Visual Communications and Image Processing 2004

Video Coding with Lifted Wavelet Transforms and Complementary Motion-Compensated Signals

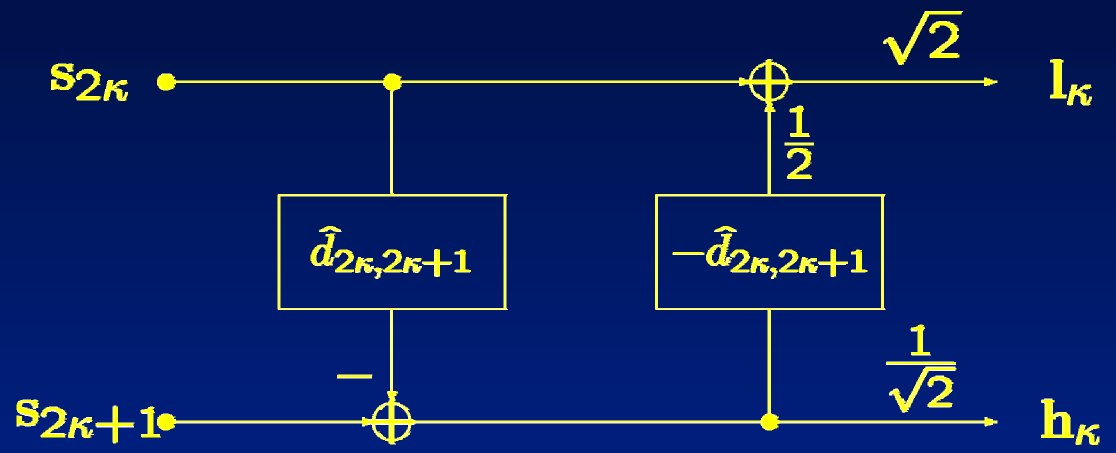
Markus Flierl, Pierre Vandergheynst
Signal Processing Institute
Swiss Federal Institute of Technology, Lausanne

Bernd Girod
Department of Electrical Engineering
Stanford University

Outline

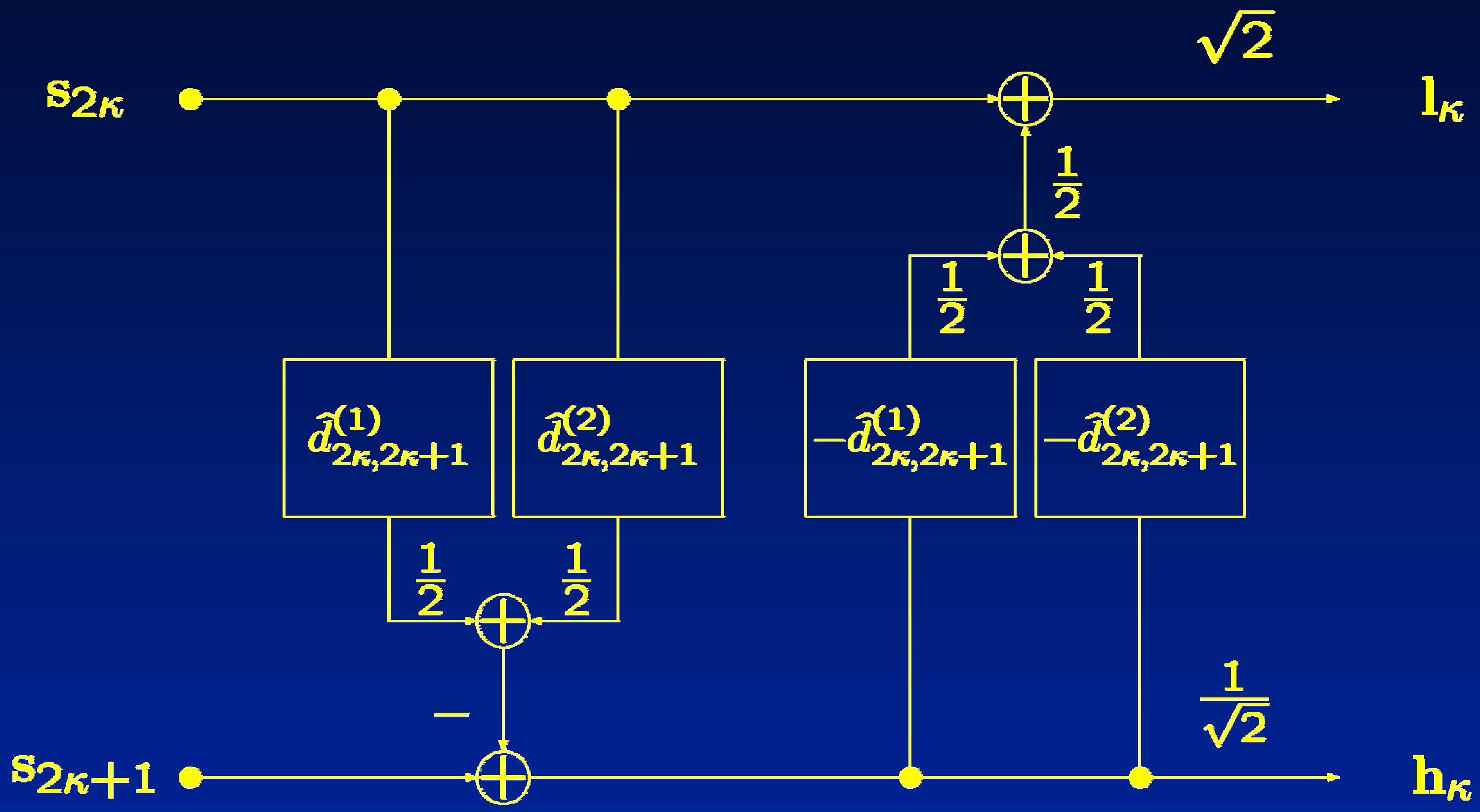
- Wavelets with lifting scheme and motion compensation
- Wavelets and superposition of multiple motion-compensated signals
- Video coding scheme and experimental results for temporal Haar and 5/3 wavelets
- Signal model and performance bounds for complementary motion-compensated signals
- Comparison to predictive coding with complementary motion-compensated signals

