The Promise of Infrared Spectroscopy

Chandrajit Bajaj

Sponsored in part by NIH- R01GM117594, NIH-R41GM116300 and a grant from SETON-Dell Medical 201602388
Diversity of Infrared Imaging

- Measures scalar intensities (absorbance, reflectance, transmittance, fluorescence,…) across various spectral channels
Visible vs Mid-Infrared

- Cell/Tissue Classification for Cancer Detection

H&E stained image

Classified and pathologically verified map of the Tissue Sample: Breast Tissue BR1003,
Tissue (Chemical) Imaging at Cellular Resolution

Unstained Biopsy Sample

H&E Staining

Pathologist

Histology
Grade Therapy

Infrared Spectroscopic Imaging / Stimulated Raman Tomography

Computation

Absorbance (au)

Wavenumber (cm⁻¹)

Epithelium
Fibrous Stroma
Pale Stroma
Smooth Muscle
Stone
Blood
Endothelium

Epithelium
Fibroblast
Myofibroblast
Blood
Lymphocytes
Collagen
Necrosis

Computational Visualization Center (CVC)  http://cvcweb.ices.utexas.edu
Dept. of Computer Science / Institute for Computational Engineering and Sciences
University of Texas at Austin
Small vs Big Data Spectral Analysis

Image Classification / Segmentation

Typical spectra from five different classes of the core

0.005TB/day  1.3TB/day
Need for Noise Estimation and De-Noising

Classification suffers if data is Noisy

- Classifying Malignant Epithelium using Machine Learning
- Denoised Signal: 98-99% correct prediction
- Noisy Signal: 85-86% correct prediction
- Based on well distinguished spectral markers
Combining IR Spectroscopy + Machine Learning

Diagnostic Problem @ Multiple Scales

Current Instrumentation, Uncertainty Analysis, ...

Forward Imaging Model

Determine Signal, Noise Dimensionality, Parameter Sensitivity Analysis

Reconstruct Raw Data Inverse Model Analysis

Determine Sparsest Data Sampling for S/N and UQ

Augment/Build Instrumentation

Towards Smart Diagnosis Multi-scale Data Analysis with Quantified Uncertainty

Changing the paradigm of data acquisition → modeling → visualization

Computational Visualization Center (CVC)  http://cvcweb.ices.utexas.edu
Dept. of Computer Science / Institute for Computational Engineering and Sciences
University of Texas at Austin
Noise Estimation: Model and Simulate the Image Acquisition

Goal: optimize for image quality

Given a collection $\Psi$ of simulated samples with associated reference images $I_0(x, y, \nu; S)$ and probability distribution $\sum_{S \in \Psi} p(S) = 1$, determine the instrument configuration $\hat{c}$ that gives the "best" expected image

$$\hat{c} = \arg \min_c \left( \sum_{S \in \Psi} p(S) \sum_x \sum_y [I(x, y, \nu; S, c) - I_0(x, y, \nu; S)]^2 \right)$$

- Minimize expected deviation from reference image $I_0$
- Reference samples $S \in \Psi$ and associated images $I_0$ supplied by user
- Image function $I$ computed using instrument model
- Using expected difference over collection of samples avoids overfitting
- Choice of sample collection allows optimization for certain types of samples
Computational flow of imaging forward model

1. Input: Layered sample
   \( S = (\ldots (d_0(x,y,r), x_0(x,y)) \ldots) \)
   relative permittivity

2. Focused light source
   Discretized rays of light, directions

3. Homogeneous layer
   Write angular spectrum

4. Heterogeneous layer
   Solve coupled differential eqns.
   Any number of layers (Sample is any number of homogeneous and heterogeneous layers)

