Stain Unmixing in Brightfield Multiplex Immunohistochemistry Images

Ting Chen and Chukka Srinivas

Ventana Medical Systems, Inc. A Member of the Roche Group, USA, 94043

Abstract. Multiplex immunohistochemistry (IHC) staining is a newly emerging technique for the detection of multiple biomarkers within a single tissue section and has become more popular due to its significant efficiency and the rich diagnostic information it contains. Therefore, to accurately unmix the IHC image and differentiate all the stains is of tremendous clinical importance since it is the initial key step in multiplex IHC image analysis in digital pathology. Due to the limitation of the CCD color camera, the acquired RGB image only contains three channels, and the unmixing of which into more than three colors is hence a challenging task. To the best of our knowledge, such a problem is barely studied in literature.

This paper presents a novel stain unmixing algorithm for brightfield multiplex IHC images based on a group sparsity model. The proposed framework achieves robust unmixing for more than three chromogenic dyes while preserving the biological constraints of the biomarkers. Typically, a number of biomarkers co-localize in the same cell parts. With this biological information known as a priori, the number of stains at one pixel therefore has a fixed up-bound, i.e. equivalent to the number of co-localized biomarkers. By leveraging the group sparsity model, the fractions of stain contributions from the co-localized biomarkers are explicitly modeled into one group to yield least square solution within the group. Sparse solution is obtained among the groups since idealy only one group of biomarkers are present at each pixel. The algorithm is evaluated on both synthetic and clinical data sets and demonstrates better unmixing results than the existing strategies.

1 Introduction

A multiplex immunohistochemistry (IHC) slide has the potential advantage of simultaneously identifying multiple biomarkers in one tissue section as opposed to single biomarker labeling in multiple slides. Therefore, it is often used for simultaneous assessment of multiple biomarkers in cancerous tissue. For example, tumors often contain infiltrates of immune cells, which may prevent the development of tumors or favor the outgrowth of tumors [1]. In this scenario, multiple biomarkers are used to target different types of immune cells and the population distribution of each type of them is used to study the clinical outcome of the patients. The biomarkers of the immune cells are stained by different chromogenic dyes. In order to conduct accurate detection and classification of the cells, the correct unmixing of the IHC digital image to its individual constituent dyes for each biomarker and obtaining the proportion of each dye in the color mixture is a prerequisite step for multiplex IHC image analysis.

Typically, a tissue slide is stained by the multiplex assay. The stained slide is then imaged using a CCD color camera mounted on a microscope or a scanner. The acquired RGB color image is a mixture of the underlying co-localized biomarker expressions. Several techniques have been proposed in the literature to decompose each pixel of the RGB image into a collection of constituent stains and the fractions of the contributions from each of them. Ruifrok *et al.* developed an unmixing method called color deconvolution [2] to unmix the RGB image with *up to three stains* in the converted optical density space. Given the reference color vectors $x_i \in R^3$ of the pure stains, the method assumes that each pixel of the color mixture $y \in R^3$ is a linear combination of the pure stain colors and solves a linear system to obtain the combination weights $b \in R^M$. The linear system is denoted as $y = Xb$, where $X = [x_1, \ldots, x_M], M \leq 3$ is the matrix of reference colors. This technique is currently most widely used in digital pathology domain, however, the maximum number of stains which can be resolved is limited to three, as the linear system is deficient for not having enough equations when there are more than three stains. A multilayer perceptron learning based technique has been proposed in [4] for three color brightfield image unmixing. In [3], Rabinovich *et al.* formulated the color unmixing problem into non-negative matrix factorization and proposed a system capable of performing the color decomposition in a fully automated manner, wherein no reference stain color selection is required. Again, these methods have the same limitation in dealing with large stain numbers due to solving $y = Xb$. *To the best of our knowledge, the method of unmixing brightfield IHC image with more than three stains is not available in literature.* In order to compare with the Ruifrok's method, we divide the color space into several systems with up to three colors in each system based on nearest color matching of each pixel to one of the systems. Ruifrok's method can therefore be used in solving each individual system. Due to the independent assignment of each pixel into different systems, the spatial continuity is lost in the unmixed images and artifacts such as holes are observed. However, this is the most straightforward modification of Ruifrok's method to work on more than three color multiplex brightfield image unmixing.

