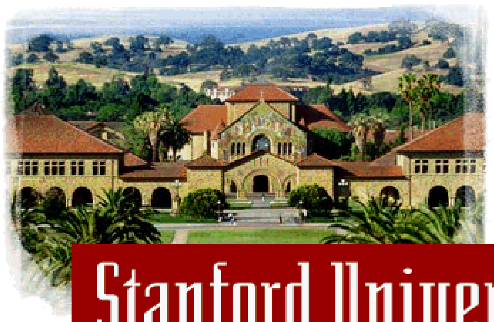
International Conference on Image Processing 2006

Coding of Multi-View Image Sequences with Video Sensors



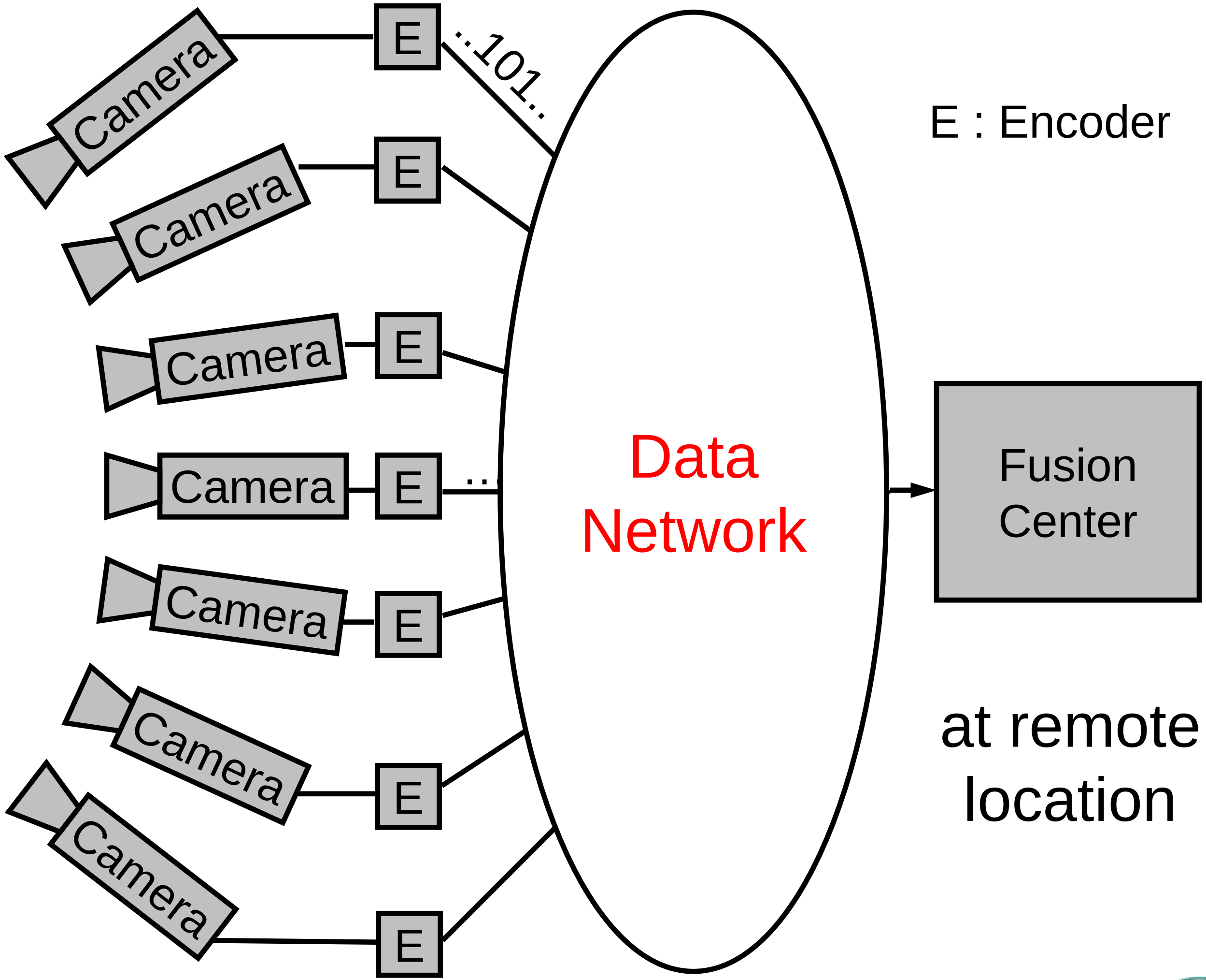
Stanford University

Markus Flierl and Bernd Girod

Max Planck Center for Visual Computing and Communication

Communicating Dynamic 3D Scenes

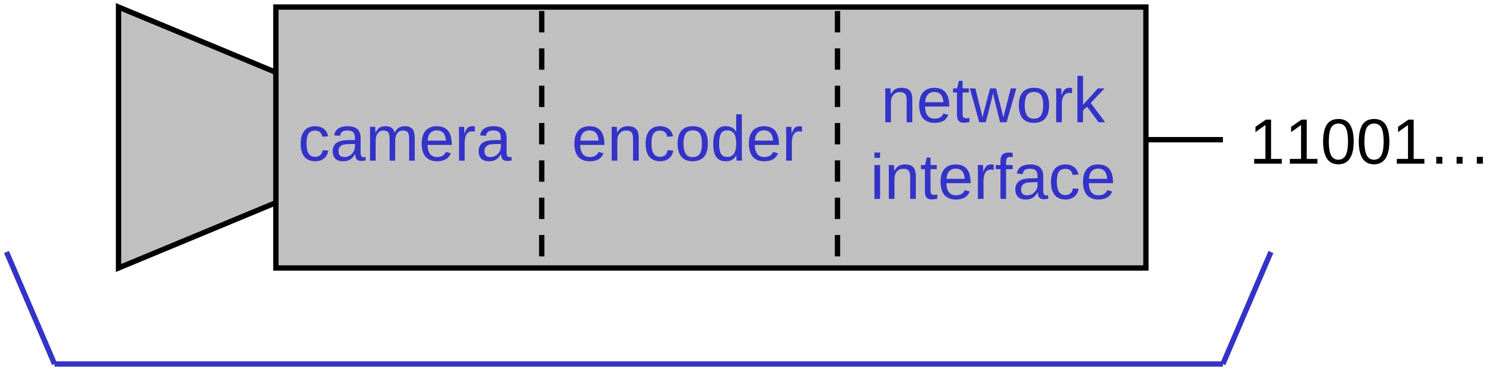
