Motion Estimation and Classification in Compressive Sensing from Dynamic Measurements
Aug 28, 2014

Vijay Rengarajan, A.N. Rajagopal, R. Aravind

Department of Electrical Engineering
Indian Institute of Technology Madras
http://www.ee.iitm.ac.in/ipcvm
Compressed sensing (CS) - an overview

Motion in CS

Motion estimation using CS measurements
  - Estimation of a single warp
  - Blockwise motion estimation

Application: Compressive classification
DMD : Digital micromirror device
RNG : Random number generator
Measurements:

\[ y = \Phi x \quad (1) \]

Reconstructions:

\[ \hat{x} = \Psi \hat{\alpha}, \]
\[ \hat{\alpha} = \arg \min_{\alpha} \| \alpha \|_1 \text{ subject to } y = \Phi x \text{ and } x = \Psi \alpha \quad (2) \]

\[ x: \text{ signal/image } \in \mathbb{R}^N \]
\[ y: \text{ measurement vector } \in \mathbb{R}^M, M \ll N \]
\[ \Phi: \text{ measurement matrix } \in \mathbb{R}^{M \times N} \]
\[ \Psi: \text{ basis, in which } x \text{ is sparse } \]
\[ \alpha: \text{ basis coefficients } \in \mathbb{R}^N \]
Sequential acquisition of measurements in a compressed sensing (CS) camera, such as a single pixel camera, lures temporal artifacts. Manifestation of such artifacts is different from that of a conventional optical camera. Relative motion between the camera and the scene renders the CS reconstruction process noisy. We estimate the relative motion using a CS measurement vector against a CS reference vector (which is acquired when there is no motion). We assume a planar scene captured with only a global relative motion.
Davenport et al. (2007) *Electronic Imaging*
- Compressive classification problem is discussed in which the probe measurement is considered to be affected by only a single warp
- Gallery contains CS measurements of all possible warps of images

Duarte et al. (2007) *ICIP*
- CS measurements are acquired at various image resolutions, and a multi-scale classification is discussed

Both the methods consider the case where the observed CS vector is affected by a single warp
MOTION IN CS

\[ x_1 \rightarrow \phi_1 \rightarrow x_2 \rightarrow \phi_2 \rightarrow \ldots \rightarrow x_M \rightarrow \phi_M \]

\[ y_1 \rightarrow \phi_1 \rightarrow y_2 \rightarrow \phi_2 \rightarrow \ldots \rightarrow y_M \rightarrow \phi_M \]

no motion

CS

conventional
WARP ESTIMATION

- Given an image $x$ and its warped version $x_p$, it is possible to estimate the warp $p$ using an energy minimisation based on Taylor series

$$\hat{p} = \arg\min_p \|x_p - x(p)\|_2^2$$

with $x(p) = x + \nabla x \cdot J(x) \cdot p$

- An iterative algorithm could be developed such that the residual energy $\|x_p - x(p)\|_2^2$ reduces in each iteration (Baker and Matthews, 2003)

- **Problem**: Given $y$, CS measurement with no motion, and $y_p$, CS measurement with a single warp, how to estimate the parameter vector $p$?

- Is it possible to develop an algorithm to iteratively estimate $p$? Will it converge?
Image domain

\[ x_p = x + D(\nabla x) p + e_x \]  \hspace{1cm} (4)

where uth row, \( D(\nabla x)[u,:) = \nabla x[u,:] J(u) \), for \( u = 1, \ldots, N \)

CS domain

\[ y_p = \Phi x_p \]
\[ = \Phi (x + D(\nabla x) p + e_x) \]
\[ y_p = y + \Phi D(\nabla x) p + e_y \]  \hspace{1cm} (5)

CS problem

\[ \hat{p} = \arg \min_p \|e_y\|_2 \]  \hspace{1cm} (6)

Will this converge?
A set of affine transformed images of the same scene forms a 6D manifold.

Suppose images $x_1$ and $x_2$ are points on this manifold $\mathcal{M}$. If $\Phi$ is an orthoprojector from $\mathbb{R}^N$ to $\mathbb{R}^M$, then the projections of all images in the affine set using $\Phi$ will form another manifold $\Phi \mathcal{M}$.

The vectors $y_1 = \Phi x_1$ and $y_2 = \Phi x_2$ are points on this manifold $\Phi \mathcal{M}$.