Alternatively, there exists another class of methods for multi-spectral image unmixing that works for a larger number of stain colors [5–9]. In fact, the multi-spectral image differs from the RGB image in terms of image acquisition. Multi-spectral imaging system is used to capture the image using a set of spectral narrow-band filters instead of the CCD color camera. The number of filters K can be as many as dozens or hundreds, leading to a mutli-channel image that provides much richer information than the bright field RGB image. The linear system constructed from it is always an over-determined system with X being a $K \times M (K \gg M)$ matrix that leads to a unique solution. However, the scanning process in the mutli-spectral imaging system is very time consuming and only a single field of view, manually selected by the technician, can be scanned instead of the whole slide, the usage of which is thus limited. As an example of the multi-spectral imaging unmixing, the two-stage methods [6, 7] are developed in the remote sensing domain to first learn the reference colors from the image context and then use them to unmix the image. More recently, a sparse model is proposed by Greer in [9] for high dimensional multi-spectral image unmixing. It adopts the L_0 norm to regularize the combination weights b of the reference colors hence leads to a solution that only a small number of reference colors are contributed to the stain color mixture. This serves as a valuable source of inspiration for selecting regularization terms for the linear system. However, the method proposed in [9] is also designed for multi-spectral image and no prior biological information about the biomarkers are used in that framework which may lead to undesired solution for real data.

In this paper, we propose a novel color unmixing algorithm for multiplex IHC image (scanned using CCD color camera) that can handle more than three stain colors and maintain the biological properties of the biomarkers. Intuitively, the unmixing algorithm for the multiplex IHC image should work as following. (1) Only one group of stains has non-zero contribution in the color mixture for each pixel. (2) Within that group, the fractions of the contributions from each constituent stain should be correctly estimated. These conditions motivate us to model the unmixing problem within the group sparsity [10] framework so as to ensure the sparsity among the group but non-sparsity within the group.

2 Methodology

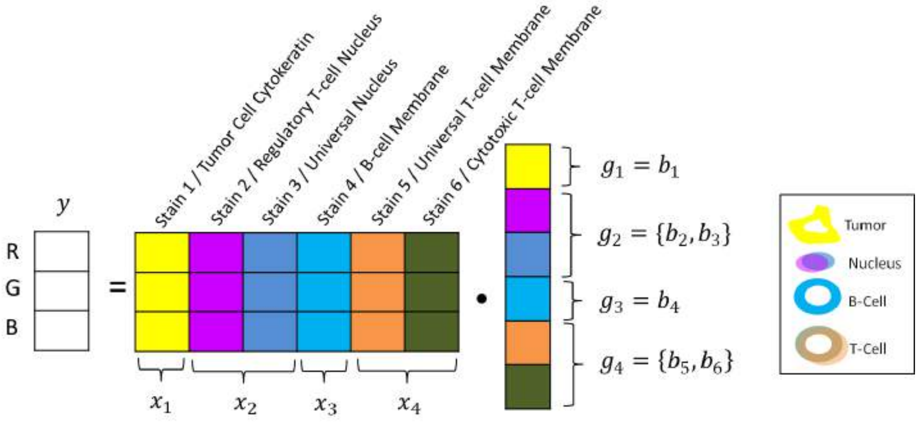
Stain 1 / Tumor Cell Cytokeratin
Stain 2 / Regulatory T-cell Nucleus
Stain 3 / Universal Nucleus
Stain 4 / B-cell Membrane
Stain 5 / Universal T-cell Membrane
Stain 6 / Cytotoxic T-cell Membrane
y
R
G
B
=
•
x_1
x_2
x_3
x_4
$g_1 = b_1$
$g_2 = \{b_2, b_3\}$
$g_3 = b_4$
$g_4 = \{b_5, b_6\}$
Tumor
Nucleus
B-Cell
T-Cell

Fig. 1: The group sparsity framework of the unmixing algorithm.

