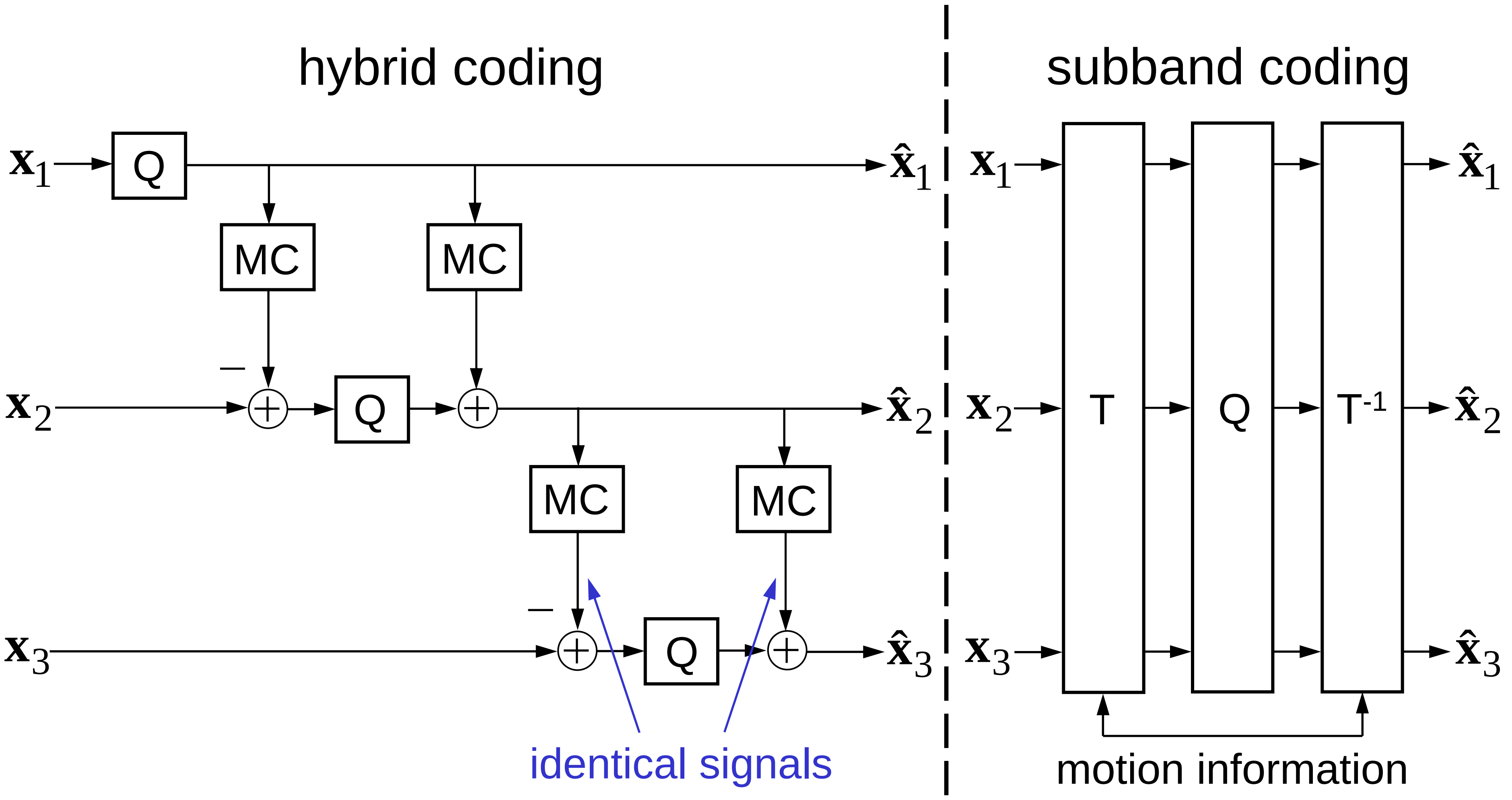
Subband Video Coding

- Hybrid video coding with motion-compensated prediction
- Temporal DPCM is challenging when aiming for
  - transmission over lossy channels
  - scalable video representations
- Alternative technique: Subband Video Coding

Architectural Comparison

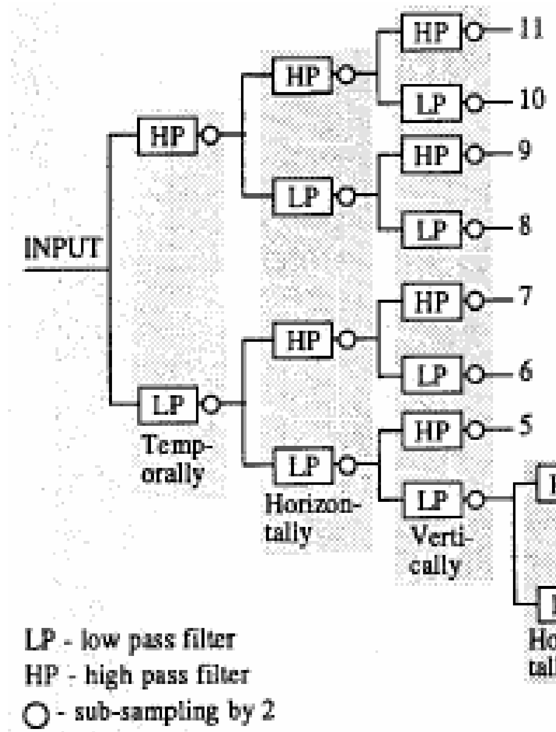


Outline

- Subband coding without motion compensation
- Motion-compensated subband coding
  - Motion-compensated pre- and post-processing
  - Motion-compensated temporal filtering
  - Motion-compensated lifted wavelets
- Model for motion-compensated subband coding
- Open problems

