© 2009 Spencer Brady

SPECTRAL AND SPATIAL RECONSTRUCTION OF FOURIER TRANSFORM INFRARED IMAGING DATA

BY

SPENCER BRADY

B.S., Purdue University, 2007

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2009

Urbana, Illinois

Adviser:

Associate Professor Minh N. Do

ABSTRACT

Fourier transform infrared (FT-IR) spectroscopic imaging is a potentially valuable tool for diagnosing breast and prostate cancer, but its clinical deployment is limited by long data acquisition time and vast storage requirements. To counter the data acquisition problem, we develop a sparse representation for FT-IR absorbance spectra using a learned dictionary and use it to regularize the compressed sensing inverse problem. Another limitation in FT-IR imaging is insufficient spatial resolution of classified images. To counter this problem, we introduce an algorithm termed piecewise linear blockwise reconstruction (PLBR). This algorithm reconstructs high-resolution discretizations of the classified image by utilizing images of per-pixel tissue concentrations (obtained from unmixing analysis) to estimate the smooth boundaries between regions of homogeneous tissue. PLBR significantly outperforms the standard approach of quantizing the per-pixel tissue concentrations and is robust to noise.

To my parents

ACKNOWLEDGMENTS

I would like to thank my adviser, Associate Professor Minh N. Do, for his enthusiastic support and steadfast willingness to offer help and guidance. Without his wisdom and advice this work would not have been possible. I would also like to thank my co-adviser from Bioengineering, Assistant Professor Rohit Bhargava, for his patient encouragement and helpful suggestions. Lastly, I would like to thank my parents for their support throughout my time in graduate school.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

FT-IR	Fourier-transform infrared
K-SVD	K-singular value decomposition
PLBR	Piecewise linear blockwise reconstruction
SNR	Signal-to-noise ratio
PSNR	Peak signal-to-noise ratio
SIDE	Stabilized inverse diffusion equation
SAR	Synthetic aperture radar
FFT	Fast Fourier transform
DFT	Discrete Fourier transform
MP	Matching pursuit
OMP	Orthogonal matching pursuit
POCS	Projection onto convex sets
VCA	Vertex component analysis

CHAPTER 1

INTRODUCTION

1.1 FT-IR Imaging

Fourier transform infrared (FT-IR) imaging is at the forefront of an exciting era in molecular imaging in which tissue structures can be characterized in quantitative, objective ways by acquiring infrared absorbance spectra using a large focal plane array [1]. Unfortunately, this data set is inherently large because of its three-dimensional (spatial and spectral) nature, as shown in Figure 1.1. This large size hinders data acquisition, which can take days for a single tissue sample. Furthermore, data set sizes are typically on the order of 100 gigabytes, which significantly strains storage resources.

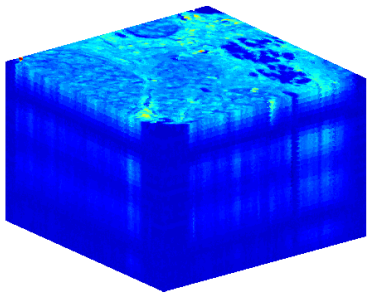


Figure 1.1: The 3-D FT-IR data cube for 1 mm diameter breast tissue is pictured above. The vertical dimension is the spectral dimension, and the two horizontal dimensions are spatial dimensions.

Another limitation of FT-IR imaging is spatial resolution. Currently, the maximum spatial resolution is around 6 micrometers. However, cancer diagnosis relies upon imaging cellular structures whose features are smaller than a pixel at this resolution. Thus, there is a great need for computational super-resolution

algorithms to provide more detailed pictures of cellular structure.

In this paper we first exploit sparse representation and compressed sensing to address the problem of large data size. Then we introduce a novel super-resolution algorithm termed piecewise linear blockwise reconstruction (PLBR), that relies upon prior knowledge of geometric regularity to obtain super-resolved images of tissue structure.

1.1.1 Data acquisition

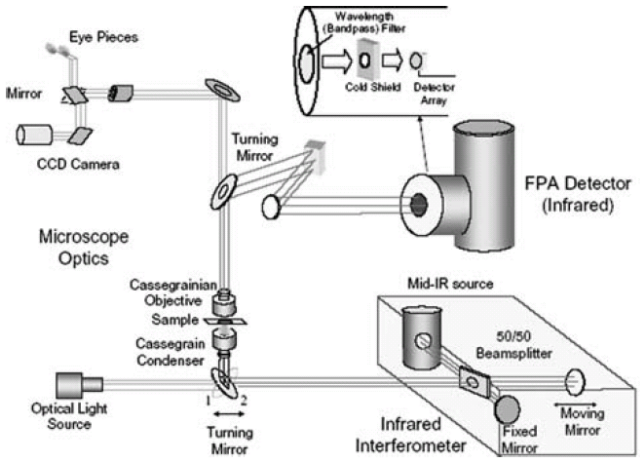


Figure 1.2: The FT-IR imaging system [2].

To acquire absorbance spectra, a mid-infrared broadband source is combined with a delayed split of itself by a Michelson interferometer [2]. This combination modulates the mid-infrared spectrum proportionally to the duration of the delay, which is governed by the retardation of a movable mirror inside the interferometer. The modulated source beam is passed through a tissue sample, which attenuates it according to that particular tissue's absorbance spectrum. Then, the spectral intensities are summed together at the detector, which produces a scalar value. Since this physical process corresponds to multiplying the tissue spectrum by a sinusoid and then summing the result, the scalar value at the detector is simply a Fourier coefficient. An illustration of FT-IR data acquisition is shown in Figure 1.2.

The mirror is moved symmetrically about its origin to produce an array of Fourier coefficients, which is called the interferogram. Because the mirror is moved symmetrically about its origin, and because the mid-infrared spectrum is always real-valued, the interferogram is also real and symmetric. Thus, it has the following simplified Fourier transform representation

$$I(\delta) = \int_{-\infty}^{\infty} S(\nu) \cos(2\pi\nu\delta) d\nu. \tag{1.1}$$

In the above equation, $I(\delta)$ is the interferogram, $S(\nu)$ is the tissue transmittance spectrum, δ is the mirror retardation (cm), and ν is the wavenumber (cm^{-1}). To obtain the tissue transmittance spectrum from the interferogram, we take the inverse of (1), which is

$$S(\nu) = \int_{-\infty}^{\infty} I(\delta) \cos(2\pi\nu\delta) d\delta. \tag{1.2}$$

A transmittance spectrum of the background (no tissue sample in beam path) is also acquired in the same manner. The logarithm of the ratio of the transmittance spectrum of the background to that of the sample provides the absorbance spectrum.

