Motion-Adaptive Transforms based on Vertex-Weighted Graphs

Du Liu and Markus Flierl

School of Electrical Engineering
KTH Royal Institute of Technology
100 44 Stockholm, Sweden

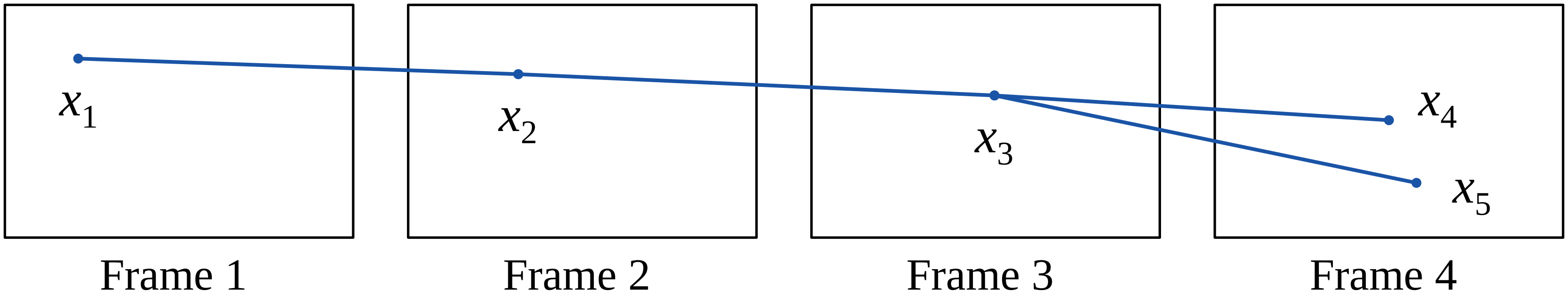
Motivation

- Limitation of motion-compensated predictive coding for packet-based networks
- Our approach:
  - Design motion-adaptive temporal transforms
  - Use vertex-weighted graphs to represent the motion

Outline

- Motion and vertex-weighted graphs
- Constrained energy compaction
- Motion-adaptive transforms
- Experimental results
- Conclusion

Motion and Vertex-Weighted Graphs



- $x_1, x_2, \dots, x_n$ are pixels connected by block-based motion estimation
- A graph is formed by the connection of motion vectors
- Vertex weights is given by the values of $x_1, x_2, \dots, x_n$
- Vector of actual pixel values: $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$

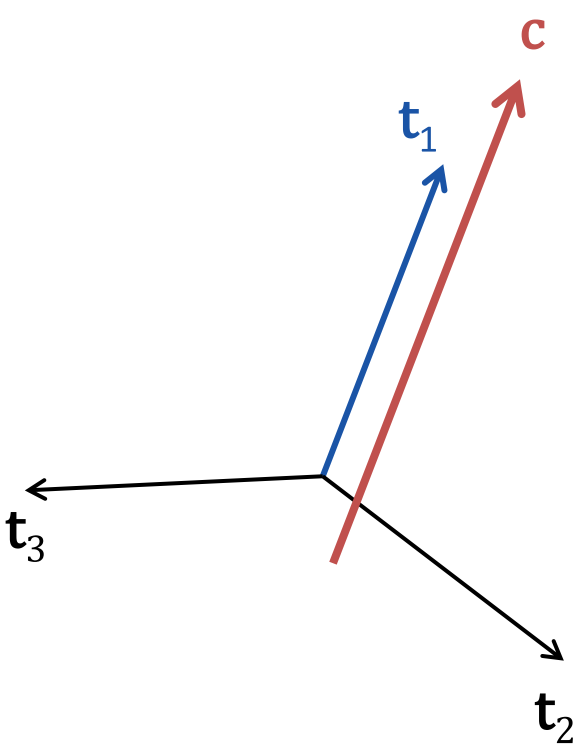
Ideal Motion

- Energy compaction changes the magnitude of the pixels
- Ideal motion implies constant intensity of connected pixels
- Energy compaction + ideal motion
  - Use actual values for the vertex-weighted graph
  - Use scale factors to accommodate energy compaction[1]
  - $x_k = c_k x'_k$ for $k = 1,2,\ldots,n$

    where x'_k is the original pixel value, c_k is the scale factor.
- For ideal motion: $x'_1 = x'_2 = \cdots = x'_n$
- Vector of scale factors: $\mathbf{c} = [c_1, c_2, \ldots, c_n]^T$

[1] M. Flierl and B. Girod, "A motion-compensated orthogonal transform with energy-concentration constraint, " *IEEE International Workshop on Multimedia Signal Processing*, 2006.

