Kalman Filtered Compressed Sensing for Real-time Dynamic MRI

Namrata Vaswani
Dept. of Electrical and Computer Engineering
Iowa State University
http://www.ece.iastate.edu/~namrata
Why Real-time Dynamic MRI?

• MRI has few harmful effects on patient and operator & provides very good tissue contrast [Martin et al, MedicaMundi, 2002]

• Excellent candidate to be used in image-guided surgery or other interventional radiology applications
  – This requires real-time imaging capability
MRI Signal Processing

- MRI acquires an incomplete set of samples of the 2D Fourier transform (k-space) of the image

- Acquisition is sequential: scan time proportional to the number of Fourier coefficients needed

- Goal of signal processing research:
  - Achieve the best possible reconstruction quality using as few Fourier coefficients as possible
  - Real-time apps: also keep reconstruction time small
Current MRI Technology

- Most existing algorithms
  - Scan time (no. of measurements required) large

- Compressed Sensing for MRI [Lustig et al, MRM, Dec’07]
  - Uses fewer measurements: reduced scan time

- Batch CS for Dynamic MRI [Gamper et al, MRM, Jan’08]
  - Uses much fewer measurements per frame than CS: real-time acquisition possible
  - Reconstruction takes very long
  - Offline method: get all measurements first

KF-CS for Real-time dynamic MRI
Our Work

• Our work:
  – Kalman filtered CS
  – Its non-Bayesian version (Least Squares CS)

• Replaces the offline, high complexity reconstruction algorithm for dynamic MRI with a real-time one
  – Reconstruct the current image using only the current set of measurements and the optimal prediction based on all past measurements
Outline

• Background, Existing Work

• Some reconstruction results

• KF-CS and non-Bayesian KF-CS (LS-CS)

• Discussion of results

• Conclusions, Other Related Work
Compressed Sensing [Candes et al, Donoho]

• It is possible to exactly reconstruct an $S$-sparse signal of size $M$ from a much smaller number, $K$, of its measurements, using simple linear/convex programming techniques,
  – if the measurements are “incoherent” enough w.r.t. the sparsity basis of the signal, and
  – if $K \sim (S \log M)$

• Reconstruction complexity is $O(M^3)$
• Reconstruct with small error if approximately sparse signal or noisy measurements
Sparsity

- An $M$-length vector, $x$, is \textbf{S-sparse} if only $S \ll M$ of its components are nonzero
  - Approximately $S$-sparse: $S$ components significantly nonzero

- An $M$-length signal, $z$, is \textbf{sparse in the $\Phi$ domain} ($\Phi_{M \times M}$: orthonormal matrix) if $x = \Phi^*z$ is sparse

- Human organ images (or any piecewise smooth images) are approximately sparse in the wavelet transform or the spatial gradient domain
Examples

Piecewise constant image (spatial gradient is sparse)

Piecewise smooth image (approx. sparse in wavelet domain)

Approx. piecewise constant (spatial gradient approx. sparse)
CS for MRI [Lustig et al, MRM, 2007]

- Signal: cross-sectional image of heart, brain, etc
  - Piecewise smooth: wavelet transform approx. sparse

- Measurements: 2D Fourier coefficients (k-space)
  - Satisfy “incoherency” w.r.t. wavelet basis [Candes et al]

- $M$: number of pixels in the image

- $S$: no. of edges / no. of large wavelet coefficients
CS for MRI (contd.)

- CS theory: possible to accurately reconstruct using $K \sim S \log M$ measurements
  - e.g., if $M = 128 \times 128 = 16384$, $S \sim 2-3\%$ of $M$, need $K \sim 28-42\%$ of $M$
  - In practice: able to use even fewer measurements by using intelligent k-space sampling patterns

- **CS reconstruction time is small enough but scan time still not small enough for real-time**
Batch-CS for Dynamic MRI [Gamper et al]

- Jointly reconstruct the entire spatio-temporal signal (image sequence), which is sparse in wx-wy-f, from the entire MRI sequence (kx-ky-t data)
  - Joint signal: much more sparse than each individual one
  - Needs much fewer measurements per frame than static CS: scan time much smaller
  - But reconstruction is offline and very slow: needs “tens of hours” using linear/convex programming (time complexity is $O(t^3M^3)$)
Our Goal