5. Solve electric/magnetic field
   Solve linear system of equations (banded matrix)

6. Objective lens
   Geometric optics: matrix multiply

7. Array detector
   Discretization with sensitivity and PSF; upgradable to CCD model

8. Output: \( I(x,y,z) \)
   image function

Parameter: \( \gamma \)
Iterate over wavenumbers
1. Input - Layered sample

- Model represents sample as layered medium
- Each layer may be modeled either homogeneously or heterogeneously
  - Homogeneous layers: $\varepsilon^{(l)}(\nu)$ depends only on wavenumber
  - Heterogeneous layers: $\varepsilon^{(l)}(x, y, \nu)$ depends on $x, y$ position as well
- Variation along the $z$ direction modeled through layers (discretization)
- Sample is described by relative permittivity $\varepsilon$ — a four-dimensional $(l, x, y, \nu)$ complex-valued function — and layer boundary positions $z^{(l)}$
- Sample $S \approx ((\varepsilon^{(1)}, z^{(1)}), (\varepsilon^{(2)}, z^{(2)}), \ldots, (\varepsilon^{(L)}, z^{(L)}))$
3-5. Light-sample interaction

The sample interaction model (steps 3–5) computes the electric and magnetic field in the sample.

- Generate constraint equations for each sample layer boundary
  - For homogeneous layers, get constraint equations directly (step 3)
  - For heterogeneous layers, solve differential system first (step 4)
- Solve linear system of $4(L - 1)N_F$ equations (step 5) to determine light leaving sample towards detector

$$E_{x,y}^{(l)}(x, y, z^{(l)}, \nu) = E_{x,y}^{(l+1)}(x, y, z^{(l)}, \nu)$$
$$H_{x,y}^{(l)}(x, y, z^{(l)}, \nu) = H_{x,y}^{(l+1)}(x, y, z^{(l)}, \nu)$$
4. Heterogeneous layer computation

- System of differential equations solved via eigenvalue method

\[
\begin{bmatrix}
\frac{dX(z, \nu)}{dz} \\
\frac{dY(z, \nu)}{dz} \\
\frac{dI(z, \nu)}{dz} \\
\frac{dJ(z, \nu)}{dz}
\end{bmatrix}
= i2\pi \nu \Phi(\nu)
\begin{bmatrix}
X(z, \nu) \\
Y(z, \nu) \\
I(z, \nu) \\
J(z, \nu)
\end{bmatrix}, \\
\Phi(\nu) = \begin{bmatrix}
* & * & * & * \\
* & * & * & * \\
* & * & * & * \\
* & * & * & *
\end{bmatrix} \in \mathbb{C}^{4N_F \times 4N_F}
\]

- Block-antidiagonal structure gives eigenvalues in opposite-sign pairs

\[0 = \det(\Phi - \lambda I) = \det \begin{bmatrix}
-\lambda I & \Phi_1 \\
\Phi_2 & -\lambda I
\end{bmatrix} = \det(\lambda^2 I - \Phi_1 \Phi_2)\]

- Solution is linear combination of eigenvalue pairs, eigenvectors
- Need eigenvalues, eigenvectors returned in opposite-sign pairs, not arbitrary order
Preliminary results: toluene

Top: Simulated by my implementation (solid line), reference implementation (dotted line).
Bottom: Observed spectrum (Coblentz No. 10130).

Preliminary results: hexane

Top: output from my implementation (solid line), reference implementation (dotted line).
Bottom: Observed spectrum (Coblentz No. 10118).
Given Noisy data \( Y = D + \epsilon \) under an unknown noise model \( \epsilon \), obtain an approximation \( \hat{D} \):

1. Maximize the Signal to Noise Ratio (SNR)
   \[ \hat{D} = \underset{D}{\text{argmax}} \frac{\text{var}(D)}{\text{var}(\epsilon)} \]

2. Preserve structural details in the image domain

3. Preserve spectral markers \( F \) (peak position, peak location, relative peak spacing)
   \[ F = [j, L] \quad j \in [1 : S], \; L \in \{0, 1\} \]
Metrics for De-Noising evaluation

- Spatial
  - Classification Accuracy
- Spectral
  - Preservation of spectral features
  - Spatial profile analysis
Challenges for De-Noising in Mid IR

