Blind Decomposition of Transmission Light Microscopic Hyperspectral Cube Using Sparse Representation

Grigory Begelman*, Michael Zibulevsky, Ehud Rivlin, and Tsafrir Kolatt

***Abstract*—In this paper, we address the problem of fully automated decomposition of hyperspectral images for transmission light microscopy. The hyperspectral images are decomposed into spectrally homogeneous compounds. The resulting compounds are described by their spectral characteristics and optical density. We present the multiplicative physical model of image formation in transmission light microscopy, justify reduction of a hyperspectral image decomposition problem to a blind source separation problem, and provide method for hyperspectral restoration of separated compounds. In our approach, dimensionality reduction using principal component analysis (PCA) is followed by a blind source separation (BSS) algorithm. The BSS method is based on sparsifying transformation of observed images and relative Newton optimization procedure. The presented method was verified on hyperspectral images of biological tissues. The method was compared to the existing approach based on nonnegative matrix factorization. Experiments showed that the presented method is faster and better separates the biological compounds from imaging artifacts. The results obtained in this work may be used for improving automatic microscope hardware calibration and computer-aided diagnostics.**

***Index Terms*—Blind source separation, hyperspectral imaging, microscopy, sparse analysis.**

I. Introduction

MODERN biological research and pathology diagnostics tend to rely on computerized microscopes and automatic specimen analysis. Many routine tasks can already be efficiently performed by automated systems: microscope slide scanning, cell segmentation and classification, etc. Most of the automated systems are based on gray-scale or color imaging.

The domain of automated biological analysis of data obtained through microscopes suffers from several intrinsic problems that cannot be easily solved by using grey-scale or color image modalities. The first problem is large color variance of the same dyes attached to the same specimen types. The second problem concerns the physical overlapping of different kinds of objects resulting in color mixtures.

The first problem arises from varying parameters of slide preparation and staining artifacts. The multiparameter (temperature, staining duration, dye solution concentration, etc.) slide preparation process inhibits prediction of the colored slide appearance. Colors of the same tissue elements, stained by the same method, may change from one slide batch to another due to varying parameters of the slide preparation process. Even within the same slide, there can be regions of artifacts where colors are more pronounced or have faded away. Inability to know *a priori* the exact colors of a stained specimen can result in erroneous specimen element identification. Human beings cope with the color variation problem, as they know how to "calibrate" themselves according to many hints available on the specimen. However, this calibration method does not apply for machine guided analysis due to its complexity and vaguely defined nature.

The second intrinsic problem of the biological image analysis arises from a three-dimensional nature of specimens on microscope slides. The composition of biological samples often consists of overlapping biological elements. Cells that pile one on top of another in cytology, three-dimensional tissue section, entangled chromosomes within nuclei are only few examples of such overlapping elements. Since the elements of interest are usually stained by distinct colors, the overlap of microscopic objects results in overlapping colors namely in color mixtures. The color mixtures make different sample entities indistinct. Therefore, automatic segmentation of the overlapping objects becomes complicated. In addition, the overlapping objects cannot be distinguished automatically based on a spectral library.

In this paper, we address problems of automatic tissue separation and identification. Our solution is based on the hyperspectral imaging modality. Decomposition of hyperspectral images obtained from stained biological specimens may contribute substantially to various fields. One such field is computational pathology [1]. In computational pathology, scanning and imaging of pathological slides is followed by computer aided diagnosis.

The spectrally homogenous compounds provided by the method described in our paper can be interpreted directly in most cases and connected to actual biological entities. Robust decomposition may contribute to reliable element segmentation and classification. The method proposed in this paper can become a crucial tool in automated analysis of samples in which multiple markers are used simultaneously to classify biological objects (like in [2]). The biological objects' features are

Manuscript received December 31, 2008; revised February 03, 2009. First published February 27, 2009; current version published July 29, 2009. *Asterisk indicates corresponding author.*

*G. Begelman is with the Department of Computer Science, Technion, Israel Institute of Technology, 32000 Haifa, Israel (e-mail: gbeg@cs.technion.ac.il).

M. Zibulevsky and E. Rivlin are with the Department of Computer Science, Technion, Israel Institute of Technology, Haifa, Israel, 32000 (e-mail: mzib@cs.technion.ac.il; ehudr@cs.technion.ac.il)

T. Kolatt is with Applied Spectral Imaging, Ltd., 10551 Migdal Ha'Emek, Israel (e-mail: tsafrir@spectral-imaging.com).

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/TMI.2009.2015145

directly connected to the amount of stain attached to them, and overlapping dyes are common problem in such measurements. Therefore, separation of spectrally homogenous compounds can greatly improve object segmentation results and help reach the correct diagnosis.

The rest of this paper is organized as follows. In Section II we present an overview of the related work. Section III addresses the physical model of a microscope slide and hyperspectral image formation. In addition, it relates the problem of spectrally homogenous compounds decomposition to the problem of blind source separation. In Section IV we present our solution to the problem of spectrally homogenous compounds separation and provide method for obtaining hyperspectral images of each single separated compound. In Section V we provide experimental validation of the proposed method. Finally, Section VI concludes our results.

