

Denoising and Deblurring of Fourier Transform Infrared Spectroscopic Imaging Data

Tan H. Nguyen^{c,d}, Rohith K. Reddy^{a,d}, Michael J. Walsh^{a,d}, Matthew Schulmerich^{a,d},
Gabriel Popescu^{c,d}, Minh N. Do^{a,c,d} and Rohit Bhargava^{a,b,c,d,e}

^aDepartment of Bioengineering, Urbana, IL, USA 61801;

^bDepartment of Mechanical Science and Engineering, IL, USA 61801;

^cDepartment of Electrical and Computer Engineering, Urbana, IL, USA 61801;

^dBeckman Institute for Advanced Science and Technology, Urbana, IL, USA 61801;

^eUniversity of Illinois Cancer Center, University of Illinois at Urbana-Champaign,

Urbana, IL, USA 61801

ABSTRACT

Fourier transform infrared (FT-IR) spectroscopic imaging is a powerful tool to obtain chemical information from images of heterogeneous, chemically diverse samples. Significant advances in instrumentation and data processing in the recent past have led to improved instrument design and relatively widespread use of FT-IR imaging, in a variety of systems ranging from biomedical tissue to polymer composites. Various techniques for improving signal to noise ratio (SNR), data collection time and spatial resolution have been proposed previously. In this paper we present an integrated framework that addresses all these factors comprehensively. We utilize the low-rank nature of the data and model the instrument point spread function to denoise data, and then simultaneously deblur and estimate unknown information from images, using a Bayesian variational approach. We show that more spatial detail and improved image quality can be obtained using the proposed framework. The proposed technique is validated through experiments on a standard USAF target and on prostate tissue specimens.

Keywords: FT-IR spectroscopic imaging, linear mixture model, optics modeling, deconvolution, mid-infrared spectroscopy

1. INTRODUCTION

Fourier transform infrared (FT-IR) spectroscopic imaging combines the benefits of optical microscopic imaging with the chemical selectivity of spectroscopy. Spectral recording in this modality is similar to that of conventional spectroscopy and rigorous models^{1,2} are available to understand these spectra. However, the quality of microscope images from absorbance data typically lags that of optical microscopy for a few important reasons. First, the physics of imaging at longer wavelengths in the infrared (IR) is not conducive to high spatial detail. Recent efforts^{3,4} have reported the highest image quality obtainable from diffraction limited mid-infrared imaging systems. Second, focal plane array (FPA) detectors in the mid-IR are of considerably smaller size and lower sensitivity than those available in the visible spectral range. Although the field-of-view limitations can be addressed by raster scanning the sample, the influence of digitization, noise and optical effects on the quality of images need careful examination. In this paper, a framework incorporating all these effects is presented to improve the quality of images from FT-IR imaging.

We utilize the fact that the recorded data is low-rank. This means that the data can be represented as a linear combination of a small number of basic components.^{5,6} This is due to the low number of constituent endmembers at every pixel location. Although the data is of low rank, the presence of noise makes estimation of the rank imprecise. Thus, our first step is to denoise the data. In section 2, we adapt a hyper-spectral data model developed by Akgun⁷ and propose a method of denoising FT-IR data by projecting the acquired data onto a low dimension subspace spanned by absorbance spectra of constituent endmembers found by the principal component analysis

Corresponding author: Dr. Rohit Bhargava: E-mail: rxb@illinois.edu, Telephone: 1 217 265 6596

(PCA) algorithm. Next, we model the denoised data as a convolution of high resolution data with a blurring kernel, or a point spread function. This kernel depends not only on limited-aperture optical diffraction but also on spatial digitization. In section 3, we propose a variational Bayesian deconvolution algorithm based on the closed form of the blurring kernel. The algorithm is capable of estimating pixels' dimensions, prior distributions' unknown parameters of the high resolution image and the model mismatching term at each wavenumber. In subsequent sections, we validate our proposed algorithm by presenting experimental results for denoising and deblurring FT-IR imaging data of a USAF target and of a prostate tissue sample. Finally, conclusion and further research directions are indicated.

2. OPTICAL SETUP AND MODELING

2.1 Optical Setup

The optical set up for an FT-IR imaging system is shown in Figure 1. In this system, light from a broad-band source is divided into two paths using a beam splitter. Upon recombination, the beam from one arm of the Michelson interferometer interferes with its path-delayed counterpart in the other arm. The resulting beam then passes through a sample where it is attenuated due to the sample's absorbance. This beam is finally focused on the detector.