1.1.2 Image formation and analysis

An interferogram is collected at each spatial location of the focal plane array, resulting in a three-dimensional interferogram cube. To ensure even symmetry, each interferogram is phase-corrected using either the Mertz or the Foreman method [3]. Next, the discrete Fourier transform (DFT) is computed via fast Fourier transform (FFT), which produces a spectrum. The last step is to take the log ratio with respect to a background spectrum to convert the recorded data to absorbance. The result of all this processing is to obtain the absorbance data cube, as shown in Figure 1.1.

But the story does not end here. With this large data cube, tissue

classification can be performed. In a preliminary step known as baseline correction, a piecewise linear function is subtracted from the spectra, which effectively "zeros out" the spectra at points known to have no light absorbance. Next, a set of nonlinear features is measured. With these features, a Bayesian classifier estimates the tissue classes at each spatial location based on prior distributions of each class. Tissue classes of interest include epithelium (in which more than 85% of cancer arises) and many other cell types including fibroblasts, myofibroblasts, lymphocytes, and endothelium.

Another type of analysis, known as unmixing, seeks to decompose each spectrum into a nonnegative linear combination of known (supervised unmixing) or unknown (unsupervised unmixing) component spectra. When the component spectra are unknown, hyperspectral unmixing algorithms such as vertex component analysis are of great use [4], [5].

1.2 Compressed Sensing

Much progress has been made recently on the theoretical problem of solving underdetermined systems of linear equations in which the solution is known to be sparse. Work in [6], [7] has shown that if a length n signal is compressible by transform coding, then it can be reconstructed from $m \ll n$ nonadaptive measurements by a nonlinear programming procedure known as ℓ_1-minimization. Candes et al. addressed the problem of reconstructing a discrete-time signal from a randomly chosen set of frequencies and provided sufficient conditions under which, with probability one, the signal can be reconstructed exactly as the solution to the ℓ_1-minimization problem [8].

On the practical aspects of compressive sensing, Parmar et al. developed sparse representation for the visible-light spectra of natural images and then used this to recover a spatio-spectral data cube from measurements obtained with a conventional digital camera [9]. Their approach is similar to ours in that they are reconstructing multi-spectral data from a few measurements using

compressed sensing techniques. One difference is that our data consists of infrared light absorbance of biological tissues on the scale of millimeters as opposed to visible light reflection of natural objects on the scale of meters. A more important difference is that our measurement process is intimately tied to the physics of FT-IR image formation (in particular, interferogram truncation), while their measurement process consists of obtaining a conventional RGB image.

1.3 Super-Resolution

The goal of spatial super-resolution in FT-IR imaging is to increase the spatial resolution of discretizations of the so-called classified image (to be defined). Our approach is to utilize images of per-pixel tissue concentrations obtained from unmixing analysis to estimate the smooth boundaries between regions of homogeneous tissue. The simplest approach is to assign to the entire pixel the tissue class with the highest concentration, which limits the spatial resolution of the classified image to that of the concentration image and inevitably results in staircase-like contours between tissue regions. Instead, we seek to find subpixel contours that correctly partition each pixel to match the known per-pixel tissue concentrations while also ensuring that the union of these contours is smooth.

This problem is related to the problem of image segmentation. However, image segmentation algorithms do not generally provide higher resolution than that of the input image. For example, Pollak et al. [10] proposed using stabilized inverse diffusion equations (SIDEs) to segment images, as shown in Figure 1.3. Note that the segmented image has the same spatial resolution as the noisy input image.

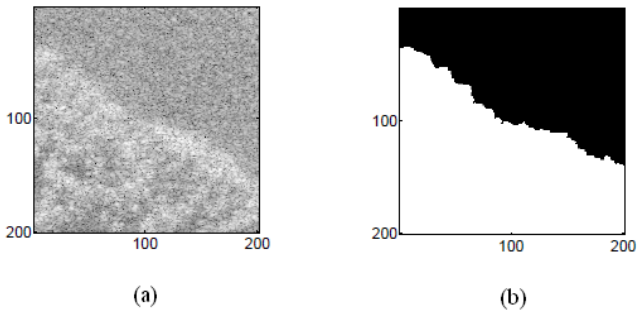


Figure 1.3: An example of image segmentation using SIDEs. (a) Noisy SAR image; (b) two-region segmentation.

CHAPTER 2

SPARSE SPECTRAL RECONSTRUCTION

2.1 Formulation

The first step in training a spectral dictionary is to transform the three-dimensional data cube $S_{3d} \in \mathbb{R}^{x \times y \times m}$ into a matrix $S \in \mathbb{R}^{m \times n}$ where m is the number of spectral samples and $n = xy$ is the number of spatial locations. We index the spatial locations columnwise lexicographically, and let $S = [\underline{s}_1 \quad \cdots \quad \underline{s}_n]$ where $\underline{s}_i$ is the spectral absorbance at spatial index i. Now we can formulate our dictionary representation as the matrix factorization $S = DX$ where $D \in \mathbb{R}^{m \times p}$ is the dictionary, $X \in \mathbb{R}^{p \times n}$ are the coefficients, and p is the number of atoms in the dictionary. Furthermore, we wish to impose a sparsity constraint on X: that for every column $\underline{x}_i$ in X, $||\underline{x}_i||_0 \leq \ell$. This is formulated as the following optimization problem

$$\underset{D,X}{\operatorname{argmin}} ||S - DX||_F^2 \quad s.t. \quad ||\underline{x}_i||_0 \leq \ell \;\; for \;\; i = \{1, \ldots, n\}, \tag{2.1}$$

and can be solved by the K-singular value decomposition (K-SVD) algorithm, which is a generalization of the K-means clustering algorithm and relies on an auxiliary basis pursuit algorithm [11]. To run K-SVD, we must provide it with three variables: the training data S, the number of dictionary atoms p, and the number of atoms permitted in each linear combination ℓ. The algorithm returns two variables: the dictionary D and the sparse coefficients X.

2.2 K-SVD Algorithm

As mentioned above, the K-SVD algorithm seeks to minimize $||S - DX||_F^2$ subject to a sparsity constraint on the columns of X. When the dictionary D is known, X is obtained in a decoupled fashion by solving for each column one at a time using any standard compressed sensing algorithms such as the greedy algorithms matching pursuit (MP) and orthogonal matching pursuit (OMP), ℓ_1-minimization (also known as basis pursuit or relaxation), or a variety of iterative methods such as shrinkage or projection onto convex sets (POCS). The method included in the K-SVD software is OMP.

Once X is known, the k^{th} column of D and the k^{th} row of X are updated simultaneously. To obtain the new d_k and x_T^k, the K-SVD algorithm performs singular value decomposition (SVD) on the following error matrix,

$$E = \left(S - \sum_{j \neq k} d_j x_T^j \right) \Omega_k = U \Delta V^T, \tag{2.2}$$

where the matrix Ω_k has the property that, when multiplied by S, it removes the columns from S that are not derived from a linear combination containing d_k.

