

Graphical User Interface for Single-Pixel Spectroscopy

Mary M. Yang^{*}, Michael R. Dilworth, and Douglas C. Youvan

KAIROS Scientific Inc., Bldg. 62, 3350 Scott Blvd., Santa Clara, CA 95054 USA

ABSTRACT

Digital imaging spectroscopy (DIS) combines image processing and optical spectroscopy such that each picture element (pixel) or group of pixels in a two dimensional scene also includes additional dimension(s) of spectral and/or temporal information. Imaging spectrophotometers developed at KAIROS are literally equivalent to hundreds of thousands of conventional spectrophotometers running in parallel, i.e., each pixel is functionally equivalent to a single 'instrument'. Because of the amount of data acquired by this massively parallel technique – often in the gigabyte range for a single experiment – it was necessary to develop a new generation of graphical tools for data analysis and display. A unifying theme in our software development is the use of multispectral color contour plots to concisely represent data from a variety of different imaging experiments. Typically, the rows of a color contour plot represent encoded information for individual pixels. The columns of a contour plot represent the values of the variable that are being measured: absorbance, reflectance, fluorescence, FRET, or time course (kinetic) data. Contour plots are very powerful for classifying pixels with similar properties; therefore, these areas within the image can be 'back-painted' with a specific pseudocolor. Numerous sorting algorithms can be applied to the contour plot to map spectral and kinetic information back onto the spatial representation in run-time. Guided by the heuristics of parameters input by the user, the application of sorting algorithms to contour plots results in a near optimal interface between machine and human intelligence for the visualization of hyperspectral and/or hypertemporal images. In addition to a general introduction to the DIS graphical user interface, we demonstrate this new technology on H&E stained tissue.

Keywords: Imaging, spectroscopy, biotechnology, absorption, kinetics, software, algorithms, hematoxylin, eosin, H&E

1. INTRODUCTION

Twenty years ago, imaging spectroscopy was just beginning to be applied to problems in biotechnology. At that time, some of the authors on this paper were particularly interested in identifying and sequencing the genes encoding the reaction center and light harvesting antennae from photosynthetic bacteria. This was a fortuitous choice of a biological 'target' for imaging, because photosynthetic pigment-protein complexes have essentially 'built-in' colorimetric indicators for protein structure and function. Mutagenized colonies of the photosynthetic bacterium *Rhodobacter capsulatus* turned out to be an excellent model for the co-development of both genetic and imaging techniques (Youvan, 1994).

These early imaging efforts began during the transition from film to video techniques, and then they were followed later by another transition to CCD-based technologies. In about 1980, initial efforts in imaging bacterial colonies of photosynthetic mutants utilized near infrared (NIR) sensitive photographic film. Point mutants in the photosynthetic reaction center and light-harvesting proteins were first identified by fluorescence NIR photography (Youvan et al., 1983). Several years later, video imaging replaced film, and newly invented site-directed mutagenesis techniques superseded our study of spontaneous point mutants. One of the first embodiments of this technology utilized video pixel-pipeline processing in a 'Filter Array Spectrophotometer' to acquire and process spectral data from individual colonies on Petri plates (Yang & Youvan, 1988). Fluorescence and absorption images were digitized from each plate, and these digital images were combined by software to identify mutants based on their aberrant spectra.

About ten years ago, we became very interested in obtaining the full absorption spectrum for every colony on a Petri dish. At that point, it became difficult to display and process the data in a manner that the end-user could assimilate readily. A typical Petri plate (generated in the course of a mutagenesis experiment) might carry as many as 100 colonies, each ~ 3 mm in diameter. Continuing the research on the structure and function of light harvesting and reaction center proteins

^{*}Correspondence: Email: myang@kairos-scientific.com; <http://www.kairos-scientific.com>; Tel: 408 567 0400; Fax: 408 567 0440

required comparing each colony on these Petri dishes for very subtle spectral differences – such as shifts in the wavelength of maximal absorption, presence of peak shoulders, or differences in peak bandwidth. Conventional plotting of hundreds of spectra resulted in difficult to read graphs, in which the numerous spectra began to overlay and obscure each other. Therefore, we began experimenting with various spectral display methods – such as 'tiling' single spectra within a collage or generating various 3-dimensional surface plots. Ultimately, we arrived upon the idea of plotting all of the data within a 2-dimensional color contour plot, where a pseudocoloring technique was used to encode the spectral dimension of the data.

