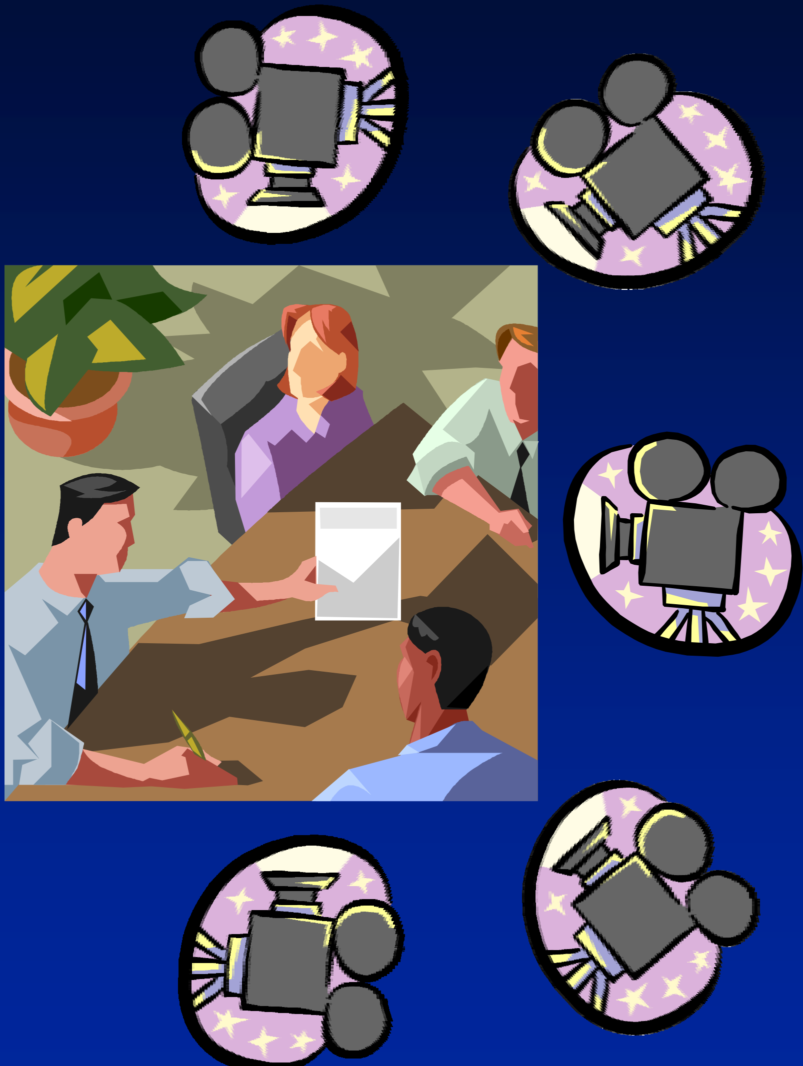
Workshop on Multimedia Signal Processing 2004

Distributed Coding of Dynamic Scenes with Motion-Compensated Wavelets

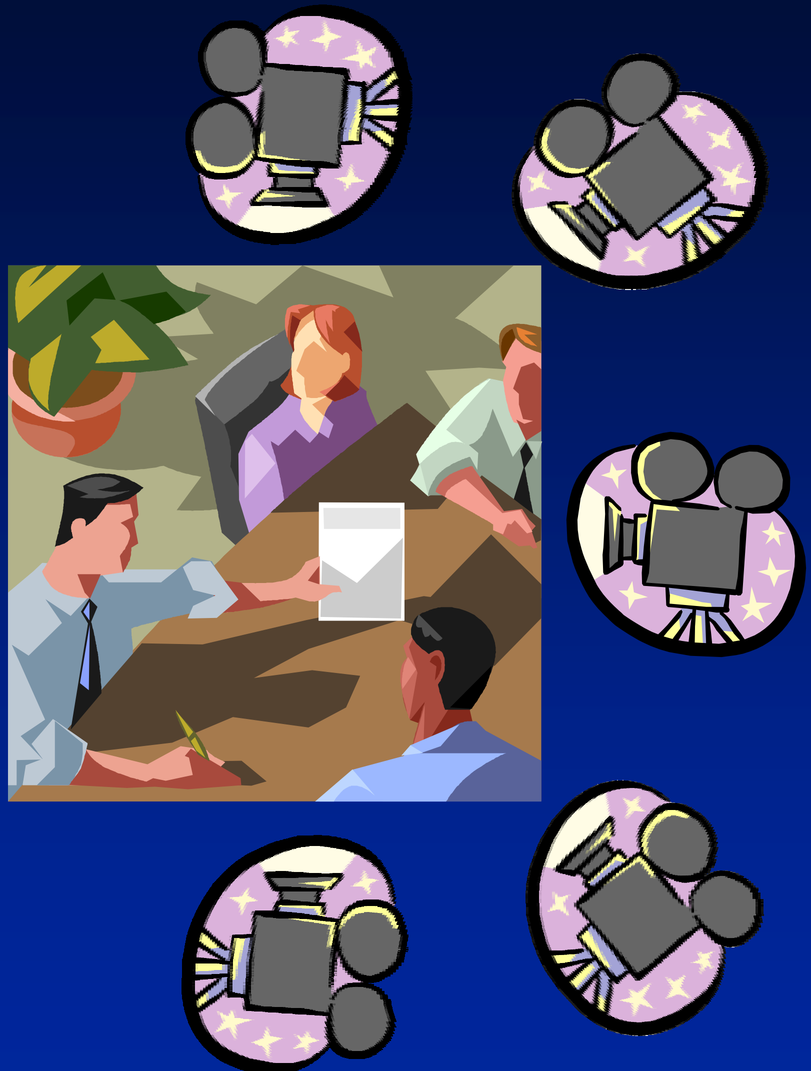
Markus Flierl and Pierre Vandergheynst
Signal Processing Institute
Swiss Federal Institute of Technology, Lausanne

