A METHOD FOR NORMALIZING HISTOLOGY SLIDES FOR QUANTITATIVE ANALYSIS

Marc Macenko [1], *Marc Niethammer* [1,2], *J. S. Marron* [3,4], *David Borland* [5], *John T. Woosley* [6], *Xiaojun Guan* [5], *Charles Schmitt* [5], and *Nancy E. Thomas* [4,7]

Departments of [1]Computer Science,[2]Biomedical Research Imaging Center,
[3]Statistics and Operations Research, [4]Lineberger Comprehensive Cancer Center,
[5]Renaissance Computing Institute, [6]Pathology and Laboratory Medicine, [7]Dermatology
University of North Carolina, Chapel Hill, NC

ABSTRACT

Inconsistencies in the preparation of histology slides make it difficult to perform quantitative analysis on their results. In this paper we provide two mechanisms for overcoming many of the known inconsistencies in the staining process, thereby bringing slides that were processed or stored under very different conditions into a common, normalized space to enable improved quantitative analysis.

***Index Terms*—** Biomedical microscopy, Biomedical image processing, Biomedical signal detection, Melanoma, Pathology

1. INTRODUCTION

In many biological fields, tissue samples are taken from a subject for analysis. One common way of analyzing the tissue sample is to treat it with stains that have selective affinities for different biological substances. The majority of stains only absorb light, and the stained slides are therefore viewed using a microscope with a light illuminating the sample from below. If no stain is present, all of the light will pass through, appearing bright white. Areas where the stain has adhered to a substance in the tissue will absorb some of the light. The amount of light absorbed depends on many factors. For a given unit of stain, a certain amount of light in each spectrum will be absorbed. In the case of multispectral imaging, this process can be quite complicated. This paper focuses on the use of standard 24-bit RGB cameras to obtain images, so the methodology is restricted to those three wavelengths of light. The proportion of each wavelength absorbed forms the *stain vector*. The stain vector not only varies greatly among different stains but can also vary significantly for the same stain depending on such factors as the manufacturer, the storage conditions prior to use, and the method of application [1, 2]. These variations will be discussed further in sections 3 and 5.

The overall amount of light absorbed also varies between slides prepared differently. The two most prominent factors that affect the intensity of a slide are the relative amounts of stain added in the original treatment and the subsequent storage and handling of the slide, as stains can fade when exposed to light. This stain *intensity* is expounded on further in Section 4.

Figure 1 shows two examples of skin histology slides treated at different times in different laboratories. The absolute color values of a slide have many influences, only one of which is the biological component that we wish to retrieve. This biological component is the actual amount of the cellular substance to which a particular stain will attach. For example, in the most popular staining method for medical diagnosis, hematoxylin selectively stains nucleic acids a blue-purple hue while eosin stains proteins a bright pink color. Other variations result from staining compounds that do not absorb the exact same amounts of light, therefore exhibiting slightly different colors. This phenomenon is addressed in Section 3.

This research was funded by The University Cancer Research Fund, North Carolina

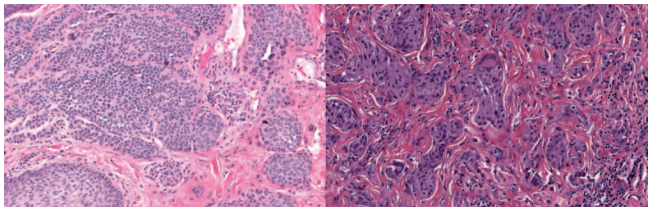


Fig. 1. Example of two histopathology slides of melanomas, both stained with hematoxylin and eosin, but with drastically different appearances. The images were obtained by scanning the slides at 20X using an Aperio Scanscope.

Following standard practice [3], all color values discussed in the paper are converted to their corresponding optical density (OD) values

$$OD = -log_{10}(I) \tag{1}$$

where I is the RGB color vector with each component normalized to $[0,1]$. This transformation provides a space where a linear combination of stains will result in a linear combination of OD values [3]. The relationship between intensity and OD is demonstrated in figures 2(a) and 2(b), using data acquired from images of hematoxylin and eosin stained melanoma slides.