Motion-Compensated Haar Wavelet



Update step uses negative motion vector of corresponding prediction step

Haar Wavelet with N=2 Motion-Compensated Signals

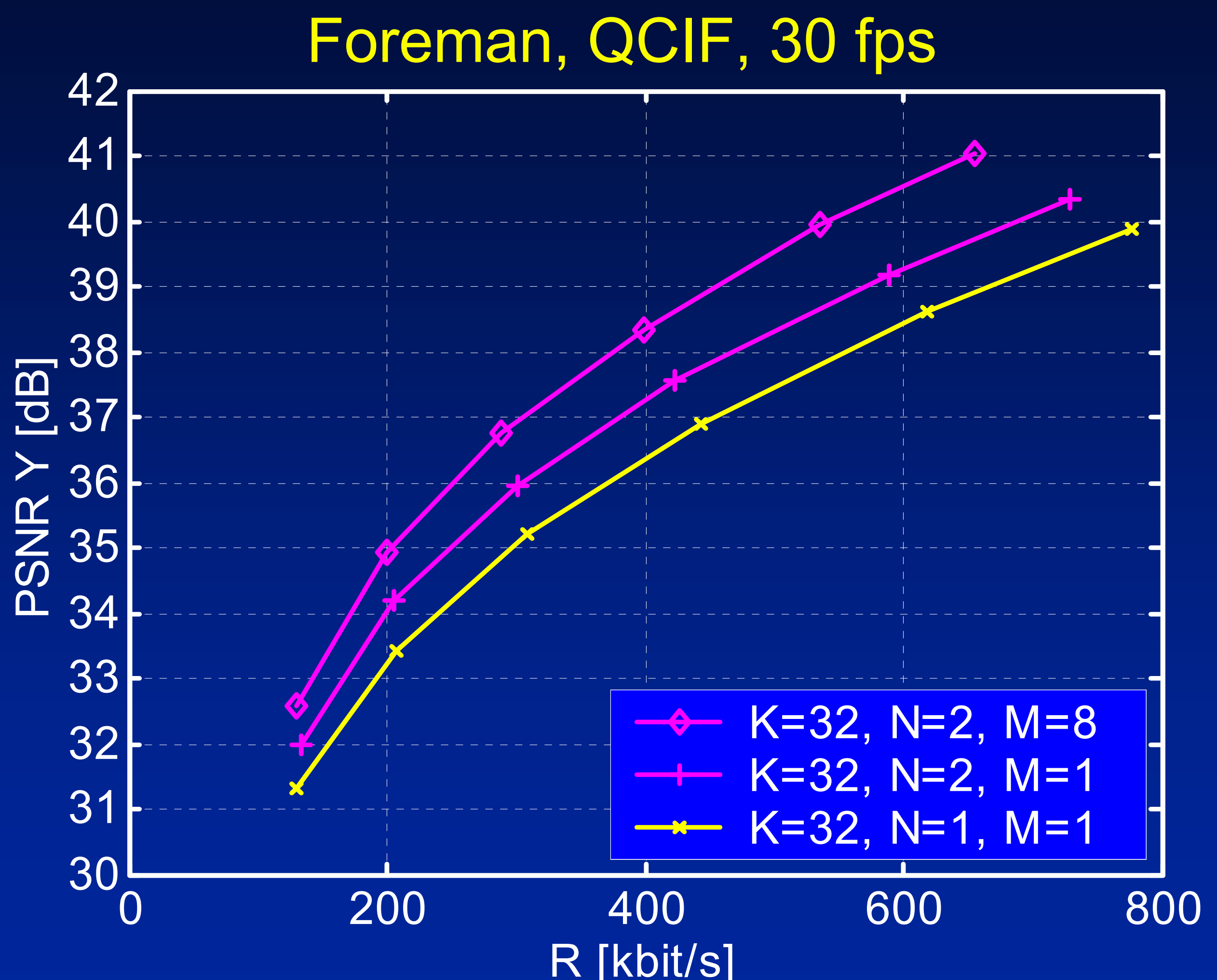


Update step uses negative motion vectors of corresponding prediction step

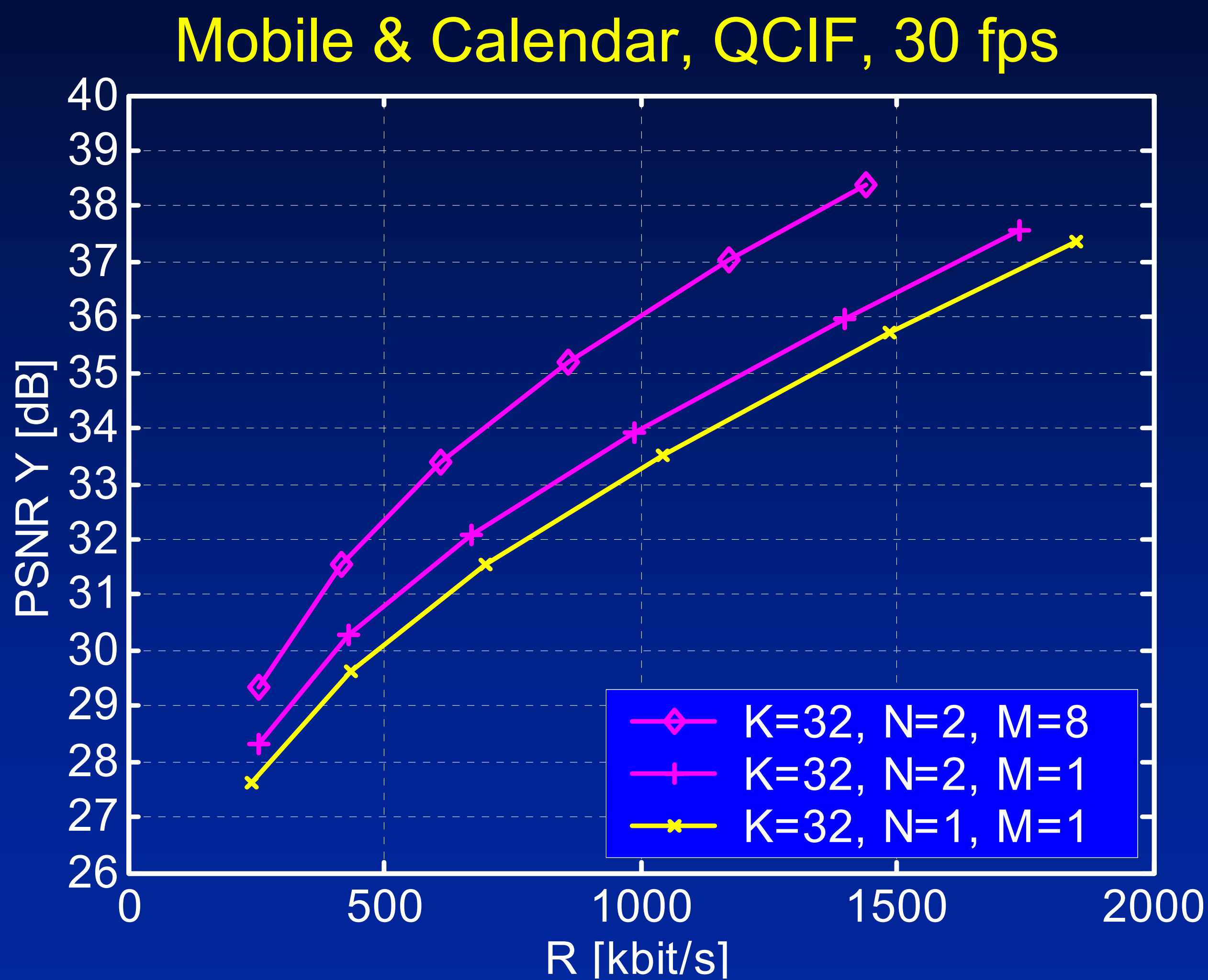
Coding Scheme

- Dyadic decomposition for each group of K pictures
- Motion-compensated Haar wavelet
