UTILITY OF MULTISPECTRAL IMAGING FOR ANALYSIS OF ROUTINE CLINICAL HISTOPATHOLOGY IMAGERY

Laura E. Boucheron[†§*], *Neal R. Harvey*[†], *B. S. Manjunath*[§]

[†]Los Alamos National Laboratory
Space and Remote Sensing Sciences
Mail Stop B244, Los Alamos, NM 87545

[§]University of California Santa Barbara
Electrical and Computer Engineering
Santa Barbara, CA 93106-9560

ABSTRACT

Our paper will present an analysis of the utility of multispectral imagery versus standard RGB imagery for routine H&E-stained histopathology imagery, in particular for the classification of histologic classes, with a focus on nuclei detection. Our multispectral data consists of 29 spectral bands, spaced 10 nm within the visual range of 420-700 nm. It is hypothesized that the additional spectra contains further information useful for classification as compared to the 3 standard bands of RGB microscopy imagery; this has been established in other application domains, e.g., remote sensing. We will present analyses of our data designed to test this hypothesis. In brief, we will use several standard classification techniques (maximum likelihood, spectral angle mapper, minimum Euclidean distance, and an automated feature extraction tool using evolutionary computation) on our multispectral image stacks, as well as on several types of derived RGB imagery. We will analyze the classification performance, as well as other applicable and informative metrics, such as the statistics of spectral bands chosen for classification and the entropy of different spectral bands within the class of interest.

1. INTRODUCTION

Our datasets are relatively unique in histopathology in that they are imaged multispectrally using the VariSpec™ (CRi, Woburn, MA), a liquid crystal tunable filter coupled with image acquisition software. While some researchers have found advantages to spectral imaging for some applications in medical image analysis [1–4], the value of multispectral analysis for routine histopathology preparations is still unknown. This work is motivated in particular by the demonstration of the slight superiority of similar multispectral imagery and classification tasks shown in [1, 5], as well as our own preliminary results that have indicated again a slight superiority of multispectral data for a nuclei classification task [6].

*L. E. Boucheron is a Ph.D. student at UC Santa Barbara, where her first two years were funded by NSF IGERT Grant DGE-0221713. Her thesis research is funded by Los Alamos National Laboratory.

2. DATA

The dataset that we analyze consists of 58 H&E-stained histopathology images of breast tissue from the Yale Tissue Microarray Facility (inventory at `http://tissuearray.org/facility/inventory/list.php`). Our 58 images are not microarray images in the general sense since we are dealing with single histopathology images as might be obtained from standard clinical biopsy specimens. The data was captured from 5 microarrays, with an average of 3 images per case. The multispectral images have 29 bands, spaced 10 nm apart, ranging from 420 to 700 nm. Each band is represented in an image stack as an 8 bit, 768×896 grayscale image.

To compare classification performance of multispectral imagery versus standard RGB imagery, we derive three kinds of RGB imagery directly from the multispectral image stacks:

1. **rgbequal:** created by (approximately) equally allocating the 29 bands to R, G, and B, similar to the approach in [4], reflecting a rough approximation of the three spectral ranges associated with the three colors red, green, and blue, albeit with some ambiguity in allocation of intermediate colors (e.g., yellow).

2. **truecolor:** created by converting the illumination wavelength for each band into the constituent RGB values as perceived by humans, then averaging the contribution to R, G, and B for each band. This method utilizes the MatlabCentral (`http://www.mathworks.com/matlabcentral/fileexchange/`) function *spectrumRGB* .

3. **ccd:** a modification of truecolor imagery to better match the spectral response of common 3-CCD color cameras used in microscopy setups for biomedical research. This method also utilizes the *spectrumRGB* function.

3. ANALYSIS

3.1. Classification Performance

In previous work, based on empirical observations of the relative contrast of nuclei within various spectral bands, we developed a nuclear classification algorithm which thresholds the 22nd multispectral band (640 nm). Comparing the segmentation performance of this method using band 22 versus the red channel of the derived RGB images, we found that there was a slight increase in performance using the 22nd multispectral band (9.4% classification error), compared with 12.3%, 12.1%, and 11.0% error using the red band of the rgbequal, truecolor, and ccd RGB imagery, respectively. Additionally, using automated feature extraction on two training images, bands 21 and 22 (630 and 640 nm) yielded the best performance (4.8% error), of any single multispectral band.