Coding Scenario I



- 3-dimensional scene that evolves in time
- Observed by multiple video cameras located at different positions
- Each camera signal is coded locally
- The cameras are connected directly to the network
- One remote decoder is able to reconstruct arbitrary views

Coding Scenario II

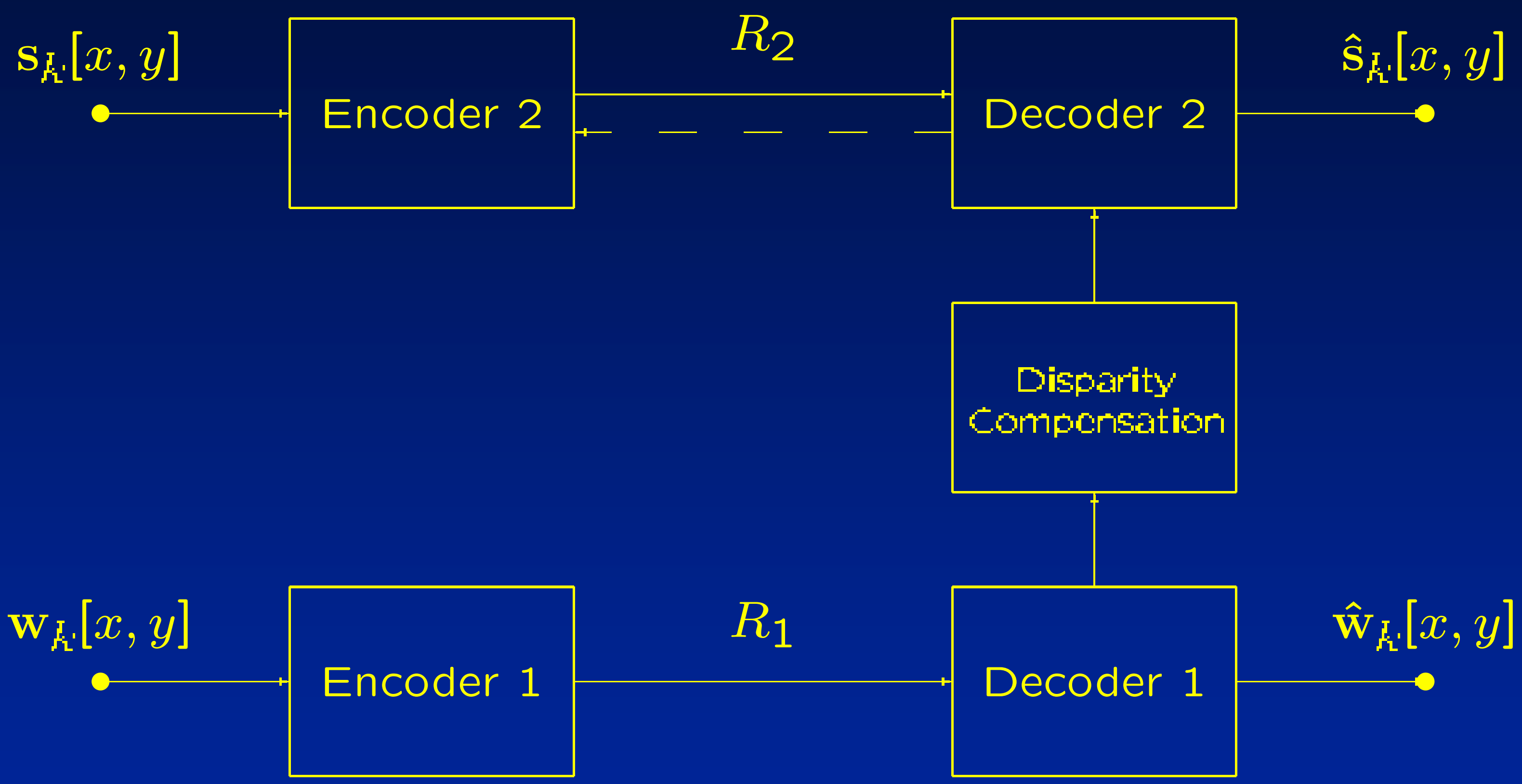


- How to use the information of neighboring cameras to improve the efficiency of the current video encoder?
- Obviously, side information can be used at encoder and decoder
- But what if the encoder do not communicate directly?

