Hyper-Spectral Analysis of Normal and Malignant Colon Tissue Microarray Sections Using a Novel DMD System

G L Davis[2,3], M Maggioni[1,3], F J Warner[1,3,4], F B Geshwind[4], A C Coppi[4], RA DeVerse[4], R R Coifman[1,3,4]. Departments of [1]Mathematics & [2]Pathology, [3]Program in Applied Mathematics; Yale University, New Haven, CT and [4]Plain Sight Systems, Hamden, CT.

DESIGN: We have developed and applied a unique tuned light source based on a Digital Mirror Device (DMD) (Plain Sight Systems Inc.) which transmits any combination of light frequencies, range 450 nm – 850 nm, transilluminating H&E stained micro-array tissue sections of normal and malignant colon through a Nikon Biophot microscope. Hyper-spectral pictures of tissues obtained with a CCD camera (Sensovation) are captured by a computer & analyzed mathematically to discriminate between normal & abnormal (carcinoma) tissues.
147 hyper-spectral pictures are collected at 400X magnification: 76 pictures of normal colon tissue and 71 pictures of abnormal malignant colon tissue from 130 different patients. The spectra of each pixel are normalized & analyzed to discriminate between gland nuclei, gland cytoplasm and lamina propria/lumens. Pixel spectra are automatically extracted & classified as nuclei using local discriminant bases and nearest neighbor classification; after denoising and normalization, spectral features separating normal nuclei from abnormal nuclei are found by using partial least squares regression.
RESULTS: Spectral analysis discriminated between single normal & abnormal nuclei (adenoma & carcinoma) with diagnostic efficiency of 93.6%., and between normal & abnormal tissue samples with a diagnostic efficiency of 100.0%.
CONCLUSION: Hyperspectral microscopic analysis does discriminate normal and abnormal populations. We are attempting to refine the feature extraction algorithms to permit the discrimination of benign (adenoma) and malignant (carcinoma) populations for diagnostic purposes.

Funding: National Science Foundation: NSF DMS-01399; Defense Advanced Research Projects Agency (DARPA)
Correspondence: Dr. G.L. Davis, Yale Pathology, 310 Cedar St., New Haven CT 06520. E-mail: gustave.davis@yale.edu
Web: www.math.yale.edu/~mmm82/

INTRODUCTION

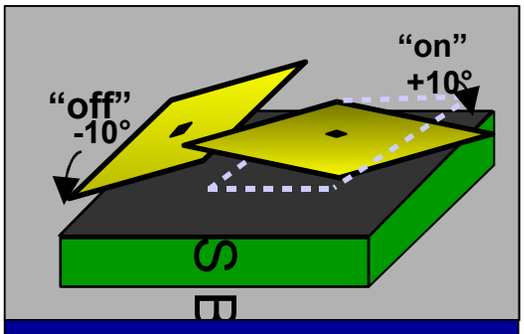
With light sources of increasingly broader ranges, spectral analysis of tissue sections has evolved from 2 wavelength image subtraction techniques to hyperspectral mapping. A variety of proprietary spectral splitting devices[6,9,13], including prisms and mirror, interferometer, variable interference filter-based monochromometer & tuned liquid crystals, mounted on microscopes in combination with CCD cameras and computers, have been used to discriminate among cell types[2,7,8,9,13] & endogenous & exogenous pigments[1,13] .

GOALS

We use a prototype unique tuned light source, a digital mirror array device (Plain Sight Systems) based on micro-optoelectromechanical systems[5], in combination with analytic algorithms developed in the Yale Program in Applied Mathematics[2,3,11], to evaluate the diagnostic efficiency of hyperspectral microscopic analysis of normal & neoplastic colon biopsies prepared as microarray tissue sections[14].
We compare the results to our previous spectral analysis of colon tissues[10] and to other spectral studies of tissues and cells.

EXPERIMENTAL DETAILS

Platform: The prototype tuned light digital mirror array device[5] (Figure 1) trans illuminates H & E stained micro-array tissue sections with any combination of light frequencies, range 440 nm – 700 nm, through a Nikon Biophot microscope. Hyper-spectral tissue images, multiplexed with a CCD camera (Sensovation), are captured & analyzed mathematically with a PC.



Mirrors rotate along their diagonal axis by exactly ± 10°, which is what makes the DMA a digital devise

A complete DMA contains 848 columns and 600 rows of mirrors and measures 10.2 mm x 13.6 mm. Here, a DMA is shown with its glass cover removed.