- Block-adaptive selection between single (N=1) and complementary (N=2) signals
- 16x16 block motion compensation with half-pel accuracy
- Spatial coding with 8x8 DCT and run-length coding (H.263 compatible)
- Same quantizer step-size for all K intra-frame encoder

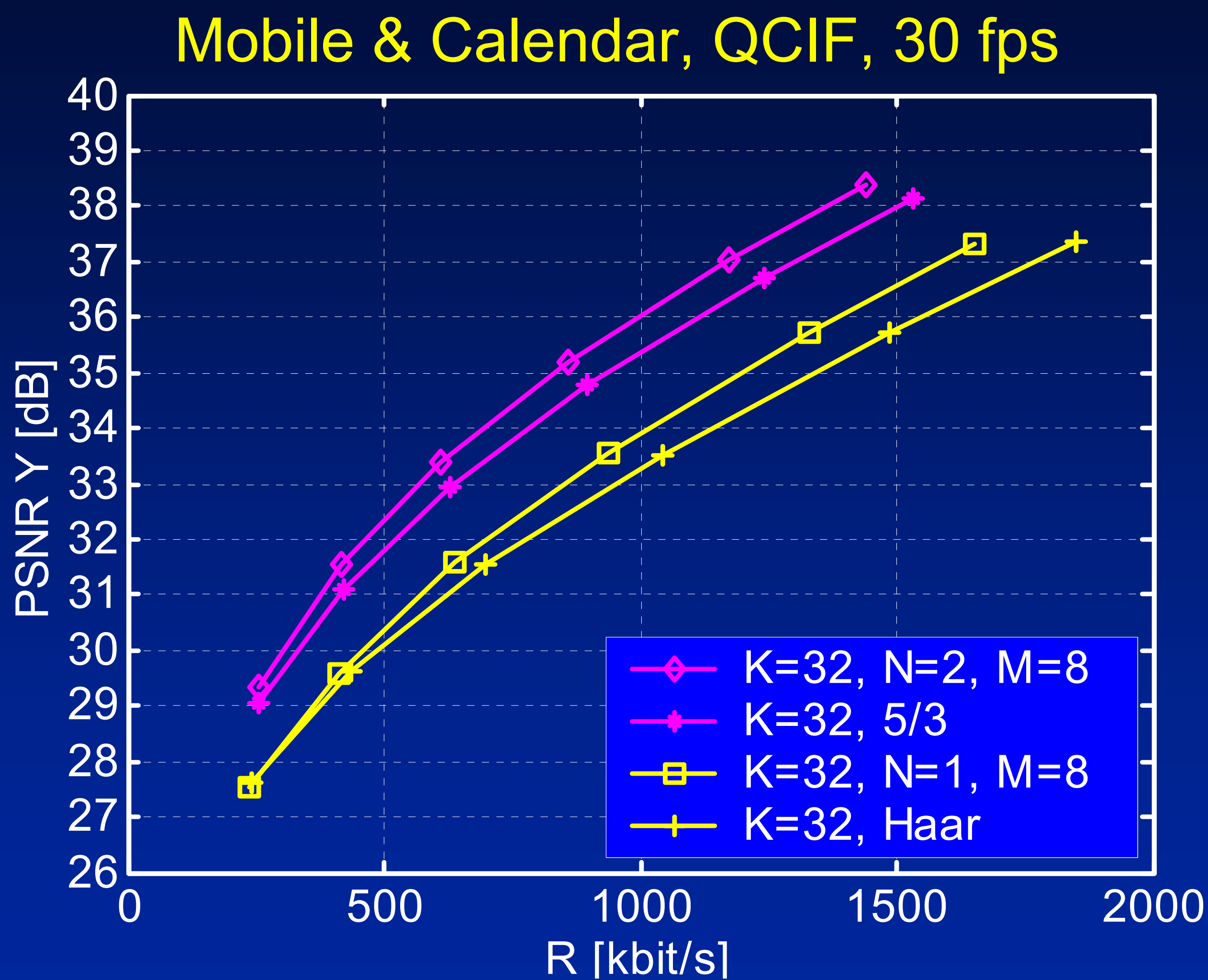
Haar Wavelet with N=2 Complementary Signals



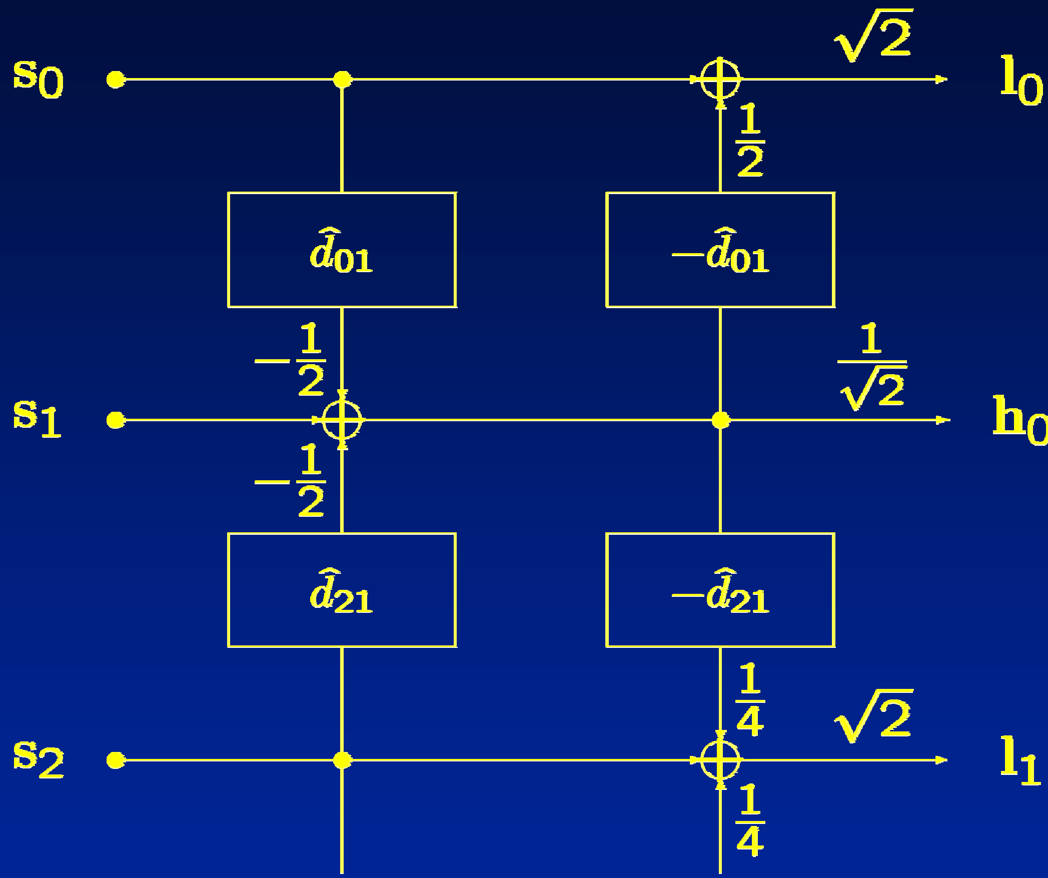
Haar Wavelet with N=2 Complementary Signals