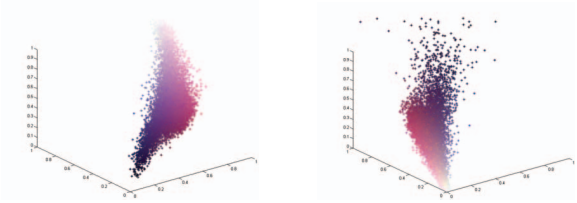
Once the correct vectors are determined by some method, a simple color deconvolution scheme similar to [4] is used to transform the color values into quantitative values of interest:

$$OD = VS \quad \Longrightarrow \quad S = V^{-1}OD \tag{2}$$

where OD is the optical density value observed and V and S are the matrices of the stain vectors and the saturations of each of the stains, respectively.

2. PREVIOUS WORK

Most histology slides are viewed in isolation by a skilled pathologist. These inspections focus on relative color differences and morphology of biological features. In this environment, there is seldom



(a) Note that the blue and the pink seem to be separable, but in a curved way.

(b) These pixels have been transformed to optical density(OD) space. Note that the blue and the pink are now separable via a straight line from the origin.

Fig. 2. Each point in 3-space corresponds to a randomly selected pixel from the image and is colored with its original color.

a need for the pathologist to directly compare different slides, and no need for quantitative analysis. However, with increased computational capabilities, there is a new focus on statistical analysis of various features for the purposes of differentiation, diagnosis, and prognosis. Analysis of large numbers of slides makes it imperative to automate the process as much as possible.

Most current applications concentrate on shape features and are thus not affected by the color irregularities between various slides [5] except when it interferes with segmentation on which the shape features are based. When color information is utilized, the easiest method is to simply use the raw color values obtained from the scanner. This approach adds some information, but differences in staining are not accounted for. A more appropriate method is to ascertain the actual quantities of the stains used.

In [4], a simple algorithm is given for obtaining stain saturation values when the stain vectors describing how the color is affected by the stain concentration are given. The remaining issue is determining which stain vectors should be used. Precalculated approximations for each type of stain exist but these approximations ignore variations between specific stains. The recommended method in the original paper requires manual selection of an area on the slide that contains only one stain, and then calculates an average stain vector from this area. In order to avoid contamination, it is recommended that a slide be stained with only one stain at a time. Obviously, individually staining a slide is not a viable option for slides that have already been processed or when only a scan of the slide is available. This leaves the inferior option of manually selecting an area with a minimal amount of other stains. Although this approach will lead to better results than using a precalculated approximation, it is very tedious for large datasets.

The need for individualized stain vectors for each slide is acknowledged in [6], but their method for determining the stain vectors does not use the fact that the stains are linear in the OD space instead of the RGB colorspace. Also, they use the peaks of the hue histograms to find the stain vectors which leads to nonoptimal solutions as discussed in Section 3.

In [7], non-negative matrix factorization (NMF) is proposed to solve the general color unmixing problem [8]. The work in [7] was motivated by the high levels of user interaction needed to determine the stain vectors. Although methods based on NMF are more general, there are no closed-form solutions for these problems, forcing the solutions to be computed numerically. The NMF-based algorithms attempt to factor the OD matrix into V and S with the constraint that all elements be non-negative, since stains cannot absorb a negative amount of light and there cannot be a negative amount of stain. It also forces all columns of V (the stain vectors in our examples) to be unit length to keep from having an infinite number of solutions that differ only by a constant. These methods work well with an abundance of information as in the multispectral images they used in their paper, but tend to have inconsistent convergence when only three spectral components are available. In [6], the data had to be manipulated to avoid ending with two purple stain vectors, even when initializing V individually for each image.

In Section 3, we present an algorithm that automatically finds the correct stain vectors for the image and then performs the color deconvolution. This is a fully automatic method that is suitable for rapid analysis of multiple slides. It is also a direct method with very few parameters and no optimizations needed.