II. Related Work

The existing methods for color unmixing and compounds separation can be divided into two categories: methods that are based on a set of known reference spectra (library) of dyes and methods that perform blind unmixing. The first group of methods requires knowledge of the spectral profiles of the dyes which are attached to the tissues. A comprehensive overview of spectral unmixing, in which known reference spectra of dyes are used, is given by [3]. Many microscope manufacturers (Zeiss, Leica, BioRad, Olympus) provide software packages that implement spectral unmixing based on reference spectral profiles. As stated earlier, the outcome color of the dye after it is chemically bound to a tissue element may be very different from the original pure dye. Preferably, measurements for the compilation of reference spectra library should be performed on selected slide regions that contain only one type of tissue and one chromophore. Such regions are difficult to locate and in practice the measurements are performed in the presence of multiple tissue types and chromophores. Due to the interrelations between the tissue elements and the various participating dyes, such compilations may result in poor unmixing and hence poor element distinction and object identification.

The second group of methods does not require the *a priori* compilation of reference spectra. Instead, in these methods, spectra of the input data are automatically reconstructed. Shirakawa *et al.* presented in [4] an experimental system for identification of the spectral signature of fluorescent probes embedded in living cells. The system is based on blind spectral decomposition of fluorescent data by method of parallel factor analysis. For separation of the fluorescent probes on one slide, their system uses nine hyperspectral images (each image is captured under different illumination conditions). Rabinovich *et al.* in [5] introduced a method of unsupervised color decomposition by blind source separation as applied to histology slides of stained tissues. The method presented in this work is based on additive image formation model. The additive model is more applicable to the fluorescent rather than to the transmission light microscopy. Hereby we propose a different method that is based on a physically correct multiplicative model. We demonstrate that the method described in our paper is more sensitive to imaging and instrument artifacts and more robust to variations in compound number than the method proposed in [5].

III. Physical Model

A. Structure of a Microscope Slide

In order to examine tissues microscopically, high quality slide preparations are required. The slide preparation process includes tissue fixation and staining the tissue components in contrasting colors. Staining is a biomedical technique of adding a component-specific dye to tissue. The selection of particular dyes depends on the tissue components that should be analyzed. For example, gram staining is used to classify bacteria according to their gram status. Haematoxylin and Eosin staining is used frequently in histology to examine thin sections of tissue: haematoxylin stains cell nuclei, while eosin stains cytoplasm, connective tissue and blood cells. Biological molecules can be tagged with fluorophores by a simple chemical reaction, and the fluorescence of the tag enables detection of the molecule. For example, ethidium bromide stains DNA of the cells in their certain life cycle stage (apoptosis). Each dye has its unique spectral curve, however, the spectral curves of dyes attached to tissue components vary because of variations of the staining process.

Suppose, there is only one type of tissue component in the field of view of the microscope. Suppose further, that this tissue component was stained by one dye, and there are no staining artifacts that may cause variability of the dye absorption spectrum in the microscope field of view. In this case, the absorption spectrum of the dye attached to the tissue component will be spatially constant. We define such tissue component as a spectrally homogeneous compound.

The specimen mounted on a slide usually contains several overlapping spectrally homogeneous compounds of different objects. Object overlapping occurs both in histological and cytological slides. For example, the tissue for histological analysis is usually cut to 3–10 μm thick. Such compounds' thickness allows several compounds of objects such as cell nuclei, cytoplasm, stroma, and various artifacts to be present on the slide. The cells prepared for cytological analysis frequently overlap each other as well.

Overlapping of the analyzed objects may cause difficulties for slide analysis. First, selection of regions for reference spectra measurements may be difficult or even impossible. Second, segmentation of overlapping objects is considerably more challenging than segmentation of standalone objects.

B. Beer–Lambert Law and Its Application to the Problem

Let us consider a small region S on a slide that will be mapped to one pixel on a hyperspectral image. Before emerging from the region S, light passes several compounds on the microscope slice. In absence of fluorescence, Beer–Lambert law relates the absorption of light to the properties of the material through which the light travels. We use the following notation.

$I_0(\lambda)$	The intensity of the incident light as function of wavelength.
$\alpha_i(\lambda)$	The absorption spectrum for the compound i $(i = 1 \dots M)$.

l_i The ith compound thickness.

c_i The concentration of the absorbing species in the ith compound.

According to the Beer–Lambert law, after traveling through the first compound, the light intensity is

$$I_1(\lambda) = I_0(\lambda)e^{-\alpha_1(\lambda)l_1c_1}.$$

After passing M compounds, the light intensity is described by

$$I_M(\lambda) = I_0(\lambda)e^{-\sum_{i=1}^{M}\alpha_i(\lambda)l_ic_i}.$$

We take into account spatial distribution of the species on the slide ($l_i = l_i(j,k)$, $c_i = c_i(j,k)$, where j and k are the row and column pixel coordinates correspondingly) and denote the spatial distribution of compounds' optical density by $\beta_i(j,k) = l_i(j,k)c_i(j,k)$

$$I_M(j,k,\lambda) = I_0(j,k,\lambda)e^{-\sum_{i=1}^{M}\alpha_i(\lambda)\beta_i(j,k)}. \quad (1)$$

