

Circular Mixture Modeling of Color Distribution for Blind Stain Separation in Pathology Images

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

Abstract—In digital pathology, to address color variation and histological component colocalization in pathology images, stain decomposition is usually performed preceding spectral normalization and tissue component segmentation. This paper examines the problem of stain decomposition, which is a naturally nonnegative matrix factorization (NMF) problem in algebra, and introduces a systematical and analytical solution consisting of a circular color analysis module and an NMF-based computation module. Unlike the paradigm of existing stain decomposition algorithms where stain proportions are computed from estimated stain spectra using a matrix inverse operation directly, the introduced solution estimates stain spectra and stain depths via probabilistic reasoning individually. Since the proposed method pays extra attentions to achromatic pixels in color analysis and stain co-occurrence in pixel clustering, it achieves consistent and reliable stain decomposition with minimum decomposition residue. Particularly, aware of the periodic and angular nature of hue, we propose the use of a circular von Mises mixture model to analyze the hue distribution, and provide a complete color-based pixel soft-clustering solution to address color mixing introduced by stain overlap. This innovation combined with saturation-weighted computation makes our study effective for weak stains and broad-spectrum stains. Extensive experimentation on multiple public pathology datasets suggests that our approach outperforms state-of-the-art blind stain separation methods in terms of decomposition effectiveness.

Index Terms—Circular mixture model, color analysis, pathology image, stain decomposition, von Mises distribution.

I. INTRODUCTION

PATHOLOGY serves as the gold standard for cancer diagnosis nowadays. To facilitate pathologists' examination, chemical staining, which highlights histological objects of interest, is usually performed in tissue preparation routine. Consequently, color in a pathology image becomes a good indicator of histological substance distribution.

With advances of technology, digital pathology diagnosis is attracting more interests for better medical service. Essentially, a digital pathology diagnosis system is a pattern recognition system whose schematic framework is shown in Fig. 1. By

comparing numerical features of a pathology image against the stored feature sets in the database, a diagnosis system generates a positive/negative result. Aware that pathology images may appear in different colors, noise levels, and storage formats due to the inconsistency in image preparation, preprocessing an image is necessary to mitigate effects of those inconsistent factors on subsequent diagnosis. Particularly, to take advantage of color information in digital pathology, stain separation/decomposition, a signal analysis process that estimates stain spectra and corresponding stain proportions/depths in a pathology image, is frequently performed in the preprocessing block in Fig. 1. The resulting single-stained images facilitate descriptive feature extractions, eventually improving the overall performance of a diagnosis system. Specifically, in a diagnosis pipeline, stain decomposition is usually performed for two major purposes. First, stain separation is an important building block of color normalization for pathology images [1]–[5]. Due to differences of stain manufactures, stain concentrations, storage conditions, or staining duration, chemical dyes in the same type may have different light-absorption spectra, resulting in color variation in images [6]. The so-called spectral variation in stains, or simply stain variation, distorts quantitative features extracted from pathology images, leading to inaccurate image analysis. To limit the impact of color variation on pathology image analysis, stain separation is usually performed to obtain spectra of stains in a pathology image, followed by spectral normalization. Second, stain decomposition is capable of addressing colocalization of tissue substances for histological component detection/segmentation [7]–[10]. Identification of certain histological structures is significant for cancer diagnosis [11]. However, applying grayscale-based segmentation techniques directly to pathology images is prone to error due to histological substance overlap. One solution to address this colocalization issue in tissue component segmentation is to apply classical algorithms to single-stain images (each containing one type of stain) obtained by stain decomposition.

Existing stain separation solutions proposed for pathology images have their own limitations, which are discussed in Section II-C and Table II in Section VI-D. In this paper, we propose an effective and robust blind stain separation approach based on an imaging model of transmission light microscopy introduced in our previous study [5]. Compared to [5] which introduces a general and complete color normalization framework for pathology images, this paper focuses on stain decomposition particularly, and presents a systematic and analytic solution based on the observation that a stain

Manuscript received August 27, 2015; revised November 4, 2015; accepted November 21, 2015. Date of publication November 25, 2015; date of current version January 31, 2017.

The authors are with the Multimedia Lab, The Edward S. Rogers Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada (e-mail: xingyu.li@mail.utoronto.ca; kostas@ece.utoronto.ca).

Digital Object Identifier 10.1109/JBHI.2015.2503720

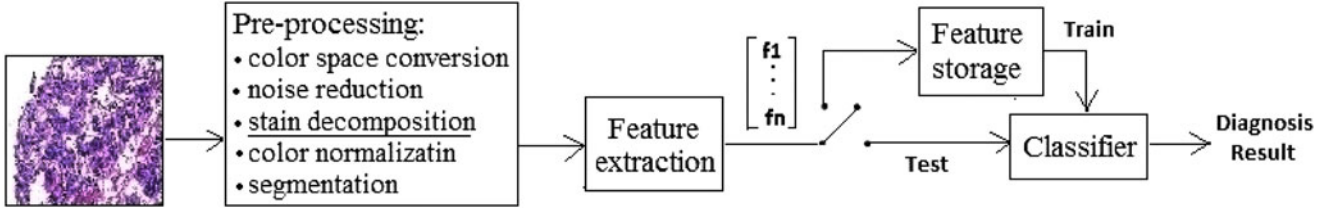


Fig. 1. Schematic framework of a pathology diagnostic system, where the pre-processing block may include operations of color space conversion, stain separation, color normalization, etc. By comparing features of a query image to the stored feature sets, a classifier makes a diagnosis decision.

decomposition problem with its physical constraints is naturally an NMF problem [12] in algebra. Due to the close relation between image color and stains, the introduced solution obtains color-based pixel clustering via a circular mixture model, and passes resulting color cues to the NMF-based computation. Experimentation on public image sets suggests that our method is superior to existing blind stain decomposition approaches.

The major contributions of this study are summarized in both the application and theoretical aspects as follows:

Contribution to Stain Decomposition: 1) This work examines the stain decomposition problem and possible challenges systematically, and provides a holistic and adaptive solution. In particular, the proposed solution mitigates effects of achromatic pixels on subsequent image analysis by saturation-weighted mechanism, addresses stain co-occurrence by a circular mixture model and pixel soft clustering, and adopts NMF computation to adaptively obtain an optimized solution which conforms to the physical constraints on stain separation. Compared to previous works addressing challenges in stain decomposition partially (discussed in Section VI-D), the proposed solution is more robust and effective. 2) Unlike previous stain separation solutions that obtain stain proportions simply by a matrix inverse operation on an estimated stain spectrum matrix, this study introduces a method to estimate stain depths via probability reasoning, which provides insightful information on stain decomposition.

Contribution to Color Processing: The circular color processing module proposed in this study focuses on the analysis of hue signals, which is one of the most informative sources in human color perception. Taking advantage of the close relation between saturation and hue, the distribution of reliable hue signals are summarized. As hue signals are periodic on the unit circle, we innovate to model a hue distribution using a circular mixture distribution, and provide a complete solution to pixel clustering based on it. To the best of our knowledge, this study with our preliminary study¹ [13] constitutes the first attempt to address the periodicity of a hue distribution by a circular mixture model in the color analysis literature.

The rest of this paper is organized as follows. The state of the art in stain separation for pathology images including problem formulation and previous works is reviewed in Section II. In Section III, we introduce a blind stain decomposition method for images generated by light-absorbing stains, and present its block diagram. Section IV and Section V specify the two modules of the proposed method, which are circular analysis of

image color distribution, and NMF-based stain decomposition. Experimental results and discussions are presented in Section VI, followed by conclusions in Section VII.