Subband Coding without Motion Compensation

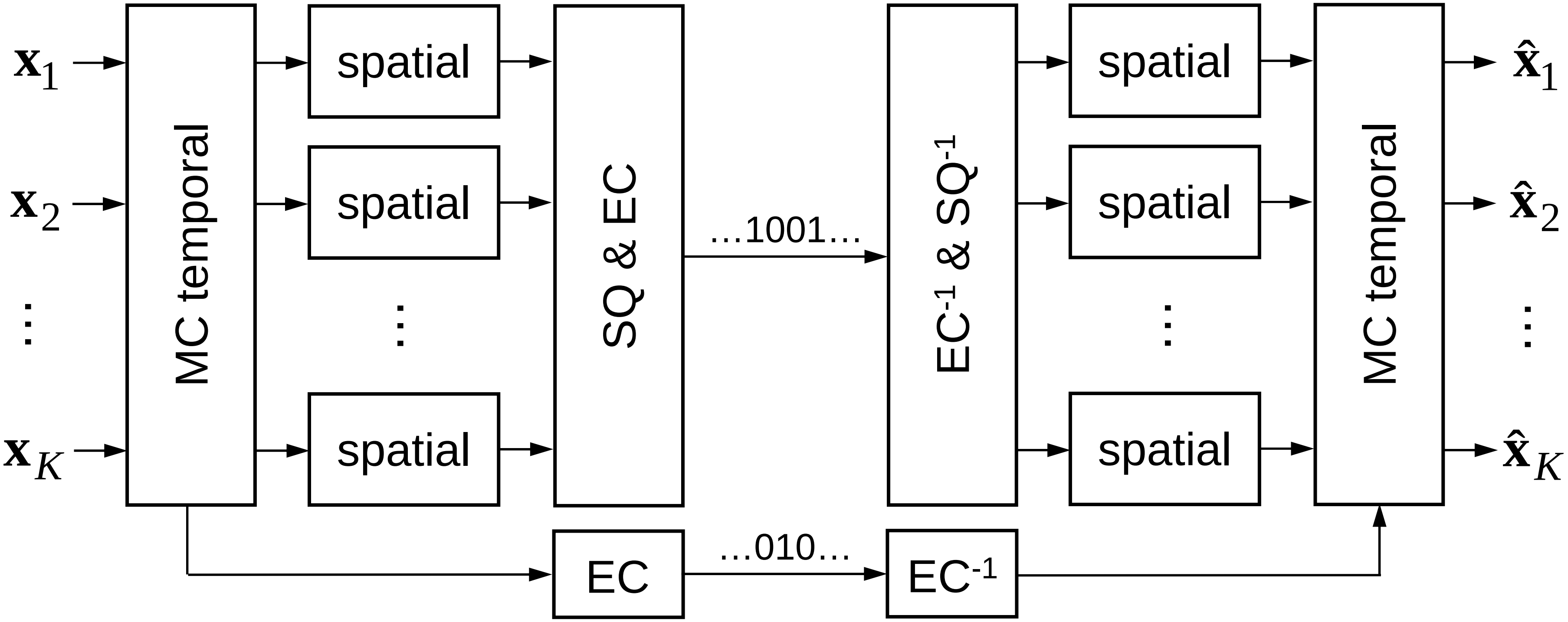
- 3D subband coding
  - Temporally
  - Horizontally
  - Vertically
- Perfect reconstruction
- Temporal filters:
  - Lowpass: $H(z) = \frac{1}{2}\left(1 + z^{-1}\right)$
  - Highpass: $H(z) = \frac{1}{2}\left(1 - z^{-1}\right)$



Band	Temp	Hor.	Vert.
1	LP	LP/LP	LP/LP
2	LP	LP/LP	LP/HP
3	LP	LP/HP	LP/LP
4	LP	LP/HP	LP/HP
5	LP	LP	HP
6	LP	HP	LP
7	LP	HP	HP
8	HP	LP	LP
9	HP	LP	HP
10	HP	HP	LP
11	HP	HP	HP

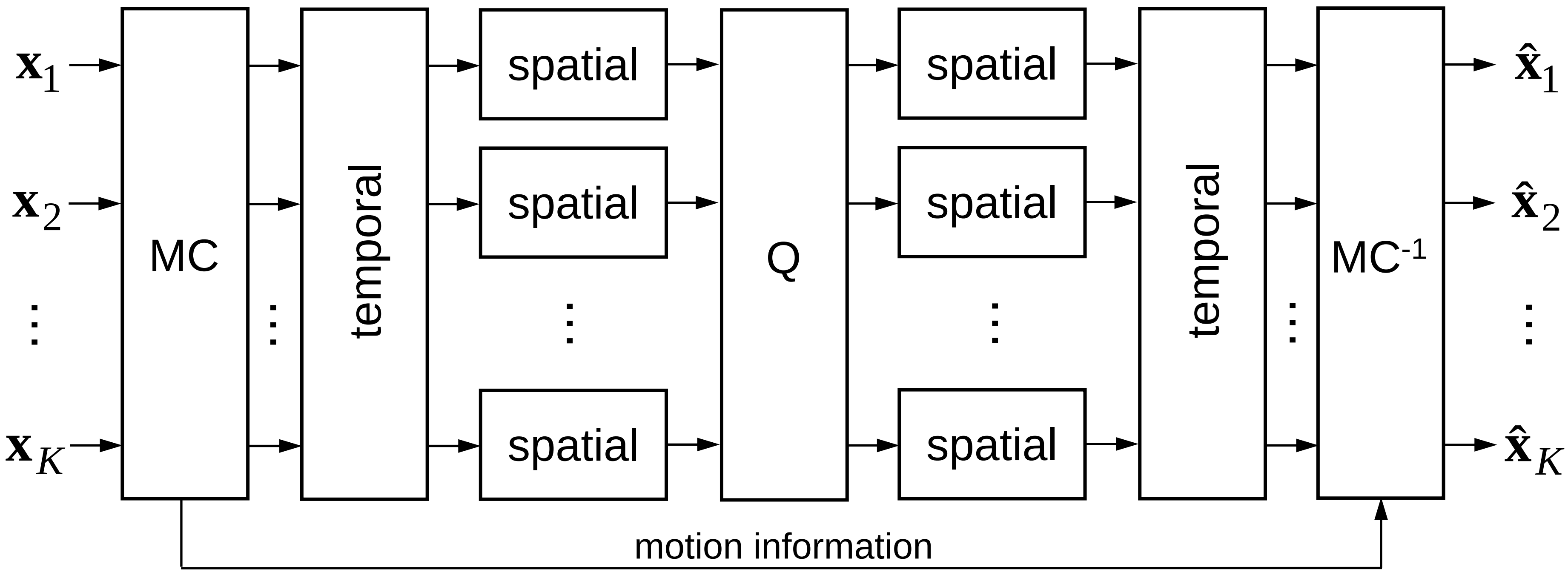
[Karlsson, Vetterli, 88]