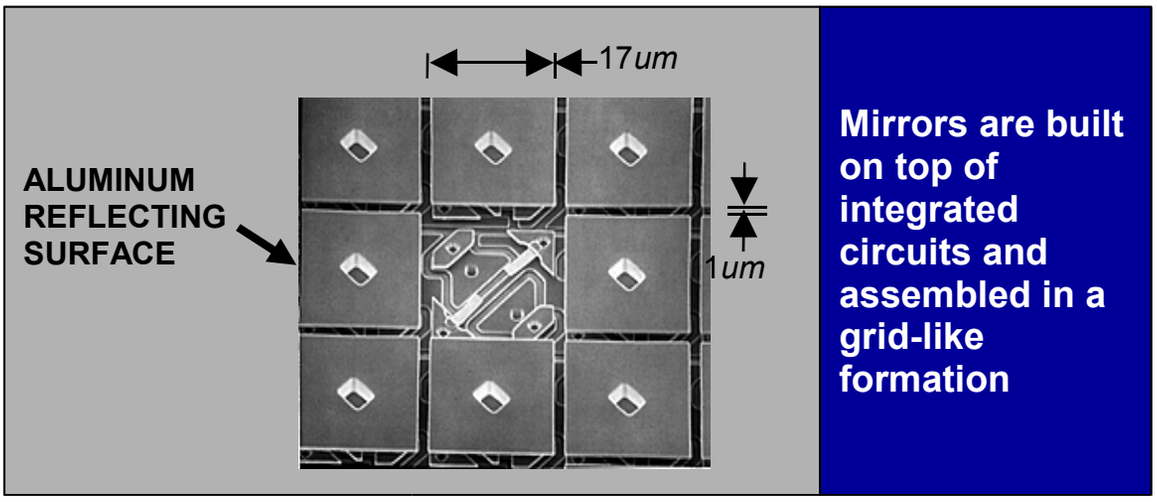


Figure 1: Three illustrations of the digital mirror array device

Image Source: 147 (76 normal & 71 malignant) hyperspectral gray scale 400X images are derived from 68 normal and 62 malignant colon biopsies selected from @ 200 normal and @600 malignant H & E stained biopsies arrayed respectively on two different slides[12] (Figure 2).

Cube: Each hyperspectral image is a 3-D data cube (Figure 3) with spatial coordinates x - 491 pixels, y - 653 pixels & spectral coordinates z - 128 pixels, a total of 41 million transmitted spectra.

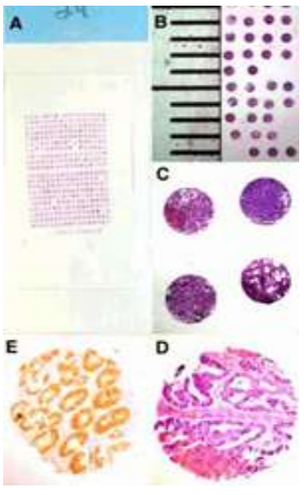


Figure 2: Microarray biopsies[14].

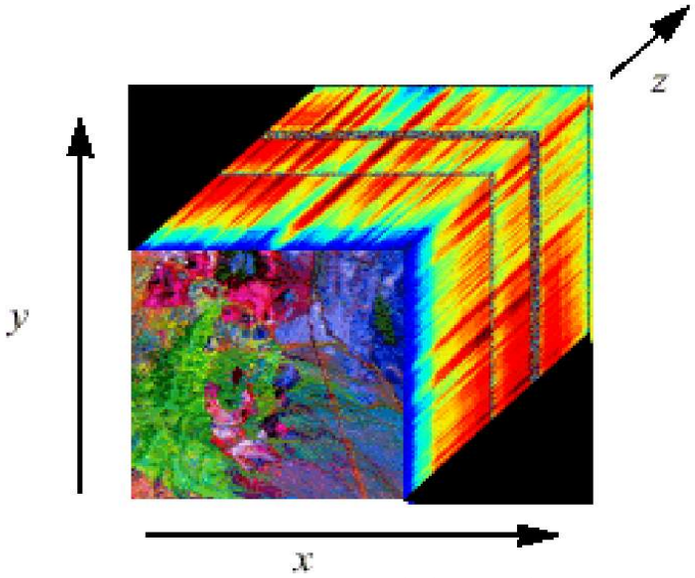


Figure 3: Hyperspectral data cube. (DataFusionCorp.)