C. Problem Statement in Terms of BSS

Suppose, the slide contains M species, and each specie has its own spatial distribution of optical density $\beta_i(j,k)$, $i = 1,\ldots,M$, j and k are the row and column pixel coordinates correspondingly. We consider a hyperspectral image of the slide, captured by a hyperspectral imaging system at N different wavelengths λ_i, $i = 1,\ldots,N$. We suppose that the spectrum of the microscope light $h_0(\lambda)$ is spatially constant, and the spatial distribution of the microscope light intensity (illumination pattern) is $\hat{I}_0(j,k)$. Then the incident light intensity can be presented as $I_0(j,k,\lambda) = \hat{I}_0(j,k)h_0(\lambda)$. For the sensor signal at wavelength λ_i, according to the (1) we have

$$\begin{aligned} H_i(j,k) &= H(j,k,\lambda_i) \\ &= \hat{I}_0(j,k)h_0(\lambda_i)e^{-\alpha_1(\lambda_i)\beta_1(j,k)}\ldots \\ &\quad \times e^{-\alpha_M(\lambda_i)\beta_M(j,k)} \\ &= \hat{I}_0(j,k)h_0(\lambda_i)\prod_{m=1}^{M}e^{-\alpha_m(\lambda_i)\beta_m(j,k)}. \end{aligned} \quad (2)$$

We transform the equation above from multiplicative into additive

$$\log H_i(j,k) = \log\left(\hat{I}_0(j,k)h_0(\lambda_i)\right) - \sum_{m=1}^{M}\alpha_m(\lambda_i)\beta_m(j,k). \quad (3)$$

We denote: $\alpha_{M+1}(\lambda) \equiv -\log h_0(\lambda)$, $\beta_{M+1}(x,y) \equiv \mathbf{1}(j,k)$, $\alpha_{M+2}(\lambda) \equiv \mathbf{1}(\lambda)$, $\beta_{M+2}(j,k) \equiv -\log \hat{I}_0(j,k)$, and $\tilde{M} = M + 2$. Finally, we unite the first and the second terms of the previous equation

$$\log H_i(j,k) = -\sum_{m=1}^{\tilde{M}}\alpha_m(\lambda_i)\beta_m(j,k).$$

Note, that two additional compounds correspond to nonuniform illumination compound with constant spectrum $\mathbf{1}(\lambda)\log\hat{I}_0(j,k)$ and uniform illumination with spectrum of incident light $\mathbf{1}(j,k)\log h_0(\lambda)$.

To present the above equation in form of matrix multiplication we denote the following.

- Rows of matrix X are composed of the N images $X_i = \log H_i(x,y)$, $i = 1\ldots N$, captured at N different wavelength. The images are settled in rows in lexicographical order.
- Matrix A is composed of $a_{im} = \alpha_m(\lambda_i)$, $m = 1\ldots\tilde{M}$.
- Rows of the matrix S are composed of $s_m = \beta_m(x,y)$, settled in rows in lexicographical order.

Taking into account additive noise ξ, we obtain

$$X = AS + \xi. \quad (4)$$

The hyperspectral image X consists of measurements for a number of wavelengths N. N is usually much greater than the number of significant spectrally homogeneous compounds in the slide M. We assume that for our problem $N \geq \tilde{M}$.

Given the sensor measurements X, we would like to find the matrices S and A. This is a classical statement of overdetermined problem of blind source separation [6]. We describe our solution to this problem in the following sections.

IV. Proposed Solution

Several algorithms have been developed for blind source separation (for example, JADE [7], FastICA [8], BS Infomax [9], [10]). According to the results presented in [11], sparse component analysis outperforms these algorithms on real-world signals and images. In this paper, we rely on sparse component analysis. The overall scheme of the solution to the problem of spectrally homogeneous compounds separation is as follows.

1) Select some transformation Ψ so that the source signals $C = S\Psi$ are sparse. Apply this sparsifying transformation to the sensor measurements: $Y = X\Psi$.
2) Apply a subspace projection to the sparsified sensor measurements, thus transform the overdetermined blind source separation problem to a "square" blind source separation problem $Y_S = \bar{A}C$. The matrix Y_S contains mixtures of the significant compounds, the matrix $\bar{A}$ is square.
3) Use Relative Newton algorithm [12] to find the mixing matrix $\bar{A}$ and the sparsified source signals matrix C.
4) Restore the original mixing matrix A and the original source signals up to scale ambiguity. Model the appearance of each source signal alone on the sensors and thus resolve the scale ambiguity.

A. Sparse Decomposition of the Signal

Solution quality of BSS may be greatly improved by using sparse representation of the signals (see [11] for discussion about influence of sparsity on blind source separation). Sparsity means that there exists some decomposition of the source signal

$$s_i(t) = \sum_k C_{ik}\phi_k(t)$$

so that only a small number of the coefficients C_k differ significantly from zero (for example, see histogram of Contourlet-SD [13] decomposition coefficients of a sample hyperspectral microscopy image on Fig. 1). The scalar functions $\phi_k(t)$ are called atoms or elements of the dictionary. The atoms do not have to

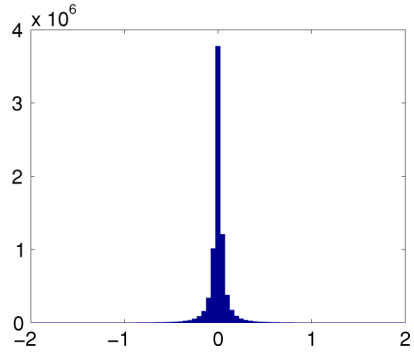


Fig. 1. Histogram of the contourlets-SD coefficients for the first hyper-spectral image.