In this section, we present the methodology of our algorithm. We begin with illustrating the basic framework in Fig.1 using the following example. In the analysis of cancerous tissues, different biomarkers are specified to one or more types of immune cells. For instance, CD3 is a known universal marker for all T-cells and CD8 only stains the membranes of the cytotoxic T-cells. FoxP3 marks the regulatory T-cells in the nuclei and Hematoxylin (HTX) stains all the nuclei. Therefore, the co-localization information of the markers can be inferred from the biological knowledge, i.e. CD3 and CD8 co-locate in the membrane while FoxP3 and HTX may appear in the same nucleus. We can also have tumor marker on the tumor cell's cytoplasm region and B-cell marker on the B-cell's membrane. The framework of our proposed algorithm is shown in Fig.1 using the aforementioned immune cell example. Based on this biological co-localization information of the biomarkers, it is straightforward to conclude that only two colors can co-exist at each pixel for this case. The six chromogenic stains are therefore grouped into four different groups where co-localized stains are in the same group, as shown in the right panel of Fig.1.

2.1 Optical Density Transform

For the preprocessing, the RGB image I is converted into the optical density (OD) space using the following formula derived from Beer's law based on the fact that the optical density is proportional to the stain concentration.

$$O_c = -\log(\frac{I_c}{I_{0,c}}) \tag{1}$$

where c is the index of the RGB color channels, I_0 is the RGB value of the white points and O is the optical density image obtained. As in [2], O will be image to work with in the rest of the paper.

2.2 Group Sparsity Unmixing

We begin with illustrating the notations used in this paper. Let $\mathbf{y}$ be a pixel of O and it is a 3-dimensional column vector corresponding to the OD values converted from RGB. Assume there are M biomarkers available in the multiplex IHC slide. We have M stain colors. Let $\mathbf{b}$ be the combination weight vector of the stains and $b_m, m = 1, \ldots, M$ is the m_{th} element of $\mathbf{b}$. The typical unmixing problem thus is formulated as the following:

$$\min_{\mathbf{b}} ||\mathbf{y} - X\mathbf{b}||_2^2. \tag{2}$$

Each column of X corresponds to a reference stain color sampled from the control slide of pure stain. As we discussed before, this linear system has solution only when the column of X is less than or equal to 3 for $\mathbf{y} \in R^3$. Therefore, meaningful regularization is needed for the linear system to have a solution.

The biomarker co-localization information provides a partition of $\mathbf{b}$ into a set of groups $g_1, g_2, \ldots, g_N$, N being the total number of groups. Within each group, the biomarkers are known to have the co-localization possibility. We adopt this biological information to formulate the regularization term of the cost funciton. Let g_i be a q_i-dimensional column vector representing the combination weights of the stains within the i_{th} group and q_i be the number of stains within the group

g_i. We thus have $q_1 + q_2 \ldots + q_N = M$. x_i denotes the i_{th} group of reference colors, which is a $3 \times q_i$ matrix. Fig.1 shows an example of the stain group setting. Six stains are available in this example ($M = 6$). Two of them are co-localized membrane stains and two are co-localized nucleus stains. One is tumor cytokeratin stain and the rest is a membrane stain but only for B-cell. This information allows us to divide the stains into four groups ($N = 4$) as shown in Fig.1. For instance, g_2 contains b_2 and b_3 that are corresponding to the two co-located nucleus stains and x_2 contains the reference color vectors for all the stains within the 2_{nd} group. However, the 4_{th} stain of B-cell marker does not co-localize with other biomarkers, so g_3 only has one single member b_4 and x_3 is its reference color vector.

