Cardiac Motion Estimation by Using Convolutional Sparse Coding

PhD(c) Nelson Diaz
Ph.D Adrian Basarab
Ph.D Jean-Yves Tourneret
Advisor: Ph.D Henry Arguello

1Department of Electrical and Computer Engineering,
2Department of Computer Science
Universidad Industrial de Santander, Bucaramanga, Colombia.
3University of Toulouse, Toulouse, Francia.

March 5, 2019
Motivation

- Optimal treatment of cardiac disease requires early detection of abnormalities and accurate monitoring tools.
Introduction to convolutional sparse coding (CDL)

Fig. 1: The convolutional model description, and its composition in terms of the local dictionary $D_L$.

Fig. 2: Stripe Dictionary
\( s_k \) is modeled as a convolution between the coefficient maps \( x_m \) and a set of \( M \) filters \( d_m \).

\[
s_k \approx \sum_{m=1}^{M} d_m \ast x_m
\]
Dictionary filters and coefficient maps

A dictionary is estimated off-line by using a set of training cardiac motions denoted as $\hat{s}$.

$$\arg\min_{d_m,x_{k,m}} \frac{1}{2} \sum_k \left\| \sum_m x_{k,m} * d_m - s_k \right\|^2_2 + \lambda \sum_m \sum_k \|x_{k,m}\|_1$$  \hspace{1cm} \text{(2)}

$$\text{s.t.} \quad \|d_m\| = 1 \quad \forall m = 1, \ldots, M$$

The coefficient maps $x_m$ are computed from cardiac motions $s_k$

$$\arg\min_{x_m} \frac{1}{2} \left\| \sum_m x_m * d_m - s_k \right\|^2_2 + \lambda \sum_m \|x_m\|_1.$$  \hspace{1cm} \text{(3)}

Eq (2) and (3) can be solved using alternating direction method of multipliers (ADMM).
Motion estimation

- A pair of successive frames \((r_k, r_{k+1}) \in \mathbb{R}^{J \times N}\) acquired at time instants \(k\) and \(k + 1\)

- \((s_x, s_y) \in \mathbb{R}^{J \times N}\) where \(s_x\) and \(s_y\) are the motions along the \(x\) and \(y\) axes.

- Since the motion estimation problem is considered independently the displacement vector is equal to \(s = s_x\) or \(s = s_y\).

- The motion estimation field is formulated as the minimization of a cost function with energy \(E_{\text{data}}(s)\) penalized by spatial and sparse regularizations, i.e.,

\[
\arg\min_{x,s}\{E_{\text{data}}(s) + \lambda_d E_{\text{sparse}}(s, x) + \lambda_s E_{\text{spatial}}(s)\}
\]

(4)
Data fidelity

The ML estimator is classically computed in the negative log-domain

$$\arg\min_s - \ln[p(r_{k+1}|r_k(n))]. \quad (5)$$

Straightforward computations exploiting the Rayleigh statistics of ultrasound imaging detailed in [1] lead to the following data fidelity term

$$E_{\text{data}}(s) = -2d(s) + 2 \log[e^{2d(s)} + 1] + C \quad (6)$$

where

$$d(s) = \frac{1}{b} \sum_{n=1}^{N} [r_{k+1}(n + s(n)) - r_k(n)],$$

$n$ indicates the pixel index, $s = [s(1), \ldots, s(N)]^T$ is the vectorized motion, and $r_k = [r_k(1), \ldots, r_k(N)]^T$ is the vectorized ultrasound image in frame $k$, and $C = - \log(2\sigma^4/b)$ is a known constant.
Regularization term

The spatial regularization term promotes the smoothness of the motion estimation field and is defined as

$$E_{\text{spatial}}(s) = \| \nabla s \|^2_2$$ (7)

The proposed sparse regularization determines the motion $s_k$ that is best represented as a convolution between $M$ filters $d_m$ and the coefficient maps $x_m$, i.e,

$$E_{\text{sparse}}(s) = \left\| s_k - \sum_{m=1}^{M} x_m \ast d_m \right\|^2_2.$$ (8)
Motion estimation algorithm