From the SVD, we obtain the rank-one approximation matrix minimizing the Frobenius norm between E and its rank-one approximation, which is

$$E \approx u_1 \Delta(1,1) v_T^1, \tag{2.3}$$

from which we obtain $\tilde{d}_k = u_1$ and $\tilde{x}_T^k = \Delta(1,1) v_1$.

The coefficients X are again computed via OMP to reinforce the sparseness constraint given the fixed dictionary D. These two steps are iterated until a desired error criterion is reached.

2.3 Practical Considerations

In practice, it is important that $\ell/p \leq 0.1$ to ensure that the representation is indeed sparse. Furthermore, it is important to choose an appropriate value for the number of dictionary atoms p in relation to the number of spectral samples m. If $p > m$, the representation becomes trivial. For $m = 100$, we set $\ell/p = 0.1$ and let p be 10, 20, 30, 40, and 50 which yielded sparse dictionary representations having PSNRs of 38.22, 43.55, 43.27, 44.26, and 44.92 dB respectively. The matrix factorization is shown in Figure 2.1.

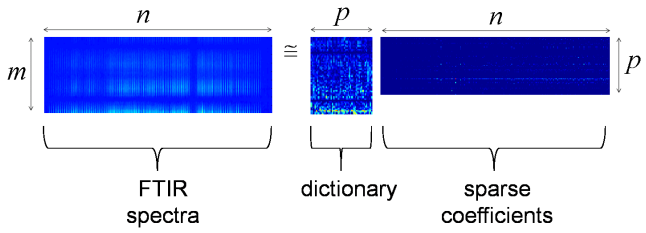


Figure 2.1: The matrix factorization $S = DX$ obtained with K-SVD for $p = 50$ and $\ell = 5$.

2.4 Measurement Process

Recall that data acquisition time and storage size are determined by the length of the interferogram. Hence, we model our measurement process as an interferogram truncation, resulting in a new spectrum $S'(\nu)$. This is equivalent to multiplying the interferogram $I(\delta)$ by the rectangular window

$$W(\delta) = \begin{cases} 1, & |\delta| \leq \frac{T}{2} \\ 0, & else \end{cases} .$$

Plugging $I'(\delta) = W(\delta)I(\delta)$ into Equation (1.2) yields

$$
\begin{aligned}
S'(\nu) &= \int_{-\infty}^{\infty} W(\delta) I(\delta) \cos(2\pi\nu\delta) d\delta \\
&= S(\nu) * \int_{-\frac{T}{2}}^{\frac{T}{2}} \cos(2\pi\nu\delta) d\delta \\
&= S(\nu) * \frac{\sin(\pi\nu T)}{\pi\nu}.
\end{aligned}
$$

Since we are dealing with discrete data, we reformulate the problem using the DFT, which yields

$$
W[j] = \begin{cases} 1, & |j| \in \{0, 1, \cdots, \frac{J}{2}\} \\ 0, & else \end{cases} \tag{2.4}
$$

$$
\begin{aligned}
S'[k] &= \sum_{j=-\frac{m}{2}}^{\frac{m}{2}} W[j] I[j] \cos(2\pi kj/m) \\
&= S[k] * \sum_{j=-\frac{J}{2}}^{\frac{J}{2}} \cos(2\pi kj/m) \\
&= S[k] * \frac{\sin(\pi kJ/m)}{\pi kJ/m}.
\end{aligned} \tag{2.5}
$$

Because $S'[k]$ is bandlimited to $2\pi J/m$, we can safely downsample by factor m/J without losing any information. Thus, our final measurement is $S'[mk/J]$, which conveniently can be obtained via matrix multiplication. Let us define the matrix Y whose columns are the measurements of the columns of the matrix S. We describe the relationship between Y, S, and X as

$$
\begin{aligned}
Y &= MS \\
&= MDX
\end{aligned} \tag{2.6}
$$

where M is a matrix that filters the columns of S with the sinc filter in

Equation (2.5) and then downsamples that by m/J. The measurement process for translating sparse coefficients X into low-resolution spectra $Y = MDX$ is shown in Figure 2.2. Given the columns of Y and the matrix $A = MD$, we seek to recover the columns of X in a decoupled fashion by solving the compressed sensing problem.

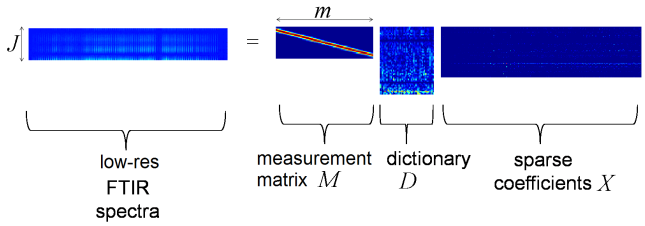


Figure 2.2: The measurement process $Y = MDX$.

2.5 Sparse Linear Inversion

The compressed sensing problem,

$$\underset{\underline{x}}{\text{argmin}} \; ||\underline{y} - A\underline{x}||_2^2 \quad s.t. \quad ||\underline{x}||_0 \leq \ell, \tag{2.7}$$

is an NP-hard search problem, which, for large data sets, is impractical to solve. Instead, we are interested in *approximate* solutions. There are three main approaches to finding approximate solutions to the compressed sensing problem. All exhibit high accuracy when the solution is sparse enough, while providing lower computational complexity, making them feasible for large data sets.

The first approach, often called greedy, includes MP and OMP. These methods start by finding $\underline{a}_i$, the column of A having the greatest correlation (inner product) with $\underline{y}$. Then the i^{th} entry of the vector $\underline{x}$ is assigned the value $< \underline{a}_i, y > /||\underline{a}_i||_2^2$. Next, the projection of $\underline{y}$ onto $\underline{a}_i$ is subtracted from $\underline{y}$ to obtain the residual signal $\underline{r}$. The original signal $\underline{y}$ is replaced with $\underline{r}$ and the process is repeated $\ell - 1$ more times, until $\underline{x}$ contains ℓ nonzero entries. These methods are called greedy because they are optimal at each step of the algorithm but are not

generally globally optimal.

The next approach is called relaxation, because we relax the norm from $||\underline{x}||_0$ to $||\underline{x}||_1$. While ℓ_0-minimization is NP-hard, ℓ_1-minimization can be cast as a standard linear programming problem, having a guaranteed solution given by the simplex algorithm. Under sparse enough conditions, the solutions to ℓ_0-minimization and ℓ_1-minimization are the same, as shown in Figure 2.3.