• Make the dynamic MR reconstruction real-time while still using as few, or slightly more, measurements per frame, K, as batch-CS
  – (or equivalently: get the least recon. error for a given small enough K)

• Develop algorithms for causal (online) reconstruction of a time sequence of sparse signals with smallest possible error for a given K
Comparing KF-CS with CS (simulated data)

- If want to keep reconstruction time equal to that of static CS (cannot do batch-CS), how does recon. error of static CS compare with our method (KF-CS), for a given K

M = 256
K = 72
$S_t$ varied from 8 to 26, but number of new additions per unit time is only 2 or less
Comparing KF-CS with CS (simulated cardiac MRI)

- **KF-CS reconstruction is sharper**, CS more smoothed
- True image itself is actually standard MRI reconstructed
Comparing LS-CS with CS

Original sequence

LS-CS with updating detected set reconstructed sequence

KF-CS for Real-time dynamic MRI
LS-CS for Brain

Original sequence

LS-CS-LS reconstructed sequence

KF-CS for Real-time dynamic MRI
Notation

• $z_t$: signal at time $t$ (approx. sparse in basis $\Phi$)
• $x_t$: sparse coefficients’ vector ($x_t := \Phi^* z_t$)
• $N_{Z_t}$: set of indices of the (significantly) nonzero components of $x_t$
  $\Delta_t$: new additions to the $N_{Z_t}$ set at $t$ ($\Delta_t := N_{Z_t} \setminus N_{Z_{t-1}}$)
• $y_t$: measurement ($y_t := H z_t + w_t = A x_t + w_t$)
  – $H$: selected rows from Fourier transform matrix
  – $A := H \Phi$: resulting measurement matrix for CS
  – $w_t$: measurement noise
Problem Formulation

• At time $t$, get the best possible estimate of $x_t$ (and consequently of the signal $z_t = \Phi x_t$) from the measurements $y_1, y_2, \ldots, y_t$

• Use the fact that the sparsity pattern changes very slowly over time (due to strong temporal correlations)
  – the set $\Delta_t := NZ_t \setminus NZ_{t-1}$ much smaller than $NZ_\tau$
Slowly-varying Sparsity Pattern

- Cardiac sequence: size of NZ set ~168 - 174 (less than 6% of image size), maximum size of the changed set is less than 4
- Brain volume sequence: size of NZ set ~ 448 - 463 (less than 2% of image size), maximum size of changed set is less than 10
Motivation

• Kalman Filtered CS:
  – If the nonzero set, $NZ = NZ_t$, was known for each $t$, the measurement model would be observable
  – One could simply use a size($NZ_t$) dim. Kalman filter at $t$ to get the best estimate of $x_t$ (or of $z_t$) from $y_1, y_2, \ldots y_t$

• Least Squares CS (non-Bayesian KF-CS):
  – If the nonzero set, $NZ = NZ_t$, was known for each $t$, one could simply compute the Least Squares estimate of $(x_t)_{NZ}$ from $y_t$, i.e.
    $$(\hat{x}_t)_{NZ} = (A^*_{NZ}A_{NZ})^{-1}A^*_{NZ}y_t, \ (\hat{x}_t)_{NZ^c} = 0$$

• $NZ_t$ unknown, time-varying: use CS to estimate it
• Two options:
  – Option 1: CS on observation
  – Option 2: CS on LS residual (or Kalman filtering error)

• It is possible to show that **CS on the LS residual (or Kalman filtering error)** to estimate only the new additions to the NZ set at t is much more accurate than performing CS on the observation [Vaswani, Allerton’08, submitted]
  – Reason: the LS residual is much more sparse than the observation
Computing the nonzero set $NZ_t$

CS on the observation:
\[ y_t := A_{NZ_t}(x_t)_{NZ_t} + w_t \]
\[ \hat{\beta} = \arg \min \|y_t - A\beta\|^2 + \gamma\|\beta\|_1 \]
\[ NZ_t = \{ i : |\hat{\beta}_i|^2 > \alpha \} \]