DATA COLLECTION

To minimize the noise-to-signal ratio of the collected spectra, we multiplex the collected hyperspectral data by shining Hadamard light packets through the samples (Figure 4). The energy of each of these frequency packets is 1/2 the energy of the light source. Since these Hadamard patterns are independent the whole physical spectrum can be reconstructed from these packet measurements & the energy of the signal is uniformly spread in the spectral range (Figure 5).

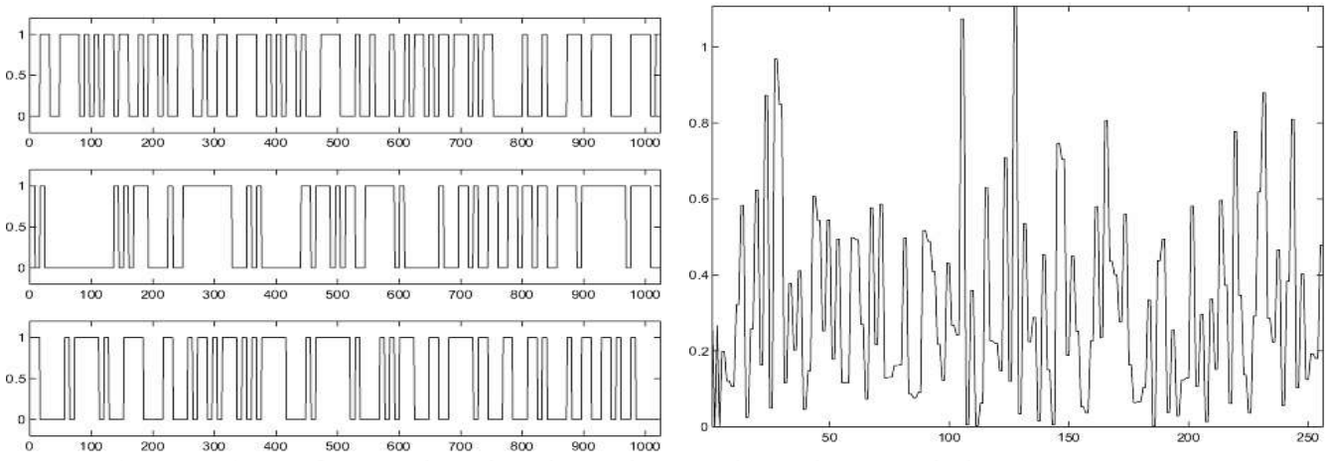


Figure 4: Left: randomized Hadamard packets. Right: spectrum representation on randomized Hadamard packets: the energy is approximately uniformly spread on the packets.

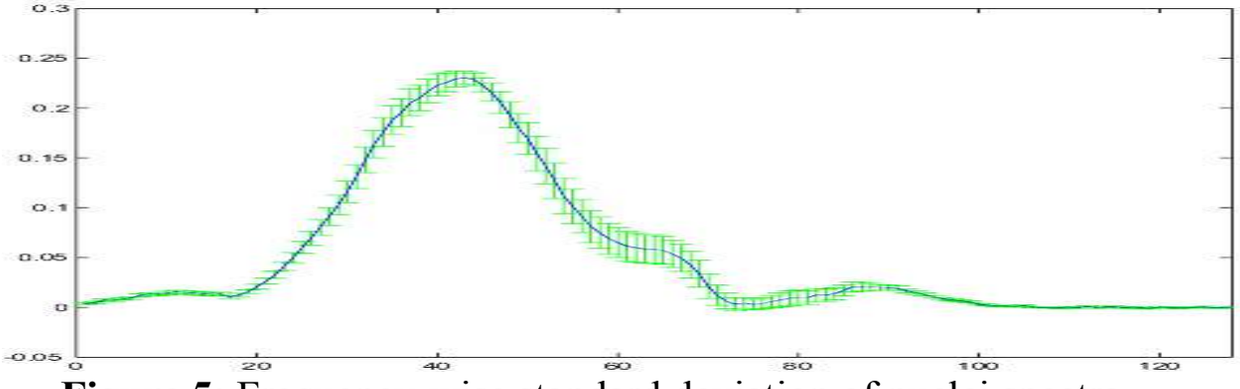


Figure 5: Frequency-wise standard deviation of nuclei spectra

Spectra to be analyzed are selected from biopsies samples of normal (Figure 5 left) and adenocarcinoma/malignant (Figure 5 right) samples.

