A Motion-Compensated Orthogonal Transform with Energy Concentration Constraint

Markus Flierl and Bernd Girod
mflierl@stanford.edu

1 Introduction

Problem

- Motion-compensated (MC) lifted Haar wavelet deviates substantially from orthonormality due to motion compensation

Why Orthogonal Transforms?

- Optimal for certain transform coding schemes at high rates
- Provide highly robust video representations

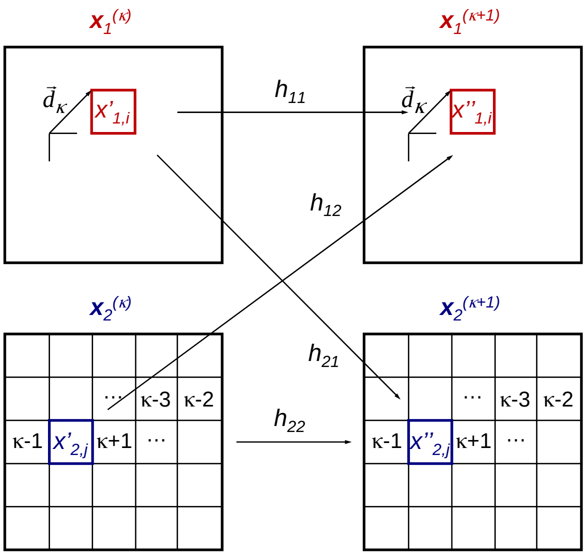
Goal

- MC transform that is orthogonal for any motion field

2 MC Orthogonal Transform

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = T \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \quad \text{with} \quad T = T_k T_{k-1} \cdots T_\kappa \cdots T_2 T_1 \quad \text{where} \quad T_\kappa T_\kappa^T = I$$

Incremental Transform T_κ



$$T_\kappa = \begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots \\ \cdots & 0 & h_{11} & 0 & \cdots & 0 & h_{12} & 0 & \cdots \\ \cdots & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \cdots & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots \\ \cdots & 0 & h_{21} & 0 & \cdots & 0 & h_{22} & 0 & \cdots \\ \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

$$H = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}$$

$$H = \frac{1}{\sqrt{1+a_n^2}} \begin{bmatrix} 1 & a_n \\ -a_n & 1 \end{bmatrix}$$

decorrelation factor

Energy Concentration Constraint

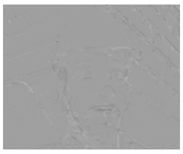
Consider previous incremental transforms by **scale factor *v***

$$\begin{bmatrix} v_m x_{1,i} \\ 0 \end{bmatrix} = \frac{1}{\sqrt{1+a_n^2}} \begin{bmatrix} 1 & a_n \\ -a_n & 1 \end{bmatrix} \begin{bmatrix} v_{n_1} x_{1,i} \\ v_{n_2} x_{1,i} \end{bmatrix} \qquad a_n = \frac{v_{n_2}}{v_{n_1}} \qquad v_m = \sqrt{v_{n_1}^2 + v_{n_2}^2}$$

Scale counter: $n = v_n^2 - 1$ Scale counter update rule: $m = n_1 + n_2 + 1$

3 Experimental Results

Example: Two Decomposition Levels

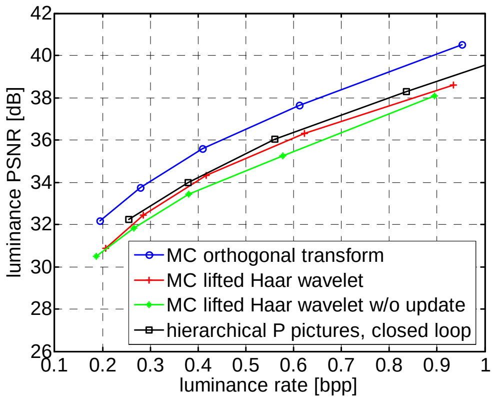


temporal high band, 1st level

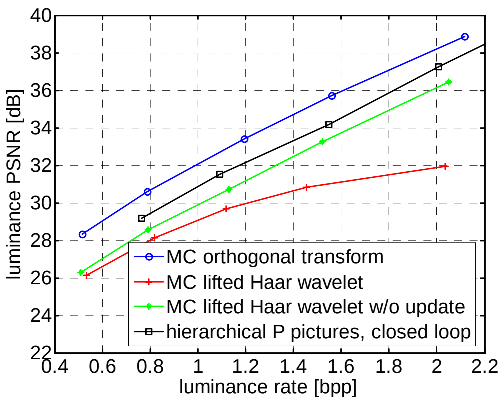
temporal low band, 2nd level

rescaled temporal low band, 2nd level

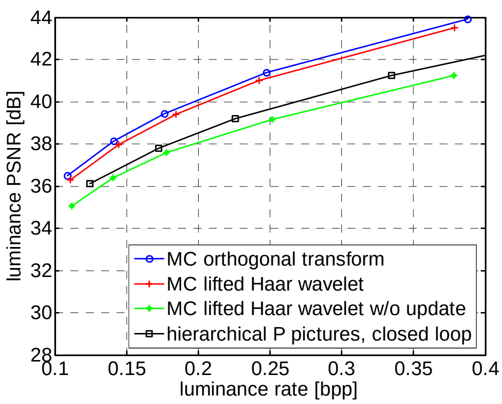
Assessment of Energy Compaction



Foreman
QCIF, 30 fps
288 frames
K=16 GOP
8x8 block motion



Bus
QCIF, 15 fps
64 frames
K=16 GOP
8x8 block motion



Mother&Daughter
QCIF, 30 fps
64 frames
K=16 GOP
8x8 block motion

4 Conclusions

Orthonormality improves energy compaction and provides highly robust video representations

Max Plank Center for Visual Computing and Communication **Stanford University**