More specifically, the unmixing problem is formulated as the following convex optimization problem with the aforementioned notations:

$$\min_{\mathbf{b}} ||\mathbf{y} - \sum_{i=1}^{N} x_i g_i||_2^2 + \lambda \sum_{i=1}^{N} \sqrt{q_i} ||g_i||_2 \tag{3}$$

where $\mathbf{b} = [b_1, b_2, \ldots, b_M]^t = [g_1^t, g_2^t, \ldots, g_N^t]^t$ and $|| \cdot ||_2$ is the Euclidean norm with out squared. The first term in Eqn.3 solves for the linear system that is equivalent to [2], which minimize the least square error between the intensity of the raw image and the possible linear combination of the reference colors that approximates the raw image. λ is the regularization parameter that controls the amount of the group sparsity constraint in the second term. This model will act like LASSO at the group level. The entire groups will be dropped out when optimal $\mathbf{b}$ (or $\mathbf{g}$) is found, that is only a small number of g_i are non-zero.

Note that when the size of each group $q_i = 1$, the model becomes equivalent to lasso. In this case, no biological co-localization information is used in this model however the system remains to be solvable due to the sparsity constraints. The background noise is suppressed in this setting, comparing to the conventional Ruifrok's method. In the experiment section, we'll also demonstrate the efficacy of lasso unmixing by limiting the size of the group to 1.

Alternative direction method of multipliers (ADMM) algorithm [11] is used to solve Eqn.3. We implemented the algorithm in C++ to provide fast computation. It costs about 7 seconds to unmix a 750 by 1400 image on an Intel Core i7 1.87GHZ PC.

3 Experiments

3.1 Synthetic Data Experiment

As ground truth unmixing results are not available for real clinical data, we created a synthetic multiplex image from ground truth unmixed channels to validate our algorithm. We first synthetically generated six unmixed images as shown in the first row of Fig.2 C, following the stain co-localization and grouping rule in the example framework (Fig.1). The vectorized binary masks of the unmixed

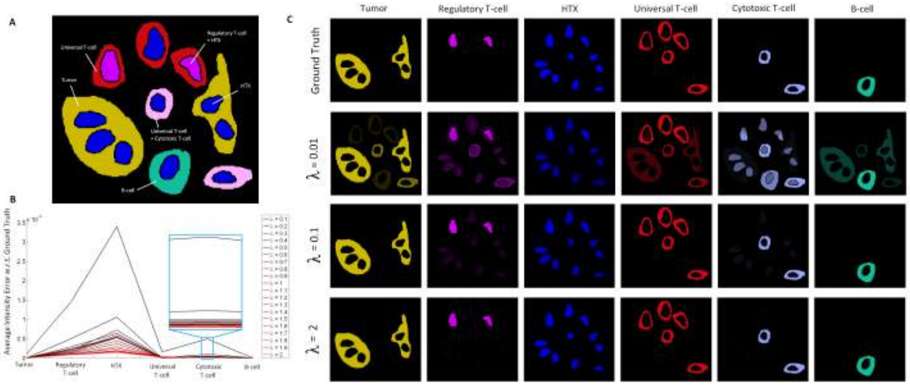

Fig. 2: Toy example. **A:** Image to be unmixed. **B:** The average intensity error of increasing λ for each channel. **C:** Unmixing results with different λ.

channels were multiplied by the reference color matrix to create the multiplex image in Fig.2 A. To demonstrate the algorithm performance w.r.t. the group sparsity regularization parameter λ variation, we plotted the average intensity error between the algorithm outputs and the ground truth unmixed channels in Fig.2 B for λ with in the range 0 to 2.The plot shows that the system has stable solutions when $\lambda > 0.3$. In Fig. 2 C, we also show the unmixing results for increasing λ. Note that when $\lambda = 0.01$, the system is close to deficient as in Eqn.2, hence unmixing errors are observed as shown in the second row of Fig.2 C.

3.2 Clinical Data Experiment

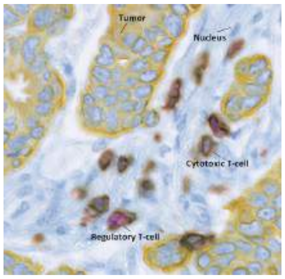

Fig. 3: Multiplexed tissue image real data example.