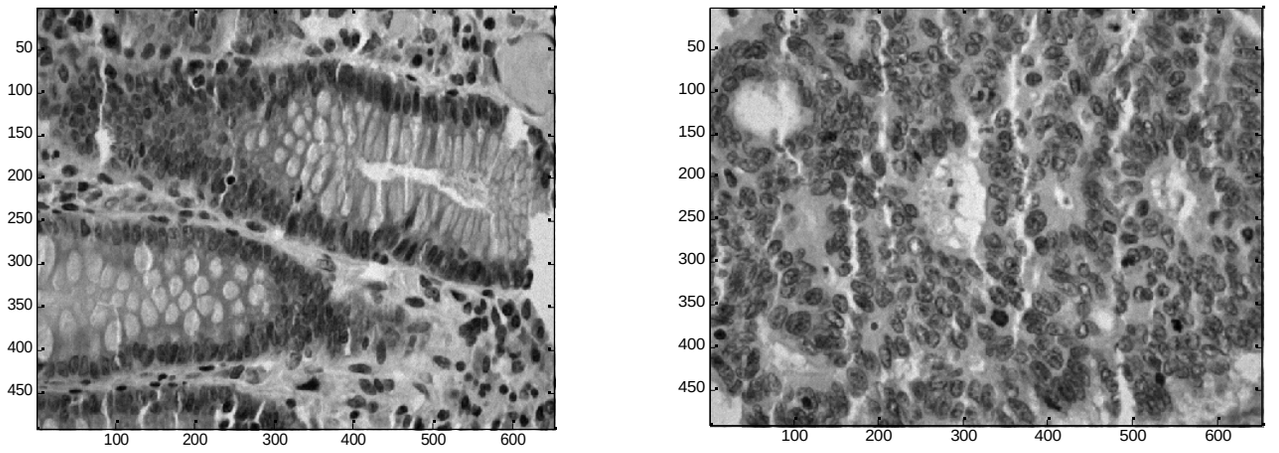


Figure 6: Gray scale spectral slices from normal (left) & abnormal (adenocarcinoma) (right) biopsies.

MATHEMATICAL ALGORITHMS[3,4,11]

Data Preprocessing and Normalization

De-noise the original 128 spectra.
Normalize the pixel spectra.

Tissue Type Classification

The first step is to differentiate among tissue types: gland nuclei & cytoplasm, and lamina propria.

Local Discriminant Bases

This algorithm identifies spectral features (Figure 7) that discriminate between the tissue types & projects the spectra onto these features.

Nearest-Neighbor classification

This algorithm acts on that projection and classifies each spectrum as one of the tissue types (Figure 8).

