounds, or blank areas, of histopathology images correspond

Algorithm 1: Illuminant Normalization

Input data : Linear RGB-format query image I, reference image I_s
Output data: Normalized query image I'
for *(each channel i, $i = 1, 2, 3$)* **do**
  $I_s^b(\lambda_i) \leftarrow$ Intensity estimation on $I_s(p, \lambda_i)$ (Eqn. (9));
  $I^b(\lambda_i) \leftarrow$ Intensity estimation on $I(p, \lambda_i)$ (Eqn. (9));
  if $I^b(\lambda_i) \geq th_b$ **then**
    $I'(p, \lambda_i) \leftarrow$ Intensity matching on $I(p, \lambda_i)$ (Eqn. (8));
  else
    return I with a notification.
  end
end

to tissues that are not bound by any stains, $D(p) = 0$. Hence

$$I^b(\lambda_i) = E(\epsilon_1) \int_{\lambda_i - \delta}^{\lambda_i + \delta} f_i(\lambda) d\lambda \tag{6}$$

$$I_s^b(\lambda_i) = E_s(\epsilon_2) \int_{\lambda_i - \delta}^{\lambda_i + \delta} f_i(\lambda) d\lambda \tag{7}$$

where I^b denotes background intensities in images. Thus, $I^b(\lambda_i)/I_s^b(\lambda_i) = E(\epsilon_1)/E_s(\epsilon_2)$ leads to

$$I_s(p, \lambda_i)/I(p, \lambda_i) = I_s^b(\lambda_i)/I^b(\lambda_i). \tag{8}$$

(8) is a formula to remove color disagreement caused by illuminant variation in histopathology images. Although it is simple, intuitively appealing and probably easy to infer, to the best of our knowledge, this study constitutes the first attempt to rigorously derive it using an imaging model.

2) Illuminant Estimation: According to (8), image intensities in blank areas need to be estimated for illuminant normalization. As $D(p) = 0$ in image blank areas, $I^b(\lambda_i) = \int_{\lambda_i - \delta}^{\lambda_i + \delta} f_i(\lambda) E(\lambda) d\lambda \geq \int_{\lambda_i - \delta}^{\lambda_i + \delta} f_i(\lambda) E(\lambda) e^{-M_i(\lambda) D(p)} d\lambda = I(p, \lambda_i)$. That is, $I^b(\lambda_i)$ is no smaller than any pixel values in the ith channel. Therefore, after removing image noise (e.g., shot noise) by a N-by-N square-shaped median filter h_N, the largest pixel value in the ith color channel is picked to estimate $I^b(\lambda_i)$. We formulate this illuminant estimation approach as

$$I^b(\lambda_i) = \max_p [I(p, \lambda_i) \otimes h_N] \tag{9}$$

where $\otimes$ denotes a convolution operation.

Though most histopathology images have blank areas, to ensure system robustness, we propose the use of a threshold th_b (e.g., $th_b = 200$ for eight bit data) to identify images without blank areas. In specific, if $I^b(\lambda_i)$ in (9) is smaller than th_b, it is likely that the image does not contain any blank area. For such a case, our algorithm will return the image with a notification to operators/pathologists for further analysis.

Combining the proposed intensity estimation in (9) and intensity matching formula in (8), the whole illuminant normalization module is described in Algorithm 1.

B. Spectral Normalization Module

If a set of biopsy samples stained by the same types of chemical dyes are imaged using one scanner, color variation in images are mainly caused by disagreement in biopsy staining. Though stains have their own diagnostic colors, such as hematoxylin usually appearing in blue while eosin in pink visually, due to inconsistency in stain manufacture, stain concentration, or storage condition, absorption spectra of stains, denoted as $M_i(\lambda)$ in (1), may differ, resulting color variation in images. Therefore, before quantitative analysis, color variation caused by inconsistent stain spectra $M_i(\lambda)$ should be removed.

When normalizing spectral variation in stains, care should be taken to preserve histological information. Though color variation among images is caused by different factors, within one image, these factors are constant and histological information is conveyed by colors. As various colors in an image are generated by different combinations of stains, histological information is actually conveyed by stain depths at each pixel, which is denoted by $D(p)$ in (1). To normalize inconsistent stain spectra $M_i(\lambda)$ while to maintain stain proportion $D(p)$ unchanged, we concatenate a NMF-based spectral estimation and spectral matching.

To simplify the problem, in this section, we assume images are generated under standard imaging condition. Otherwise, the proposed illuminant normalization will be applied first.

1) Spectral Matching: We apply the first mean value theorem for integration to (1) and factor out $e^{-M_i(\lambda) D(p)}$

$$I(p, \lambda_i) = e^{-M_i(\epsilon_3) D(p)} \int_{\lambda_i - \delta}^{\lambda_i + \delta} f_i(\lambda) E(\lambda) d\lambda \tag{10}$$

where $\epsilon_3 \in [\lambda_i - \delta, \lambda_i + \delta]$. Observing that the integral part of (10) corresponds to background intensities in images

$$I(p, \lambda_i) = I^b(\lambda_i) e^{-M_i(\epsilon_3) D(p)}. \tag{11}$$

(11) has the exact same form as the Beer–Lambert law [29], whereas has different physical meaning. The Beer–Lambert law describes an optical phenomenon that light is absorbed by materials when it travels through; while (11) relates a digital quantity, image intensity, to stain absorption spectra. For a biopsy stained by chemical dyes that have standard spectra $M_{si}(\lambda)$, its corresponding image has intensity

$$I_s(p, \lambda_i) = I^b(\lambda_i) e^{-M_{si}(\epsilon_4) D(p)} \tag{12}$$

for $\epsilon_4 \in [\lambda_i - \delta, \lambda_i + \delta]$. For histological information preservation, (11) and (12) share the same $D(p)$. Hence

$$I_s(p, \lambda_i) = I^b(\lambda_i) \times \left[\frac{I(p, \lambda_i)}{I^b(\lambda_i)} \right]^{\frac{M_{si}(\epsilon_4)}{M_i(\epsilon_3)}}. \tag{13}$$

(13) formulates a spectral matching method to remove stain variation in histopathology images, meanwhile preserving histological information conveyed by stain proportions. Then the left problem is to estimate stain spectra from images.