- Most algorithms fail if the data is noisy
- Spectral markers get corrupted or lost
- Signal and noise maybe correlated
- Noise correlated across multiple channels*

\[ \varepsilon_s \sim \rho \cdot \varepsilon_{s-1} + \epsilon, \epsilon \sim \mathcal{N}(\mu, \sigma^2) \]

Prior De-Noising Techniques

- Filtering based approaches on spectrum *$+$
- Coefficients chosen by human experimentation
- Can handle any kind of noise
- Initial analysis is very time consuming
- Recalculate for every new dataset
- Good Classification accuracy

Prior De-Noising Techniques

- Signal is much stronger than noise
  - Signal and noise are independent
  - Noise across channels are independent
- Recover a low rank structure from the data where all the major signal contributions lie, and with with sparse noise structure
- Perform multi-modal iteration using Parallel Factor Analysis or Multidimensional Weiner based filters on the HSI cube
- Use a rank-1 approximated tensor decomposition approach
- Model the signal and noise as gaussian random fields and solve using Bayesian approach

\[ \delta_i = \exp(-c\sum_\delta(i, i)) \]

Minimum Noise Fraction (MNF)

- Developed by Green et al. *
- Can handle cases when signal and noise are uncorrelated
- Noise can be correlated within channels
- Orders data in terms of SNR in the MNF transformed space


**Algorithm 1 MNF\( (Y \in \mathbb{R}^{N \times S}) \)**

1. Estimate Data Covariance Matrix: \( \Sigma_Y = \text{Cov}\{Y\} \)
2. If \( \delta \) is unknown, estimate: \( \delta \sim Y(j, :) - Y(j + 1, :) \), \( \forall j = 1 : N - 1 \)
3. Estimate Noise Covariance Matrix: \( \Sigma_\delta = \frac{1}{2} \text{Cov}\{\delta\} \)
4. Calculate MNF projection vectors: \( (\Sigma_Y \Sigma_\delta^{-1})\Phi = \Lambda_{\text{MNF}} \Phi \)
5. Project data along MNF vectors: \( Y_{\text{MNF}} \in \mathbb{R}^{N \times S} = Y\Phi \)
6. Initialize \( R \in \mathbb{R}^{S \times S} \) as top rank-K matrix
7. Project data back: \( \tilde{D} \in \mathbb{R}^{N \times S} = Y_{\text{MNF}} R(\Phi^{-1}) = Y \Phi R(\Phi^{-1}) \)
Minimum Noise Fraction: Geometry

\[ \Sigma_Y = \Sigma_D + \Sigma_\delta \]

Data Covariance Structure

\[ \Sigma_\delta = E \Lambda E^T \]

Spectral Decomposition of Noise Covariance

\[ \Sigma_W = (E \Lambda_\delta^{-1/2})^T \Sigma_Y (E \Lambda_\delta^{-1/2}) \]

\[ = (E \Lambda_\delta^{-1/2})^T \Sigma_D (E \Lambda_\delta^{-1/2}) + (E \Lambda_\delta^{-1/2})^T \Sigma_\delta (E \Lambda_\delta^{-1/2}) \]

\[ = \Sigma_{W(D)} + I_\delta \]

Whitened Noise Covariance

\[ \Sigma_W = G \Lambda_{MNF} G^T \]

Spectral Decomposition of Whitened Data

**Figure**: Geometric Interpretation of Whitening. Left: Orientation of noise. Middle: Orientation of noise after multiplication with \( E \). Right: Orientation of noise after multiplication with \( \Lambda_\delta^{-1/2} \).
Minimum Noise Fraction: Fast MNF

\[ D = Y \ast \Phi \ast R \ast \Phi^{-1} = Y \ast (E \ast \Lambda^{-1/2}_\delta \ast G) \ast R \ast (E \ast \Lambda^{-1/2}_\delta \ast G)^{-1} = Y \ast E \ast \Lambda^{-1/2}_\delta \ast G \ast R \ast G^T \ast \Lambda^{1/2}_\delta \ast E^T = Y \ast E \ast \Lambda^{-1/2}_\delta \ast G \ast R \ast R^T \ast G^T \ast \Lambda^{1/2}_\delta \ast E^T = Y \ast \hat{\Phi} \ast \tilde{\Phi}^T \] //fastMNF

