STRUCTURE-PRESERVED COLOR NORMALIZATION FOR HISTOLOGICAL IMAGES

Abhishek Vahadane[1,3], *Tingying Peng*[1], *Shadi Albarqouni*[1], *Maximilian Baust*[1], *Katja Steiger*[4], *Anna Melissa Schlitter*[4], *Amit Sethi*[3], *Irene Esposito*[4,5], *Nassir Navab*[1,2].

[1] Computer Aided Medical Procedures (CAMP), Technical University of Munich, Germany

[2] Computer Aided Medical Procedures (CAMP), Johns Hopkins University, USA

[3]Indian Institute of Technology Guwahati, India

[4]Institute of Pathology, Technical University of Munich, Germany

[5]Institute of Pathology, Medical University of Innsbruck, Austria

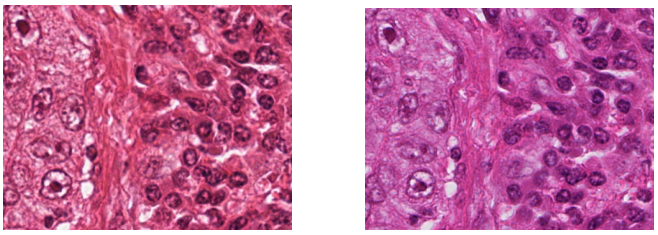
ABSTRACT

Automated image processing and quantification are increasingly gaining attention in the field of digital pathology. However, a common problem that encumbers computerized analysis is the color variation in histology, due to the use of different microscopes/scanners, or inconsistencies in tissue preparation. In this paper, we present a novel color normalization technique to bring a histological image (source image) into a different color appearance of a second image (target image), which therefore standardizes the color representation of both images. In particular, by incorporating biological stain-sparse regularized stain separation, our color normalization technique preserves the structural information of the source image after color normalization, which is very important for subsequent image analysis. Both qualitative and quantitative validation demonstrates the superior performance of our stain separation and color normalization techniques.

Index Terms— Histopathology image, color normalization, stain separation, non-negative matrix factorization

1. INTRODUCTION

Traditional histopathological study is commonly conducted by a pathologist, based on visual examination of tissue under a light microscope. The ability to visualize or differentially identify microscopic structures is facilitated by coloring the tissue by histological stains. In this process, color variation in the obtained histology images is almost inevitable due to: 1) use of different microscopes or scanners for image acquisition (as shown in Fig.1); 2) different staining solutions from different chemical manufacturers; 3) variation of staining protocols across different labs. In general, color variation is not a big obstacle for professional pathologists, as they are trained to cope with this visual variability in their routine diagnoses. However, it has become more of an issue for many automatic image processing and quantification techniques. For example, it has been shown in two recent studies that the performance of learning based tumor segmentation techniques deteriorates when the color of the test images differs from that of the training dataset [1] [2].



(a) Aperio scanner (b) Hamamatsu scanner

Fig. 1. The same tissue section scanned by different scanners

In order to improve automatic image analysis, many researchers have developed techniques to change histology images that were acquired under different conditions into a common color space. These existing normalization methods can be generally categorized into: 1) Histogram specification in RGB space [3]; 2) Matching histogram statistics (i.e. mean and standard deviation) after transforming the RGB image into $l\alpha\beta$ color space [4]; 3) Normalization performed after separating the RGB image into channels of stain concentrations [1] [5] [6]. Approaches 1) and 2) implicitly assume similar proportion of stains presented in two histology images, which is not always correct. Therefore, they could easily introduce visual artifact and alter the structural information of an image after normalization (as shown in Fig. 4 and detailed explanation will be provided in Section 3.2). By contrast, approaches 3) are on a promising track as they relate color appearance of histology to stain concentration, which is biologically valid as color in histology stems from stains. Yet drawbacks exist for their stain separation techniques and hence an invariant biological structure is not guaranteed in their normalization processes.

Stain separation is a classic research topic in the field of digital pathology. Color deconvolution [7] has been used

A.V. thanks to HERITAGE Erasmus Mundus cohort-2 fellowship and anonymous reviewers

extensively in histopathology image analysis to separate an RGB image into the concentration maps of stains presented in each pixel. However, this method requires the use of experimentally measured stain color matrix and hence a sample set is only valid for images using the same staining protocol. In the above color normalization techniques, although [1] and [6] have both derived image-specific staining color matrices, they rely on supervised color decomposition, which is therefore inevitably limited to the color variation of their training dataset. A common application for color normalization is to improve the color appearance/contrast of a low-quality histology slide (e.g. faded stains). It would be very challenging for a training dataset to cover histology slides in all possible negative scenarios. Instead, Macenko [5] uses unsupervised singular value decomposition to extract stain vectors and hence is more robust to image variations. However, color normalization assumes similar proportion of stains presented in two images and hence limits the performance.

