Half-Pel Accurate Motion-Compensated Orthogonal Video Transforms

Markus Flierl and Bernd Girod

Max Planck Center for Visual Computing and Communication
Motivation

- Motion-compensated 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

Motion-adaptive transform that strictly maintains orthonormality while permitting flexible
- Integer-pel accurate motion compensation and
- Sub-pel accurate motion compensation
Outline

- Motion-Compensated Orthogonal Transform (MCOT)
- Single MC incremental transform
  - Energy concentration constraint
  - Example for a dyadic decomposition of a group of pictures
- Double MC incremental transform
  - Euler rotations
  - Energy concentration constraint
- P-hypothesis MC incremental transform
- Example: Half-pel MC with averaging filter
- Experimental results
Orthogonal Video Transform

- Orthogonal transform for pairs of input images:

\[
\begin{align*}
\text{low band image} & \quad \rightarrow \quad \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\
\text{high band image} & \quad \rightarrow
\end{align*}
\]

- Factor \( T \) into a sequence of \( k \) incremental transforms:

\[
T = T_k T_{k-1} \cdots T_\kappa \cdots T_2 T_1
\]

- Each incremental transform is orthogonal: \( T_\kappa T_\kappa^T = I \)

- Incremental transforms generate a sequence of transformed image pairs:

\[
\begin{align*}
\begin{pmatrix} x_1^{(\kappa+1)} \\ x_2^{(\kappa+1)} \end{pmatrix} &= T_\kappa \begin{pmatrix} x_1^{(\kappa)} \\ x_2^{(\kappa)} \end{pmatrix}
\end{align*}
\]
Single MC Incremental Transform

\[ x_1^{(\kappa)} \]

\[ x_1^{(\kappa+1)} \]

\[ \vec{d}_\kappa \]

\[ h_{11} \]

\[ x_{1,i}^{'} \]

\[ x_{1,i}^{''} \]

\[ \text{low band image to-be} \]

\[ x_2^{(\kappa)} \]

\[ x_2^{(\kappa+1)} \]

\[ h_{12} \]

\[ h_{21} \]

\[ h_{22} \]

\[ \kappa-1 \]

\[ \kappa \]

\[ \kappa+1 \]

\[ \kappa-2 \]

\[ \kappa-3 \]

\[ \cdots \]

\[ \text{high band image to-be} \]
Single MC Incremental Transform

\[ T_{ki} = \begin{pmatrix}
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
\ldots & 1 & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots \\
\ldots & 0 & h_{11} & 0 & \ldots & 0 & h_{12} & 0 & \ldots \\
\ldots & 0 & 0 & 1 & \ldots & 0 & 0 & 0 & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
\ldots & 0 & 0 & 0 & \ldots & 1 & 0 & 0 & \ldots \\
\ldots & 0 & h_{21} & 0 & \ldots & 0 & h_{22} & 0 & \ldots \\
\ldots & 0 & 0 & 0 & \ldots & 0 & 0 & 1 & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
\end{pmatrix} \]

- \( i \)-th pixel in \( x_1 \)
- \( j \)-th pixel in \( x_2 \)

\[ H = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} = \frac{1}{\sqrt{1 + a^2}} \begin{pmatrix} 1 & a \\ -a & 1 \end{pmatrix} \]

\[ HH^T = I \]

decorrelation factor
Example: Single MC Orthogonal Transform

\[
\begin{bmatrix}
  x''_{1,i} \\
  x''_{2,j}
\end{bmatrix} = \frac{1}{\sqrt{1 + a_K^2}} \begin{bmatrix} 1 & a_K \\ -a_K & 1 \end{bmatrix} \begin{bmatrix} x'_{1,i} \\
  x'_{2,j}
\end{bmatrix}
\]
SMCOT: Energy Concentration Constraint

