

Cleveland State University
EngagedScholarship@CSU



Chemistry Faculty Publications

Chemistry Department

5-2002

Chemical Imaging Comes of Age

John F. Turner II

Cleveland State University, J.F.TURNER@csuohio.edu

Follow this and additional works at: https://engagedscholarship.csuohio.edu/scichem_facpub

 Part of the [Chemistry Commons](#)

How does access to this work benefit you? Let us know!

Recommended Citation

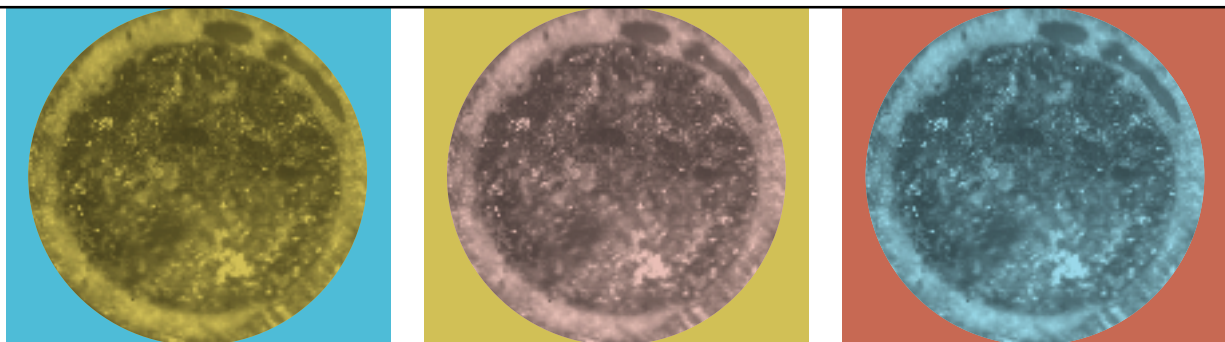
Turner, John F. II, "Chemical Imaging Comes of Age" (2002). *Chemistry Faculty Publications*. 469.
https://engagedscholarship.csuohio.edu/scichem_facpub/469

This Article is brought to you for free and open access by the Chemistry Department at EngagedScholarship@CSU. It has been accepted for inclusion in Chemistry Faculty Publications by an authorized administrator of EngagedScholarship@CSU. For more information, please contact library.es@csuohio.edu.

Sophisticated analysis techniques allow researchers to extract meaningful results from spectral imaging data.

CHEMICAL IMAGING

comes of age



By John Turner II, Cleveland State University

Before the introduction of technologies like the CCD camera and electronically tunable wavelength filters that ushered in the age of chemical imaging, optical spectroscopy was an immensely important technique that had established itself as a premier analytical tool for determining sample composition and concentration. The advantages of optical spectroscopy include minimal sample preparation, nondestructive analysis, high sensitivity, fast acquisition times, and, depending on the wavelength region and spectroscopic technique employed, excellent chemical specificity.

challenges of inhomogeneity

While the breadth of analytical problems within the scope of conventional optical spectroscopy seems endless, there are drawbacks to the technique that continue to spawn efforts to advance the field. Among them is the inherent pro-

clivity of conventional spectrometers to record bulk spectra that can retain contributions from all regions of the sample that lie within the optical path. Spectra acquired from liquid samples that are inhomogeneous or that contain air bubbles can yield erroneous concentration values. The situation worsens when mixtures of solids need to be analyzed; packing density and sample mixing can strongly affect the recorded spectrum.

The validity of the bulk spectrum for heterogeneous samples is often verified through conventional means by repeating the analysis on another sample region (a point mapping experiment, for example). Recording specific spectral features such as band shape or peak height as a function of the sampling position can be used to construct chemical maps of the sample heterogeneity. The result is an image, however crude, exhibiting chemically based image contrast.

data analysis

Although researchers coined the term chemical imaging several decades ago, a spectral imaging revolution began to rapidly spread throughout the analytical sciences in the early 1990s. Key technologies that advanced the field were enhanced multichannel array detectors like CCD cameras and electronically tunable wavelength filters. The popularity of chemical imaging stemmed from its ability to capture large numbers of spectra at spatial resolutions of better than 200 nm across the sample surface.