Motion-Compensated Subband Coding



- Temporal subband decomposition is motion dependent
- Temporal subbands are spatially decomposed and coded

Motion-Compensated Pre- and Post-Processing



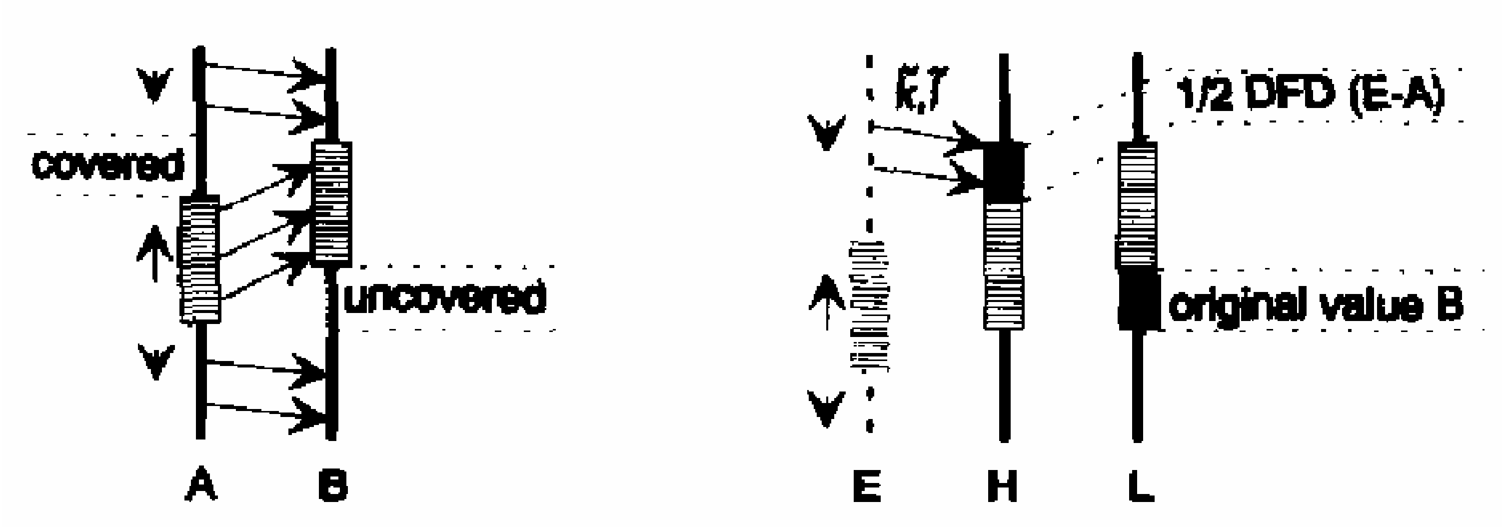
- Pre-processing compensates motion w.r.t. reference image
- **Problem:** Inversion of motion compensation operation

  "Imagine that the image is painted on a rubber sheet, then a good compensation algorithm should distort the rubber but not fold it"

[Kronander, 89]

Motion-Compensated Temporal Filtering

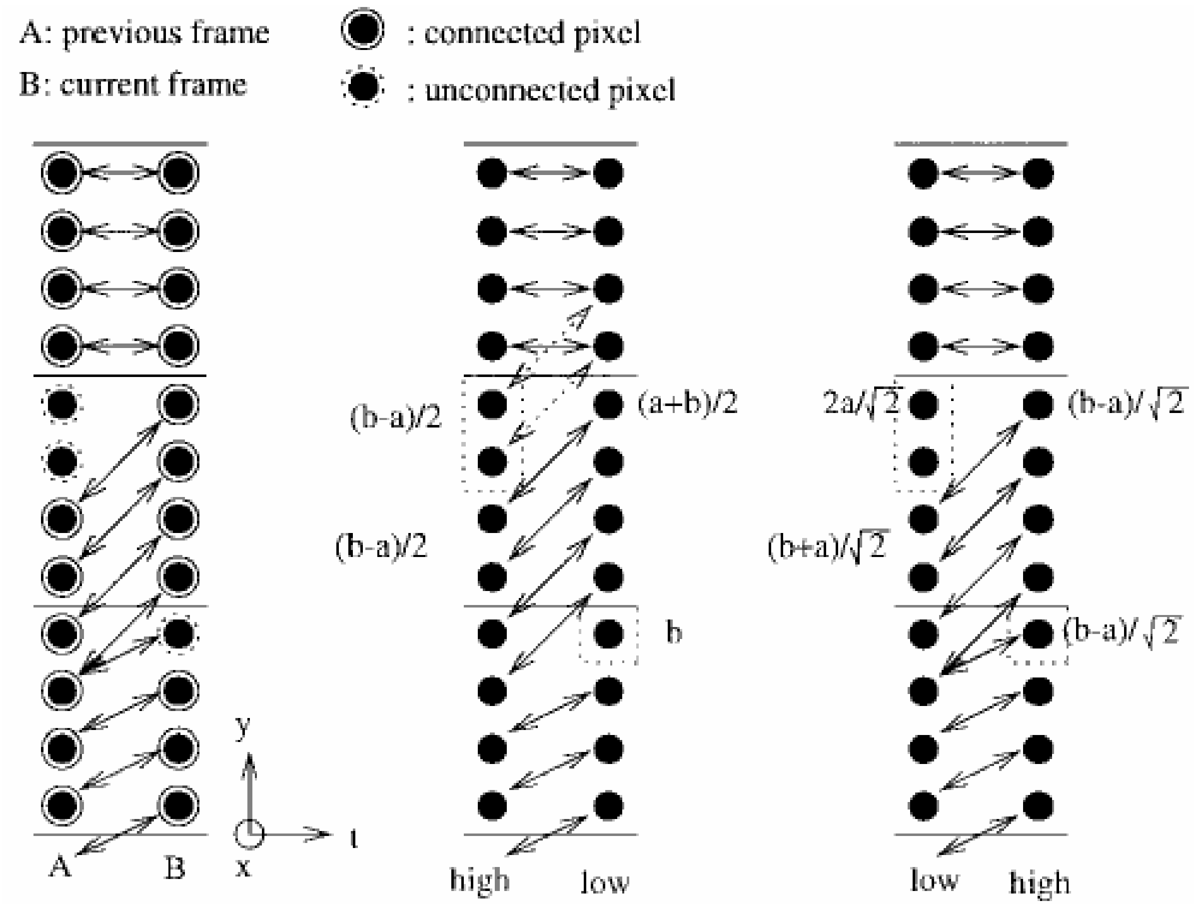
- Problem of inhomogeneous motion vector fields
- Temporal subband decomposition with Haar wavelet and MC that distinguishes between:
  - connected pixels
  - covered pixels
  - uncovered pixels



