Multimodal, multiplex, Raman spectroscopy of alcohol in diffuse, fluorescent media

Scott T. McCain, Michael E. Gehm, Yanqia Wang,
Nikos P. Pitsianis, Michael E. Sullivan and David J. Brady

Duke University Fitzpatrick Center for Photonics and Communication Systems and
Department of Electrical and Computer Engineering, Box 90291, Durham, NC, 27708, USA

ABSTRACT

Optical diagnostics in biological materials are hindered by fluorescence and scattering. We have developed a multimodal, multiplex, coded-aperture Raman spectrometer to detect alcohol in a lipid tissue phantom solution.

Keywords: Raman Spectroscopy, Tissue Spectroscopy, Optical Multiplexing

1. INTRODUCTION

Raman spectroscopy is a powerful diagnostic tool due to its high specificity and possibility for *in vivo* applications. At the same time, its very weak signal strength and incoherent scattering properties make it a challenging signal to detect. In biological systems where components scatter strongly and fluoresce, this detection problem is compounded even further. We have designed and constructed a multimodal multiplex Raman spectrometer which uses coded aperture spectroscopy to have a large collection area and multi-wavelength excitation to better detect signals in the presence of fluorescence. While such spectrometers have been studied before, ours is novel due to its static design and use of a 2-dimensional coding pattern. We have demonstrated the detection of ethanol in solutions of a lipid tissue phantom. At high lipid concentrations, our instrument detects Raman signatures which are not observable in a research grade instrument. Through principle component regression, concentration analysis of alcohol is shown to have the possibility of detection levels as low as 0.01% by volume.

2. LIGHT SCATTERING AND ABSORPTION BY TISSUE

In optical tissue diagnostics, the NIR regime is typically chosen due to the low absorption by tissue components. In Raman spectroscopy it is also advantageous due to the reduced intensity of background fluorescence. In this wavelength region in tissue such as skin, scattering rather than absorption dominates photon transport and thus the "diffusion approximation" can be made. The time-independent diffusion can be expressed as follows[1]:

$$\nabla^2 \Phi_d(r) - \kappa_d^2 \Phi_d(r) = 0 \, , \tag{1}$$

where Φ_d is the photon fluence, and $1/\kappa_d$ is the diffusion length. The diffusion length can be expressed as:

$$\frac{1}{\kappa_d} = \sqrt{\frac{1}{3\mu_a(\mu_a + (1-g)\mu_s)}} \, . \tag{2}$$

The constants for scattering and absorption, μ_s and μ_a respectively, and g, an anisotropic scattering parameter are readily available for various biological components. In order to roughly estimate the photon distribution in tissue, the solution for the diffusion equation for a point source was used:

$$\frac{\Phi_d(r)}{h\nu c_m} = \frac{N\kappa_d^2}{4\pi} \frac{\exp[-\kappa_d r]}{r} \, , \tag{3}$$
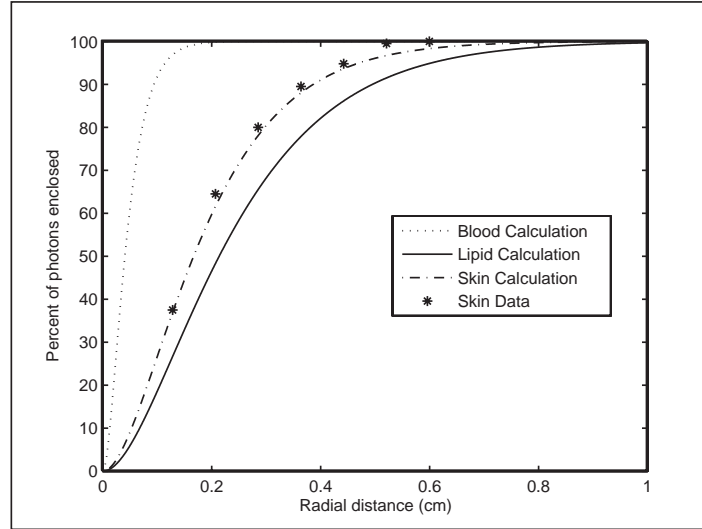
Figure 1. The percentage of enclosed photons as a function of radial distance from a point source is shown above for the scattering-dominant limit in a uniform medium. The tissue phantom has a μ_a of 0.23 cm^{-1}, a μ_s of 217.5 cm^{-1}, and a g of 0.589.[2] The blood model has a μ_a of 12.5 cm^{-1}, a μ_s of 22.2 cm^{-1}, and a g of 0.99.[3] The skin model has a μ_a of 0.63 cm^{-1}, a μ_s of 265 cm^{-1}, and a g of 0.8.

