Object Tracking and Image Reconstruction from Compressive Sensing Cameras
Henry Braun, Pavan Turaga, Andreas Spanias, Cihan Tepedelenloglu

Summary
Compressive imaging is a novel sensing paradigm which promises lower data rates and lower cost sensing hardware, particularly in non-visible wavelengths where per-pixel costs are high. However, current methods for reconstructing compressively sensed videos do not fully exploit prior knowledge of motion and frame-to-frame correlations. We have extended typical l1 reconstruction to incorporate optical flow information when simultaneously reconstructing multiple frames of video. Although estimating generalized dense optical flow from compressively sensed data is a difficult and unsolved problem, we demonstrate the usefulness of this approach for two special cases: known optical flow and unknown but constant optical flow. In both cases, a clear perceptual improvement in reconstructed frames is evident, as is an increase in Peak Signal-to-Noise Ratio (PSNR).

Compressive Sensing Camera Architecture

• Rather than sensing a grid of m small pixels, take a series of linear measurements y over the whole image using n different pseudorandom masks.
  \[ n << m \], hence the signal is compressively sensed.
• The original image is known to be sparse (i.e. compressible) using some basis B, for example a wavelet basis:

\[ \mathbf{y} = \mathbf{Mx} = \mathbf{MBx} \]

Video Reconstruction Incorporating Optical Flow

• Modern video codecs exploit correlations between neighboring frames when encoding video from conventional cameras – how can we do the same when reconstructing compressively sensed video sequences?
• If we know the optical flow vectors u and v between two frames, we can reconstruct by solving the following optimization problem:

\[ \hat{\mathbf{u}}, \hat{\mathbf{v}} = \arg\min_{\mathbf{u}, \mathbf{v}} \| \mathbf{y} - \mathbf{M} \mathbf{F}_{\mathbf{u} \mathbf{v}} \mathbf{B} \mathbf{M}^T \mathbf{B} \|_2 + \tau \| \mathbf{u} \|_1 \]

• If optical flow is known to be constant across the image, we can perform single-frame reconstruction to estimate the shift between frames, and use this shift as before to perform 2-frame reconstruction.

Reconstruction with known motion
• The reconstruction algorithm above was tested using the Middlebury optical flow dataset, which includes known optical flow data.
• Clear improvements in both perceptual quality and PSNR across a wide range of sensing rates.

Reconstruction with constant shift
• The algorithm was tested on a single cropped frame from the PETS2000 dataset, with a shift introduced to simulate a moving camera.

Ground Truth
Single-frame reconstruction
PSNR = 24.4 dB

2-frame reconstruction with optical flow
PSNR = 27.21 dB

Results

Conclusion
This work is an important first step toward fully incorporating temporal information in compressively sensed video. Relative to single-frame reconstruction, the algorithm described here shows clear benefits in the quality of reconstructed compressively sensed images. This knowledge may speed the commercialization of cost-effective imaging systems in a variety of wavelengths beyond the visible range.

Acknowledgements
This work is funded by Raytheon Missile Systems and supported in part by the SenSIP site of the NSF NCSS I/UCRC. Additional funding is provided by the ARCS foundation.