Therefore, in this paper, we propose a novel color normalization technique to match a given histology image (source) to the color appearance of another image (target). Taking the promising approach of [1] [5] [6], the first step of our color normalization is also transforming an RGB image into stain concentrations. We propose sparse stain separation by adding sparsity constraints on stain channels, inspired by the biological fact about stain binding. Subsequently, we normalize the source image by changing its color appearance to that of target image while preserving its own stain concentration. Here, the stain concentration is preserved by ensuring accurate stain separation. In this way, the biological significance of tissue structure is kept during the normalization process, which is a big advantage over other competing methods.

2. PROPOSED STAIN SEPARATION AND COLOR NORMALIZATION

The Beer-Lambert law relates the attenuation of transmitted light through the medium (here, tissue specimen) with the concentration of the attenuating substance (here stains). Let I be $m \times n$ matrix ($m = 3$ RGB channels, $n =$ number of pixels) of RGB intensities in an acquired bright field microscopic image, I_0 be the illuminating light on the sample, W be $m \times r$ matrix (r = number of stains) called stain color matrix whose columns represent RGB color of each stain, and H be $r \times n$ stain concentration matrix, rows of which represent total amount of stained tissue, then by Beer-Lambert law, (for good proof of Beer-Lambert law to eq.1 for histology see [8])

$$I = I_0 \exp(-WH). \tag{1}$$

Let V be the relative optical density, defined as:

$$V = \log I_0 - \log I. \tag{2}$$

Hence, equation (1) becomes,

$$V = WH. \tag{3}$$

Different stain separation methods take different approaches to estimate W and H. For example, color deconvolution [7] measures W experimentally and subsequently use pseudo inverse transform to obtain H. By contrast, blind color decomposition methods estimate both unknown W and H [8]. Non-negative matrix factorization (NMF) is an excellent matrix factorization technique [9]. It has been used for stain separation, with non-negativity constraints on W and H [10]:

$$\underset{W,H}{min} \ \frac{1}{2}\|V - WH\|_F^2, \ \text{ such that } \ W, H \geq 0. \tag{4}$$

We improve the above NMF objective by including L1 sparseness regularization on stains H, by the fact that a type of stain is only bound to certain biological structures. As example, in H&E, H binds to the nuclei and E to the cytoplasm & stroma, hence both stains when separated should look sparse. This promote us to include sparsity constraints on H:

$$\underset{W,H}{min} \ \frac{1}{2}\|V - WH\|_F^2 + \lambda \sum_{j=1}^{r} \|H(j,:)\|_1, \ \ W, H \geq 0. \tag{5}$$

This optimization problem is solved by alternating between W and H which optimizes one set of parameters whilst keeping other fixed. This (eq.5) is closely related to the well-established dictionary learning objective, but with additional non-negative constraints on dictionary atoms W and coefficients H. Thus, we use the publicly available SPAMS toolbox [11] in our implementation. A similar formulation of sparse promoted stain separation was recently reported in [12]. However, their method is only an extension of color deconvolution method [7] and still requires W to be experimentally determined, which limits its applicability over large histological dataset with considerable color variation. Hence proposed sparse-regulated NMF is novel for stain separation. After an accurate separation of color appearance and staining concentration for both source and target images, we change the color appearance of source W_s to be as target W_t, which is

Inputs: source & target image, number of stains r, sparsity regularization parameter λ (default $\lambda = 0.1$)
1. Convert source & target into optical densities using eq.(2)
2. Source stain separation: $V_s = W_s H_s$ using eq.(5)
3. Target stain separation: $V_t = W_t H_t$ using eq.(5)
4. Adjust the dynamic range of H_s to be same as that of H_t to form the normalized source stains H_{snorm}. The dynamic range is robustly estimated with a pseudo maximum (99%).
5. Color exchange: $V_{snorm} = W_t H_{snorm}$
6. Project the V_{snorm} into RGB color space using eq.(2) to get the normalized source image

Output: normalized source image

Table 1. Proposed color normalization algorithm

subsequently combined with a normalized version of the stain concentration H_s of source image, to generate the (linearly) normalized source image. In this way, our color normalization techniques only change the stain color appearance whilst preserving structures (as we do not change the relative stain concentration). The detailed algorithm is shown in Table 1.