where N is the number of injected photons and c_m is the phase velocity in the medium. This equation is explicitly integrated to determine the approximate photon distribution of light in tissue, blood, and a tissue phantom. The solution can be expressed as:

$$N(r) = [1 - (1 + \kappa_d r) \exp[-\kappa_d r])] \tag{4}$$

To accurately calculate the distribution of Raman photons, a simulation would have to be done including the Raman cross sections of tissue components and the boundary conditions of the medium. To augment the simulation, an experiment was done using a Ti:Sapphire laser and a photodiode array to measure the remitted photon distribution of a human forearm. A calibrated linear array of CCD detectors pushed up against the skin was used to measure the remitted light of the laser at 840 nm. The percentage of the total remitted light distribution is shown, along with calculated photon distributions in Fig. 1.

While in the blood model, the photons stay relatively localized, with 90% of the photons staying within 1 millimeter of the excitation source, in the tissue models and data the spot size of the radiation vastly increases. For the skin like data and model, only 25% of the photons are distributed within 1 millimeter of the source.

Another important aspect of the source is its incoherent nature, meaning it can not be focused or have its local brightness increased.[4] Thus, in order to couple a large source that has been scattered by tissue into a dispersive spectrometer, it must be either spatially filtered, resulting in a loss of photons, or fiber coupled, resulting in fiber coupling losses and increased detector and imaging optics sizes, to have its spectrum measured with a reasonable resolution and collection angle. Previous Raman studies involving blood samples[5] used a parabolic concentrator non-imaging optic to match the étendue of the source to a fiber bundle for a 1 mm^2 surface area. However as can be seen from Fig. 1, when dealing with tissue samples a way to couple a spatially extended diffuse source of several millimeters could drastically increase photon collection and thus allow for systems with shorter exposure times or less costly components.

3. CODED APERTURE SPECTROSCOPY

In order to increase the photon throughput of spectrometers, we have designed systems based on coded apertures which provide high resolution without sacrificing photon throughput. Such systems have been studied before[6,7], however previous system described were either dynamic[8] or poorly conditioned computationally.[9] Our design

utilizes a 2-D code based on the Hadamard matrices, which allows static measurement of $2N$ aperture patterns in order to utilize a size N Hadamard matrix for inversion. Thus, a spectrum can be reconstructed very quickly with high fidelity due to the well-conditioned nature of the Hadamard matrices.

3.1. Spectral Measurement

A spectrometer projects the cross-spectral density onto a set of spectral measurements of the form

$$
\begin{aligned}
m_i &= \int h_i(\nu, \mathbf{r}) W(\mathbf{r}, \mathbf{r}, \nu) d\mathbf{r} d\nu \\
&= \sum_n \int h_i(\nu, \mathbf{r}) S(\nu) |\phi_n(\mathbf{r})|^2 d\mathbf{r} d\nu
\end{aligned}
\tag{5}
$$

where $\mathbf{r}$ is integrated over the surface of the ith detector element.

Spectral resolution is determined by both the spatial and the spectral distributions of $h_i(\nu, \mathbf{r})$. One can limit the impact of the spatial distribution by reducing the spatial size of the detector, but this has the impact of reducing the light gathering capacity of the system. Alternatively, one may attempt to code $h_i(\nu, \mathbf{r})$ to break the spectral-spatial ambiguity and obtain both high spectral resolution and high throughput.

