Residual Reconstruction for Block-Based Compressed Sensing of Video

Sungkwang Mun and James E. Fowler

Department of Electrical & Computer Engineering
Geosystems Research Institute
Mississippi State University

Mar 30 2011

MISSISSIPPI STATE
UNIVERSITY™

Outline

1. **Motivation**
2. **Background**
   - CS Overview
   - BCS-SPL
3. **Block Based CS Reconstruction of Video**
   - Extension of Still-Image
   - MC-BCS-SPL
   - Results
4. **Conclusion**

Traditional Video Compression

BCS of Video

S. Mun

Motivation

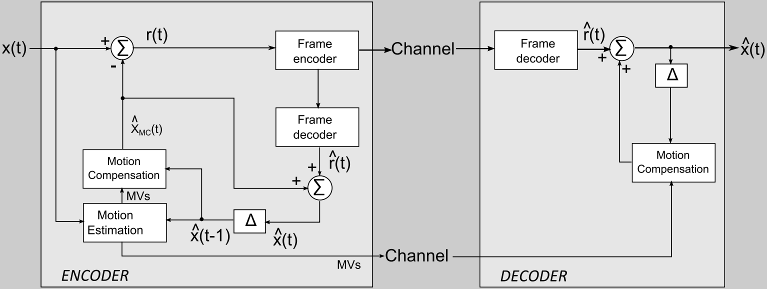
Background

MC-BCS-SPL

Conclusion

Traditional Video Compression

- Smart encoder and fast, simple decoder



- Well suited for once encode, decode many times
- Not the best solution for resource-limited encoding
- Recent theory gives possibility that we can hand over the burden to the other side

Compressed Sensing Video Compression

Compressed Sensing (CS)

- Sampling at sub-Nyquist rates via linear projection onto a measurement basis of lower dimension
- Exact (or approximate) reconstruction when signal is sparse under certain condition
- Simultaneous sampling and dimension reduction

CS Video Conception

- Using still-image-based CS sampling, i.e, single pixel camera
- Joint reconstruction exploiting temporal redundancy
- Fast, simple encoder and smart decoder

Outline

CS Overview

Goal

Recover $\mathbf{x} \in \Re^N$ from

$$\mathbf{y} = \mathbf{\Phi x} \in \Re^M$$

- **$\mathbf{\Phi}$: $M \times N$ random measurement matrix, $M \ll N$**
- **$S = \frac{M}{N}$: Sub-sampling ratio or subrate**

Fundamental Tenet of CS

If x is sufficiently sparse, recovery is exact from

$$M \geq O(K \cdot logN)$$

measurements by solving tractable program.

- **K: number of nonzero coefficients in some transform Ψ, $\check{\mathbf{x}} = \mathbf{\Psi x}$**

Outline

Block Compressed Sensing

Block Compressed Sensing (BCS)

Image partitioned into small blocks ($B \times B$)

$$\mathbf{y}_j = \mathbf{\Phi}_B \mathbf{x}_j$$

$\mathbf{\Phi}_B$: $\lfloor \frac{M}{N} B^2 \rfloor \times B^2$, $\mathbf{x}_j$: block j of image

Smooth Projected Landweber (SPL)

$$\min_{\mathbf{\Psi x}} \|\mathbf{\Psi x}\|_{\ell_1} \ \ \textbf{subject to} \ \ \|\mathbf{y} - \mathbf{\Phi x}\|_{\ell_2} \leq \epsilon$$

- Wiener filter to smooth blocky artifacts (L. Gan, 07)
- Directional Transform and statistically determined threshold (Mun & Fowler, 09)
- Practical recovery (simple, fast)

Outline

Extension of Still-image CS to Video

Block Based CS Video Reconstruction

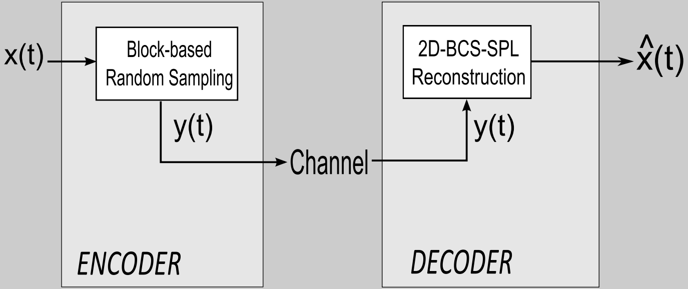
- 2D image recovery enables frame-by-frame reconstruction
- Unlike images, video has temporal redundancy as well as spatial redundancy
- Goal: Finding a good CS reconstruction method for video in terms of quality and speed