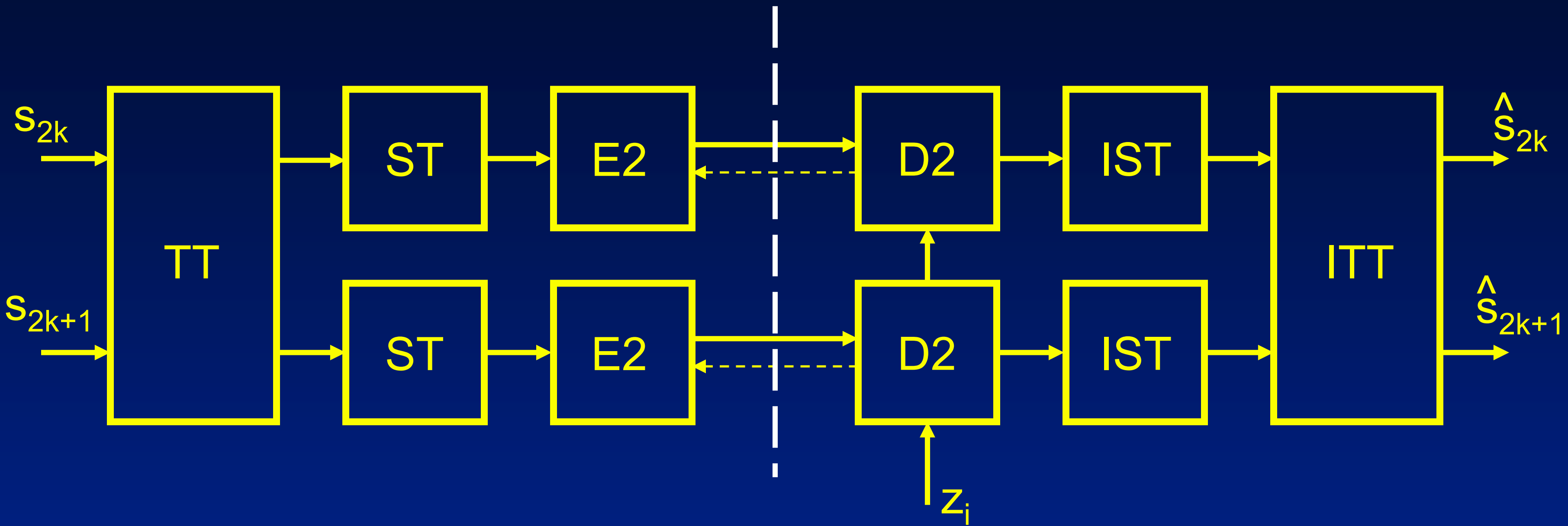
Outline

- Coding scheme with disparity compensation at the decoder
- Motion-compensated temporal Haar wavelet
- Nested lattice codes for transform coefficients
- Decoding with side information
- Experimental results
- Investigate the relation between the level of temporal decorrelation and the efficiency of multi-view side information