For the coded aperture, dispersive grating spectrometer case in 1-D with dispersion direction x, the intensity on the detector plane can be described by

$$
I(x') = \int \int A(x) \, S(\lambda) \, \delta(x' - (x - \alpha\lambda)) \, dx \, d\lambda ,
\tag{6}
$$

where α is the linear dispersion of the diffraction grating in terms of detector length per wavelength. Simplifying through the delta function term leads to

$$
I(x') = \int A(x) S((x - x')/\alpha) dx .
\tag{7}
$$

For the mask patterns involving transmissive and blocking apertures, the aperture function can be described by

$$
A(x) = \sum_j \delta(x - x_j) ,
\tag{8}
$$

giving a modified detector intensity of

$$
I(x') = \sum_j S((x_j - x')/\alpha) .
\tag{9}
$$

Thus combinations of the spectral density are measured at every detector, allowing for a system with both high resolution and throughput. Deconvolution to estimate the spectral density is possible in this case, however the conditionality of this transformation is very poor. By measuring combinations of codes, A_k, a greater family of transformations is possible, with some being very well conditioned. This can be described by

$$
I_k(x') = \sum_j S((x_{k,j} - x')/\alpha) .
\tag{10}
$$

Inversion of Eq. 10 leads to j spectral estimates corresponding to the aperture positions sampled by $A(x)$.

3.2. Spectrum Reconstruction Techniques

The source spectrum is assumed to be uniform along the non-dispersive direction, and can be treated as a vector $S(\nu)$ with k spectral channels that fall onto the CCD. The width of each of spectral channel on the detector, Δ_k, is used to define the relative widths of the mask elements, Δ_m, and detector elements, Δ_d, where k is defined in Eq. 11, with R being the ratio of mask element width to CCD pixel pitch

$$
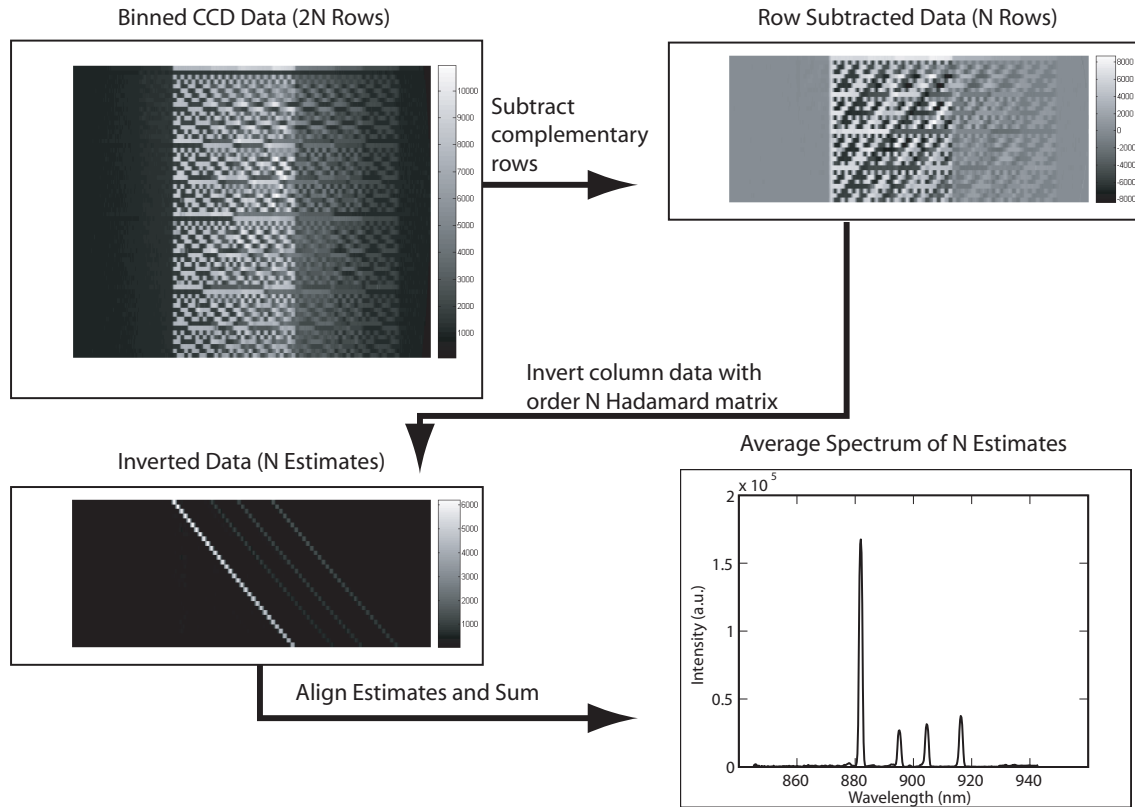k = M + N R .
\tag{11}
$$

Figure 2. Reconstruction process is shown for a xenon discharge lamp. The calibration procedure parameters are saved and used for the reconstruction of Raman spectra.