Experimental Assumption

- Block based sampling frame-by-frame by using same 2D sampler
- Measurements of multiple frames are available (decoder)
- Key frames and non key frames

Frame-by-frame Reconstruction

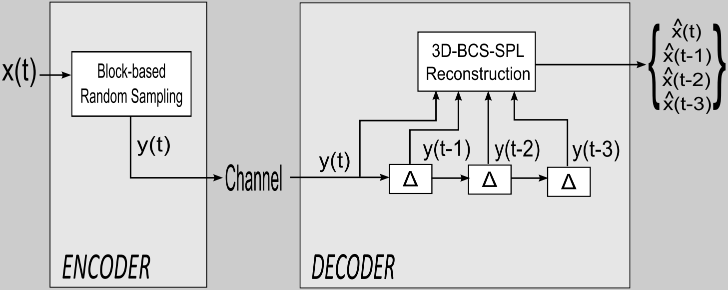
2D BCS-SPL reconstruction



- Straightforward, but ignores the correlation in consecutive frames

Joint Reconstruction

3D BCS-SPL reconstruction

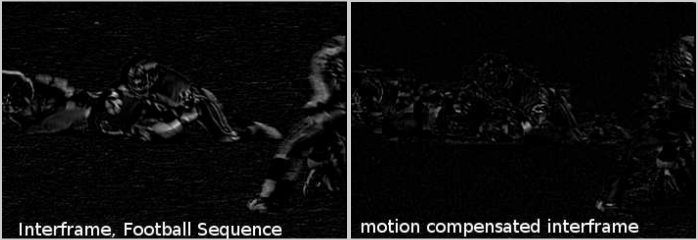


- Wakin et al. 2006
- Group multiple frames together and reconstructing using 3D block-based transform
- Only pursuing the joint sparsity that occurs in the 3D transform

Outline

Motion Compensated Reconstruction

Residual Reconstruction



- Using motion estimation/compensation to capture objects motion in frames
- Similar to *motion compensated interframe coding* in standard video compression
- Residual is *random projected* motion compensated interframe

Residual Reconstruction

$$\begin{aligned}\mathbf{y}_r &= \mathbf{y} - \mathbf{\Phi}\hat{\mathbf{x}}_{mc} \\ &= \mathbf{\Phi}(\mathbf{x} - \hat{\mathbf{x}}_{mc})(\because \mathbf{y} = \mathbf{\Phi}\mathbf{x}) \\ &= \mathbf{\Phi}\mathbf{x}_r\end{aligned}$$

After solving

$$\min_{\mathbf{\Psi}\mathbf{x}_r} \|\mathbf{\Psi}\mathbf{x}_r\|_{\ell_1} \ \ \textbf{subject to} \ \ \|\mathbf{y}_r - \mathbf{\Phi}\mathbf{x}_r\|_{\ell_2} \leq \epsilon$$

$$\therefore \hat{\mathbf{x}} = \hat{\mathbf{x}}_{mc} + \hat{\mathbf{x}}_r$$

- **Predicted interframe $\mathbf{x}_r$ is more compressible than the target $\mathbf{x}$**
- **Therefore, the solution for residual $\mathbf{x}_r$ can be obtained more accurately**

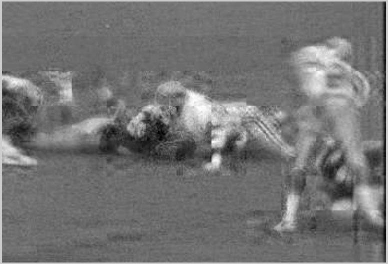
Single-Frame Results(Perfect MVs)

BCS of Video

S. Mun

- Motivation
- Background
- MC-BCS-SPL
  - Extension of Still-Image
  - **MC-BCS-SPL**
  - Results
- Conclusion

