

# Offset-Sparsity Decomposition for Enhancement of Color Microscopic Image of Stained Specimen in Histopathology: Further Results

Ivica Kopriva<sup>1</sup>, Marijana Popović Hadžija<sup>2</sup>, Mirko Hadžija<sup>2</sup> and Gorana Aralica<sup>3,4</sup>

<sup>1</sup>Division of Laser and Atomic Research and Development

<sup>2</sup>Division of Molecular Medicine

Ruđer Bošković Institute, Bijenička cesta 54, P.O. Box 180, 10002, Zagreb, Croatia

<sup>3</sup>Department of Pathology and Cytology, Clinical Hospital Dubrava, Avenija Gojka Šuška 6, Zagreb 10000, Croatia

<sup>4</sup>School of Medicine, University of Zagreb, Šalata 3, Zagreb 10000, Croatia

e-mail: ikopriva@irb.hr, Marijana.Popovic-Hadzija@irb.hr, Mirko.Hadzija@irb.hr, garalica@kdb.hr

## Abstract

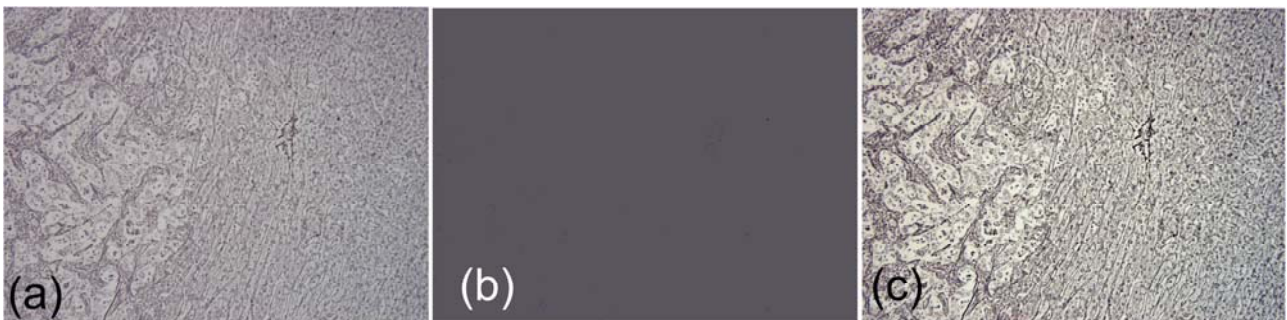
Recently, novel data-driven offset-sparsity decomposition (OSD) method was proposed by us to increase colorimetric difference between tissue-structures present in the color microscopic image of stained specimen in histopathology. The OSD method performs additive decomposition of vectorized spectral images into image-adapted offset term and sparse term. Thereby, the sparse term represents an enhanced image. The method was tested on images of the histological slides of human liver stained with hematoxylin and eosin, anti-CD34 monoclonal antibody and Sudan III. Herein, we present further results related to increase of colorimetric difference between tissue structures present in the images of human liver specimens with pancreatic carcinoma metastasis stained with Gomori, CK7, CDX2 and LCA, and with colon carcinoma metastasis stained with Gomori, CK20 and PAN CK. Obtained relative increase of colorimetric difference is in the range [19.36%, 103.94%].

**Keywords:** Histopathological image enhancement, colorimetric difference, offset-sparsity decomposition.

## 1 INTRODUCTION

In setting diagnoses pathologists look for problem-specific visual cues, or features, in histopathological images to distinguish between healthy and diseased tissue. These features, as an example, may come from specific characteristics of cells such as size, shape or texture [1, 2]. To visualize them various stains and tags are attached to biological tissues to improve the colorimetric difference between the tissue components (histological structures), thereby improving their visibility [3]. As an example, it is known that in hematoxylin-eosin (H&E)-stained slides, color information is essential to discriminate between healthy and diseased tissue [4]. However, the tissue color and texture can vary considerably between specimens because of the variations in the tissue preparation processes. These non-biological experimental variations are also known as batch