CS on the LS residual or Kalman filtering error:
\[ \tilde{y}_{t,f} := A_{\Delta_t}(x_t)_{\Delta_t} + A(x_t - \hat{x}_t, NZ_{t-1}) + w_t \]
\[ \tilde{\beta} = \arg \min \|\tilde{y}_{t,f} - A\beta\|^2 + \gamma\|\beta\|_1 \]
\[ \Delta_t = \{ i \in NZ_{t-1}^c : |\tilde{\beta}_i|^2 > \alpha \} \quad NZ_t = NZ_{t-1} \cup \Delta_t \]

It is possible to show that the error in the first estimate is larger than that in the second one [Vaswani, Allerton’08, sub.]
Least Squares CS (LS-CS) [Vaswani, ICIP’08]

\[ \tilde{y}_{t,f} = y_t - A_{\text{NZ}}(A_{\text{NZ}}^*A_{\text{NZ}})^{-1}A_{\text{NZ}}^*y_t \]

(\text{MR measurements at } t)

\[ y_t \xrightarrow{\Delta_t} \tilde{y}_{t,f} \]

Delay \( t \xrightarrow{\text{t+1}} \)

\[ \beta = \arg\min ||\tilde{y}_{t,f} - A\beta||^2 + \gamma||\beta||_1 \]

\[ \Delta_t = \{ i \in N\text{Z}^c : |\tilde{\beta}_i|^2 > \alpha \} \]

\[ (\hat{x}_t)_{\text{NZ}} = (A_{\text{NZ}}^*A_{\text{NZ}})^{-1}A_{\text{NZ}}^*y_t, \quad (\hat{x}_t)_{\text{NZ}^c} = 0 \]

(\text{reconstructed image at } t)

\[ \hat{z}_t = \Phi \hat{x}_t \]

KF-CS for Real-time dynamic MRI
Kalman Filtered CS (KF-CS) [Vaswani, ICIP'08]

\[
\tilde{y}_{t,f} = y_t - A(\tilde{x}_{t-1} + K_{t,mp}(NZ)(y_t - A\tilde{x}_{t-1}))
\]

Kalman filtering error using \(NZ=NZ_{t-1}\)

\[
\tilde{y}_{t,f} = y_t - A(\tilde{x}_{t-1} + K_{t,mp}(NZ)(y_t - A\tilde{x}_{t-1}))
\]

CS on KF error + thresholding

\[
\tilde{x}_{t-1}, y_t \rightarrow \Delta_t
\]

Delay \(t \leftarrow t+1\)

\[
\tilde{x}_{t-1}, y_t \rightarrow \Delta_t
\]

KF estimate using \(NZ=NZ_t = NZ_{t-1} \cup \Delta_t\)

\[
\tilde{x}_{t} = \tilde{x}_{t-1} + K_{t}(NZ)(y_t - A\tilde{x}_{t-1})
\]

\[
\tilde{x}_{t} \rightarrow NZ_t
\]

\[
\tilde{x}_{t} \rightarrow NZ_t
\]

\[
\tilde{x}_{t} \rightarrow \tilde{z}_t = \Phi \tilde{x}_t
\]

KF-CS for Real-time dynamic MRI

(MR measurements at time \(t\))
KF-CS versus LS-CS

• Reconstruction performance
  – If a reliable prior model is available and if additions to the nonzero set occur slowly enough: KF-CS should outperform LS-CS

• Computational complexity for 2D problems
  – It is possible to implement a 2D version of LS-CS which is much more efficient than 1D LS-CS (or KF-CS)
Comparison with CS (simulated data)

- If want to keep reconstruction time equal to that of CS for one image (cannot do batch-CS), how does recon. error of CS compare with KF-CS (for a given K)

M = 256
K = 72
$S_t$ varied from 8 to 26, but number of new additions per unit time is only 2 or less
Simulated MRI Experiments

- Used a cardiac image sequence and a brain functional image sequence

- Computed their 2D Fourier transform, selected a random set of Fourier coefficients and added i.i.d. Gaussian noise to get the simulated MR measurements

- Computed the error in the KF-CS and LS-CS estimate of the image and compared it with that of the CS estimate

- Very preliminary experiments
  - Not using the most sparsifying wavelet transform
  - Not using good k-space sampling schemes
1D KF-CS: MSE comparison