2) NMF-Based Spectral Estimation: Let us reexamine (11). By moving $I^b(\lambda_i)$ to the left in (11) and taking logarithm on both sides, the optical density (OD) domain equivalent expression is as follows:

$$\log(I^b(\lambda_i)) - \log(I(p, \lambda_i)) = M_i(\epsilon_3) \times D(p). \tag{14}$$

For a biopsy sample stained by N types of stains and imaged via K sensors, we have K equations in the form of (14) for its corresponding histopathology image. Then a K-by-N spectrum

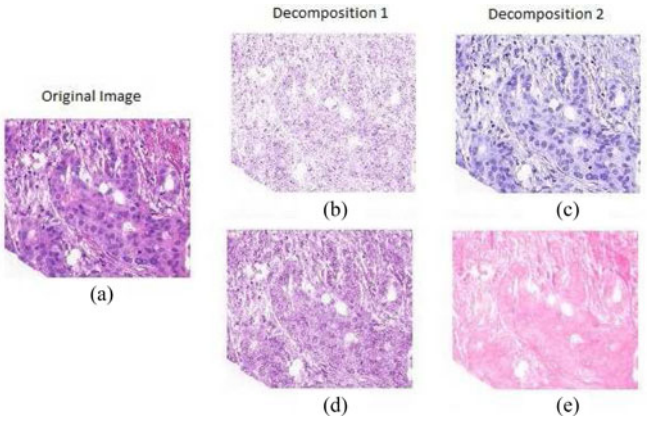

Fig. 2. (a) H&E stained breast cancer image. (b)–(e) NMF decompositions with random initializations. The first row of the decomposition results corresponds to hematoxylin images, while the second row to eosin images.

matrix M is formed by using $M_i(\lambda)$ as the ith row. Hence, an element m_{ij} in a spectrum matrix M represents the jth stain's absorption spectrum around particular wavelength λ_i and spectrum of the jth stain, M_{ij} can be characterized by the jth column of M. For instance, an RGB-format H&E stained image has a three-by-two spectrum matrix, whose first column corresponds to spectrum vector of hematoxylin and the second column to eosin. Let $I^b = [I^b(\lambda_1), \ldots, I^b(\lambda_K)]'$ and $I(p) = [I(p, \lambda_1), \ldots, I(p, \lambda_K)]'$, then

$$\log(I^b) - \log(I(p)) = M \times D(p). \tag{15}$$

Following (15), given a query image, we need to factorize $\log(I^b) - \log(I(p))$ into spectrum matrix M and stain depth $D(p)$. The fact that both spectrum matrix M and stain proportion $D(p)$ at any location should be nonnegative motivates us to exploit NMF [30] to achieve an optimal spectral estimation in terms of minimizing the mean-squared error

$$\arg \min_{M, D>=0} \text{Dist}_2 \left[\log(I^b) - \log(I(p)) \| M \times D(p)\right] \tag{16}$$

where $\text{Dist}_2(Y \| AX) = \|Y - AX\|^2$. As the quadratic function in (16) is a nonconvex function of both M and $D(p)$ simultaneously, depending on initial matrix M^{in}, solutions of NMF may converge to any local minima, resulting in incorrect solutions [31]. For example, Fig. 2 shows two spectral estimation results computed by NMF using two randomly generated spectrum matrices. The first row of decomposition results are supposed to be images of histological components stained by hematoxylin, and the second row to eosin-stained images. However, the two decompositions differ a lot. Therefore, a good initial spectrum matrix for NMF iteration is significant.

To reduce the probability of converging to local optimum instead of a global optimal solution for NMF computation, the initial spectrum matrix should be image dependent and close enough to the true stain spectra. For this end, we propose a novel SW statistical method in the HSV color space to identify reliable colors in an image implicitly for stable estimation of stain spectra M^{in}. The reason behind our innovation is that achromatic pixels in histopathology images are usually noise or

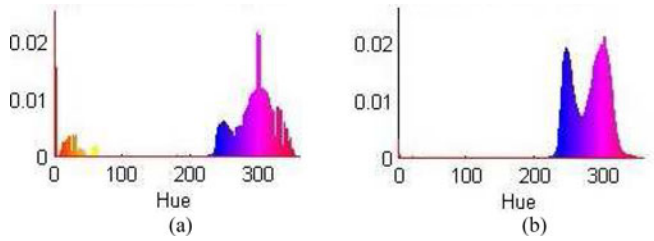

Fig. 3. (a) Standard hue histogram of the reference image in Fig. 1. (b) SW-hue histogram of the same image. The SW-hue histogram is much smoother due to its strong resistance to achromatic colors.

locations bounded by few stains, and spectral estimation using these less-saturated colors is unreliable, or even meaningless. However, at pixels where spectra of imaging light are absorbed by stains, colors are usually highly saturated and more reliable for estimation.

In specific, colors in histopathology images are generated by transmitted light that is not absorbed by stains [11]. Hence, a hue histogram of an image actually shows what spectra of light are not absorbed, and thus can be used to estimate matrix M^{in}. To limit impacts of achromatic colors on spectral estimation, rather than using standard histogram, we propose the use of an SW hue histogram [32] defined as follows:

$$\text{Hist}_\theta^{\text{sw}} = \sum_p s_p \delta(\theta, h_p) \tag{17}$$

$$\text{where} \quad \delta(\theta, h_p) = \begin{cases} 1, & \text{if } \theta = h_p \\ 0, & \text{otherwise} \end{cases}$$

$\theta \in [0^\circ, \ldots, 360^\circ)$ represents a bin in histogram, s_p and h_p are the saturation and hue at pixel p in the HSV color space. Fig. 3 shows histograms computed using standard statistics and the SW method in (17) for the reference image in Fig. 1. The standard hue histogram in Fig. 3(a) has more spikes contributed by achromatic pixels; while the SW-hue histogram in Fig. 3(b) depicts smooth dominant colors which are more coherent to human's color perception.

With the SW-histogram Hist^{sw} showing the light spectra not absorbed by any stains, we proceed to estimate the typical color for each stain. As clusters in the SW-color histogram bias highly-saturated colors associated with different chemical dyes, k-mean clustering is applied to Hist^{sw} to estimate the N representative hues $h_1^{\text{sw}}, \ldots, h_N^{\text{sw}}$. Aware that k-mean clustering has an inconsistent convergence problem due to poor starters, we use prior knowledge of stain spectra as its initialization. For instance, hue values of purple-to-blue and pink can be used as cluster centers of H&E initially for an H&E stained image. Then for each hue h_i^{sw}, we propose SW means of value v_i^{sw} and saturation s_i^{sw} in the HSV domain, defined as

$$v_i^{\text{sw}} = \Big[\sum_p s_p v_p \delta(h_i^{\text{sw}}, h_p)\Big] / \Big[\sum_p s_p \delta(h_i^{\text{sw}}, h_p)\Big] \tag{18}$$

$$s_i^{\text{sw}} = \Big[\sum_p (s_p)^2 \delta(h_i^m, h_p)\Big] / \Big[\sum_p s_p \delta(h_i^{\text{sw}}, h_p)\Big] \tag{19}$$