Ideal Motion



- Let the transform matrix be

$$T = \left[\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_n\right]$$

- Output

$$\mathbf{y} = T^T\mathbf{x} = T^T \begin{bmatrix} c_1 x'_1 \\ c_2 x'_1 \\ \vdots \\ c_n x'_1 \end{bmatrix} = \begin{bmatrix} \sqrt{\mathbf{c}^T\mathbf{c}} \cdot x'_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\Rightarrow \sqrt{\mathbf{c}^T\mathbf{c}} = \mathbf{t}_1^T\mathbf{c}$$

$$0 = \mathbf{t}_k^T\mathbf{c}, k = 2, \ldots, n$$

- Subspace constraint

$$\mathbf{t}_1 = \frac{\mathbf{c}}{\sqrt{\mathbf{c}^T\mathbf{c}}} = \frac{\mathbf{c}}{\|\mathbf{c}\|_2}$$

Design of Motion-Adaptive Transform

- Classic transform coding
  - Energy compaction
  - Karhunen–Loève Transform (KLT)
- Transform along motion trajectory
  - Invertibility
  - Example: Motion-Compensated Orthogonal Transform (MCOT)[1]
- Our design goal
  - Invertible motion-adaptive transform
  - Optimal energy compaction given the graph

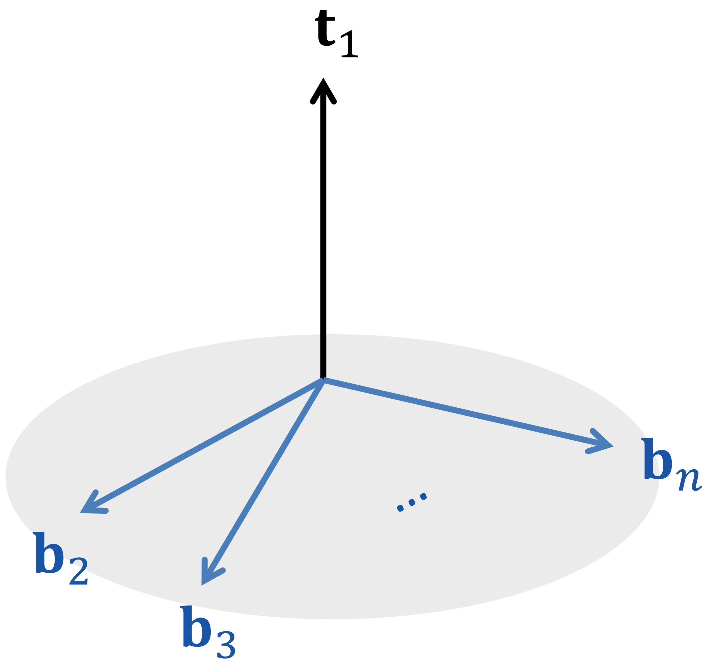
Constrained Energy Compaction

- Let L be the autocorrelation matrix of x

$$\begin{aligned} \min_{\mathbf{t}_k} \quad & -\mathbf{t}_k^T L \mathbf{t}_k, \quad k = 2, \ldots, n, \\ \text{s.t.} \quad & \mathbf{t}_k^T \mathbf{t}_k = 1, \\ & \mathbf{t}_k^T \mathbf{t}_j = 0, \quad j = 1, \ldots, k-1, \\ & \mathbf{t}_1 = \frac{\mathbf{c}}{\|\mathbf{c}\|_2}. \end{aligned}$$

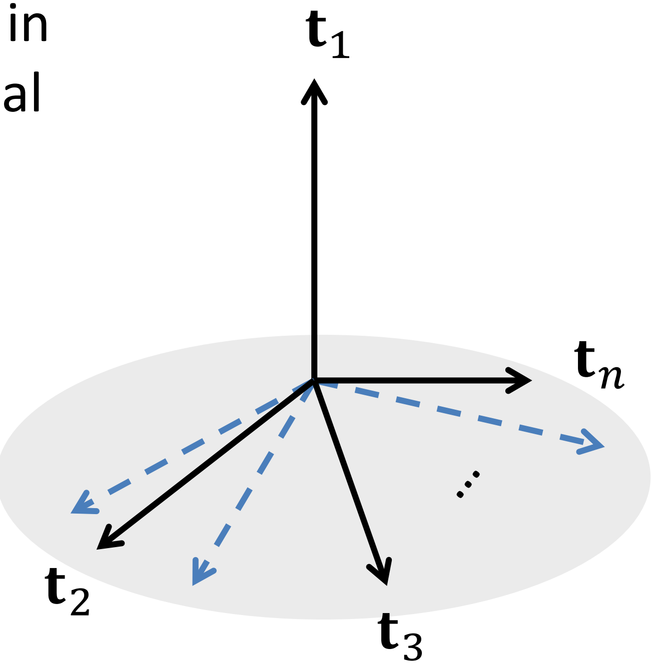
- Unconstrained formulation using a Lagrangian cost

Subspace-Constrained Transform (SCT)



$$B' = [\mathbf{b}_2, \mathbf{b}_3, \ldots, \mathbf{b}_n]$$

Find rotation A' in (n-1)-dimensional subspace



$$T' = [\mathbf{t}_2, \mathbf{t}_3, \ldots, \mathbf{t}_n] = B'A'$$

Subspace-Constrained Transform (SCT)