effects [5, 6]. These variations can change the quantitative morphological image features, and this makes it difficult to reach an accurate diagnosis [5], e.g. in the field of digital pathology, i.e. computerized image analysis [7], that has entered an era of computer-assisted diagnosis and treatment of medical conditions based on an analysis of medical images [1, 2, 8]. As an example, diagnosis of hepatocellular carcinoma that is based on segmentation of trabecula is affected by the variation of color and/or texture [9]. Thus, problems related to the variations of quality of the staining process set the motivation for development of an automated image enhancement method, particularly for enhancing the colorimetric difference between the histological structures present in the images of a stained specimen. To this end, offset-sparsity decomposition (OSD) method has been recently developed, [10], for enhancement (increase of colorimetric difference between the tissue structures) of the image quality of stained specimen in histopathology. The OSD method is a special (degenerative) case of the rank-sparsity decomposition (RSD) method that decomposes a matrix into a sum of low-rank and sparse matrices [11, 12]. As opposed RSD, the OSD method [10] decomposes vectorized spectral images into a sum of an approximately constant offset vector and a sparse vector. Thereby, the offset term corresponds to the  $\ell_2$ -norm-based regularization, and the sparse term corresponds to the  $\ell_1$ -norm-based regularization in an optimization problem related to the minimization of the difference between the vectorized spectral images and the model. Since for a vector, the nuclear norm equals the  $\ell_2$ -norm, related optimization problem in the OSD case is simpler than in the RSD case. That is, thresholded singular value decomposition (SVD) required by nuclear norm (low-rank) regularization in the RSD problem is trivial to compute for  $\ell_2$ -norm regularization in the OSD problem [10]. The OSD method is applied on decomposition of vectorized grayscale intensity images that correspond to red, green, and blue (RGB) colors by modeling each vectorized image as additive superposition of the offset term, noiseless image and noise. The OSD method is illustrated in Figure 1 through enhancement of the histopathological image of human liver with pancreatic carcinoma metastasis stained by Gomori, Figure 1a. The color offset term is shown in Figure 1b, while Figure 1c shows an enhanced image that is captured by the “sparse” term and is of actual interest. Relative improvement of colorimetric difference, measured by colorfulness measures estimated from the image [13], is 76.1%.



**Figure 1.** OSD approach to the enhancement of a color microscopic image of the specimen of human liver with pancreatic carcinoma metastasis stained by Gomori: (a) original image of stained specimen with colorfulness 0.3092; (b) image adapted color offset; (c) enhanced image with colorfulness 0.5445. Relative improvement is 76.1%.

In [10] we have demonstrated the OSD-based increase of colorimetric difference between tissue structures in the color histological slide images of the human liver specimens with primary malignant tumor and metastasis from colon and pancreatic carcinoma stained by H&E and fatty liver stained by anti-CD34 antibody and Sudan III. The purpose of the present paper is to demonstrate further the OSD-based increase of colorimetric difference between tissue structures in the images of human liver specimens with metastasis from pancreatic carcinoma stained with Gomori, CK7, CDX2 and LCA, and with metastasis from colon carcinoma stained with Gomori, CK20 and PAN CK. Obtained relative increase of colorimetric difference is in the range [19.36%, 103.94%], see Figures 1 to 3. Thus, application of the proposed method on different type stains further contributes to the relevance of the OSD method for the field of digital pathology.

## 2 MATERIALS AND METHODS

Within the present paper, and following [10], we use the following notation. An underlined upper-case bold letter, e.g.,  $\underline{\mathbf{X}} \in \mathbb{R}_{0+}^{I_1 \times I_2 \times 3}$ , denotes a three-dimensional (3D) RGB image tensor consisting of three spectral images corresponding with the red, green, and blue colors, where each image measures  $I_1 \times I_2$  pixels. An upper-case bold letter, e.g.,  $\mathbf{X}$ , denotes a matrix; a lower-case bold letter, e.g.,  $\mathbf{x}$ , denotes a vector; and an italicized lower-case letter, e.g.,  $x$ , denotes a scalar. The random variable  $e$  that follows the Gaussian distribution with zero mean and variance  $\sigma^2$  is denoted as  $e \sim N(0, \sigma^2)$ . Standard model of the observed image  $\{\mathbf{b}_n \in \mathbb{R}_{0+}^{I_1 \times I_2}\}_{n=1}^3$  assumed by many image denoising/enhancement methods, [14, 15], involves noiseless image  $\{\mathbf{s}_n \in \mathbb{R}_{0+}^{I_1 \times I_2}\}_{n=1}^3$  and noise  $\{\mathbf{e}_n \in \mathbb{R}_{0+}^{I_1 \times I_2}\}_{n=1}^3$ , where  $\mathbf{e}_n \sim N(\mathbf{0}, \sigma_n^2 \mathbf{I})$ , stands for the additive white Gaussian noise (AWGN). The OSD method, in addition to that, also includes small  $\ell_2$ -norm offset term  $\{\mathbf{a}_n \in \mathbb{R}_{0+}^{I_1 \times I_2}\}_{n=1}^3$  to represent the intensity of the observed spectral images of a color microscopic image  $\underline{\mathbf{B}} \in \mathbb{R}_{0+}^{I_1 \times I_2 \times 3}$  is used in [10]:

$$\mathbf{b}_n = \mathbf{s}_n + \mathbf{a}_n + \mathbf{e}_n \quad n \in \{1, 2, 3\} \quad (1)$$