The inversion problem can be formulated in the following equation, where $S_1 \ldots S_N$ are the unknown vectors of length M, the spectra that would be acquired for a system with only one aperture at each position along the dispersion direction. $R_1 \ldots R_N$ are the data vectors obtained after binning and subtracting the CCD row data as described earlier. This simplifies the problem to M matrix inversions of size $N \times N$. Non-negative enforcing Matlab algorithms were used, since the detector only measures positive counts.

$$\mathbf{H}_{N \times N} \begin{bmatrix} S_1 \\ . \\ . \\ . \\ S_N \end{bmatrix}_{N \times M} = \begin{bmatrix} R_1 \\ . \\ . \\ . \\ R_N \end{bmatrix}_{N \times M} . \tag{12}$$

The vectors $S_1 \ldots S_N$ then form a 1-D spectral image across the dispersion direction of the system. Alternatively, these N vectors can be shifted and then averaged in order to achieve a multiplex gain by averaging the different estimates of the spectrum.

3.3. Principal Component Regression for Concentration Estimation from Raman Spectra

The Principal Component Regression (PCR) method is used for the concentration estimation. PCR consists of two major steps, i) Principal Component Analysis (PCA) of the reconstructed Raman spectra and ii) Multiple Linear Regression (MLR) between the spectral principal components and the concentrations of the training data.

More specifically, let X be the $n \times p$ array consisting of p training spectra sampled at n wavelengths and vector c containing the corresponding p concentrations. Matrix X is row-centered, that is, the measurements at each wavelength have zero means. This is the most common preprocessing procedure for spectroscopic data.

The singular value decomposition (SVD) algorithm is used for the determination of the principal components of X. Let $X = U\Sigma V^T$ be the SVD of X where U and V are orthogonal matrices and Σ is a diagonal matrix with elements on the diagonal $\sigma_1 \geq \sigma_2 \geq \cdots \sigma_p \geq 0$.

The number of principal components r is determined by leave-one-out cross-validation. Let X_r denote the "denoised" data, with

$$X_r = \sum_{k=1}^{r} \sigma_k U_k V_k^T = U_r \Sigma_r V_r^T \ ,$$

with u_k and v_k denoting the kth left and right singular vector (the kth columns of matrices U and V respectively).

MLR between the training spectral principal components and their concentrations determines the PCR vector w that minimizes $\min \|X_r^T w - c\|_2$. The corresponding normal equations are

$$
\begin{aligned}
w &= (X_r X_r^T)^{-1} X_r c & (13) \\
&= (U_r \Sigma_r V_r^T V_r \Sigma_r U_r^T)^{-1} U_r \Sigma_r V_r^T c & (14) \\
&= (U_r \Sigma_r 2 U_r^T)^{-1} U_r \Sigma_r V_r^T c \ . & (15)
\end{aligned}
$$

The concentration estimation of a new spectrum s is determined by the inner product $w^T(s - m)$ where m is the vector of the mean spectral values for each wavelength.

4. PROTOTYPE SYSTEM

For work in tissue, the near-infrared (NIR) region is attractive due to the low photon absorption and availability of high performance CCD detectors in this wavelength range. Excitation at 808 nm provides reduced tissue fluorescence and produces Raman peaks of ethanol at 870-930 nm. With our aperture coding and multi-wavelength excitation, the goal is to have a system that can have good performance in the presence of noise associated with a higher temperature detector and lower grade CCDs.

4.1. Focal Plane Detector

The system is built around an astromonic CCD, with a 765×510 pixel detector cooled by a single-stage Peltier cooler. The CCD is not front-illuminated, but contains a micro-lens array in order to achieve quantum efficiencies of 10-20% in the NIR. In the right of Fig. 3, the complete system is shown. The camera was chosen due to its low cost ($1500), 16-bit digitization, and compact form factor. Typical bench top laboratory Raman spectrometers utilize multi-stage cooling and scientific-grade CCDs which dramatically increase the cost of a system.

