Dimensionality reduction of MALDI Imaging datasets using non-linear redundant wavelet transform-based representations

Luis Mancera\textsuperscript{1}, Lyna Sellami\textsuperscript{2}, Jamie Cunliffe\textsuperscript{2}, Luis González\textsuperscript{1}, Omar Belgacem\textsuperscript{2}

\textsuperscript{1}Clover Bioanalytical Software, Spain
\textsuperscript{2}Shimadzu, UK
Content

• The problem of meaningful reduction
• The sparse representation problem
• Iterative hard thresholding algorithm
• Avoiding local minima
• Results on data compression
• Conclusions
• Future Work
Meaningful reduction

• Mass Spectrometry Imaging (MSI) experiments very often need a reduction in dimensionality

• Meaningless vs. meaningful reduction:

Original message

What are you gonna eat?

A nice dish of yellow seasoned rice with lots of toppings

What your compression algorithm does

What are you gonna eat?

A nice dish of yello ssnd rice w lots of toppings
Meaningful reduction

• Mass Spectrometry Imaging (MSI) experiments very often need a reduction in dimensionality

• Meaningless vs. meaningful reduction:

Original message

What are you gonna eat?

A nice dish of yellow seasoned rice with lots of toppings

Meaningful compression

What are you going to eat?

Paella
Meaningful reduction

- The target is to have a *sparse* representation, e.g., binning...

- Four features, four meaningful components
Redundant wavelets

\textbf{A MALDI image}  \hspace{1cm}  256 \times 256

\textbf{Wavelet} [Mallat 89]:
- Well adapted to represent some typical features of the images (scale invariance, local orientations)
- Band-pass, multi-scale, multi-orientation

\textbf{X-lets (Redundancy)}:
- Invariance to translation
- Better orientation selectivity
- Better compaction of energy
- Curvelets [Candes et al 99], DT-CWT [Kingsbury 01], Steerable Pyramid [Simoncelli et al 92],...
The sparse representation problem

\( \mathbf{x} \): A spectrum as a vector of \( N \) elements

\( \Phi^T \): A \( M \times N \) linear transformation \( M > N \)

\( \text{range}(\Phi^T) = N \)

\( \Phi^T \) is a Parseval frame

Are there any solutions in \( \mathbf{a} \) to the equation:

\[ \Phi \mathbf{a} = \mathbf{x} \quad \Rightarrow \quad \text{YES, INFINITE} \]

Solving the **sparse representation problem** means looking for the sparsest solution:

\[ \hat{\mathbf{a}} = \min_{\mathbf{a} \in \mathbb{R}^M} \| \mathbf{a} \|_0 \; \text{s.t.} \; \Phi \mathbf{a} = \mathbf{x} \]

The \( l_0 \)-norm counts the number of non-zero elements in a vector
The sparse representation problem in MSI

• 3 approaches for MSI experiments:
  – **Spectrum-based**
    • Each single spectrum gets compressed independently
  – **Image-based**
    • Each single MALDI image from a selected mass list gets compressed independently
  – **Dataset-based**
    • The whole dataset is represented as a 2D matrix with pixels as rows and selected masses as columns
    • This dataset is compressed as a whole
    • Advantage: Takes into account dependencies across all directions
### Iterative hard thresholding

- **Example using:**
  1. $B_0(1)$: $L0$-ball of dimension 1. All those vectors with just one non-zero element
  2. $S(\Phi, x)$: set of perfect reconstruction

- Alternated orthogonal projections: local minimum of the distance from $B_0(K)$ to $S(\Phi, x)$
  - Convergence can be proved
    [Blumensath & Davies, 09]

- The solution provides a perfect representation of the input as close as possible to the limits of the $L0$-ball of radius $K$. 

\[
N = 2, \quad M = 3, \quad K = 1
\]
Iterative Hard Thresholding algorithm

1. Set a level of sparseness $K$
   This is equivalent to setting a threshold $\Theta$

2. Get the minimum energy solution $a^{(0)} = a_{LS} = \Phi^T x$

3. Repeat until convergence:
   
   1. $b^{(k+1)} =$ Project $a^{(k)}$ onto $B_0(K)$:
      
      Set to zero all elements with intensity lower than $\Theta$

   2. Project $b^{(k+1)}$ onto $S(\Phi, x)$:

      $a^{(k+1)} = b^{(k+1)} + \Phi^T (x - \Phi b^{(k+1)})$
Using dynamic thresholding

Intuitive justification of dynamic thresholding:

Multiple minima:
Difficult global optimization

Find favourable optima searching from good candidates at higher $\theta$s
Compressing single spectra

• $\beta = 0.9$, Dual-Tree Complex Wavelets, 8 levels, x2 redundancy factor

• Basepeak spectrum from a mouse brain tissue
  – Basepeak spectrum chosen for illustration because it typically shows more number of peaks and is typically more noisy than the ROI spectrum