Wavelets with Frame-Adaptive Motion Compensation



Motion-Compensated 5/3 Wavelet

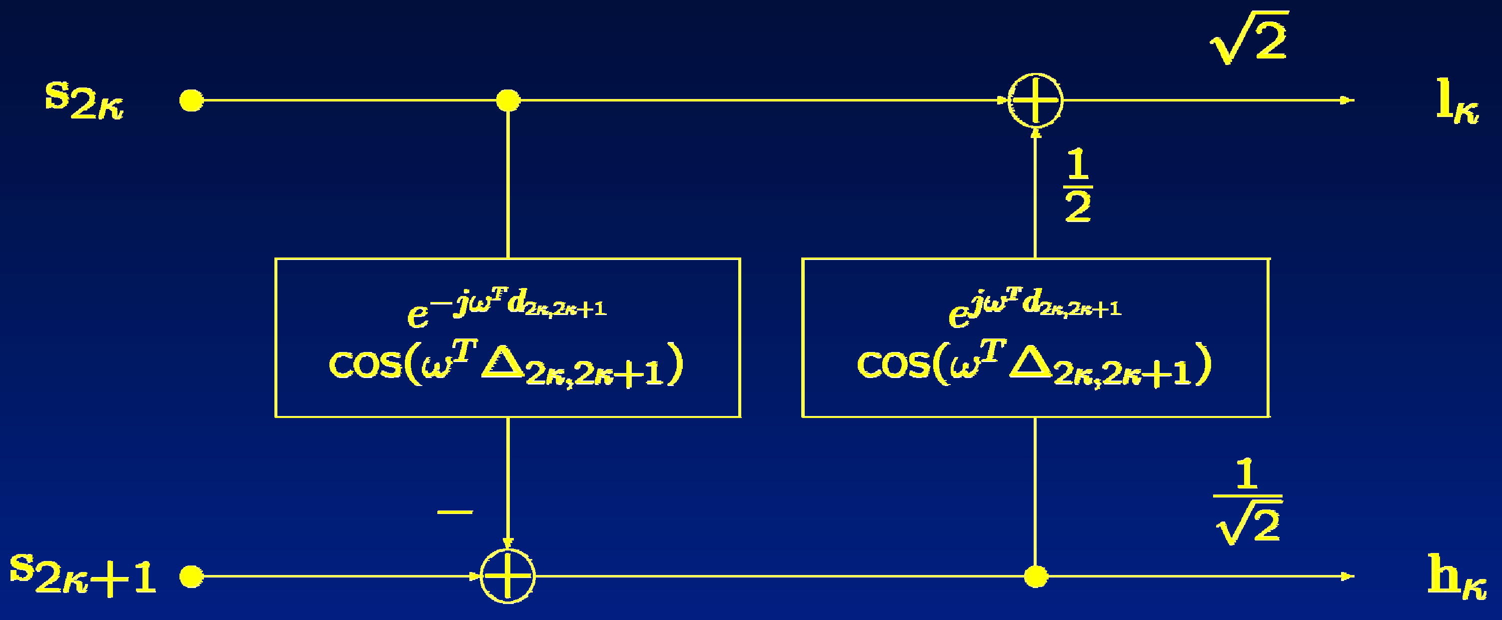


Update steps use negative motion vectors of corresponding prediction steps

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 $d_{\mu\nu}$
- Displacement operation is invertible

Haar Wavelet with N=2 Complementary Signals



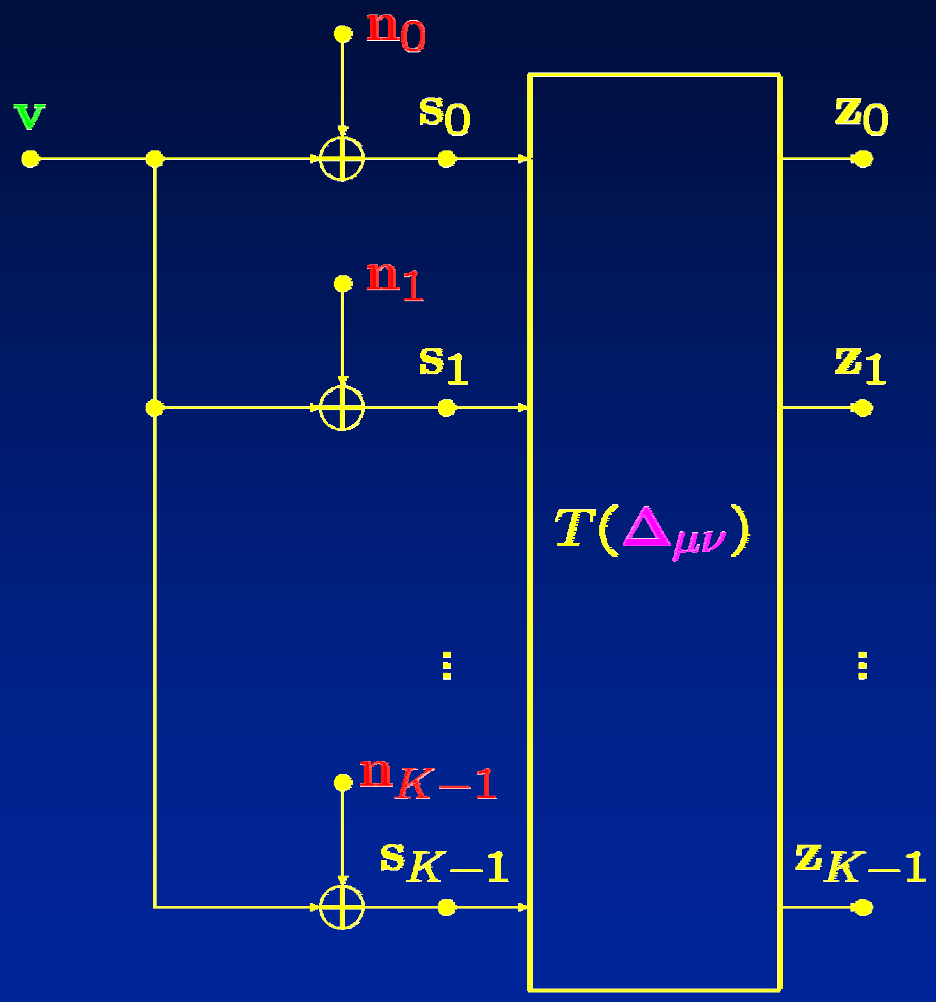
Displacements are complementary such that