\[ \Rightarrow R = \begin{bmatrix} I_K & 0 \\ 0 & 0 \end{bmatrix} \]

\[ \Rightarrow \hat{\Phi} = E \ast \Lambda^{-1/2}_\delta \ast G \ast R \] //forward MNF transform
\[ \Rightarrow \tilde{\Phi} = E \ast \Lambda^{1/2}_\delta \ast G \ast R \] //inverse MNF transform

- Previously slow due to inversion of large matrices
- Current formulation avoids any inverse
- Uses rank-K approximation of G

Algorithm 1 MNF(\( Y \in \mathbb{R}^{N \times 5} \))

1. Estimate Data Covariance Matrix: \( \Sigma_Y = \text{Cov}(Y) \).
2. If \( \delta \) is unknown, estimate: \( \delta \sim Y(j, :) - Y(j + 1, :) \), \( \forall j = 1 : N - 1 \).
3. Estimate Noise Covariance Matrix: \( \Sigma_e = \frac{1}{2} \text{Cov}(\delta) \).
4. Calculate MNF projection vectors: \( (\Sigma_e \Sigma^{-1}_\delta) \Phi = \Lambda_{\text{MNF}} \Phi \).
5. Project data along MNF vectors: \( Y_{\text{MNF}} \in \mathbb{R}^{N \times 5} = Y \Phi \).
6. Retain \( R \in \mathbb{R}^{5 \times 5} \) as top rank-K matrix.
7. Project data back: \( D \in \mathbb{R}^{N \times 5} = Y_{\text{MNF}} R(\Phi^{-1}) = Y \Phi R(\Phi^{-1}) \).
Variations on Minimum Noise Fraction: Fast MNF
- Block Lanczos $\#$ method to compute a rank $K'$ SVD
- Guarantees $(1+\epsilon)$ Frobenius and $(1+\epsilon)$ spectral norm approximation
- Guarantees $\epsilon$ per vector norm approximation
- Although the Block Lanczos algorithm can attain machine precision
- Block Lanczos is rel.slow when the matrix is large
- A faster randomized and memory efficient version &
- Computes the $K'$-SVD up to $(1+\epsilon)$ Frobenius norm relative error
- Only keeps a $N \times O(K'/\epsilon)$ sketch in memory

- Approx MNF
- Rand MNF

---


---

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Time</th>
<th>Space</th>
</tr>
</thead>
<tbody>
<tr>
<td>MNF</td>
<td>$O(S^3 + NS^2)$</td>
<td>$O(NS + S^2)$</td>
</tr>
<tr>
<td>Fast MNF</td>
<td>$O(S^3 + NSK)$</td>
<td>$O(NK + S^2)$</td>
</tr>
<tr>
<td>Approx MNF</td>
<td>$O(S^2K + NSK)$</td>
<td>$O(NK + SK)$</td>
</tr>
<tr>
<td>Rand MNF</td>
<td>$O(nnz(W)K + NSK)$</td>
<td>$O(NK + SK)$</td>
</tr>
</tbody>
</table>

Minimum Noise Fraction: Profile Analysis
Minimum Noise Fraction: Classification Analysis

A

B

C

D

E

F

- Normal Epithelium
- Malignant Epithelium
- Loosestoma
- Reactivestroma
- Densestoma
- Others

FPR (1 - sensitivity) vs Sensitivity

- Standard MNF: 0.98
- Fast MNF: 0.99
- Approx MNF: 0.98
- Rand MNF: 0.98
- No MNF: 0.86
Next Steps: Super-Resolution IR

Breast Tissue
Visualization and chemical characterization of intralobular stroma.