Input: \(r_{b,1}, r_{b,2}, \lambda_s, \lambda_d, K, J, \lambda, \rho, \tilde{s}_0 = \text{LADdist motions}, \hat{s}_0 = \text{LADprox motions}\)

Output: \(s\)

1: function \(\text{MEFCDL}(r_{b,1}, r_{b,2}, \lambda_s, \lambda_d, K, J, \lambda, \rho, \tilde{s}_0, \hat{s}_0)\)
2: \(d_m \leftarrow \text{Computes the dictionary by solving (2)}\)
3: \(x_m \leftarrow \text{Computes the coefficient maps by solving (3)}\)
4: for \(k \leftarrow 1, K\) do
5:    for \(j \leftarrow 1, J\) do
6:       \(\arg\min_s \left\{ E_{\text{data}}(r_{b,1}, r_{b,2}, s_{j-1}) + \lambda_s \| \nabla s_{j-1} \|^2_2 + \lambda_d(k) \| s_{j-1} - \sum_m d_m * x_m \|^2_2 \right\} \)
       \(\text{s.t.} \quad \| d_m \| = 1 \ \forall m \) \quad \triangleright \text{Motion estimation}
7:          return \(s\) \quad \triangleright \text{(Estimated motion field)}
## Experimentation setup

**Table 1:** Parameters for each step of algorithm 7, dictionary learning, sparse coding, and cardiac motion estimation.

<table>
<thead>
<tr>
<th>Step</th>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dictionary learning</strong></td>
<td>Database</td>
<td>LADdist</td>
</tr>
<tr>
<td></td>
<td>Filter size</td>
<td>16 × 16</td>
</tr>
<tr>
<td></td>
<td>Filters number</td>
<td>( M = 48 )</td>
</tr>
<tr>
<td></td>
<td>Sparsity term</td>
<td>( \lambda = 0.001 )</td>
</tr>
<tr>
<td></td>
<td>Number of iteration</td>
<td>500</td>
</tr>
<tr>
<td><strong>Sparse coding</strong></td>
<td>Database</td>
<td>LADprox</td>
</tr>
<tr>
<td></td>
<td>Number of iteration</td>
<td>500</td>
</tr>
<tr>
<td><strong>Cardiac motion estimation</strong></td>
<td>Regularization parameter</td>
<td>( \lambda_s = 0.75 )</td>
</tr>
<tr>
<td></td>
<td>Sparsity term (Systole)</td>
<td>( \lambda_d = {1 \times 10^{-6} \times 10^{-3}} )</td>
</tr>
<tr>
<td></td>
<td>Sparsity term (Diastole)</td>
<td>( \lambda_d = {1 \times 10^{-9} \times 10^{-2}} )</td>
</tr>
</tbody>
</table>
### Parameters $\lambda$ and $\rho$

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$\rho_0$</th>
<th>$\rho_1$</th>
<th>$\rho_2$</th>
<th>$\rho_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_0 = 25$</td>
<td>28.9992</td>
<td>29.2101</td>
<td>28.6905</td>
<td><strong>29.3802</strong></td>
</tr>
<tr>
<td>$\rho_1 = 50$</td>
<td>24.9587</td>
<td>24.9932</td>
<td>25.2910</td>
<td>25.2395</td>
</tr>
<tr>
<td>$\rho_2 = 100$</td>
<td>20.9748</td>
<td>21.1248</td>
<td>21.2963</td>
<td>21.3028</td>
</tr>
<tr>
<td>$\rho_3 = 150$</td>
<td>18.7247</td>
<td>18.8688</td>
<td>18.8703</td>
<td>18.8667</td>
</tr>
<tr>
<td>$\rho_4 = 0.5$</td>
<td>16.1208</td>
<td>16.1271</td>
<td>15.9824</td>
<td>16.0115</td>
</tr>
</tbody>
</table>

**Table 2:** Image quality (PSNR) with 100 iteration of (CBPDN) for different choices of $\lambda$ and $\rho$ by using the motion horizontal ground-truth.