Coding Scheme with Disparity Compensation

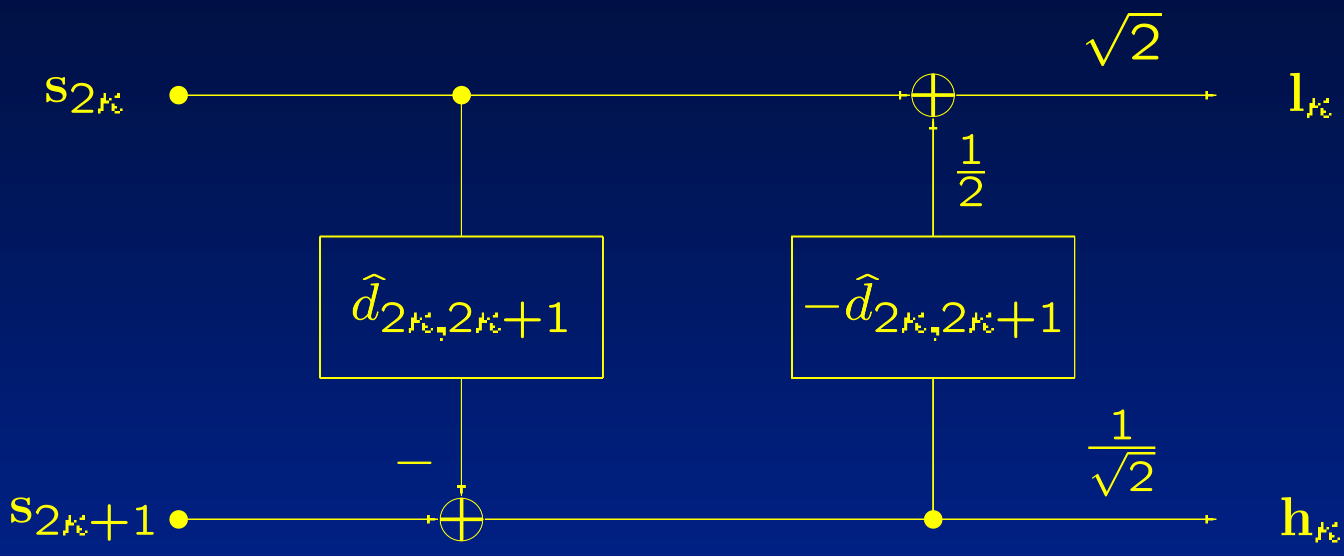


Transform Coding Scheme



- Temporal Transform: Motion-compensated Haar wavelet
  - *Dyadic decomposition for each group of K pictures*
- Spatial Transform: 8x8 DCT
- Coefficient Coder: Nested lattice code
  - *Same minimum distance for all codes*
  - *Coefficient decoder uses side information z_i*

Motion-Compensated Temporal Haar Wavelet

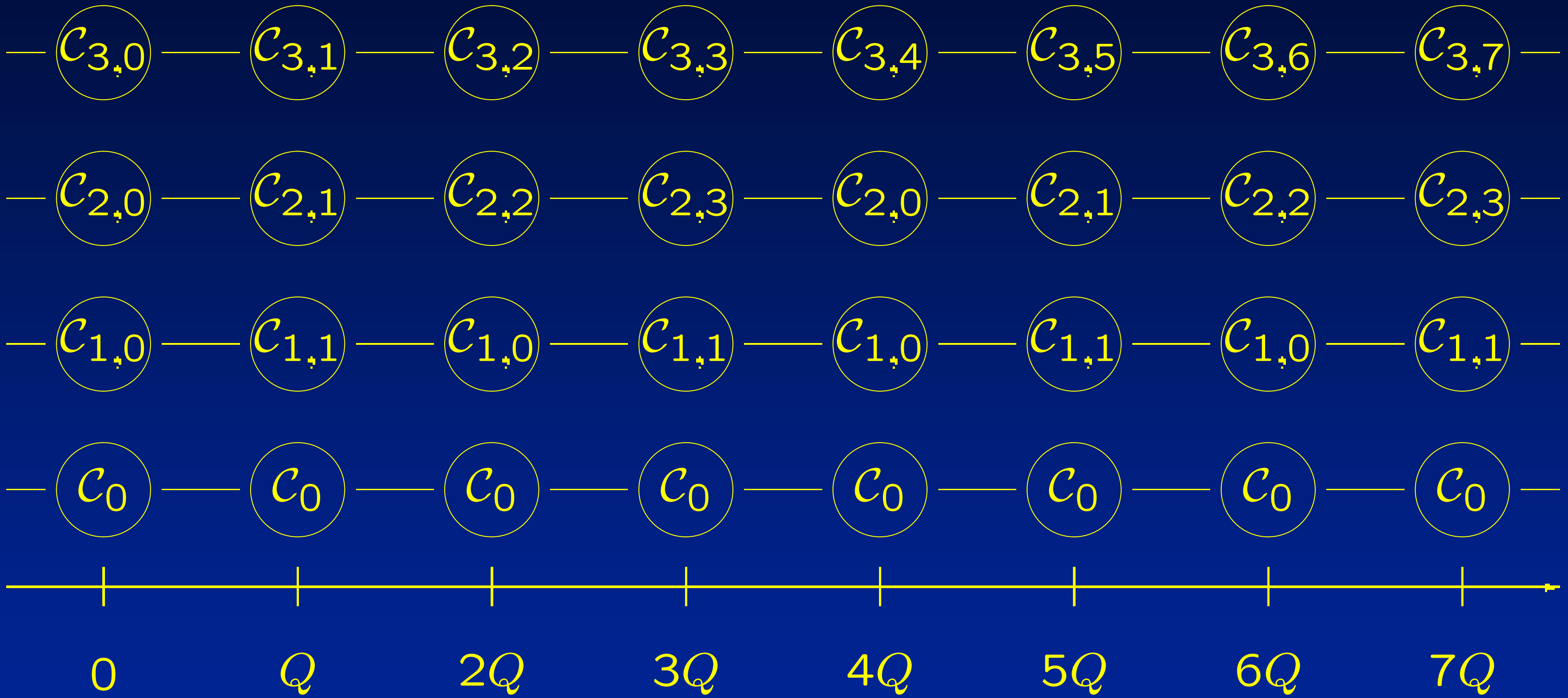


- Haar wavelet with motion-compensated lifting steps
- 16x16 block motion compensation with half-pel accuracy