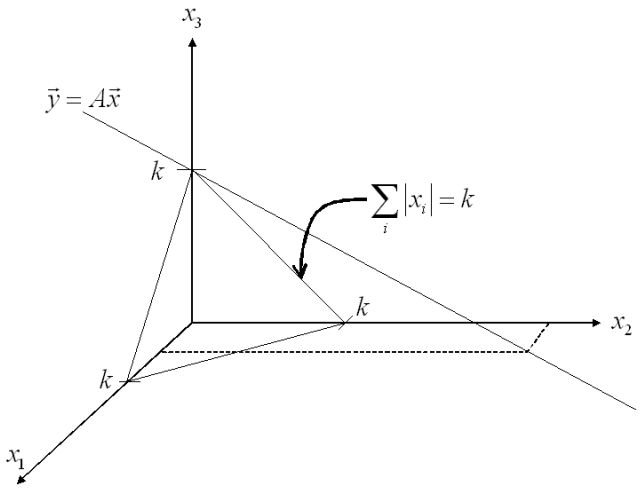


Figure 2.3: An illustration showing the case when ℓ_0-minimization and ℓ_1-minimization are the same.

The third main approach to approximately solving the compressed sensing problem is to iteratively project onto two solution sets by orthogonal projection. If the solution sets are convex, this procedure is guaranteed to converge to a solution. For example, one solution set enforces the equality $\underline{y} = A\underline{x}$, while the other solution set enforces sparsity on $\underline{x}$. Projection onto the first set is implemented by replacing the reconstructed $\underline{y}$ with the given $\underline{y}$, while projection onto the second set is implemented by soft or hard thresholding of the coefficients $\underline{x}$, a procedure known as shrinkage.

2.6 Reconstruction

We are given the low-resolution measurements $Y = MS$, where M is our measurement matrix, and S are the high-resolution spectra in matrix form, and we wish to recover $\hat{S}$ from Y by exploiting our knowledge that $S = DX$, where D is the K-SVD dictionary and X are sparse coefficients. On a pixel-by-pixel basis, this problem is formulated as the compressed sensing problem, which we attack using ℓ_1-minimization

$$\underline{x}_{opt} = \underset{\underline{x}}{\operatorname{argmin}} \, ||\underline{y} - MD\underline{x}||_2^2 + \lambda||\underline{x}||_1. \tag{2.8}$$

Once the sparse coefficients $\underline{x}_{opt}$ are obtained, a reconstructed spectrum $\hat{\underline{s}}$ is obtained by $\hat{\underline{s}} = D\underline{x}_{opt}$.

2.7 Experimental Results

We acquired a data set for breast tissue from a Perkin Elmer imaging spectrometer (Spotlight 300). The image size was 300×300, pixels corresponding to an area of 1 mm^2, and the spectral size was 100×1, corresponding to a wavenumber range of 1100 to 1300 cm^{-1}. Half of the spectra were used to train the K-SVD dictionary, while the other half were used for the measurement and reconstruction. For M we used a symmetrically extended support-preserving convolution matrix generated with a sinc filter of length 21 and then removed rows to implement downsampling. The dictionary D was obtained from K-SVD with p and ℓ depending on the downsample factor. The ℓ_1-minimization was implemented with the software package ℓ_1-magic [8]. To initialize the algorithm, for our initial guess $\underline{x}_0$, we chose the minimum (ℓ_2) norm least squares solution obtained via pseudo-inverse $\underline{x}_0 = (MD)^\dagger \underline{y}$.

In our first experiment, we trained the K-SVD dictionary with $p = 50$ and $\ell = 5$. We constructed the measurement matrix M to have a downsample factor of four. We generated the low-resolution measurements $Y = MS$. Standard

interpolation of the low-resolution measurements yielded a PSNR of 36.93 dB, while sparse reconstruction yielded a PSNR of 41.92 dB. A comparison of the original, reconstructed, and interpolated measurement spectra is shown in Figure 2.4. In comparison to the interpolated spectrum, the sparse reconstruction retains details necessary for tissue characterization.

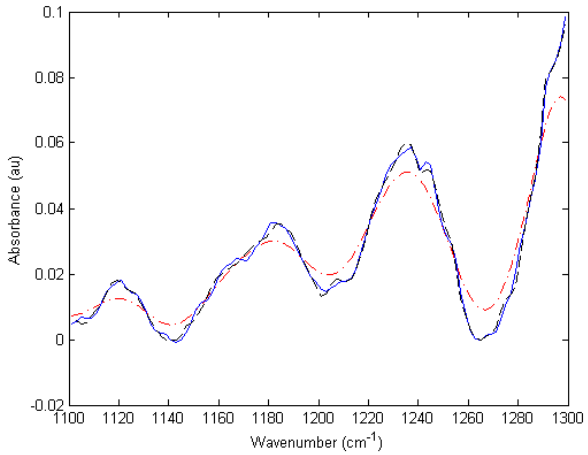


Figure 2.4: Downsample factor = 4. Original high-resolution spectrum (solid blue), sparsely reconstructed spectrum (dashed black), and interpolated low-resolution spectrum (dash-dot red) at spatial index $i = 8000$.

Our second experiment was identical to the first, except that we set $p = 25$ and $\ell = 3$ and then designed M to have a downsample factor of eight. Standard interpolation of the low-resolution measurements yielded a PSNR of 34.37 dB, while sparse reconstruction yielded a PSNR of 36.02 dB. A comparison of the original, reconstructed, and interpolated measurement spectra is given in Figure 2.5.

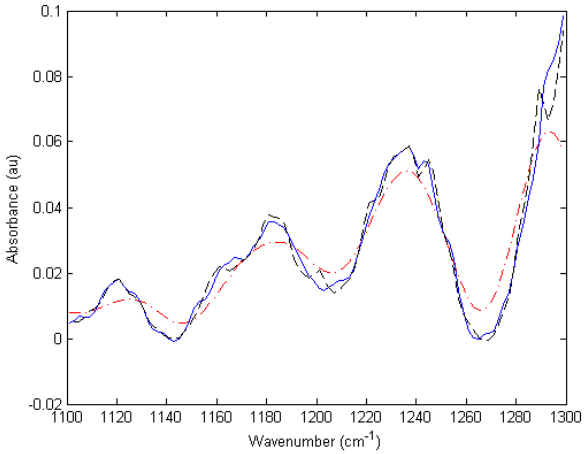


Figure 2.5: Downsample factor = 8. Original high-resolution spectrum (solid blue), sparsely reconstructed spectrum (dashed black), and interpolated low-resolution spectrum (dash-dot red) at spatial index $i = 8000$.