A clinical data set containing several different cancer tissue samples was used to demonstrate the proposed algorithm, including colorectal cancer, non small cell lung cancer and breast cancer that consist of 32 fields of view (FOV). The tissues were stained with the following assay as shown in Fig. 3: yellow chromogen for tumor cell cytokeratin, purple for regulatory T-cell nucleus, blue for universal nucleus, light blue for B-cell membrane, orange for universal T-cell membrane and dark green for cytotoxic T-cell membrane. Fig.5 shows the unmixing examples of decomposing the multiplexed image into single stain channels using modified Ruifrok's method based on nearest neighbor color assignment and the proposed group sparsity method. Note that λ is set to be 0.5 through the clinical data experiments. Pixel discontinuities, unmixing errors and artifacts are observed from the modified Ruifrok's method by solving multiple three color systems using the color similarity for system assignment. The proposed method instead solves one single system for all the pixels hence leads to a smoother unmixed images, meanwhile maintains the biological constraints as wells as reduces the background noises due to the group sparsity regularization.

Since the cytotoxic T-cell is a subset of the universal T-cell, the green cytotoxic T-cell membrane marker always co-localizes with the orange universal T-cell membrane marker, but the orange marker can present alone. Fig. 6 shows an example of the orange only cell and the green and orange co-localized cell. We can see that the aglorithm is able to handle both cases. This demonstrates that the L_2 norm constraint is used within the group to linearly separate the color mixture into different stain channels. Meanwhile, the modified Ruifrok's method is prone to unmixing errors due to the hard assignment of the unmixing system based on color similarity.

As a special case example, the algorithm can also be used for less than or equal to three color unmixing. When the group size becomes 1, the algorithm is equivalent to Ruifrok's unmixing plus a sparse constraint on the combination weights. The system can be solved by LASSO. We set the group size to 1 and compared to Ruifrok's method [2] for two-stain unmixing on a clinical breast cancer data set containing 217 FOVs. The proposed technique consistently shows better performance than Ruifrok's method. Example results are shown in Fig.4 and much less background noise is observed using the proposed sparse unmixing method.

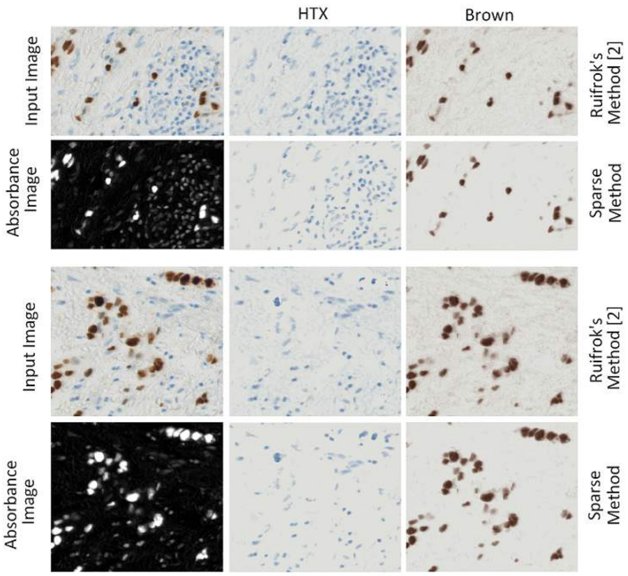

Fig. 4: Two-stain unmixing result comparisons when group size is 1.

4 Conclusion

In this paper, we introduce a novel color unmixing strategy for multiplexed bright field histopathology images based on a group sparsity model. The biological co-localization information of the biomarkers is explicitly defined in the regularization term to produce biologically meaningful unmixing results. The

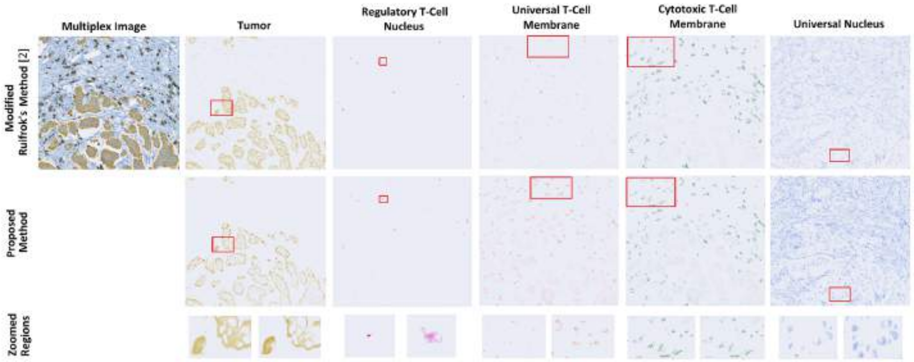
Multiplex Image
Tumor
Regulatory T-Cell Nucleus
Universal T-Cell Membrane
Cytotoxic T-Cell Membrane
Universal Nucleus
Modified Ruifrok's Method [2]
Proposed Method
Zoomed Regions

