Compressive Sensing: Opportunities and Perils for Computer Vision

Rama Chellappa
And
Volkan Cevher (Rice)

Joint work with
Aswin C. Sankaranarayanan
Dikpal Reddy
Dr. Ashok Veeraraghavan (MERL)
Prof. Rich Baraniuk (Rice)
Compressed outline

• **Computer Vision**
  – Given images and videos from single or multiple cameras (sensors) tell me what you find in them.

• **Always looking for**
  – image formation theories
    • Reflectance functions, scattering theory, phenomenology
  – Sparse representations
    • feature extraction, sparsity in motion coding, object representations and
  – Optimization techniques
    • Regularization, calculus of variations, $l1$, $l\text{-}infinity$, …
Compressed outline

• Four examples
  – Compressive sensing of reflectance fields (Piers, et al, ACM Trans. on Graphics (Poster by Aswin))
  – SAR image formation (Poster by Vishal).
Compressive Sampling

- When data is sparse/compressible, can directly acquire a condensed representation with no/little information loss through linear dimensionality reduction

\[ y = \Phi x \]

\[ M \times 1 \]
measurements

\[ \Phi \]

\[ M \times N \]

\[ N \times 1 \]
sparse signal

\[ K < M \ll N \]
nonzero entries
Compressive sensing for background subtraction

- Background subtraction (BS)
  - Automatically detecting and tracking moving objects with applications in surveillance, teleconferencing and 3D modeling
Compressive sensing for background subtraction

- **Background subtraction (BS)**
  - Automatically detecting and tracking moving objects with applications in surveillance, teleconferencing and 3D modeling
Compressive sensing for background subtraction

• Background subtraction (BS)
  – Automatically detecting and tracking moving objects with applications in surveillance, teleconferencing and 3D modeling

• State of the art
  – The background and test images are fully sampled using a conventional camera. After foreground estimation the images are discarded or embedded into the background model
Compressive sensing for background subtraction

- **Background subtraction (BS)**
  - Automatically detecting and tracking moving objects with applications in surveillance, teleconferencing and 3D modeling

- **State of the art**
  - The background and test images are fully sampled using a conventional camera. After foreground estimation the images are discarded or embedded into the background model

- **We perform BS on compressed images**
  - Not new but now the images are sensed in a compressed format (e.g. SPC). *Motivation*: especially useful at other wavelengths
Sparsity of BS images and measurement statistics

• **KEY IDEA:** Image foreground sparser than the image. Implies lesser measurements to reconstruct the foreground.

• **SIMPLE APPROACH:**
  • The foreground/silhouette can be constructed from $y_t - y_b$. The appearance of the object can be obtained by reconstructing an auxiliary background image using more measurements.

• **OUR APPROACH:**
  • Background can be naturally adapted to drifts and shifts.
  • Reconstruct the sparse BS silhouette from compressive measurements directly without intermediate background reconstruction.
Adaptation to background changes

Our approach

Simple update rule

Background subtraction results on a sequence with changing illumination
Camera Network: tracking on the CS background subtracted images
Camera Network: tracking on the CS background subtracted images

20% Compression
No performance loss in tracking
Beyond Sparse Models

- Assumption: sparse/compressible signal

background subtracted image
Beyond Sparse Models

- Sparse/compressible signal model captures \textit{simplistic primary structure}

sparse image
Beyond Sparse Models

- Sparse/compressible signal model captures **simplistic primary structure**

- Modern compression/processing algorithms capture **richer secondary coefficient structure**

- wavelets: natural images
- Gabor atoms: chirps/tones
- pixels: background subtracted images
Ex: Background Subtraction

- **Graphical model** encodes dependencies
- Model **clustering of significant pixels** in space domain using Ising MRF
- Ising model approximation performed efficiently using **graph cuts**
- Details: **Model-based Compressive Sensing**

![Images of target, Ising-model recovery, CoSaMP recovery, LP (FPC) recovery]
Gradient domain processing: vision and graphics

- Images/Videos/Meshes/Surfaces
- Estimation of Gradients
- Manipulation of Gradients
- Non-Integrable Gradient Fields
- Reconstruction from Gradients
- Images/Videos/Meshes/Surfaces