Sampled
Dynamic 3D Scene



Outline

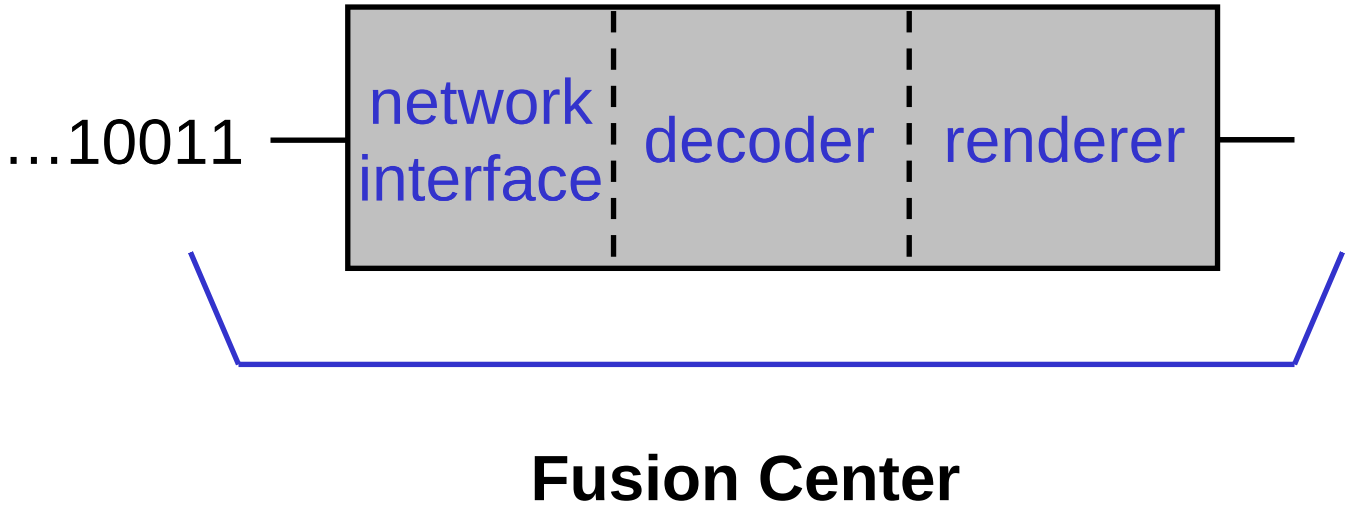
- Video sensor and fusion center
- Correlated multi-view image sequences
- Distributed coding with video sensors
  - Encoder of a video sensor
  - Decoding with side information
  - Disparity-compensated side information
- Experimental results
- Model study for stereoscopic images