Coefficient Encoder

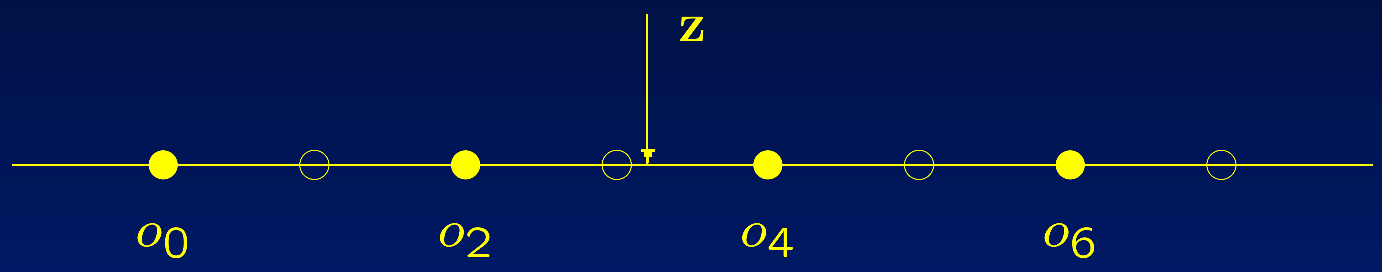
- Each transform coefficient is associated with its side information coefficient
- Correlation is varying for the coefficients
- For weakly correlated coefficients, a higher transmission rate R_{TX} is necessary
- This adaptation is accomplished with a nested lattice code
  - *The fine code has minimum distance Q*
  - *The nested codes are coarser*
  - *The union of the cosets of a coarse code gives the fine code*
  - *The transmission rate R_{TX} determines the number of necessary cosets*

Nested Lattice Code



C_μ are nested codes with the ν-th coset $C_{\mu,\nu}$ relative to C_0

Example: Coset-Coding at R_{TX}=1 bit



- *Encoder 2* transmits at a rate R_{TX} = 1 bit per coefficient
- Coset $C_{1,0} = \{o_0, o_2, o_4, o_6\}$
- Coset $C_{1,1} = \{o_1, o_3, o_5, o_7\}$
- Decoding with side information coefficient **z**

Decoding with Side Information

- *Encoder 2* provides bit-planes of transform coefficients and sends weighted bit-planes
- *Decoder 2* decodes with feed-back and returns a bit-plane mask
- *Decoder 2* attempts to decode the transform coefficients given the received bit-planes and the side information

$$\hat{\mathbf{c}}_i = \underset{\mathbf{c}_i \in \mathcal{C}_{\mu,\nu}}{\text{argmin}} \, [\mathbf{c}_i - \mathbf{z}_i]^2 \quad \text{given} \quad \mu = R_{TX}[i]$$

- No decoding error beyond the critical transmission rate

Algorithm for Error-Free Decoding

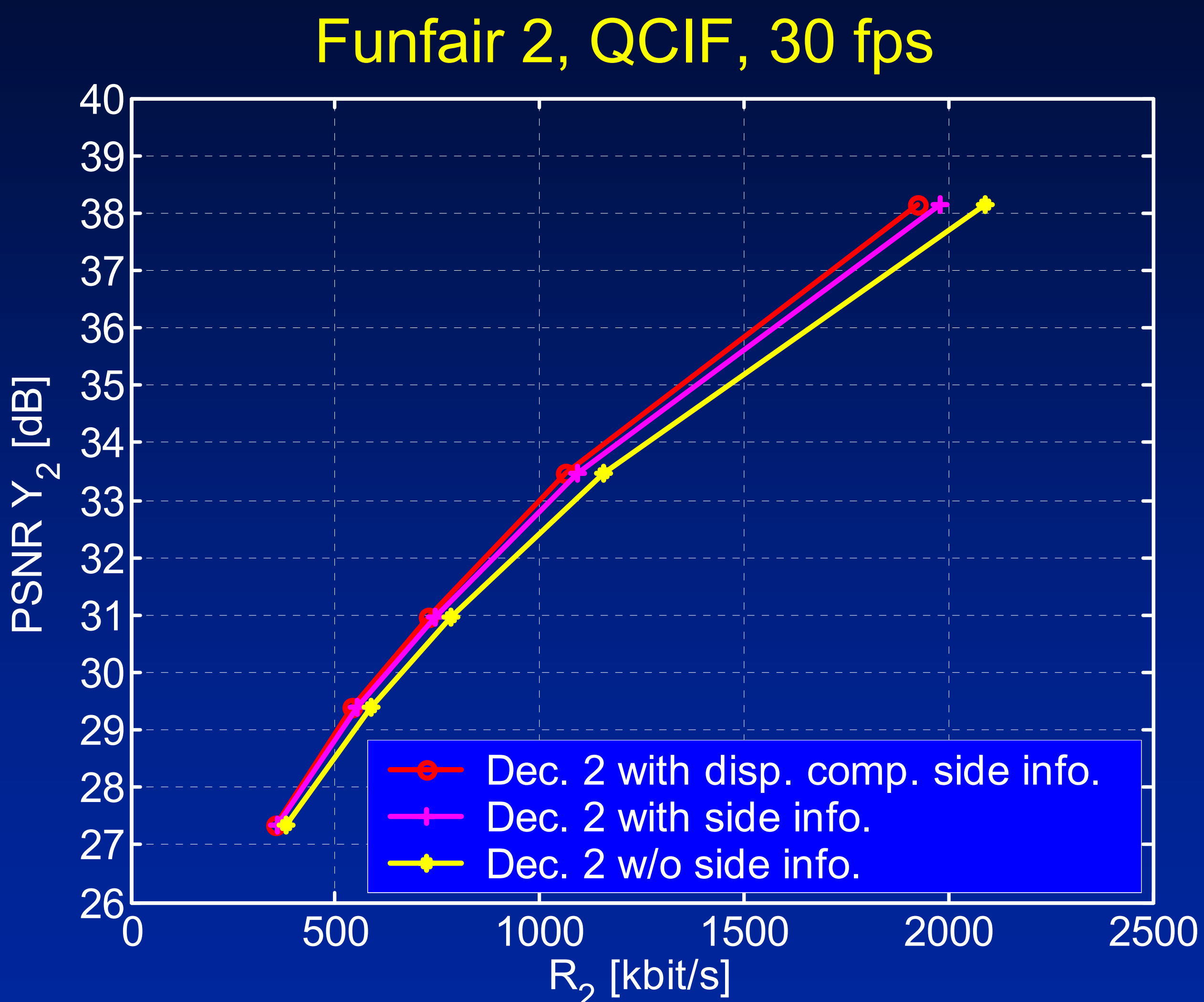
With increasing transmission rate R_{TX}, the coefficient estimate gets more accurate and stays constant for rates beyond the critical transmission rate R^{*}_{TX}.