Algorithm 2: Initialization using SW-Statistics

Input data : Linear RGB-format image I', stain type N
Output data: Initial stain spectrum matrix M^{in}
Convert I' to the HSV domain via non-linear RGB space [27] ;
$Hist^{sw} \leftarrow$ Compute SW-hue histogram (Eqn. (17)) ;
$h_i^{sw} \leftarrow$ K-mean clustering on $Hist^{sw}$ to find N centers ;
for *(each* h_i^{sw}, $i = 1, .., N$*)* **do**
  $v_i^{sw} \leftarrow$ SW-statistics of value (Eqn. (18));
  $s_i^{sw} \leftarrow$ SW-statistics of saturation (Eqn. (19));
  $c_i \leftarrow$ Convert $[h_i^{sw}, s_i^{sw}, v_i^{sw}]$ to the linear RGB space [27];
  $M_{:i} \leftarrow$ Convert c_i to the OD domain via $log(I'^b/c_i)$;
end
$M^{in} \leftarrow$ Form matrix using all $M_{:i}$ as column;
Normalize M^{in}.

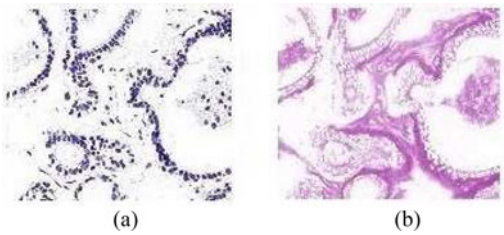

Fig. 4. Stain decomposition results merely rely on the initial spectrum matrix $M^{in} = [0.6322\ 0.3023; 0.7354\ 0.9045; 0.2438\ 0.3008]$ estimated through SW-statistics for the reference H&E stained images in Fig. 1.

to fully determine the representative stain colors. In (18) and (19), saturation values are used as filter parameters and resulting representative colors bias saturated image colors. Finally, we convert representative colors $[h_i^{\text{SW}}, s_i^{\text{SW}}, v_i^{\text{SW}}]$ of stains from the HSV color space to the OD domain and form an initial spectrum matrix M^{in} for NMF. The whole initialization algorithm is outlined in Algorithm 2.

We use the reference image in Fig. 1 as an example to demonstrate the performance of the proposed initialization algorithm. The resulted three-by-two H&E spectrum matrix is $M^{in} = [0.6322\ 0.3023; 0.7354\ 0.9045; 0.2438\ 0.3008]$, where the first column corresponds to a spectrum vector of hematoxylin, and the second to eosin spectrum vector. To visualize this result, the corresponding stain decomposition result is shown in Fig. 4. Though the resulting M^{in} is just an initialization for subsequent NMF computation, the original image is well separated into two stain channels which correspond to H&E, respectively.

The overall spectral normalization module on the basis of NMF-based spectral estimation is descried in Algorithm 3.

IV. Experimental Results and Discussions

We design three experiments to evaluate the proposed normalization approach. First, we assess the robustness of our normalization method to system settings, and demonstrate that our method is stable under various parameter settings. In the second experiment, since spectral estimation plays a significant role in spectral normalization, effectiveness and consistency of the proposed stain decomposition is examined and compared to state-of-the-art adaptive stain estimation algorithms. In the last experiment, performance of the proposed normalization approach, especially the capability of histological information preservation, is evaluated. This experiment is also performed on four representative normalization approaches for comparison.

Algorithm 3: Spectral Normalization

Input data : Linear RGB-format query image I', reference image I_s, stain type N
Output data: Normalized query image I_n
$M_s^{in}, M^{in} \leftarrow$ Apply Algorithm 2 to I_s and I';
$M_s \leftarrow$ NMF on I_s with initial M_s^{in} [see (16)];
$M \leftarrow$ NMF on I' with initial M^{in} [see (16)];
$I_n \leftarrow$ Apply spectral matching to I' [see (13)].

Two public histopathology image sets are used as experimental data in this study. They are selected in different experiments depending on experimental objectives.

1) The NIA malignant lymphoma dataset can be accessed from the IICBU Biological Image Repository [33]. It contains 20x magnified H&E stained images with resolution of 1388×1040, including 113 chronic lymphocytic leukemia images, 140 follicular lymphoma cases, and 122 mantle cell lymphoma slides. Because biopsies in this dataset were prepared by different pathology laboratories, significant variations in image colors are observed. This dataset is believed more representative of histopathology images commonly received in clinics [34].
2) The UCSB breast cancer cell dataset [35] was published for an objective of cell segmentation in both benign and malignant cell images. The dataset consists of 26 cancerous cell images and 32 normal cases cut from ten H&E stained breast cancer biopsies. All images are scanned in the same laboratory, with resolution of 896×768.

As the NIA dataset and UCSB dataset store images in 24-b nonlinear RGB format, all images are converted to the linear RGB domain first [27]. All algorithms in this paper are implemented in Matlab.

A. Experiment 1: System Robustness to Parameter Settings

There are two sets of parameters predetermined in our normalization method: median filter dimension in the illuminant normalization module, and NMF convergence condition in the spectral estimation algorithm. Usually, performance of an algorithm varies with parameter selection. In medical-related applications, we want this performance variation small. Hence, we perform two tests in this experiment and assess sensitivity of the proposed approach to the two sets of parameters.

Dataset: Both the NIA image set and the UCSB dataset are used as testing sets in Experiment 1.

1) Test 1: Dimension of Median Filter: The objective is to assess sensitivity of the illuminant normalization module to the median filter with neighbor-supporting in N-by-N square area. Following common evaluation procedure, filter dimension N is changed while the square-shape of the filter is maintained.

Experimental Design: In the illuminant normalization module, a median filter is applied to an image to remove potential

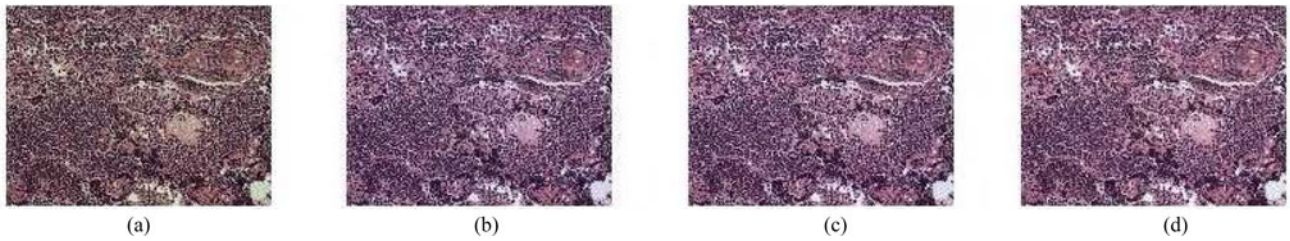

Fig. 5. Example of illuminant normalization with median filters in various sizes, where reference imaging illumiant is defined explicitly by CIE D_{65}. (a) Query image randomly selected from the NIA lymphoma dataset. (b)–(d) Illuminant-normalized images generated using median filters with $N = 3, 5, 7$.