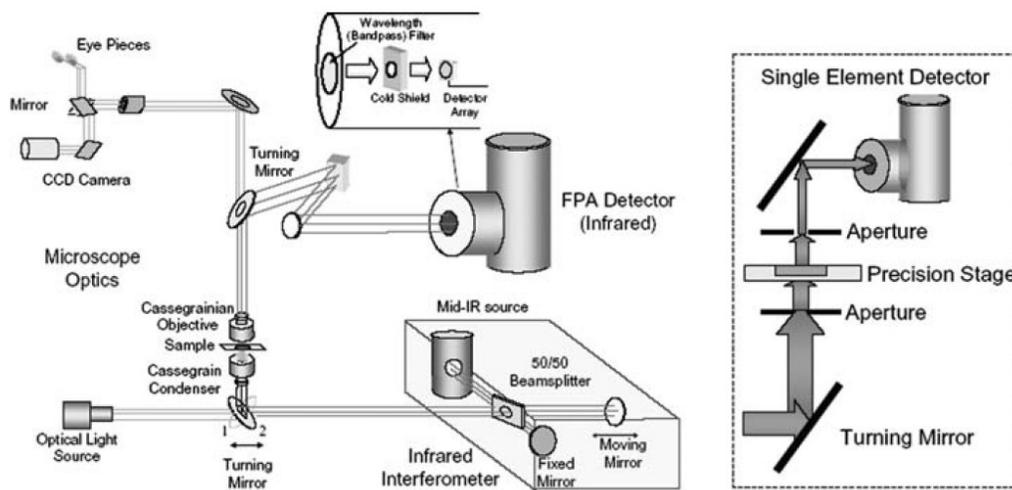


Figure 1. Conventional FT-IR optical setup with FPA detector (left) and single element detector (right) [Figure reproduced from Bhargava.⁸]

For a single element detector, there is a Fourier relation⁹ between the detected light intensity $I(\delta)$ as a function of path difference δ and the spectrum of light $S(\bar{\nu})$ incident on the detector. Here, $\bar{\nu}$ denotes wavenumber.

$$I(\delta) = \int_{-\infty}^{+\infty} S(\bar{\nu}) \cos(2\pi\bar{\nu}\delta) d\nu. \quad (1)$$

The spectrum can be obtained from the recorded data $I(\delta)$ using the relation

$$S(\bar{\nu}) = 2 \int_{-\infty}^{+\infty} I(\delta) \cos(2\pi\bar{\nu}\delta) d\delta. \quad (2)$$

In performing an experiment, the spectrum in the presence of the sample $S_{spl}(\bar{\nu})$ and the spectrum in the absence of a sample (i.e. background) $S_{bg}(\bar{\nu})$ are obtained. The absorbance of the sample at wavenumber $\bar{\nu}$ is defined as

$$A(\bar{\nu}) = \log_{10}(S_{bg}(\bar{\nu})/S_{spl}(\bar{\nu})). \quad (3)$$

2.2 Data model

An adapted hyper-spectral data model developed by Akgun, et. al.⁷ is applied to FT-IR imaging as shown in Figure 2. We make the following assumptions. First, it is possible to recover the continuous absorbance signal $A_{corr,b}(x, \bar{\nu})$ from data if the concentration images are band limited and appropriately discretized. Second, sampling the continuous absorbance spectra will give us discrete noisy images at L bands $g_1[n], g_2[n], \dots, g_L[n]$. Third, these noisy images belong to a low dimensional subspace whose basis is constructed from discrete concentration images, the reconstruction filter, the constituent spectra and the optical blurring kernel. This is under the approximation that the noise images result from a convolution between the reconstruction filter and the optical blurring kernel.