Although important successes were achieved, the long-term growth of chemical imaging was unmistakably tethered to complicated data analyses that were needed to extract meaningful chemical details about the sample. The number of datum points acquired in a typical wide-field spectral imaging experiment can easily exceed 100 million and is a

strong impetus for automated data processing. While modern desktop computers are up to the task, the difficulty remains generating effective algorithms to sift through the numerical data.

Conventional chemometric techniques such as factor analysis, least squares fitting, principal components analysis, and principal components regression are powerful tools for determining the composition and concentration of samples with known

faces. An underlying theme in all of this work is data processing, and we are working on novel multivariate methods for generating chemically meaningful image contrast without the need for *a priori* sample information or training sets. In particular, we have developed cosine correlation analysis¹ (CCA) and cosine histogram analysis (CHA), which generate qualitative and quantitative chemical image contrast, respectively.

In its simplest form, CCA is a qualitative technique that can be easily tailored to more rigorous quantitative analyses. The objective of CCA is to rapidly reveal changes in sample composition by correlating the shape of each spectrum in the chemical image data set to a reference spectrum. The calculated correlation values are mapped as colors to the pixel locations where the spectra originated. The result is a single image that exhibits chemically relevant image contrast.

In order to calculate the correlation values, the reference spectrum and the image spectra are represented as vectors in n dimensional space, where n is the

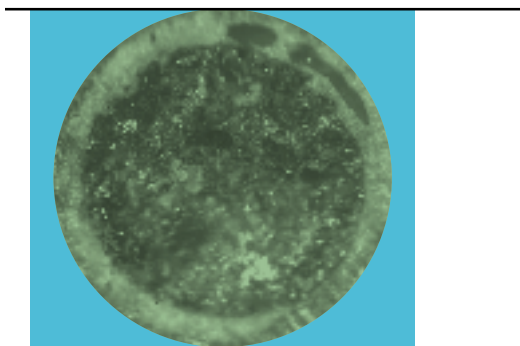
number of wavelengths (see figure 1). The correlation between the reference spectrum, D_R^* , and the i^{th} spectral vector, D_i^* , is given by

$$r_{iR} = \frac{D_i^* \cdot D_R^*}{n\sigma_i\sigma_R}$$

where σ_i and σ_R are the standard deviations of the intensity values in the i^{th} spectral vector and the reference vector, respectively. By substituting $|D_i^*||D_R^*|\cos\theta$ for the inner product, $D_i^* \cdot D_R^*$, and expanding σ_i and σ_R , the statistical correlation becomes

$$r_{iR} = \cos\theta$$

where θ is the angle between D_i^* and D_R^* . Consequently, the correlation between two spectra can be thought of as the cosine of the angle between their vector representations in n dimensional space. As θ becomes smaller, the spectra appear more similar, and $\cos\theta$ approaches unity. Spectra that are orthogonal to one another have $\cos\theta$ values equal to zero and exhibit no correlation.



constituents. Neural networks and data segmentation approaches have also been successful.

The chief disadvantage of conventional methods is that they rely on training data or learning algorithms. In fact, most chemometric analyses require a carefully constructed data model on which to base the results. In spectroscopy, the training model is typically a set of spectra from samples having known composition that span the data space. The spectra from these training samples are used to estimate the chemical composition of the unknown sample.

The difficulty arises when the sample composition increases from several components to many components, some of which are unknown. This scenario is typical in chemical imaging experiments in which an understanding of sample heterogeneity is sought. Likewise, chemical images are often the means for discovering unexpected heterogeneity.

