A Double 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

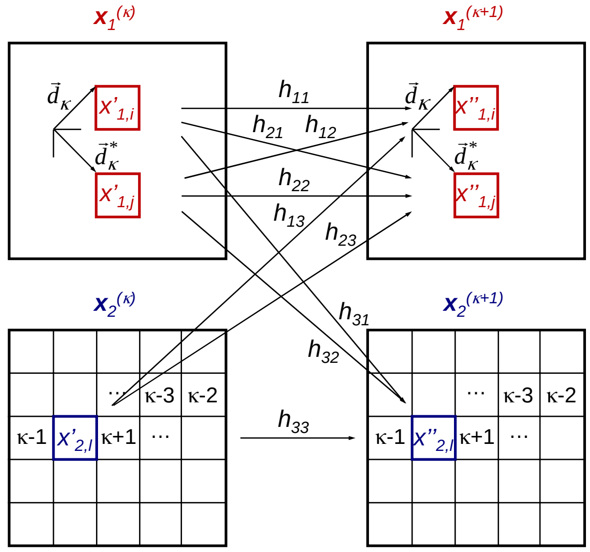
Goal

- Extend integer-pel accurate MC orthogonal transform in [1]
- MC transform that is orthogonal for any 2-motion field

2 Double 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



$$T_\kappa = \begin{bmatrix}
\ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\
\cdots & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots\\
\cdots & 0 & h_{11} & \cdots & 0 & h_{12} & \cdots & 0 & h_{13} & \cdots\\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\
\cdots & 0 & 0 & \cdots & 1 & 0 & \cdots & 0 & 0 & \cdots\\
\cdots & 0 & h_{21} & \cdots & 0 & h_{22} & \cdots & 0 & h_{23} & \cdots\\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots\\
\cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 1 & 0 & \cdots\\
\cdots & 0 & h_{31} & \cdots & 0 & h_{32} & \cdots & 0 & h_{33} & \cdots\\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{bmatrix}$$

$$H := \{h_{\mu\nu}\} = H_3 H_2 H_1$$

$$H_3 = \begin{bmatrix}\cos(\psi) & \sin(\psi) & 0\\ -\sin(\psi) & \cos(\psi) & 0\\ 0 & 0 & 1\end{bmatrix}$$

$$H_2 = \begin{bmatrix}1 & 0 & 0\\ 0 & \cos(\theta) & \sin(\theta)\\ 0 & -\sin(\theta) & \cos(\theta)\end{bmatrix}$$

$$H_1 = \begin{bmatrix}\cos(\phi) & \sin(\phi) & 0\\ -\sin(\phi) & \cos(\phi) & 0\\ 0 & 0 & 1\end{bmatrix}$$

Energy Concentration Constraint

Pixels are connected by 2-motion: $x_{2,l} = x_{1,i} = x_{1,j}$

Consider previous incremental transforms by **scale factors u_α, v_β**

$$x'_{1,i} = v_1 x_{1,i} \qquad x''_{1,i} = u_1 x_{1,i}$$
$$x'_{1,j} = v_2 x_{1,j} \qquad x''_{1,j} = u_2 x_{1,j}$$
$$x'_{2,l} = v_3 x_{2,l}$$

$$\begin{bmatrix}u_1 x_{1,i}\\ u_2 x_{1,i}\\ 0\end{bmatrix} = H_3 H_2 H_1 \begin{bmatrix}v_1 x_{1,i}\\ v_2 x_{1,i}\\ v_3 x_{1,i}\end{bmatrix}$$

Energy conservation: $u_1^2 + u_2^2 = v_1^2 + v_2^2 + v_3^2$

Energy concentration:

$$\tan(\phi) = -\frac{v_1}{v_2}, \quad \tan(\theta) = \frac{v_3}{\sqrt{v_1^2 + v_2^2}}, \quad \tan(\psi) = \frac{u_1}{u_2} \quad \text{s.t.} \quad u_1^2 = v_1^2 + \frac{v_3^2}{2}, \quad u_2^2 = v_2^2 + \frac{v_3^2}{2}$$

Scale counter: $m_\alpha = u_\alpha^2 - 1$, $n_\beta = v_\beta^2 - 1$

Scale counter update rule: $m_1 = n_1 + \frac{n_3 + 1}{2}$ and $m_2 = n_2 + \frac{n_3 + 1}{2}$

3 Experimental Results

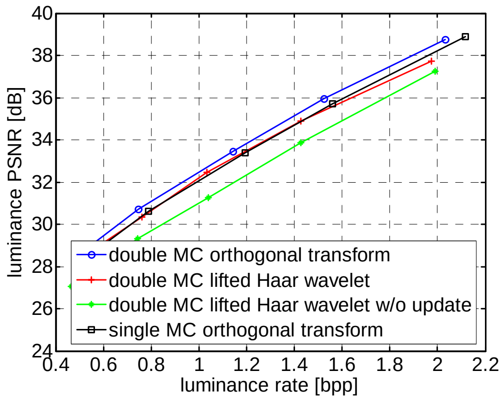
Example: Two Decomposition Levels

temporal high band, 1st level

temporal low band, 2nd level

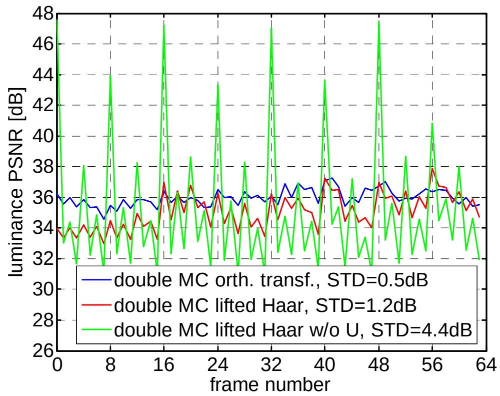
rescaled temporal low band, 2nd level

Assessment of Energy Compaction



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

Reconstructed Image Quality



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

4 Conclusions

Orthonormality improves energy compaction, provides highly robust video representations, and permits 2-motion compensation

References

[1] M. Flierl, B. Girod, "A motion-compensated orthogonal transform with energy-concentration constraint," IEEE MMSP, Victoria, BC, Oct. 2006.

Max Plank Center for Visual Computing and Communication Stanford University