Near Infrared Spectroscopic Imaging: A Paradigm Shift in Quantitative Analysis

E. Neil Lewis, Linda H. Kidder, and Eunah Lee

Spectral Dimensions, Inc., 3416 Olandwood Court, Suite 210, Olney, MD 20832

Data collected using NIR imaging instrumentation has the potential to deliver quantitative information on sample constituents without first creating a calibration model. In contrast to conventional, single point NIR spectrometers, the imaging system uses an infrared focal-plane array (FPA) to generate image cubes and collect up to 76,800 complete spectra, one for each pixel. These images and spectra document the type and distribution of individual chemical constituents comprising the sample. This fact makes spectroscopic imaging uniquely suited for this performing *a priori* quantitative analysis.

An NIR imaging data set of an over-the-counter tablet containing the active pharmaceutical ingredient (API) acetaminophen, was collected in reflectance mode and divided by a 100% reflectance background data set (Spectralon, Lab Sphere, North Sutton, NH) to generate reflectance image cubes. The data were collected using a Matrix NIR imaging system (Spectral Dimensions, Olney, MD) over the spectral range 1100-1700 nm with 10 nm increments. Each data set, containing 76,800 individual spectra, takes approximately 2 minutes to collect. The resulting reflectance image cubes are transformed using $\log10(1/R)$ and then further processed with a two-point baseline correction and a multiplicative scatter correction. Figure 1(a) is an image at 1640 nm of the tablet. The image contrast is based on the strength of the absorption band at 1640 nm, which us assigned to acetaminophen, and so is bright where the concentration of acetaminophen is greatest. The remainder of the tablet is composed of a mixture of several inactive ingredients, called excipients, which do not absorb as strongly at 1640 nm. As the relative strength of different absorption bands vary across the spectral range, images derived from these wavelengths will highlight different chemical components. Representative single pixel spectra extracted from the image cubes corresponding to regions rich in the API and excipient mixture respectively are shown in figure 1(b).

Since bright areas in figure 1(a) indicate the distribution of the API, after determining an appropriate threshold it is possible to estimate its 'abundance' by simply counting bright pixels. Pixels containing mostly the API in this image are determined to have a $\log10(1/R)$ value greater than 0.146. A histogram representation of the image, shown in figure 2, is a convenient way to count these pixels. This graph is a series of bins, in which the *x*-axis is spectral intensity, and the *y*-axis is number of pixels. For the image at 1640 nm, the sum of pixels with a $\log10(1/R)$ threshold value of 0.146 or greater, divided by the total number of pixels in the sample, gives an approximate API concentration. This value was determined to be (5,702/23,674) or 24.09%. This is higher than the manufacturer's specification of 20.89% by weight, but provides a reasonably close approximation even though no *a priori* information was available about the sample. No calibration standards were used and no pure component reference spectra were made available. The estimation is made simply using NIR imaging to provide sufficient chemical contrast to distinguish one component from another and then relying on statistical sampling and particle counting to estimate the relative abundance of each species.

This concept of parallel statistical sampling and particle counting to determine component concentration can be made more robust when pure component spectra are available. By utilizing a multivariate technique such as partial least-squares (PLS) the somewhat subjective, thresholding step needed in the above method to determine population membership can be avoided. For the same sample, a library of the pure component spectrum of acetaminophen and a "pure component" spectrum of the excipients, were recorded and used to perform a PLS analysis on every pixel in the imaging data set. The result of this calculation is that each pixel is now assigned a real concentration for each component rather than a simple binary value (pixel is either API *or* excipient) as determined previously. The total concentration of any component in an imaging data set is then calculated by summing the concentration values for all the pixels in its PLS score image and normalizing for the total number of pixels. Using this approach on the same data set, a concentration value of 21.55% was determined for the acetaminophen, a value within the manufacturers' tolerance for this product. Further details of these novel hybrid approaches for quantitative NIR analysis will be presented.

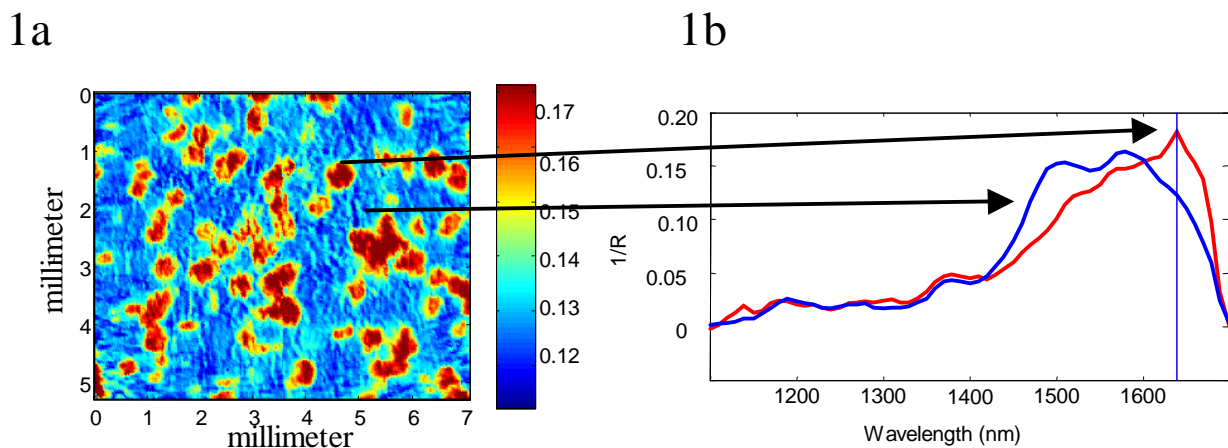1a   1b

Figure 1: NIR spectroscopic image of an over-the-counter pain medication at 1640 nm showing the spatial distribution of the chemical components within the tablet. Bright areas show primarily the active ingredient, acetaminophen. Darker areas correspond to regions rich in excipients. Representative single pixel spectra are shown in 1(b). The solid red spectrum corresponds to the active ingredient, and the dashed blue spectrum corresponds to the excipients.
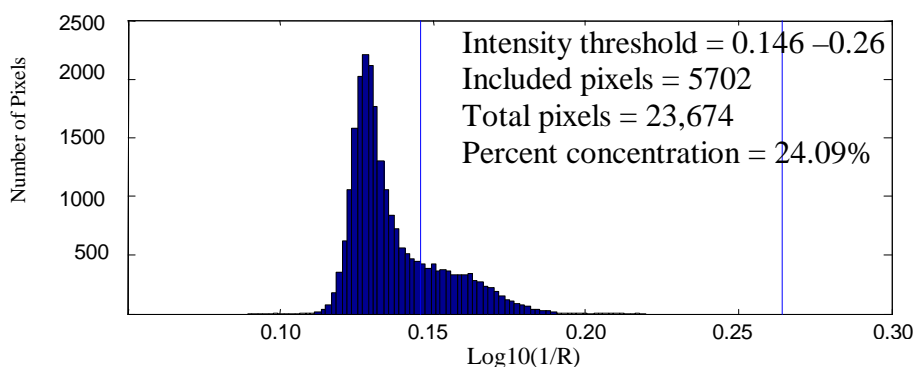

Figure 2: Histogram of a single image plane at 1640nm showing the segmentation of pixels corresponding to populations of the API and the excipient mixture. The dotted lines indicate the threshold range for calculation of the concentration of the API.