Such color contour plots (and their ability to be sorted in run-time) form the centerpiece of many of our present-day Graphical User Interfaces (GUIs) for the analysis of multispectral and kinetic datasets. This form of data visualization will be used here to demonstrate various applications of single-pixel spectroscopy. We begin with a pedagogical device for demonstrating such GUIs, and then we will present an application to H&E-stained tissue. Two other papers in these proceedings show how this technology can also be applied to multispectral bacterial identification (using fluorescent probes) and for isolating kinetic variants of genetically engineered enzymes (using chromogenic dye indicators); see Tanner et al., 2000; Bylina et al., 2000, respectively.

2. DISCUSSION

The graphical user interface for KAIROS' Digital Imaging Spectroscopy software is diagrammed in Figure 1. As shown, this GUI consists of four main interactive windows: Workspace, Image, Plot, and Contour. The workspace window has a look and feel similar to the Microsoft Windows Explorer, and it is used to organize projects and files. Typically, a project contains a data stack of images taken under conditions in which a spectral or temporal variable is incremented by wavelength or time, respectively. Each project can be identified with a single experiment, sample, or mode of data acquisition. In the latter case, our imaging spectrophotometers can be used to acquire absorption and fluorescence spectra for the same sample. In directed evolution and enzyme kinetics studies, both absorbance and kinetic data for a particular assay disk can be acquired. In each instance, these data types can be stored in separate projects which enables the simultaneous or combined analysis of fluorescence, absorption and/or kinetics for every pixel in the scene. A detailed discussion of this technology is presented in two concurrent SPIE proceeding papers: 1) mutant enzyme screening is presented by Bylina et al., and 2) bacterial ID using fluorescent probes is presented by Tanner et al.

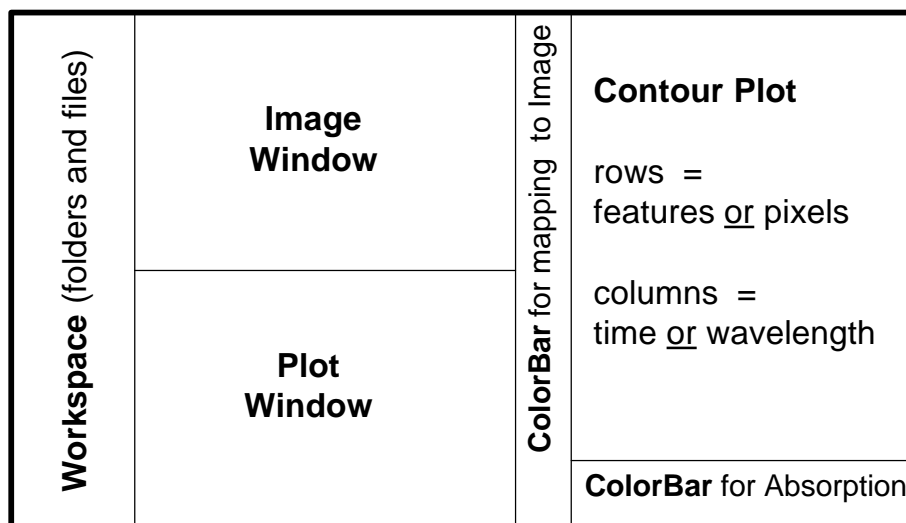


Figure 1. KAIROS' patented (Yang, 1999) graphical user interface for digital imaging spectroscopy utilizes coordinately controlled windows that are interactive in run-time. Several of the figures in this paper, and two other papers in these proceedings (Tanner, et al., 2000; Bylina et al., 2000) use this GUI. Note that two separate color bars are used to concisely encode information: (1) The horizontal bar under the contour plot provides a color-coded scale for the variable plotted within the contour plot. (2) The vertical bar (optional) to the left of the contour plot indicates rows that have been grouped and backcolored onto the image.

2.1 Single Pixel versus Blob Analysis

Using the KAIROS GUI (described above) pixels can be grouped within features by conventional image processing techniques, and then all four of the windows within the GUI act to coordinate feature- (rather than pixel-) based information. While feature-based analysis can increase the signal-to-noise ratio in certain low-light applications, we find feature extraction to be inferior to pixel-based analysis. This is due largely to problems associated with separating neighbor pixels into different features which may be adjacent or overlapping. Another reason for basing analyses on pixels rather than features is that problems with 'edge' pixels can be minimized (Figure 2). This is especially important when one attempts to identify objects in a target with the highest or lowest (spectroscopic) parameter – which would be otherwise averaged-out within a merged feature.