In our work at the Advanced Chemical Imaging Facility at Cleveland State University, we use wide-field Raman, UV-visible, and fluorescence chemical imaging to investigate the nature of interactions that occur at tissue-biomaterial inter-

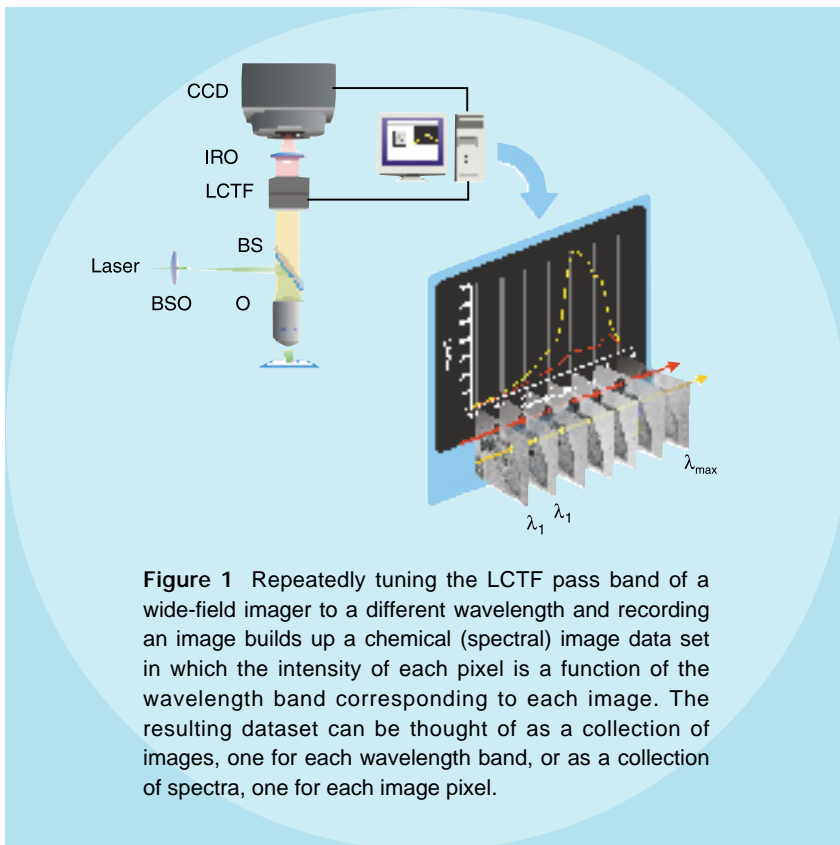


Figure 1 Repeatedly tuning the LCTF pass band of a wide-field imager to a different wavelength and recording an image builds up a chemical (spectral) image data set in which the intensity of each pixel is a function of the wavelength band corresponding to each image. The resulting dataset can be thought of as a collection of images, one for each wavelength band, or as a collection of spectra, one for each image pixel.

Since the direction of each vector corresponds to a specific spectral shape, the cosine correlation values are qualitative indicators for chemical composition. Similarly, the vector lengths, each corresponding to a specific spectral intensity, are quantitative indicators for concentration. Because $\cos\theta$ remains unchanged if the lengths of the spectral vectors are altered, qualitative CCA is scale-invariant and is relatively immune to illumination nonuniformities. A practical limitation exists in cases where noise begins to dominate the overall shape of the spectrum.

One of the interesting consequences of wide-field imaging is that large chemical and morphological features within the sample can yield statistically relevant quantities of spectral data. In the visible region of the spectrum, for example, the obtainable spatial resolution (~ 200 nm) is higher than what is required for imaging most biological cells (~ 1 to 5 μm). The result is that major components within cells, such as the nucleus, are sampled many times across their spatial extent, yielding potentially hundreds of spectra. While the average spectrum fluctuates little from nucleus to nucleus, the variability in their spectral populations can differ substantially.