The offset term models "shadows" that are present in the image due to various nonbiological experimental variations, a.k.a. batch effects, [10]. The OSD method derived in [10] estimates  $\{\mathbf{a}_n\}_{n=1}^3$  and  $\{\mathbf{s}_n\}_{n=1}^3$  using  $\{\mathbf{b}_n\}_{n=1}^3$  only. After matricization,  $\{\mathbf{a}_n \in \mathbb{R}_{0+}^{I_1 \times I_2} \rightarrow \mathbf{A}_n \in \mathbb{R}_{0+}^{I_1 \times I_2}\}_{n=1}^3$ , and tensorisation  $\{\underline{\mathbf{A}}(:, :, n) = \mathbf{A}_n\}_{n=1}^3$ ,  $\underline{\mathbf{A}} \in \mathbb{R}_{0+}^{I_1 \times I_2 \times 3}$  represents the image-adapted color offset.  $\underline{\mathbf{S}} \in \mathbb{R}_{0+}^{I_1 \times I_2 \times 3}$ , formed analogously from  $\{\mathbf{s}_n \in \mathbb{R}_{0+}^{I_1 \times I_2}\}_{n=1}^3$ , represents an enhanced color microscopic image that, comparably with the original image  $\underline{\mathbf{B}}$ , has an improved colorimetric difference between the histological structures present in the specimen. Under the AWGN assumption, the optimal estimates of  $\mathbf{a}_n$  and  $\mathbf{s}_n$  are obtained by solving the log-likelihood problem regularized by the addition of  $\ell_2$ - and  $\ell_1$ -penalties as follows:

$$(\hat{\mathbf{a}}_n, \hat{\mathbf{s}}_n) = \min_{(\mathbf{a}_n, \mathbf{s}_n)} \left\{ \frac{1}{2} \|\mathbf{b}_n - \mathbf{a}_n - \mathbf{s}_n\|_2^2 + \mu \|\mathbf{a}_n\|_2 + \mu\lambda \|\mathbf{s}_n\|_1 \right\} \quad n \in \{1, 2, 3\} \quad (2)$$

Optimization problem (2) is solved by the OSD Algorithm 1 in [10]. It admits a unique solution with the value of the regularization parameter set to  $\lambda = 1/\sqrt{\max(I_1, I_2)}$ . To avoid the color artifacts, the enhancement of the color images is preferably executed in the CIE  $L^*a^*b^*$  color space instead of the RGB color space. In the case of the OSD approach to color image enhancement, which operates independently on the color channels in RGB color space, color artifacts are avoided because of the following reason: Even though the optimization problem implied by Eq. (2) is solved for each channel independently, the data fidelity terms,  $\left\{ 2^{-1} \|\mathbf{b}_n - \mathbf{a}_n - \mathbf{s}_n\|_2^2 \right\}_{n=1}^3$ , prevent enhanced spectral images  $\{\mathbf{s}_n\}_{n=1}^3$  to deviate significantly from the experimental spectral image  $\{\mathbf{b}_n\}_{n=1}^3$ . Because of the same reason, the intensity offsets  $\{\mathbf{a}_n\}_{n=1}^3$  that are extracted independently in each spectral channel yield, when merged together, the image adapted color offset term. In addition to Fig. 1, this can be seen in Figs. 2 and 3 in Section 3. We summarize the OSD algorithm in Algorithm 2 that is reproduced from [10].

**Algorithm 2.** The OSD\_rgb algorithm for the enhancement of a color microscopic image of the stained specimen.

**Input:**  $\mathbf{B} \in \mathbb{R}_{0+}^{3 \times I_1 I_2}$  unfolded color image of the stained specimen with vectorized grayscale images  $\{\mathbf{b}_n \in \mathbb{R}_{0+}^{1 \times I_1 I_2}\}_{n=1}^3$  measuring  $I_1 \times I_2$  pixels. Sparsity regularization constant in Eq. (2):  $\lambda = 1/\sqrt{I_1 \times I_2}$ , regularization constant in Eq. (2):  $\mu = 10^{-3}$ , thresholding constant in Eqs. (13) and (14) in [10]:  $L = 2$ .