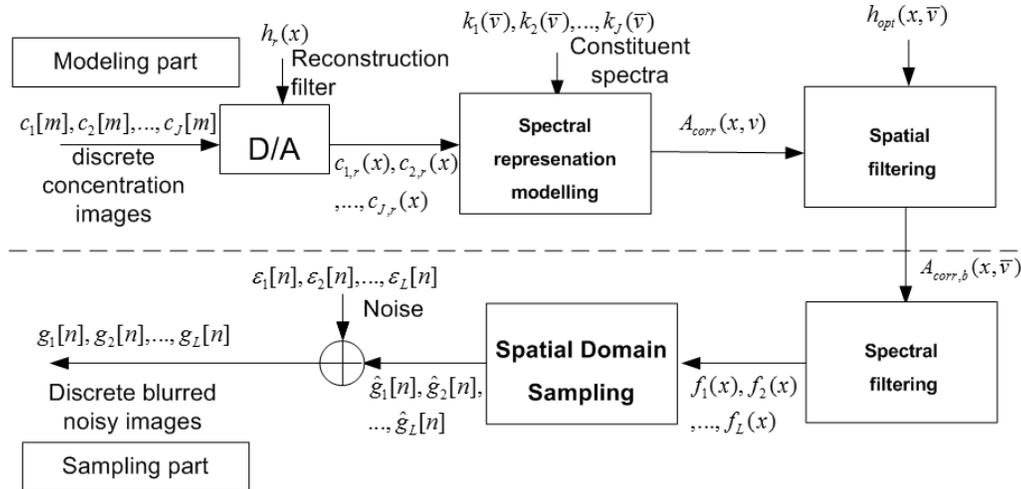


Figure 2. Adapted hyper-spectral data model for FT-IR imaging

Using results from Akgun, et. al.,⁷ the relation between each a discrete blurred and noisy image at band i , which is $f_i[n]$, and the discrete concentration images $c_1[m], c_2[m], \dots, c_J[m]$ can be written as

$$g_i[n] \approx \sum_{j=1}^J w_{j,i} \left(\sum_m c_j[m] h_b[n, m, \mu_{\bar{\nu}}] \right) + \epsilon_i[n], \quad (4)$$

where $w_{j,i} = \langle \mathbf{k}_j, \mathbf{r}_i \rangle$ expresses the effects of the filter of band i to absorbance spectra of endmember j . $\langle \mathbf{a}, \mathbf{b} \rangle$ denotes the dot product of the vector \mathbf{a}, \mathbf{b} . $h_b[n, m, \bar{\nu}] = h_r(n\Delta - m\Delta) * h_{opt}(n\Delta, \bar{\nu})$ denotes an approximation of the convolution, where h_r, h_{opt} represents the reconstruction filter, the optical blurring kernel respectively and $\bar{\nu}$ has been replaced by its mean $\mu_{\bar{\nu}}$ over the wavelength region under consideration. The above approximation is justified since the width of the main lobe $h_{opt}(x, \bar{\nu})$ does not change significantly with wavenumber. For example, moving from 3500 cm^{-1} to 4000 cm^{-1} would result in a change of $\approx 14\%$ in the width the main lobe (from $0.57 \mu\text{m}$ to $0.5 \mu\text{m}$).

Eqn. (4) can be written in matrix form as follows.

$$\mathbf{G}_i \approx (\mathbf{H}_b \mathbf{C}) \mathbf{W}_i + \mathbf{E}_i, \quad (5)$$

where $\mathbf{G}_i = [g_i[1], g_i[2], \dots, g_i[N]]^T$, $\mathbf{E}_i = [\epsilon_i[1], \epsilon_i[2], \dots, \epsilon_i[N]]^T$, $\mathbf{H}_b \in \mathbb{R}^{N^2 \times N^2}$, $\mathbf{C} \in \mathbb{R}^{N^2 \times J}$, $\mathbf{W}_i \in \mathbb{R}^{J \times 1}$ and N is the number of pixels. Note that in the absence of the error term \mathbf{E}_i , \mathbf{G}_i stays in the column space $\mathfrak{R}(\mathbf{H}_b \mathbf{C})$ of $\mathbf{H}_b \mathbf{C}$. It can be seen that the dimension of this subspace is upper-bounded by the number of endmembers J which is much smaller than the number of bands L indicating the low-rank nature of the problem. Next, we proposed an orthogonal-projection based denoising algorithm to utilize these low-rank properties.

2.3 Orthogonal-Projection based Denoising

The matrix $\hat{\mathbf{G}}_i = [\hat{g}_1, \hat{g}_2, \dots, \hat{g}_L]$ in (5) is a low-rank approximation of the original matrix of discrete noisy images $\mathbf{G}_o = [g_1, g_2, \dots, g_L]$ since each of its columns belongs to the column space of $\mathbf{H}_b \mathbf{C}$. To denoise images, we obtain $\hat{\mathbf{G}}$ from \mathbf{G}_o by utilizing its low-rank estimate, which minimizes the variance of projected data. The estimated basis s_1, s_2, \dots, s_J is obtained from PCA analysis. The low-rank estimate $\hat{\mathbf{G}}$ is given as

$$\hat{\mathbf{G}} = \mathbf{S} \mathbf{S}^\dagger \mathbf{G}_o, \quad (6)$$

where $\mathbf{S} = [s_1, s_2, \dots, s_J]$, the pseudo-inverse \mathbf{S}^\dagger is defined as $\mathbf{S}^\dagger = (\mathbf{S}^T \mathbf{S})^{-1} \mathbf{S}^T$.