The purpose of CHA is to extract meaningful statistical information from CCA results. The procedure is a two-step process that begins by identifying the spatial extents of chemical domains within the CCA image. Once identified, each domain is treated as a discrete data population, and trends are sought between the pattern of variability in their CCA values and other forms of empirical data. For instance, the independently measured metabolic activity in hepatocytes might correlate with a specific pattern of variability in the shapes of their absorption spectra. After revealing the trend, CHA can estimate the metabolic activity of other hepatocytes based on the statistical criteria obtained from their CCA data populations. The CHA result is an image of semi-quantitative values for a particular data classification.

Although CHA requires training samples in order to establish the data trends, the criteria needed to establish effective training samples are consid-

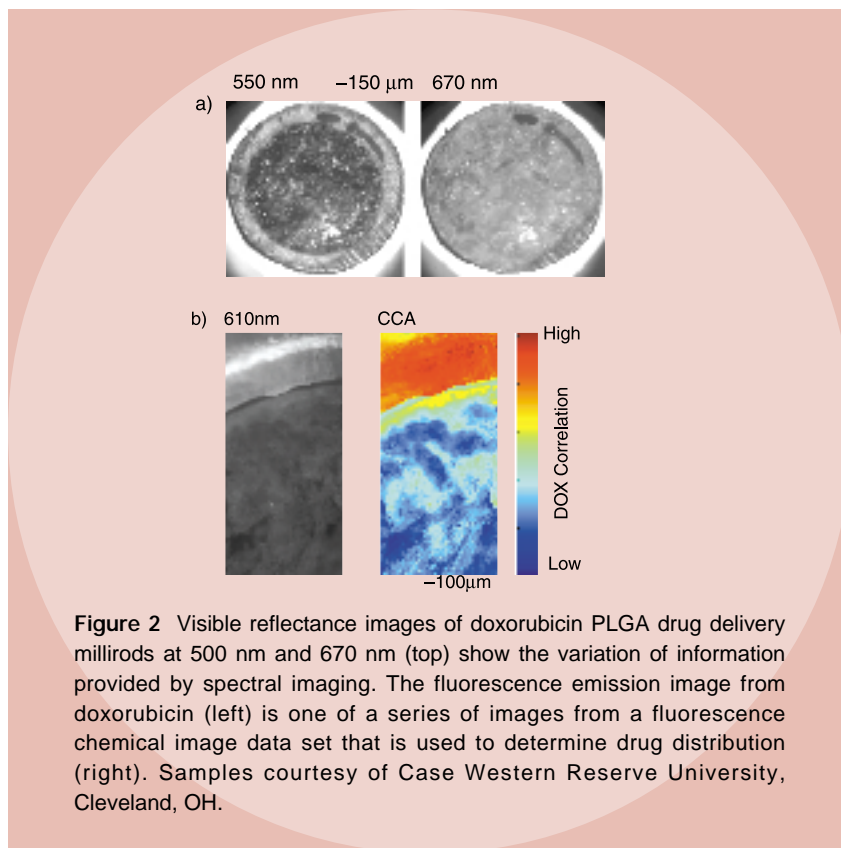


Figure 2 Visible reflectance images of doxorubicin PLGA drug delivery millirods at 500 nm and 670 nm (top) show the variation of information provided by spectral imaging. The fluorescence emission image from doxorubicin (left) is one of a series of images from a fluorescence chemical image data set that is used to determine drug distribution (right). Samples courtesy of Case Western Reserve University, Cleveland, OH.

erably relaxed. In fact, meaningful CHA image contrast can be generated on the basis of the statistical criteria calculated for each chemical domain in the absence of training data. Such analyses are often useful for surveying preliminary chemical image data.

biomedical applications

Perhaps the most dramatic influence of chemical imaging is in the areas of bioanalysis and biotechnology. Researchers are applying chemical imaging techniques to investigations of cellular function, disease processes, protein interactions, DNA, biomaterials, and pharmaceuticals. The unique advantage of chemical imaging over traditional analytical tools is that it provides morphological and compositional information simultaneously (see figure 2). Drug distribution inside the polymer matrix, composition and structure of the devices, and drug stability during and after fabrication are important areas of investigation that accompany the development of advanced drug-delivery devices.