II. STATE OF THE ART

A. Formulation of Stain Decomposition

Stain separation or stain decomposition, is a process that estimates stains' absorption spectra and their proportions at each pixel in a pathology image. To formulate the stain decomposition problem, we follow the image model of transmission light microscopy in [5], and have

$$I(p, \lambda_i) = I^b(\lambda_i) e^{-\sum_{j=1}^N m_{ij}(\epsilon_1) d_j(p)} \quad (1)$$

where $I(p, \lambda_i)$ and $I^b(\lambda_i)$ represent image values in the i th color channel at pixel p and background where no stains exists, respectively. N is the number of chemical dyes used at staining, $m_{ij}(\lambda)$ represents the spectral mass absorptivity of the j th stain to an incident light in wavelength λ_i , and $d_j(p)$ is so-called stain depth, describing amount of the j th stain at location p . (1) connects a digital quantity, image intensity, to stain absorption spectra, and can be viewed as a digital version of the Beer-Lambert law [14]. By moving $I^b(\lambda_i)$ to the left in (1) and taking logarithm on both sides, the optical density (OD) domain equivalent expression is as follows:

$$\log(I^b(\lambda_i)) - \log(I(p, \lambda_i)) = \sum_{j=1}^N m_{ij}(\epsilon_1) d_j(p). \quad (2)$$

For a biopsy sample stained by N types of stains and imaged via K sensors, we have K equations in the form of (2) for its pathology image. Hence, we form a K -by- N spectrum matrix M using m_{ij} as a matrix element in the i th row, j th column. Then, spectrum of the j th stain, $M_{\cdot j}$, can be characterized by the j th column of M . For instance, an RGB-format H&E stained image has a 3×2 spectrum matrix, whose columns correspond to spectrum vectors of hematoxylin and eosin. Let $I^b = [I^b(\lambda_1), \dots, I^b(\lambda_K)]'$, $I(p) = [I(p, \lambda_1), \dots, I(p, \lambda_K)]'$, and $D(p) = [d_1(p), \dots, d_N(p)]'$, then

$$\log(I^b) - \log(I(p)) = M \times D(p). \quad (3)$$

According to (3), given a pathology image, stain decomposition is an image analysis process to factorize $\log(I^b) - \log(I(p))$ into a spectrum matrix M and stain depths $D(p)$ which are non-negative. Consequently, a set of single-stain images $D = [D_1, \dots, D_N]$ is generated, where D_i is a density map consisting of stain depth $d_i(p)$ of the i th stain over all pixels. It should

¹Compared to [13], this study considers the circumstance where noise exists, and adds a noise term in the mixture model for model robustness.

be noted that $I(p, \lambda_i)$ represents a pixel value in a color channel without gamma correction, i.e., $\gamma = 1$. Correspondingly, an RGB-format image in this paper refers to an image represented in the linear RGB color space.

B. Challenges in Stain Decomposition

There are two major challenges in stain decomposition. First, limited information is available for stain separation. Though stains have their own diagnostic colors, such as hematoxylin usually appearing in blue while eosin in pink visually, both stain variation due to operational inconsistency in staining and color mixing due to stain co-occurrence are usually observed. As the profile of chemical dyes may not be available for each pathology image, quantification of each stain from mixing color in a pathology image is challenging.

Second, achromatic noise is a big challenge for weak stain estimation. In pathology images, a weak stain is defined as a chemical dye that has small intensity or weak color. The weak stained pixels are prone to be confused with unreliable achromatic pixels in background where no stains exist, leading to an inaccurate stain separation.

C. Previous Works on Stain Decomposition

Existing pathology image centered stain separation solutions usually follow the paradigm that estimates a stain spectrum matrix M first and then generates stain density maps D by matrix inverse computation. They can be categorized into supervised and unsupervised groups depending on whether control slides are required.

Supervised approaches achieve stain separation using knowledge on stain spectra obtained from training images. Early stain decomposition works [15]–[17] quantify stain spectra at manually-selected pixels in a controlled slide. Recently, stain decomposition achieved by supervised learning on a set of training pathology images was proposed in a stain normalization method [3]. By soft-classifying color features of training images into categories, the supervised solution computes stain spectra from different groups. As generating well-controlled training samples for pathology images generated in different laboratories is time consuming, or even impossible, our study focuses on the scenario of blind stain separation.

Unsupervised stain decomposition methods usually estimate a spectrum matrix M from a query image directly. Following a color linear unmixing model for fluorescence images where color is superposition of stains in the linear RGB color space, independent component analysis or NMF is applied to an RGB-format pathology image for stain decomposition [18], [19]. However, pathology imaging follows the Beer–Lambert law, rather than the additive staining model of fluorescence imaging. Later, sparseness analysis followed by relative Newton method was proposed to achieve adaptive stain decomposition in [20]. Since this method is designed for hyperspectral images whose channel number is much larger than the number of stains on biopsy samples, it is inapplicable to RGB-format images. Based on an assumption that single-stained pixels exist in a two-stain

image, the plane fitting (PF) method proposes a geometric solution achieved by singular value decomposition (SVD) and thresholding [1]. Due to its promising results, variants of the PF method were proposed in recent publications. In [21], prior knowledge on stains is exploited in a modified PF process for weak stain estimation. In [4], the enhanced-PH (EPH) method tried to address H&E stain interaction based on the PF method. However, PF-based solutions have two major limitations. First, the thresholds for stain spectral estimation is prefixed and cannot be adjusted automatically to an optimal values according to image local content, which results in unstable stain decomposition. Second, the PF method targets for images having two stains only, hence it's nontrivial to generalize it to the scenario of multiple stain separation. For accurate spectrum matrix estimation, spectra of chemical dyes were detected using dedicated spectral sensing hardware [22]; whereas dependence on hardware limits its adoption. Recently, blind color decomposition (BCD) is achieved by performing expectation-maximization (EM) on color distributions in the Maxwell color triangle [23]. Though a heuristic randomization function tries to select stable colors for spectrum matrix estimation, BCD method is prone to be affected by achromatic pixels when estimating weak stains' spectra. In [5] and [24], either K-mean clustering or circular thresholding (CrT) is exploited for pixel clustering for subsequent stain spectrum estimation. As these studies ignore color mixing introduced by stain overlap and hard classify one pixel to one stain cluster, resulting stain separation is less accurate.

In summary, though unsupervised stain separation algorithms are more advanced over supervised solutions in terms of requiring no control slides, they may fail to generate reasonable stain decomposition due to various limitations (See examples and discussion in Section VI-D). Hence, a new robust blind stain separation approach is needed.

III. OVERVIEW OF PROPOSED STAIN DECOMPOSITION

Stain decomposition for a pathology image generated by absorbing stains is an analysis process to factorize $\log(I^b) - \log(I(p))$ into a spectrum matrix M , which is composed of N stain absorption vectors, and stain depths $D(p)$. Geometrically, stain vectors intersect at the origin in the OD domain and form a pyramid, within which all image colors lie. Hence, a stain separation problem can be phrased as searching a pyramid that contains all image colors inside.

Unlike the PF work [1] proposing a geometric stain separation solution², we notice that in algebra stain decomposition formulated in (3) with constraints of non-negative M and $D(p)$ forms an NMF problem naturally, and thus, make use of the generic NMF technique to automatically search the N stain vectors of the targeting pyramid. It should be noted that NMF is unstable in the sense that factorization results may converge to different stationary points due to different initializations. Hence, our study proposes to analyze color distribution in a query image

²Targeting a simplified circumstance that a pathology image is generated by two stains, the PF method [1] tries to find a wedge plane, instead of a pyramid, to accommodate most color samples in the OD domain.

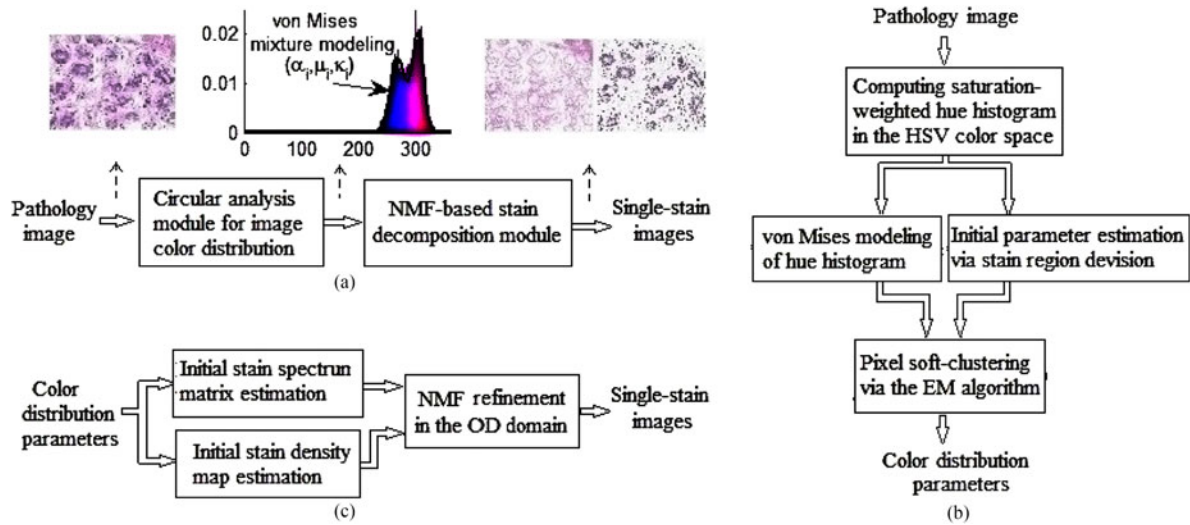


Fig. 2. (a) Block diagram of the proposed stain decomposition solution, which is composed of (b) circular analysis module and (c) stain separation module. In (a), a dash arrow represents information transferred between modules, and histogram bins are filled by corresponding colors for visualization.

first, so that resulting information leads subsequent NMF-based computation to an optimal convergence point. The diagram of the proposed blind stain decomposition solution, which is composed of a circular analysis module and an NMF-based computation module, is presented in Fig. 2(a). In this study, to visualize image color distribution, bins in a histogram are filled by their corresponding colors.