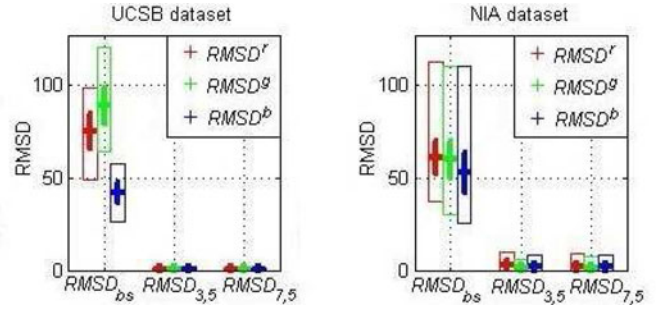

Fig. 6. RMSDs statistics of both datasets. Bars and crosses represent ranges and means (with variances) of RMSD between normalized image. RMSD_{bs} characterizes expected difference between distinct images within a dataset.

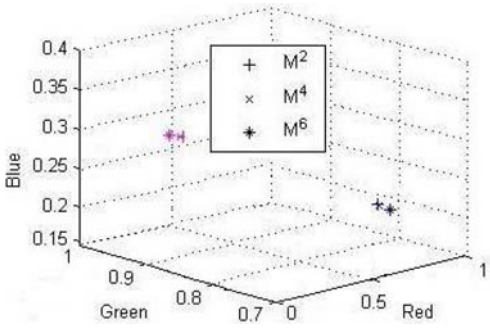

Fig. 7. Stain spectral estimations using different NMF convergence conditions for the reference image in Fig. 1. M^i represents spectra estimated under $T_y = T_x = 10^{-i}$, and blue and pink correspond to H&E.

noise, while image local statistics should be maintained. To this end, $N = 3, 5, 7$ are selected as candidates of filter size. To quantify sensitivity of the proposed illuminant normalization to N, we compute the root mean square difference (RMSD) for each image, $\text{RMSD}^c_{i,j} = \sqrt{E[(I'^c_i - I'^c_j)^2]}$, where I'^c_i represents color component in the c channel of an illuminant-normalized image I' generated using filter parameter $N = i$. For example, $\text{RMSD}^r_{i,j} = 0$ indicates two normalized images having identical red image components. Then statistics of $\text{RMSD}_{i,j} = [\text{RMSD}^r_{i,j}, \text{RMSD}^g_{i,j}, \text{RMSD}^b_{i,j}]$ over a dataset are summarized and compared to the baseline RMSD_{bs}, which is defined as RMSDs of distinct normalized images generated by the same filter in a dataset. In this test, reference imaging light is defined explicitly by CIE illuminant D_{65} [26].

Results and Discussion: Fig. 5 shows an example of illuminant-normalized images generated using median filters with $N = 3, 5, 7$, where $\text{RMSD}_{3,5} = [5.090, 3.013, 1.982]$ and $\text{RMSD}_{7,5} = [4.873, 3.201, 1.379]$. For an RGB-formant image with 8-b per channel, the image difference is very small. Fig. 6 summarizes RMSD statistics in both testing image sets. Compared to RMSD_{bs} which quantifies typical difference between two 24-b normalized images generated by a filter with $N = 5$, small $\text{RMSD}_{3,5}$ and $\text{RMSD}_{7,5}$ imply that the illuminant normalization module is insensitivity to the filter size, and we use $N = 5$ in all of our experiments.

2) Test 2: Convergence Conditions of NMF Iterations: In the proposed spectral estimation algorithm, NMF iteration terminates when either the root mean square residue of factorization $\sqrt{E(\text{Dist}_2(Y\|AX))}$ is smaller than a predetermined value T_y, or the relative changes of factors A or X in NMF computation is small enough to achieve a prefixed tolerance T_x. In this test, effect of different NMF convergence conditions (T_y, T_x) on our spectral estimation is examined.

Experimental Design: Convergence conditions $(T_y = T_x = 10^{-i})$ for $i = 2$, 4, 6 are applied to the NMF computation of stain spectrum matrices M^i on testing image sets. Specifically, for the kth testing image I_k in a dataset, absolute differences between resulting spectrum matrices, $dif_{k,k}(i, j) = |M^i_k - M^j_k|$ for $i, j = 2$, 4, 6, are computed. Then statistics of $dif_{k,k}(i, j)$ over a testing set are summarized and compared to baseline $dif_{k,l}(i, i) = |M^i_k - M^i_l|$ for $k \neq l$, which quantifies stain variation estimated from distinct biopsies in a dataset.

Results and Discussion: Fig. 7 depicts spectrum vectors estimated under different NMF termination conditions for the reference image in Fig. 1. Vectors colored in blue and pink, which represent spectra of H&E, respectively, form two very tight clusters. We observe similar results for other testing images. Table II summaries the statistics of $dif_{k,k}(i, j)$ for $i, j = 2$, 4, 6 in the UCSB dataset and the NIA dataset, where a 3×2 H&E spectrum matrix is formed by three row vectors grouped by ";". Compared to $dif_{k,l}(4, 4)$, the typical spectral difference between distinct H&E stained biopsies in the order of 10^{-2}, expectations of both $dif_{k,k}(2, 4)$ and $dif_{k,k}(6, 4)$ in the order of 10^{-3} are much smaller, indicating that M^i_k for $i = 2, 4, 6$ convey the same information on stains' spectra in an image. Therefore, we conclude that the proposed spectral estimation is insensitive to NMF convergence conditions, and $T_y = T_x = 10^{-4}$ are used in all experiments.