Football frame (352×240), quarter-pixel ME, $S = 0.3$

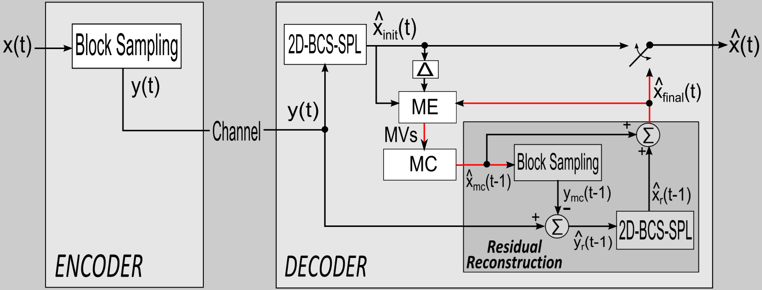


2D-BCS-SPL
24.02 dB
8s

MC-BCS-SPL
27.64 dB
20s

Motion Compensated BCS-SPL

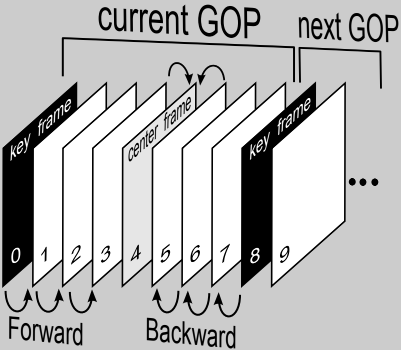
MC-BCS-SPL Diagram



- For MVs, 2D SPL reconstruction is used as initial step
- Iteration on ME/MC helps to obtain more accurate MVs, resulting in better final reconstruction

Multiple-Frame Processing

Forward/Backward Processing



- Keyframes could have better quality than non-keyframes by sampling more
- Forward processing for first half of the GOP; backward processing for last half of the GOP

Outline

Multiple-Frame MC-BCS-SPL

Experiment Setup

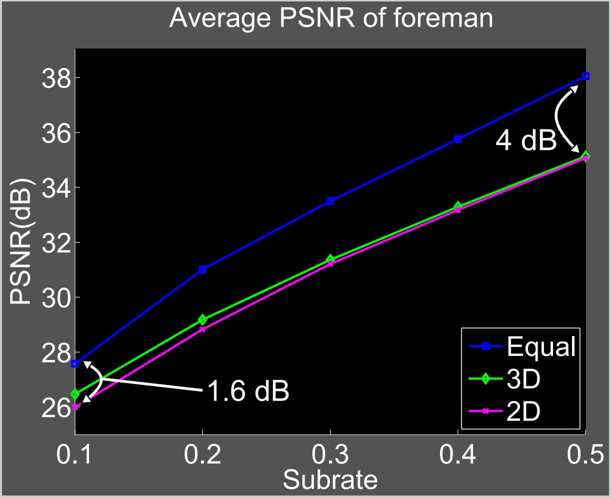
- Using DCT as sparsity basis for simplicity
- Subrate-distortion performance is observed
- Case 1: All frames are equally subsampled
- Case 2: Keyframes are sampled with higher subrate
  - Performance is measured over non-key frames only

Compared algorithms

- Frame-by-frame reconstruction: 2D-BCS-SPL
- 3D joint reconstruction: 3D-BCS-SPL

Multi-Frame Average Results

Case 1: $S_K = S_{NK}$, Foreman, 296 frames



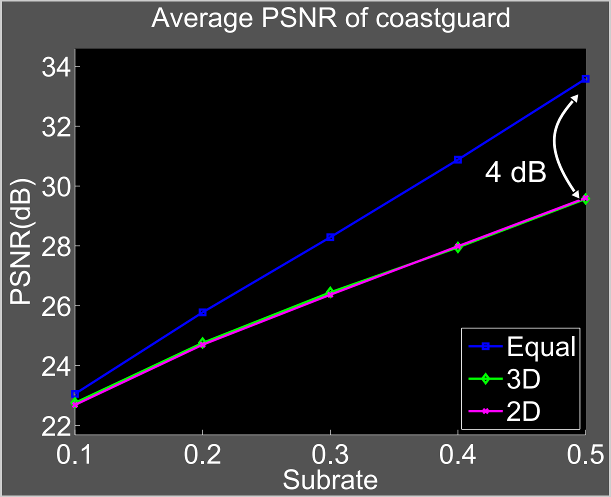
Multi-Frame Average Results

BCS of Video

S. Mun

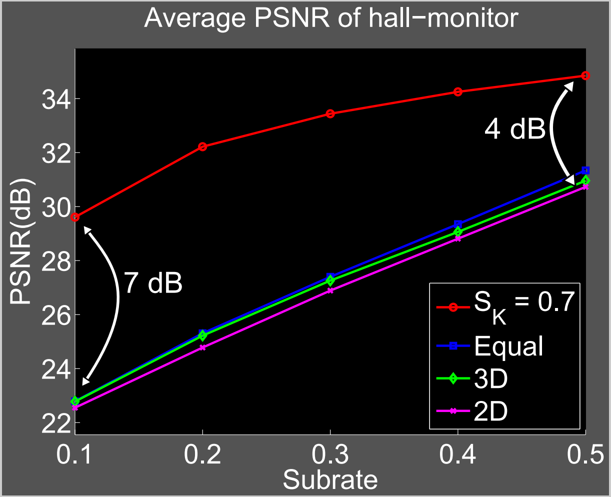
- Motivation
- Background
- MC-BCS-SPL
  - Extension of Still-Image
  - MC-BCS-SPL
  - **Results**
- Conclusion