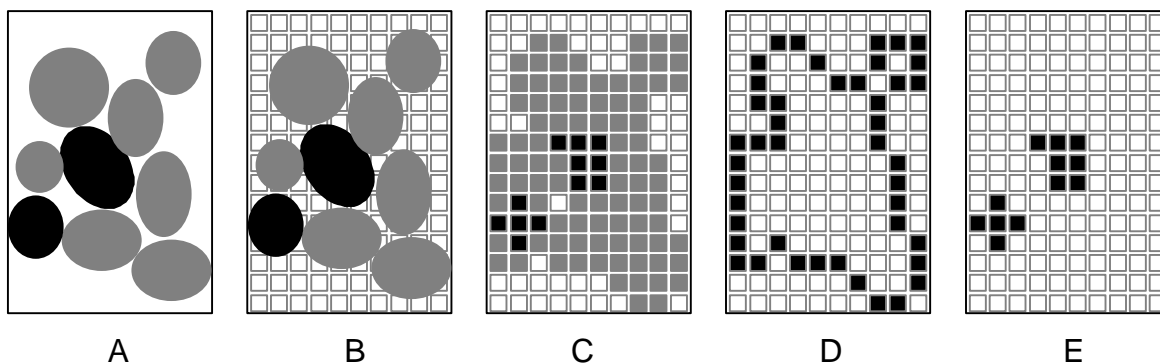


Figure 2. Picking a 'positive' (pixel or feature) is best done by pixel-based rather than by feature-based analysis. For example, in complex images such as confluent groups of microcolonies (Panel A), feature extraction (Panel D) is inferior to single-pixel analysis (Panel E) for identification of the 'fastest' colonies (represented in black) because of edge effects and other artifacts. This effect becomes more apparent as the features within a target become smaller and approach the apparent pixel size, i.e., each feature in Panel B covers only a few pixels and appears as a single 'blob' (Panel C) if it is processed based on feature extraction.

2.2 Pedagogical Test Target

Rather than using example targets that are either esoteric or difficult to recognize, we have opted to use a 'pedagogical target' in order to introduce this technology to the reader. Figure 3 shows a screen capture of the KAIROS GUI, where we have acquired a 'stack' of images of a plate of mini-M&M's candies (9 mm in diameter, seven colors) imaged in a reflectance mode on a black background. The image window has been magnified (zoomed) so that two individual pixels falling on a yellow and on a green M&M can be more easily identified. The absorption plots corresponding to the reflectance spectra of the target area ($\sim 25\mu \times 25\mu$) under these two pixels are shown in the plot window. As can be seen in the actual spectrum, the 'yellow pixel' encodes data for a microscopic region within a single yellow M&M; its spectrum exhibits high absorbance in the blue. The 'green pixel' encodes data for a green M&M; its spectrum exhibits high absorbance in both the blue and red, with minimal absorption in the green. Both of these spectral plots are linked by the GUI to a yellow and a green tick mark to the left of the contour plot (left side of color contour plot window). Each row of the contour plot corresponds to a single-pixel absorption spectrum which is color-encoded according to the rainbow-like colorbar directly beneath. Using the row indicated by the yellow tick mark and referencing it with the yellow-colored conventional plot in the plot window, we see white pseudocoloring of this row in the contour plot corresponding to high absorbance in the 450 - 490 nm region, a rainbow of hues between 495 - 540 nm as the absorbance decreases, followed by black as we look from left to right in the row.

The KAIROS GUI is highly interactive in run-time. The computer's mouse can be used to point to a pixel in the image window which causes two actions: 1) the pixel spectrum is displayed in the plot window, and 2) a tick mark appears next to the associated row in the contour window. Alternatively, pointing to a row in the contour plot causes two analogous actions: 1) the corresponding pixel is highlighted in the image window, and 2) the associated conventional plot is updated for that particular pixel. With the mouse button down, vertical movement over the contour plot selects multiple rows whose spectra are then plotted simultaneously. Likewise, dragging out a box in the image window will simultaneously indicate the corresponding spectra in both the plot and contour windows. Once these selections are made, numerous options for color display, spectral averaging, and further processing are provided by various drop-down menus. Most of these functions are enabled by keyboard and mouse manipulations well-known to MS Windows' users.