B. Experiment 2: Consistency of Spectral Estimation

To identify contribution of each stain to color variation, accurate spectral estimation is crucial. As histopathology images

TABLE II
STATISTICS OF SPECTRUM MATRIX DIFFERENCES DUE TO VARIOUS NMF CONVERGENCE CONDITIONS IN THE UCSB DATASET AND THE NIA DATASET

UCSB dataset			
Statistics	$\mathrm{dif}_{k,l}(4,4)$	$\mathrm{dif}_{k,k}(2,4)$	$\mathrm{dif}_{k,k}(6,4)$
mean ($\times 10^{-2}$)	[6.92 11.47; 5.65 3.58; 2.04 3.17]	[0.10 0.34; 0.09 0.49; 0.02 0.51]	[0.99 2.00; 0.80 0.74; 0.26 0.44]
std ($\times 10^{-2}$)	[5.19 8.87; 4.09 3.12; 1.52 3.37]	[0.23 1.11; 0.23 3.16; 0.04 3.78]	[0.47 0.66; 0.40 0.75; 0.12 1.70]
min ($\times 10^{-2}$)	[0.02 0.01; 0.05 0.02; 0.00 0.01]	[0.02 0.05; 0.02; 0.01; 0.00 0.00]	[0.42 0.92; 0.42 0.12; 0.09 0.01]
max ($\times 10^{-2}$)	[24.34 55.51; 19.48 26.12; 6.67 29.78]	[1.77 8.59; 1.79 24.15; 0.31 28.83]	[3.50 3.68; 3.50; 5.87; 0.77 13.04]
NIA Lymphoma dataset			
Statistics	$\mathrm{dif}_{k,l}(4,4)$	$\mathrm{dif}_{k,k}(2,4)$	$\mathrm{dif}_{k,k}(6,4)$
mean ($\times 10^{-2}$)	[4.04 6.76; 2.75 2.51; 2.53 4.26]	[0.20 0.75; 0.10 0.24; 0.10 0.22]	[0.83 2.06; 0.39 0.54; 0.51 0.83]
std ($\times 10^{-2}$)	[4.3 7.31; 2.11 2.06; 7.31 4.86]	[0.17 0.73; 0.07 0.24; 0.09 0.13]	[0.28 0.94; 0.14 0.28; 0.23 0.44]
min ($\times 10^{-2}$)	[0.00 0.01; 0.01 0.01; 0.00 0.00]	[0.03 0.02; 0.02 0.04; 0.01 0.00]	[0.10 0.02; 0.03 0.01; 0.02 0.00]
max ($\times 10^{-2}$)	[27.58 43.59; 17.14 15.38; 19.36 30.25]	[1.96 6.75; 0.78 2.13; 1.08 0.91]	[2.86 8.39; 0.93 2.94; 2.43 3.62]

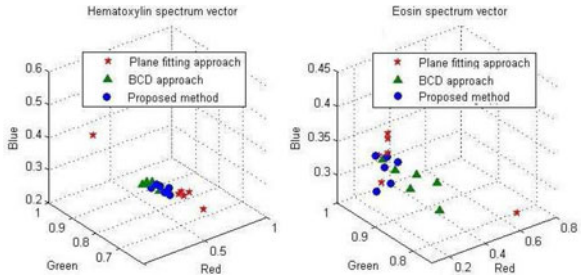

Fig. 8. Stain spectrum vectors of six H&E stained images estimated by the plane fitting algorithm [7], the BCD approach [25], and the proposed method. Since the six images were cut from a biopsy labeled as ytma10-010704 in the UCSB dataset, estimated stain spectra should be very close.

may contain different contents, we want the spectral estimation is not affected by content difference. Therefore, this experiment evaluates consistency of our spectral estimation against different image contents.

1) Dataset: The UCSB dataset is selected as an evaluation set for reasons as follows. The UCSB images were cut from ten biopsies. For images from the same biopsy, a spectral estimation algorithm should generate consistent, or at least similar, spectrum matrices. Thus, we use consistency of spectrum matrices estimated from images cut from the same biopsy as a metric to measure performance of a stain estimation algorithm.

2) Experimental Design: We form ten sets of images in the UCSB dataset, each containing all images from the same biopsy. Then standard deviation σ of estimated matrices in each set is computed. A smaller σ indicates more consistent estimation. This experiment is also performed on state-of-the-art blind stain estimation algorithms: the plane fitting approach [7] and BCD approach [25]. For a fair comparison, we set parameters $\alpha = 0.05$ and $\beta = 0.15$ in the plane fitting approach and $\sigma_k^{\infty} = \sqrt{1/12}$ and $N = 10$ in BCD method following the original papers, and tried our best to faithfully reproduce the experimental results in both papers.

3) Results and Discussion: Spectral estimation results of the first image set are shown in Fig. 8. Since similar results were obtained, estimation results of other nine sets were omitted here. As the first set contains six H&E stained images cut from a biopsy labeled as ytma10-010704, each approach generates six pairs of spectrum vectors, among which spectra of hematoxylin are depicted in the left panel and eosin spectra are in the right panel. In the figure, both H&E spectrum vectors estimated by the proposed method form tight clusters, while estimations computed by the plane fitting approach show the largest diversity. Table III summarizes standard deviations of spectral estimation over the ten image sets, where the smallest σ of each matrix element is marked black. The data illustrate that our spectral estimation method generates more consistent results than the other two methods.

The relatively inconsistent performance of the plane fitting method [7] is attributed to the thresholding mechanism in stain estimation, where only image pixels whose colors are close to predetermined thresholds contribute to spectral estimation. However, for images containing different contents, the predetermined values may be inappropriate, resulting inaccurate and inconsistent spectral estimation. The estimation consistency of the BCD algorithm [25] lies between the plane fitting method and our method. In the BCD approach, a heuristic randomization function assigns large weights to colors in high OD values. Then colors with large weights are selected for stain estimation. Compared to colors associated with hematoxylin that are usually in larger OD values, eosin-stained tissues often have weak stains, and thus colors mainly associated with eosin usually have smaller weights. For images containing large blank areas, significance of eosin-stained pixels may be overwhelmed by achromatic colors. This explains the less consistency in eosin spectrum vectors estimated by the BCD algorithm. The consistent performance of our estimation method is attributed to the proposed saturation-weight statistical method. Since SW-statistics identifies saturated colors implicitly, impacts of achromatic pixels and noise on spectral estimation are alleviated. The strong noise-resistant capability of the proposed algorithm inherited from the SW-statistics ensures consistency of spectral estimation, and will benefit effectiveness and reliability of the proposed normalization method.