[Ohm, 94]

Motion-Compensated Temporal Filtering

- Alternative method for unconnected pixels:

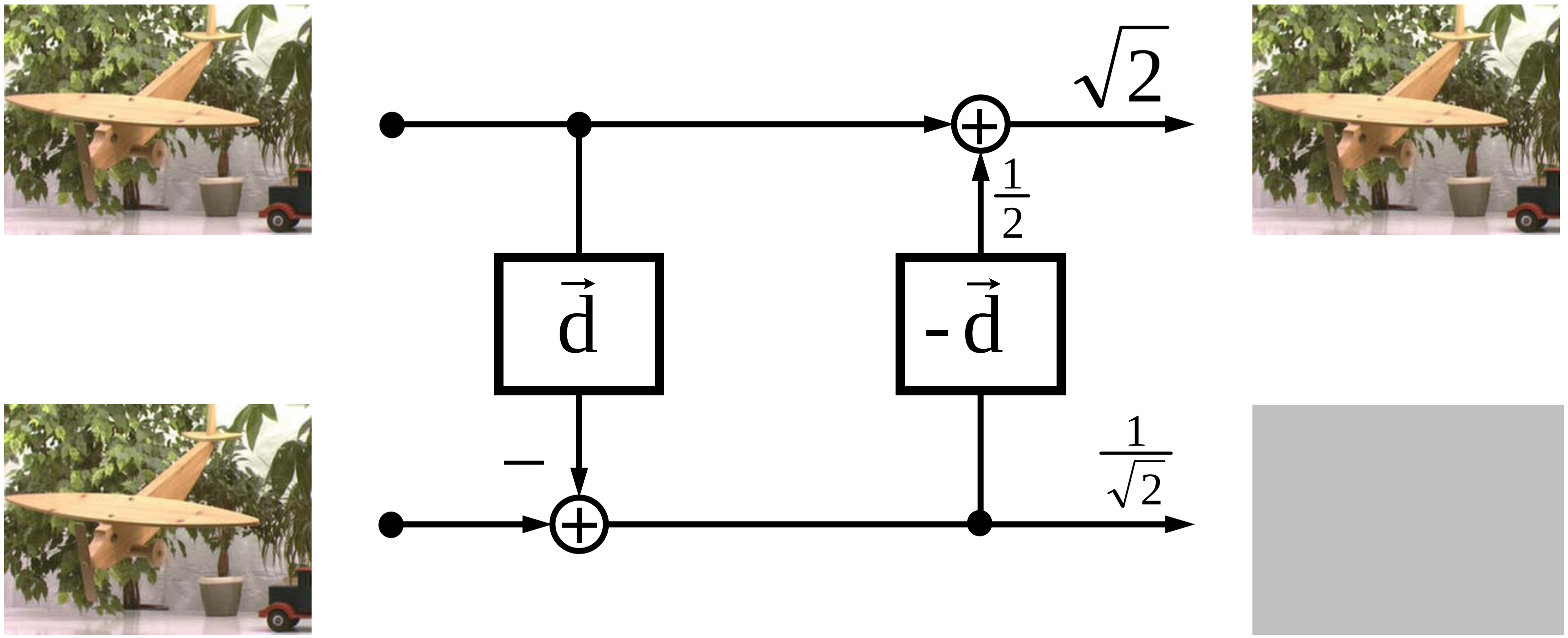


[Ohm, 94]

[Choi, Woods, 99]

Motion-Compensated Lifted Wavelets

- **Example:** Motion-compensated lifted Haar wavelet



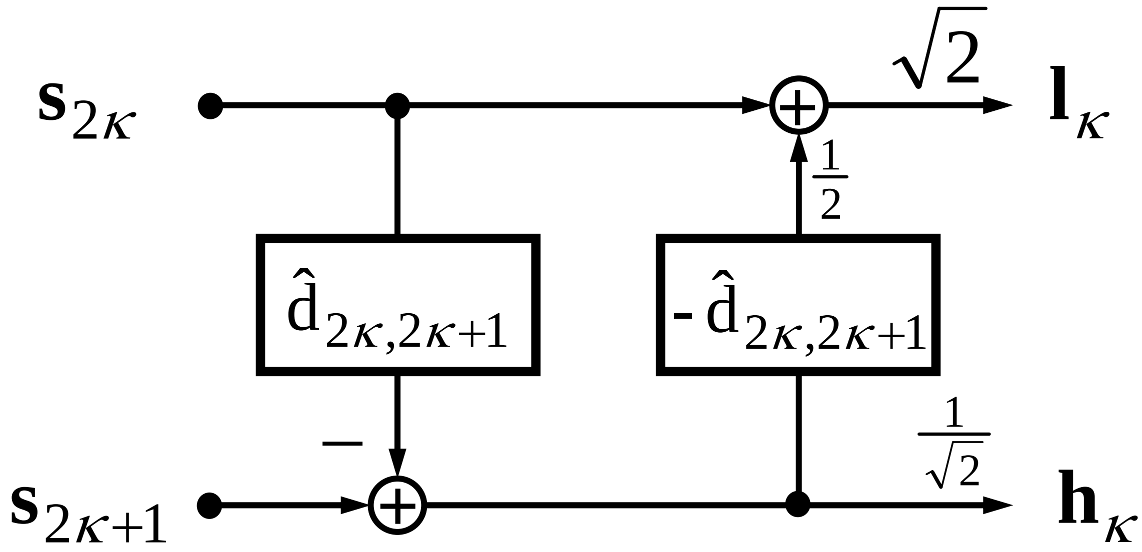
- Perfect reconstruction due to reversible lifting structure
- Strictly invertible motion fields not necessary for perfect reconstruction

[Pesquet-Popescu, Bottreau, 01]