2.4 Point Spread Function

In the above section, denoising was performed and a low-rank approximation to absorbance spectra was obtained based on the assumption that the blurring matrix \mathbf{H}_b is wavelength independent within a small wavenumber range. However, to obtain high-resolution images at every band, the blurring matrix \mathbf{H}_b is required to be wavelength dependent. To perform deblurring, we make following assumption. The denoised image \hat{g}_i at band i is expressed as

$$\hat{\mathbf{G}}_i = \mathbf{H}_{b,i} \mathbf{G}_{hi,i} + \mathbf{T}_i, \quad (7)$$

where $\mathbf{G}_{hi,i}$ and \mathbf{T}_i are the matrix form of the high resolution image and the modeling error term at band i respectively. Equivalently, we can rewrite equation (7) in convolution form as

$$\hat{g}_i[m] = h_{opt,i}[m] * g_{hi,i}[m] + t_i[m], \quad (8)$$

Using Fourier optics,¹⁰ the point spread function for an incoherent source at band i is

$$h_{opt,i}[m] = \left\{ \begin{array}{l} NA_{out,l}^2 \text{jinc} \left[2\bar{\nu}_i \Delta N A_{out,l} (m_1^2 + m_2^2)^2 \right] \\ -NA_{in,l}^2 \text{jinc} \left[2\bar{\nu}_i \Delta N A_{in,l} (m_1^2 + m_2^2)^2 \right] \end{array} \right\}^2, \quad (9)$$

where $\bar{\nu}_i$ is the central wavenumber of band i , $\text{jinc}(x) = J_1(\pi x)/(2x)$ in which $J_1(x)$ is the first order Bessel function of the first kind, Δ is the sampling interval, which depends on the net magnification factor, and NA_{in}, NA_{out} denote the inner and outer numerical aperture of the objective.

It has been shown⁴ that the quality of recorded spectroscopic imaging data $f_i s$ depends on the resolution of the system, which is related to the numerical aperture, and the sampling interval Δ . Thus, an effective deconvolution algorithm to recover high resolution image needs to take all the above factors into account. Such an algorithm is presented in the next section.

3. VARIATIONAL BAYESIAN DEBLURRING

A method to recover the “deblurred” image $g_{hi,i}[m]$ using the PSF $h_{opt,i}[m]$ and denoised images $\hat{g}_i[m]$ is presented here. Since $\mathbf{H}_{b,i}$ depends on Δ , we can rewrite equation (7) as:

$$\hat{\mathbf{G}}_i = \mathbf{H}_{b,i}(\Delta) \mathbf{G}_{hi,i} + \mathbf{T}_i, \quad (10)$$

Here, we adapt a blind deconvolution framework^{11,12} to simultaneously estimate the sampling interval Δ , the deblurred image $\mathbf{G}_{hi,i}[m]$ and the unknown parameters of the prior distributions of $\mathbf{G}_{hi,i}[m]$ and $\mathbf{T}_i[m]$. We use the Simultaneous Auto-Regression (SAR)¹³ for the distribution of $\mathbf{G}_{hi,i}$ and a degenerate distribution for the sampling interval Δ . The error term $\mathbf{T}_i[m]$ is modeled as a zero-mean multivariate Gaussian distribution. Mathematically, expressions of these distributions can be written as

$$\begin{aligned} p(\mathbf{G}_{hi,i} | \alpha_{\mathbf{G}_{hi,i}}) &\propto \alpha_{\mathbf{G}_{hi,i}}^{(N-1)/2} e^{-(1/2)\alpha_{\mathbf{G}_{hi,i}} \|\mathbf{C} \mathbf{G}_{hi,i}\|^2}, \\ p(\mathbf{H}_{b,i} = \mathbf{H}_{b,i}(\Delta) | \Delta) &= 1, \\ p(\mathbf{T}_i | \Delta, \alpha_{\mathbf{G}_{hi,i}}, \beta) &\propto \beta^{N/2} e^{-\beta \|\hat{\mathbf{G}}_i - \mathbf{H}_{b,i}(\Delta) \mathbf{G}_{hi,i}\|^2}, \end{aligned} \quad (11)$$

where C is the Laplacian operator, $\alpha_{\mathbf{G}_{hi,i}}, \beta, \Delta$ are random variables defining these distribution. They are assumed to have following prior distributions

$$\begin{aligned} p(\alpha_{\mathbf{G}_{hi,i}}) &= \Gamma(\alpha_{\mathbf{G}_{hi,i}} | a_{\alpha_{\mathbf{G}_{hi,i}}}, b_{\alpha_{\mathbf{G}_{hi,i}}}), \\ p(\beta) &= \Gamma(\beta | a_{\beta}, b_{\beta}), \\ p(\Delta) &= N(\Delta | \mu_{\Delta}, \sigma_{\Delta}^2), \end{aligned} \quad (12)$$

where $\Gamma(\cdot), N(\cdot)$ denote the Gamma and Gaussian distribution. Here, $\mu_{\Delta} = 5.5\mu m$ for the setup with the 15x magnification and 0.5 N.A collecting objective and $\mu_{\Delta} = 1.1\mu m$ for that with the 74x magnification and 0.65 N.A collecting objective.