Video Sensor

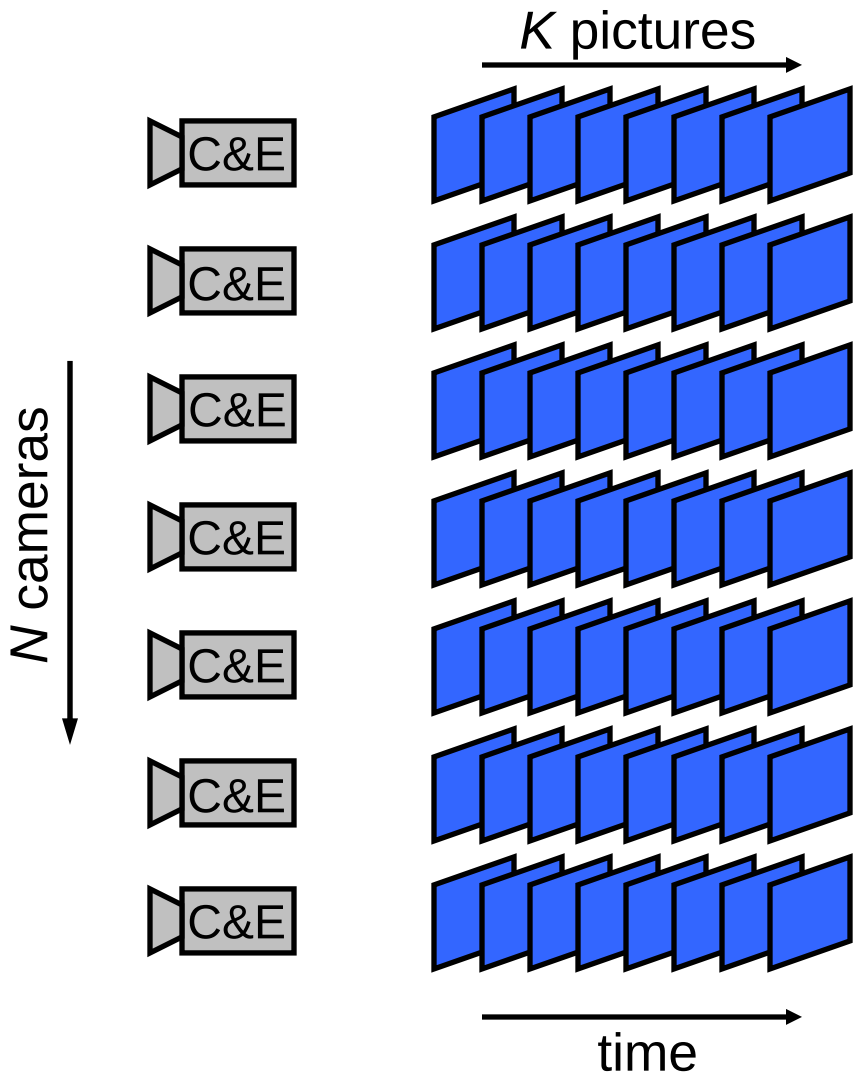


Video Sensor

Fusion Center at Remote Location

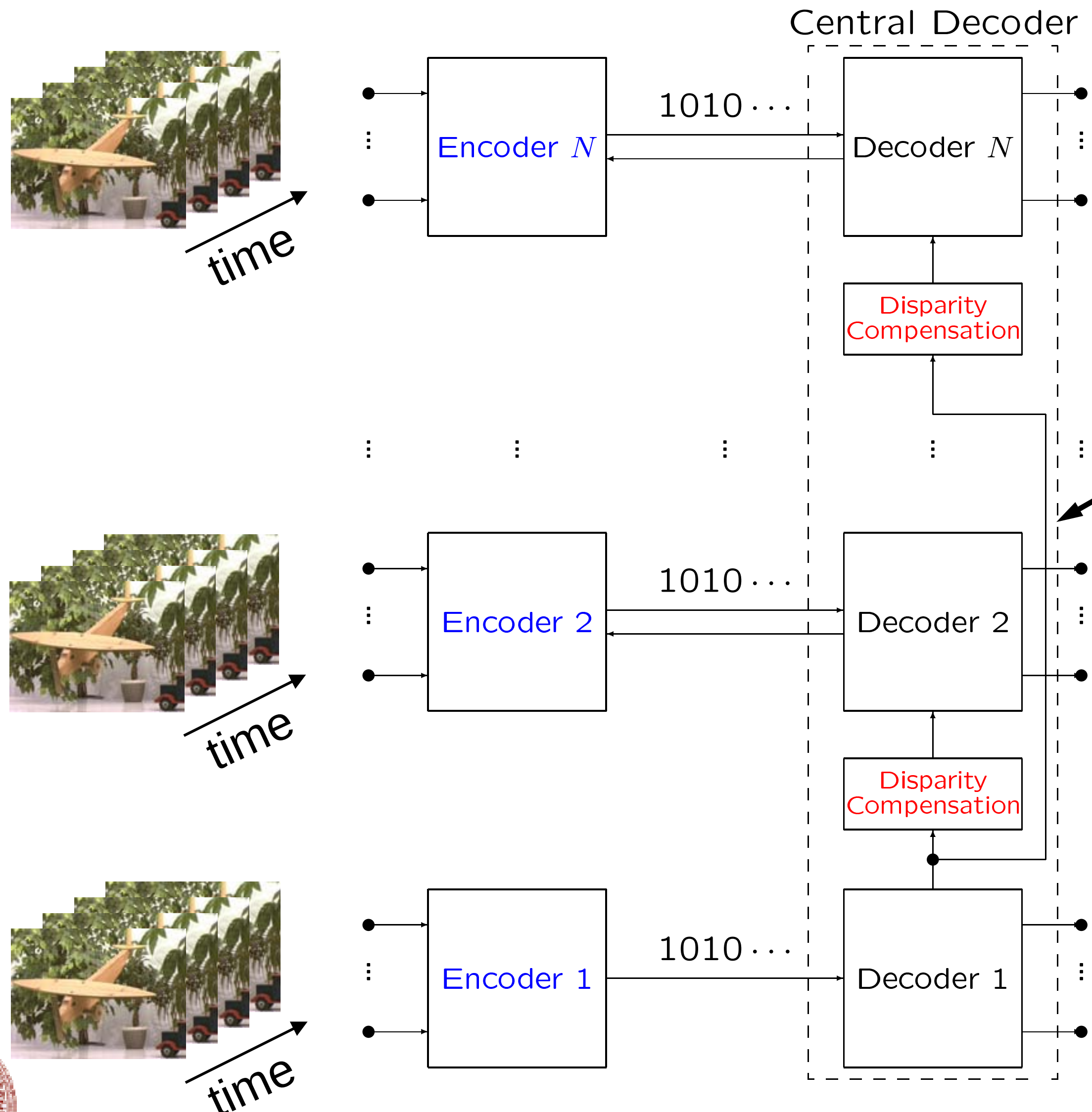


Correlated Multi-View Image Sequences



Each camera exploits the temporal correlation among *K* successive pictures

Distributed Coding with Video Sensors

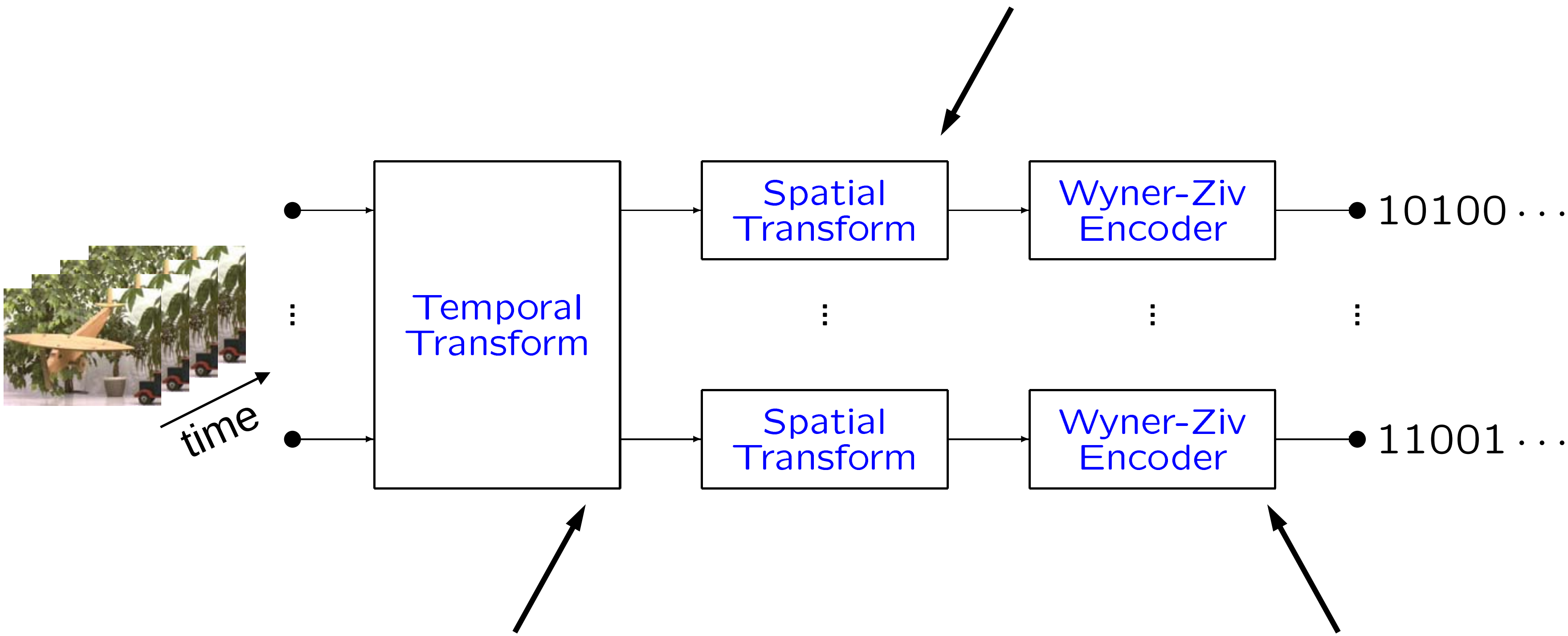


decoder of sensor signal using side information