3. RESULTS AND VALIDATION

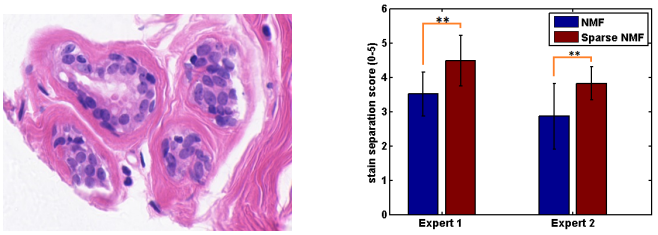
3.1. Stain Separation

The quality of proposed stain separation was evaluated by expert pathologists, in comparison with NMF based stain separation [10]. 50 H&E ($r = 2$) stained images are randomly selected from breast cancer cell database (http://bioimage.ucsb.edu/research/bio-segmentation). Two experts independently evaluate the quality of separation using scores in the range of 0-5 (0 poorest, 5 perfect). Fig. 2b shows bar plot of statistics (mean±standard deviation) of expert scores. It shows that our method achieves an average scores of 4.48 (expert1) and 3.82 (expert2), which are acceptable for clinical practices. In comparison, NMF only achieves an average score of 3.52 (expert1) and 2.86 (expert2), which are significantly lower than ours ($p < 0.001$ using Wilcoxon signed-rank test). An example of the separation results shown in Fig. 2 illustrates that our method achieves a much clearer separation with less disturbing background stains (e.g. less protein in H channel and less nuclei in E channel), suggesting the crucial role of sparse regularization.

3.2. Color Normalization

Validation for color normalization in histology images is not an easy task, especially in a quantitative manner. The results presented in previous studies ([1] [5] [6]) are mostly qualitative or indirect validation (e.g. improvement in the segmentation performance). Here we use a direct benchmark validation for color variation using the ICPR2014 contest dataset (http://mitos-atypia-14.grand-challenge.org/). In this dataset, each breast cancer section is acquired using two different scanners, which results in color variation between them (an example is shown in Fig.1).

We randomly sample 300 images from Aperio scanner as our source images, and 300 images from Hamamatsu scanner as target images. We perform color normalization to change the source color appearance into the target color appearance and subsequently compare the normalized source to the corresponding target image of the same tissue section (ground-truth). We choose multi-channel normalized mutual information and Pearson correlation as similarity metrics between the normalized source and ground-truth (multi-channel measure is taken as the average of three RGB channel-wise measure). The quantitative results shown in Fig. 3 demonstrates our superior performance over other competing methods ($p < 0.05$ by Wilcoxon signed-rank test).

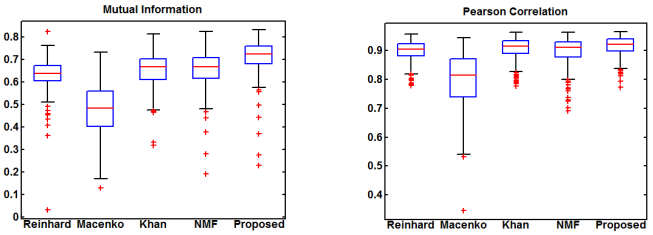


(a) H&E image (b) Expert scores ($^{**}p < 0.001$)

(c) NMF H stain (d) NMF E stain

(e) Sparse NMF H stain (f) Sparse NMF E stain

Fig. 2. NMF and sparse NMF based stain separation



(a) (b)

Fig. 3. Quality (similarity) metrics between normalized source and ground truth for different normalization methods.

The proposed color normalization method can also be used to improve contrast of low-quality histology images, such as faded stain. As shown in Fig. 4, Reinhard method [4] fills lumens in the source image with stains (Fig. 4c); Macenko method [5] generates unrealistic stain colors (Fig. 4d); in comparison, both Khan's method [1] and our method achieves better results yet a close inspection of the two (Fig.4e vs Fig.4f) illustrates that our method preserves tissue structure better than Khan's method: e.g. we maintain the color contrast of the two nuclei (green circles) whilst Khan's method make them almost identical in color; the small weakly-stained

nucleus (pointed out by the yellow arrow) is preserved after our normalization whilst disappears after Khan's.

4. DISCUSSIONS AND CONCLUSIONS

Color normalization is an essential step to remove inherent color variation in histopathological images and hence is useful for analyzing disease and its progression on large datasets from different pathology labs. A potential danger of such a technique is that tissue structures presented in the original image could be altered after normalization (many pathologists are concerned about this). In this paper, we propose a novel structure-preserved color normalization scheme, which changes the color of one image (source) into another (target) but reliably keeps the invariance of the structural information of the source image. One key step of our normalization scheme is an accurate stain separation of both source and target images, which is achieved by a sparse regularized non-negative matrix factorization. Both qualitative and quantitative validation demonstrates the superior performance of proposed stain separation and color normalization techniques.

