A Survey on Color Normalization Approach to Histopathology Images

Hasna Panikkaveettil K., Beena M.V.

Department of Computer Science & Engineering, Vidya Academy of Science &Technology, Thrissur, Kerala, India

Abstract— Automated image processing has become very popular in the field of digital histopathology. Digital histopathology is an upcoming area of research where diagnosis decisions are made by quantitative analysis of image content. The problem of color variation is a serious issue in this field. Color variation is a result of inconsistencies in staining procedure or due to the use of different scanners. Color normalization aims at bringing a histological image into a different color appearance of another image which standardizes the color representation of both the images. This literature gives a description of various color normalization approaches used to address the problem of color variation in histopathology images. It focuses mainly on methods used in normalization of breast cancer images.

Keywords—Quantitative analysis, color normalization.

I. INTRODUCTION

Color normalization is a topic in computer vision concerned with artificial color vision and object recognition. The distribution of color values in an image depends on illumination which varies depending on lighting conditions, cameras and other factors. Color normalization has been used for object recognition on color images in the field of robotics, bioinformatics and artificial intelligence. It is important to remove all intensity values from the image while ensuring preservation of color values. This avoids loss of information obtained from source image. When color normalization is applied to images, minimum information should be guaranteed.

II. OVERVIEW

When preparing histopathology slides, multiple dyes are used for staining. The use of multiple dyes may result in color variation. It may be caused due to different scanners used while digitizing the image. It may result in loss of histopathological information like tissue texture. In order to address all these problems of color variation, color normalization can be used. Images prepared are subjected to varying conditions such as illumination, stain concentration etc. All these factors make color normalization very challenging. It must be ensured that there is no loss of histological information.

III. QUANTITATIVE IMAGE ANALYSIS

Quantitative image analysis involves using information within a scientific image to extract data. Example of quantitative information includes number of cells in an image. Manual analysis is very slow. This problem can be overcome by automating image analysis. In [1] the steps involved in quantitative image analysis are discussed. Six steps are involved in quantitative analysis namely:

- Sample preparation and image acquisition
- Preprocessing of acquired images
- Segmentation of preprocessed images
- Feature extraction based on segmented images
- Validation of feature extraction
- Data analysis

The author has also presented case studies regarding quantification of buds, peroxisomes, lymph nodes etc.

In [2] the authors have presented two mechanisms for overcoming many of the known inconsistencies in the staining process namely: stain vector variation and correction intensity and variation and correction.

Stain vector variation and correction:

Slide preparation can vary widely due to different stain manufacturers, different staining procedures, and difference in storage times. The author assumes 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 staining vectors. Since there has to be a non-negative weight on each component, every value must exist between the two stain vectors.

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 how much of the cellular substance of interest is present in the material as well. The last quantity is measured. Removing the confounding factors that

degrade the signal is necessary for direct analytical analysis of these samples.

Here the authors performed tests on melanoma sides. The results of the tests were accurate. There were some inconsistencies when three or more stains were present in a single slide.

IV. VARIOUSCOLOR NAORMALIZATION METHODS

4.1 Histogram Based Method

The early normalization techniques were based on histogram matching in RGB color space. After the preprocessing steps, matching is performed separately in red, blue and green channels respectively. In [3] and [4] the authors have presented application of such a normalization scheme. As local differences of image content are ignored in this method, matching of irrelevant stains may take place. This will result in loss of histological information.

4.2 Color Transfer

The author has presented color transfer technique in [5]. Images stained with multiple dyes will have different color distributions. The image is divided by manual segmentation or by automatic segmentation to avoid mixing of color components. The mean and variance in each channel of a query image are matched with that of a reference image. The major drawback involved in this method is that, the tissue details may not be preserved after normalization.It is overcome in the methods mentioned later in this literature.