Fig. 2(b) and (c) depicts the block diagrams of the circular analysis module and the NMF-based computation module, respectively. In the first module, since pathology images may contain many achromatic pixels, i.e., white background pixels, reliable image colors are selected implicitly based on pixel saturation to build a hue histogram. With circular modeling of the color distribution, image pixels are clustered into groups via maximum likelihood estimation (MLE), each cluster corresponding to one stain. In the second module, based on circular analysis clues, initial stain vectors and stain depths are computed individually and then optimized by NMF. Generally, the EM algorithm for pixel clustering in the color analysis module and the NMF refinement in the decomposition module are computationally intensive. However, as the NMF computation has good initializations in our method, it converges quickly. Hence, computation overhead of the method is mainly introduced by the EM algorithm. Details of the two modules are given in Section IV and Section V, respectively.

Since side information on stains, such as information about stain concentration, is hardly available, our work 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. It should be noted that to ensure reliable stain decomposition, a query image should satisfy two constraints. 1) A query image should contain light-absorbing stains only. Any light-scattering stains, such as Diaminobenzidine, should be removed from images beforehand [23]. 2) A query image should be generated/scanned under a standard illuminant, such

as CIE D_{65} [25], so that image color is not biased by the microscopic light. Otherwise, color normalization [5] is required to compensate illuminate variation first.

IV. CIRCULAR MODELING OF COLOR DISTRIBUTION FOR PIXEL CLUSTERING

Theoretically, chemical dyes used in pathological staining have different absorption spectra, and color in a pathology image is generated by transmitted light that is not absorbed by stains [26]. As color distribution of a pathology image suggests spectra of incident light that are not absorbed, it can be used to estimate stains absorption spectra and corresponding stain proportions. To this end, this section introduces a module, which is depicted in Fig. 2(b), to analyze color distribution in an image.

A. Saturation-Weighted Hue Histogram

Though pixel values of a color image are usually represented by vectors in a 3-D space, we propose to study color distribution in the hue channel of the HSV color space for reasons of robustness and computational efficiency. First, microscope setup and sample preparation may cause uneven illumination and vignetting in pathology images [27], [28]. As the HSV color space separates chroma (hue and saturation) from luma, color analysis in the hue component is invariant to uneven illumination. Second, though hue is a 1-D angular variable on the chromatic circle, it is a major property of color in human vision. Compared to color analysis in high-dimensional spaces, for example the RGB domain, analysis in hue channel is much easier. Therefore, a pathology image in the linear-RGB color space is first converted into the HSV domain for color analysis in this module.

When summarizing the hue distribution of a pathology image by a histogram, we notice that most standard hue histograms have many spikes contributed by achromatic pixels. This

phenomenon is attributed to singularities in the HSV color space. That is, hue value is ill-defined for achromatic color, and unreliable for a pixel that has a low saturation value. As a result, color analysis on a standard hue histogram corrupted by these unreliable pixels is unreliable, or even meaningless. Since achromatic color is characterized by low saturation in the HSV space, to mitigate their effects on hue distribution, we follow our study in [5] and compute saturation-weighted hue histogram [29] defined as follows:

$$H_{\theta}^{\text{sw}} = \sum_p s(p)\delta(\theta, h(p)) \quad (4)$$

where

$$\delta(\theta, h(p)) = \begin{cases} 1 & \text{if } \theta = h(p) \\ 0 & \text{otherwise.} \end{cases}$$

$\theta = 360l/L$ represents the l th bin in an L -bin histogram, and $s(p)$ and $h(p)$ are saturation and hue at pixel p in the HSV color space. Consequently, less-saturated pixels have small contributions to a hue histogram, and subsequent analysis is less sensitive to achromatic pixels in image local content.

B. Pixel Clustering via Circular Mixture Model

As demonstrated in Fig. 2(a), the resulting saturation-weighted hue histogram is composed of multiple hue clusters, each corresponding to one stain in an image. Hence, we try to divide a histogram into stain regions for pixel clustering. To this end, we have two options.

- Nonparametric thresholding methods, such as multilevel Otsu' method [30] and CrT [31], can be used to partition a histogram into nonoverlapping regions. However, hard-classifying a pixel to one stain cluster is inappropriate for pathology images, because color at a pixel may be contributed by more than one stains due to overlap of tissue substances.
- To address stain co-occurrence in a pathology image, mixture distributions, for instance the well-known Gaussian mixture model (GMM) [32], can be used to model color distribution for pixel clustering. However, due to the circular nature of hue in the HSV color space, applying a mixture model that consists of linear distributions to a hue histogram is inappropriate. For instance, hue values $h(p) = 360$ are equivalent to $h(p) = 0$ on the chromatic plane, and their average is still 0 (or 360). However, when linear computation is applied, the wrong value 180, instead of 360 (or 0), is generated.

To address stain co-occurrence and angular hue signals, this study proposes the use of a circular mixture distribution to model a hue histogram for soft clustering of image pixels.

1) Circular Modeling of Hue Distribution: We assume that hue is a random variable drawn from a circular distribution whose probability density function (pdf) is given by the resulting saturation-weighted histogram. Then for a pathology image containing N types of stains, the color mixture distribution has a pdf defined as $f(x; \alpha_i, \phi_i) = \sum_{i=1}^{N+1} \alpha_i f_i(x; \phi_i)$ subject to (s.t.) $\sum_{i=1}^{N+1} \alpha_i = 1$. α_i represents a probability that

a hue value x is drawn from the i th population $f_i(x; \phi_i)$ with a parameter set ϕ_i . In the model, the first N populations correspond to the N stain clusters, and we use one more distribution to model color distribution of achromatic pixels.

With very-limited prior knowledge on color distribution (i.e., stains' typical color), a circular population, called von Mises (vM) distribution [33], is used to model each cluster of hue for two reasons as follows:

- The maximum entropy property of the vM distribution [33] implies that we have minimum prior information on the data distribution.
- The shape of a vM distribution varies with different selection of parameters. The flexibility in shape enables the vM model to approximate many circular distribution [33]. For instance, a vM distribution converges to the uniform distribution on the unit circle when its concentration parameter is 0; while when the concentration parameter goes to infinity, it tends to an impulse.

A vM distribution has the density function defined as

$$vM(\mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(x-\mu)} \quad (5)$$

where x is a circular random variable with a unit length, $0 \leq \mu < 2\pi$ and $\kappa \geq 0$ represent its mean direction and concentration, and $I_0(\kappa)$ is the zero-order modified Bessel function. Hence, our von Mises mixture model (vMMM) with a parameter set $\Psi = \{\alpha_i, \mu_i, \kappa_i\}$ for a hue histogram is

$$f(x; \Psi) = \sum_{i=1}^{N+1} \alpha_i vM(\mu_i, \kappa_i), \text{ s.t. } \sum_{i=1}^{N+1} \alpha_i = 1. \quad (6)$$

Noted that the histogram H_{θ}^{sw} is computed from weighted hue samples and differs from the standard hue distribution of an image. Hence, a set of synthetic hue data $X = \{x_1, \dots, x_M\}$ is generated following H_{θ}^{sw} for parameter estimation. Assuming that $X = \{x_j\}$ are independent and identically distributed following (6), the joint distribution of X is

$$g(X; \Psi) = \prod_{j=1}^M f(x_j; \Psi) = \prod_{j=1}^M \left[\sum_{i=1}^{N+1} \alpha_i f_i(x_j; \phi_i) \right] \quad (7)$$

where $f_i(x; \phi_i) = vM(\mu_i, \kappa_i)$.