4.2. Excitation source

The excitation source consists of 4 single longitudinal and transverse mode semiconductor diode lasers, with peak CW powers of 100 mW, and spectral output of 0.5-1 nm FWHM. The lasers are temperature tuned to allow for peak wavelengths between 803-810 nm. The lasers and small thermistors are thermally epoxied into machined copper mounts, which sit on top of a Peltier cooler. The mount is secured to the copper base with nylon screws to prevent thermal conduction from the top of the Peltier cooler and to allow for good thermal contact between the copper base and the bottom of the cooler. The laser diodes have an epoxied band-pass filter with transmission from 800-820 nm to suppress out of band spontaneous emission and a lens to form a small spot of 100 um on the sample of interest, for this system a liquid filled cuvette.

4.3. Optical system

The optical system is shown in the left of Fig. 3. Due to the isotropy of Raman scattering, a low f/#, 2.1, system was custom designed, which combined with the large input aperture of the coding mask, 4 mm × 2 mm, allows for a high étendue of the entire system. To keep the system compact, short focal lengths are used to keep the lens diameters under 25 mm. The relay optic provides 2:1 magnification of the source onto the mask and a long-pass filter to reject Rayleigh scattered laser light.
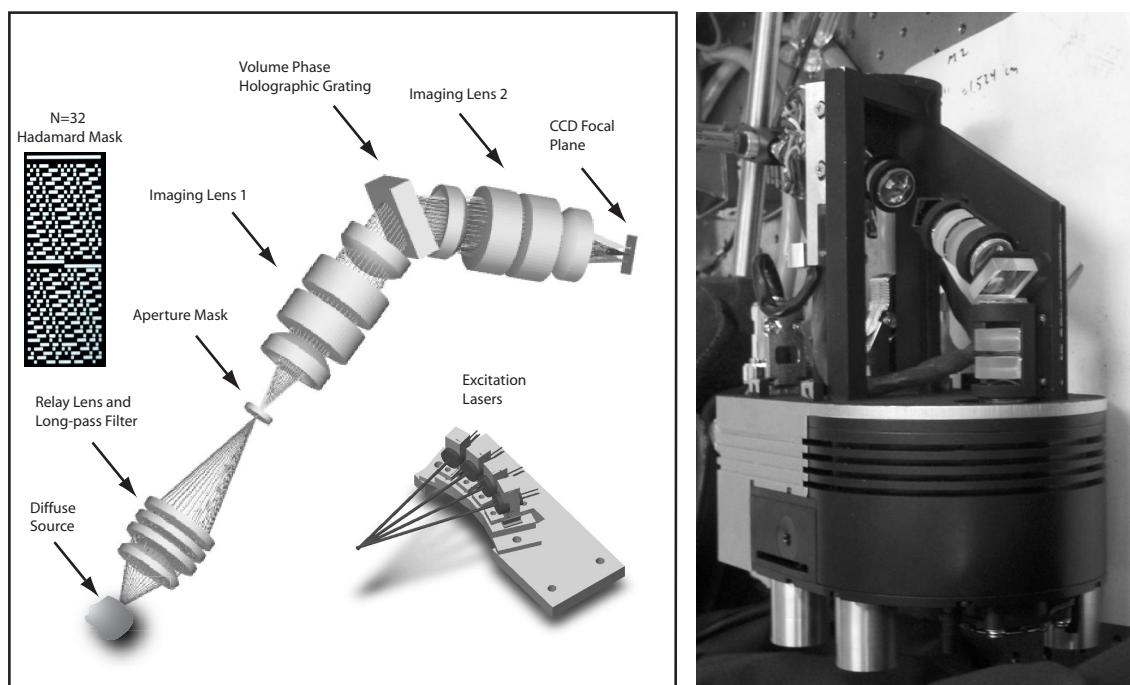
Figure 3. Left, optical system for prototype. Short focal lengths and multiple elements provide spot sizes of the order of the 9 μm CCD pixels with an effective f/# of 2.1 and input aperture size of 4x2 mm. 4 excitation laser beams converge on a sample cuvette in a 45 degree geometry relative to the optical axis of the spectrometer. Right, picture of assembled system with baffles and cover removed. The housing and mounts were rapid prototypes by stereo-lithography.

A volume phase holographic (VPH) grating of 1200 lpmm yields high diffraction efficiencies ($> 80\%$) over the wavelengths and angles present in the system. The system described yields a spectral range of 850-950 nm.