Our method: Fit surface in 1-norm sense
Gradient fields and integrability

Image or surface: \( S(x, y) \)

Gradients: \( \nabla S = \left\{ \frac{\partial S}{\partial x}, \frac{\partial S}{\partial y} \right\} \)

Divergence: \( \text{Div}(\nabla S) = \nabla \cdot \nabla S \)

Curl: \( \text{Curl}(\nabla S) = \nabla \times \nabla S \)

Integrability: Conservative vector field

For a scalar field \( S(x, y) \)

\[
\nabla \times \nabla S = 0
\]

\[
\text{Curl}(\nabla S) = S_{yx} - S_{xy} = 0
\]
Non-integrable gradient fields

- Estimation of gradients
  - E.g. Shape from Shading, Photometric Stereo
  - Noise and outliers in estimation

- Manipulation of integrable gradients
  - Synthesize new gradient field

Input Images  \[\rightarrow\]  Surface Normals/Gradients  \[\rightarrow\]  Not-integrable

Image  \[\rightarrow\]  \(S_x\)  \[\rightarrow\]  \(S_y\)  \[\rightarrow\]  Gradient Manipulations  \[\rightarrow\]  New Gradients  \[\rightarrow\]  Not-integrable
Discrete domain

- $S(x, y)$ surface of size $H \times W$ vectorized to give $S$

- Gradients $g$ obtained from estimation/manipulation
  $$g = Ds + e \text{ of size } (H - 1)W + (W - 1)H$$

- Non-integrable gradient field: Non-zero curl values $d$
  $$Cg = CDs + Ce$$
  $$d = Ce$$

Curl matrix $C$ of size $(H - 1)(W - 1) \times (H - 1)W + (W - 1)H$

$$C = \begin{bmatrix}
-1 & 1 & 1 & -1 \\
-1 & 1 & 1 & -1 \\
-1 & 1 & 1 & -1 \\
-1 & 1 & 1 & -1 \\
\end{bmatrix}$$
Poisson solver

Least squares: $\hat{e} = C^\dagger d$

- Works well and optimal when error in gradients is Gaussian noise
- When corrupted by outliers, Poisson solver (Simchony, et al) and least squares based methods (Frankot et al., Kovesi) perform poorly
$l_1$- minimization

$$\hat{e} = \arg \min_{e} \|e\|_1 \text{ s.t. } d = Ce$$

- Known to correct outliers. Performs well in noise too.
- Confines error locally
CS and reconstruction from gradient fields

- Recover s from  \( g = Ds + e \)
- D is the coding matrix
- g is the gradient field
- E is the unknown vector of errors
- C does not obey RIP
- C obeys RIP-1

Properties

- How many outliers can $l_1$ minimization fully correct?
- How should they be distributed?
- If large number of outliers then what outliers does $l_1$ find and correct?

Answers can be found in CS literature by answering

- What properties does the curl matrix $C$ have?
- How well does $C$ satisfy RIP? How well does it satisfy RIP-1?
- Can we use structure of $C$ for faster error correction?
Other examples

• Sparse representations for face recognition – Yang, et al, PAMI 12/08.
• Compressive Sensing of Reflectance Fields
  – Poster by Aswin at this mtg
• SAR image formation
  – Patel, Easley, Healy and Chellappa
  – Poster by Patel
• Many others presented at this meeting
Problems under study

• Image-based modeling (L. Quan, HKUST, Marc Pollefeys use traditional sfm and multi-view geometry)
  – Often 20,000 photographs are need for 3D modeling and visualization (Urbanscape)

• CS approaches for multi-view tracking, object and activity recognition.
  – Joint work with Baraniuk
Perils

• May open the flood gates for papers with compressive sensing for .. titles
• Vision researchers are open to the beauty of mathematics and statistics.
  – Regularization, simulated annealing, anistropic diffusion, differential geometry, MCMC, manifolds, …
• Often driven by techniques than the problems at hand.
• Focus on approaches well ground on image formation theories, sparse representations and \( l1 \) optimizations.
• The need is there!