2) Parameter Estimation via EM Algorithm: Based on the vMMM model and hue data X , we proceed to compute the unknown parameters Ψ in (7) via MLE. By maximizing the log-likelihood function $\log g(X; \Psi)$ over Ψ , we will get

$$\Psi^{\text{mle}} = \arg \max_{\Psi} \sum_{j=1}^M \log \left[\sum_{i=1}^{N+1} \alpha_i f_i(x_j; \phi_i) \right], \text{ s.t. } \sum_{i=1}^{N+1} \alpha_i = 1. \quad (8)$$

Due to the complexity and the singularities of this MLE problem, the EM algorithm [34] is exploited to obtain Ψ^{mle} in (8). By introducing a set of latent variables $Z = \{z_{ji}\}$ to represent the information that a sample x_j is drawn from the i th population, where z_{ji} can be either 0 or 1 with a constraint $\sum_{i=1}^{N+1} \delta(z_{ji}, 1) = 1$, the EM algorithm defines $g_c(X, Z; \Psi)$ as

the joint distribution of observation X and latent data Z given a parameter set Ψ ,

$$g_c(X, Z; \Psi) = \prod_{j=1}^M \prod_{i=1}^{N+1} [\alpha_i f_i(x_j; \phi_i)]^{z_{ji}}. \quad (9)$$

Following the conventional expression in the EM algorithm, we denote the expectation of $\log g_c(X, Z; \Psi)$ based on observation samples X and current estimation of $\Psi^{(k)}$ as

$$Q[\Psi; \Psi^{(k)}] = E_{\Psi^{(k)}} [\log g_c(X, Z; \Psi) | X]. \quad (10)$$

Then, Ψ is updated via maximizing $Q[\Psi; \Psi^{(k)}]$ iteratively. Specifically, in the $(k+1)$ th iteration, the EM algorithm first computes $Q[\Psi; \Psi^{(k)}]$ using the current $\Psi^{(k)}$ in the E-step,

$$\begin{aligned} Q[\Psi; \Psi^{(k)}] &= \sum_{j=1}^M \sum_{i=1}^{N+1} \bar{z}_{ji}^{(k+1)} \log [\alpha_i^{(k)} f_i(x_j; \phi_i^{(k)})], \\ &= \sum_{j=1}^M \sum_{i=1}^{N+1} \bar{z}_{ji}^{(k+1)} \left[\log \alpha_i^{(k)} - \log 2\pi I_0(\kappa_i^{(k)}) \right. \\ &\quad \left. + \kappa_i^{(k)} \cos(x_j - \mu_i^{(k)}) \right], \end{aligned} \quad (11)$$

where

$$\begin{aligned} \bar{z}_{ji}^{(k+1)} &= E[z_{ji} | x_j, \Psi^{(k)}] = \sum_{z_{ji}=0}^1 z_{ji} P(z_{ji} | x_j, \Psi^{(k)}) \\ &= \alpha_i^{(k)} f_i(x_j; \phi_i^{(k)}) / \sum_{i=1}^{N+1} \alpha_i^{(k)} f_i(x_j; \phi_i^{(k)}). \end{aligned} \quad (12)$$

Then, the parameter set $\Psi = \{\alpha_i, \mu_i, \kappa_i\}$ is chosen so that $\Psi^{(k+1)} = \arg \max_{\Psi} Q[\Psi; \Psi^{(k)}]$ in the M-step as follows:

- 1) Under the constraint that $\sum_{i=1}^{N+1} \alpha_i = 1$, we apply the Lagrange method to (11) with respect to α_i and get

$$\alpha_i^{(k+1)} = \sum_{j=1}^M \bar{z}_{ji}^{(k+1)} / (N+1). \quad (13)$$

- 2) Let partial derivative of (11) over μ_i equal 0, leading to $\sum_{j=1}^M \bar{z}_{ji}^{(k+1)} \sin(x_j - \mu_i) = 0$. Hence,

$$\mu_i^{(k+1)} = \arctan \left(\frac{\sum_{j=1}^M \bar{z}_{ji}^{(k+1)} \sin(x_j)}{\sum_{j=1}^M \bar{z}_{ji}^{(k+1)} \cos(x_j)} \right). \quad (14)$$

- 3) Differentiating (11) with respect to κ_i , we obtain $\sum_{j=1}^M \bar{z}_{ji}^{(k+1)} [\cos(x_j - \mu_i^{(k+1)}) - I_0'(\kappa_i) / I_0(\kappa_i)] = 0$. Since the Bessel function has the property of $I_0'(\kappa_i) = I_1(\kappa_i)$, we let $A(\kappa_i) = I_1(\kappa_i) / I_0(\kappa_i)$ and obtain

$$\kappa_i^{(k+1)} = A^{-1} \left(\frac{\sum_{j=1}^M \bar{z}_{ji}^{(k+1)} \cos(x_j - \mu_i^{(k+1)})}{\sum_{j=1}^M \bar{z}_{ji}^{(k+1)}} \right) \quad (15)$$

where $A^{-1}(\cdot)$ can be solved by numerical methods [33].

According to the convergence theory for an EM sequence [35], as $Q[\Psi; \Psi^{(k)}]$ in (10) is continuous in both Ψ and $\Psi^{(k)}$,

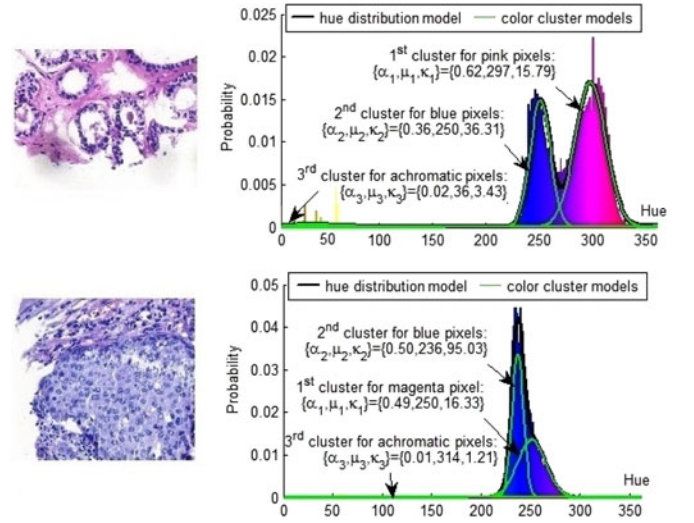


Fig. 3. Examples of histogram modeling by the vM mixture distribution. Images are shown on the left, and hue histograms and corresponding modeling parameters are on the right.

\log -likelihood value $\log g(X; \Psi)$ converges monotonically to $\log g(X; \Psi^*)$ for some stationary point Ψ^* . That is, starting from a good initial parameter set $\Psi^{(0)}$ (which is discussed in Section IV-D), the EM algorithm will finally obtain Ψ^{mle} via updating (11) in the E-step and (13), (14), and (15) in the M-step iteratively. In this study, the EM algorithm is regarded as convergence when $|g(X; \Psi^{(k+1)}) - g(X; \Psi^{(k)})| < 10^{-6}$.

Examples of the proposed circular mixture modeling of saturation-weighted hue histograms for pathology images are presented in Fig. 3. It is noteworthy that the $N+1$ th cluster has very small α_{N+1} and κ_{N+1} . On one hand, the small α_{N+1} implies that the saturation-weighted hue histogram contains very small amount of achromatic pixels. On the other hand, the small κ_{N+1} suggests that the color of achromatic pixels tends to distribute uniformly over the entire color band.

C. Soft Clustering of Pixels

Since stains may overlap on a biopsy sample, an image pixel with hue $h(p)$ is assigned a probability vector $P^{h(p)} = \{P_1^{h(p)}, \dots, P_{N+1}^{h(p)}\}$ s.t. $\sum_{i=1}^{N+1} P_i^{h(p)} = 1$, where

$$P_i^{h(p)} = \alpha_i^{\text{mle}} f_i(h(p); \phi_i^{\text{mle}}) / \sum_{i=1}^{N+1} \alpha_i^{\text{mle}} f_i(h(p); \phi_i^{\text{mle}}) \quad (16)$$

representing the probability that the pixel belongs to the i th stain based on the circular mixture model with Ψ^{mle} .

D. EM Initialization Through Stain Region Separation

As the log-likelihood $\log g(X; \Psi)$ is not unimodal in Ψ for the vMMM model, it may converge to $\log g(X; \Psi^*)$ for any stationary point Ψ^* depending on the starting point $\Psi^{(0)}$. Thus, extra care needs to be taken in choosing $\Psi^{(0)}$.