3. STAIN VECTOR VARIATION AND CORRECTION

Slide preparation can vary widely due to different stain manufacturers, different staining procedures, and different storage times. We assume that there is a specific stain vector corresponding to each of the two stains present in the image, and that the resulting color (in OD space) of every pixel is a linear combination of these stain vectors. Since there has to be a non-negative weight on each component, every value must exist between the two stain vectors. This is why the following algorithm finds the fringe of the pixel distribution instead of searching for peaks. If noise were not a factor, the minimum and maximum along the found direction would be used. Instead, robust versions of the minimum and maximum are used by taking the α^{th} and $(100-\alpha)^{th}$ percentile. Empirically, $\alpha = 1$ provides robust results.

In this section, an algorithm is presented to find these particular stain vectors for each image based on the colors that are present. An OD value of 0 corresponds to a pixel that was all white where nothing on the slide absorbed any light. For stability reasons, the pixels with nearly no stain (low OD) were thresholded. After much empirical analysis, a threshold value of $\beta = 0.15$ was found to provide the most robust results while removing as little data as possible. Acceptable results are achieved for a wide range of both α and β .

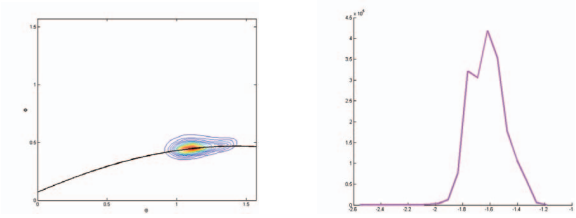
The shortest path between two unit-norm color vectors on the sphere is the geodesic path. This line appears to be curved in a spherical coordinate decomposition unless it would correspond to change in only one direction or the other. By finding this specific geodesic direction, we can project the OD transformed pixels onto it in order to find the endpoints that correspond to the stain vectors.

The first step in this process is to calculate the plane that the vectors form. This is done by forming a plane from the two vectors corresponding to the two largest singular values of the SVD decomposition of the OD transformed pixels. All of these OD transformed pixels are then projected onto this plane, and subsequently normalized to unit length. The projection line is shown to be curved in Figure 3(a). The angle with respect to the first SVD direction is calculated for each point, thus mapping the directions in the plane to a scalar. The histogram of these angles is shown in Figure 3(b). This whole process is outlined in Algorithm 1.

This method was performed on twelve different slides with considerable variation. Before the use of this method, standard vectors were computed using the manual methods described in Section 2, attempting to find vectors that could adequately describe all twelve slides. The results of these computations, along with the standard vectors, are shown in Figure 4 where the color of each symbol corresponds to what would be produced by that vector. The stars are the standard vectors used without regard to the specific slide. The circles

Input: RGB Slide
1 Convert RGB to OD
2 Remove data with OD intensity less than β
3 Calculate SVD on the OD tuples
4 Create plane from the SVD directions corresponding to the two largest singular values
5 Project data onto the plane, and normalize to unit length
6 Calculate angle of each point wrt the first SVD direction
7 Find robust extremes (α^{th} and $(100-\alpha)^{th}$ percentiles) of the angle
8 Convert extreme values back to OD space
Output: Optimal Stain Vectors

Algorithm 1: SVD-geodesic method for obtaining stain vectors.



(a) Contours of the pixel histogram. The ridge line is calculated by singular value decomposition (SVD). This results in a geodesic on the sphere and is overlaid in black.

(b) This is the histogram of angles that the points form with the geodesic line shown in Figure 3(a). The color corresponds to the pixel color that would contribute to that bin.

Fig. 3. An example of the data analysis using the SVD geodesic method.

are the stains automatically generated with the proposed method. Notice that all the recovered stain vectors form tight clusters. While not completely coinciding with the stain vectors chosen manually (which is expected, since they were chosen as a compromise to represent all twelve slides simultaneously), stain vectors are similar, yet distinct for each of the slides, showing the need for repeatable automatic methods.

Figure 5 shows the results of deconvolving with the automatically determined stain vectors. The top left image is the original image. The two bottom images show a good separation into the two stains. The top right image shows the values orthogonal to the plane created by the two stain vectors and includes pigmentation and noise. The fact that this image is nearly empty is evidence that the stain vectors were well chosen.

4. INTENSITY VARIATION AND CORRECTION

The intensity of a particular stain depends on the original strength of the stain, the staining procedure, how much fading has occurred since the sample was originally processed, and finally how much of the cellular substance of interest is present in the material. The last quantity is what we actually want to measure. Removing the confounding factors that degrade the signal is necessary for direct analytical analysis of these samples.