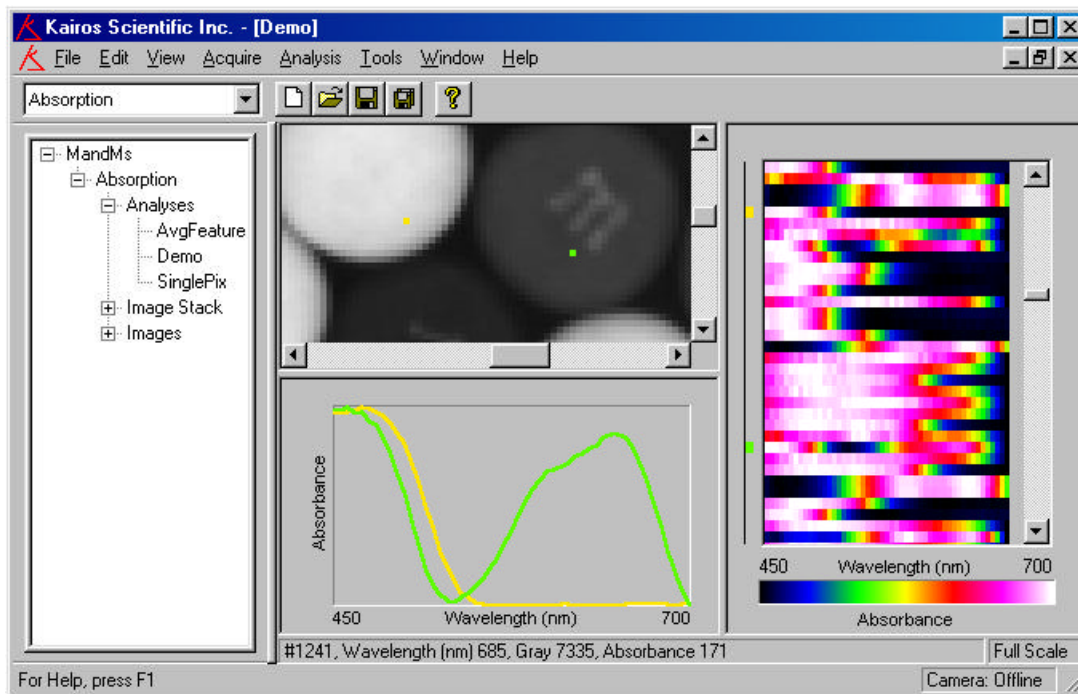


Figure 3. Screen dump of the KAIROS GUI demonstrating reflectance imaging spectroscopy and single pixel analyses of M&M's candies. The four windows in this image correspond to those depicted in Figure 1. Two, single-pixel spectra have been highlighted for microscopic regions within a yellow and a green candy.