$$\hat{d}^{(1)}_{2\kappa,2\kappa+1} = d_{2\kappa,2\kappa+1} - \Delta_{2\kappa,2\kappa+1}$$

$$\hat{d}^{(2)}_{2\kappa,2\kappa+1} = d_{2\kappa,2\kappa+1} + \Delta_{2\kappa,2\kappa+1}$$

Motion-compensated signals are averaged

Model for Coding with Complementary Signals



v	model picture
n_k	k-th noise signal
s_k	k-th input picture
z_k	k-th transform signal
$\Delta_{\mu\nu}$	displacement error between pictures μ and ν

Any input picture can be the reference picture

Performance Measure

- Rate difference for each picture k

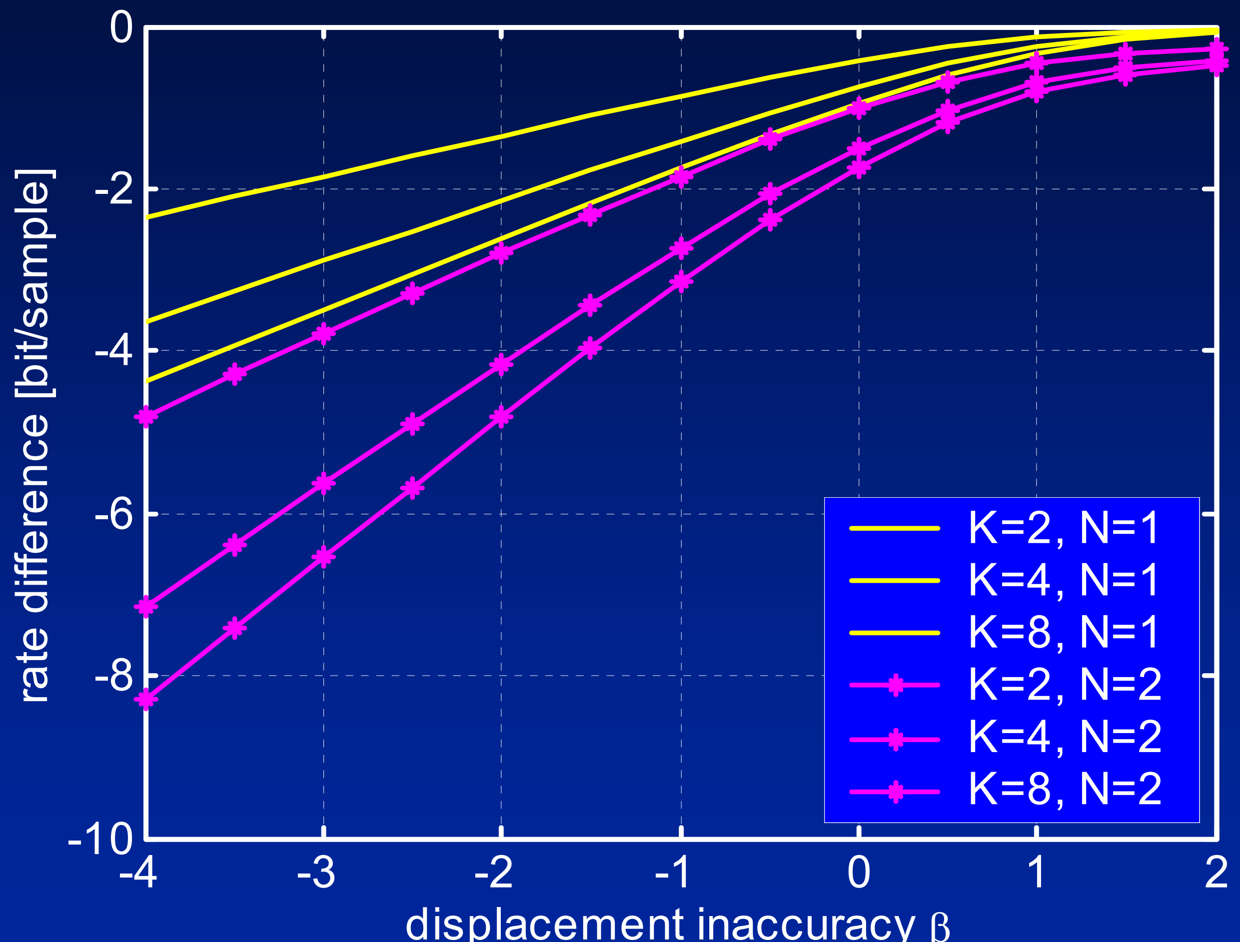
$$\Delta R_k = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{1}{2} \log_2 \left(\frac{\Phi_{\mathbf{z}_k \mathbf{z}_k}(\omega)}{\Phi_{\mathbf{s}_k \mathbf{s}_k}(\omega)} \right) d\omega$$

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

- Average rate difference

$$\Delta R = \frac{1}{K} \sum_{k=0}^{K-1} \Delta R_k$$

Rate Difference with Negligible Noise



Calibration:

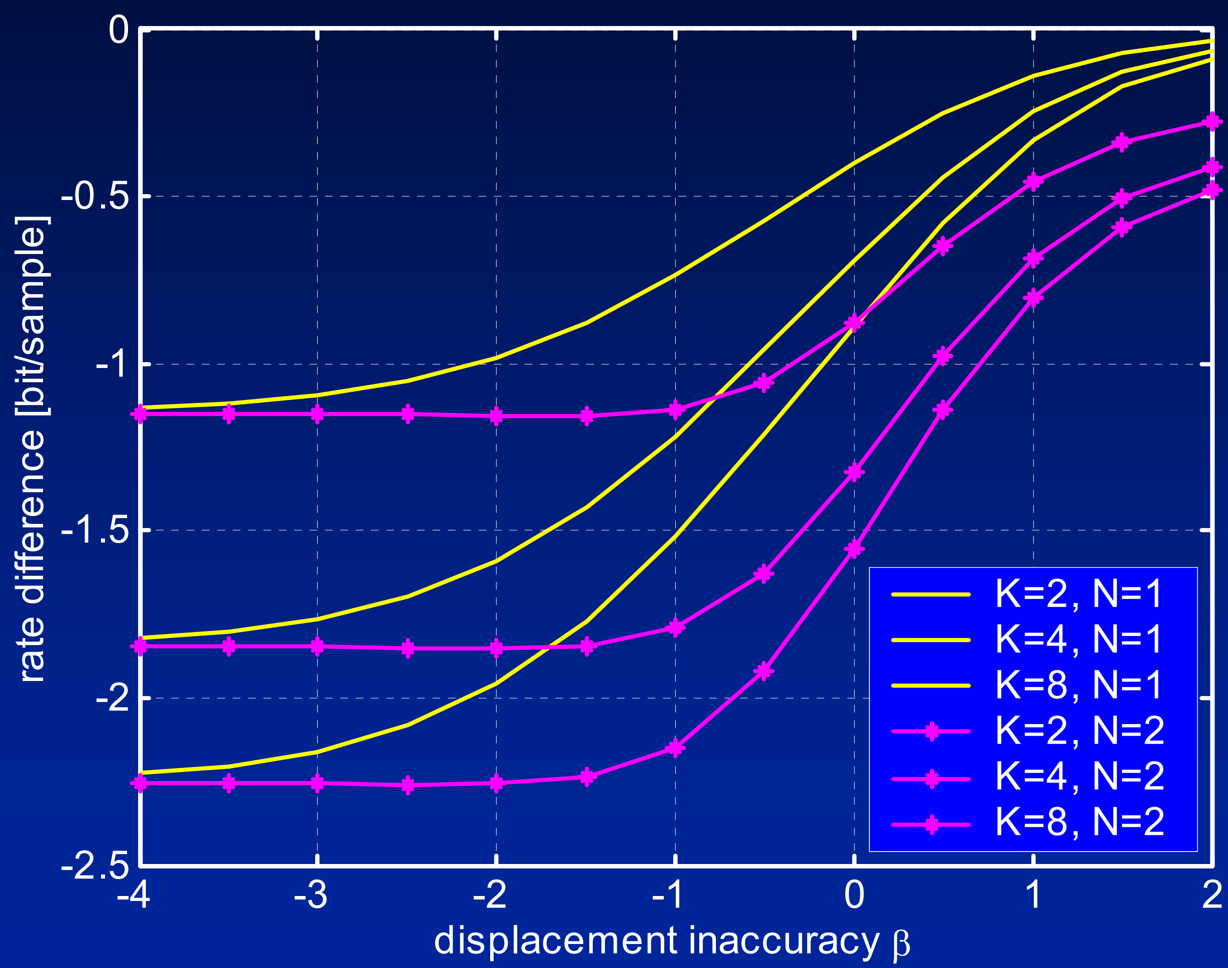
$\beta = 0.5 \log_2(12\, \sigma^2_{\Delta})$

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 $\beta=-2$

Important Observations

- For single motion-compensated signals in the lifting steps, bit-rate savings at high bit-rates are bounded by 1 bit per sample per displacement inaccuracy step when compared to optimum intra-frame coding of the input pictures

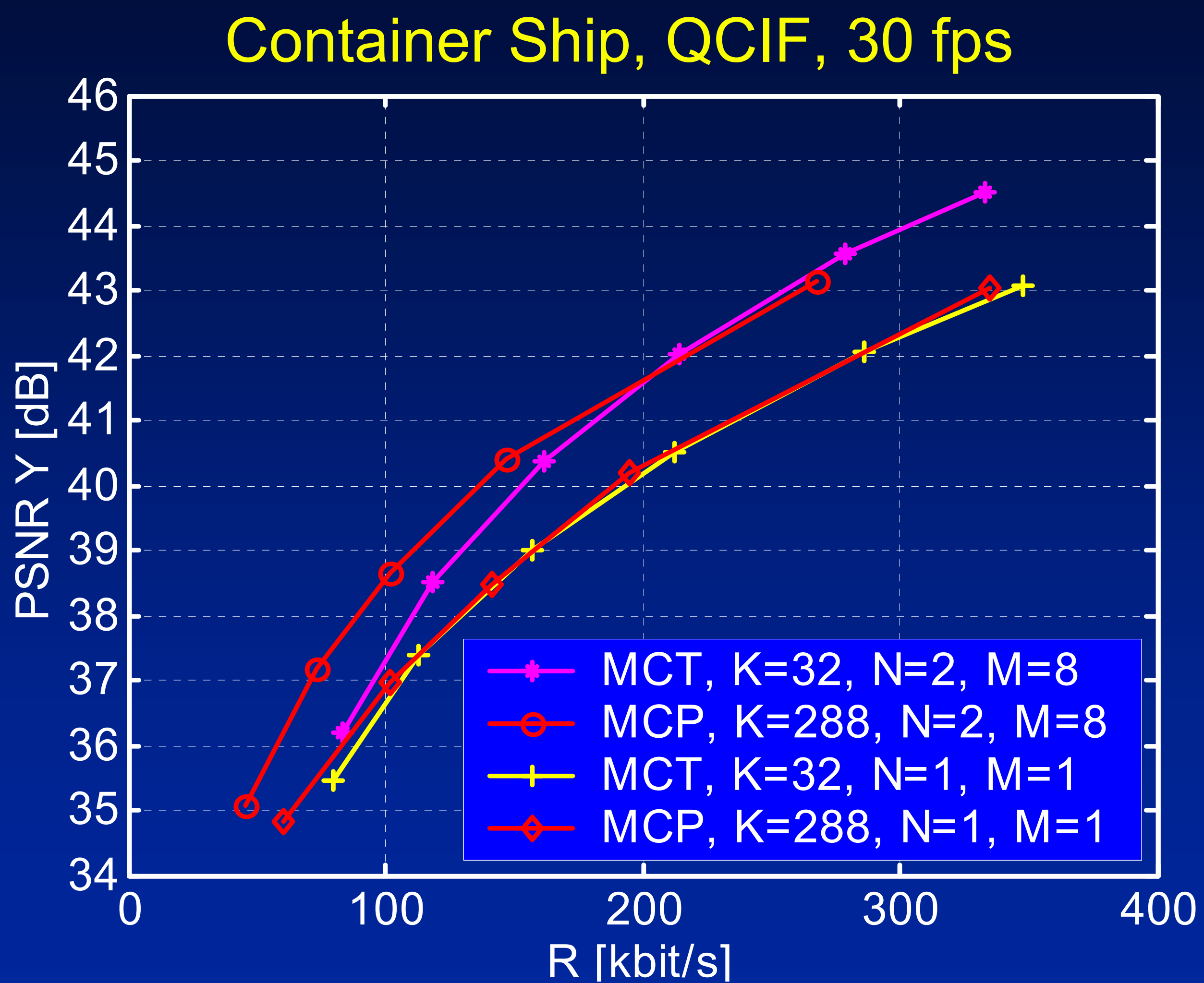
- Superimposed complementary signals in the lifting steps improve compression efficiency: For GOPs of size K, bit-rate savings at high bit-rates are bounded by $\frac{K-1}{K}2$ bits per sample per displacement inaccuracy step