CHAPTER 3

SPATIAL SUPER-RESOLUTION

3.1 Linearity and Beer's Law

Beer's law [3] is a fundamental law in spectroscopy that states that for an absorbance spectrum containing p different component spectra in different concentrations, the total absorbance at wavenumber ν is given by

$$s(\nu) = \sum_{i=1}^{p} a_i e_i(\nu). \tag{3.1}$$

Thanks to Beer's law, FT-IR data can be modeled as a linear mixture of component spectra. This is formulated as the matrix equation

$$S = EA, \tag{3.2}$$

where columns of $S \in \mathbb{R}^{m \times n}$ are the absorbance spectra, the columns of $E \in \mathbb{R}^{m \times p}$ are termed the component spectra, and the columns of $A \in \mathbb{R}^{n \times p}$ are termed coefficients. Note that when infrared light is incident on the boundary between tissues having different indices of refraction, a nonlinear phenomenon occurs, known as scattering, which in some cases violates the linear mixing model assumption.

3.2 Vertex Component Analysis

Vertex component analysis (VCA) [4] is an unmixing algorithm that, when given the absorbance spectra matrix S and the number of component spectra p,

extracts the component spectra E and the corresponding concentrations A. The algorithm assumes that there exist spectra in S that are pure (only one column of E), and then finds the vertices in the simplex formed by the columns of S in $\mathbb{R}^n$. With the component spectra known, VCA obtains A through linear inversion of the equation $S = EA$ subject to the constraint that each column of A sums to the same constant.

3.3 Preliminary Definitions

3.3.1 Classified images

A classified image $r(x, y)$ is a function that maps from continuous coordinate points in the tissue sample $(x, y) \in \mathbb{R}^2$ to the set of p discrete tissue classes $\{1, \ldots, p\}$. Classified images are used by pathologists to diagnose cancer.

When we want to consider only tissue class i, we define the marginally classified image

$$r_i(x, y) = \begin{cases} 1, & r(x, y) = i \\ 0, & r(x, y) \neq i \end{cases}.$$

Note that the marginally classified image is a bi-level (binary or black-and-white) image. Since classified images have continuous support, we aspire to reconstruct only high-resolution discretizations of the classified images, rather than the classified images themselves. Nonetheless, we shall model the classified images in the continuous domain.

3.3.2 Concentration images

Given a marginally classified image $r_i(x, y)$, the corresponding marginal concentration image $t_i(m, n)$ is a function mapping from a discrete array points in the tissue sample $(m, n) \in \mathbb{Z}^2$ to the set of real numbers between zero and one

$[0,1]$. The relationship between $r_i(x,y)$ and $t_i(m,n)$ is

$$t_i(m,n) = \int_{\frac{m}{k_1}}^{\frac{m+1}{k_1}} \int_{\frac{n}{k_2}}^{\frac{n+1}{k_2}} r_i(x,y)dxdy. \tag{3.3}$$

In this definition, $r_i(x,y)$ has support on $[0,1] \times [0,1]$, $1/k_1 \times 1/k_2$ is the size of each pixel, and m and n take on values from zero to $k_1 - 1$ and $k_2 - 1$, respectively. Intuitively, the concentration image tells us the concentration of the tissue i at pixel (m,n) by integrating the marginally classified image over that pixel. We call this process area sampling.

We hope that, in the absence of nonlinear scattering affects, concentration images can be obtained from FT-IR data using VCA. For now we focus on synthetically mixed multi-spectral data, from which we extract concentration images using VCA and then use to perform super-resolution.

3.4 Problem Statement

Our goal is to extract high-resolution discretizations of the classified image $r(x,y)$ from the concentration images $t_i(m,n)$. The simplest approach is to assign to each pixel in the discretized classified image the tissue class having the highest concentration. However, this limits the resolution of the classified image to the resolution of the concentration image. By utilizing prior knowledge that the sub-pixel boundaries between adjacent regions of homogeneous tissue are smooth, we hope to increase the spatial resolution of the classified image beyond that of the concentration image.

3.5 Piecewise Linear Blockwise Reconstruction

3.5.1 An invertible transform

First we present an invertible transform that will serve as the foundation for our super-resolution algorithm. Consider the case when we have only two tissue classes ($p = 2$) and let $k_1 = k_2 = 2$. In this case, note the following simplifications: $r(x, y)$ is a bi-level image; since $t_1(m, n) = \frac{1}{4} - t_2(m, n)$, the information content in $t_1(m, n)$ and $t_2(m, n)$ is the same, thus we only need to consider one marginal concentration image.

Furthermore, suppose that $r(x, y)$, assumed to have support $[0, 1] \times [0, 1]$, has the following parameterization:

$$r(x, y) = \begin{cases} 1, & y \geq mx + h \\ 0, & y < mx + h \end{cases} . \tag{3.4}$$

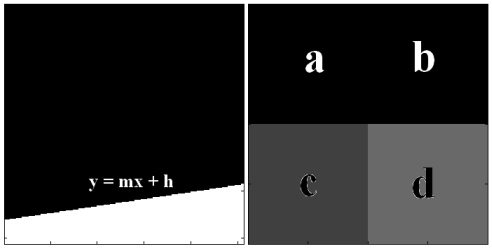
Consider the concentration image $t_1(m, n)$ obtained from $r(x, y)$ when k_1 and k_2 are both equal to two. The concentration image $t_1(m, n)$ is a 2×2 matrix, which we shall denote

$$t = \begin{bmatrix} a & b \\ c & d \end{bmatrix} . \tag{3.5}$$

Note that each entry in t is between zero and $\frac{1}{4}$ because the total concentration must be less than or equal to one. Let us distinguish entries in t that are not equal to zero or $\frac{1}{4}$ by calling them *significant*.

We claim that the mapping between the line parameters $\{m, h\}$ and the concentrations $\{a, b, c, d\}$ is bijective whenever t has at least two significant entries. The equations relating $\{m, h\}$ and $\{a, b, c, d\}$ will be respectively linear or nonlinear, depending on whether the line $y = mx + h$ intersects opposite or adjacent sides of the support boundary of $r(x, y)$.

First let us examine the linear case, where $y = mx + h$ intersects opposite



(a) (b)

Figure 3.1: (a) A discretization of $r(x, y)$ when the boundary is parameterized by a line intersecting opposite sides of the support boundary; (b) the concentration image $t(m, n)$ corresponding to $r(x, y)$ for $k_1 = k_2 = 2$.

sides of the support boundary of $r(x, y)$ as shown in Figure 3.1. Let us assume that the only significant entries in t are c and d. In this case, we can write the following expressions

$$c = \int_0^{\frac{1}{2}} (mx + h)dx = \frac{m}{8} + \frac{h}{2}$$

$$d = \int_{\frac{1}{2}}^1 (mx + h)dx = \frac{3m}{8} + \frac{h}{2},$$

which we can reformulate this as the following matrix expression

$$\begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} \frac{1}{8} & \frac{1}{2} \\ \frac{3}{8} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} m \\ h \end{bmatrix}, \tag{3.6}$$

with the inverse equation

$$\begin{bmatrix} m \\ h \end{bmatrix} = \begin{bmatrix} -4 & -4 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix}. \tag{3.7}$$

This forward matrix has condition number 5.05 and is thus invertible and robust to small perturbations in c and d.