Here, we define $\Theta = (\mathbf{G}_{hi,i}, \beta, \alpha_{\mathbf{G}_{hi,i}}, \Delta)$ be a set of deblurred and unknown parameters to be estimated and the joint probability distribution is defined as:

$$p(\Theta, \hat{\mathbf{G}}_i) = p(\hat{\mathbf{G}}_i | \beta, \alpha_{\mathbf{G}_{hi,i}}, \Delta, \mathbf{G}_{hi,i}) p(\mathbf{G}_{hi,i} | \alpha_{\mathbf{G}_{hi,i}}) p(\Delta) p(\beta) p(\alpha_{\mathbf{G}_{hi,i}}). \quad (13)$$

The inference is based on the posterior distribution $p(\Theta | \hat{\mathbf{G}}_i) = p(\Theta, \hat{\mathbf{G}}_i) / p(\hat{\mathbf{G}}_i)$ which cannot be obtained in closed form. Therefore, we resort to the Variational distribution approximation to a simpler distribution $q(\Theta) = q(\mathbf{G}_{hi,i}) q(\Delta) q(\beta) q(\alpha_{\mathbf{G}_{hi,i}})$ such that the Kullback-Leiber divergence $C_{KL}(q(\Theta) || p(\Theta | \mathbf{G}_i))$ is minimized. Following previously described approach, we obtain following algorithm to simultaneously estimate the deblurred image and unknown parameters:

1. Initialize parameters of the prior distributions $(\alpha_{\mathbf{G}_{hi,i}}, b_{\alpha_{\mathbf{G}_{hi,i}}}), (a_{\beta}, b_{\beta}), (\mu_{\Delta}, \sigma_{\Delta}^2)$, initial estimations of deblurred image $\mathbf{G}_{hi,i}$ (maybe chosen to be the blurred image) and those of random variables $\beta^o, \alpha_{\mathbf{G}_{hi,i}}^o, \Delta^o$ where the superscripts are used to denote the iteration index.
2. For each $k = 0, 1, 2, \dots$, sequentially update the distributions

- Update $q^{k+1}(\mathbf{G}_{hi,i}) = N(\mathbf{G}_{hi,i} | E^{k+1}(\mathbf{G}_{hi,i}), \text{cov}^{k+1}(\mathbf{G}_{hi,i}))$,
update the estimation $\mathbf{G}_{hi,i}^{k+1} := E^{k+1}(\mathbf{G}_{hi,i})_{q^{k+1}(\mathbf{G}_{hi,i})}$, where

$$\begin{aligned} E^{k+1}(\mathbf{G}_{hi,i}) &= \text{cov}^{k+1}(\mathbf{G}_{hi,i}) \beta^k \mathbf{H}_{b,i}(\Delta^k)^T \hat{\mathbf{G}}_i, \\ \text{cov}^{k+1}(\mathbf{G}_{hi,i}) &= \left(\beta^k \mathbf{H}_{b,i}(\Delta^k)^T \mathbf{H}_{b,i}(\Delta^k) + \alpha_{\mathbf{G}_{hi,i}}^k C^T C \right)^{-1}, \end{aligned} \quad (14)$$

- Update $q^{k+1}(\Delta) = N(\Delta | E^{k+1}(\Delta), \text{cov}^{k+1}(\Delta))$,
update the estimation $\Delta^{k+1} := E^{k+1}(\Delta)_{q^{k+1}(\Delta)}$, where

$$\begin{aligned} E^{k+1}(\Delta) &= [(\mu_{\Delta^k} / \sigma_D^2) + \varphi'(\Delta^k) - \varphi''(\Delta^k) \Delta^k] / [1 / \sigma_D^2 - \varphi''(\Delta^k) \Delta^k], \\ \text{cov}^{k+1}(\Delta) &= [1 / \sigma_D^2 + \varphi''(\Delta^k) \Delta^k]^{-1} \\ \varphi(\Delta^k) &= -0.5 \beta^k \left\| \hat{\mathbf{G}}_i - \mathbf{H}_{b,i}(\Delta^k) \mathbf{G}_{hi,i}^{k+1} \right\|^2 - 0.5 \text{trace} \left[\mathbf{H}_{b,i}(\Delta^k)^T \mathbf{H}_{b,i}(\Delta^k) \text{cov}(\mathbf{G}_{hi,i}^{k+1}) \right] \end{aligned} \quad (15)$$