We notice that most achromatic pixels in a pathology image belong to image background characterized by large intensity

values $v(p)$ in the HSV model. Hence, with a predetermined threshold T , an image is divided into stain region $R_s = \{p : v(p) \leq T\}$ and achromatic pixel region $R_{ap} = \{p : v(p) > T\}$. In this paper, we set $T = 240$ for a 24-bit RGB-format image. Then, the EM initial parameters are computed from saturation-weighted hue histograms $H_{\theta,s}^{sw}$ and $H_{\theta,ap}^m$, which summarize color distributions in R_s and R_{ap} , respectively.

We first proceed to compute $\{\mu_i^{(0)}, \kappa_i^{(0)}, \alpha_i^{(0)}\}$ from $H_{\theta,s}^{sw}$ for the N stain clusters. As we know the typical color of each stain in an image (for instance, eosin is pink in general), this prior information is used to initiate $\mu_i^{(0)}$, $1 \leq i \leq N$. To estimate $\kappa_i^{(0)}$ and $\alpha_i^{(0)}$, the histogram $H_{\theta,s}^{sw}$ is divided into N nonoverlap regions R_s^i using cosine distance:

$$R_s^i = \left\{ H_{\theta,s}^{sw} : \cos(\theta - \mu_i^{(0)}) < \cos(\theta - \mu_j^{(0)}), \forall j \neq i \right\} \quad (17)$$

where θ is a bin in a hue histogram. It is noteworthy that rather than a linear distance metric such as Euclidean distance, cosine distance is more efficient to address the circular nature of hue. For a vM distribution, $\kappa = A^{-1}(\bar{R})$ in the MLE sense, where \bar{R} is the mean resultant length of circular samples. Hence, assuming that a stain region R_s^i in $H_{\theta,s}^{sw}$ contains all stain samples drawn from the i th vM distribution,

$$\begin{aligned} \kappa_i^{(0)} = A^{-1} & \left(\frac{1}{\sum_{\theta \in R_s^i} H_{\theta,s}^{sw}} \right. \\ & \times \sqrt{\left(\sum_{\theta \in R_s^i} H_{\theta,s}^{sw} \cos \theta \right)^2 + \left(\sum_{\theta \in R_s^i} H_{\theta,s}^{sw} \sin \theta \right)^2} \Big) \end{aligned} \quad (18)$$

$$\alpha_i^{(0)} = \sum_{\theta \in R_s^i} H_{\theta,s}^{sw} / \sum_{\theta} (H_{\theta,s}^{sw} + H_{\theta,ap}^{sw}). \quad (19)$$

Similarly, $\{\mu_{N+1}^{(0)}, \kappa_{N+1}^{(0)}, \alpha_{N+1}^{(0)}\}$ of the achromatic pixel cluster are estimated from the histogram $H_{\theta,ap}^{sw}$ as follows:

$$\mu_{N+1}^{(0)} = \arctan \left(\frac{\sum_{\theta} H_{\theta,ap}^{sw} \cos \theta}{\sum_{\theta} H_{\theta,ap}^{sw} \sin \theta} \right) \quad (20)$$

$$\begin{aligned} \kappa_{N+1}^{(0)} = A^{-1} & \left(\frac{1}{\sum_{\theta} H_{\theta,ap}^{sw}} \right. \\ & \times \sqrt{\left(\sum_{\theta} H_{\theta,ap}^{sw} \cos \theta \right)^2 + \left(\sum_{\theta} H_{\theta,ap}^{sw} \sin \theta \right)^2} \Big) \end{aligned} \quad (21)$$

$$\alpha_{N+1}^{(0)} = \sum_{\theta} H_{\theta,ap}^{sw} / \sum_{\theta} (H_{\theta,s}^{sw} + H_{\theta,ap}^{sw}). \quad (22)$$

Note, since stain region division in a histogram does not address stain overlap, resulting parameters $\{\mu_i^{(0)}, \kappa_i^{(0)}, \alpha_i^{(0)}\}$ may not be very accurate. However, as long as these initial parameters are reasonable and close to the true values, the EM algorithm will converge to the MLE solution.

V. NMF-BASED STAIN DECOMPOSITION

In algebra, stain decomposition combined with its physical constraints that spectrum matrix and stain depths are non-negative forms an NMF problem naturally. Hence, as demonstrated in Fig. 2(c), based on the vMMM model of color distribution, initial stain vectors and stain depths are generated independently. Then, NMF computation is performed with generated initializations to obtain the optimal stain decomposition.

A. Estimation of Spectrum Matrix

The estimation of stain spectra starts from the computation of representative color for each stain cluster. Since the vM distribution has the maximum likelihood characterization that its sample mean direction is the MLE of the distribution mean direction, we take the obtained μ_i^{mle} , rather than computing a sample mean, as the representative hue h_i for the i th stain. Then, saturation-weighted statistics of color components $v(p)$ and $s(p)$ [5] are exploited to obtain stain representative color:

$$s_i = \left[\sum_p (s(p))^2 \delta(h_i, h(p)) \right] / \left[\sum_p s(p) \delta(h_i, h(p)) \right], \quad (23)$$

$$v_i = \left[\sum_p s(p) v(p) \delta(h_i, h(p)) \right] / \left[\sum_p s(p) \delta(h_i, h(p)) \right]. \quad (24)$$

In (23) and (24), as saturation is used as filter parameters, the color estimation is insensitivity to achromatic pixels.

A stain's absorption spectrum characterizes the deduction amount of light in the OD domain when light travels through the stain [16]. Hence, for a stain whose representative color is $I_i^{rgb} = f_{hsv}^{rgb}([h_i, s_i, v_i]^T)$ where $f_{hsv}^{rgb}(\cdot)$ is the function to convert a color vector in the HSV color space to the RGB model, its corresponding absorption vector in the OD domain is $M_{:,i}^{in} = \log(I^b / I_i^{rgb})$. For an RGB-format image having N types of stains, the $3 \times N$ initial spectrum matrix M^{in1} is formed by using $M_{:,i}^{in}$ as the i th column. Hence, given a query image, one candidate for stain decomposition is

$$\begin{cases} M^{in1} = [M_{:,1}^{in}, \dots, M_{:,N}^{in}] \\ D^{in1} = (M^{in1})^{-1} [\log(I^b) - \log(I)]. \end{cases} \quad (25)$$

It should be noted that D^{in1} may be negative in (25) and violates the physical constraints in stain separation. Hence, refinement on (M^{in1}, D^{in1}) (Section V-C) is required.

B. Estimation of Stain Density Maps

Initial stain density maps are computed from the point of view of probabilistic reasoning. As stain co-occurrence may exist in an image, this study achieves pixel soft-clustering by assigning a probability vector $P^{h(p)} = \{P_1^{h(p)}, \dots, P_{N+1}^{h(p)}\}$ to a pixel with hue $h(p)$, which provides insightful information on relative proportions of stains at pixel p . In this study, we generate the i th density map $D_i^{in2} = \{P_i^{h(p)}, \forall p\}$. Hence, another

candidate for stain separation is

$$\begin{cases} M^{\text{in}2} = [\log(I^b) - \log(I)](D^{\text{in}2})^{-1} \\ D^{\text{in}2} = \{D_1^{\text{in}}, \dots, D_N^{\text{in}}\}. \end{cases} \quad (26)$$

Similarly, as the resulting $M^{\text{in}2}$ may contain negative absorption coefficients which is impossible in practice, parameter refinement is performed later.

C. Optimal Stain Separation Generation

In algebra, image colors are non-negative combination of N stain vectors in the OD domain. Due to the NMF-property of stain decomposition, we exploit the NMF technique to refine the two candidate solutions with respect to minimizing the mean-squared decomposition error over all pixels,

$$\begin{aligned} \{M, D\} &= \arg \min_{M, D \geq 0} \text{MSE}_{M \times D} \\ &= \arg \min_{M, D \geq 0} E_p [\|\log(I^b/I(p)) - M \times D(p)\|^2]. \end{aligned} \quad (27)$$

Consequently, two local optimal results (M^1, D^1) and (M^2, D^2) are obtained in the areas around $(M^{\text{in}1}, D^{\text{in}1})$ and $(M^{\text{in}2}, D^{\text{in}2})$, respectively. Note, as $(M^{\text{in}1}, D^{\text{in}1})$ and $(M^{\text{in}2}, D^{\text{in}2})$ usually fall in the same convergence area of (27), (M^1, D^1) and (M^2, D^2) tend to converge to the same stationary point. In this study, the NMF iteration is terminated when the difference of two successive $\text{MSE}_{M \times D}$ or M is less than 10^{-6} . Then, the stain decomposition result with a smaller $\text{MSE}_{M \times D}$ is taken as the final solution.