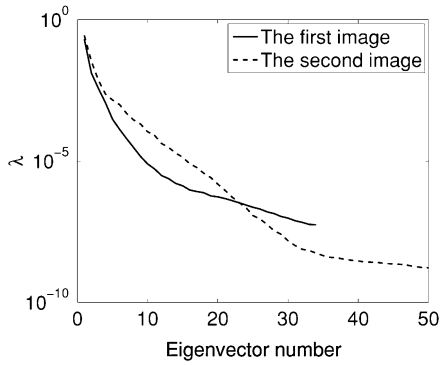


Fig. 2. Eigenvalues of matrix Y (for more details, see the Section IV-B) for the first and the second hyperspectral images.

be linearly independent, instead they may form an overcomplete dictionary. Many natural signals can be sparsely represented in a proper signal dictionary. Important examples of the dictionaries that allow sparse decomposition of natural signals are wavelet-related dictionaries (see [14], for example), or learned dictionaries ([15]–[17]).

Suppose, there exists some transformation Ψ so that $s_i\Psi$ is typically sparse, and the coefficients C_{ik} of the source decomposition $C = S\Psi$ are independent random variables with a probability density function of an exponential type

$$p_i(C_{ik}) \propto e^{-\lambda_i h(C_{ik})} \tag{5}$$

where h may be chosen, for example, as $h(c) = |c|^{1/\gamma}$, $\gamma \geq 1$, or a smooth approximation of this function.

Applying the transformation Ψ to both sides of (4), we get $X\Psi = AS\Psi$. We denote $Y = X\Psi$ and $C = S\Psi$, and obtain $Y = AC$.

The matrix Y is composed of N rows, the matrix A is composed of $\tilde{M}$ columns and N rows, and the matrix C is composed of $\tilde{M}$ rows.

B. Transformation of Over-Determined Problem to a "Square" Problem

The next step is to transform the original overdetermined problem to a square problem. In absence of noise the columns of the matrix Y are embedded into a linear subspace $\mathcal{L}$ of dimension $M + 2$ at most, where M is the number of stains used to color the slide, and it is usually known. Two additional compounds correspond to nonuniform illumination compound with constant spectrum and uniform illumination with real spectrum. However, according to our experiments, there can be more sources in the mixtures. These sources do not correspond to biological objects, but to imaging artifacts. To ensure that the number of sources is not larger than subspace dimension, we require that the subspace dimension d is higher than $M + 2$.

In microscopy images with reasonably small noise may be obtained if long exposure is used. For such images, the principal component analysis (PCA) provides good estimation of the subspace $\mathcal{L}$. Our goal is to separate compounds corresponding to meaningful biological objects. Due to physical considerations, the meaningful objects have large variances. Therefore, if we take enough principal components, projection to subspace provided by PCA will retain the compounds corresponding to the meaningful biological objects. In our experiments, the energy of the discarded components was very small (see Fig. 2 for the scree plot of eigenvalues of the matrix Y).

The basis vectors of the subspace $\mathcal{L}$ are the first d eigenvectors of the covariance matrix $C_Y = E\{YY^T\}$. If v_i are the normalized eigenvectors of the covariance matrix C_Y, the projection matrix B from the original space of Y to the subspace $\mathcal{L}$ equals to: $B = ((v_1)/(\lambda_1)|\ldots|(v_d)/(\lambda_d))^T$ (where λ_i are eigenvalues of C_Y). Likewise, the back-projection matrix from the subspace $\mathcal{L}$ to the original space of Y equals to: $B_{\text{back}} = (\lambda_1 v_1|\ldots|\lambda_d v_d)$.

Finally, we convert the overdetermined problem $Y = AC$ to the "square" problem by multiplying the sparsified sensor measurements by the matrix B: $Y_S = BY = BAC = \bar{A}C$.

C. Classical BSS

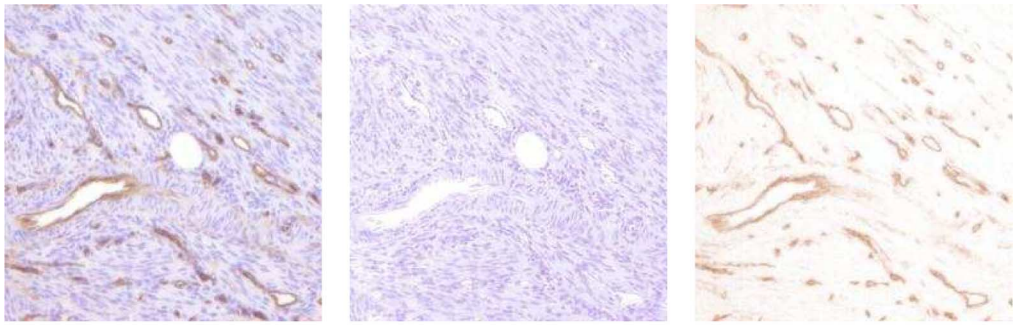
In this section, we present a maximum a-posteriori solution to the problem $Y = \bar{A}C$. Suppose that the mixing matrix is uniformly distributed, and the noise n is negligible ($\sigma \approx 0$). We wish to maximize $P(\bar{A}, C \,|\, Y)$—the conditional probability of observing Y given $\bar{A}$ and C. According to the Bayes Law: $P(\bar{A}, C \,|\, Y) \propto P(Y \,|\, \bar{A}, C)P(\bar{A})P(C)$, and we maximize

$$\max_{\bar{A},C} P(Y \,|\, \bar{A}, C)P(\bar{A})P(C). \tag{6}$$

As prior p.d.f. $P(\bar{A})$ is uniform, it can be dropped from the equation above. As noise is negligible, Y can be expressed as $Y = \bar{A}C$ and the term $P(Y \,|\, \bar{A}, C)$ can be dropped as well. Due to independence of the coefficients C_{ik}, the prior pdf of C is

$$P(C) = \Pi_{ij} \exp(-\lambda_i h(C_{jk})). \tag{7}$$

Further, we denote $W = A^{-1}$ and express $C = A^{-1}Y = WY$. We obtain the following objective function by substituting (7) into (6), taking the logarithm of both sides and inverting the sign: $L(W) = \sum_{j,k} h(WY)_{jk}$. Minimizing this objective can lead to degenerate solution ($W = 0$). To avoid this, we need to enforce nonsingularity of W. This can be done in several ways, e.g., the minimal singular value of W can be restricted from below: $r_{\min}(W) \geq 1$. Another possibility is to subtract $K \log\det(W)$ from the objective function. The term $\log\det(W)$ can be obtained rigorously by maximum likelihood or mutual information considerations [9], [18].