The nonlinear equations arise when $y = mx + h$ intersects adjacent sides of the support boundary of $r(x, y)$ as shown in Figure 3.2. Let us again assume that the only significant entries in t are c and d. In this case, we can write the

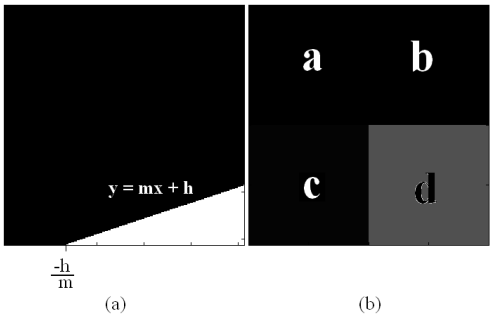


Figure 3.2: (a) A discretization of $r(x, y)$ when the boundary is parameterized by a line intersecting adjacent sides of the support boundary; (b) the concentration image $t(m, n)$ corresponding to $r(x, y)$ for $k_1 = k_2 = 2$.

following expressions

$$c = \int_{\frac{-h}{m}}^{\frac{1}{2}} (mx + h)dx = \frac{m}{8} + \frac{h}{2} + \frac{h^2}{2m}$$

$$d = \int_{\frac{1}{2}}^{1} (mx + h)dx = \frac{3m}{8} + \frac{h}{2},$$

which gives us a system of nonlinear equations that can be solved with the quadratic formula to yield the following two solutions

$$m = 16c + 8d \mp \sqrt{c^2 + dc}$$

$$h = -4d - 12c \pm \sqrt{c^2 + dc}.$$

Note that m and h have two possible values each, but that only two of the four possible combinations of m and h are solutions to the above system of equations. Among the two valid solutions, only one will satisfy $0 \leq -h/m \leq \frac{1}{2}$, as shown in Figure 3.3.

Given these two different sets of equations (linear and nonlinear), we need to know which formulae are appropriate to use to extract $\{m, h\}$ from $\{c, d\}$. If the x-intercept, $-h/m$, is less than zero, then the line must intersect opposite sides of the support boundary of $r(x, y)$, indicating the need to use the linear

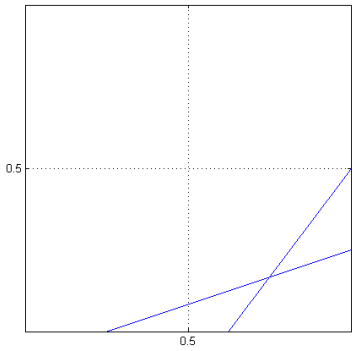


Figure 3.3: An illustration of two possible solutions for m and h in the nonlinear system of equations corresponding to one concentration image t. Note that only one of the lines intersects both pixels c and d.

equations. Otherwise, the nonlinear equations should be used. This condition can be formalized as

$$\frac{-h}{m} < 0.$$

Plugging in our linear solutions for h and m allows us to write this condition in terms of c and d,

$$\frac{d - 3c}{4(d - c)} < 0. \tag{3.8}$$

Thus, given a 2×2 matrix t with two adjacent, significant entries, we can flip and/or rotate t so that both significant entries are on the bottom, and then we can use the above inequality to decide which set of equations to use to exactly recover m and h.

In the case where t has three significant entries, we combine two adjacent pixels to form a super-pixel, and the proceed as if there were only two pixels.

In the case where t has only one significant entry, one can convince themselves that there is no unique mapping between $\{m, h\}$ and t. For example, the lines $y = 0.5x - 0.25$ and $y = 2x - 1.5$ both produce the same concentration matrix t, as shown in Figure 3.4.

Finally, the case where t has four significant entries is not possible if the

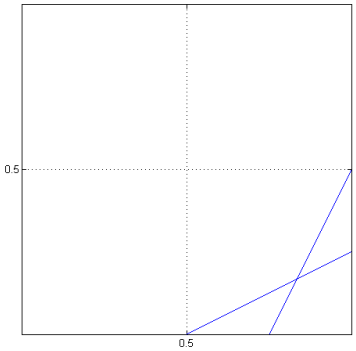


Figure 3.4: The lines $y = 0.5x - 0.25$ and $y = 2x - 1.5$ both produce the same concentration matrix with only one significant entry $d = \frac{1}{16}$.

boundary between adjacent regions is a straight line.

Thus, we have shown that the mapping between $\{m, h\}$ and t is bijective as long as t has at least two significant entries (that is, entries that are not equal to zero or $\frac{1}{4}$).

3.5.2 Noise-free reconstruction

We model the classified image $r(x, y)$ as being well approximated by piecewise linear contours. We reconstruct our approximation of $r(x, y)$ from $t(m, n)$ by replacing 2×2 blocks with solid black, solid white, or high-resolution linear contour boundaries obtained from our invertible transform. The first step is to count the number of significant entries in a 2×2 block. If it is zero, then we replace the 2×2 block with an image of entirely black or white. If there is one significant entry, we do a local search around that pixel for 2×2 blocks containing that pixel and having at least two significant entries. When the block contains at least two significant entries, we perform the inverse transform to obtain the line parameters m and h, and then replace the 2×2 block with a high-resolution image of a linear contour boundary.

3.5.3 Reconstruction with noise

Adding noise to the concentration image compounds the inverse problem in two respects. First, in the block search, when counting the number of significant pixels in each 2×2 block, we must allow for the possibility that insignificant pixels may not be exactly equal to 0 or $\frac{1}{4}$. Second, if we incorrectly identify the number of significant pixels in a block, then we risk selecting the wrong set of inverse equations to solve for the line parameters, which can result in spurious solutions.

As a remedy to these difficulties, we make two modifications. First, we use a tolerance when identifying significant pixels. We set this tolerance to be equal to the noise standard deviation. Second, we replace the closed-form solution for the invertible transform with an optimization problem. We seek the line parameters that minimize the Frobenius norm between the given concentration image and the one computed using the line parameters. This can be written as

$$p^* = \{m^*, h^*\} = \underset{m,h}{\operatorname{argmin}} \, ||t - \hat{t}||_F^2, \tag{3.9}$$

where

$$\hat{t}_(m, n) = \int_{\frac{m}{k_1}}^{\frac{m+1}{k_1}} \int_{\frac{n}{k_2}}^{\frac{n+1}{k_2}} r(x, y) dx dy,$$

and

$$r(x, y) = \begin{cases} 1, & y \geq mx + h \\ 0, & y < mx + h \end{cases} . \tag{3.10}$$

However, a brute force search for $p = \{m, h\}$ is difficult without restricting the possible values of $\{m, h\}$. It is additionally challenging because we are searching over an infinite set. To overcome this, we propose using a new parameter set, $p = \{\ell_1, \ell_2\}$, defined as the counterclockwise distances along the image perimeter to the points where the line intersects the image boundary. The advantage of this parameterization is that we only have to search over $0 \leq \ell_1 \leq 4$ and $0 \leq \ell_2 \leq 4$. This parameterization is illustrated in Figure 3.5.