VI. EXPERIMENTAL RESULTS AND DISCUSSIONS

In this study, five experiments are performed to evaluate the circular analysis module and our stain separation approach. All experiments run on Matlab, and datasets and experimental design are specified in each experiment as follows.

A. vMMM for Stain Color Estimation

Through estimating stain colors in a pathology image, this experiment conducts a comparative evaluation on the proposed vMMM-based pixel clustering.

Test Data: Experimental images are randomly picked from the thyroid image atlas [36] published by Papanicolaou Society of Cytopathology. As the lesion images are stained by different chemical dyes, they have distinct color distributions.

Experimental Design: The vMMM-based pixel clustering is performed on the saturation-weighted histogram of a test image. For comparison, linear thresholding methods (Otsu's method and GMM model) and state-of-art CrT method [31] are also applied to the same histogram, replacing the circular pixel clustering. To better visualize the performance, for each pathology image, stains' representative colors are estimated following (14), (23), and (24) afterward.

Results and Discussion: Examples of stain color estimation are presented in Fig. 4, where the two colored bins represent stain colors estimated from an image. Because the vMMM model

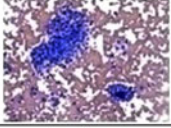
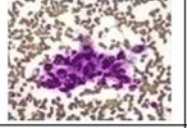








	Query image		
Linear processing	Otsu's [30]		
	GMM [32]		
Circular processing	CrT [31]		
	vMMM		

Fig. 4. Examples of stain color estimation using different thresholding methods for pathology images, where the colored bins represent the two estimated stain colors in an image.

addresses both the circular nature of hue and stain co-occurrence simultaneously, it generates most accurate stain colors. By contrast, Otsu's method and the GMM model fail as they ignore the periodicity of hue in computation. Performance of the CrT method lays between the linear methods and the vMMM model for two reasons. 1) As an extension of Otsu's method for directional data, CrT is achieved by minimizing the intra-class variance only, rather than by considering all statistics properties behind observed samples. 2) Particularly for pathology image, soft clustering of image pixels is preferable due to stain mixing, whereas CrT can only achieve hard thresholding, which makes color estimation less accurate.

B. System Robustness to Parameter Settings

In the histogram-based analysis, low resolution of a histogram may lead to inaccuracy, while too many bins may result in a sparse histogram and unstable analysis. In this experiment, we change the number of bins, L , in a hue histogram and examine whether the final stain separation is sensitive to this parameter.

Test Data: The UCSB breast cancer cell dataset [37], consisting of 58 H&E stained breast cancer pathology images stored in 24-bit nonlinear RGB format, is selected for two reasons. 1) It is important to examine the proposed method in H&E stained images as H&E staining is the dominant staining protocol in pathology. 2) Evaluation on 58 images is more solid as it does not depend on a particular image. Note, a UCSB image was converted to the linear RGB format first to follow the imaging model in (1).

Experimental Design: The method is said to be robust to L if the resulting stain spectrum matrix are consistent given various values of L . Hence, we first compute absolute matrix difference for the i th UCSB image, $\Delta M^i = |M_L^i - M_{L'}^i|$, where M_L^i is a spectrum matrix associated with a histogram having L bins. Then, the expected value of ΔM^i over the UCSB set is summarized. The expectation of spectral variation in H&E stains among UCSB biopsies, $\Delta M^{\text{sv}} = E[|M_L^i - M_{L'}^j|]$ for $i \neq j$, is used as comparison baseline. Since the matrix difference (rather than the matrix itself) matters, we set $L = 360$ and $L' = 90l$ for $l = 1, \dots, 8$ in this experiment.

Results and Discussion: The average matrix difference ΔM over the UCSB set are depicted in Fig. 5, where dash lines represent the comparative baseline ΔM^{sv} . ΔM keeps 1/10 smaller

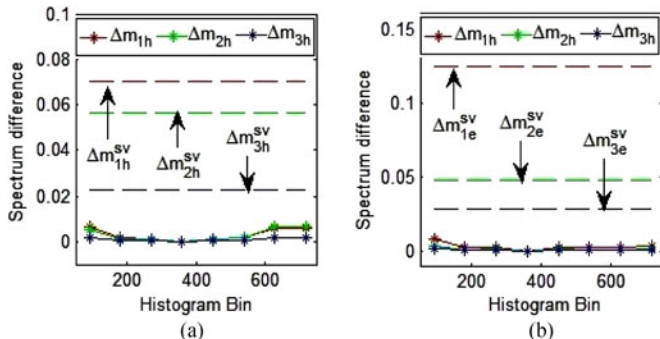


Fig. 5. Matrix differences $\Delta M = [\Delta m_{1h}, \Delta m_{1e}; \Delta m_{2h}, \Delta m_{2e}; \Delta m_{3h}, \Delta m_{3e}]$ against the parameter, histogram bin L , over the UCSB dataset. $\Delta M^{sv} = [\Delta m_{1h}^{sv}, \Delta m_{1e}^{sv}; \Delta m_{2h}^{sv}, \Delta m_{2e}^{sv}; \Delta m_{3h}^{sv}, \Delta m_{3e}^{sv}]$ denotes the comparison baseline: expectation of H&E spectral variation in the UCSB set.

than ΔM^{sv} within a large range of L , which suggests that M_L^i and M_L^j deliver consistent information on stains' spectra. Hence, we conclude that our system is robust against the parameter L . In this paper, a histogram with 360 bins is used in all other experiment.

C. Stain Separation on CMU Benchmark Dataset

Recently, a public stain separation benchmark is provided by the bimagicLab in Carnegie Mellon University [4]. We will evaluate the proposed method on this benchmark.

Test Data: The CMU dataset contains three H&E stained images in 48-bit linear RGB format, on which strong H&E interaction is observed. Decomposition groundtruth, the H-stained image and E-stained image, are provided in this set.

Experimental Design: We apply stain separation methods to each CMU image and compare the decomposition result to the groundtruth. As the dataset does not provide numerical information on either stain vectors nor registration parameters between H&E images and corresponding groundtruth, the Kullback–Leibler divergence $D_{KL}(P||Q)$, which measures information lost when Q is used to approximate P in information theory, is used as a quantitative metric. Specifically, in this experiment, Q is the color histogram of a decomposition image, and P is the histogram computed from the corresponding decomposition groundtruth. Hence, $D_{KL}(P||Q)$ quantifies the similarity between the decomposition results and corresponding groundtruth from the perspective of information theory, and a smaller $D_{KL}(P||Q)$ suggests a better decomposition.

Results and Discussion: Fig. 6(c) presents our stain separation on the CMU images. As other separation methods generate similar results visually, the obtained stain spectra, instead of their decomposition images, are depicted in Fig. 6(d) to save space. Table I reports $D_{KL}(P||Q)$ for each method, where the H(E) represents the H(E)-stained image. For one thing, average $D_{KL}(P||Q)$ of the proposed method are close to the EPF method [4] which has the smallest average Kullback–Leibler divergence on the benchmark. For another, we notice that the EPF method may generate negative stain spectra [for instance, the eosin vector for the second CMU image demonstrated in

TABLE I

$D_{KL}(P||Q)$ OF STAIN DECOMPOSITION METHODS ON THE UMC SET [4]

		PF [1]	EPF [4]	BCD [23]	HTN [5]	proposed
Image1	E	1.3007	0.9946	1.0361	1.5472	0.8485
	H	2.5625	2.3687	2.9449	2.8020	2.7777
Image2	E	1.0959	0.8297	0.9916	1.3364	0.8885
	H	2.6274	2.3506	2.9217	2.5936	2.4855
Image3	E	1.4278	1.1563	1.1678	1.7036	1.3593
	H	2.5523	2.3683	2.9506	2.5033	2.4559
average	E	1.2748	0.9935	1.0652	1.5291	1.0324
	H	2.5807	2.3625	2.9391	2.6330	2.5730

Fig. 6(d)], which violates the non-negative constraint on stain separation. Taking both the metric values $D_{KL}(P||Q)$ and the constraint into consideration, we conclude that in general the proposed method generates more reliable decomposition on the UMC image set.