- Update $q^{k+1}(\alpha_{\mathbf{G}_{hi,i}}) = \Gamma \left(\alpha_{\mathbf{G}_{hi,i}} \left| \begin{array}{l} a_{\alpha_{\mathbf{G}_{hi,i}}} + N/2, b_{\alpha_{\mathbf{G}_{hi,i}}} + 0.5 \left\| C \mathbf{G}_{hi,i}^{k+1} \right\|^2 \\ 0.5 \text{trace}(C^T C \text{cov}^{k+1}(\mathbf{G}_{hi,i})) \end{array} \right. \right)$,
update the estimation $\alpha_{\mathbf{G}_{hi,i}}^{k+1} := E^{k+1}(\alpha_{\mathbf{G}_{hi,i}})_{q^{k+1}(\alpha_{\mathbf{G}_{hi,i}})}$.

- Update $q^{k+1}(\beta) = \Gamma \left(\beta \left| \begin{array}{l} a_{\beta} + N/2, b_{\beta} + 0.5 \left\| \mathbf{G}_i - \mathbf{H}_{b,i}(\Delta^{k+1}) \mathbf{G}_{hi,i}^{k+1} \right\|^2 \\ + 0.5 \text{trace}(\mathbf{H}_{b,i}(\Delta^{k+1})^T \mathbf{H}_{b,i}(\Delta^{k+1}) \text{cov}^{k+1}(\mathbf{G}_{hi,i})) \end{array} \right. \right)$,
update the estimation $\beta^{k+1} := E^{k+1}(\beta)_{q^{k+1}(\beta)}$.

4. EXPERIMENTS AND SIMULATIONS

A Varian 7000 Spectrometer coupled to UMA-400 microscope was used to perform experiments with parameters listed in Table 1. The interferometer is operated in the step-scan mode at a stepping rate of 200 Hz. Data were acquired at every other zero-crossing of a He-Ne laser for a free-scanning spectral range of $7900 - 0 \text{ cm}^{-1}$. A Fourier transform of the recorded data was carried out after using the Norton-Beer medium apodization function. Data were truncated and stored as absorbance after a ratio against an appropriate background data. Simulations were carried out using the same set of parameters.

A standard polymer-substrate USAF 1951 target was used as a sample for the first experiment. Prostate tissue microarrays (TMAs) were used as a platform for high throughput sampling⁸ in the second experiment. TMAs consist of a large number of small tissue sections arranged in a grid pattern on a substrate (BaF₂).

Parameter	Value
$NA_{out}^{Condenser}$	0.5
$NA_{out}^{Objective}$	0.5
$NA_{out}^{Detector}$	0.5
Magnification of Condenser	15×
Magnification of Objective	15×
Magnification of Detector	3×
Pixel Size (effective)	$5.5\mu\text{m}$
Spectral Resolution	2cm^{-1}
# Scans per pixel	8
# Scans (Background)	128
Detector size	128×128
Undersampling ratio	2

Table 1. Experimental parameters used for data acquisition.

5. RESULTS AND DISCUSSION

5.1 Experimental results with a USAF target

In order to validate the algorithm, we perform experiments using a commercial FT-IR imaging system. The first experiment is conducted on a USAF target and uses parameters listed in table 1. An absorbance image from acquired data at 1008cm^{-1} and that obtained after processing are shown in Figure 3A and Figure 3B respectively. The number of constituents endmembers is chosen to be 4. It may be observed qualitatively that the deblurred image in Figure 3B is of higher quality than Figure 3A. To quantify performance of the algorithm across different spatial frequencies, we calculate the absorbance contrast functions of the acquired and processed images over different spatial frequencies. The absorbance contrast function is defined as the ratio of the difference between the absorbance of the bar and the absorbance of the region between the bar to the maximum absorbance of each image. Values of the absorbance contrast functions of the acquired and processed images are shown in Figure 3C. The processed image have higher value of the absorbance contrast function over all spatial frequency demonstrating that the deblurring algorithm improves contrast, which is an important measure of image quality.