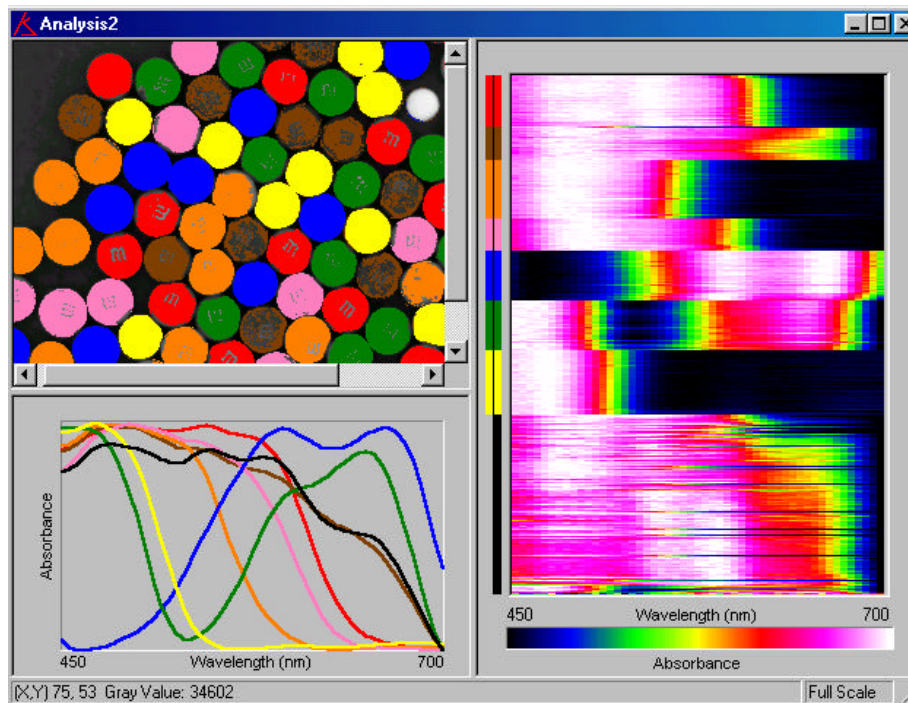


Figure 4. Backcoloring of M&M's candies based on spectral sorting of single-pixel data. A pseudocolor bar (vertical; left of the contour plot) uses colors approximating the real color of the candies. This color code is maintained in the plot window for each of the seven different colored candies, along with a eighth plot for the black background. All spectra are plotted with full-scale (variable) deflection. A 24-bit high resolution version of this GUI demonstration is available on-line at www.et-al.com.

2.3 Contour Plot Sorting Algorithms and Display

As shown above, contour plots are very effective for visualizing and extracting useful information from massive amounts of spectral data. The concise display of data in Figure 4 is possible because of the application of a series of sorting algorithms that group pixels with similar properties – in this case, visible light absorption. Aspects of this sorting process are shown below (Figure 5). Similar to Figure 4, Panel A displays the desampled contour plot in a variable scale mode, wherein each spectrum has been stretched to full scale, so that each row displays absorbance intensity ranging from black to white on the absorption color-code bar. Panel B shows this same contour in a fixed scale mode, wherein each spectrum has been normalized to the minimum and maximum absorbances for the entire dataset. Both Panels A and B are unsorted and represent spectral information as it was initially acquired. In Panel C, we have sorted fixed scale spectra by the maximum absorbance, so that the spectra with the highest optical density (represented by white to reddish colors) are grouped at the top of the contour plot.

In Panel D we have sorted the spectra within the contour plot by the sum-of-the-square-of-differences (SSD) using an initial yellow target spectrum. The SSD equation is:

$$SSD = \sum_j (I_A(j) - I_B(j))^2 \quad \text{Eqn.1}$$

Where I_A represents the intensity of the spectrum A at wavelength j and I_B is a target spectrum. As can be seen in Panel D, single pixel spectra corresponding to yellow M&M's have been isolated at the top of the contour plot. In Panel E, the SSD calculation has been recursively performed on all spectra to determine their similarity. Numerous groupings and categories of spectra are apparent. Panel E is similar to the contour plot shown in Figure 4, except that it has been thresholded to remove the black background.