We will show in-depth results for classification of nuclei over our entire dataset using several common classification methods, including maximum likelihood, spectral angle mapper, minimum Euclidean distance, and an automated feature extraction tool based on evolutionary computation [7], for our multispectral and RGB data, using the entire image stacks as well as individual spectral bands.

3.2. Spectral Bands Used in Classification

Another method that provides insight into the utility of certain spectral bands is to statistically analyze which bands are more often chosen for classification. We have found, using an automated feature extraction tool for classification and the two training images mentioned above, a statistical preference for bands 21 and 22. Similarly, if limited to a choice of a single input band, bands 21 and 22 are the only ones chosen. These results seem to correlate with these same bands yielding the least classification error. We will demonstrate similar results over our entire dataset for all proposed classifiers.

3.3. Entropy of Nuclear Regions

Based on the results discussed previously, we will investigate the entropy of nuclear regions in our various data. The average maximum entropy of the multispectral images is 7.30 bits/pixel, with bands 21-26 (630-680 nm) most likely to be the maximum entropy bands. In comparison, the average maximum entropy of the derived RGB data is 7.14 (rgbequal), 6.96 (truecolor), and 7.05 (ccd), where the maximum entropy band is almost always the red channel. Again, note the apparent correlation between least classification error and maximum entropy for single bands. Note also that poorer performance of the derived RGB imagery may correlate with a slightly smaller entropy for the corresponding nuclear regions. We will demonstrate relationships between classification performance, spectral bands used in classification, and the entropy of the bands for our nuclear classification task.

4. CONCLUSIONS

In this work we seek to test the hypothesis that multispectral data contains more useful information than standard RGB, enabling better classification performance. Our particular application is H&E-stained histopathology images of breast biopsies, but our analysis methods should be applicable to other multispectral imagery applications. Our use of histopathology imagery is motivated by the recent use of multispectral imaging in the field, using liquid crystal tunable filters, along with the lack of comprehensive understanding of the utility of such multispectral imagery for automating routine clinical pathology analysis. To focus our efforts, we will concentrate on a nuclei classification task, with the possibility to extend this analysis to further histologic classes (e.g., cytoplasm and connective tissue).

5. ACKNOWLEDGMENTS

Many thanks to the Department of Pathology at Yale University, particularly Carola Zalles, Greg Tedeschi, and David Rimm, for the images used in this study.

6. REFERENCES

[1] R. Levenson, P. J. Cronin, and K. K. Pankratov, "Spectral imaging for brightfield microscopy," in *Proc SPIE*, 2003, vol. 4959, pp. 27–33.

[2] R. Jaganath, C. Angeletti, R. Levenson, and D. L. Rimm, "Diagnostic classification of urothelial cells in urine cytology specimens using exclusively spectral information," *Cancer Cytopathol*, vol. 102, no. 3, pp. 186–191, June 2004.

[3] C. Angeletti, N. R. Harvey, V. Khomitch, A. H. Fischer, R. M. Levenson, and D. L. Rimm, "Detection of malignancy in cytology specimens using spectral-spatial analysis," *Lab Invest*, vol. 85, pp. 1555–1564, 2005.

[4] S. M. Gentry and R. Levenson, "Biomedical applications of the information-efficient spectral imaging sensor (ISIS)," in *Proc SPIE*, 1999, vol. 3603, pp. 129–142.

[5] M. A. Roula, A. Bouridane, F. Kurugollu, and A. Amira, "A quadratic classifier based on multispectral texture features for prostate cancer diagnosis," in *Proc Int Symp on Signal Processing and Its Applications*, July 2003, vol. 2, pp. 37–40.

[6] L. E. Boucheron, "Interactive object-level quantitative analysis of multispectral histo- and cyto-pathology images for detection and characterization of cancer," Ph.D. Dissertation Proposal, Dept. of ECE, UCSB, June 2006.

[7] N. R. Harvey, J. Theiler, S. P. Brumby, S. Perkins, J. J. Szymanski, J. J. Bloch, R. B. Porter, M. Galassi, and A. C. Young, "Comparison of GENIE and conventional supervised classifiers for multispectral image feature extraction," *IEEE T Geosci Remote*, vol. 40, pp. 393–404, 2002.