- When the transmission rate increases by 1 and the estimated transform coefficient changes its value, an additional bit is required
- An unchanged estimate is just a necessary condition for having achieved the critical transmission rate
- In this case, a sufficient condition for error-free decoding is not available at the decoder side and *Encoder 2* determines when to stop sending additional bits

Experimental Results

- *Encoder 1* encodes the side information (left view of a stereoscopic sequence) at high quality
- *Encoder 2* encodes the right view of a stereoscopic sequence
- The GOP sizes for *Encoder 1 & 2* are identical
- The side information is disparity compensated in the image domain
- The disparity is estimated for 24 blocks on the first frame pair
- The camera positions are unaltered in time and the disparity estimates are used for all images

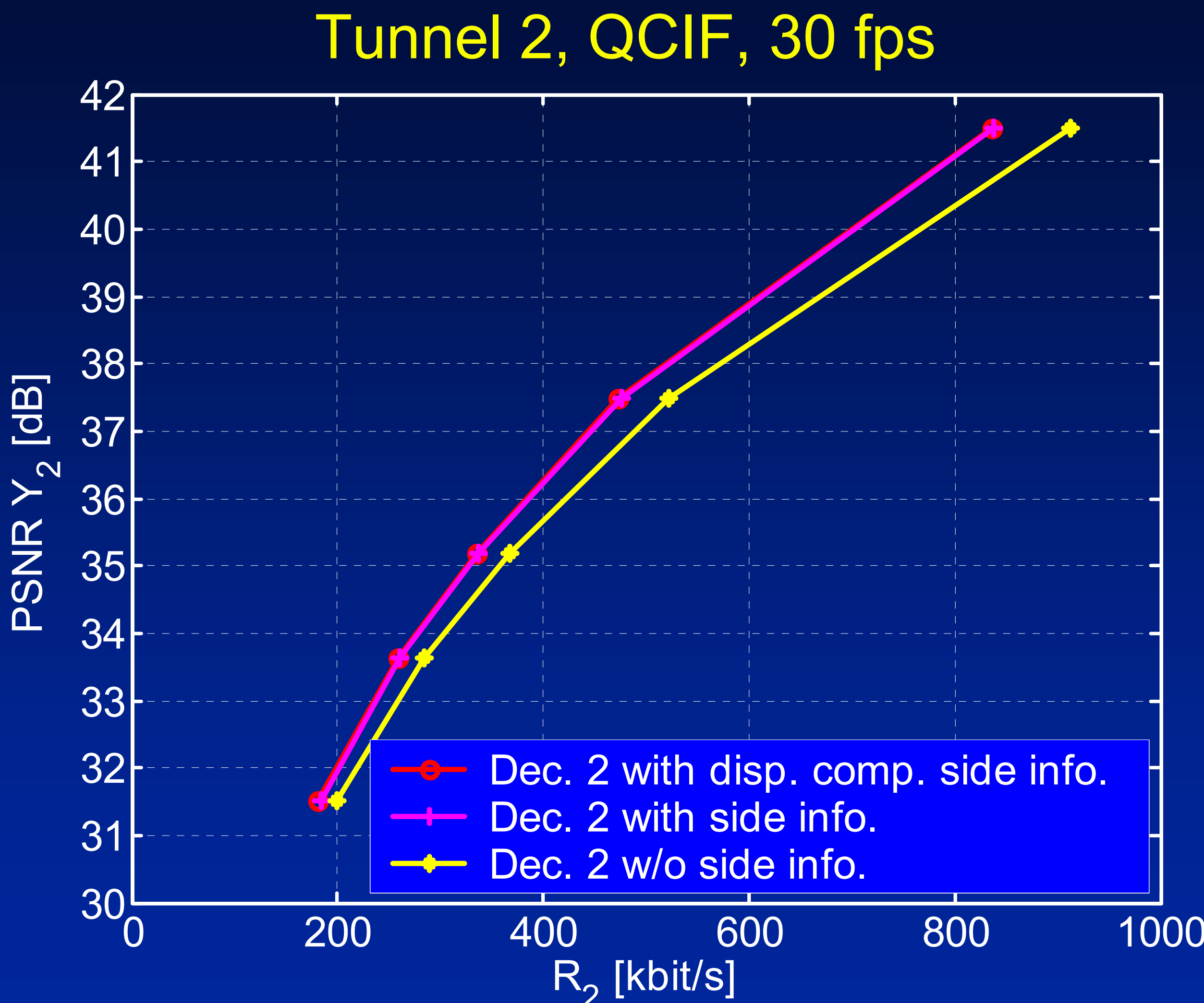
Decoding with Video Side Information: GOP=32