D. Qualitative Comparison of Stain Separation Solutions

Test Data: This experiment performs a qualitative comparison among blind stain separation solutions on public pathology images stained by different chemical dyes. Hence, both the UCSB images [37] and the thyroid image [36] are used.

Experimental Design: Stain separation solution examined in this experiment includes the original PF method [1], EPF method [4], BCD method [23], HTN method [5], and the proposed method. First, they are assessed on the UCSB H&E images. Then, these methods are examined on randomly picked thyroid images which are stained by other chemical dyes. For a fair comparison, we tried our best to faithfully reproduce experimental results in corresponding original papers.

Results and Discussion: We observe that the four previous blind stain decomposition methods fail in some pathology images, but that the proposed method always achieves satisfactory results. Fig. 7 gives four typical stain separation examples, where the first two images are from the UCSB dataset, and the rest are from the thyroid image set.

The PF method [1] estimates stain spectra by thresholding along the geodesic line in the SVD domain. As thresholds are predetermined, inaccurate stain vectors may be generated. The separation failure due to the image-independent thresholds is more frequently observed in weak stain estimation which is sensitive to achromatic pixels. We notice that for most examined images, the PF method can obtain similar results to the proposed method if the thresholds are manually selected.

The EPF method [4] tried to improve the original PF method for images with H&E stain interaction. Hence, it inherits PF's limitations. Besides, as modifications in EPF is particularly for H&E stains, they have side effects on other stains.

The BCD method [23] estimates stain vectors by applying the GMM model to selected pixels with high optimal density in the Maxwell color triangle. However, as less pixels associated with weak stains are selected in the estimation process, spectra of weak stains are biased by strong stains, and thus, less reliable. This explains why the decomposed E-stained images have hematoxylin-similar color for UCSB images.

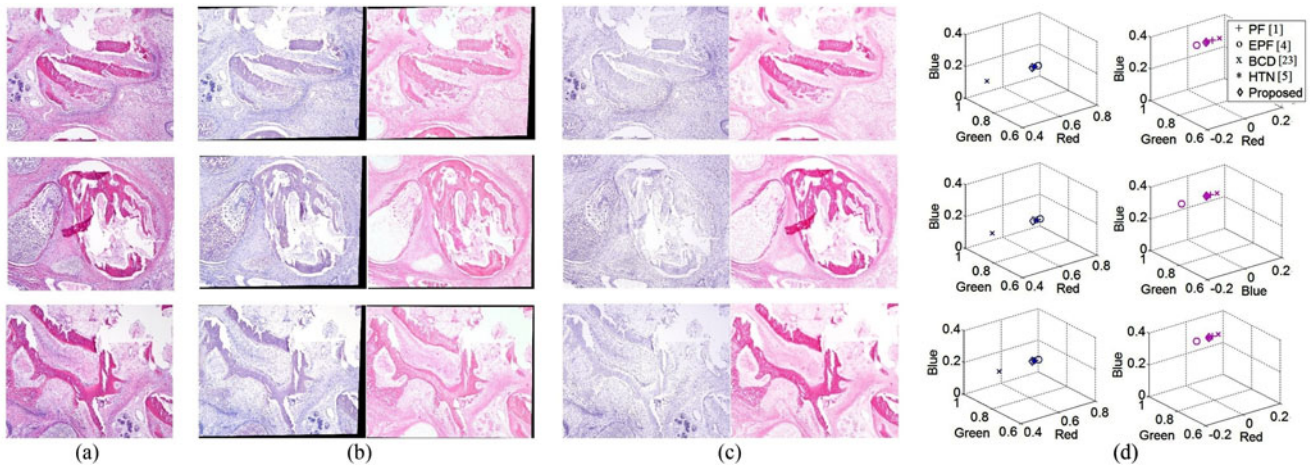


Fig. 6. Stain separation on the CMU benchmark. (a) Benchmark images; (b) Groundtruth; (c) Stain decomposition achieved by the proposed method; (d) Stain vectors obtained by various stain separation solutions: PF method [1], EPF method [4], BCD method [23], HTN method [5] and the proposed solution.

	Query images	PF [1]	EPF [4]	BCD [23]	HTN [5]	Proposed
UCSB images						
Thyroid images						

Fig. 7. Example of stain decomposition achieved by the PF method [1], EPF method [4], BCD method [23], HTN method [5] and our solution.

TABLE II
BLIND STAIN DECOMPOSITION METHOD COMPARISON

Addressing:	(E)PF[1], [4]	BCD[23]	HTN[5]	proposed
achromatic pixels in weak stain estimation	×	×	✓	✓
stain overlap in pixel clustering	N/A	✓	×	✓
physical constraints in optimization	×	×	✓	✓

TABLE III
RUN TIMES OF BLIND STAIN DECOMPOSITION OVER THE UCSB SET

Methods	PF [1]	BCD [23]	HTN [5]	proposed
Run time (sec)	0.3352	9.6344	0.6849	5.3041 ³

The HTN method [5] uses the K-mean algorithm to divide a hue histogram into nonoverlap stain regions for stain estimation. Thus, it is incapable of addressing stain colocalization and hue periodicity, and may fail in images either containing³ large areas of stain overlap (first UCSB image) or generated by stains with broad color bands (second thyroid image).

As briefly summarized in Table II, since the proposed method addresses issues related to achromatic pixels, stain overlap, and the problem's physical constraints systematically and holistically, it is able to generate a good stain separation result with minimized decomposition error.

E. Computational Complexity

This experiment quantitatively compares the computation efficiency of the proposed method to the PF method [1], the BCD method [23], and the HTN method [5].

Test Data: The UCSB dataset [37] is used, because contents in the 58 images pose various difficulties for stain separation.

Experimental Design: For quantitative comparison, we run all examined algorithms on the test data, and summarize their average run times in second (s). To clarify, convergence conditions (of the EM method and NMF computation) in the experiment are all set to 10^{-4} , and all timings are calculated on a 3.4 GHz Windows 7 computer running Matlab R2014a.

Results and Discussion: Table III reports the average run times for various stain decomposition over the UCSB set. The PF method [1] is fast as its two building blocks, thresholding and SVD, are computationally efficient. As both the BCD method [23] and the proposed method adopt the EM algorithm for pixel clustering, they are much slower. However, compared to the 2-D EM iteration in the BCD method, the 1-D parameter estimation and the NMF computation in the proposed method converge much faster because of good initializations. The computation complexity of the HTN method [5] mainly comes from the

NMF computation. Hence, its efficiency resides between the proposed method and the PF method.

VII. CONCLUSION

This study introduced an effective blind stain decomposition method for pathology images. Since challenges of achromatic pixels in color analysis and stain estimation, stain co-occurrence due to histological substance overlap, and non-negative constraints on stain decomposition were carefully addressed, the proposed method achieved optimized result with minimal decomposition error. Extensive experimentation on publicly available pathology images indicates that the proposed solution performs more consistently compared to state-of-the-art blind stain separation methods.

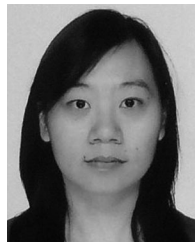
The introduced circular color processing module made use of a vMMM distribution for saturation-weighted hue histogram modeling. Complete solution to color-based pixel cluster was provided. Since hue is a periodic angular signal on the unit chromatic circle, the innovation of circular modeling of hue distribution led to more accurate analysis, which is demonstrated in our experiment.