4.3 Spectral Matching

Spectral matching aims at removing stain variation. In this method stain spectra is evaluated either by using adaptive estimators as in [2] or by using dedicated hardware as in [6]. It estimates stain spectra and corresponding stain proportions at each pixel in a histopathology image. This method has an advantage over other methods as it can preserve histological information. It can be applied in situations where stain variation is the only reason behind color mismatch. The regions of interest for estimating stain spectra are selected manually. It makes this method a time consuming one. There are many other methods to find the stain spectra other than the above mentioned methods. This is applied in blind color decomposition as in [7]. It is obtained by performing expectation – maximization on color distributions in the Maxwell color triangle. Stain estimation may not be precise when spectral variation in stains used occurs between a query image and training image.

The major drawbacks involved are that the method fails if the stain absorption is poor or if a stain is not present. The specimen preparation andimaging artifacts may result in variations in the estimation.

4.4 Color Deconvolution

Colour deconvolution is a method to obtain stain concentration values when the stain matrix, describing how the colour is affected by the stain concentration, is given. Instead of depending on standard stain matrices, which may be inappropriate for a given image, a colour based classifier that incorporates a novel stain colour descriptor to calculate image-specific stain matrix. This method accurately separates up to three stains. In [8] the author presents a stain normalization technique based on colour deconvolution. A stain decomposition algorithm consisting of four modules: stain matrix estimation, CD, nonlinear mapping of channel statistics, and reconstruction is also presented.

Compared to other methods the color deconvolution method performs better in many aspects. It involves accurate stain separation which is a major advantage. Color deconvolution separates out the effect of variation of each stain so it can be corrected independently.

V. CONCLUSION

This literature presented several color normalization approaches applied to histopathology images. In histogram matching method an unwanted bias is caused due to mismatches. In color transfer technique, segmentation of image is performed before color transfer making color transfer technique a time consuming process. Stain variation can be removed by spectral matching. Its success depends on accuracy of stain spectral estimation. It is advantageous over the other methods as it enables normalization even if only stain variation is the reason behind color variation. It can be understood from the literature that color normalization approach helps in handling problems due to color variation in histopathology images. It is a very useful concept in medical field.

REFERENCES

[1] AnttiNiemisto, “Quantitative Image Analysis Methods for Applications in Biomedical Microscopy,” *TampereenTeknillinenYiliopsto-Tampere University of Technology*, 2006.

[2] *Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 6, pp. 713–724, Jun. 2003.

[3] Marc Macenko , Marc Niethammer , J. S. Marron , David Borland , John T. Woosley

[4] XiaojunGuan , Charles Schmitt , and Nancy E. Thomas , “A Method For Normalizing Histology Slides For Quantitative Analysis,” *in Proc. ISBI,2009.*

[5] S. Kothari et al., "Automatic batch-invariant color segmentation of histological cancer images," in *Proc. ISBI, 2011*.

[6] A. Tabesh et al., "Multifeature prostate cancer diagnosis and gleason grading of histological images," *IEEE Trans. Med. Imag., vol. 26, no. 10, pp. 1366 – 1378*, Oct. 2007.

[7] E. Reinhard et al., "Color transfer between images," *IEEE Comput. Graph.Appl.*, vol. 21, no. 5, pp. 34 – 41, 2001.

[8] T. Otsuka et al., "Color standardization," in *Pathology Informatics*, 2010, [Session #4 Color and Compression , Scientific Session].

[9] *M. Gavrilovic et al., "Blind color decomposition of histological images,"* IEEE Trans. Med. Imag., vol. 32, no. 6, pp. 983 – 994, Jun. 2013.

[10] A.M. Khan et al., "A non-linear mapping approach to stain normalization in digital histopathology images using image-specic color deconvolution," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 6, pp. 1729 – 1738, Jun. 2014.

[11] *A. Koschan and M. Abidi, Digital Color Image Processing, Wiley*, 2008.

[12] A. Hanbury, "Circular statistics applied to color images," in *8th Computer Vision Winter Workshop*, 2003, vol. 91, pp. 53 – 71.