For $M = O(d \log (\mu N \epsilon^{-1})/\epsilon^2) < N$, where $\mu$ depends on the properties of the manifold such as volume and curvature, $d$ is the dimension of the manifold, and $0 < \epsilon < 1$, the following holds with high probability (Baraniuk and Wakin 2009):

$$
(1 - \epsilon) \sqrt{\frac{M}{N}} \leq \frac{\|\Phi x_1 - \Phi x_2\|_2}{\|x_1 - x_2\|_2} \leq (1 + \epsilon) \sqrt{\frac{M}{N}}
$$

(7)
In our case, we have

\[(1 - \epsilon) \sqrt{\frac{M}{N}} \|x_p - \hat{x}_p\|_2^2 \leq \|\Phi x_p - \Phi \hat{x}_p\|_2^2 \leq (1 + \epsilon) \sqrt{\frac{M}{N}} \|x_p - \hat{x}_p\|_2^2\]

For \(\epsilon \approx 0\),

\[\|\Phi x_p - \Phi \hat{x}_p\|_2^2 \approx \lambda \|x_p - \hat{x}_p\|_2^2\]

for some constant \(\lambda\). Hence, a monotonic decrease of residual energy is assured with high probability.

Since

\[M = O(d \log (\mu N \epsilon^{-1}) / \epsilon^2)\]

\[\propto (\log \epsilon^{-1}) / \epsilon^2 = O(1 / \epsilon^2),\]

choose \(M\) sufficiently large to ensure that the algorithm converges.
Algorithm 1: \((\hat{p}, e) = \text{estimate\_motion}(y_p, y, \Phi)\)

Initialise \(\hat{p} = [1, 0, 0, 1, 0, 0]^T\). Determine \(x\) from \(y\)

repeat
- Warp \(x\) and \(\nabla x\) by \(\hat{p}\) to get \(x_{\hat{p}}\) and \(\nabla x_{\hat{p}}\) respectively
- Calculate descent matrix, \(S = \Phi \mathcal{D}(\nabla x_{\hat{p}}) \in \mathbb{R}^{M \times 6}\)
- Calculate Hessian, \(H = S^T S \in \mathbb{R}^{6 \times 6}\)
- Calculate \(\hat{y} = \Phi x_{\hat{p}}\)
- Calculate \(\Delta p = H^{-1} S^T (y_p - \hat{y})\)
- Update \(\hat{p} = \hat{p} + \Delta p\)
until \(\hat{p}\) converges

return \(\hat{p}\) and residual energy \(e = \|y_p - \hat{y}\|_2^2\)

In real cases, sequential acquisition affects the CS measurement vector continuously.
Algorithm 2: 
\((\{\hat{p}^{(j)}\}, \{e^{(j)}\}) = \text{recursive}_{-}\text{estimator}(y_p, y, \Phi)\)

Let \(L = \text{length}(y_p), j = 0\) 
\((\hat{p}, e) = \text{estimate}_{-}\text{motion}(y_p, y, \Phi)\)

if \(e > \tau \text{ and } L \geq 2B_{\text{min}}\) then
    recursive_{-}\text{estimator}(y_p[1 : \frac{L}{2}], y, \Phi[1 : \frac{L}{2}, :])
    recursive_{-}\text{estimator}(y_p[\frac{L}{2} + 1 : L], y, \Phi[\frac{L}{2} + 1 : L, :])
else
    \(j = j + 1\)
    return \(\hat{p}^{(j)} = \hat{p} \text{ and } e^{(j)} = e\)
end if
ADAPTIVE BLOCKWISE ESTIMATION
RESULTS

- $x$: 64 \times 64$ image from FERET database
- $\Phi$: scrambled Hadamard matrix
- $y_p$: horizontal translatory motion

- Constant warp
- Linear motion
- Oscillatory motion

- x-axis: measurement number
- y-axis: camera translation in pixels
- red line: actual camera motion
- blue shading: divided blocks
- blue circle: estimated motion for a block
x : 64 × 64 image from FERET database
Φ : scrambled Hadamard matrix
yp : 2D translation camera motions

red line : actual camera motion
blue circle : estimated motion
COMPRESSIVE CLASSIFICATION

- 200 persons in FERET database
- One CS gallery vector per person
- To assign the incoming probe vector to a gallery person

\[ c^* = \arg \min_{1 \leq c \leq C} \{ e_c \}, \text{ where } e_c = \frac{1}{M} \sum_{j=1}^{Q} e_c^{(j)} \]

and \( j = 1, \ldots, Q \) represents the block number in Algorithm 2.

<table>
<thead>
<tr>
<th>Type of motion</th>
<th>Adaptive</th>
<th>No motion estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>((r_z))</td>
<td>93.5</td>
<td>55.0</td>
</tr>
<tr>
<td>((t_x, t_y))</td>
<td>95.5</td>
<td>52.5</td>
</tr>
<tr>
<td>((t_x, t_y, r_z))</td>
<td>95.0</td>
<td>53.0</td>
</tr>
<tr>
<td>Affine</td>
<td>88.0</td>
<td>48.5</td>
</tr>
</tbody>
</table>

Recognition results (in %) on FERET database for continuous camera motion during acquisition

\( M = 200 \)
CONCLUSION

- Proposed an algorithm to estimate relative motion between camera and scene during acquisition of CS measurements
- Discussed how a descent algorithm can be formulated to estimate the motion parameter vector from these measurements
- Demonstrated the utility of our motion estimation framework in the CS domain for the face recognition problem