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$\rho_0$</th>
<th>$\rho_1$</th>
<th>$\rho_2$</th>
<th>$\rho_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_0 = 25$</td>
<td>26.9361</td>
<td>27.2072</td>
<td>27.2159</td>
<td><strong>27.3542</strong></td>
</tr>
<tr>
<td>$\rho_1 = 50$</td>
<td>22.8906</td>
<td>22.8575</td>
<td>22.9179</td>
<td>23.1302</td>
</tr>
<tr>
<td>$\rho_2 = 100$</td>
<td>19.2978</td>
<td>19.2049</td>
<td>19.3376</td>
<td>19.3352</td>
</tr>
<tr>
<td>$\rho_3 = 150$</td>
<td>17.3294</td>
<td>17.4144</td>
<td>17.4043</td>
<td>17.3599</td>
</tr>
<tr>
<td>$\rho_4 = 0.5$</td>
<td>15.1404</td>
<td>15.2372</td>
<td>15.2145</td>
<td>15.1526</td>
</tr>
</tbody>
</table>

**Table 3:** Image quality (PSNR) with 100 iteration of (CBPDN) for different choices of $\lambda$ and $\rho$ by using the motion horizontal ground-truth.
3 dictionaries, 3 scenarios

(a) (b)
Dictionary for each frame
Horizontal Vertical

(c) (d)
Dictionary for all sequence

(e) (f)
Dictionary for systole motion

(g) (h)
Dictionary for diastole motion
This error is defined for the $n$th pixel as

$$e_n = \sqrt{[s_x(n) - \hat{s}_x(n)]^2 + [s_y(n) - \hat{s}_y(n)]^2}, \quad (9)$$

where $s_x(n)$, $s_y(n)$, $\hat{s}_x(n)$, $\hat{s}_y(n)$ are the true and estimated (horizontal and vertical) motions at pixel $n$. 
Error maps estimation

Frame 5 motion error DL method

Frame 5 motion error CDL proposed
Estimation motion comparison

Frame 5 ground-truth zoom version

Frame 5 DL zoom version

Frame 5 CDL zoom version

Frame 5 ground-truth

Frame 5 DL

Frame 5 CDL
Clustering the PC of $x$ (kmeans with 2 classes)
PCA for multispectral imaging

- Implement the principal component transform to compress and reconstruct multispectral images.
PCA for multispectral imaging 2

- Understand the applications of satellital images in agriculture, geology and surveillance (you will need to understand which is the structure multispectral images)
Understand the mathematics that PCA involves.

Compute the mean vector \( \mathbf{m}_x = \frac{1}{K} \sum_{k=1}^{K} \mathbf{x}_k \) and the covariance matrix \( \mathbf{C}_x = \frac{1}{K} \sum_{k=1}^{K} \mathbf{x}_k \mathbf{x}_k^T - \mathbf{m}_k \mathbf{m}_k^T \). Let \( \mathbf{A} \) a matrix whose rows are formed from the eigenvectors of \( \mathbf{C}_x \) ordered in descending order,

\[
\mathbf{y} = \mathbf{A}(\mathbf{x} - \mathbf{m}_x) \tag{10}
\]

Because the rows of \( \mathbf{A} \) are orthonormal vectors, it follows that \( \mathbf{A}^{-1} = \mathbf{A}^T \), and any vector \( \mathbf{x} \) can be recovered from its corresponding \( \mathbf{y} \) by using the expression

\[
\mathbf{x} = \mathbf{A}^T \mathbf{y} + \mathbf{m}_x \tag{11}
\]

Using just \( k \) eigenvectors

\[
\hat{\mathbf{x}} = \mathbf{A}_k^T \mathbf{y} + \mathbf{m}_x \tag{12}
\]
Principal components

- Six principal component images, obtained using equation (5).
Clustering the PC of $\mathbf{x}$ (kmeans with 2 classes)
High-correlation between $x$ and $s$
Clustering the PC of $\mathbf{x}$ (kmeans with 6 classes)
Products

- It was submitted the paper entitled: Cardiac motion estimation by using convolutional sparse coding.
Conclusions and future work

- A new method for cardiac motion estimation in ultrasound imaging was presented.
- In the future will be studied the sparse decomposition for anomaly detection and cardiac tissue classification.
Questions?
Thanks for your attention!