reconstructed reference view

Encoder of a Video Sensor

8x8 DCT



dyadic decomposition of K pictures
with motion-compensated Haar wavelet

coefficient coder uses
a nested lattice code

Decoding with Side Information

- **Encoder n** uses a nested lattice code and transmits R_{TX} syndrome bits for each transform coefficient *c*
- **Decoder n** decodes R_{TX} syndrome bits for each transform coefficient *c with feed-back:*
  - **Encoder n** sends the initial R_{TX} syndrome bits
  - **Decoder n** attempts to decode the transform coefficient *c* given the received R_{TX} syndrome bits and the coefficient side information *z*

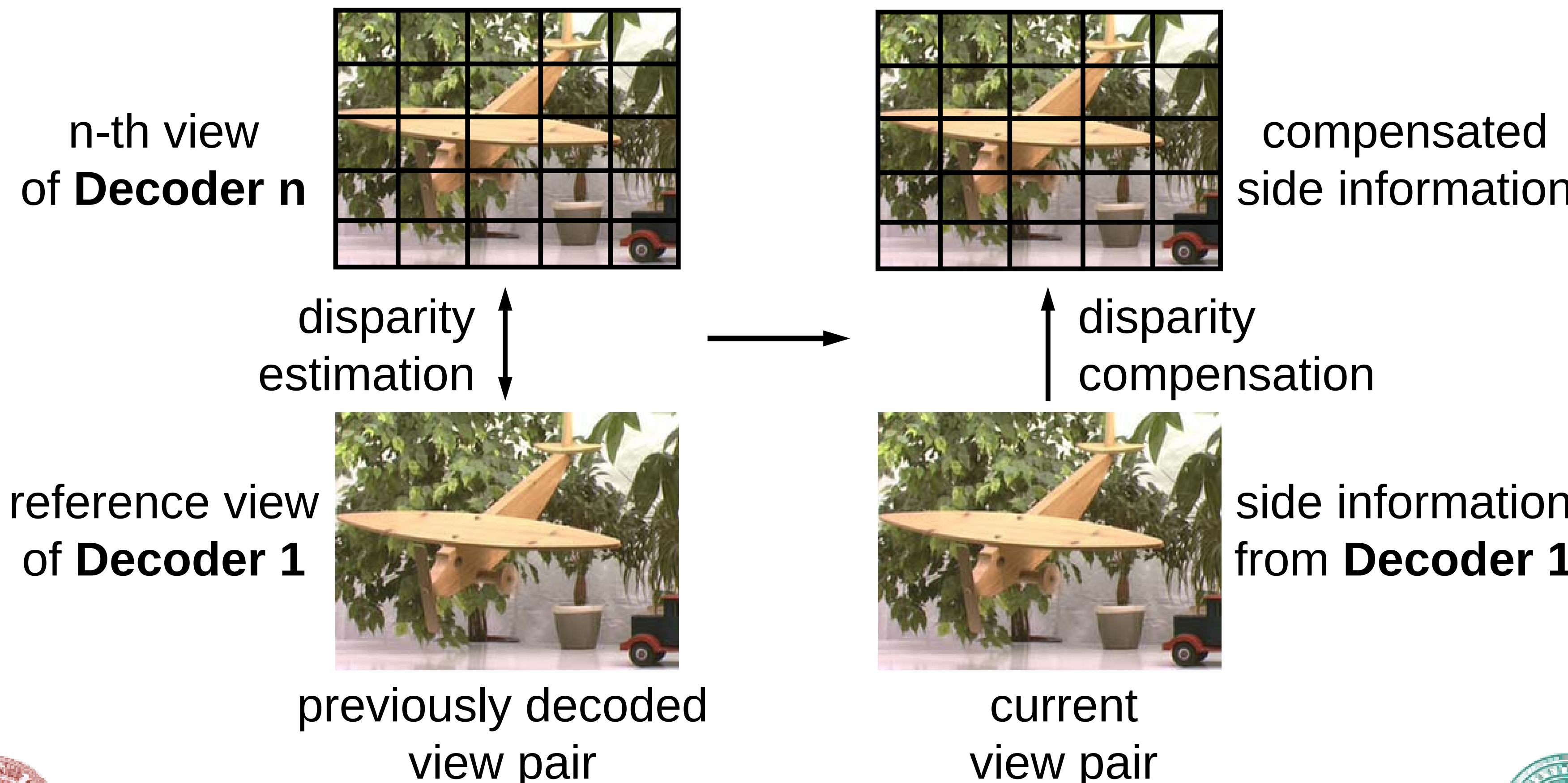
$$\hat{\mathbf{c}} = \underset{\mathbf{c}\in\mathcal{C}_{\mu,\nu}}{\text{argmin}} \; [\mathbf{c} - \mathbf{z}]^2 \quad \text{given} \quad \mu = R_{TX}$$