Our method assumes that the amount of protein or nucleic acid is a random variable that is scaled by the confounding factors mentioned previously. For each stain in question, we calculate the intensity histograms for all pixels that have a majority of that stain. We

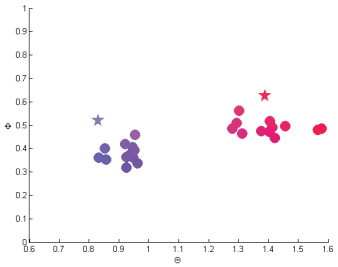


Fig. 4. This shows the calculated stain vectors for twelve hematoxylin and eosin stained test slides. The color of each symbol corresponds to what would be produced by that vector. The stars are the standard vectors used without regard to the specific slide. The circles are the automatically computed stain vectors. Notice that all recovered vectors are significantly different from the standard vectors.

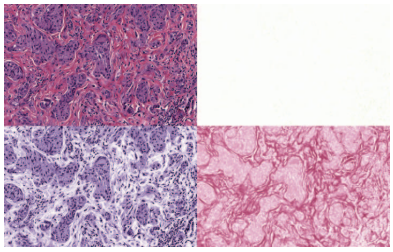


Fig. 5. This example shows the results of deconvolving with the automatically determined stain vectors. The top left image is the original image. The two bottom images show a good separation into the two stains. The top right image shows the values orthogonal to the plane created by the two stain vectors and includes pigmentation and noise. The fact that this image is nearly empty is evidence that the stain vectors were well chosen.

then find the 99^{th} percentile of these intensity values and use this as a robust approximation of the maximum. This value was shown experimentally to be a good simple descriptor of the histogram by analyzing several patches of each slide. All intensity histograms are then scaled to have the same pseudo-maximum and are then able to be compared with each other.

As can be seen from Figure 6, each stain has its own distribution, which is to be expected from the previously described theory. We currently assume that fading affects each stain identically, and that the only difference is the amount of stain added. We are currently studying the effects of fading and investigating whether this is a correct assumption.

5. RESULTS

Figure 7 shows an example where the images from Figure 1 were transformed into the same colorspace by the method discussed in Section 3. Visually, they appear quite different, but to the stain quantization algorithms and subsequent statistical analysis, there is a strong beneficial effect. Figure 8 shows the images from Figure 7, but both are now at the same average intensity level by using the method discussed in Section 4. This correction also leads to a much improved visual consistency of the two slides.

Analysis of five slides diagnosed with melanoma and seven

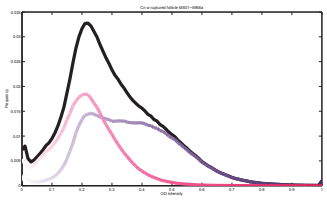


Fig. 6. Histogram of pixel saturation values. The black line includes all values while the two colored histograms correspond to the two stains. The saturation of the color corresponds to the saturation of the pixels that corresponded to that histogram bin.

slides containing benign nevi (common moles) was performed using a variety of shape and stain-based features. The slides had all been stained with hematoxylin and eosin and scanned at 20X. For each slide, a large number of nuclei are segmented and features calculated for each of them. The statistical method known as Distance Weighted Discrimination (DWD) [9] was used to find the optimal separation direction between melanoma and nevi based on this feature-space. Figure 9(a) shows a graphical representation of where the values from individual slides without correction are located on this DWD direction. There is clearly a difference between the two groups (melanoma is colored red; nevi are colored blue), but there is a large overlap. Figure 9(b) shows the same results after being corrected with the methods described in sections 3 and 4, and the increased separation of the groups is evident.

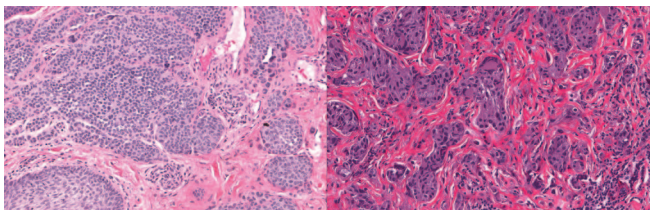


Fig. 7. Same images as Figure 1, but the second slide has been transformed into the same colorspace as the first slide by the method discussed in Section 3.

