Adaptive Compressed Sensing for Depthmap Compression Using Graph-based Transform

Ming Hsieh Department of Electrical Engineering

USC Viterbi School of Engineering

Sungwon Lee and Antonio Ortega

Signal and Image Processing Institute, University of Southern California
Los Angeles, CA, USA

Depthmap Compression

- Depthmap characteristics
  - ▷ Gray-scaled video
  - ▷ Piecewise constant signal
  - ▷ Sharp edges where depth changes
- Precise edge information crucial for view synthesis
  - ▷ The error in edge information
    - → geometry error in interpolated view
    - → significant degradation of the quality [1, 2]
  - ▷ Traditional DCT (H.264/AVC) not efficient for arbitrary edges
    - → Requires edge-adaptive transform; avoid filtering across the edges

Related Work

- Graph-based transform [3]
  - ▷ Represent depth map as a graph
  - ▷ Present edge-adaptive GBT construction
  - ▷ Propose a greedy algorithm to find an optimal transform kernel of GBT
  - ▷ Outperforms H.264/AVC
    - - comparison to this approach will be shown later
- CS-based depthmap compression [4, 5]
  - ▷ Use fixed DCT as sparsifying basis:
    - - inefficient for arbitrary edges
    - - Not adaptive to edges in blocks
  - ▷ Performance was not evaluated with interpolated view quality

Proposed Algorithm

- Propose a design of optimal GBT for CS
  - ▷ Residual depthmap block after intra/inter prediction
  - ▷ Find optimal edgemap (adjacency matrix) with cost function
  - ▷ Generate GBT from edgemap
- GBT construction from edgemap (adjacency matrix)
  - ▷ Given $\mathbf{G(V,E)}$, nodes (pixels) and links (between pixels)
  - ▷ Construct adjacency matrix, $\mathbf{A}$:
    $\mathbf{A(i,j) = 1}$ if no edge between pixels and 0 otherwise
  - ▷ Construct Laplacian matrix, $\mathbf{L}$:

$$\mathbf{L = D - A} = \begin{cases} -1 & \text{if } (i,j) \in E \\ d_i & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

  - ▷ Sparsifying basis $\mathbf{\Psi = [e_1, \ldots, e_N]}$, $\mathbf{e_i}$ is $\mathbf{i^{th}}$ eigenvector of $\mathbf{L}$
- Propose greedy algorithm to find optimal edgemap
  - ▷ Assume 4-neighbor connectivity
    - → 24 possible edges in $\mathbf{4 \times 4}$ block
  - ▷ Define cost:

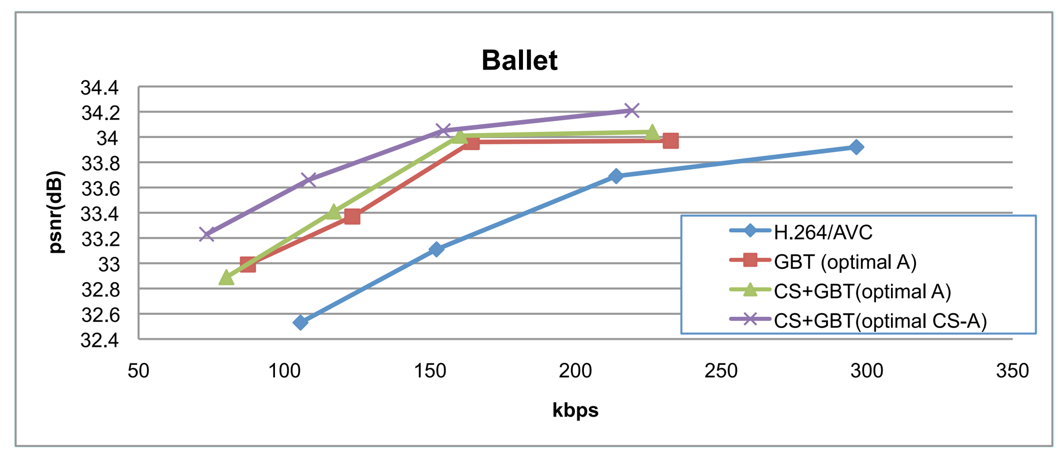
$$\begin{aligned} \mathrm{Cost} &= \mathrm{Cost}_{\mathrm{measurement\_rate}} + \mathbf{k}\mathrm{Cost}_{\mathrm{edge\_rate}} \\ &= \log_2\left(\frac{(\sum_i \mathbf{a_{ij}(x_i - x_j)^2})}{\mathbf{2Q^2}} \mu^2_{\mathbf{avg}}\right) + \mathbf{km} \end{aligned}$$

  - ▷ Initial state: put edges between pixels with different values
    - ▸ Find an edge between smallest pixel difference
    - ▸ Remove the edge if updated cost decreases
    - ▸ Keep doing it until scanning all the edges