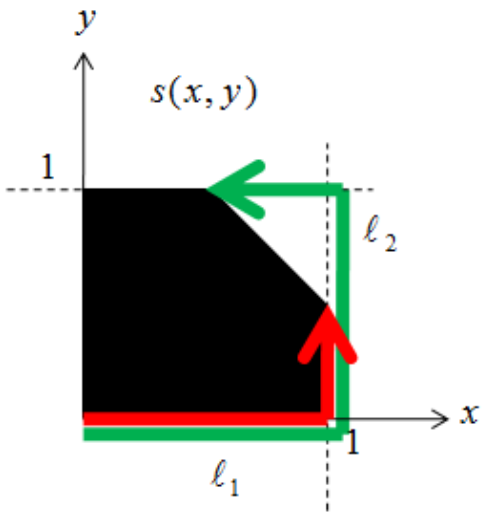


Figure 3.5: The proposed line parameterization $p = \{\ell_1, \ell_2\}$.

We impose the convention that the image polarity is determined by the inequality $\ell_2 > \ell_1$. That is, if we exchange ℓ_2 and ℓ_1, we simply invert the image polarity. In addition, we note that $\ell_2 \geq \lceil \ell_1 \rceil$ and $\ell_1 \geq \lceil \ell_2 \rceil$ to further reduce the parameterization search set, shown in Figure 3.6. We search the parameter set in Figure 3.6 for the optimal parameter set $p^* = \{\ell_1^*, \ell_2^*\}$ using a multi-resolution approach shown in Figure 3.7.

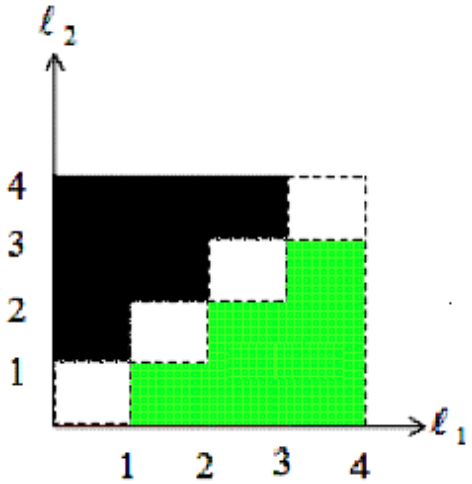


Figure 3.6: The search set of image parameters.

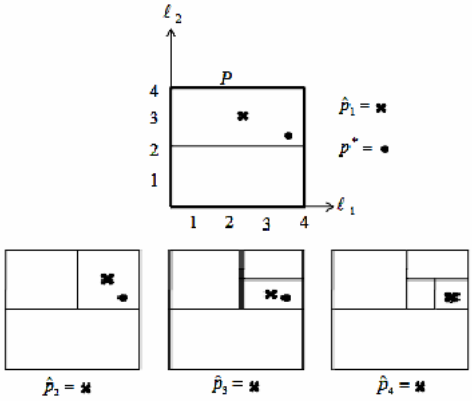


Figure 3.7: The multi-resolution search algorithm. p^* is the optimal solution, and p^i is the i^{th} guess.

3.6 Experimental Results

3.6.1 Noise-free case

For the noise-free case, we synthesized a discretization of the bi-level classified image $r(x, y)$, and then performed area sampling on it to obtain the concentration image $t(m, n)$. First, as a benchmark, we reconstructed $r(x, y)$ by uniformly quantizing the concentration images to two levels. Then we reconstructed the $r(x, y)$ by applying our piecewise linear blockwise reconstruction algorithm. The results were significantly better with PLBR than with quantization, as shown in Figures 3.8–3.11.

3.6.2 Noisy case

We synthesized noisy concentration images by adding white, zero-mean, Gaussian noise with standard deviation 0.05. We tested our numerical search PLBR algorithm on synthetic noisy concentration images and then reconstructed the classified images, shown in Figures 3.12–3.14.

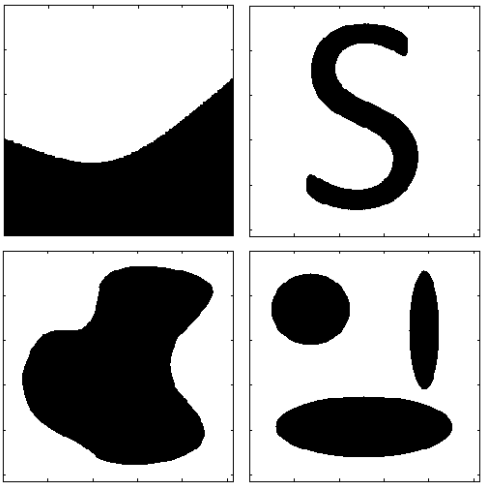


Figure 3.8: The discretized classified images $r(x, y)$.

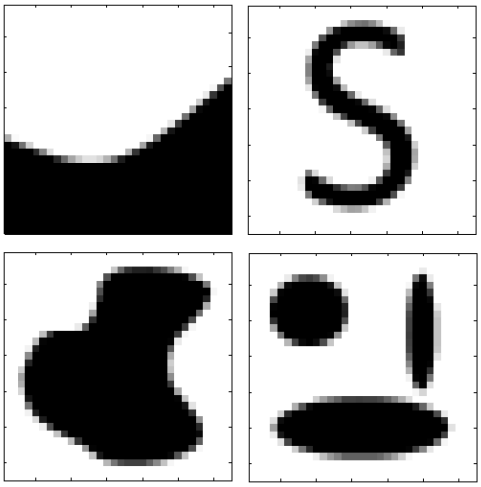


Figure 3.9: The concentration images $t(m, n)$ obtained by area sampling of the classified images in Figure 3.8.

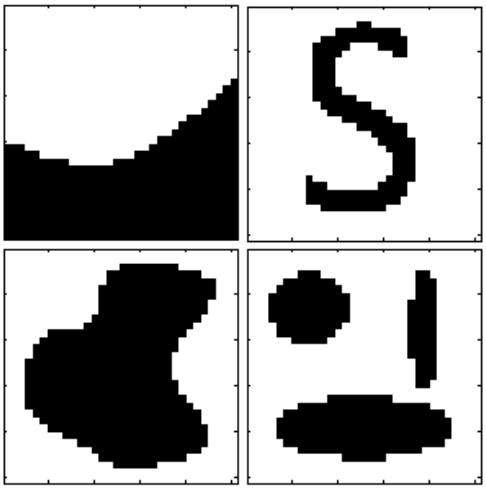


Figure 3.10: The quantized concentration image.