Model for Motion-Compensated Subband Coding

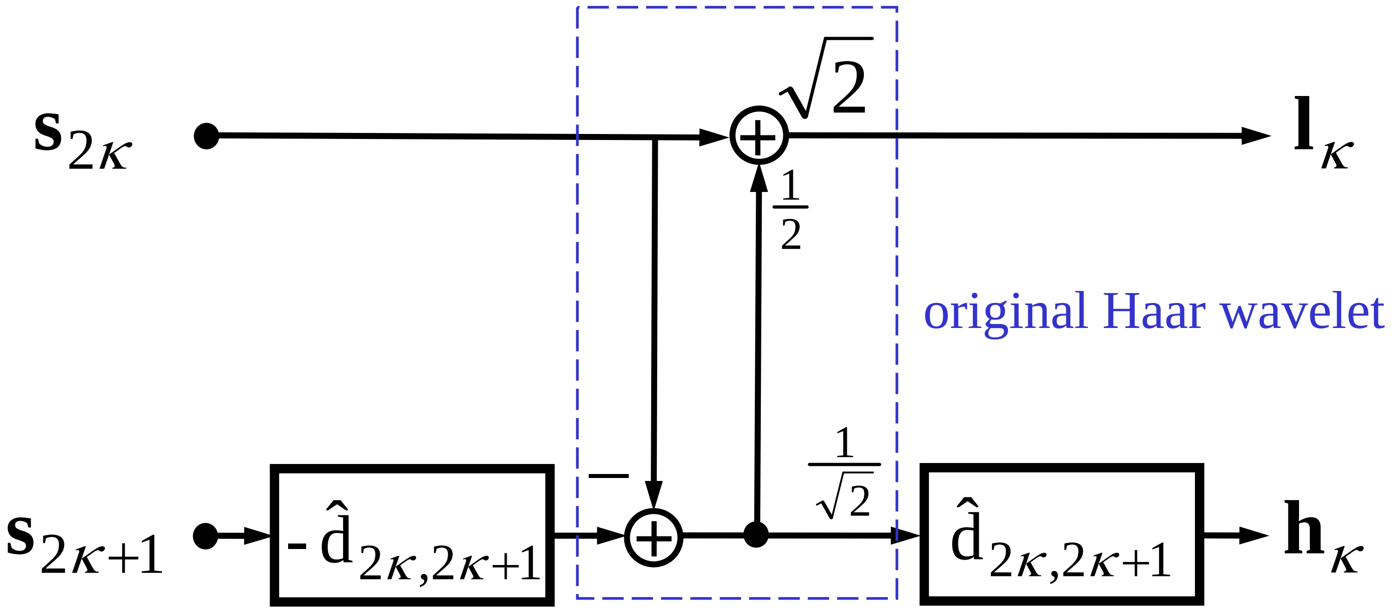
- Consider K temporally successive images
- Let $\mathbf{s}_k[x,y]$ be the k-th picture at pel-location x,y
- The signals are space-discrete and band-limited
- Ideal reconstruction is used for sub-pel accurate displacements
- Displacement operation is invertible

Motion-Compensated Lifted Haar Wavelet



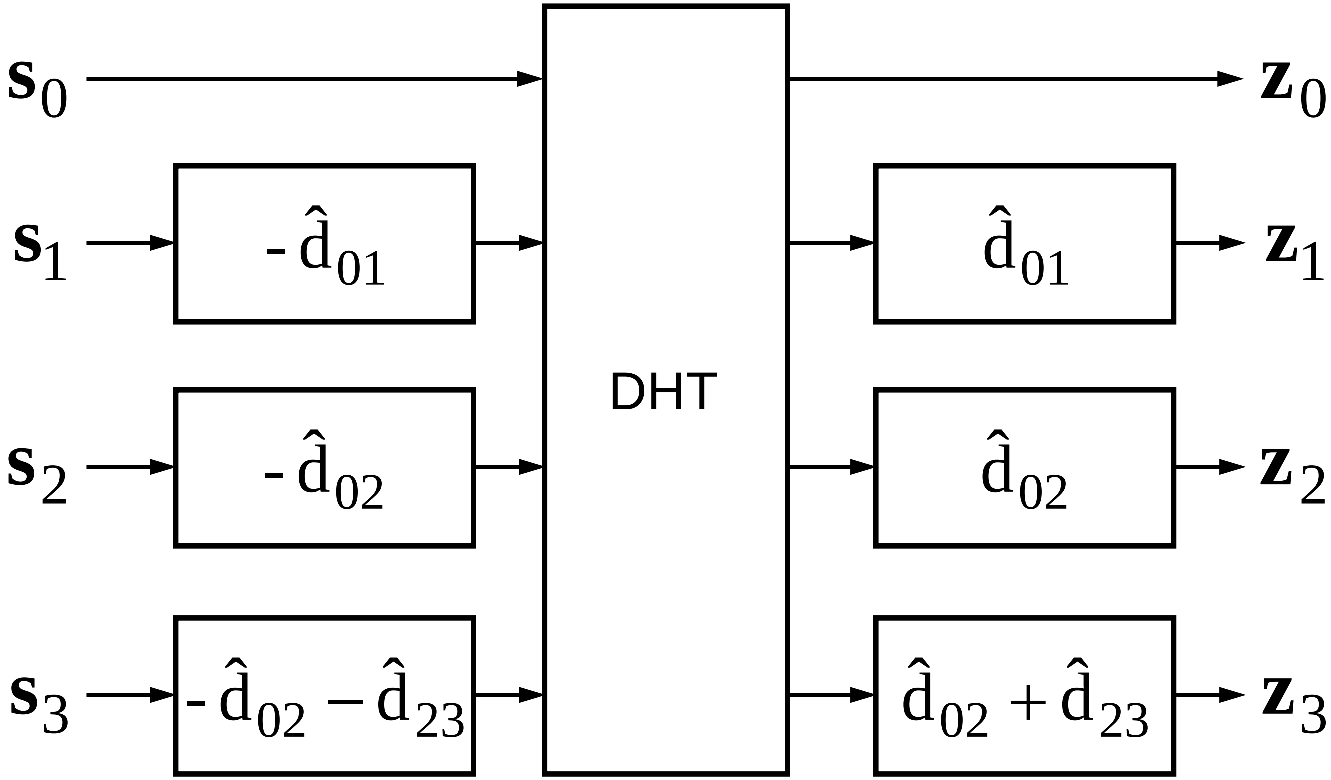
Update step uses negative motion vector of corresponding prediction step