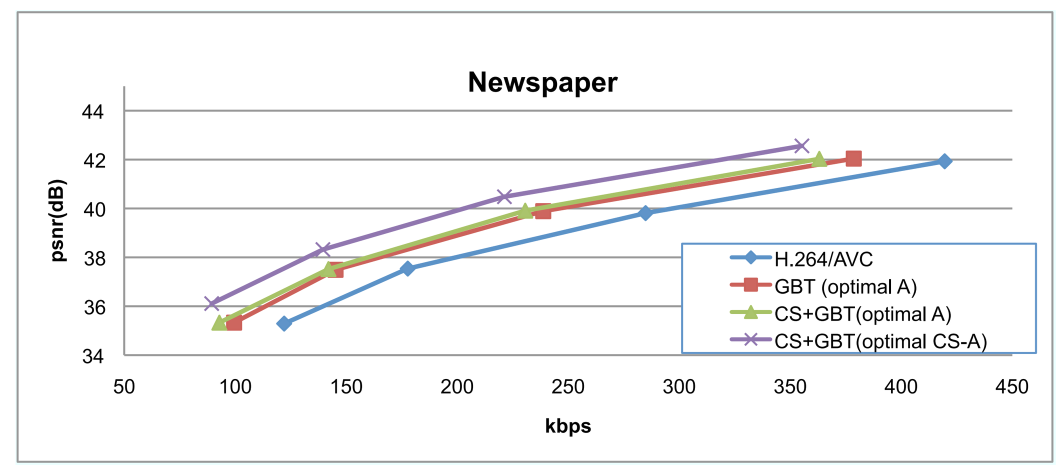
Average Mutual Coherence, μ_{avg}

- $\mu_{\mathbf{avg}}$ can be computed with Hadamard, $\mathbf{\Phi}$, and GBT matrix, $\mathbf{\Psi}$ [6]
  - → complex GBT construction
  - → Increase complexity especially when
  - ▷ depth map block size increases
  - ▷ involved in the greedy algorithm
- Upper bound of $\mu_{\mathbf{upper}}$
  - ▷ $\mu_{\mathbf{upper}} = \sqrt{\frac{\max_{\forall i} N_{G_i}}{N}}$
  - ▷ Depends on the maximum size of group
- Lower bound of $\mu_{\mathbf{lower}}$
  - ▷ $\mu_{\mathbf{lower}} = \max_{\forall k} \sqrt{\frac{\sum_{(l,m)\in E}(\Phi(k,l) - \Phi(k,m))^2}{2|E|}}$
  - ▷ $|\mathbf{E}|$: No. of links = 24 - no. of edges in $\mathbf{4 \times 4}$ blocks
  - ▷ Depends on Hadamard matrix and edge map
- $\mu_{\mathbf{avg}} = \frac{\mu_{\mathbf{lower}} + \mu_{\mathbf{upper}}}{2}$ independent of explicit GBT matrix