**for**  $n = 1 : 3$

$(\mathbf{a}_n, \mathbf{s}_n) = \text{Algorithm 1}(\mathbf{b}_n, \lambda, \mu, L)$ .

**end for**

**Set:**  $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \end{bmatrix}$ ,  $\mathbf{S} = \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \mathbf{s}_3 \end{bmatrix}$ .

**Output:**  $\mathbf{S} \in \mathbb{R}_{0+}^{3 \times I_1 I_2}$  unfolded enhanced color image of the stained specimen.  $\mathbf{A} \in \mathbb{R}^{3 \times I_1 I_2}$  unfolded image with the color offset term. The enhanced color image is obtained by tensorizing  $\mathbf{S}$ :  $\underline{\mathbf{S}} \in \mathbb{R}_{0+}^{I_1 \times I_2 \times 3}$ .

To quantify the performance of OSD image enhancement algorithm, a measure of the colorimetric difference between the histological structures is estimated directly from the image, [13]. It measures the amount of chrominance information that humans perceive. The colorfulness measure is defined as follows [13]:

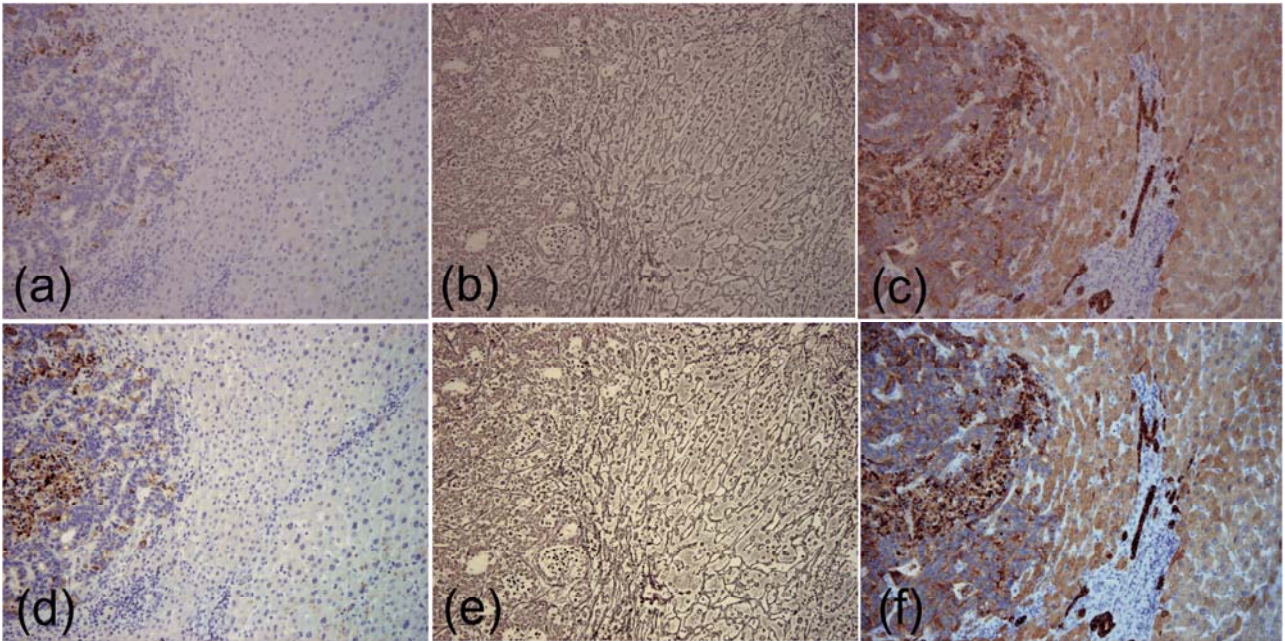
$$colorfulness = 0.02 \times \log \left( \frac{\sigma_{\alpha}^2}{|\mu_{\alpha}|^{0.2}} \right) \times \log \left( \frac{\sigma_{\beta}^2}{|\mu_{\beta}|^{0.2}} \right) \quad (3)$$

where  $\alpha$  = Red - Green color images;  $\beta$  =  $0.5 \times (\text{Red} + \text{Green})$  - Blue color images; and  $\sigma_{\alpha}^2$ ,  $\sigma_{\beta}^2$ ,  $\mu_{\alpha}$ , and  $\mu_{\beta}$  represent the variance and mean along the  $\alpha$  and  $\beta$  opponent color axes, respectively.