- Residual noise limits the efficiency for very accurate motion compensation

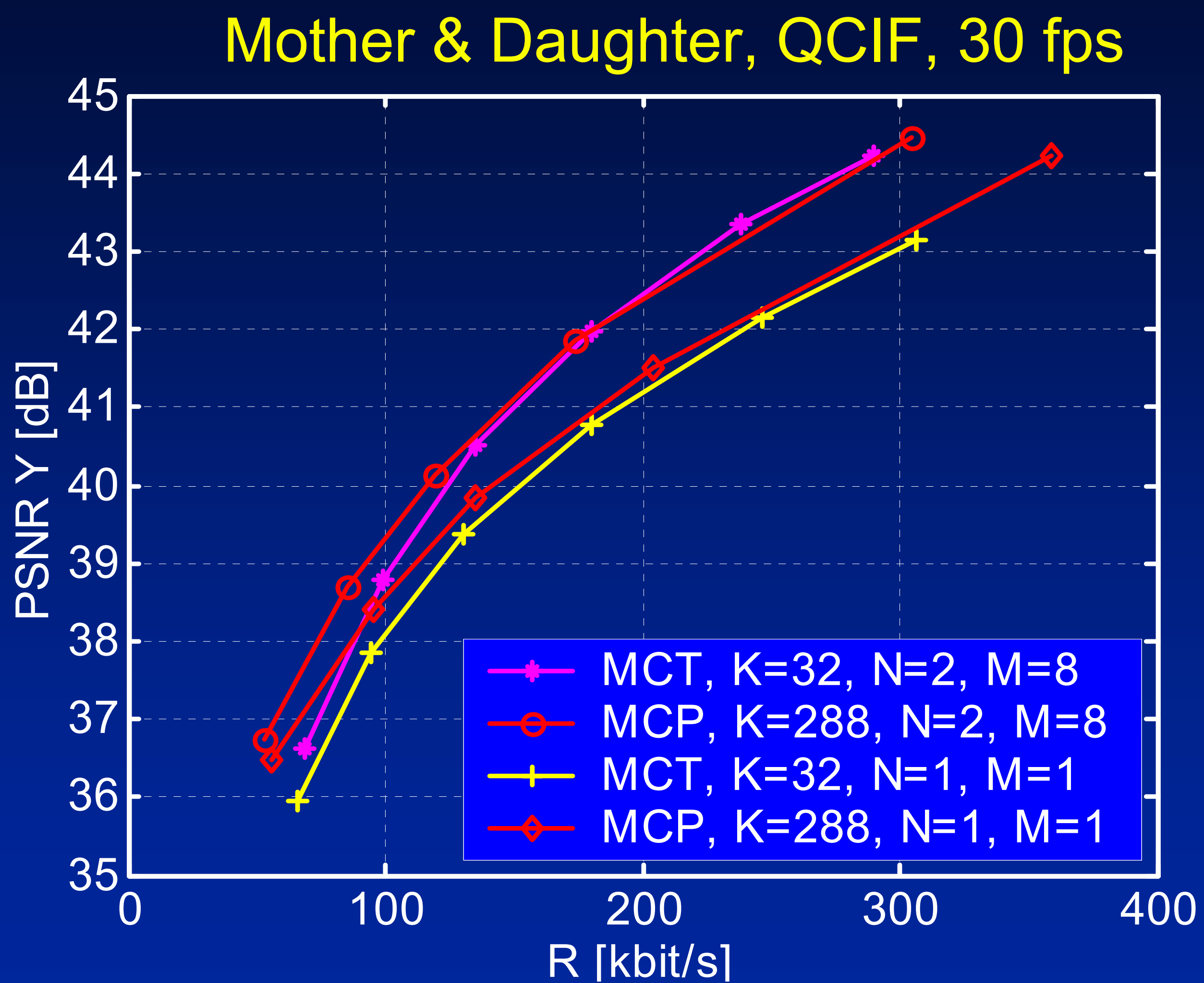
Comparison to Predictive Coding

- Predictive coding scheme:
  - *16x16 block motion compensation with half-pel accuracy*
  - *Single (N=1) or two complementary (N=2) motion-compensated signals*
  - *M previous reference frames*
  - *Spatial 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