- Let $M' = B'^T L B'$ be the autocorrelation matrix based on B'. M' is an $(n-1) \times (n-1)$ matrix.
- We show that the rotation

  A' is a matrix of eigenvectors for M'

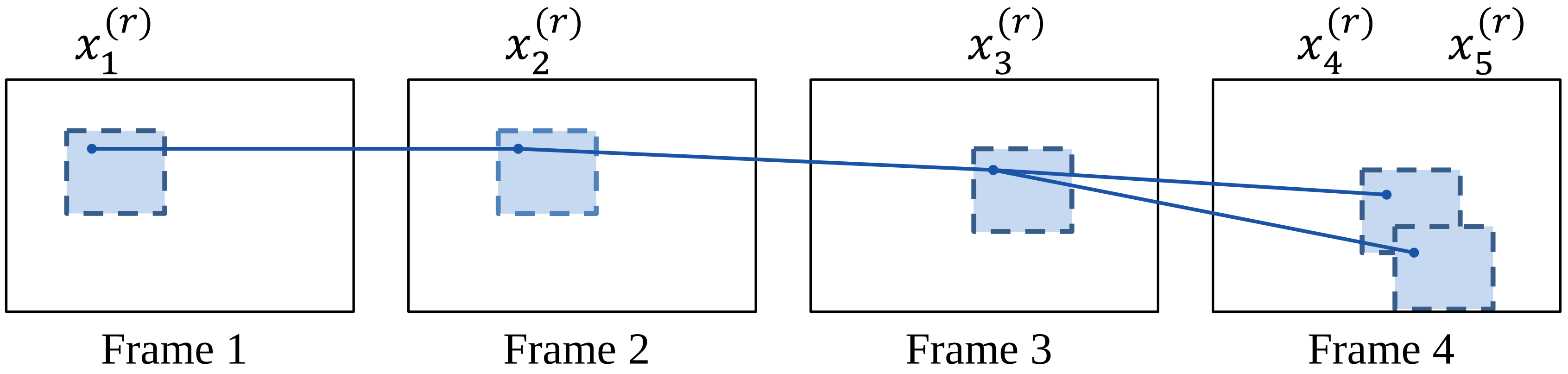
- We are free to choose any basis B' to construct our T
- $T = [\mathbf{t}_1, T'] = [\mathbf{t}_1, B'A']$

Discussion

- If $\mathbf{t}_1$ is identical to the first basis vector of the KLT, SCT is the same as the KLT.
- In general, SCT approximates the KLT for a given graph.
- The following experiments illustrate
  - $\mathbf{t}_1$ approximates the first basis vector of the KLT
  - the energy compaction of SCT is close to that of the KLT

Experimental Setup

- Estimate the autocorrelation matrix L



  - Consider samples that use the same vertex-weighted graph
  - $L_{ij} = \sum_{r=1}^{N} x_i^{(r)} x_j^{(r)}$, where N is the total number of samples, r is an instance of a graph

Experimental Setup

- A hierarchical decomposition is performed on each GOP
- Results on energy compaction are given for comparison
  - MCOT: basis vectors are dependent on the motion vectors only
    two tap, hierarchical Haar
  - SCT: basis vectors $\mathbf{t}_2, \dots, \mathbf{t}_n$ are signal dependent
  - KLT: basis vectors $\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_n$ are signal dependent

Experimental Results

- QCIF Foreman, relative energy in the first four temporal subbands, 1st decomposition level

	Lowband	Highband 1	Highband 2	Highband 3
MCOT	99.36%	0.12%	0.42%	0.10%
SCT	99.36%	0.54%	0.08%	0.02%
KLT	99.42%	0.49%	0.07%	0.02%

- QCIF City, relative energy in the first four temporal subbands, 1st decomposition level

	Lowband	Highband 1	Highband 2	Highband 3
MCOT	93.99%	1.54%	3.13%	1.34%
SCT	93.99%	4.08%	1.40%	0.53%
KLT	94.18%	3.95%	1.35%	0.52%

Experimental Results

- QCIF Foreman, GOP = 16, relative energy in the 2nd decomposition level

	Lowband	Highband 1	Highband 2	Highband 3
MCOT	98.32%	0.37%	0.72%	0.59%
SCT	98.32%	1.29%	0.30%	0.09%
KLT	98.35%	1.30%	0.28%	0.07%

- QCIF City, GOP = 16, relative energy in the 2nd decomposition level

	Lowband	Highband 1	Highband 2	Highband 3
MCOT	90.12%	2.95%	4.27%	2.66%
SCT	90.12%	7.47%	1.86%	0.55%
KLT	91.22%	6.50%	1.77%	0.51%

Conclusion

- We present a class of motion-adaptive transforms that is based on vertex-weighted graphs.
- The vertex-weighted graph determines uniquely the first basis vector of the linear transform.
- This first vector defines a subspace that constrains the energy compaction of our transform.
- SCT achieves optimal energy compaction, given our subspace constraint.