ν-th coset of the μ-th nested lattice

  - In case of decoding error, **Decoder n** requests further syndrome bits
  - No decoding error beyond the critical syndrome rate

Disparity-Compensated Side Information

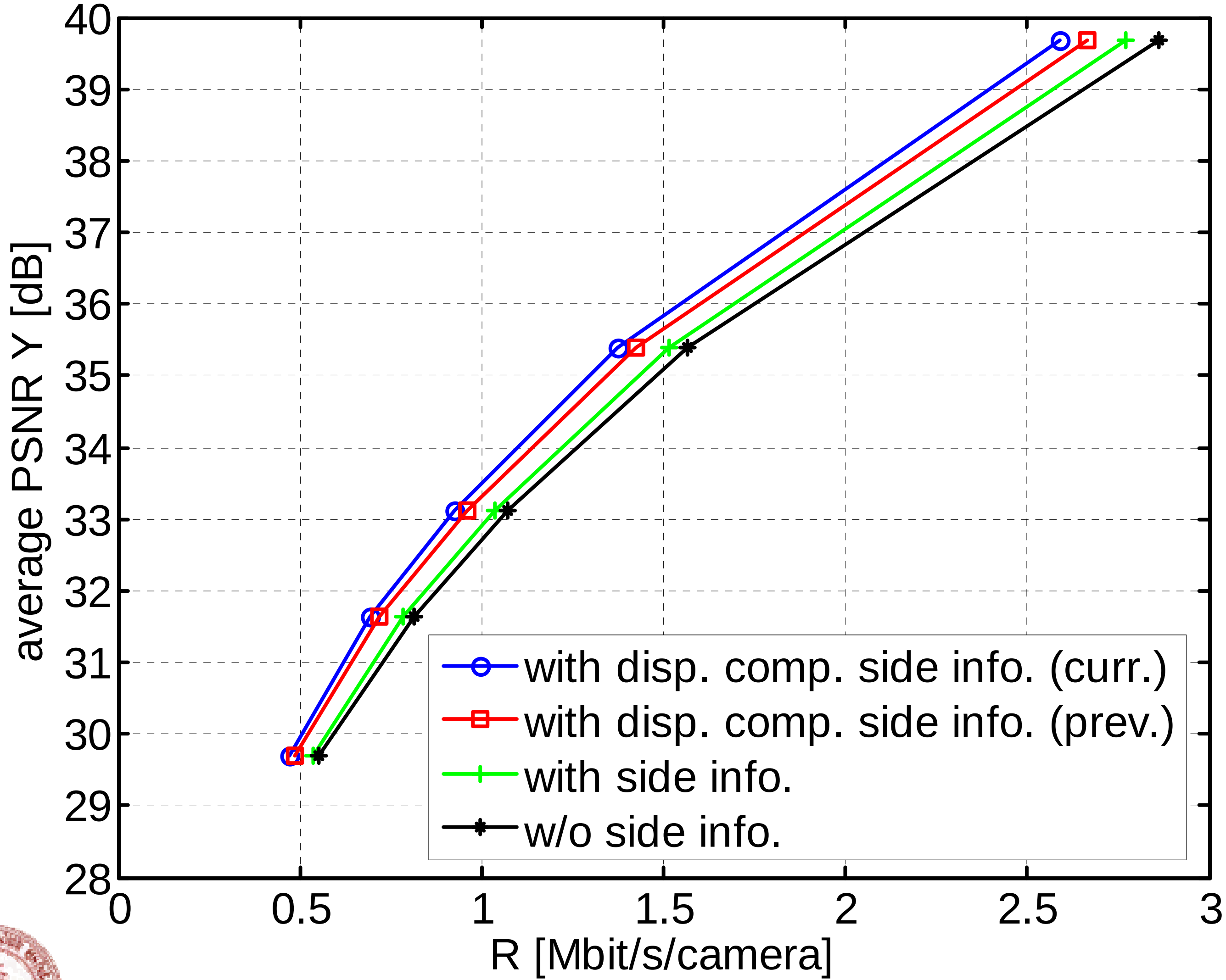
Side information from **Decoder 1** is disparity-compensated in the image domain



Example Test Sequence *Jungle*

[3DTV Network of Excellence]