- Choose decorrelation factor for each incremental transform such that the energy in the high band to-be is removed.
- Assume that pixel $x_{2,j}$ is connected to pixel $x_{1,i}$, i.e., $x_{2,j} = x_{1,i}$.
- Note that pixel $x_{1,i}$ may have been processed previously!
- Therefore, let $v_1$ be the scale factor for pixel $x_{1,i}$.
- After processing, let $u_1$ be the scale factor for pixel $x_{1,i}$.
- For higher levels of temporal decomposition, $x_{2,j}$ is a low band coefficient that carries a scale factor.
- Therefore, let $v_2$ be the scale factor for pixel $x_{2,j}$.
- Now, resulting high band pixel to-be $x''_{2,j}$ shall be zero:

$$
\begin{pmatrix}
  u_1x_{1,i} \\
  0
\end{pmatrix} = H
\begin{pmatrix}
  v_1x_{1,i} \\
  v_2x_{1,i}
\end{pmatrix}
$$
Definition of Scale Counters

- Let $n_1, n_2$ be the scale counters for pixel $x_{1,i}, x_{2,j}$
- $n_1, n_2$ simply count how often the pixel $x_{1,i}, x_{2,j}$ are used as reference for motion compensation
- In the beginning, the scale counter is $n = 0$ and the scale factor is $v = 1$
- Let $m_1$ be the scale counters for pixel $x_{1,i}$ after being processed by the incremental transform
- For arbitrary scale counter $m$ and $n$, the scale factors are
  \[ u = \sqrt{m + 1} \quad \text{and} \quad v = \sqrt{n + 1} \]
- Example: Scale counter update rule for SMCOT:
  \[ m_1 = n_1 + n_2 + 1 \]
IP-MCOT Experimental Results

- temporal high band first decomposition level
- temporal high band second decomposition level
IP-MCOT Experimental Results

temporal low band second decomposition level

rescaled temporal low band second decomposition level

\[ v = \sqrt{n + 1} \]
IP-MCOT Experimental Results

![Graph showing luminance PSNR vs. luminance rate for different motion-compensated orthogonal transforms. The graph includes the following lines:
- Blue line: MC orthogonal transform
- Red line: MC lifted Haar wavelet
- Green line: MC lifted Haar wavelet w/o update
- Black line: Hierarchical P pictures, closed loop

The graph is labeled with the following details:
- Foreman QCIF
- 30 fps
- 288 frames
- GOP size K=16
- 8x8 block motion]
Double MC Incremental Transform

\[ x^{(\kappa)}_1 \]

\[ x^{(\kappa+1)}_1 \]

\[ x^{(\kappa)}_2 \]

\[ x^{(\kappa+1)}_2 \]

Low band image to-be

High band image to-be

\[ h_{11} \]
\[ h_{12} \]
\[ h_{13} \]
\[ h_{21} \]
\[ h_{22} \]
\[ h_{23} \]
\[ h_{31} \]
\[ h_{32} \]
\[ h_{33} \]
Double MC Incremental Transform

\[ T_k = \begin{pmatrix} \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \end{pmatrix} \]

\[ \ldots \begin{array}{cccccccccccccccc} 1 & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots \end{array} \]

\[ \ldots \begin{array}{cccccccccccccccc} 0 & h_{11} & 0 & \ldots & 0 & h_{12} & 0 & \ldots & 0 & h_{13} & 0 & \ldots \end{array} \]

\[ \ldots \begin{array}{cccccccccccccccc} 0 & 0 & 1 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots \end{array} \]

\[ \vdots \]

\[ \ldots \begin{array}{cccccccccccccccc} 0 & 0 & 0 & \ldots & 1 & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots \end{array} \]

\[ \ldots \begin{array}{cccccccccccccccc} 0 & h_{21} & 0 & \ldots & 0 & h_{22} & 0 & \ldots & 0 & h_{23} & 0 & \ldots \end{array} \]