Case 1: $S_K = S_{NK}$, Coastguard, 296 frames



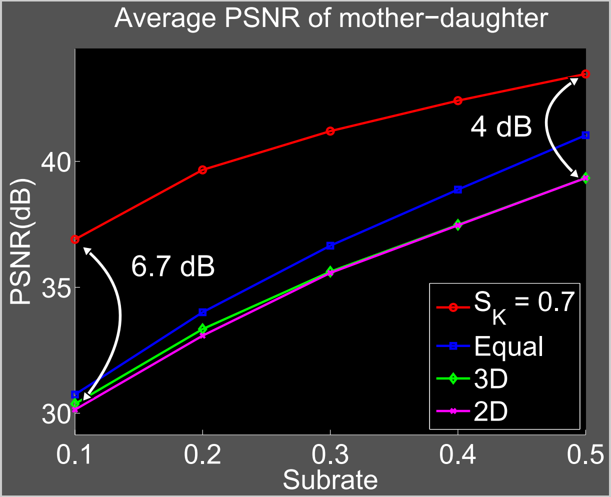
Multi-Frame Average Results

Case 2: S_K = 0.7, Hall monitor, 88 frames



Multi-Frame Average Results

Case 2: S_K = 0.7, Mother and Daughter, 296 frames



Comparison with Other Video Technique

Do et al, 2009,“Distributed Compressed Video Sensing”

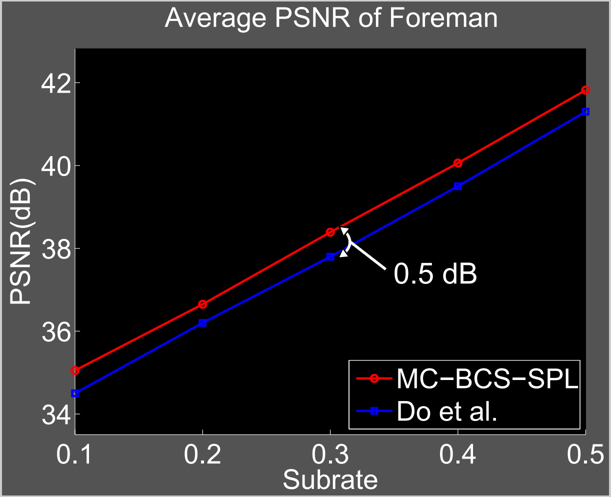
- Proposed hybrid CS reconstruction of video using H.264 type intra coding together
- Find best linear combination of multiple best matching blocks from the two key frames nearby
- Similar residual reconstruction, but not iterative
- Full frame measurement

Setup

- GOP length: 4 frames
- Perfect Key frame, i.e. $S_K = 1.0$

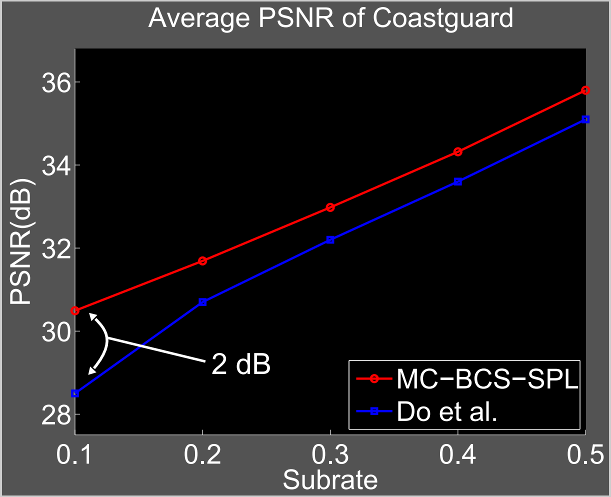
Comparison with Other Video Technique

Do et al, 2009, Foreman, 88 frames, S_K = 1.0



Comparison with Other Video Technique

Do et al, 2009, Coastguard, 100 frames, S_K = 1.0



Conclusion

Motion Compensated BCS-SPL

- Block based sensing ME/MC based reconstruction
- ME/MC in residual reconstruction helps pursue the sparsity along the sequences
- Enhancing MVs through iteration

Matlab Source Code

- MC-BCS-SPL Version 1.0
  `http://www.ece.msstate.edu/~fowler/BCSSPL`