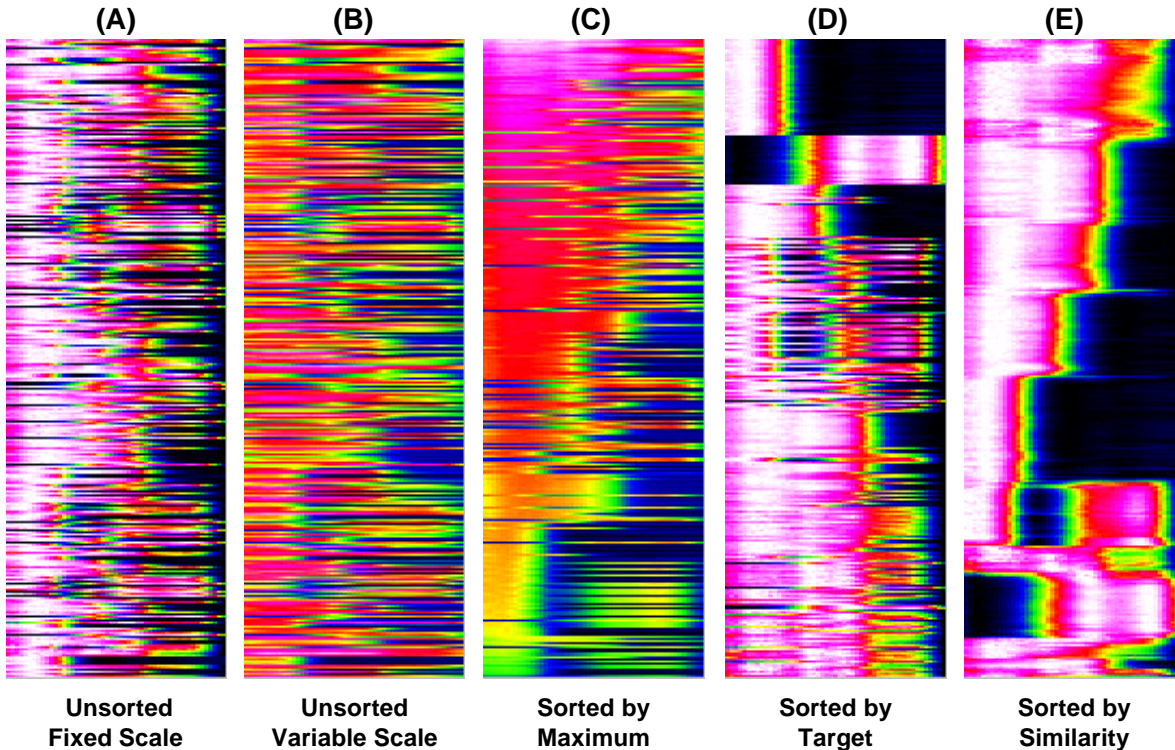


Figure 5. Contour plots can be sorted to group single-pixel spectra by similarity or other criteria. Unsorted data (Panel A) can be processed to group spectra with similar attributes (Panel E). User-input heuristics is important in guiding the particular series of algorithms that are required to sort a dataset to a level of refinement appropriate for backpainting the image window.

2.4 Analyses of H&E Stained Tissue

Hematoxylin and eosin (H&E) stains are performed on almost all biopsied tissues before any other special stain or immunochemical analysis is considered. As a result, there are on the order of 10,000 H&E stained thin sections analyzed per day in the US. H&E's widespread use in histology (Prento & Schulte, 1991) has as much to do with the pathologist's familiarity and experience with this stain as it does with cost considerations. H&E is informative, inexpensive, and utilizes relatively low cost instrumentation (i.e., microtomes and brightfield microscopes). H&E stains cell nuclei, basophilic cytoplasmic structures, some elastic fibers and mucins a blue-purple, while most other structures are stained shades of red.

Despite its widespread and established use, there is no 'standard H&E' stain. In fact, pure hematoxylin is white to yellowish in color (Green, 1990). In order to be used as a stain, hematoxylin is oxidized to hematein which then forms a colored metal complex. Procedures for this reversible oxidation reaction can be carried out by "ripening" or oxidation with exposure to air (for up to a period of 8 weeks (Delafield)), boiling the solution with HgO (Harris), addition of NaIO₃ (Ehrlich, Mayer's, Clark), or addition of ferric salts, peroxides or perchlorates. In many cases, a further oxidized product, oxyhematein (brown) is irreversibly formed. Metal chelates of hematein range in color from red-violet to blue-black, depending on the metal. The result of stains using Al-hematein is strongly dependent on the molar ratio of Al:hematein, pH and ionic strength (Prento & Schulte, 1991). Al-hematein solutions are red at low pH and blue at high pH. This acid/base indicator property is used to advantage in H&E staining. However, the exact pH is difficult to control and will be a function of the redox potential as well as the stepwise association of SO₄⁼ (from the alum added), among other factors.

Eosin Y is the most commonly used counterstain for hematoxylin. According to *Staining Procedures Used by the Biological Stain Commission*, (Puchtler, 1981) "the range in individual preferences for the intensity of the eosin stain is so great that only approximate staining and dehydration times can be suggested ... procedures are tentative and must be adjusted to suit the individual preference." In fact, eosin B is often substituted for eosin Y to give a slightly pinker cytoplasmic stain lacking the yellowish tinge of eosin Y. Furthermore, phloxine is sometimes added for additional "red" colors. Our literature searches of the Medline and Biosis databases reveals some uncertainty as to the actual chemical formulae of hematoxylin derivatives (e.g., Alum-hematein) and the structure of possible protein or DNA adducts. Except for a spectrum of hematoxylin in methanol given in the *Sigma-Aldrich Handbook of Stains, Dyes and Indicators* (Green, 1990), no purification or spectral characterization of these putative hematein derivatives can be found. We suspect that several different reactions occur between Al-hematein and biological substances and that this affects the spectrum of the chromophore.