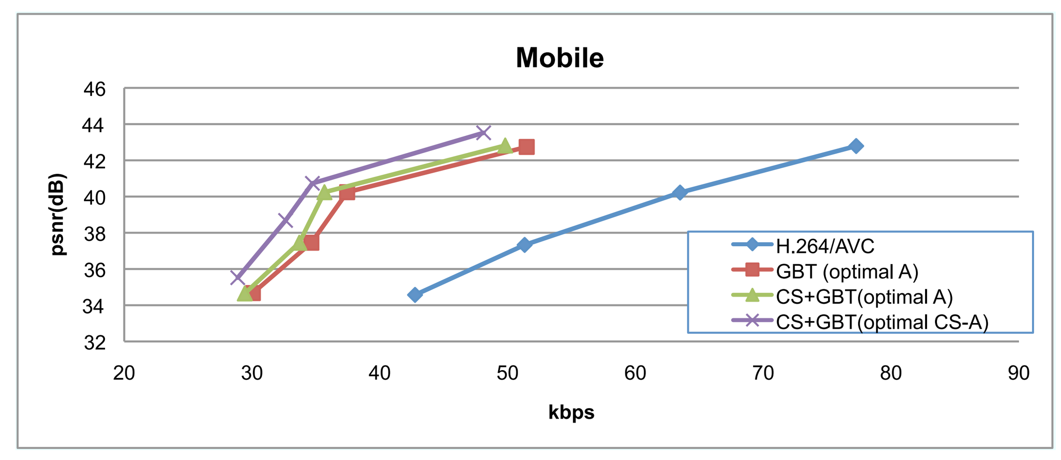
Simulation Result



(a) Ballet

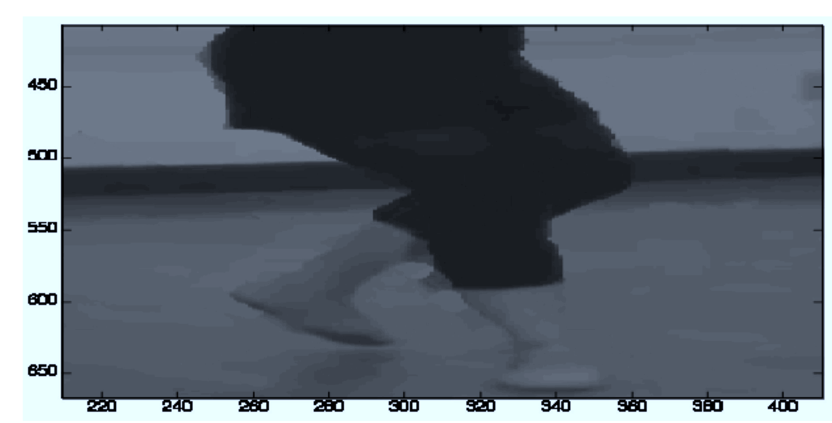


(b) Newspaper

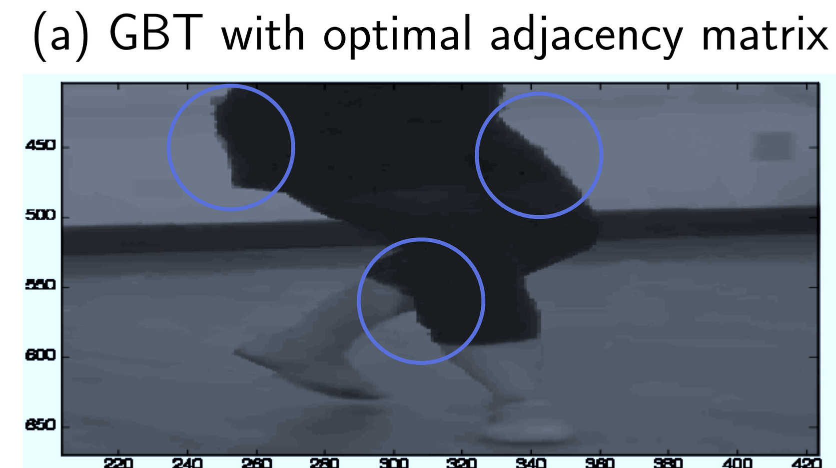


(c) Mobile

Figure: RD curve comparison of i) H.264/AVC ii) GBT and CS-GBT with optimal adjacency matrix [3] iii) CS-GBT with optimal adjacency matrix with averaged mutual coherence for different sequences: (a) Ballet (b) Newspaper (c) Mobile



(a) GBT with optimal adjacency matrix



(b) CS+GBT with optimal adjacency matrix for CS

Figure: Perceptual improvement in Ballet sequence: comparison of (a) GBT and CS-GBT with optimal adjacency matrix [3] (b) CS-GBT with optimal adjacency matrix with averaged mutual coherence

Sequence	BD-PSNR	BD-bitrate
Ballet	0.9	-49.4
Newspaper	1.5	-26.8
Mobile	9.2	-42.8

Table: BD-PSNR/bitrate results of CS-GBT compared to H.264/AVC.

Sequence	BD-PSNR	BD-bitrate
Ballet	0.3	-7.8
Newspaper	0.9	-16.1
Mobile	2.4	-9.7

Table: BD-PSNR/bitrate results of CS-GBT with optimal adjacency matrix for CS compared to CS-GBT with optimal adjacency matrix [3]

Conclusion

- Propose a block-based CS framework applicable to depthmap compression
- Approximate average mutual coherence without constructing GBT
- Find an optimal edgemap (adjacency matrix) for CS
- Proposed approach achieves a significant gain over H.264/AVC

References

- W.-S. Kim, A. Ortega, P. Lai, D. Tian, and C. Gomila, "Depth map distortion analysis for view rendering and depth coding," in *Proc. of IEEE Int. Conf. Image Proc., ICIP 2009*, Cairo, Egypt, Nov. 2009.
- ——, "Depth map coding with distortion estimation of rendered view," in *Proc. of IS&T/SPIE Electronic Imaging, VIPC 2010*, San Jose, CA, USA, Jan. 2010.
- W.-S. Kim, "3-d video coding system with enhanced rendered view quality," in *Ph.D. dissertation, University of Southern California*, 2011.
- M. Sarkis and K. Diepold, "Depth map compression via compressed sensing," in *Proc. of IEEE Int. Conf. Image Proc., ICIP 2009*, Cairo, Egypt, Nov. 2009.
- J. Duan, L. Zhang, R. Pan, and Y. Sun, "An improved video coding scheme for depth map sequences based on compressed sensing," in *International Conference on Multimedia Technology (ICMT)*, Hangzhou, Aug. 2011.
- M. Elad, "Optimized projections for compressed sensing," *IEEE Trans. on Signal Processing*, vol. 55, no. 12, pp. 5695–5702, 2007.

http://biron.usc.edu/wiki/index.php?title=CompressionGroup

sungwonL@usc.edu, ortega@sipi.usc.edu

Ming Hsieh Department of Electrical Engineering