New strategies for non-invasive and nondestructive disease diagnosis and assessment are also benefiting from

chemical imaging. For example, our laboratory is investigating the role of wide-field Raman and visible reflectance chemical imaging in the diagnosis of cervical cancer. Because identification of abnormal cells is based on conventional histological examination using an optical microscope, failure to identify dysplastic cells and misinterpretation of the degree of dysplasia are risks that can avert a proper diagnosis and course of treatment. Currently, the Papanicolaou (Pap) smear is the most common screening method for uterine cervical dysplasia and carcinoma *in situ*. Although the Pap smear is a simple and effective method for identifying patients at risk, it has little diagnostic value. The false negative rate for the Pap procedure is typically between 10% and 20%, and can be as high as 55%—hence the need for an alternative.

The challenge in spectroscopic diagnostics is in developing robust multivariate algorithms to sort out and classify abnormal cell types from the range of normal cells, which vary considerably in morphology and optical properties. The focus of our work

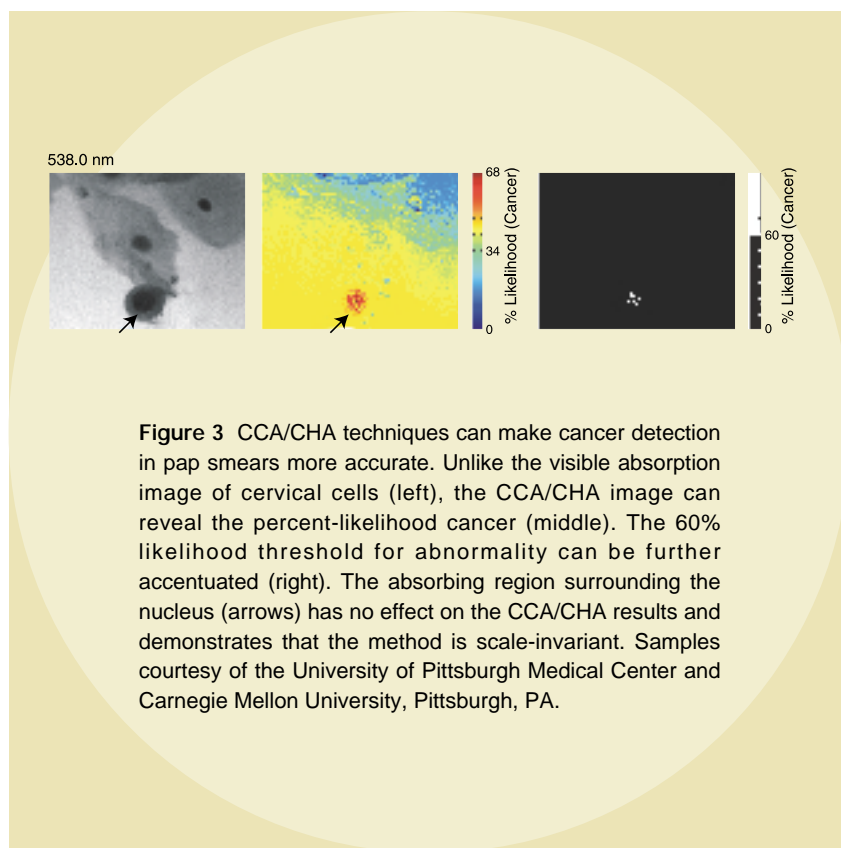


Figure 3 CCA/CHA techniques can make cancer detection in pap smears more accurate. Unlike the visible absorption image of cervical cells (left), the CCA/CHA image can reveal the percent-likelihood cancer (middle). The 60% likelihood threshold for abnormality can be further accentuated (right). The absorbing region surrounding the nucleus (arrows) has no effect on the CCA/CHA results and demonstrates that the method is scale-invariant. Samples courtesy of the University of Pittsburgh Medical Center and Carnegie Mellon University, Pittsburgh, PA.

has been to develop effective strategies for analyzing multispectral chemical image data sets acquired from suspect Pap preparations (see figure 3).