The discussions above indicate that the staining process is difficult to control, and that the information obtained from a stained thin section is often based on very subtle color differences. Standardization and visual enhancement of such differences would benefit the entire histology community. Parameters to consider include:

- Type of tissue
- Normal or abnormal tissue
- Degree of autolysis of the tissue
- Method of processing
- Type of mounting medium, slides and coverslips (refractive index)
- Type of "Paraffin" used for infiltration
- Thickness of tissue section cut on the microtome
- Type of Hematoxylin (Mayers, Harris, or Gill)
- Type of Eosin (alcoholic, aqueous, or combined with phloxine)
- Alternative methods using frozen sections

Subtle differences in pink colors may be used by pathologists in differentiating amyloid from normal connective tissue if these color differences can be accentuated or quantified. This is a case where morphological differences may be inadequate for differentiation and a colorimetric aid could be definitive. In this particular case, current technology would require a Congo Red stain to verify the presence of amyloid. This specialized stain would be delayed relative to the initial observation and could not be taken on exactly the same group of cells originally stained with H&E. Our ability to detect amyloid in cardiac tissue would be a prelude to the analysis of other tissues wherein the presence of amyloid is indicative of disease. We have begun to explore this phenomenon in the next two DIS analyses of H&E-stained normal tissue using a brightfield microscope configuration (Yang, 1999; Yang et al., 1998).

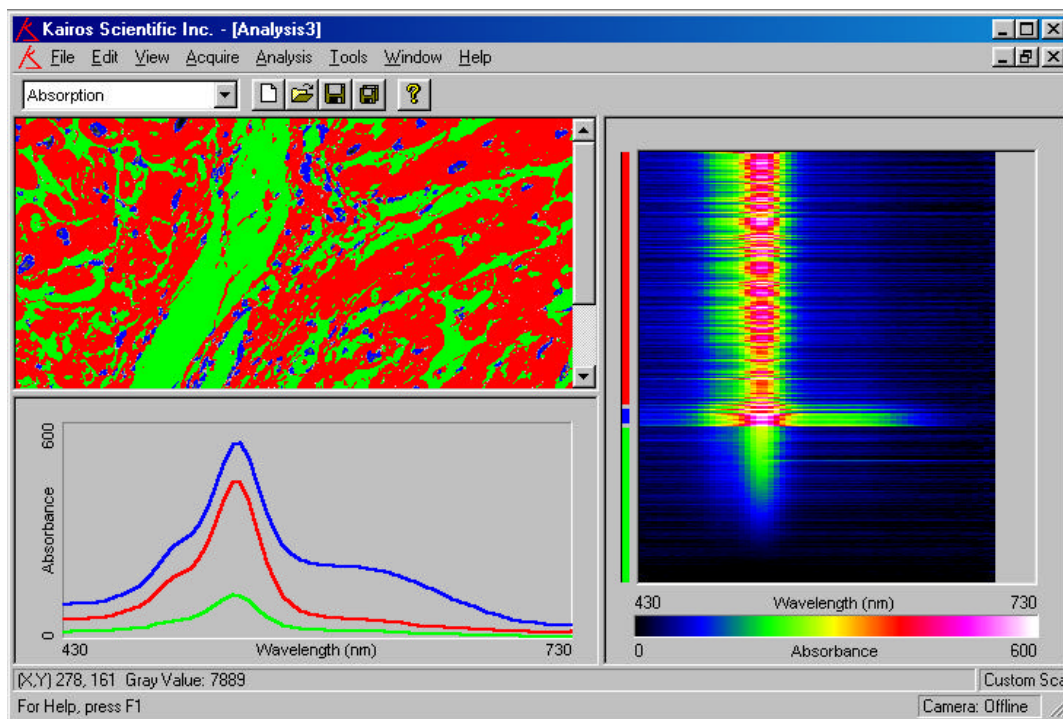


Figure 6. Single pixel absorption spectra from a slide stained with hematoxylin and eosin (H&E). The image window shows a transverse section of cardiac tissue back-painted based on the sorted single-pixel spectra. Clear areas in the slide are sorted to the bottom of the contour plot and pseudocolored green. Pixels indicating hematoxylin absorption (610 nm) are painted purple and coincide with nuclei. Eosinophilic tissue is painted red.