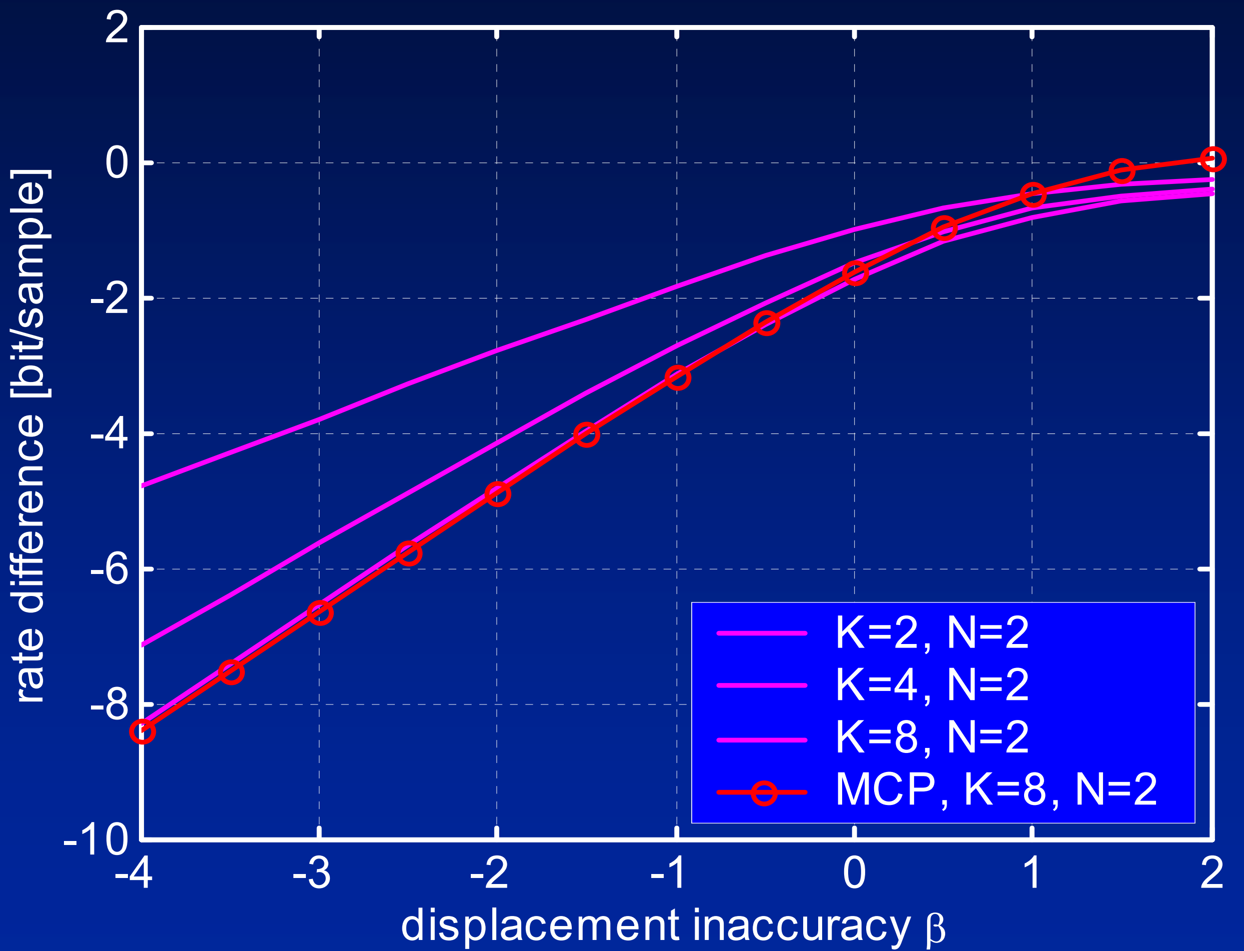
Comparison to Predictive Coding



Comparison to Predictive Coding



Comparison to Predictive Coding with GOP 8



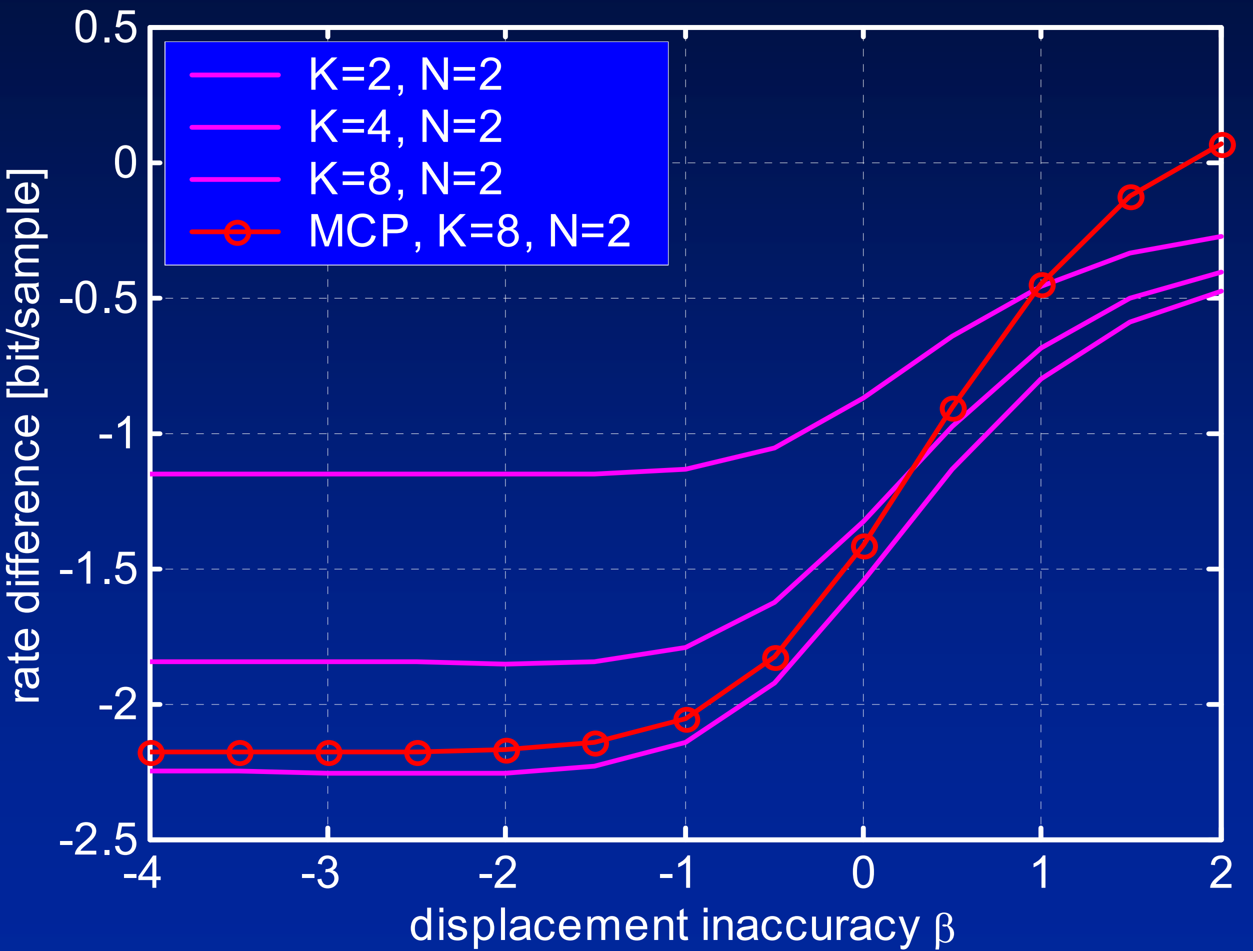
RNL = -100 dB

Calibration:

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

Integer-pel	β=0
Half-pel	β=-1
Quarter-pel	β=-2

Comparison to Predictive Coding with GOP 8



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

- For single motion-compensated signals, the rate difference is limited to 1 bit per sample per displacement inaccuracy step
- For N=2 complementary motion-compensated signals, the rate difference approaches the limit of 2 bits per sample and displacement inaccuracy step
- Complementary motion-compensated lifted wavelet transforms achieve the same bounds for compression efficiency as predictive coding with complementary signals and permit additionally efficient scalable representations