The side information is always quantized finely

Lower gains for coarsely quantized signals

Decoding with Video Side Information: GOP=32



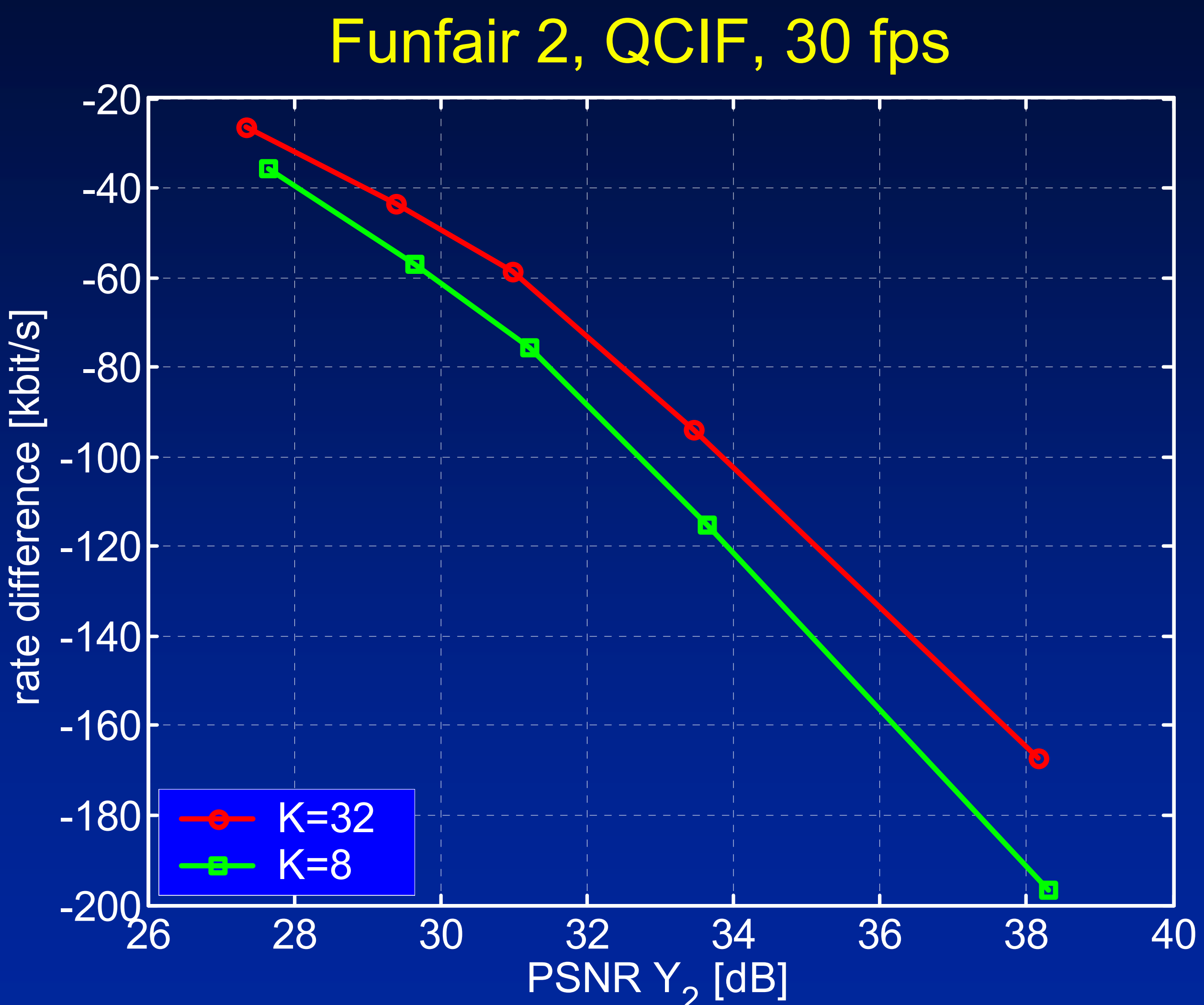
The side information is always quantized finely