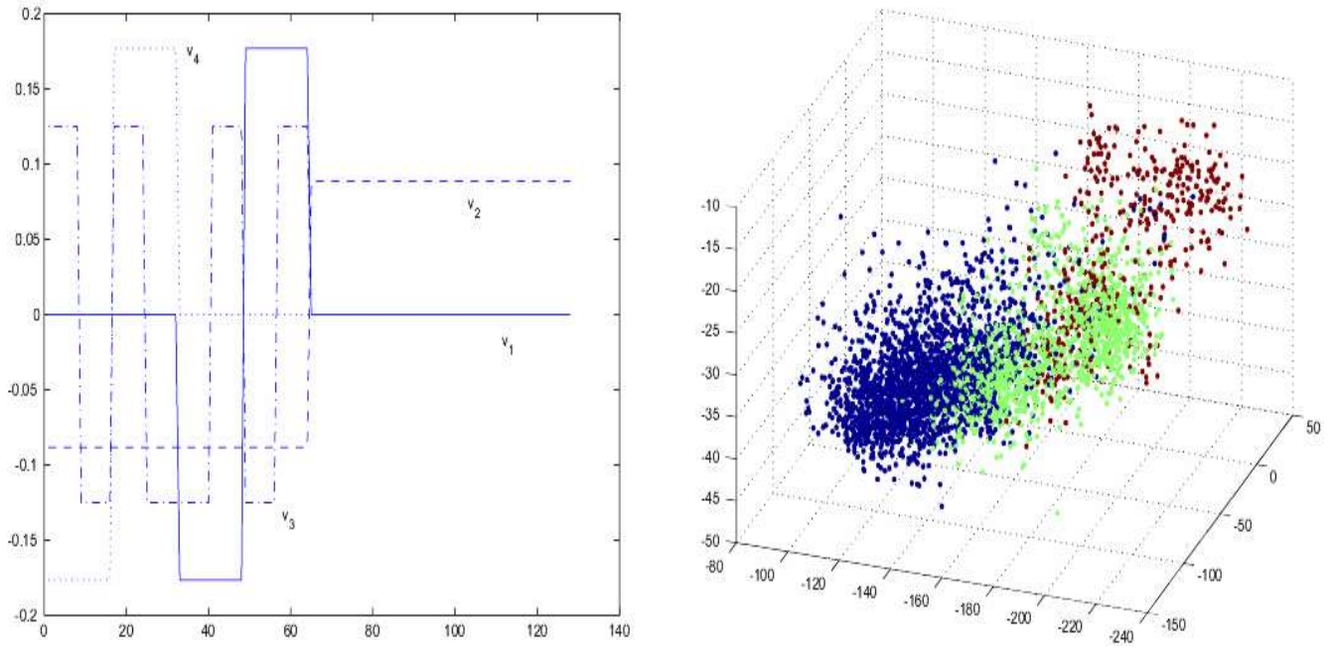


Figure 7: (Left) - Local Discriminant Bases discriminates among normal tissue types. (Right) - Projection of the normal training set onto these tissue features: nuclei (red), cytoplasm (green) lumens/lamina propria (blue).

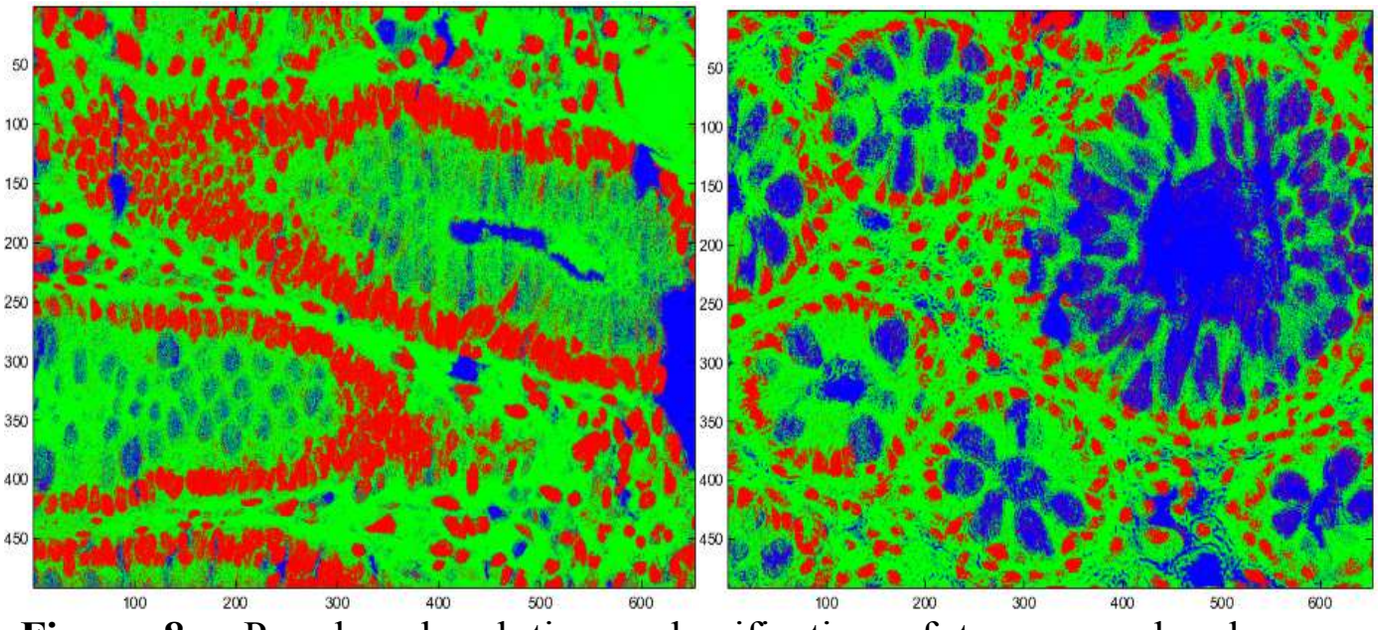


Figure 8: Pseudo-colored tissue classification of two normal colon biopsies. We automatically detect and extract the nuclei spectra (red).

Collection of Nuclei Patches

After tissue classification we collect 8688 patches, 32 by 32 pixels each (at pixel resolution & magnification each patch = @ one normal gland nucleus) & compute the average nucleus spectrum/patch.

Spectral Classifier of the Average Nucleus Spectrum in each Patch

We build a classifier on the average, de-noised nucleus spectrum of each patch by using **Partial Least Square Regression**.

The algorithm returns a spectral vector θ such that the projection of a nucleus spectrum onto θ is positive when the nucleus spectrum is likely to be normal, and negative when the nucleus spectrum is likely to be malignant.