C. Experiment 3: Color Normalization Performance

One unique issue of histopathology image normalization is to preserve histological information. This experiment examines

TABLE III
STANDARD DEVIATIONS OF SPECTRAL ESTIMATIONS ON THE TEN IMAGE SETS OF THE UCSB DATASET

Biospy id	Plane fitting algorithm [7]	BCD approach [25]	Proposed method
ytma10-010704	[3.04 **2.72**; 2.23 2.01; 2.67 5.79]	[**2.54** 7.64; **1.41** 1.84; **0.97** 2.66]	2.67 3.25; 1.68 **0.97**; 1.14 **2.43**
ytma12-010804	[1.99 **3.79**; 2.16 2.82; 1.72 6.81]	[2.14 7.02; 1.58 3.03; 0.67 1.66]	**1.37** 5.79; **1.07 1.82**; **0.55 1.30**
ytma23-022103	[1.77 **3.43**; 1.95 **0.46**; 1.31 3.34]	[**1.47** 8.08; **1.33** 3.78; 0.89 2.26]	2.39 5.36; 2.25 1.59; **0.70 1.57**
ytma49-042003	[**3.11** 2.11; 2.98 2.02; 1.09 3.67]	[5.08 7.35; 3.38 3.00; 0.95 1.60]	3.17 **1.86**; **2.35 0.45**; **0.55 0.98**
ytma49-042203	[**2.29 2.18**; 2.95 1.40; 1.88 3.72]	[3.64 5.69; **2.94** 2.33; **0.79 1.97**]	4.04 3.13; 3.81 **0.92**; 1.15 2.12
ytma49-042403	[**3.08 3.13**; 3.16 3.18; 1.04 5.38]	[3.82 6.77; 2.70 2.89; **0.61** 1.37]	3.42 3.67; **2.55 0.96**; 0.63 **0.47**
ytma49-072303	[**2.09** 2.63; 2.48 1.22; 0.57 1.43]	[2.95 **2.19**; 2.84 1.25; 0.42 **0.64**]	2.49 2.82; **0.95**; **0.38** 0.80
ytma49-0111003	[1.90 13.17; 2.60 11.29; 3.62 10.73]	[1.97 6.96; 2.03 4.80; **0.57 1.96**]	**1.64 6.94 1.61 3.49**; 1.23 2.79
ytma49-0111303	[2.15 **6.12**; 2.52 **1.75**; 1.50 4.30]	[1.28 6.24; 1.32 3.64; 0.26 **1.60**]	**1.20** 6.71; **1.30** 2.46; **0.24 2.34**
ytma55-030603	[2.18 7.85; 3.39 3.90; 2.73 8.20]	[**1.52 3.59**; **1.36** 2.46; 0.72 **1.15**]	1.88 16.56; 1.81 **2.09**; **0.67** 8.32

All data are in the scale of 10^{-2}, and the smallest σ of each matrix element is marked black.

performance of the proposed normalization method, particularly the capability of histological information preservation.

Dataset: All images in the UCSB dataset and the NIA dataset are selected as testing images in this experiment.

1) Experimental Design: We first select a slide in the UCSB dataset as a reference image, from which standard imaging light and absorption spectra of stains are estimated. Then the proposed normalization pipeline is tested in three scenarios.

1) In the first test, query images are those slides cut from the same biopsy as the reference in the UCSB dataset. Thus, they should not be modified after normalization.
2) Then, UCSB breast cancer images other than those used in the first test are used as query images. After normalization, spectral variation in H&E stains should be removed, while histological features should be maintained.
3) The last test is performed on the NIA images, which are more representative of histopathology images received in clinics. Color variation jointly caused by inconsistency in biopsy staining and imaging is expected to be removed.

The three tests are also performed on histogram matching method [8], color map quantile matching method [9], color transfer method [11], and spectral matching method [7].

2) Results and Discussion: Fig. 9 shows color normalization results of the three tests. As similar observations are obtained in each test, one example in each test is presented here. In the figure, normalized images in test 1 are shown in the first column. Images in the middle column correspond to the normalized results in test 2, and images in the last column give examples of color normalization in the third test. Our observations are summarized as follows.

1) In test 1, normalized images should be the same as the query images. However, this is achieved by the proposed method only. Other normalization methods modify the query image in various degrees.
2) The example query image in test 2 has its eosin-stained components mainly located in the low half image, and nuclei distributed in the upper left area. Due to stain variation, colors associated with eosin are more magenta and nuclei appear more blue compared to the reference. After normalization, histological components stained by eosin and hemotoxylin should appear in pink and purple, respectively. Among normalized images in the middle column, only our method succeeds in color normalization.
3) Test 3 is a more complicated normalization problem. Again, the first four methods fail to preserve histological information in the query image, and only the proposed normalization solution provides a good normalized result.

We examine these normalized methods and investigate reasons for our observations. In normalized images generated by histogram matching approach [8], a large amount of cytoplasm stained by eosin in query images is normalized to pure white and disappears. This phenomenon is due to an implicit assumption behind histogram matching that all images have the same color distribution. However, since colors in histopathology images are dependent on tissue contents, this assumption does not hold. For example, the reference image contains more blank areas than query images. Therefore, the histogram matching method modifies query images such that the normalized images have a similar number of pure white pixels. In the sense of histological information preservation, histogram matching-based methods have poor performance.

Color map quantile matching method discards information on color frequency in images and adjusts the color map of a query image according to the color map of a reference [9]. However, as a variation of histogram matching, this method has a similar limitation as the histogram matching approach. The nuclei in the resulting normalized images either appearing in pink or fading into background demonstrates its weak capability of histological information preservation.

Color transfer method in [11] matches the mean and variance in every decomposition channel of a query image to the statistics of a reference image. Since images have different content, stain depths in each decomposition channel are modified after color transfer, consequently introducing color distortions in normalized images. Color of nuclei normalized to cyan rather than purple in the first two tests, and some cytoplasm disappearing in test 3 are such examples.