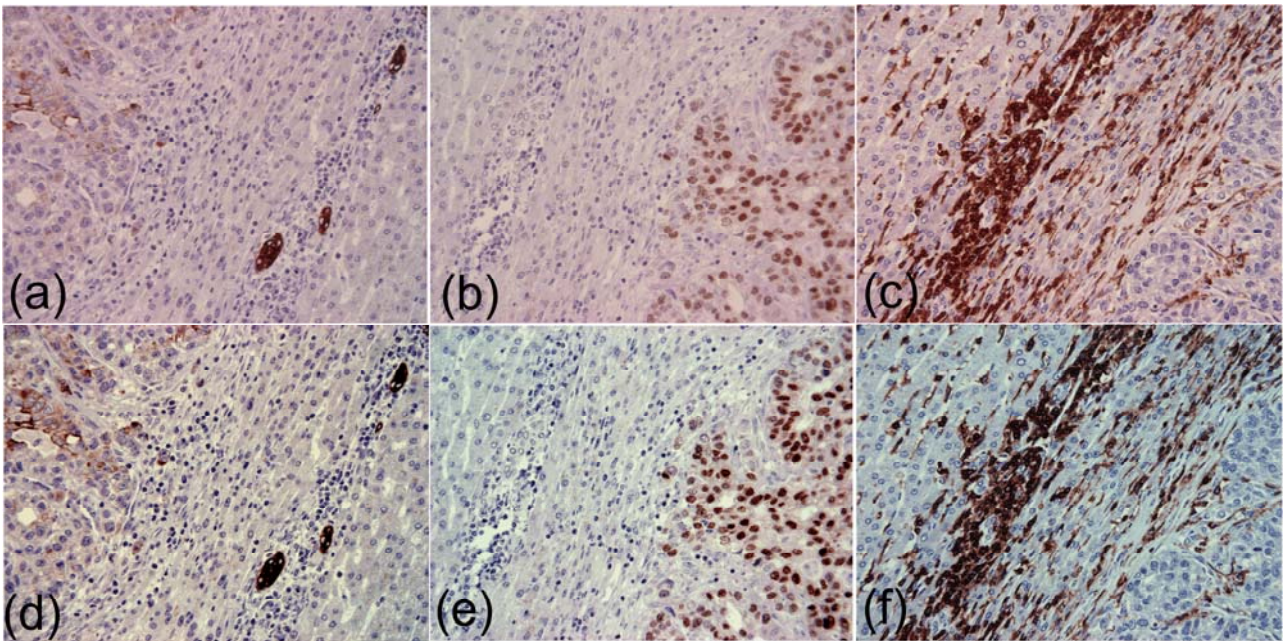
### 3 RESULTS

In Figure 1 we have illustrated the enhancement of color microscopic image obtained by applying the OSD-based approach on the image of the human liver specimen with pancreatic carcinoma metastasis stained with Gomori. Increase of the colorfulness attribute relative to the original image is 76.1%. Furthermore, Figure 2a to 2c shows images of the human liver specimen with colon carcinoma metastasis stained respectively with CK20, Gomori and PAN CK. The OSD-enhanced images are shown in Figure 2d to 2f in respective order. Figure 3a to 3c shows images of the human liver specimen with pancreatic carcinoma metastasis stained respectively with CK7, CDX2 and LCA. The OSD-enhanced images are shown in Figure 3d to 3f in respective order. Relative increase of the colorfulness measure varies between 19.36% and 103.94%. In particular, reticulin fiber net is perceived better in enhanced image in Figure 1c than in the original image in Figure 1a. Furthermore, in Figures 3d and 3f membranous positivity for CK7 and LCA and nuclear positivity for CDX2 in Figure 3e are more pronounced than in Figures 3a and 3c, respectively Figure 3b. Also, background tissue is non-stained and more distinctive of immunohistochemically positive cells in Figures 3d to 3f than in Figures. 3a to 3c.





**Figure 2.** Images of the specimen of human liver with colon carcinoma metastasis stained with: (a) CK20, (b) Gomori, and (c) PAN. Colorfulness value in respective order: 0.2792, 0.2544, 0.4393. (e) to (f): OSD-enhanced images corresponding to images (a) to (c). Colorfulness value in respective order: 0.5694, 0.4877, 0.7631. Relative improvement: 103.94%, 91.71%, 73.71%.