REFERENCES

- [1] M. Macenko, M. Niethammer, J. Marron, D. Borland, J. Woosley, X. Guan, C. Schmitt, and N. Thomas, "A method for normalizing histology slides for quantitative analysis," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2009, pp. 1107–1110.
- [2] M. Saraswat and K. Arya, "Color normalisation of histopathological images," *Comput. Methods Biomech. Biomed. Eng. Imag. Vis.*, vol. 1, no. 4, pp. 185–197, 2013.
- [3] A. Khan, N. Rajpoot, D. Treanor, and D. Magee, "A non-linear mapping approach to stain normalization in digital histopathology images using image-specific color deconvolution," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 6, pp. 1729–1738, Jun. 2014.
- [4] M. McCann, J. Majumdar, C. Peng, C. Castro, and J. Kovacevic, "Algorithm and benchmark dataset for stain separation in histology image," in *Proc. IEEE Int. Conf. Image Process.*, 2014, pp. 3953–3957.
- [5] X. Li and K. Plataniotis, "A complete color normalization approach to histo-pathology images using color cues computed from saturation-weighted statistics," *IEEE Trans. Biomed. Eng.*, vol. 7, no. 62, pp. 1862–1873, Jul. 2015.
- [6] M. Veta, J. Pluim, P. van Diest, and M. Viergever, "Breast cancer histopathology image analysis: A review," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 5, pp. 1400–1411, May 2014.
- [7] S. Cataldo, E. Ficarra, A. Acquaviva, and E. Macii, "Automated segmentation of tissue images for computerized IHC analysis," *Methods Programs Biomed.*, vol. 100, pp. 1–15, 2010.
- [8] K. Lesack and C. Naugler, "Performance of a simple chromatin-rich segmentation algorithm in quantifying basal cell carcinoma from histology images," *BMC Res. Note*, vol. 5, pp. 1–9, 2012.
- [9] E. Ozdemir and C. Gunduz-Demir, "A hybrid classification model for digital pathology using structural and statistical pattern recognition," *IEEE Trans. Med. Imag.*, vol. 32, no. 2, pp. 474–483, Feb. 2013.
- [10] G. Olgun, C. Sokmensuer, and C. Gunduz-Demir, "Local object patterns for the representation and classification of colon tissue images," *IEEE J. Biomed. Health Informat.*, vol. 18, no. 4, pp. 1390–1396, Jul. 2014.
- [11] M. Gurcan, L. Boucheron, A. Can, A. Madabhushi, N. Rajpoot, and B. Yener, "Histopathological image analysis: A review," *IEEE Rev. Biomed. Eng.*, vol. 2, pp. 147–171, 2009.
- [12] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *Proc. Adv. Neural Inf. Process. Syst.*, 2001, pp. 556–562.
- [13] X. Li and K. Plataniotis, "Diagnostic color estimation of tissue components in pathology images," in *Proc. IEEE Int. Conf. Image Process.*, 2015.
- [14] J. Ingle and S. Crouch, *Spectrochemical Analysis*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1988.

³The proposed method is currently implemented for validating our idea and has not been optimized. Its computation efficiency can be much improved using vector-based computation and parallel computing.

- [15] R. Zhou, E. Hammond, and D. Parker, "A multiple wavelength algorithm in color image analysis and its application in stain decomposition in microscopy images," *Med. Phys.*, vol. 23, no. 12, pp. 1977–1986, 1996.
- [16] A. Ruifrok and D. Johnston, "Quantification of histochemical staining by color deconvolution," *Anal. Quant. Cytol. Histol.*, vol. 23, no. 4, pp. 291–299, 2001.
- [17] A. Ruifrok, R. Katz, and D. Johnston, "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.
- [18] A. Rabinovich, S. Agarwal, C. Laris, J. Price, and S. Belongie, "Unsupervised color decomposition of histologically stained tissue samples," in *Proc. Adv. Neural Inf. Process. Syst.*, 2003, pp. 667–674.
- [19] J. Newberg and R. Murphy, "A framework for the automated analysis of subcellular patterns in human protein atlas images," *J. Proteome Res.*, vol. 7, no. 6, pp. 2300–2308, 2008.
- [20] G. Begelman, M. Zibulevsky, E. Rivlin, and T. Kolatt, "Blind decomposition of transmission light microscopic hyperspectral cube using sparse representation," *IEEE Trans. Med. Imag.*, vol. 28, no. 8, pp. 1317–1324, Aug. 2009.
- [21] M. Niethammer, D. Borland, J. Marron, J. Woosley, and N. Thomas, "Appearance normalization of histology slides," in *Proc. Mach. Learn. Med. Imag.*, 2010, pp. 58–66.
- [22] S. Tani, Y. Fukunaga, S. Shimizu, M. Fukunishi, K. Ishii, and K. Tamiya, "Color standardization method and system for whole slide imaging based on spectral sensing," *Anal. Cell. Pathol.*, vol. 35, no. 2, pp. 107–115, 2012.
- [23] M. Gavrilovic, J. Azar, J. Lindblad, C. Wahlby, E. Bengtsson, C. Busch, and I. Carlom, "Blind color decomposition of histological images," *IEEE Trans. Med. Imag.*, vol. 32, no. 6, pp. 983–994, Jun. 2013.
- [24] X. Li and K. Plataniotis, "Blind stain decomposition for histo-pathology images using circular nature of chroma components," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2015, pp. 877–881.
- [25] N. Ohta and A. R. Robertson, *CIE Standard Colorimetric System. Colorimetry: Fundamentals and Applications*. New York, NY, USA: Wiley, 2006.
- [26] D. Magee, D. Treanor, D. Crellin, M. Shires, K. Smith, K. Mohee, and P. Quirke, "Color normalization in digital histopathology images," in *Proc. MICCAI Workshop*, 2009, pp. 100–111.
- [27] F. Piccinini, E. Lucarelli, A. Gherardi, and A. Bevilacqua, "Multi-image based method to correct vignetting effect in light microscopy images," *J. Microscopy*, vol. 248, no. 1, pp. 6–22, Jan. 2012.
- [28] H. Irshad, A. Veillard, L. Roux, and D. Racoceanu, "Methods for nuclei detection, segmentation, and classification in digital histopathology: A review - current status and future potential," *IEEE Rev. Biomed. Eng.*, vol. 7, pp. 97–114, Jan. 2014.
- [29] A. Hanbury, "Circular statistics applied to color images," in *Proc. 8th Comput. Vis. Winter Workshop*, vol. 91, no. 1-2, 2003, pp. 53–71.
- [30] M. Luessi, M. Eichmann, G. Schuster, and A. Katsaggelos, "Framework for efficient optimal multilevel image thresholding," *J. Electron. Imag.*, vol. 18, no. 1, 2009.
- [31] Y.-K. Lai and P. Rosin, "Efficient circular thresholding," *IEEE Trans. Image Process.*, vol. 23, no. 3, pp. 992–1001, Mar. 2014.
- [32] T. Kurita, N. Otsu, and N. Abdelmalek, "Maximum likelihood thresholding based on population mixture models," *Pattern Recognit.*, vol. 25, no. 10, pp. 1231–1240, 1992.
- [33] N. Fisher, *Statistical Analysis of Circular Data*. Cambridge, U.K.: Cambridge Univ. Press, 1996.
- [34] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *J. Roy. Statist. Soc. Ser. B*, vol. 39, pp. 1–38, 1977.
- [35] G. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*. New York, NY, USA: Wiley, 2008.
- [36] P. S. of Cytopathology, Thyroid image atlas, pp. 013004-1–013004-10, 2004.
- [37] E. Gelasca, J. Byun, B. Obara, and B. Manjunath, "Evaluation and benchmark for biological image segmentation," in *Proc. 15th IEEE Int. Conf. Image Process.*, 2008, pp. 1816–1819.



Xingyu Li (S'09) received the B.Sc. degree in electronics engineering from the Peking University, Beijing, China, in 2007, and the M.Sc. degree in electrical and computer engineering from the University of Alberta, Edmonton, AB, Canada, in 2009. She joined the Multimedia lab in the University of Toronto, Toronto, ON, Canada, in 2012, where she is currently working toward the Ph.D. degree with the Department of Electrical and Computer Engineering.

Her research interests include color image processing, circular data analysis, and pattern recognition with applications to medical images.



Konstantinos N. Plataniotis (S'93–M'95–SM'03–F'12) is a Professor and the Bell Canada Chair in Multimedia with the ECE Department, University of Toronto, Toronto, ON, Canada. He is the Founder and Inaugural Director-Research for the Identity, Privacy and Security Institute, the University of Toronto and he has served as the Director for the Knowledge Media Design Institute, University of Toronto from January 2010 to July 2012. His research interests include knowledge and digital media design, multimedia systems, biometrics, image & signal processing, communications systems and pattern recognition. Among his publications in these fields are the recent books WLAN positioning systems in 2012 and multi-linear subspace learning: Reduction of multidimensional data in 2013.

Dr. Plataniotis is a Registered Professional Engineer in Ontario and Fellow of the Engineering Institute of Canada. He has served as the Editor-in-Chief of the IEEE SIGNAL PROCESSING LETTERS, and as Technical Co-Chair of the IEEE 2013 International Conference in Acoustics, Speech and Signal Processing. He is the IEEE Signal Processing Society Vice President for Membership (2014–2016) and the General Chair for the 2018 International Conference on Image Processing (ICIP-18).