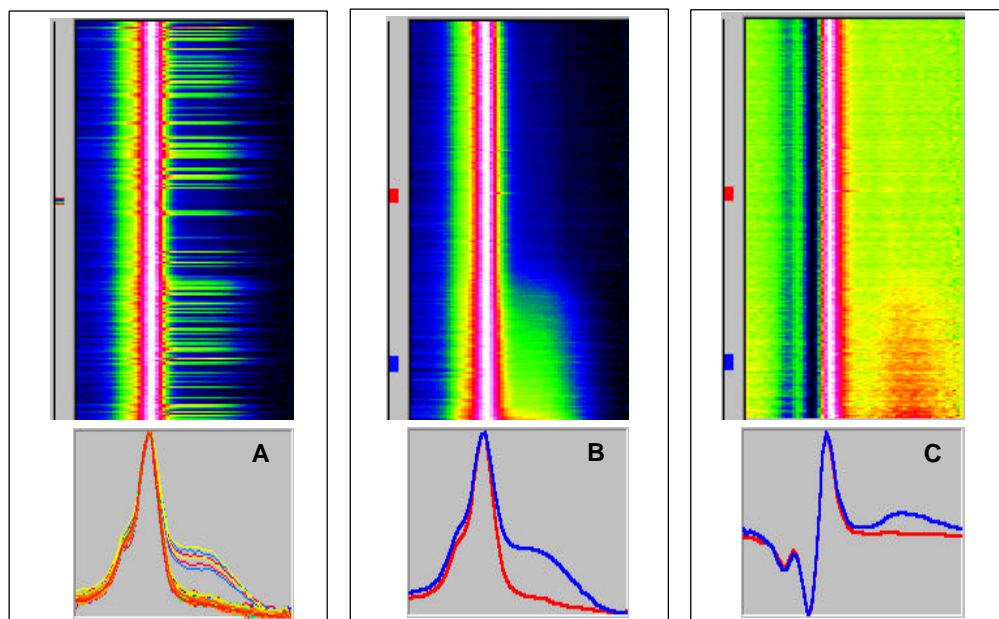


Figure 7. Desampled contour plots displaying single-pixel absorption spectra from a slide stained with H&E. Panel A shows the initial unsorted contour plot in a variable scale mode. The plot window (A) shows single-pixel spectra corresponding to the colored tick marks in the contour plot (directly above). In Panel B, we have sorted the variable scale spectra by the ratio of optical absorbances at 540 nm to 610 nm. Thick blue and red lines in the plot window (B) correspond to spectral averages of pixels grouped above in the contour plot, marked by blue and red bars, respectively. Panel C shows a derivative plot of these spectra with an eosin-specific inflection point at 530 nm.

ACKNOWLEDGEMENTS

The work reviewed in this paper has been supported by NASA NAS214386; NIH R44GM55470, R43GM60209, R43GM60073; Div. of Energy Biosciences, DOE DE-FG03-96ER20211, and by KAIROS IR&D funds.

REFERENCES

- Bylina, E.J., Grek, C.L., Coleman, W.J., and Youvan, D.C., Directed Evolution and Solid Phase Enzyme Screening. *SPIE* (this volume).
- Green, F.J. (1990) The Sigma-Aldrich Handbook of Stains, Dyes and Indicators. Aldrich Chemical Company, Milwaukee.
- Prento, P., and Schulte, E. (1991) Staining involving Metal Complex Dyes. In, Theory and Strategy in Histochemistry. Ed: H.Lyon., Springer Verlag, p107-119.
- Puchtler, G. (1981) Staining Procedures, 4th edition., Williams & Wilkins, Baltimore.
- Puchtler, H., Meloan, S. N., and Waldrop, F. S. (1986) Application of current chemical concepts to metal-hematein and brazilin stains. *Histochemistry* 85:353-364.
- Tanner, M.A., Coleman, W.J., Everett, C.L., Robles, S.J., Dilworth, M.R., Yang, M.M., and Youvan, D.C., Multispectral Bacterial Identification. *SPIE* (this volume).
- Yang, M.M. and Youvan, D.C. (1988) Applications of Imaging Spectroscopy in Molecular Biology. I. Screening Photosynthetic Bacteria. *Nature Biotechnology*, 6:939-942.
- Yang, M.M., Coleman, W.J., Silva, C.M., Dilworth, M.R., Bylina, E.J., and Youvan, D.C. (1998) High Resolution Imaging Microscope (HIRIM). *Biotechnology et alia*, <www.et-al.com>4:1-20.
- Yang, M.M. (1999) High Resolution Imaging Microscope (HIRIM) and Uses Thereof. U.S. Patent 5,859,700.
- Youvan, D.C., Hearst, J.E., and Marrs, B. (1983) Isolation and Characterization of Enhanced Fluorescence Mutants from *Rhodospseudomonas capsulata*. *J. Bact.*, 154:748-755.
- Youvan, D.C. (1994) Imaging Sequence Space. *Nature*, 369: 79-80.