Lower gains for coarsely quantized signals

Temporal Decorrelation and Video Side Information

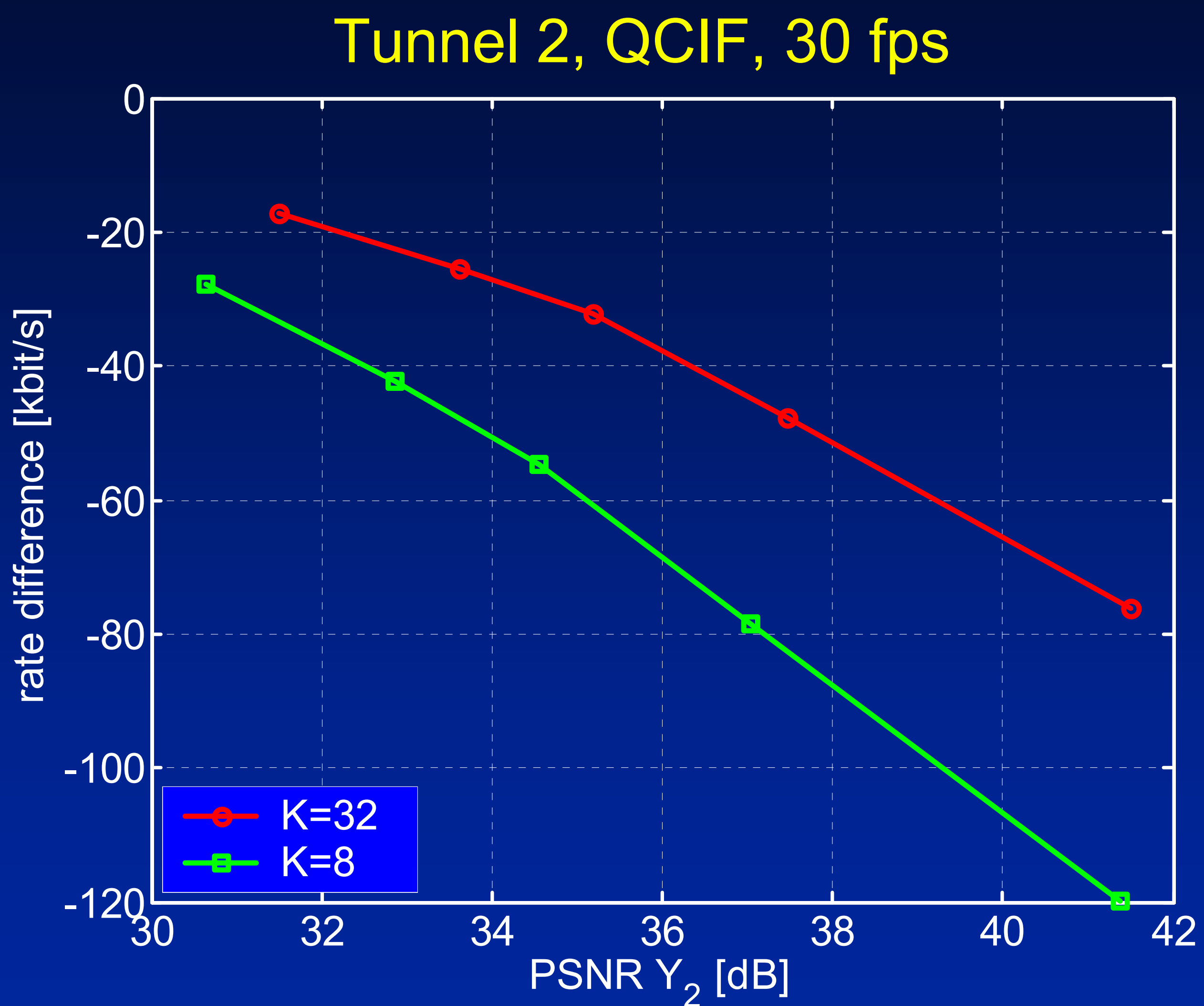
- Correlation of multi-view sequences
  - *Temporal correlation*
  - *Inter-view correlation*
- How does temporal decorrelation affect the efficiency of the video side information of the neighboring view?
- Temporal decorrelation is controlled by the length of the temporal filter …
  - *Short filter – weak temporal decorrelation*
  - *Longer filter – stronger temporal decorrelation*

Temporal Decorrelation and Video Side Information



Side information is less efficient for larger GOP sizes

Temporal Decorrelation and Video Side Information



Side information is less efficient for larger GOP sizes

Conclusions

- Decoding of a video signal given a highly correlated video side information
- Video coding with a motion-compensated temporal wavelet transform
- Use nested lattice codes to represent the transform coefficients
- Observe a trade-off between the level of temporal decorrelation and the efficiency of multi-view side information
- The efficiency of side information increases for decreasing GOP size