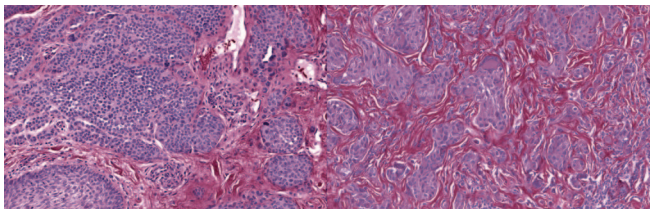
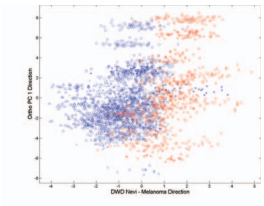


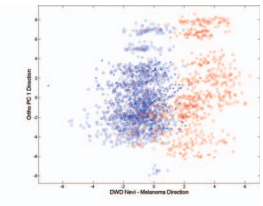
Fig. 8. Same images as Figure 7, but both are now at the same average intensity level by the method discussed in Section 4.

6. CONCLUSIONS

While the work in this paper has been performed using hematoxylin and eosin stained slides of melanomas and nevi, it is applicable to other histologic stains and tissues. The algorithm for obtaining the optimal stain vectors has been evaluated on slides with various stain



(a) Before correction. Note that while there is separation between the two classes, there is still a large degree of overlap.



(b) After correction with the method discussed in Section 3. Note the much improved separation

Fig. 9. Distance Weighted Discrimination (DWD) is utilized for cell-by-cell discrimination of melanomas from benign nevi (moles). Red and blue symbols are melanoma and nevus cells, respectively.

combinations satisfactorily. When three or more stains are present in a slide, results are sometimes inconsistent. The methods in this paper have greatly improved the ability to quantitatively analyze histology slides and have improved the results of our investigations. Automating the process can accommodate larger datasets and enable a level of reproducibility not guaranteed with manual selection methods. The methods presented are easy to implement, and computation time is much improved over the NMF methods. It is hoped that these methods will facilitate additional research into medical aspects that use stained histology slides for diagnosis, prognosis or basic research. So far, we have only used these methods on small datasets, but have a much larger dataset prepared for testing.

7. REFERENCES

[1] K Glatz-Krieger, U Spornitz, A Spatz, M Mihatsch, and D Glatz, "Factors to keep in mind when introducing virtual microscopy," in *Virchows Archiv*, 2006, vol. 448, pp. 248–255.

[2] A Ljungberg and O Johansson, "Methodological aspects on immunohistochemistry in dermatology with special reference to neuronal markers," in *The Histochemical J.*, 1993, vol. 25, pp. 735–745.

[3] JK Fawcett and JE Scott, "A rapid and precise method for the determination of urea," in *J. of Clin. Path.*, 1960, vol. 13, pp. 156–159.

[4] AC Ruifrok and DA Johnston, "Quantification of histochemical staining by color deconvolution," in *Anal. and Quant. Cytology and Histology*, August 2001, vol. 23, pp. 291–299.

[5] A Viros, J Fridlyand, J Bauer, K Lasithiotakis, C Garbe, D Pinkel, and BC Bastian, "Improving melanoma classification by integrating genetic and morphologic features," in *PLoS Medicine*, 2008, vol. 5, pp. 941–952.

[6] J Newberg and R Murphy, "A framework for the automated analysis of subcellular patterns in human protein atlas images," in *J. of Proteome Research*, 2008, pp. 2300–2308.

[7] A Rabinovich, S Agarwal, CA Laris, JH Price, and S Belongie, "Unsupervised color decomposition of histologically stained tissue samples," in *Adv. in Neural Inf. Proc. Systems*, 2003.

[8] R Levenson, "Spectral imaging and pathology: Seeing more," in *Lab. Med.*, 2004, vol. 35, pp. 244–251.

[9] JS Marron, MJ Todd, and J Ahn, "Distance weighted discrimination," in *J.of the Am. Statistical Assoc.*, 2007, vol. 102, pp. 1267–1271.