(a) (b) (c)

Fig. 3. Reconstructed color images of the first input hyperspectral image and two separated compounds. Original microscope illumination was used for color reconstruction. (a) RGB reconstruction of a hyperspectral image captured from a slide containing mouse heart tissue. (b) RGB reconstruction of the cell nuclei and cytoplasm compound. (c) RGB reconstruction of the blood vessels' border compound.

Combining the equations above, taking the logarithm, and inverting the sign, we obtain the following objective function: $L_1(W) = -K \log|\det W| + \sum_{j,k} \gamma_i h((WY)_{jk})$. We use Relative Newton algorithm [12] for minimization of this objective function since widely accepted natural gradient method [19], [10] does not work well for this type of problem (see [12] for comparison of two methods).

D. Reprojection of the Source Images to the Original Data Space

In the final step we reconstruct the original mixing matrix A and the original sources. To reconstruct the original source signals matrix S, the following transformation should be applied.

1) Project the original observation signals to the subspace $\mathcal{L}$: $\bar{X} = BX = BAS = \bar{A}S$,.
2) Restore the source signal matrix from the projected observations matrix $\bar{X}$: $S = (\bar{A})^{-1}\bar{X} = (\bar{A})^{-1}BX$.

The original mixing matrix A is obtained from the matrix $\bar{A}$ by back-projecting to the original space: $A = B_{\text{back}}\bar{A}$. Finally, we have obtained the original mixing matrix A and the separated source signals S.

BSS problem is intrinsically scale indefinite: multiplying a source by a constant and dividing the corresponding column of the mixing matrix by the same constant does not change the observed mixed data. To allow natural analysis of the separation results, we bypass the problem of scale ambiguity by representing the separation results in the "mixture" space. Representing the separation results in the mixture space models the situation when only one source, corresponding to a separated compound, presents in the mixture. This representation does not suffer from scale indefiniteness, as the unknown constant is canceled by multiplication of the source by the mixing matrix.

We note that the separated sources have form of attenuation coefficients for each pixel of the image: $S_k = \beta_k(x, y)$, and the mixing matrix $a_k = \alpha_k(\lambda_i)$ indicates degree of attenuation of the separated source for each wavelength. We denote by a_i columns of the mixing matrix A and by s_i rows of the source signals matrix S. The hyperspectral sensor measurements corresponding to the ith source is modeled according to the following equation: $H_i = I_0(x, y)h_0(\lambda_i)e^{-a_i s_i}$ (note that the indefinite constant is canceled in multiplication $a_i s_i$). The parameters I_0 and h_0 of the incident illumination light may be chosen arbitrarily. For example, parameters of the original illumination may be obtained by capturing a hyperspectral image of an empty slide. In our experiments, spectrum of the incident illumination light was obtained by averaging spectrum of empty regions of the analyzed slide (regions that do not contain tissues) and illumination pattern was uniform. Finally, the spectral properties of the incident illumination may be taken according to the spectral properties of some microscope bulb, e.g., halogen or xenon, or some standard white light.

From (3), the source number $M + 1$ is a spatial constant ($\beta_{M+1}(x, y) = 1$). This source intrinsically can not be separated from the other sources unless we know the mean values of the other sources in advance. Therefore, we know all the separated sources up to some unknown spatial constant.

It is possible to estimate this constant using physical considerations. In case when there is a hole in specimen [like on Fig. 3(a)], the attenuation of light contributed by each compound is zero in the corresponding pixels. Therefore, we can force the minimal value of each source to zero by assigning: $s_i = s_i - \min(s_i)$. If there is no hole in specimen, this strategy can still be justified: we assume that minimal attenuation of a compound is zero at some place, and assign the remaining constant in space attenuation to the incident light. Finally, if the actual spatial-spectral distribution of the incident light is known, we can attribute this overall constant attenuation to an additional compound.

V. Experimental Results

In this section we demonstrate the separation results provided by our algorithm and the algorithm described in [5] and discuss the outcome.

We implemented our algorithm in Matlab. The sources' sparsity was achieved by two methods: using the Contourlet-SD transform [13] and using the stack of horizontal and vertical derivatives of the images. We obtained similar decomposition results for both sparsifying methods, and in this paper we demonstrate results obtained using the Contourlets dictionary (see Fig. 1 for the histogram of Contourlets-SD coefficients for one of the sample images).

The method described in [5], relies on additive physical model of image formation and uses nonsparse blind source separation technique based on nonnegative matrix factorization.