5.2 Experimental results with biological samples

A second set of experiments were performed using a prostate tissue (TMA) sample using the same experimental configuration as the previous section. Results from this experiment are presented in Figure 4. Absorbance images from acquired data at 1008cm^{-1} , 1052cm^{-1} , 1650cm^{-1} , and 3308cm^{-1} are shown in Figure 4A. Corresponding images after deblurring are presented in 4B. A qualitative improvement in image quality is evident while comparing Figures 4A and B. We have used 5 constituent endmembers for denoising. All parameters of the prior Gamma distributions are set to 1 for deconvolution. The prior distribution of the sampling interval Δ is chosen to be $N(5.5, 0.01)$. Estimated sampling interval is $5.47\mu\text{m}$. Figure 4C shows the used discrete point spread func-

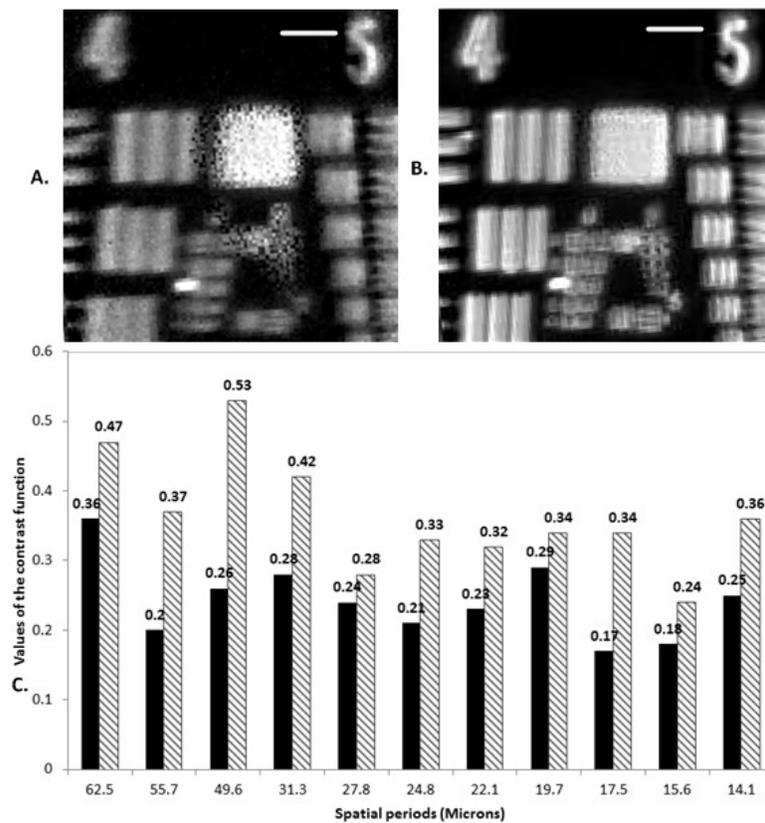


Figure 3. A. Acquired absorbance image at 1008cm^{-1} , B. Processed absorbance image at 1008cm^{-1} and C. Relation between absorbance contrast function and spatial periods of the acquired image (solid) and the process image (striped). Scale bar: $100\mu\text{m}$.

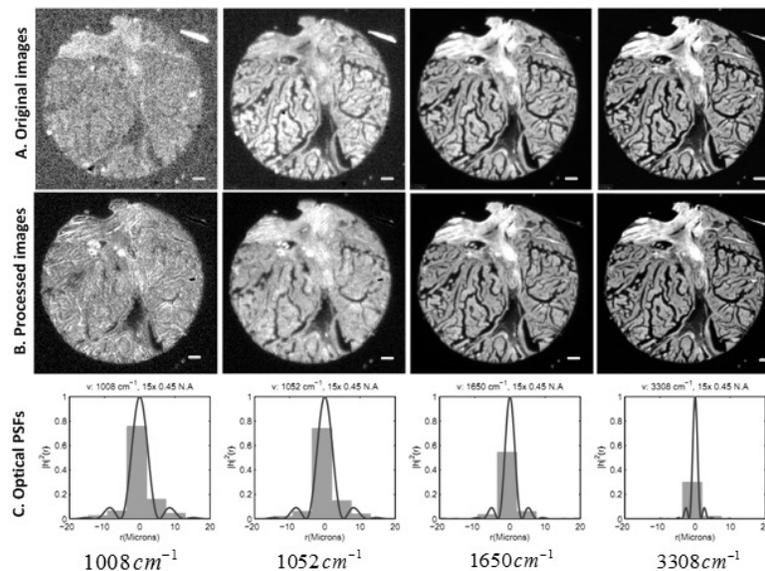


Figure 4. (A) Acquired absorbance images (Original images), (B) processed absorbance images at 1008cm^{-1} , 1052cm^{-1} , 1650cm^{-1} , and 3308cm^{-1} wavenumbers. (C) Discrete vs. continuous point spread functions. Scale bar: $100\mu\text{m}$.