Currently, the sparse regulated parameter λ, is chosen by experience and kept the same for different stains and images ($\lambda = 0.1$). In future, we plan to automatically tune this parameter and to make it adaptive to different stains. In our prospective, this can be achieved by learning from examples.

5. REFERENCES

[1] AM. Khan *et.al*, "A nonlinear mapping approach to stain normalization in digital histopathology images using image-specific color deconvolution," *IEEE TBME*, vol. 61, no. 6, pp. 1729–1738, June 2014.

[2] Loic Peter *et.al*, "Leveraging random forests for interactive exploration of large histological images," *MICCAI*, 2014.

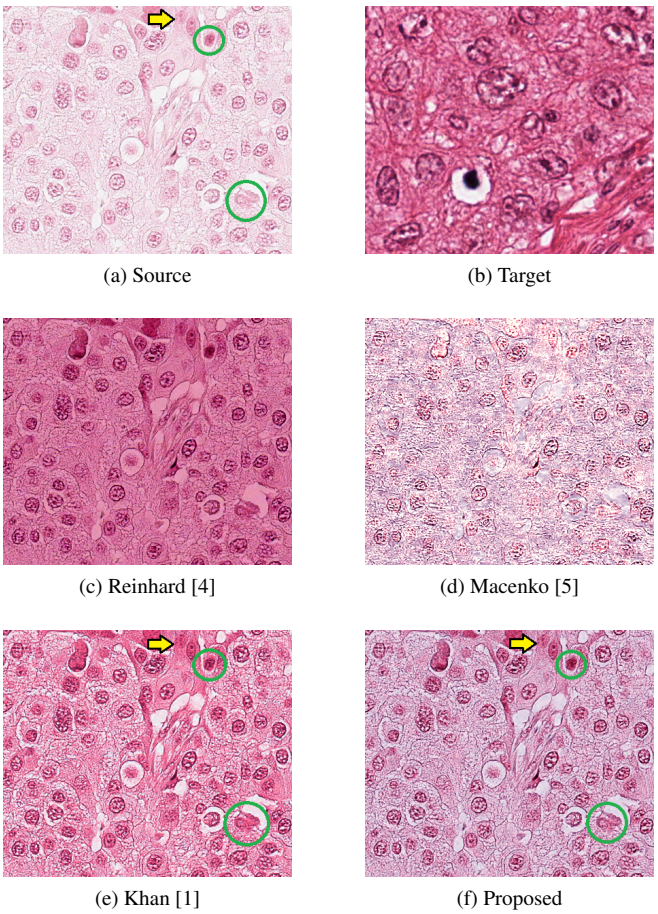
[3] Anil K. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1989.

[4] Erik Reinhard *et.al*, "Color transfer between images," *IEEE Computer graphics and applications*, vol. 21, no. 5, pp. 34–41, 2001.

[5] Marc Macenko *et.al*, "A method for normalizing histology slides for quantitative analysis.," in *ISBI*, 2009, vol. 9, pp. 1107–1110.

[6] Derek Magee *et.al*, "Colour normalisation in digital histopathology images," in *OPTIMHisE (MICCAI Workshop)*, 2009, pp. 20–24.

[7] Arnout C Ruifrok and Dennis A Johnston, "Quantification of histochemical staining by color deconvolution.," *AQCH*, vol. 23, no. 4, pp. 291–299, 2001.



(a) Source (b) Target (c) Reinhard [4] (d) Macenko [5] (e) Khan [1] (f) Proposed

Fig. 4. Comparison of different normalization techniques. Arrow and circle highlight important structures that are well preserved by our method. See text for detailed explanation.

[8] Milan Gavrilovic *et.al*, "Blind color decomposition of histological images.," *IEEE TMI*, vol. 32, no. 6, pp. 983–994, 2013.

[9] Daniel D Lee and H Sebastian Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.

[10] Andrew Rabinovich *et.al*, "Unsupervised color decomposition of histologically stained tissue samples," in *NIPS*, 2003.

[11] Julien Mairal *et.al*, "Online dictionary learning for sparse coding," in *ICML*. ACM, 2009, pp. 689–696.

[12] Ting Chen and Christophe Chefdhotel, "Deep learning based automatic immune cell detection for immunohistochemistry images," in *MLMI (MICCAI Workshop)*, pp. 17–24. Springer, 2014.