The spectral matching method in [7] adopts the plane fitting method to estimate stain spectra. As less accurate spectral estimation may be caused by inappropriate predetermined thresholds in the plane fitting method, small color distortion in normalized images may occur. For example, compared to the reference image, normalized images of test 1 and test 2 shown in

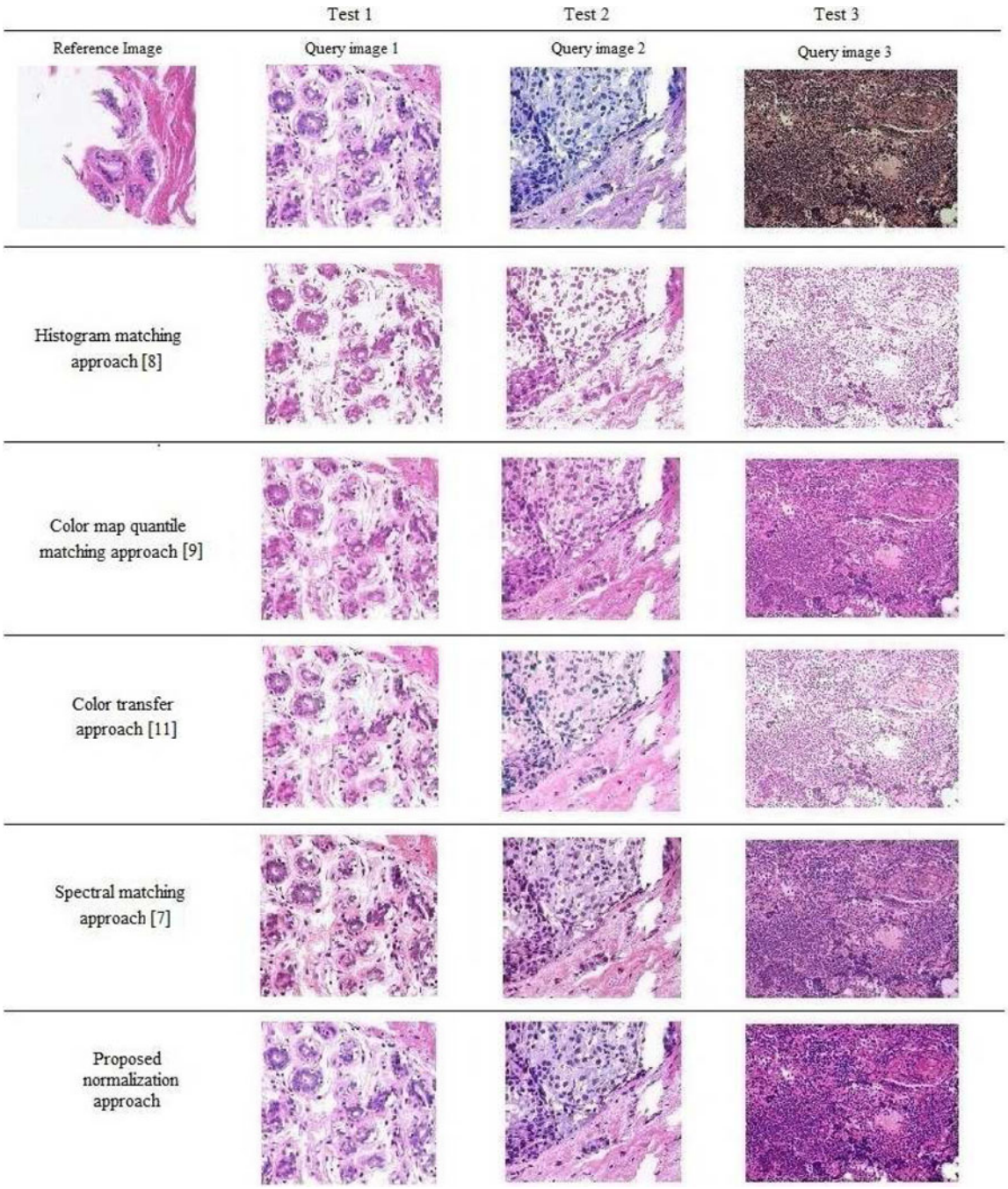

Fig. 9. Comparison of color normalization methods. The reference image is selected from the UCSB dataset. Images in the first column give examples of the normalization results in test 1. The normalized images in the second column correspond to the second test, and so on so forth.

Fig. 9 have darker nuclei. In addition, as shown in the third test, blank areas in the normalized image generated by the spectral matching approach have a bias to pink. This is because spectral matching method cannot appropriately address color variation caused by nonstandard imaging condition.

Based on discussions, we conclude that the proposed normalization method is effective to remove color variation in histopathology images. It outperforms other methods in terms of histological information preservation, and would be useful for histopathology image analysis tasks such as color-based histological component detection and segmentation.

V. Conclusion

This study introduced a robust and complete color normalization approach capable of addressing color variation in histopathology images. Based on an imaging model, the introduced solution was able to identify the source of color variation, and addressed effectively both illuminant variation and stain variation using an intensity matching algorithm and a spectral normalization module, respectively. Extensive experimentation on publicly available datasets indicates that the proposed solution outperforms state-of-the-art color normalization solutions, while preserving histological information.

The SW stain estimation method introduced by this study limited impacts of achromatic spectra on stain estimation and robustified the estimation process. The experimental results of stain estimation indicate that the proposed method delivers superior consistent performance compared to state-of-the-art blind stain decomposition solutions.

APPENDIX
IMAGING MODEL OF HISTOPATHOLOGY IMAGES

In histopathology imaging, a camera captures incident light that is not absorbed by stains and forms an image. The transmitted light arriving at each location $p = (x, y)$ on the sensor can be specified by its SPD $E'(p, \lambda)$. Assuming a sensing device consisting of a lens whose sensor spectral sensitivity is $f_i(\lambda) > 0$ for $\lambda \in (\lambda_i - \delta, \lambda_i + \delta)$, then image value at pixel p is formulated as follows [36]:

$$I(p, \lambda_i) = \int_{\lambda_i - \delta}^{\lambda_i + \delta} f_i(\lambda) E'(p, \lambda) d\lambda. \tag{20}$$

Let $E(\lambda)$ denote SPD of imaging illuminant. When chemical dyes used at staining absorb light only, according to the Beer–Lambert law, $E'(p, \lambda) = E(\lambda) T(p, \lambda)$, where $T(p, \lambda) = e^{-m_i(\lambda) d(p)}$ denotes the absorption of light by a stain and $m_i(\lambda)$ is the spectral mass absorptivity of a stain to the incident light in wavelength λ_i. If N types of such stains are used in chemical staining, $T(p, \lambda) = e^{-\sum_{j=1}^{N} m_{ij}(\lambda) d_j(p)}$. Let $M_i(\lambda) = [m_{i1}(\lambda), \ldots, m_{iN}(\lambda)]$ represent the absorbing spectra of N types of stains at wavelength λ_i and $D(p) = [d_1(p), \ldots, d_N(p)]'$ correspond to the stain depths at location p, $T(p, \lambda) = e^{-M_i(\lambda) D(p)}$. Substituting this to (20), we get a microscopic imaging formula

$$I(p, \lambda_i) = \int_{\lambda_i - \delta}^{\lambda_i + \delta} f_i(\lambda) E(\lambda) e^{-M_i(\lambda) D(p)} d\lambda. \tag{21}$$

REFERENCES

[1] K. Jafari-Khouzani and H. Soltanian-Zadeh, "Multiwavelet grading of pathological images of prostate," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 6, pp. 697–704, Jun. 2003.