Prostate Tissue
Visualization and chemical characterization of collagen bands

Colon Tissue
Visualization and chemical characterization of subcellular mucin


Collaboration with Rohit Bhargava, UIUC
Next Steps: Tunable Quantum Cascade Laser (Compressed Sensing)

**Frontier – Imaging with QCLs**
- Excellent SNR ($10^4$); $10^5$ pix/s.$\Delta v$
- Ideal for ML approaches $\Rightarrow$ use discrete frequency data for classification
- Limited frequency range
SMART DATA ANALYSIS: Tumor Cytotyping, Tracking Progression in 3D with Molecular-Cell Precision

Developing computational tools convert chemical imaging data to knowledge, as shown here for identifying all cells in prostate tissue (E).

(D) nucleic acids (left, at 1080cm⁻¹) and collagen specific (right, at 1245cm⁻¹).

Learning the Heterogeneity in cellular morphology, with a reduction in cell polarity during branching morphogenesis.

(A) acinar org. of primary organoid cultures with a well-defined lumen
(B) 200 micron x 200 micron field of view of a single 100 nm thin resin section imaged at 1 nm resolution of an organoid undergoing branching morphogenesis
(C) HMT-3522 S1 acini (left) and T4 aggregates (right).
(D) S1 cells form growth-arrested acini with extensive intercellular membrane network and reduction in cell polarity, illustrating the label-free resolving power of the electron microscope.
DETAILS
Sources for hyperspectral images

- **Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) program**
  - Spatial resolution 20 meters at 20 km altitude, 4 meters at 4 km
  - 0.4-2.5 µm

- **Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Spectral Library**
  - 2000 spectra of minerals, rocks, soils, water
  - 0.4-14 µm

- **USGS Spectral Library**
  - 500 spectra of minerals and a few plants
  - 0.2-3.0 µm
Reflectance vs. radiance

Laboratory instruments observe reflectance (or absorption), while remote sensing observes radiance, capturing several effects which must be corrected for.

- Spectrum of illuminating (solar) light
  - Additionally affected by shadows

- Light interactions with atmosphere
  - e.g. absorbance by water vapor, CO₂

- Illumination geometry (angle of incidence)
  - Varies by time of day and season

- Sensor characteristics
  - Variations between sensors, temporal changes
Hyperspectral image acquisition

- Imaging spectrometers capture hyperspectral images

- Remote sensing
  - Analysis of the surface of the Earth (or other planets)
  - Optical dispersing element to separate frequencies
  - Can have spectral resolution as fine as 0.01 µm
  - Broadband (solar) light source

- Hyperspectral images also used in microscopy
  - Focal plane array detector
  - FTIR spectroscopy using interferometer light source
Hyperspectral instrument specifications

<table>
<thead>
<tr>
<th></th>
<th>AVIRIS</th>
<th>ASTER</th>
<th>Agilent Cary FTIR</th>
<th>QCL DFIR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Image size</strong></td>
<td>677 pixels wide</td>
<td>4980 pixels wide (V/NIR)</td>
<td>128 × 128 pixels</td>
<td>128 × 128 pixels</td>
</tr>
<tr>
<td><strong>Spatial resolution</strong></td>
<td>20m (at 20km alt.)</td>
<td>15m (V/NIR)</td>
<td>1.1 μm</td>
<td>0.95 μm</td>
</tr>
<tr>
<td></td>
<td>4m (at 4km alt.)</td>
<td>30m (SWIR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>90m (MidIR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Spectral resolution</strong></td>
<td>0.01 μm in 0.4-2.5 μm</td>
<td>14 bands increments of 0.4-14 μm</td>
<td>0.0002 μm increments of 1.1-28.5 μm</td>
<td>0.002 μm in 5.25-12.87 μm</td>
</tr>
<tr>
<td><strong>Detector</strong></td>
<td>Si (Vis)</td>
<td>Si (V/NIR)</td>
<td>DLaTGS or HgCdTe (MCT)</td>
<td>HgCdTe (MCT)</td>
</tr>
<tr>
<td></td>
<td>InGaAr (NIR)</td>
<td>PtSi-Si (SWIR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>InSb (SWIR)</td>
<td>HgCdTe (MidIR)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SNR</strong></td>
<td>100 (at 0.49)</td>
<td>varies</td>
<td>10,000</td>
<td>260 (single)</td>
</tr>
</tbody>
</table>
Research and commercial imaging spectrometers