The aperture mask consists of lithographically patterned chrome on quartz, allowing for $> 90\%$ transmission and a blocking optical density (OD) of 4. An $N = 32$ Hadamard code with 36×36 μm apertures allows for a system resolution of 1 nm which is matched to the spectral width of the Raman ethanol peaks, also ~ 1 nm FWHM.

5. PERFORMANCE AND ANALYSIS

5.1. Calibration and Correction with Gas Discharge Lamp

In order to be able to calibrate various parameters of the system, a xenon pen lamp illuminates a diffuser placed where the sample cuvette normally sits. The diffuser is necessary due to the line-like nature of the pen lamp, which creates a non-uniform intensity on the mask plane. Xenon was chosen due to there being 4 strong peaks between 880 and 920 nm, near the observed Raman peaks of ethanol for 808 nm excitation. The narrow linewidths of the emission lines form crisp mask images for each emission line on the detector plane which allow for mask registration in reference to the CCD pixels. Nonlinearity of the dispersed spectra are corrected using the constant spacing of the emission lines. In order to perform vertical registration, the raw CCD data is binned along the dispersion direction into a data vector. Due to the dark rows of the mask, a sinusoidal-like form is seen in the vector. The maxima are calculated and the 2N values are stored.

To correct for the "smile" distortion in the CCD image, rows of the mask pattern that have an open aperture and then a closed aperture at the beginning of the row are found. The corresponding CCD rows are then used to find the position of the left-most maxima of the most intense xenon peak. With no distortion, these maxima would all occur in the same column, however with the "smile" distortion they are displaced quadratically
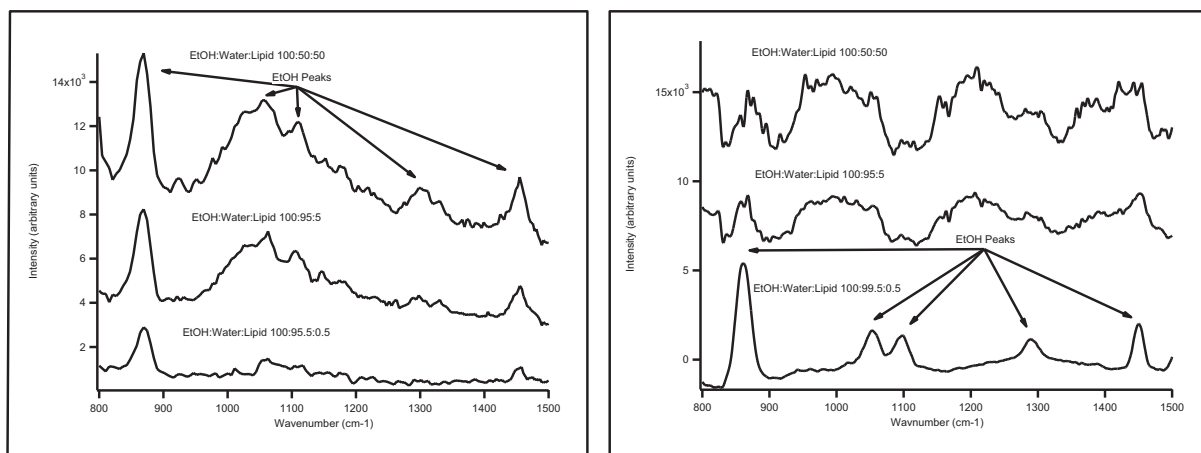
Figure 4. Reconstructed spectra from coded aperture system (left) and commercial fiber-slit based system (right). Spectra are separated vertically for visual purposes only, and lipid concentration increases vertically.

centered approximately at the center row of the CCD. The CCD data is spline interpolated 5× in order to allow for sub-pixel offsets. A polynomial fit is done between the offset and the CCD row. The interpolated data is then shifted row by row, and then an interpolation of 1/5× is done to return to the original number of columns. The shift needed is assumed to be independent of wavelength, which is an approximation, however the different wavelengths are mixed on the detector plane so a physical method such as a prism would be needed to compensate further.

The corrected data is then selectively binned in order to form a $2N$ row data matrix. The ratio of the mask aperture height to CCD pixel height, R, is known from the design of the mask, and can be scaled by the magnification of the 4-f system which in this case is 1. The rows used for the binning are the CCD rows that correspond to the mask rows $\pm R/2$.