This algorithm automatically selects 40 patches per biopsy centered around nuclei, computes the average nucleus spectrum in each patch, & classifies it by projecting onto the PLS vector (Figure 9).

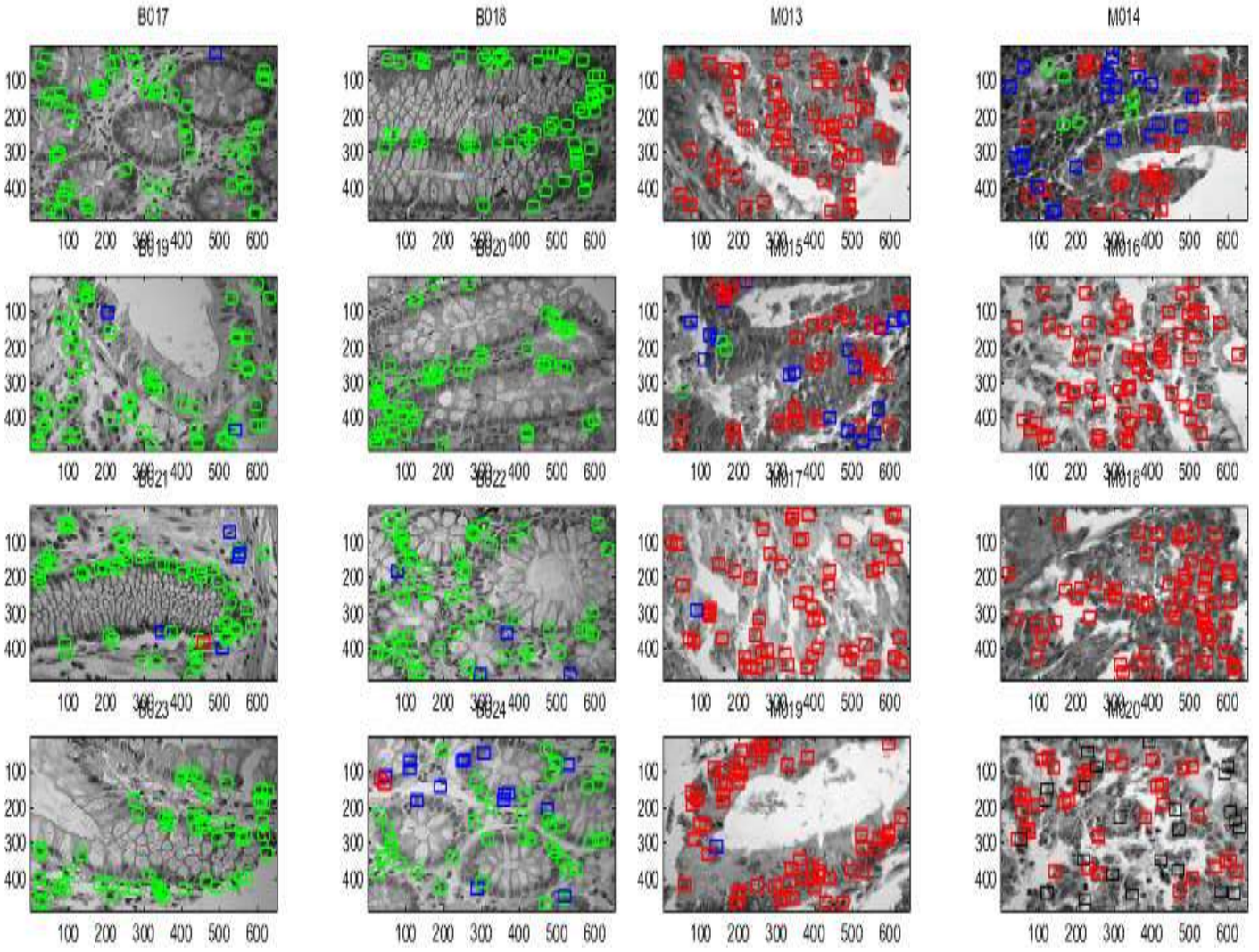


Figure 9: Spectral classifier on nuclei.
Normal (left 2 cols) :

GREEN – true negative (normal classified as normal);
BLUE – indeterminate
RED – false positive (normal classified as abnormal)

Abnormal (right 2 cols):

GREEN – false negative (abnormal classified as normal)
BLUE – indeterminate
RED – true positive (abnormal classified as abnormal)

ALGORITHM VALIDATION AND PERFORMANCE

Nuclei Patches: The algorithm was tested on the biopsies, by collecting patches (such that at least 80% of the pixels of each patch contain nucleus), and cross validating 100 times, and omitting 10% of the samples each time as testing set. Performance is summarized in Table 1.