**Figure 3.** Images of the specimen of human liver with pancreatic carcinoma metastasis stained with: (a) CK7, (b) CDX2, and (c) LCA. Colorfulness value in respective order: 0.3328, 0.4032, 0.5708. (e) to (f): OSD-enhanced images corresponding to images (a) to (c). Colorfulness value in respective order: 0.5324, 0.6098, 0.6813. Relative improvement: 59.98%, 51.24%, 19.36%.

## ACKNOWLEDGEMENTS

This work was supported in part through the FP7-REGPOT-2012-2013-1, Grant Agreement Number 316289 – InnoMol and in part through the Grant 9.01/232 funded by Croatian Science Foundation.

## REFERENCES

- [1] Gurcan, M., Boucheron, L., Can, A., Madabhushi, A., Rajpoot, N., Yener, B., "Histopathological image analysis: a review," *IEEE Rev. Biomed. Eng.* 2, 147-171 (2009).
- [2] Dundar, M. N., Badve, S., Bilgin, G., Raykar, V., Jain, R., Sertel, O., Gurcan, M. N., "Computerized classification of intraductal breast lesions using histopathological images," *IEEE Trans. Sig. Proc.* 58(7), 1977-1984 (2011).
- [3] Bautista, P. A., Yagi, Y., "Digital simulation of staining in histopathology multispectral images: enhancement and linear transformation of spectral transmittance," *J. Biomed. Opt.* 17(5), 056013 (2012).
- [4] Gavrilović, M., Azar, J. C., Linblad, J., Whalbi, C., Bentsson, E., Busch, C., Carlbom, I. B., "Blind color decomposition of histological images," *IEEE Trans. Med. Imag.* 32(6), 983-994 (2013).
- [5] Kothari, S., et al., "Removing batch effects from histopathological images for enhanced cancer diagnosis," *IEEE J. Biomed. Health Inf.* 18(3), 765-772 (2014).
- [6] Phan, J. H., Quo, C. F., Cheng, C., Wang, M. D. "Multiscale integration of -omic, imaging, and clinical data in biomedical informatics," *IEEE Rev. Biomed. Eng.* 5, 75-87 (2012).
- [7] Mendez, A. J. Tahoces, P. G., Lado, M. J., Souto, M., Vidal, J. J., "Computer-aided diagnosis: Automatic detection of malignant masses in digitized mammograms," *Med. Phys.* 25, 957-964 (1998).
- [8] Srinivas, U., Mousavi, H. S., Monga, V., Hattel, A., Jayarao, B., "Simultaneous sparsity model for histopathological image representation and classification," *IEEE Trans. Med. Imag.* 33(5), 1163-1179 (2014).
- [9] Ishikawa M, Ahi, S. T., Murakami, Y., Kimura, F., Yamaguchi, M., Abe, T., Hashiguchi, A., Sakamoto, M., "Automatic segmentation of hepatocellular structure from HE-stained liver tissue," *Proc. SPIE* 8676, 867611 (2013).
- [10] Kopriva, I., Hadžija Popović, M., Hadžija, M., Aralica, G., "Offset-sparsity decomposition for automated enhancement of color microscopic images of stained specimen in histopathology," *J. Biomed. Opt.* 20 (7), 076012 (2015).
- [11] Candès, E. J., Li, X., Ma, Y., Wright, H., "Robust principal component analysis?," *J. ACM* 58, 11 (2011).
- [12] Chandrasekaran, V., Sanghavi, S., Parrilo, P. A., Willsky, A. S., "Rank-sparsity incoherence for matrix decomposition," *SIAM J. Opt.* 21, 572-596 (2011).
- [13] Panetta, K., Gao, C., Agaian, S., "No reference color image contrast and quality measure," *IEEE Trans. Cons. Elec.* 59, 643-651 (2013).
- [14] Luisier, F., Blu, T., Unser, M., "A new sure approach to image denoising: Interscale orthonormal wavelet thresholding," *IEEE Trans. Image Process.* 16(3), 593-606 (2007).
- [15] Portilla, J., Strela, V., Wainwright, M. J. Simoncelli, E. P., "Image denoising using scale mixtures of Gaussians in the wavelet domain," *IEEE Trans. Image Process.* 12(11), 1338-1351 (2003).