TABLE I
AVERAGE OPTICAL DENSITY OF THE SEPARATED COMPOUNDS FOR THE FIRST AND THE SECOND HYPERSPECTRAL IMAGES

		Compound 1	Compound 2	Compound 3	Compound 4	Compound 5	Compound 6
The first image	Our algorithm	0.1224	0.0727	0.0070	0.0197	0.0052	0.0002
	Algorithm from [5]	0.0048	0.0112	0.0126	0.0103	0.0117	0.0092
The second image	Our algorithm	0.1142	0.1316	0.0851	0.0151	0.0059	0.0021
	Algorithm from [5]	0.0223	0.0078	0.0049	0.0177	0.0154	0.0125

We implemented this algorithm in Matlab using nonnegative matrix factorization algorithm described in [20] (the Matlab code for nonnegative matrix factorization was provided by the author of [20]). The methods based on nonnegative matrix factorization require subtraction of the incident illumination pattern from the hyperspectral image prior to separation. As the hyperspectral images used in our experiments were captured from the central part of microscope field of view, we assumed that influence of nonuniform illumination may be neglected and subtracted the illumination estimated from the specimen hole areas.

The runtime of our algorithm on the first hyperspectral image was 14 s and on the second hyperspectral image was 21 s (Intel Core 2 CPU, 2.4 GHz, 4 GB of RAM). Most of time was spent by the optimization using Relative Newton algorithm. The runtime of the algorithm from [5] was 30 min on the first hyperspectral image and about one hour on the second hyperspectral image. Time required to read the input hyperspectral images from disk was below 0.3 s per image, while time required to write the hyperspectral images of separated compound was below 2 s per image.

We experimented with hyperspectral images of biological slides. The images were captured by a Fourier-based spectroscope and provided by ASI Ltd. The first image was captured from a slide containing mouse heart tissue [see Fig. 3(a)]. The tissue was stained by Hematoxylin and DAB. The size of the image is 256×256 pixels. The hyperspectral image was captured at 52 wavelengths from 423 to 812 nm. Two main object types can be seen at this image. The small blue objects correspond to the cell nuclei. The dark areas correspond to the borders of the blood vessels. Besides that, the light blue area corresponds to the cytoplasm of the cells and the white areas correspond to the lumens.

The second image was captured from a slide containing muscle tissue abuts dermal tissue [see Fig. 7(a)]. The tissue was stained by Hematoxylin, Eosin and DAB. The size of the image is 256×256 pixels. The hyperspectral image was captured at 34 wavelengths from 487 to 767 nm. There are three main object types at this image. The blue regions (Hematoxylin) correspond mainly to cell nuclei. The red regions (Eosin), correspond mainly to cytoplasm and connecting tissue. The brown region (DAB) corresponds mainly to muscle tissue cells.

In an ideal case (very low noise, even illumination, truly linear mixing model) the number of spectrally homogeneous compounds should be equal to the number of stains used to color the tissue. However, in a real-life scenario, there are additional compounds, corresponding to imaging process artifacts. When we decomposed both hyperspectral images to the number of compounds corresponding to the number of stains, our method and method from [5] yielded similar results. To demonstrate ability of both algorithm to cope with the artifacts compounds as well as with the stained compounds, we decompose both hyperspectral images to six compounds (see Fig. 2 for the scree plot of eigenvalues of the matrix Y).

In order to compare ability of both algorithms to cope with separation of biological compounds from artifacts, we estimate the average optical density for each separated compound m:

$$D_m = 1 - \frac{\sum_{i,j,k} e^{\beta_m(j,k)\alpha_m(\lambda_i)}}{\mathrm{IJK}}. \tag{8}$$

The average optical density D_m ranges from 0-absolutely transparent compound to 1 absolutely nontransparent compound. As the compounds, corresponding to the artifacts are less visible in the source images than the compounds, corresponding to the biological objects, the former should have significantly lower optical density. The attenuation coefficients of the first and the second image decompositions are presented in [21].

First, we address the decomposition results of our algorithm for the first image. The first two compounds correspond to the stained biological objects. The compounds 3, 4, 5 containing strip patterns are likely to correspond to artifacts of the Fourier-based spectroscope: they contain periodic signals that are due to residual frequencies that were not accounted for in the Fourier decomposition. The last compound is probably caused by diffraction on the edges of the second compound. The average optical density of the decomposed compounds (see Table I) corresponds to the observations. The first two compounds are most dense, and these compounds represent biological objects colored by different stains. The rest of the compounds are significantly less dense. Experiments showed that both dictionaries used for mixtures sparsification yielded similar artifacts pattern.

Next, we address the decomposition results provided by the algorithm [5] for the first image. Clearly, this algorithm captured the compounds corresponding to the biological objects (e.g., the compounds 2 and 4). Other separated compounds represent mixture of biological compounds instead of separated biological components. In addition, it was unable to cope with the artifacts compounds. The average optical density of the decomposed compounds (see Table I) demonstrates that the algorithm did not succeed to separate the compounds: except for the first compound, the average optical densities of the rest of compounds are of the same order. However, only two stains were used to color the image, therefore we would expect two main compounds.

Next, we address the decomposition results of our algorithm for the second image. The first three compounds correspond to the stained biological objects. The last three compounds correspond to the imaging artifacts (the compounds 4 and 6) and nonuniform illumination (the compound 5). The average optical density of the decomposed compounds (see Table I) correspond to the observations. The first three compounds are most dense, and these compounds represent biological objects colored by

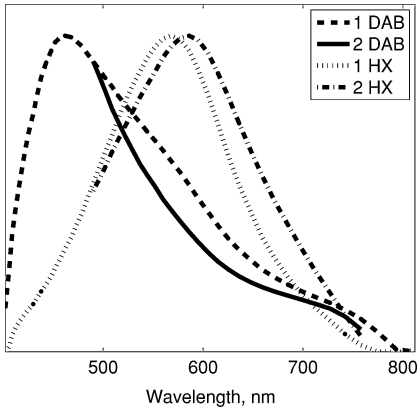


Fig. 4. Reconstructed spectra of hematoxylin and DAB of the separated compounds. "1 DAB" and "2 DAB" correspond to the DAB spectrum of the first and the second hyperspectral images correspondingly. "1 HX" and "2 HX" correspond to the hematoxylin spectrum of the first and the second hyperspectral images correspondingly.