Table 1: Performance of classifier on nuclei patches, cross-validated.

Patches/Nuclei (8688)	**True Positive (4860)**	**True Negative (3828)**
Predicted Positive (malignant)	94.0% (4568)	7.3% (280)
Predicted Negative normal)	6.0% (292)	92.7% (3548)

Sensitivity	92.79%
Specificity	93.92%
False Negative	7.21%
False Positive	6.08%
Pred Val Pos.	94.00%
Pred Val Neg.	92.70%
Diag. Eff.	93.35%

Biopsies: We run tissue classification on a biopsy instead of single nuclei & automatically select 40 nuclei patches, compute the average nucleus spectrum in each patch & classify that spectrum by using the spectral vector we found with the PLS algorithm. We average the classification over all the biopsy patches & classify the entire biopsy (Figure 10). Small errors of the nuclei classifier are completely averaged out over a whole biopsy with complete separation of normal & abnormal (malignant) biopsies yielding 100% diagnostic efficiency. Performance is summarized in Table 2.

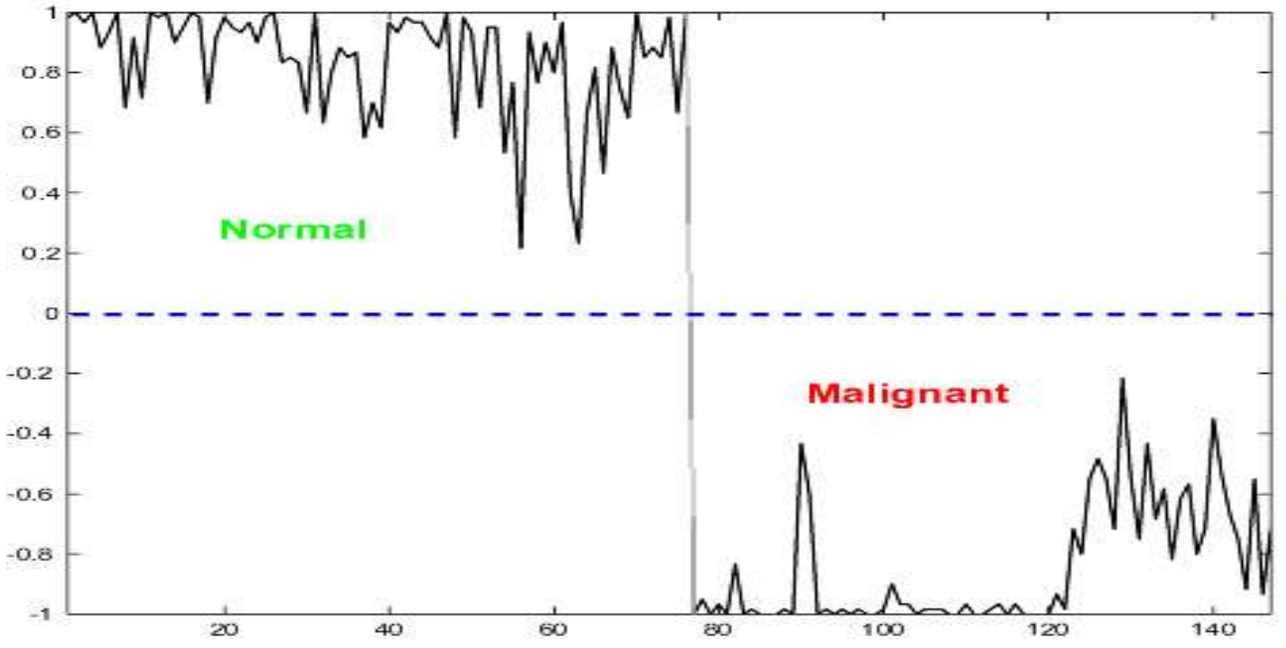


Figure 10: Classifier of biopsies obtained by averaging the classification of automatically extracted nuclei patches.

Table 2: Performance of the classifier on biopsies, cross validated.