5.2. Single Wavelength Excitation Comparison

To evaluate the performance of the system, samples of varying tissue phantom solutions are used to simulate a diffuse, fluorescent source similar to tissue. For comparison, samples are also tested on a commercial system with a fiber bundle that converts a circular input to a rectangular slit-like source and is then imaged through a slit and dispersed onto a high-performance CCD detector. The resultant spectra at three different tissue phantom concentrations are shown in Fig. 4. Spectra from the fiber system has to be background subtracted due to their high fluorescence background. This process led to difficulty in resolving the spectra at high lipid concentrations due to the extremely large fluorescence signal. The coded aperture system's spectra are shown without any post-processing after the reconstruction described earlier. The greater collection efficiency and coded-aperture sampling technique are assumed to be the reason for the higher fidelity of the Raman ethanol peaks in the higher concentrations of the lipid tissue phantom. The reduced signal strength for the low concentration lipid in the coded-aperture system is attributed to the non-uniform filling of the mask with a weakly scattering medium.

5.3. Multi-wavelength Concentration Estimation

To test the detection limits of the system, a series of measurements on three different sample solutions were investigated- a microsphere solution with weak scattering and fluorescence, a blood-like lipid solution of 10% Intra-Lipid solution, and a tissue-like lipid solution of 20% Intra-Lipid solution. Eleven different concentrations of 20%, 10%, 5%, 2%, 1%, 0.5%, 0.2%, 0.1%, 0.05%, 0.02%, and 0.01% ethanol by volume were tested. All 10 possible single and dual laser excitation schemes were tested, with six exposures at each laser setting over the course of two days. To reduce the effect of laser drift and mode-hopping, a short exposure is performed before and after each exposure to verify the laser has not changed wavelengths by correlating the very edge of the
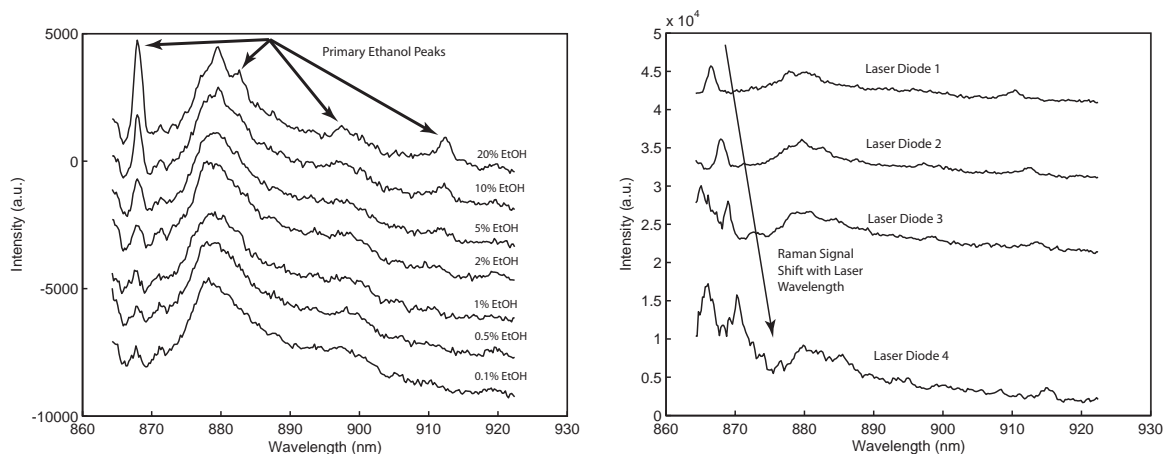
Figure 5. Left, Raman spectra at varying concentrations for one laser at varying concentrations of ethanol in a tissue-like lipid solution. Right, Raman spectra at 20% concentration of each of the 4 laser diodes. The lasers are temperature tuned to set their wavelengths 1 nm apart.
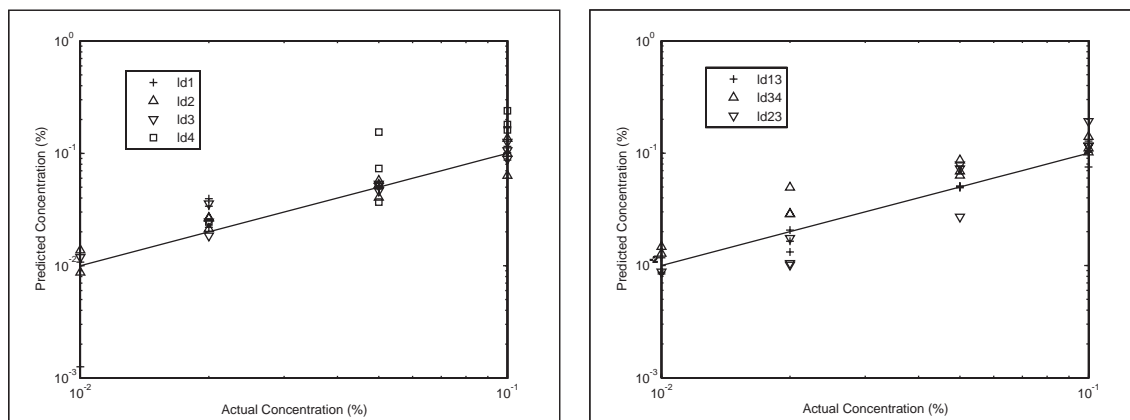