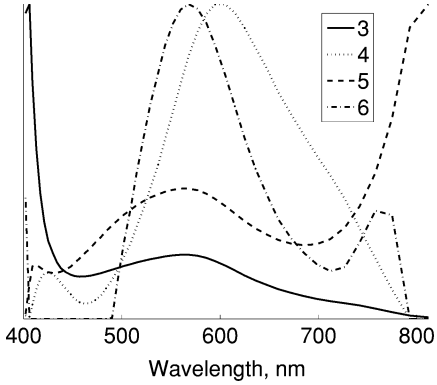


Fig. 5. Normalized spectra of separated compounds corresponding to the imaging artifacts for the first hyperspectral image.

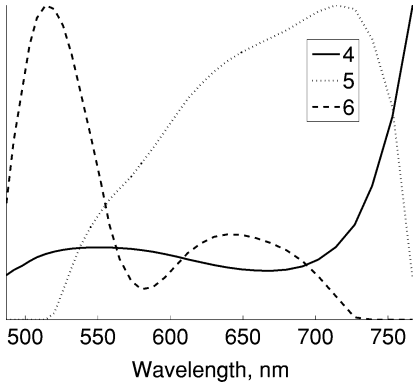


Fig. 6. Normalized spectra of separated compounds corresponding to the imaging artifacts for the second hyperspectral image.

different stains. The rest of the compounds are significantly less dense.

Finally, we consider the decomposition results provided by the algorithm [5] for the second image. In this case, the algorithm successfully separated the compound corresponding to cytoplasm and connecting tissue (the compound 3), but it failed to separate the compounds corresponding to the muscle tissue and cell nuclei. Other separated compounds represent mixture of biological compounds instead of separated biological components. The artifacts compounds were not separated as well. The average optical density of the decomposed compounds (see Table I) demonstrates that the algorithm did not succeed to separate the compounds: except for the first and the third compounds, the average optical densities of the rest of compounds are of the same order. However, only three stains were used to color the image, therefore we would expect three main compounds.

The spectra of the separated artifacts compounds for the first and the second images are presented on Figs. 5 and 6 correspondingly. The artifacts compounds 3, 5 on Fig. 5 and the artifact compounds 4, 5, 6 on Fig. 6 occur at the lowest and the highest wavelengths where imaging artifacts are most likely. While the spectrum energy of the compounds 4, 6 on Fig. 5 is localized near the center of measured wavelength range, the energy of all artifact compounds (and their optical density) is negligible compared to the energy of the compounds, corresponding to the biological tissues (see Table I for distribution of optical densities among the separated compounds for both images).

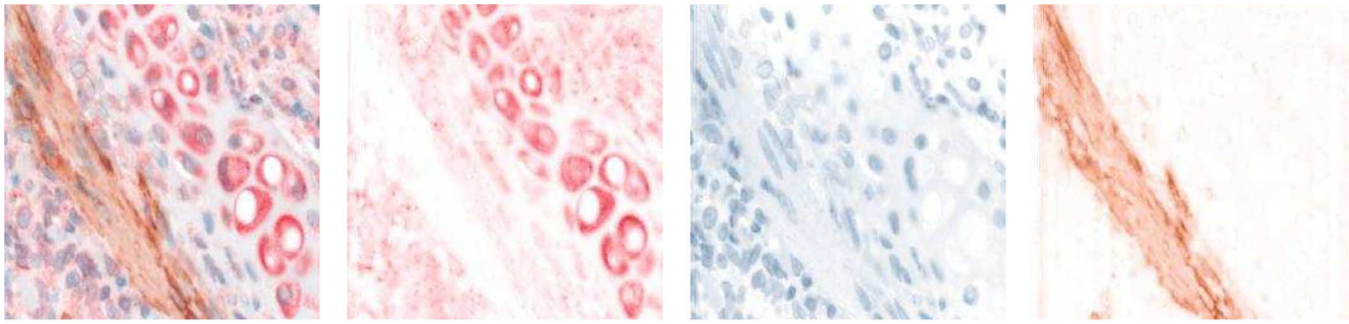
The differences in separation performance between the algorithm proposed in this paper and the algorithm presented in [5] may be explained by two reasons. First, we propose the multiplicative image formation model that is physically correct, while in [5] the additive image formation model is used. Second, our method employs sparsity for blind source separation, that is proven to provide superior results compared to nonsparse methods.

Next, we address the problem of color variance of the same dyes attached to the same specimen types. The first and the second hyperspectral images have common dies: DAB and hematoxylin. On Fig. 4 we present the hematoxylin and DAB spectra of the separated compounds for both first and second hyperspectral images. Both dies are represented by different spectra and therefore, they have different colors. However, the reconstructed spectra of hematoxylin and DAB can be easily distinguished and matched to a reference spectra by using some kind of distance between histograms like χ^2 statistics, L_2 distance or Kullback–Leibler divergence.