[2] A. Basavanhally *et al.*, "Computerized image-based detection and grading of lymphocytic infiltration in her2+ breast cancer histopathology," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 3, pp. 642–653, Mar. 2010.

[3] K. Nguyen *et al.*, "Structure and context in prostatic gland segmentation and classification," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, 2012, pp. 115–123.

[4] L. Gorelick *et al.*, "Prostate histopathology: Learning tissue component histograms for cancer detection and classification," *IEEE Trans. Med. Imag.*, vol. 32, no. 10, pp. 1804–1818, Oct. 2013.

[5] M. Veta *et al.*, "Breast cancer histopathology image analysis: A review," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 5, pp. 1400–1411, May 2014.

[6] M. N. Gurcan *et al.*, "Histopathological image analysis: A review," *IEEE Rev. Biomed. Eng.*, vol. 2, pp. 147–171, Oct. 2009.

[7] M. Macenko *et al.*, "A method for normalizing histology slides for quantitative analysis," in *Proc. Int. Symp. Biomed. Imaging*, 2009, pp. 1107–1110.

[8] A. Tabesh *et al.*, "Multifeature prostate cancer diagnosis and gleason grading of histological images," *IEEE Trans. Med. Imag.*, vol. 26, no. 10, pp. 1366–1378, Oct. 2007.

[9] S. Kothari *et al.*, "Automatic batch-invariant color segmentation of histological cancer images," in *Proc. Int. Symp. Biomed. Imaging*, 2011, pp. 657–660.

[10] Y. Wang *et al.*, "A color-based approach for automated segmentation in tumor tissue classification," in *Proc. Eng. Med. Biol. Soc.*, 2007, pp. 6577–6580.

[11] D. Magee *et al.*, "Color normalization in digital histopathology images," in *Proc. Med. Image Comput. Comput. Assisted Intervention*, 2009, pp. 100–111.

[12] A. M. Khan *et al.*, "A non-linear mapping approach to stain normalization in digital histopathology images using image-specic color deconvolution," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 6, pp. 1729–1738, Jun. 2014.

[13] S. Tani *et al.*, "Color standardization method and system for whole slide imaging based on spectral sensing," *Analytical Cellular Pathology*, vol. 35, no. 2, pp. 107–115, 2012.

[14] A. Basavanhally and A. Madabhushi, "Em-based segmentatio-driven color standardization of digitized histopathology," *Proc. SPIE*, vol. 8676, 2013, pp. G86760G-1–G86760G-12.

[15] E. Reinhard *et al.*, "Color transfer between images," *IEEE Comput. Graph. Appl.*, vol. 21, no. 5, pp. 34–41, Sep./Oct. 2001.

[16] D. L. Ruderman *et al.*, "Statistics of cone responses to natural images: Implications for visual coding," *J. Opt. Soc. Amer.*, vol. 15, no. 8, pp. 2036–2045, 1998.

[17] T. Abe *et al.*, "Color correction of pathological images based on dye amount quantification," *Opt. Rev.*, vol. 12, no. 4, pp. 293–300, 2005.

[18] M. Saraswat and K. Arya, "Color normalisation of histopathological images," *Comput. Methods Biomech. Biomed. Eng.: Imaging Vis.*, vol. 1, no. 4, pp. 185–197, 2013.

[19] M. Niethammer *et al.*, "Appearance normalization of histology slides," in *Proc. Mach. Learning Med. Imaging*, 2010, pp. 58–66.

[20] T. Otsuka *et al.*, "Color standardization," in *Proc. Pathology Informatics*, Session 4: Color and Compression, 2010.

[21] A. Ruifrok and D. Johnston, "Quantification of histochemical staining by color deconcolution," *Anal. Quant. Cytol. Histol. Int. Acad. Cytol. Amer. Soc. Cytol.*, vol. 23, no. 4, pp. 291–299, 2001.

[22] A. C. Ruifrok *et al.*, "Comparison of quantification of histochemical staining by hue-saturation-intensity (hsi) transformation and color-deconvolution," *Appl. Immunohistochem. Mol. Morphol.*, vol. 11, no. 1, pp. 85–91, 2003.

[23] G. Begelman *et al.*, "Blind decomposition of transmission light microscopic hyperspectral cube using sparse representation," *IEEE Trans. Med. Imag.*, vol. 28, no. 8, pp. 1317–1324, Aug. 2009.

[24] C. C. Bilgin *et al.*, "Digitally adjusting chromogenic dye proportions in brightfield microscopy images," *J. Microsc.*, vol. 245, no. 3, pp. 319–330, 2011.

[25] M. Gavrilovic *et al.*, "Blind color decomposition of histological images," *IEEE Trans. Med. Imag.*, vol. 32, no. 6, pp. 983–994, Jun. 2013.

[26] N. Ohta and A. R. Robertson, "CIE standard colorimetric system," in *Colorimetry: Fundamentals and Applications*. New York, NY, USA: Wiley, 2006.

[27] K. N. Plataniotis and A. N. Venetsanopoulos, *Color Image Processing and Applications*. New York, NY, USA: Springer, 2000.

[28] W. F. Trench, *Introduction to Real Analysis*. Upper Saddle River, NJ, USA: Prentice-Hall, 2003.

[29] J. Ingle and S. Crouch, *Spectrochemical Analysis*. Upper Saddle River, NJ, USA: Prentice-Hall, 1988.

[30] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *Proc. Adv. Neural Inform. Process. Syst.*, 2001, pp. 556–562.

[31] A. Cichocki *et al.*, *Nonnegative Matrix and Tensor Factorizations*. New York, NY, USA: Wiley, 2009.

[32] A. Hanbury, "Circular statistics applied to color images," in *Proc. 8th Comput. Vis. Winter Workshop*, 2003, pp. 53–71.

[33] L. Shamir *et al.*, "Iicbu 2008 proposed benchmark suite for biological image analysis," *Med. Biol. Eng. Comput.*, vol. 46, no. 9, pp. 943–947, 2008.

[34] N. Orlov *et al.*, "Automatic classification of lymphoma images with transform-based global features," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 4, pp. 1003–1013, Jul. 2010.

[35] E. Gelasca *et al.*, "Evaluation and benchmark for biological image segmentation," in *Proc. Int. Conf. Image Process.*, 2008, pp. 1816–1819.

[36] A. Koschan and M. Abidi, *Digital Color Image Processing*. New York, NY, USA: Wiley, 2008.

Authors' photographs and biographies not available at the time of publication.