Equivalent Motion-Compensated Wavelet



Invertible displacement operations are assumed

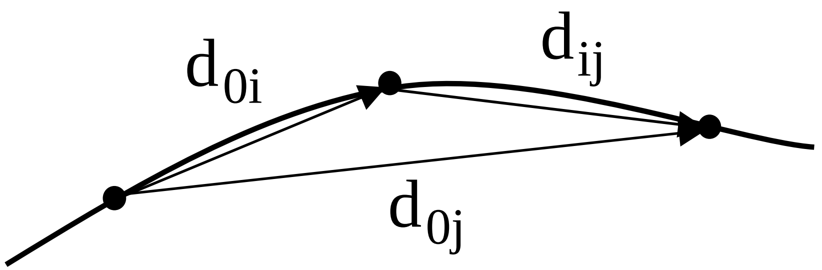
Dyadic Haar Transform



Both true displacements d and estimated displacements $\hat{d}$ are additive

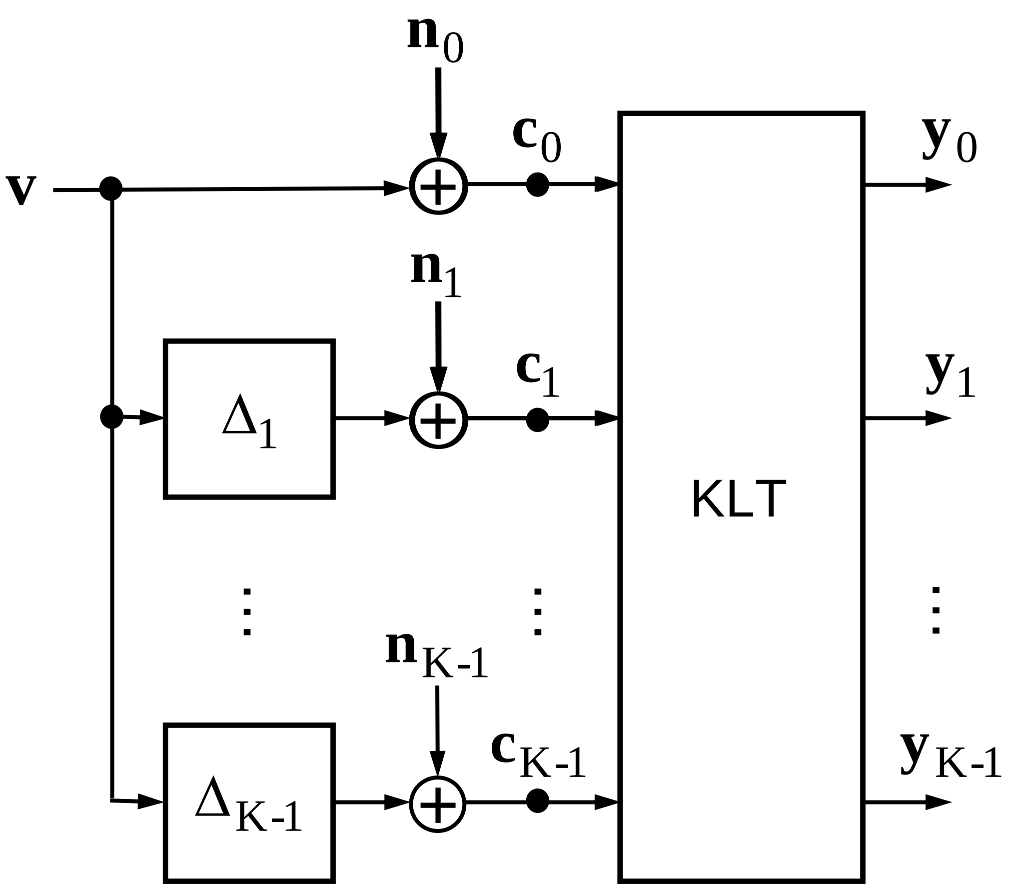
Additive Displacements

- Objects move on reversible trajectories



- True displacements: $$d_{0i} + d_{ij} = d_{0j}$$
- Estimated displacements: $$\hat{d}_{0i} + \hat{d}_{ij} = \hat{d}_{0j}$$
- Displacement error: $$d_{ij} = \hat{d}_{ij} + \Delta_{ij}$$

Generalized Signal Model



$\mathbf{v}$	model picture
Δ_k	*k*-th displacement error
$\mathbf{n}_k$	*k*-th noise signal
$\mathbf{c}_k$	*k*-th motion-compensated signal
$\mathbf{y}_k$	*k*-th transform signal

Any input picture can be reference picture

PSD Matrix of Motion-Compensated Pictures

- Cross spectral densities:

$$\Phi_{c_\mu c_\nu}(\omega) = E\left\{ e^{-j\omega^T(\Delta_\mu - \Delta_\nu)} \right\} \Phi_{vv}(\omega) + \Phi_{n_\mu n_\nu}(\omega)$$

- Absolute and relative displacement error: $\Delta_{0j} - \Delta_{0i} = \Delta_{ij}$
- Any picture can be reference, hence, absolute and relative displacement error variances are identical
- PSD Matrix:

$$\frac{\Phi_{cc}(\omega)}{\Phi_{vv}(\omega)} = \begin{bmatrix} 1+\alpha(\omega) & P(\omega) & \cdots & P(\omega) \\ P(\omega) & 1+\alpha(\omega) & \cdots & P(\omega) \\ \vdots & \vdots & \ddots & \vdots \\ P(\omega) & P(\omega) & \cdots & 1+\alpha(\omega) \end{bmatrix}$$

$$P(\omega) = E\left\{ e^{-j\omega^T \Delta_{\nu\mu}} \right\}$$

$$\alpha(\omega) = \frac{\Phi_{n_k n_k}(\omega)}{\Phi_{vv}(\omega)}$$

Rate-Distortion Bound via KLT

- KLT diagonalizes the PSD matrix:

$$\frac{\Phi_{yy}(\omega)}{\Phi_{vv}(\omega)} = \begin{bmatrix} \lambda_1(\omega) & 0 & \cdots & 0 \\ 0 & \lambda_2(\omega) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_K(\omega) \end{bmatrix}$$