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Organization</th>
<th>Country</th>
<th># bands</th>
<th>Wavelengths</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVIRIS</td>
<td>NASA</td>
<td>United States</td>
<td>224</td>
<td>0.4 - 2.5 µm</td>
</tr>
<tr>
<td>AISA</td>
<td>Spectral Imaging Ltd</td>
<td>Finland</td>
<td>286</td>
<td>0.45 - 0.9 µm</td>
</tr>
<tr>
<td>CASI</td>
<td>Itres Research</td>
<td>Canada</td>
<td>288</td>
<td>0.43 - 0.87 µm</td>
</tr>
<tr>
<td>DAIS 2115</td>
<td>GER Corp</td>
<td>United States</td>
<td>211</td>
<td>0.4 - 12.0 µm</td>
</tr>
<tr>
<td>HYMAP</td>
<td>Integrated Spectronics Pty Ltd</td>
<td>Australia</td>
<td>128</td>
<td>0.4 - 2.45 µm</td>
</tr>
<tr>
<td>PROBE</td>
<td>Earth Search Sciences</td>
<td>United</td>
<td>128</td>
<td>0.4 - 2.45 µm</td>
</tr>
</tbody>
</table>
Aerobic glycolysis inhibitors
Proapoptotic BH3 mimetics
PARP inhibitors
Inhibitors of VEGF signaling
Inhibitors of HGF/c-Met
EGFR inhibitors
Cyclin-dependent kinase inhibitors
Immune activating anti-CTLA4 mAb
Telomerase Inhibitors
Immune activating anti-inflammatory drugs
Sustaining proliferative signaling
Resisting cell death
Inducing angiogenesis
Activating invasion & metastasis
Deregulating cellular energetics
Genome instability & mutation
Avoiding immune destruction
Enabling replicative immortality
Tumor-promoting inflammation
Enabling replicative immortality
Sustaining proliferative signaling
Evading growth suppressors
Avoiding immune destruction
Enabling replicative immortality
Minimum Noise Fraction: Automatic Band Selection

- Previous works include:
  - Manual inspection of Eigen-images
  - Automatic selection of Eigen-images based on its RMSE error to a clear image
  - Computationally expensive
  - Time Consuming
- The optimal value of K can be determined from the diagonal entries of $\Lambda_{MNF} = SNR + 1$
- The Rose criteria & states that an SNR of at least 5.0 is needed to be able to distinguish image features at 100% certainty.

Algorithm 1 $\text{MNF}(Y \in \mathbb{R}^{N \times S})$

1. Estimate Data Covariance Matrix: $\Sigma_Y = \text{Cov}(Y)$
2. If $\delta$ is unknown, estimate: $\delta \sim Y(j, :) - Y(j + 1, :)$, $j = 1: N - 1$
3. Estimate Noise Covariance Matrix: $\Sigma_N = \frac{1}{2} \text{Cov}(\delta)$
4. Calculate MNF projection vectors: $(\Sigma_Y \Sigma_N^{-1})\Phi = \Lambda_{MNF} \Phi$
5. Project data along MNF vectors: $Y_{MNF} \in \mathbb{R}^{N \times S} = Y\Phi$
6. Initialize $R \in \mathbb{R}^{S \times S}$ as top rank-K matrix
7. Project data back: $D \in \mathbb{R}^{N \times S} = Y_{MNF} R(\Phi^{-1}) = Y\Phi R(\Phi^{-1})$

---


Blind Error Metric

- Absence of ground truth

- The Method Noise Image (MNI)*
  - No-reference metric, simple and easy to use
  - Based on Structural Similarity Index Metric% (SSIM)
  - Scores based on intensity, contrast, Image moments
  - Maximum score around highly structured regions