Most of our recent work focuses on Pap smear samples that have been fixed and stained using automated systems. While stained samples are less ideal for chemical imaging analyses, they are more easily obtained from clinical laboratories that sometimes discard them after the normal course of histological examination is completed. Chemical images of the Pap samples are collected using an optical microscope that is coupled to a Fourier transform visible step-scan Sagnac interferometer and CCD detector. Although visible absorption and reflectance spectroscopies lack the chemical specificity of vibrational spectroscopy, they are well suited to the analysis of stained samples.

In order to establish whether the populations of spectra from abnormal and normal nuclei differ, we divide the prepared samples into two groups. The first group contains cells that present histologically normal nuclei, and the second group contains cells with abnormal nuclei. Each group is fur-

ther subdivided into training and validation categories. The wide-field visible absorption spectra from each group are correlated to a predetermined reference spectrum using CCA. In n -dimensional wavelength space, the reference spectrum is chosen to be orthogonal to the average normal nuclei spectrum and pointing in the direction of the average abnormal nuclei spectrum. The advantage of using this carefully chosen reference vector is that larger CCA values will tend to correspond to greater degrees of abnormality. Conversely, normal nuclei will, on average, exhibit no correlation to the reference spectrum.

Once the CCA values for the normal and abnormal nuclei populations are calculated, the next step is to calculate the percent-likelihood values corresponding to each CCA value. Conceptually, this is accomplished by comparing the histograms of CCA results for both groups. The likelihood values are simply the percentage of the spectra corresponding to a particular CCA score that belong to the abnormal training group relative to the total number of occurrences found in both groups. Ideally, the number of

training spectra in each group should be similar and constitute a statistically large data population.

Unknown Pap samples are examined using the same CCA protocol. The resulting CCA image is then remapped as a CHA image by replacing each CCA value with its corresponding percent-likelihood value. The CHA image is a semi-quantitative result that reveals the relative likelihood of nuclear abnormality.

This work is the first step toward developing a diagnostic tool that reduces the need for clinical expertise and tissue biopsies. Although the visible chemical-image data is promising, Raman studies are being initiated on unstained samples to provide enhanced chemical specificity. Future work will focus on the clinical significance of the CCA/CHA results.

The chemical imaging field has begun to experience a radical shift from proof-of-principle-driven growth to application-driven growth. No longer is the issue whether or not chemical imaging is a viable analytical tool. Instead, the question is whether chemical imaging can reveal the still-elusive solutions to historically difficult problems. New instrumentation and expanding application areas are the telltale signs of coming successes. **oe**

John F. Turner II is a professor in the department of chemistry at Cleveland State University, Cleveland, Ohio. Phone: 216-687-2451; fax: 216-687-9298; e-mail: j.f.turner@csuohio.edu.

Acknowledgements

The author expresses appreciation to Mark Sparrow (Intertape Polymer Group, Marysville, MI), Jinming Gao (Case Western Reserve University, Cleveland, OH), and Elliot Wachman (Carnegie Mellon University, Pittsburgh, PA) for providing samples to be analyzed. Additionally, FTIR images were graciously acquired by Michael Schaeberle at the National Institutes of Health. Lastly, a special thanks goes to graduate students Jing Zhang and Anne O'Connor (Cleveland State University) for performing data acquisition and multivariate CCA/CHA processing.

Reference

1. Hannah Morris, John Turner II, et al., Langmuir, 15[8], 1999.