- Eigenvalues:

$$\lambda_1(\omega) = 1 + \alpha(\omega) + (K-1)P(\omega)$$

$$\lambda_{2,3,\ldots,K}(\omega) = 1 + \alpha(\omega) - P(\omega)$$

- Eigenvectors: Orthonormal basis

Performance Measure

- Rate difference for each picture k

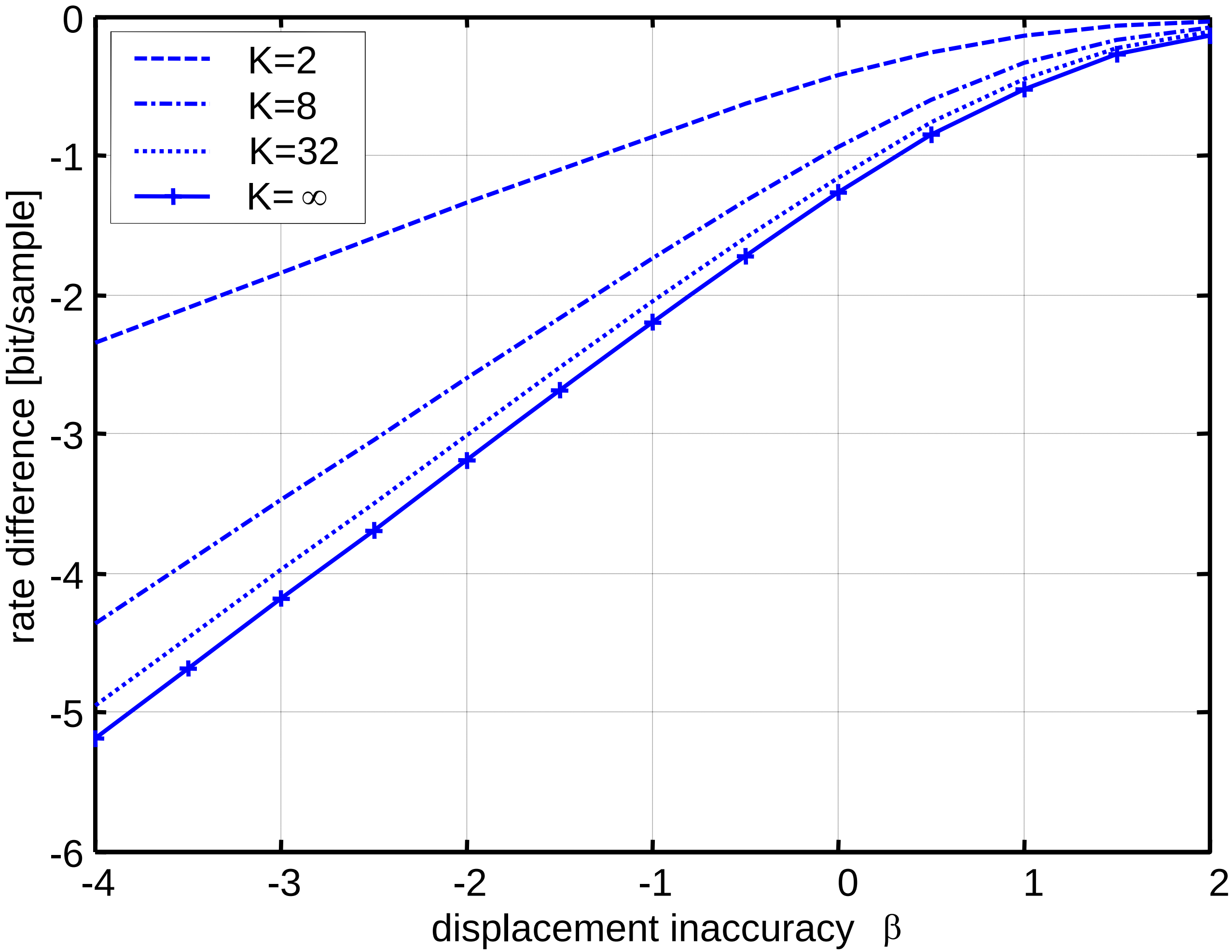
$$\Delta R_k = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{1}{2} \log_2 \left(\frac{\Phi_{\mathrm{y_k y_k}}(\omega)}{\Phi_{\mathrm{c_k c_k}}(\omega)} \right) d\omega$$

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

- Average rate difference

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

Rate Difference with Negligible Noise



Calibration:

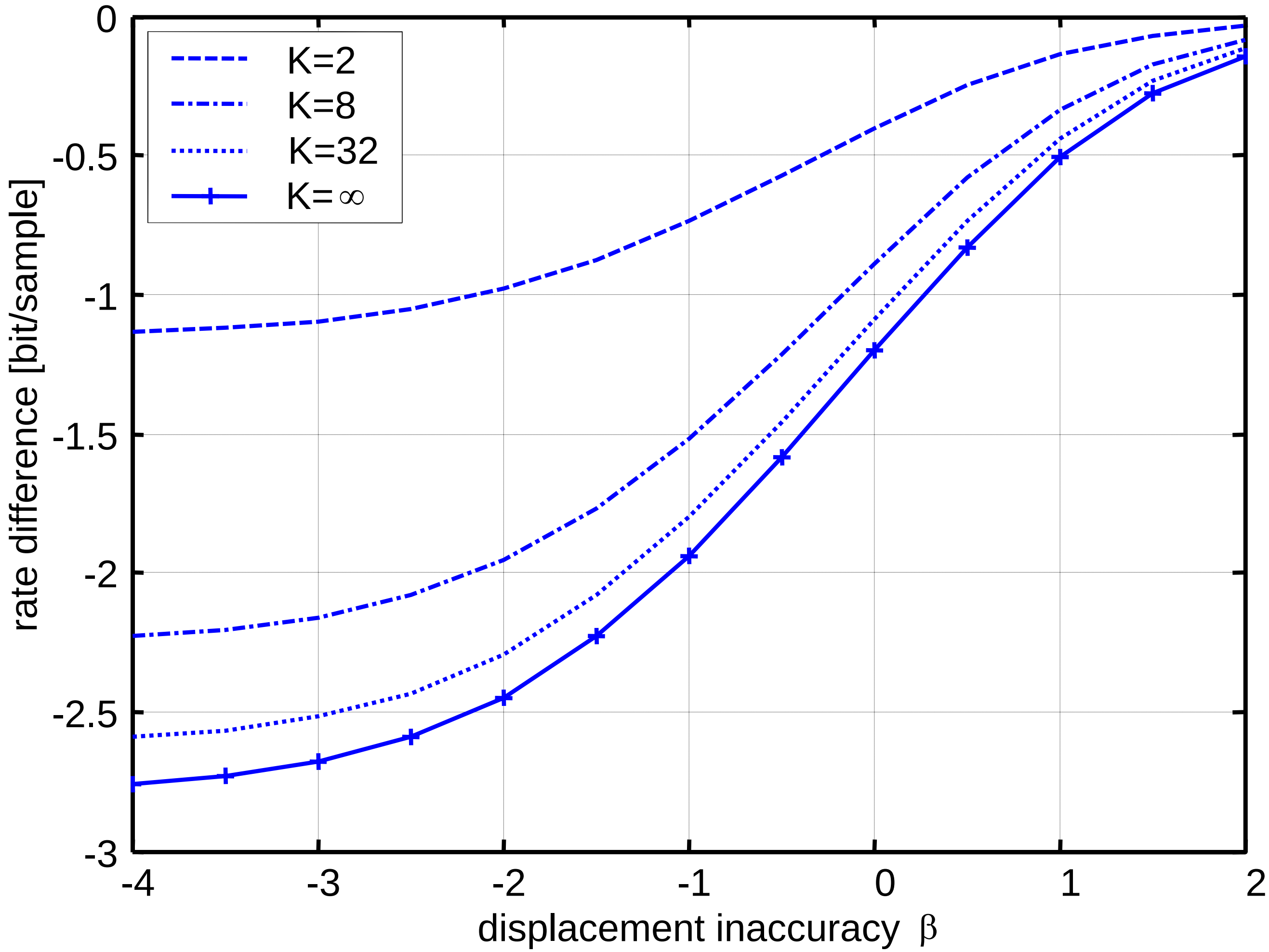
$\beta = 0.5 \log_2(12\, \sigma^2_\Delta)$