Figure 6. Left, PCR training data showing predicted and estimated ethanol concentration with each laser's data processed independently. Right, PCR training data using three different dual laser excitations independently.

Rayleigh scattered light that is dispersed onto the detector plane. Fig. 5 shows reconstructed spectra at various laser excitations and ethanol concentrations.

The singular value decomposition algorithm described is applied to the data sets to determine the principal components. Using leave-one-out cross validation on each set individually an estimator is determined, and then is applied to the original data set to observe the quality of the fit. Plots of actual versus estimated concentrations using the estimator are shown in Fig. 6. The quality of the fits for a single data set show promise for detection limits at 0.01%, corresponding to 7.89 mg/dL of ethanol. Future work will test the algorithms with cross-validation between different data sets, and the use of combinations of laser excitations to better estimate concentrations.

6. CONCLUSIONS

A coded-aperture, multi-wavelength Raman spectrometer has been demonstrated to detect alcohol in trace concentrations in lipid tissue phantoms. Through the use of a large input aperture, a custom optical system, and a shifted excitation scheme, Raman signals have been extracted. In highly scattering media, the performance is

shown to be greater than that of fiber-coupled slit-based systems. Repeatability and reliability of the measurements still need to be demonstrated, along with tests of samples with a more complex background than phantom solutions. With more sophisticated algorithms, the combination of information from multiple laser excitations should lead to a lower detection limit and possibilities of low cost, *in vivo* non-invasive optical sensors for alcohol and various other chemicals found in the human body.

7. ACKNOWLEDGMENTS

REFERENCES

1. T. Vo-Dinh, *Biomedical photonics handbook*, CRC Press, Boca Raton, Fla., 2003.
2. H. J. Vanstaveren, C. J. M. Moes, J. Vanmarle, S. A. Prahl, and M. J. C. Vangemert, "Light-scattering in intralipid-10-percent in the wavelength range of 400-1100 nm," *Applied Optics* **30**(31), pp. 4507–4514, 1991.
3. A. Roggan, M. Friebel, K. Dorschel, A. Hahn, and G. Muller, "Optical properties of circulating human blood in the wavelength range 400-2500 nm," *Journal of Biomedical Optics* **4**(1), pp. 36–46, 1999.
4. D. J. Brady, "Multiplex sensors and the constant radiance theorem," *Optics Letters* **27**(1), pp. 16–18, 2002.
5. A. M. K. Enejder, T. W. Koo, J. Oh, M. Hunter, S. Sasic, M. S. Feld, and G. L. Horowitz, "Blood analysis by raman spectroscopy," *Optics Letters* **27**(22), pp. 2004–2006, 2002.
6. M. J. E. Golay, "Multi-slit spectrometry," *JOSA* **39**(6), pp. 437–444, 1949.
7. M. Harwit and N. J. A. Sloane, *Hadamard transform optics*, Academic Press, New York, 1979.
8. R. A. DeVerse, R. M. Hammaker, W. G. Fateley, J. A. Graham, and J. D. Tate, "Spectrometry and imaging using a digital micromirror array," *American Laboratory* **30**(21), p. 112, 1998.
9. S. B. Mende, E. S. Claflin, R. L. Rairden, and G. R. Swenson, "Hadamard spectroscopy with a 2-dimensional detecting array," *Applied Optics* **32**(34), pp. 7095–7105, 1993. Article.