Finally, the RGB reconstruction of the compounds obtained by our algorithm are presented on Fig. 3 (for the first hyper-spectral image) and on Fig. 7 (for the second hyperspectral image). On Fig. 7, we see that overlapping muscle tissue, cytoplasm and cell nuclei are correctly separated by our algorithm.

VI. Conclusion

In this work we addressed a problem of blind separation of physical compounds on a microscopic slide for transmission light microscopy. First, we provided a multiplicative physical model of the problem. Second, using logarithmic transformation, we showed the reduction of the problem to the problem of overdetermined blind source separation. Sparse analysis of the sensor measurements followed by principal component analysis allowed us to reduce the overdetermined blind source separation problem to the "square" blind source separation problem. The "square" blind source separation problem is efficiently solved using Relative Newton algorithm. Finally, we demonstrated the transformation from the solution provided by Relative Newton algorithm to the hyperspectral images of the separated



(a) (b) (c) (d)

Fig. 7. Reconstructed color images of the second input hyperspectral image and three separated compounds. Original microscope illumination was used for color reconstruction. (a) RGB reconstruction of a hyperspectral image captured from a slide containing muscle tissue section abuts dermal tissue. (b) RGB reconstruction of the cytoplasm and connecting tissue compounds. (c) RGB reconstruction of the cell nuclei compound. (d) RGB reconstruction of the muscle tissue compound.

compounds. Reprojecting each source to the original dataspace avoids intrinsic ambiguities of BSS. The proposed algorithm was compared to the previous state-of-the-art algorithm. The comparison demonstrated that the proposed algorithm is faster and more robust to the imaging artifacts.

REFERENCES

[1] K. Kayser, "Introducing diagnostic pathology," *Diagn. Pathol.*, vol. 1, no. 1, p. 1, 2006.

[2] Y. Garini, M. Macville, S. du Manoir, R. Buckwald, M. Lavi, N. Katzir, D. Wine, I. Bar-Am, E. Schrck, D. Cabib, and T. Ried, "Spectral karyotyping," *Bioimaging*, vol. 4, no. 2, pp. 65–72, 1996.

[3] T. Zimmermann, "Spectral imaging and linear unmixing in light microscopy," *Adv. Biochem. Eng./Biotechnol.*, vol. 95, pp. 245–265, 2005.

[4] H. Shirakawa and S. Miyazaki, "Blind spectral decomposition of singlecell fluorescence by parallel factor analysis," *Biophy. J.*, vol. 86, pp. 1739–1752, 2004.

[5] A. Rabinovich, S. Agarwal, C. A. Laris, J. Price, and S. Belongie, "Unsupervised color decomposition of histologically stained tissue samples," in *Neural Inf. Process. Syst.*, 2003, pp. 667–674.

[6] J.-F. Cardoso, "Blind signal separation: Statistical principles," *Proc. IEEE*, vol. 86, no. 10, pp. 2009–2025, Oct. 1998.

[7] J.-F. Cardoso, "High-order contrasts for independent component analysis," *Neural Comput.*, vol. 11, no. 1, pp. 157–192, Jan. 1999.

[8] A. Hyvarinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. Neural Netw.* vol. 10, no. 3, pp. 626–634, May 1999.

[9] A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Comput.* vol. 7, no. 6, pp. 1129–1159, 1995.

[10] S. Amari, A. Cichocki, and H. H. Yang, , D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo, Eds., "A new learning algorithm for blind signal separation," in *Adv. Neural Inf. Process. Syst.*. Cambridge, MA: MIT Press, 1996, vol. 8, pp. 752–763.

[11] M. Zibulevsky and B. Pearlmutter, "Blind source separation by sparse decomposition in a signal dictionary," *Neural Comput.*, vol. 13, no. 4, pp. 863–882, 2001.

[12] M. Zibulevsky, "Blind source separation with relative newton method," in *Proc. ICA*, Nara, Japan, 2003, pp. 897–902.

[13] Y. Lu and M. N. Do, "A new contourlet transform with sharp frequency localization," in *IEEE Int. Conf. Image Process.*, Atlanta, GA, Oct. 2006, pp. 1629–1632.

[14] S. Mallat, *A Wavelet Tour of Signal Processing*. San Diego, CA: Academic, 1998.

[15] M. Lewicki and T. Sejnowski, "Learning overcomplete representations," *Neural Computat.*, vol. 12, no. 2, pp. 337–365, 2000.

[16] M. Lewicki and B. Olshausen, "A probabilistic framework for the adaptation and comparison of image codes," *J. Opt. Soc. Amer.*, vol. 16, pp. 1587–1601, 1999.

[17] B. Olshausen and D. Field, "Sparse coding with an overcomplete basis set: A strategy employed by v1?," *Vis. Res.*, vol. 37, pp. 3311–3325, 1997.

[18] D. Pham and P. Garat, "Blind separation of mixture of independent sources through a quasi-maximum likelihood approach," *IEEE Trans. Signal Process.*, vol. 45, no. 7, pp. 1712–1725, Jul. 1997.

[19] A. Cichocki, R. Unbehauen, and E. Rummert, "Robust learning algorithm for blind separation of signals," *Electron. Lett.*, vol. 30, no. 17, pp. 1386–1387, Aug. 1994.

[20] C.-J. Lin, "Projected gradient methods for non-negative matrix factorization," *Neural Comput.*, vol. 19, no. 10, pp. 2756–2779, 2007.

[21] G. Begelman, "Processing and interpretation of biological microscopical images," Ph.D. dissertation, Technion—Israel Institute of Technology, Haifa, Israel, 2009.