\[ \ldots \begin{array}{cccccccccccccccc} 0 & 0 & 0 & \ldots & 0 & h_{31} & 0 & \ldots & 0 & h_{32} & 0 & \ldots \end{array} \]

\[ \ldots \begin{array}{cccccccccccccccc} 0 & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & 0 & 1 & \ldots \end{array} \]

\[ \vdots \]

\[ \ldots \begin{array}{cccccccccccccccc} \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \end{array} \]

\[ \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \]

with \[ HH^T = I \]
Euler’s Rotation Theorem

- Any rotation in 3D can be given as a composition of rotations about three axes, i.e., \( H = H_3 H_2 H_1 \)
- We choose the following composition:

\[
H = \begin{pmatrix}
\cos(\psi) & 0 & \sin(\psi) \\
0 & 1 & 0 \\
-\sin(\psi) & 0 & \cos(\psi)
\end{pmatrix}
\begin{pmatrix}
1 & 0 & 0 \\
0 & \cos(\theta) & -\sin(\theta) \\
0 & \sin(\theta) & \cos(\theta)
\end{pmatrix}
\begin{pmatrix}
\cos(\phi) & 0 & \sin(\phi) \\
0 & 1 & 0 \\
-\sin(\phi) & 0 & \cos(\phi)
\end{pmatrix}
\]

- Euler angles \( \psi, \theta, \phi \) are determined by the energy concentration constraint
DMCOT: Energy Concentration Constraint

- Choose 3 Euler angles for each incremental transform.
- Assume that pixel $x_{2,l}$ is connected to pixel $x_{1,i}$, i.e., $x_{2,l} = x_{1,i}$.
- Assume that pixel $x_{2,l}$ is connected to pixel $x_{1,j}$, i.e., $x_{2,l} = x_{1,j}$.
- State zero-energy constraint for the high band pixel:
  \[
  \begin{pmatrix}
  u_1 x_{1,i} \\
  u_2 x_{1,i} \\
  0
  \end{pmatrix}
  = H_3 H_2 H_1
  \begin{pmatrix}
  v_1 x_{1,i} \\
  v_2 x_{1,i} \\
  v_3 x_{1,i}
  \end{pmatrix}
  \]

- Obtain Euler angles for averaging the 2 hypotheses.
- Use definition of scale counters.
- Choose scale counter update rule for double MCOT:
  \[
  m_1 = n_1 + \frac{n_3 + 1}{2}
  \quad \text{and} \quad
  m_2 = n_2 + \frac{n_3 + 1}{2}
  \]
P-Hypothesis MC Incremental Transform

- Number of hypotheses \( P \) is a power of 2
- Assume that high-band pixel to-be \( x_{2,i} \) is connected to all \( P = 2^r \) hypotheses pixel, where \( r = 0, 1, 2, \ldots \)
- Incremental transform is given as a composition of Euler rotations in \( P+1 \) dimensions
- Obtain Euler angles for dyadic averaging of pairs of hypotheses
- Hence, each of the \( P \) hypotheses is weighted by \( 1/P \)
- Choose scale counter update rule for P-MCOT:

\[
m_p = n_p + n_{P+1} + \frac{1}{P} \quad \text{for} \quad p = 1, 2, \ldots, P
\]
Half-Pel MC with Averaging Filter

- IP position via 1-hypothesis MC incremental transform
- HP positions 1 and 2 via 2-hypothesis MC incremental transform averaging IP positions A, B and A, C, respectively
- HP position 3 via 4-hypothesis MC incremental transform
- Type of incremental transform can be chosen on block level

integer-pel positions

half-pel positions
Half-Pel Accurate Motion-Compensated Orthogonal Video Transforms
Conclusions

- Class of motion-compensated orthogonal video transforms
- Highly flexible incremental transforms
- Energy concentration constraint
- Permit sub-pel accurate motion compensation
- Bidirectionally MC orthogonal transform to be presented at ICASSP 2007
Further Reading

http://www.orthogonalvideo.org