Biopsies (130)	**True Positive (62)**	**True Negative (68)**
Predicted Positive (malignant 62)	100.0%	0.0%
Predicted Negative (normal 68)	0.0%	100.0%

CONCLUSIONS

As a "proof of concept" **we previously demonstrated** the efficacy of a digital mirror array device for **hyperspectral discrimination** between normal & abnormal (adenoma & carcinoma) colon biopsies **with 70% diagnostic efficiency[8] using only spectral analysis.** The addition of **spatial analysis increased diagnostic efficiency to 85.5%**

With the same device with completely different algorithms for spectral de-noising and classification of nuclei spectra[9,] and a larger number of biopsies as supervised training and test sets **we now achieve 100% diagnostic efficiency using spectral analysis alone.**

Taking a similar approach with different proprietary algorithms and a tuned liquid crystal light source, Jaganath et al.[6], also using only spectral information, discriminated between benign and malignant urothelial cells in routine cytology specimens, with 82% sensitivity & 81% specificity.

Montiorni et al.[12], have detected subvisual chromatin differences among normal, PIN-3 and adenocarcinoma in prostate biopsies with visual digital texture analysis (karyometry) and quantitative algorithms.

Information present within the nuclei of routinely prepared and stained tissues and cells can be accurately extracted and measured with readily available new optical technology and algorithms thus bringing quantification and objectivity to diagnostic pathology.

Clinical efficacy must be established with larger data sets including various "abnormal" controls in addition to malignancy.

We are currently constructing a microarray with benign "adenomas" for spectral comparison with the above normal and malignant colon biopsies.

REFERENCES

1. Barry TS, Gown AM, Yaziji HE, Levenson RW. Use of spectral imaging analysis for evaluation of multi-color immuno-histochemistry. Abstract 1481. Mod Pathol 2004;17 Suppl 1:350A.
2. Barshack I, Kopolovic J, Malik Z, Rothman C. Spectral morphometric characterization of breast carcinoma cells. Brit J Cancer 1999;79:1613-19.
3. Coifman RR, Saito N. Local discriminant bases and their applications. J Mathematical Imaging Vision 1995;5:337-58.
4. Coifman RR,Geshwind F, Saito N, Warner N. Discriminant feature extraction using empirical probability density estimation and a local basis library. Pattern Recog 2002;35:2841–52.
5. DeVerse RA, Coifmann RR, Coppi AC, Fateley WG, Geshwind F, Hammaker RM, Valenti S, Warner FJ, Davis GL. Application of spatial light modulators for new modalities in spectrometry and imaging. Proc SPIE 2003;4959:12-22.
6. Garini Y, Katzir N, Cabib D, Buckwald RA, Soenksen D, Malik Z. Spectral bio-imaging. In Wang XF, Herman B editors. Fluorescence imaging spectroscopy and microscopy. New York, John Wiley & Sons 1996:87-124.
7. Greenspan H, Rothman C, Cycowitz T, Nissan Y, Cohen AM, Malik Z. Classification of lymphoproliferative disorders by spectral imaging of the nucleus. Histol Histopathol 2002;17:767-73.
8. Jaganath R, Angeletti C, Levenson R, Rimm DL. Diagnostic classification of urothelial cells in urine cytology specimens using exclusively spectral information. Cancer (Cancer Cytopathol) 2004;102:186-91.
9. Levenson RM, Farkas DL. Digital spectral imaging for histopathology and cytopathology. Proc SPIE 1997;2983:1123-35.
10. Maggioni M, Davis GL, Warner FJ, Geshwind FB, Coppi AC, DeVerse RA, Coifman RR. Spectral analysis of normal and malignant microarray tissue sections using a novel micro-optoelectricalmechanical system. Mod Pathol 2004;17 Suppl 1:358A.
11. Maggioni M, Davis GL, Warner FJ, Geshwind FB, Coppi AC, DeVerse RA, Coifman RR. Data extraction algorithms applied to hyperspectral diagnosis of normal and malignant colon tissue microarray tissue sections. Yale University Program in Applied Mathematics, Technical Report September 2004, www.math.yale.edu/~mmm82/
12. Montironi R, Thompson D, Scarpelli M, Mazzucchelli R, Peketi P, Hamilton PW, Bostwick DG, Bartels PH. Karyometry detects subvisual differences in chromatin organization state between cribriform and flat high-grade prostatic intraepithelial neoplasia. Mod Pathol. 2004;17:928-37.
13. Papadakis A, Stathopoulos E, Delides G, Berberides K, Nikiforides G, Balas C. A novel spectral microscope system: Application in quantitative pathology. IEEE Transactions on Biomedical Engineering 2003;50:207-17.
14. Rimm DL, Charette L, Cowan J, Harigopal M. Tissue microarray facility. http://tissuearray.org/yale.index.html