Experimental Results



Jungle

256x192

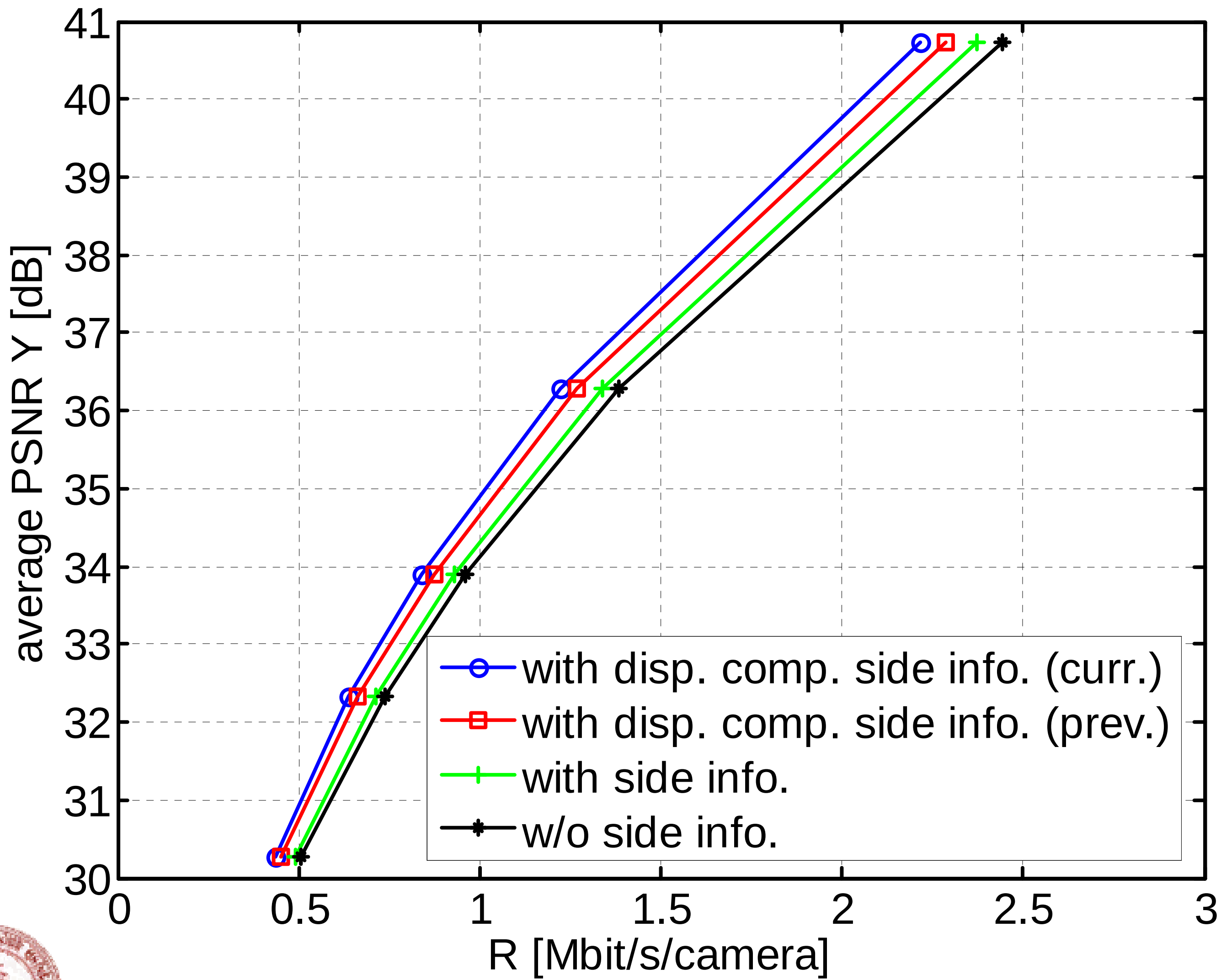
30 fps

240 frames

N=8 views

GOP size K=8

Experimental Results



Uli

256x192

30 fps

240 frames

N=8 views

GOP size K=8

Model Study for Stereoscopic Images

- Let the image **u**$[l_x, l_y]$ be a scalar Gaussian random field
- Let **u'**(x,y) be its space-continuous ideal reconstruction
- Let **w**[l] be a shifted and noisy version of the image **u**[l] with the deterministic 2D real-valued shift Θ_c

white noise

$$\mathbf{w}[l] = \mathbf{u}'(l - \Theta_c) + \mathbf{n}[l]$$

- Let **s**[l] be a shifted and noisy version of the image **u**[l] with the uncertain shift **Θ**, distributed with the PDF $f_{\boldsymbol{\Theta}}(\Theta)$

$$\mathbf{s}[l] = \mathbf{u}'(l - \boldsymbol{\Theta}) + \mathbf{n}[l]$$

- Compare conditional **differential entropy rate** differences:

H(**u**|**s**)-**H**(**u**) vs. **H**(**u**|**w**)-**H**(**u**)

Model Study for Stereoscopic Images

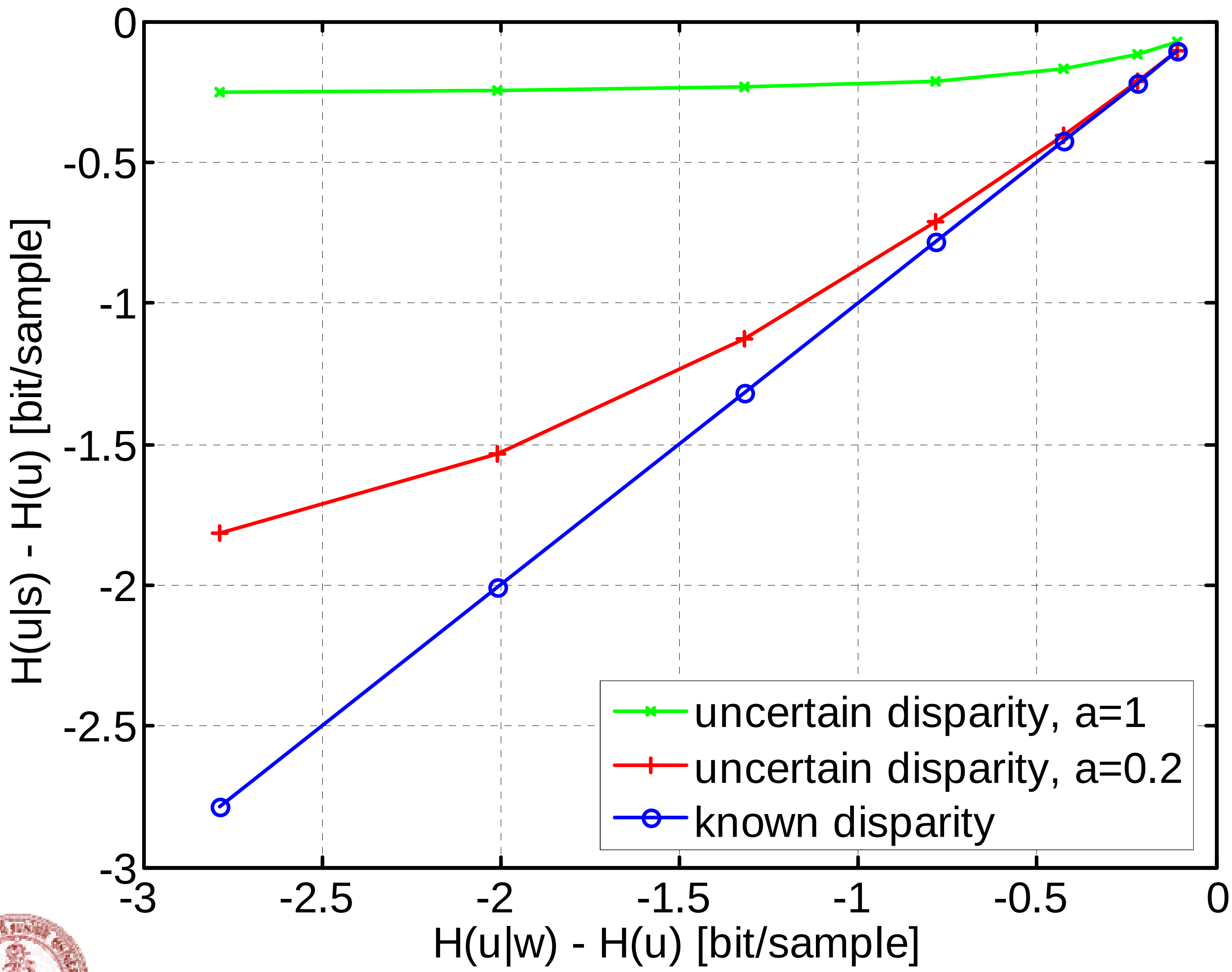


Image **u**:
scalar Gaussian random field

Image **w**:
$f_\Theta(\Theta)=\delta(\Theta - \Theta_c)$

Image **s**:
$f_\Theta(\Theta)=\frac{1}{2a}\mathbf{1}_{[-a,a]}(\Theta)$

Note:
deterministic disparity does not affect **H**(**u**|**w**)

Conclusions

- Exploit view-correlation of multi-view image sequences
- Operate video sensors in a collaborative fashion
- Centralized decoder performs disparity compensation
- Our experiments show that:
  - Disparity-compensated side information reduces bit-rate up to 10%
  - Without disparity compensation, gain is limited to 3%
- The uncertainty of the estimated disparity at the decoder causes a **entropy rate loss** when compared to centralized encoding