Investigation of Motion-Compensated Lifted Wavelet Transforms

Markus Flierl
and Bernd Girod

Department of Electrical Engineering,
Stanford University
Outline

- Coding scheme with motion-compensated wavelets
- Experimental results for temporal Haar and 5/3 wavelets
- Signal model and performance bounds
- Comparison to predictive coding
Coding Scheme

- Dyadic decomposition for each group of $K$ pictures
- Motion-compensated Haar and 5/3 wavelet
- 16x16 block motion compensation with half-pel accuracy
- Intra-frame coding with 8x8 DCT and run-length coding
- Same quantizer step-size for all K intra-frame encoder
Motion-Compensated Haar Wavelet

Update step uses negative motion vector of corresponding prediction step
Motion-Compensated 5/3 Wavelet

Update steps uses negative motion vectors of corresponding prediction steps
Motion-Compensated Haar & 5/3 Wavelet

Mother & Daughter, QCIF, 30 fps

![Graph showing PSNR Y vs R for different wavelet transforms.](image-url)
Motion-Compensated Haar & 5/3 Wavelet

Mobile & Calendar, QCIF, 30 fps
Theoretical Signal Model

- Let $s_k[x,y]$ be the $k$-th picture at pel-location $x,y$
- The signals are space-discrete and band-limited
- Ideal reconstruction is used for sub-pel accurate displacements
- Displacement operation is invertible
Equivalent Motion-Compensated Wavelets

Invertible displacement operations are assumed
Dyadic Haar Transform

Both, true displacements \( d \) and estimated displacements \( \hat{d} \) are additive
Generalized Signal Model

Any input picture can be reference picture

\[ v \]

model picture

\[ \Delta_k \]

k-th displacement error

\[ n_k \]

k-th noise signal

\[ c_k \]

k-th motion-compensated signal

\[ y_k \]

k-th transform signal
Performance Measure

Rate difference for each picture $k$:

$$\Delta R_k = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{1}{2} \log_2 \left( \frac{\Phi y_k y_k(\omega)}{\Phi c_k c_k(\omega)} \right) d\omega$$

- Measures maximum bit-rate reduction
- Compares to optimum intra-frame encoding
- For the same mean squared reconstruction error
- For Gaussian signals

Average rate difference

$$\Delta R = \sum_{k=0}^{K-1} \Delta R_k$$
Rate Difference with Negligible Noise

Calibration:

\[ \beta = 0.5 \log_2(12 \sigma^2_\Delta) \]

- Integer-pel \( \beta = 0 \)
- Half-pel \( \beta = -1 \)
- Quarter-pel \( \beta = -2 \)
Rate Difference with RNL = -30 dB

Calibration:

$$\beta = 0.5 \log_2(12 \sigma^2_\Delta)$$

- Integer-pel $\beta = 0$
- Half-pel $\beta = -1$
- Quarter-pel $\beta = -2$
Comparison to Predictive Coding

- Predictive coding scheme:
  - 16x16 block motion compensation with half-pel accuracy
  - Previous reference frame only
  - Intra-frame coding with 8x8 DCT and run-length coding
  - Only one intra-frame in the beginning of the sequence
  - Same quantizer step-size for all inter-frames

- Motion-compensated wavelet coding scheme uses the same components
Comparison to Predictive Coding

Mother & Daughter, QCIF, 30 fps

![Graph showing comparison to predictive coding](image.png)

- Blue diamonds: 5/3, K=32
- Black circles: Haar, K=32
- Red circles: Prediction, K=288

Flierl and Girod: Investigation of MC Lifted Wavelet Transforms

April 23, 2003  16
Comparison to Predictive Coding

Mobile & Calendar, QCIF, 30 fps

- 5/3, K=32
- Haar, K=32
- Prediction, K=288
Comparison to Predictive Coding

RNL = -30 dB

Calibration:

\[ \beta = 0.5 \log_2(12 \sigma^2_\Delta) \]

- Integer-pel \( \beta = 0 \)
- Half-pel \( \beta = -1 \)
- Quarter-pel \( \beta = -2 \)
Conclusions

- Rate difference is limited to 1 bit per sample per displacement inaccuracy step

- Gain by accurate motion compensation is limited by residual noise

- Motion-compensated 3-d transform coding outperforms predictive coding by at most 0.5 bits per sample