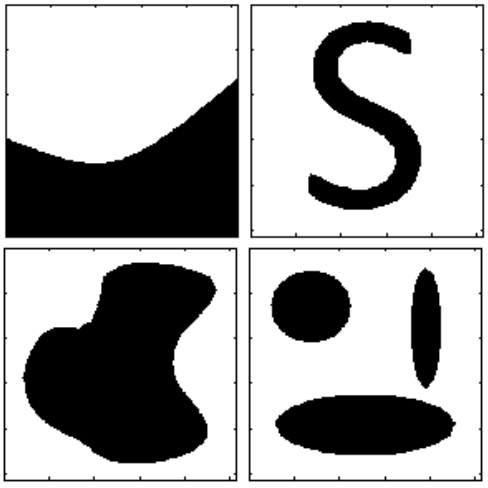


Figure 3.11: The image reconstructed by PLBR.

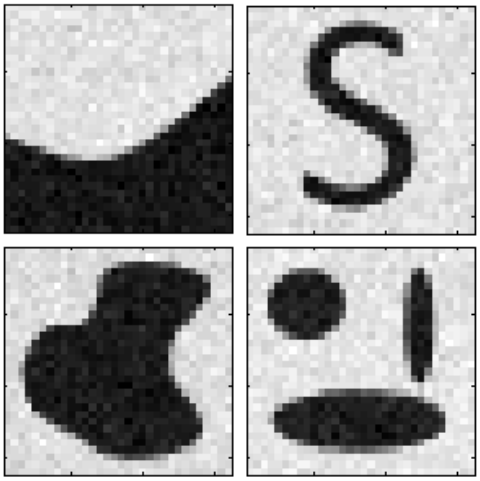


Figure 3.12: The noisy concentration images.

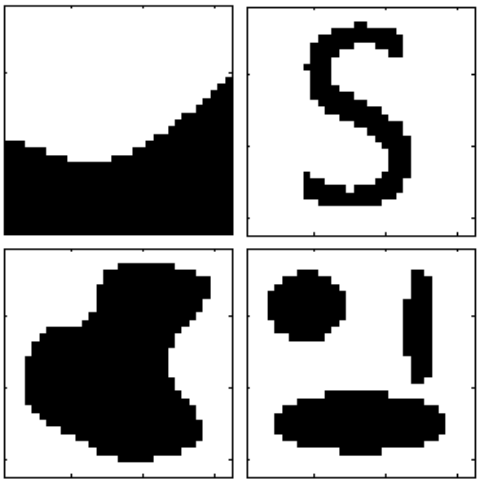


Figure 3.13: The quantized noisy concentration images.

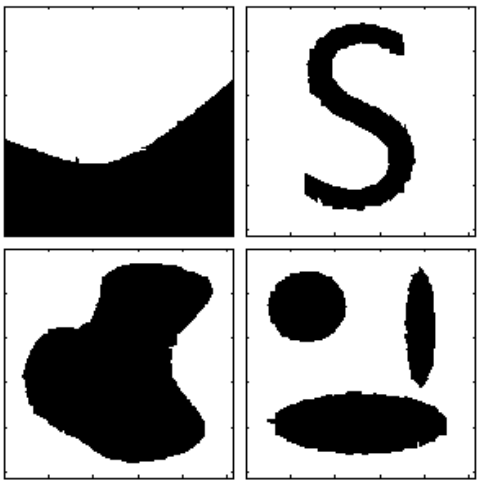


Figure 3.14: Reconstructed images obtained from noisy concentration images using modified PLBR.

CHAPTER 4

CONCLUSION

In Chapter 2, we showed that data storage and acquisition times in FT-IR imaging can be significantly reduced while preserving data quality by employing compressed sensing. We developed a sparse representation for absorbance spectra using a K-SVD learned dictionary. We modeled our measurement process as interferogram truncation, a feasible adjustment to the FT-IR data acquisition process. For downsample factors of four and eight, sparse reconstruction via ℓ_1-minimization provided PSNRs of 41.92 dB and 36.02 dB, respectively, while standard interpolation of the low-resolution measurements provided PSNRs of 36.93 dB and 34.37 dB, respectively. For future work we will additionally exploit spatial correlation in FT-IR data via transforms that sparsify the spatial domain (e.g. wavelet). Finally, sparse reconstructed FT-IR signals will be evaluated in cancer histopathology settings.

In Chapter 3, we introduced piecewise linear blockwise reconstruction (PLBR) as a tool for generating high-resolution discretizations of the classified image given images of per-pixel tissue concentrations. This allows us to generate classified images with spatial resolution greater than that of the concentration images (previously the resolution limit). In the noise-free case, PLBR significantly outperforms the standard approach of quantizing the per-pixel tissue concentrations, while the noisy case also provides higher-resolution classified images.

REFERENCES

[1] R. Bhargava, "Towards a practical Fourier transform infrared chemical imaging protocol for cancer histopathology," *Analytical Biochemistry*, vol. 389, no. 4, pp. 1155–1169, September 2007.

[2] I. Levin and R. Bhargava, "Fourier transform infrared vibrational spectroscopic imaging: Integrating microscopy and molecular recognition," *Annual Review of Physical Chemistry*, vol. 56, pp. 429–474, January 2005.

[3] P. R. Griffiths and J. A. de Haseth, *Fourier Transform Infrared Spectrometry*. Hoboken, NJ: John Wiley and Sons, Inc., 2007.

[4] J. Nascimento and J. Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, pp. 898–910, April 2005.

[5] S. Brady and M. Do, "Unsupervised unmixing of FT-IR data using vertex component analysis," poster presentation in *Biomedical Engineering Society Conference (BMES)*, St. Louis, MO, October 2008, P5.7.

[6] A. Bruckstein, D. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Review*, vol. 51, pp. 34–81, February 2009.

[7] D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, pp. 1289–1306, April 2006.

[8] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, pp. 489–509, February 2006.

[9] M. Parmar, S. Lansel, and B. A. Wandell, "Spatio-spectral reconstruction of the multispectral datacube using sparse recovery," in *15th IEEE International Conference on Image Processing (ICIP 2008)*, 2008, pp. 473–476.

[10] I. Pollak, A. Willsky, and H. Krim, "Image segmentation and edge enhancement with stabilized inverse diffusion equations," *IEEE Transactions on Image Processing*, vol. 9, pp. 256–266, February 2000.

[11] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, pp. 4311–4322, November 2006.