- **KF-CS (blue plot)** without coefficients’ removal step
- $M = 32^2 = 1024$, $K = 513$, SNR = 8.5dB
- **GA-KF**: Genie-aided KF which assumes known sparsity pattern
Reconstructed Images

- KF-CS recon. is sharper, CS more smoothed
- True image itself is actually standard MRI reconstructed
1D KF-CS: another MSE plot

- Used $K = M/3$ measurements, $M = 64 \times 64$, SNR = 40dB
- Used 4-level wavelet decomposition
- KF-CS: blue plot
1D KF-CS: Reconstructed image

Original sequence

KF-CS reconstructed sequence
2D LS-CS

- $M = 128 \times 128 = 16384$, $K = 0.4 \, M$, SNR = 30dB
- LS-CS: red plot
2D LS-CS: Reconstructed Images

Original sequence

LS-CS with updating detected set reconstructed sequence

KF-CS for Real-time dynamic MRI
2D LS-CS for Brain

Original sequence

LS-CS-LS reconstructed sequence
Conclusions

- Older methods (e.g. zero filling)
  - Scan time per frame: large, not useful for real-time
  - Reconstruction: can be real-time

- Single-frame (static) CS
  - Scan time per frame: lesser, but still not real-time
  - Reconstruction: can be real-time

- Batch-CS (joint CS for entire time sequence)
  - Scan time per frame: MUCH smaller than older methods or CS
  - Reconstruction: NOT real-time (very slow)

- Proposed methods: LS-CS and KF-CS
  - Scan time per frame: should be low enough for real-time
  - Reconstruction: can be real-time (as fast or faster than CS)
Other Related Work

• k-t FOCUSSS: uses MPEG style prediction to do online CS in dynamic MRI (very recent, still unpublished)

• Techniques that use prior information
  – k-t Blast, k-t Sense: spatio-temporal prior obtained from a low resolution scan

• Kalman filtering has been used in past work also but in very different ways

• Classical algorithms: Zero filling with DC, FBP
Real-time Dynamic Tomography

- CT, other tomographies exactly the same from a signal processing perspective
  - Acquire the Radon transform of the image along different directions
  - Fourier slice theorem applied to the Radon transform gives a sub-Nyquist set of 2D-Fourier coefficients
  - Acquisition is sequential: acquire one row at a time
Ongoing/Future Work

- Coefficient deletion schemes
  - More difficult for MRI problems because the signal is not exactly sparse (just compressible)

- When does prior information help & how much

- Use of faster greedy methods: OMP, CoSaMP

- Real MRI data experiments, clinical relevance

- Use of better wavelet transforms, better k-space sampling ideas from existing work
Acknowledgements

- Research partially supported by National Science Foundation (NSF) grant: ECCS-0725849

- Graduate students
  - Chenlu Qiu (1D KF-CS for MRI)
  - Wei Lu (2D LS-CS for MRI)

- Data obtained from
  - Prof. Vince Magnotta (Univ. of Iowa)
  - Prof. Allen Tannenbaum (Georgia Tech)
Sequential Segmentation

• Sequentially segment *deforming objects or Regions of Interest (ROIs)* from video or other image sequences

• Optimally use prior knowledge about shape change dynamics to segment noisy/ low contrast imagery
  – Prior dynamic model: guess/ learn from “training” data

• How to do this in real-time (using only current and past images for segmenting current frame)
Examples: Deforming contours

- Actual deformations: biological images
  - Human tracking: surveillance, sports videos, ...  
  - Animals such as a fish  
  - Medical sequences: ROIs in brain or heart
- Changing region of partial occlusions
  - Automatic vehicle navigation  
  - Robot navigation
- Frequently changing camera viewpoint
  - Tracking using a UAV
Low contrast + deforming contours
(large deformation per frame)

Low contrast + Frequent viewpoint changes
(small deformation per frame)
Separate clutter (multiple fishes) + deformation

Overlapping clutter (light grey object) + deformation
Partial occlusion of car by street light:
3 contour modes, 2 are deforming contours
Brain MRI (Tumor, Ventricle)

Multiple nearby modes due to low contrast

(b) Attempt to track the right ventricle (black region in the center) using Algorithm 2. Notice the low contrast imagery.