Integer-pel $\beta=0$

Half-pel $\beta=-1$

Quarter-pel $\beta=-2$

Rate Difference with RNL = -30 dB



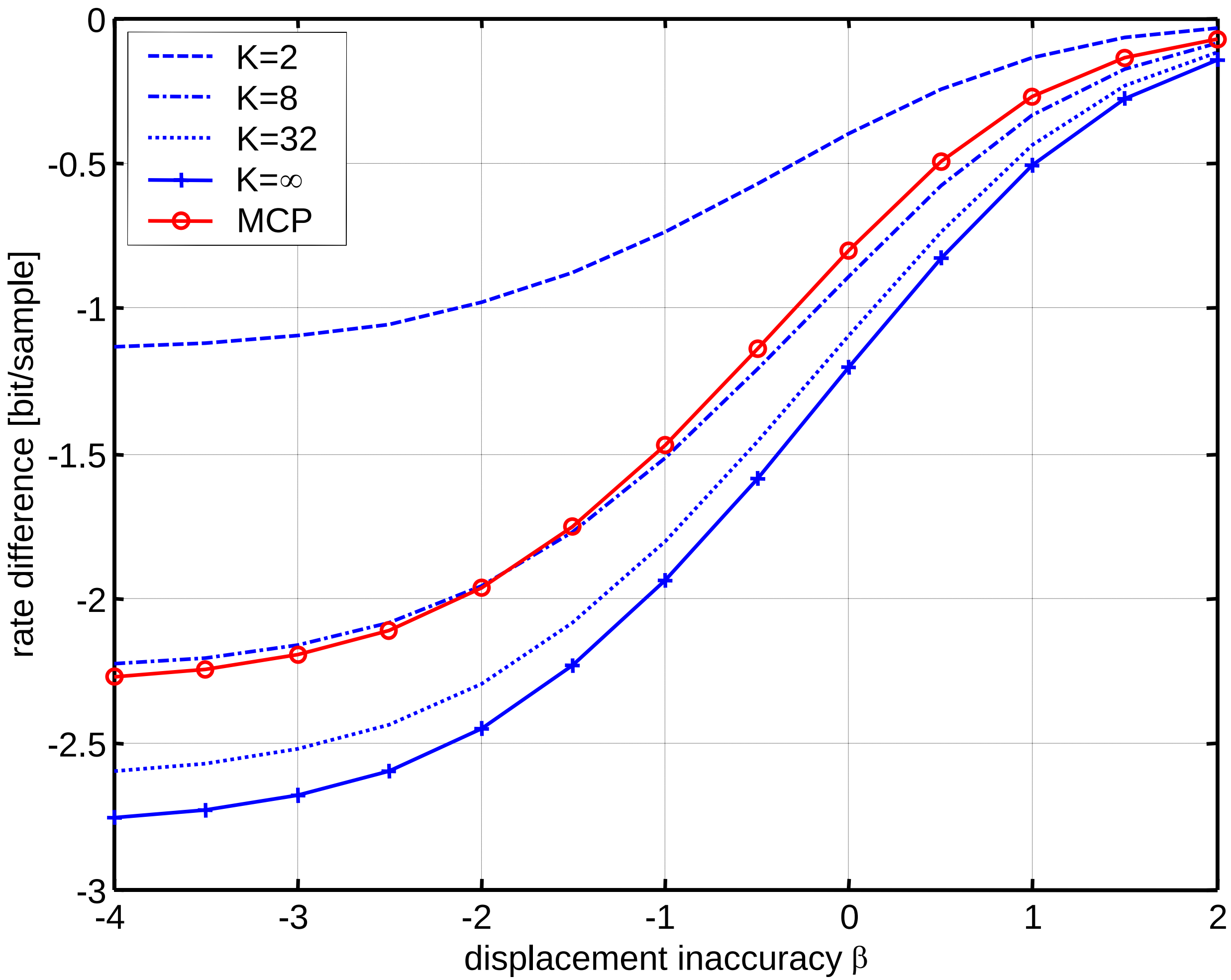
RNL = $10 \log_{10}(\sigma^2_{\mathbf{n}})$

Calibration:

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

Integer-pel	$\beta=0$
Half-pel	$\beta=-1$
Quarter-pel	$\beta=-2$

Comparison to MC Predictive Coding



RNL = -30 dB

Calibration:

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

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

Summary of Observations

- 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 3D subband coding outperforms predictive coding by at most 0.5 bits per sample

Open Problems

- Motion-compensated lifted wavelets are reversible but properties of subband decomposition are motion dependent
- For example, MC lifted Haar wavelet loses property of orthonormality for general multi-connecting motion fields
- In general, block motion fields are not invertible and burden update steps of MC lifted wavelets
- The model suggests a motion-adaptive KLT, which is orthonormal for any motion field