Fig. 5: Comparisons between the proposed group sparsity unmixing method and the modified Ruifrok's method based on nearest neighbor color assignment. More completed nuclei (purple and blue channels) are observed in group sparsity unmixing results. Incorrect universal T-cell unmixing is observed in modified Ruifrok's unmixing result due to the lack of co-localization constraint.

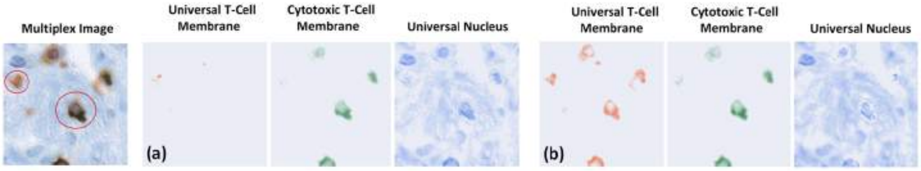
Multiplex Image
Universal T-Cell Membrane
Cytotoxic T-Cell Membrane
Universal Nucleus
(a)
Universal T-Cell Membrane
Cytotoxic T-Cell Membrane
Universal Nucleus
(b)

Fig. 6: Example unmixing of T-cell membrane co-localization case. **a:** The proposed group sparsity method without co-localization constraint (group size = 1). **b:** The proposed group sparsity method with co-localization constraint (group size = 2 for the two membrane stainings).

experiments of both synthetic and clinical data demonstrate the efficacy of the proposed algorithm in terms of accuracy and stability when compared to the existing techniques.

References

1. J. Galon, *et al.*, " Type, Density, and Location of Immune Cells Within Human Colorectal Tumors Predict Clinical Outcome", Science, 313(5795):1960-1964, 2006.
2. A. C. Ruifrok, *et al.* "Quantification of Histochemical Staining by Color Deconvolution", Anal. Quant. Cytol. Histol., 23:291-299, 2001.
3. A . Rabinovich, *et al.*, "Unsupervised Color Decomposition of Histologically Stained Tissue Samples", NIPS,2003.
4. C. Wemmert, *et al.*, "Stain Unmixing in Brightfield Multiplexed Immunohistochemistry", ICIP, 2013.
5. N. Kesheva, "A Survey of Spectral Unmixing Algorithms". Lincoln Laboratory Journal, v.14(1), pp. 55-78, 2003.
6. M. Zortea and A. Plaza, "Spatial Preprocessing for Endmember Extraction". IEEE Trans. Geosci. Remote Sens., vol.47(8), pp. 2679-2693, 2009.

7. G. Foody and D. Cox, “Sub-pixel Land Cover Composition Estimation Using A Linear Mixture Model and Fuzzy Membership Functions”. Int. J. Remote Sens., vol.15(3), pp. 619-631, 1994.
8. Z. Yang et.al, “Blind Spectral Unmixing Based on Sparse Nonnegative Matrix Factorization”. IEEE Trans. Image Proc., vol.20(4), pp.1112-1125, 2011.
9. J. B. Greer, “Sparse Demixing of Hyperspectral Images”. IEEE Trans Image Proc., vol. 21(1), pp.219-228, 2012.
10. N. Simon, *et al.*, “A Sparse Group Lasso”. Journal of Computational and Graphical Statistics, vol. 22(2), pp.231-245, 2013.
11. S. Boyd, *et al.*, “Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers”. Foundations and Trends in Machine Learning, vol. 3(1), pp.1-122, 2010.