• Results obtained using Matlab™ R2016a on an Intel® Core™ i5-3230 @ 2.60 GHz
Compressing single spectra

Original (179200 data points)

From top 230 wavelet coeffs (99.87% reduction)
Compressing single spectra

Original (zoomed in to 700-800 Da)

From top 230 wavelet coeffs

(99.87% reduction)
Compressing single spectra

Original (179200 data points)

From top 1432 wavelet coeffs (99.20% reduction)
Compressing single spectra

Original (zoomed in to 700-800 Da)

From top 1432 wavelet coeffs
(99.20% reduction)
Compressing single spectra

Original (179200 data points)

From top 2214 wavelet coeffs
(98.76% reduction)
Compressing single spectra

Original (zoomed in to 700-800 Da)

From top 2214 wavelet coeffs (98.76% reduction)
Compressing single spectra

Original (179200 data points)

From top 3672 wavelet coeffs (97.95% reduction)
Compressing single spectra

Original (zoomed in to 700-800 Da)

From top 3672 wavelet coeffs (97.95% reduction)
Compressing all spectra in a dataset

<table>
<thead>
<tr>
<th>Dataset (ROI)</th>
<th>Number of pixels</th>
<th>Number of data points per pixel</th>
<th>Size in disk (ASCII)</th>
<th>Avg. num of reduced data points</th>
<th>Reduced size in disk (ASCII)</th>
<th>Avg. % of reduction</th>
<th>Avg. accuracy of the representation (MSE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mouse brain 1</td>
<td>10397</td>
<td>292,864</td>
<td>35.4 GB</td>
<td>13,718</td>
<td>2.49 GB</td>
<td>92.97%</td>
<td>2.8·10^{-2}</td>
</tr>
<tr>
<td>Mouse brain 2</td>
<td>11976</td>
<td>179,200</td>
<td>24.95 GB</td>
<td>7,780</td>
<td>1.62 GB</td>
<td>95.66%</td>
<td>2.4·10^{-2}</td>
</tr>
<tr>
<td>Mouse brain 3</td>
<td>8976</td>
<td>154,300</td>
<td>16.10 GB</td>
<td>10,452</td>
<td>1.64 GB</td>
<td>93.23%</td>
<td>1.99·10^{-2}</td>
</tr>
<tr>
<td>Averaged</td>
<td>10450</td>
<td>208,788</td>
<td>25.48 GB</td>
<td>10,650</td>
<td>1.91 GB</td>
<td>93.95%</td>
<td>2.4·10^{-2}</td>
</tr>
</tbody>
</table>
Compressing single spectra

• Further compression with traditional algorithms is possible

• Alternative convex relaxation algorithm provides a total average of 89.54%

• Execution time averaged approx. 9 seconds per spectrum on a non-optimised script in Matlab™
  – Win64 Intel® Core™ i5-3230M CPU @ 2.60 GHz
  – Most of the time spent on Analysis/Synthesis operations

• The average compression done image-wise is lower, averaging roughly 75% for images showing “some” structure
Application to other problems

• Application to other problems need to reformulate the problem to look for the sparsest approximation

• Many applications might also need to redefine the concept of perfect reconstruction to that of compatibility to the observation
Conclusions

• We have presented a method to find sparse (compressible) representations of imaging datasets using redundant wavelet transformations

• The large dictionary forming a redundant wavelet basis provides meaningful representation of features in relatively few number of coefficients

• We have applied the Iterative Hard Thresholding algorithm, shown previously to be superior to other alternatives to find sparse representations [Blumensath & Davies 08, Mancera & Portilla 08]

• We have used a technique inspired by simulated annealing to avoid local minima of the non-convex quasi-norm used in the cost function and remove the need to estimate the level of sparsity required

• Results show high compression ratios, with increasingly better approximations as more coefficients are used
Future work

• Need to speed up the iterations
  – Approaches to reduce the number of iterations needed
  – Code optimisation
• Adaptive threshold strategy
  – Deal differently with different mass range
• Potential applications include:
  – Spectra smoothing
  – Peak detection
  – Image denoising
  – Restoration after vector quantization
  – Improving image resolution
  – Multivariate analysis
  – …
Thanks for your attention!

Acknowledgments

- Dr. Luis González,
- Dr. Emmanuel Wey,
- Dr. Gerald Stubiger

- Jamie Cunliffe
- Dr. Lyna Sellami
- Dr. Omar Belgacem
- Dr. Matt Oppenshaw
- Roberto Raso

Clover Bioanalytical Software, S.L.
Av. de la Innovación, 1
18016 Granada, Spain

Kratos Analytical Ltd
Trafford Wharf Road
M17 1GP, Manchester UK