tions. For a fixed optical setup, it can be seen that at low wavenumbers (1008cm^{-1} , 1052cm^{-1}), more sampling points are used to describe the continuous point spread function. Thus, the continuous point spread function is better approximated in this region. On the other hand, at relatively large wavenumbers (3308cm^{-1}), fewer sampling points are used. For example, at 3308cm^{-1} , main lobe of the continuous point spread function has the width equal to one sampling interval. As a result, processing efficiency decreases with increasing wavenumber. Our proposed algorithm also suggests that, to obtain good processing results across the entire spectrum, it is necessary to increase the number of sampling points at each wavenumber to at least 3 points for the main lobe so that the discrete point spread function is a good approximation of the continuous one. This can be achieved by increasing the net magnification factor of the system between the sample and the detector.

6. CONCLUSION

In this paper, we have presented a framework for improving the quality of FT-IR spectroscopic imaging data. The processing is based on three main ideas, namely, denoising based on low-rank property of the data, optical modeling of limited-aperture optical diffraction PSF and the Bayesian Variational method to estimate unknown parameters and obtain deblurred images. The algorithms were validated using a set of experiments on both standard (USAF target) and atypical (prostate tissue) samples. A qualitative and quantitative increase in image quality was demonstrated. While the algorithms show promising results, a few further improvements can be suggested. Prior knowledge on the continuous form of the point spread function can be incorporated into the model. Information from the low resolution image can be used to obtain a higher resolution image without increasing the magnification factor. A better image model for deconvolution that can incorporate boundary sharpness and background smoothness may be useful in improving deblurring results.

ACKNOWLEDGMENTS

This work was supported in part by the US National Science Foundation under Grant CHE 09-57849. The authors would like to thank Dr. Derin Babacan and Dr. David Mayerich for discussions.

REFERENCES

- [1] Davis, B., Carney, P., and Bhargava, R., "Theory of midinfrared absorption microspectroscopy: I. homogeneous samples," *Analytical Chemistry* **82**(9), 3474–3486 (2010).
- [2] Davis, B., Carney, P., and Bhargava, R., "Theory of mid-infrared absorption microspectroscopy: II. heterogeneous samples," *Analytical Chemistry* **82**(9), 3487–3499 (2010).
- [3] Nasse, M., Walsh, M., Mattson, E., Reininger, R., Kajdacsy-Balla, A., Macias, V., Bhargava, R., and Hirschmugl, C., "High-resolution fourier-transform infrared chemical imaging with multiple synchrotron beams," *Nature Methods* **8**(5), 413–416 (2011).
- [4] Reddy, R., Walsh, M., Schulmerich, M., Carney, P., and Bhargava, R., "High-definition infrared spectroscopic imaging," *Applied Spectroscopy* (Submitted Dec 2011).
- [5] Kohler, A., Kirschner, C., Oust, A., and Martens, H., "Extended multiplicative signal correction as a tool for separation and characterization of physical and chemical information in fourier transform infrared microscopy images of cryo-sections of beef loin," *Applied Spectroscopy* **59**(6), 707–716 (2005).
- [6] Kohler, A., Sul-Suso, J., Sockalingum, G., Tobin, M., Bahrami, F., Yang, Y., Pijanka, J., Dumas, P., Cotte, M., Pittius, D., Parkes, G., and Martens, H., "Estimating and correcting mie scattering in synchrotron-based microscopic fourier transform infrared spectra by extended multiplicative signal correction," *Applied Spectroscopy* **62**(3), 259–266 (2008).
- [7] Akgun, T., Altunbasak, Y., and Mersereau, R., "Super-resolution reconstruction of hyperspectral images," *IEEE Trans Image Proc.* **14**(11) (2005).
- [8] Bhargava, R., "Towards a practical fourier transform infrared chemical imaging protocol for cancer histopathology," *Anal. Bioanal. Chem.* **389**(4), 1155–1169 (2007).
- [9] Griffiths, P. and Haseth, J., [*Fourier Transform Infrared Spectrometry*], Wiley & Sons, New Jersey, 2nd ed. (2007).
- [10] Goodman, J., [*Introduction to Fourier Optics*], Roberts & Company Publishers (2005).

- [11] Molina, R., Mateos, J., and Katsaggelos, A., "Blind deconvolution using a variational approach to parameter, image, and blur estimation," *IEEE Trans. Image Proc.* **15**(12) (2006).
- [12] Babacan, D., Rafael, M., and Katsaggelos, K., "Variational bayesian blind deconvolution using a total variation prior," *IEEE Trans. Image Proc.* **18**(1) (2009).
- [13] Ripley, B., [*Spatial Statistics*], Wiley, New York (1981).