STREAMING VIDEO WITH OPTIMIZED RECONSTRUCTION-BASED DCT

Benjamin W. Wah and Xiao Su

Department of Electrical and Computer Engineering
and the Coordinated Science Lab
University of Illinois at Urbana-Champaign
Urbana, IL61801, USA
E-mail: b-wah@uiuc.edu
http://manip.crhc.uiuc.edu

July 2000

Outline

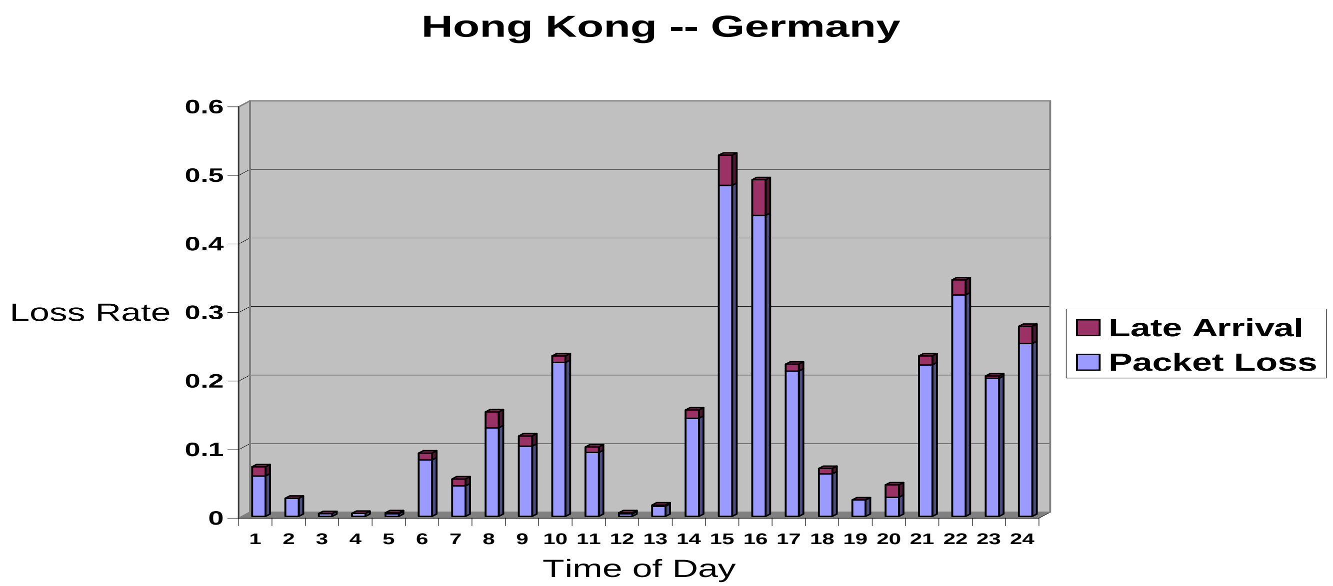
- Motivations: dealing with information loss
  - Internet loss behavior
  - Performance of commercial players under loss
- Previous work
- Proposed optimized reconstruction-based DCT (ORB-DCT)
- Experimental results
- Conclusions and future work

Three Major Challenges in Real-Time Internet Transmissions

- Limited bandwidth:
  - CIF (352x288) H.263 video at 30 f/s needs 80 Kbps-1.5 Mbps
  - Internet sustained bandwidth ranges from 30 - 800 Kbps
- Strict timing constraints:
  - Real-time streaming requires playback at prescribed time instances
  - Jitters make packets arrive too late for playback
- High playback quality:
  - High-quality real-time streaming needs robust delivery mechanisms
  - Poor quality due to information loss and limited bandwidth

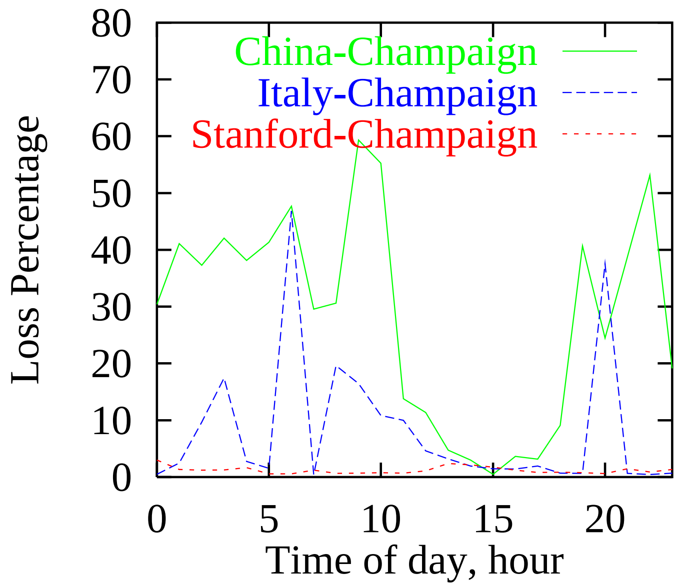
Sources of Information Loss

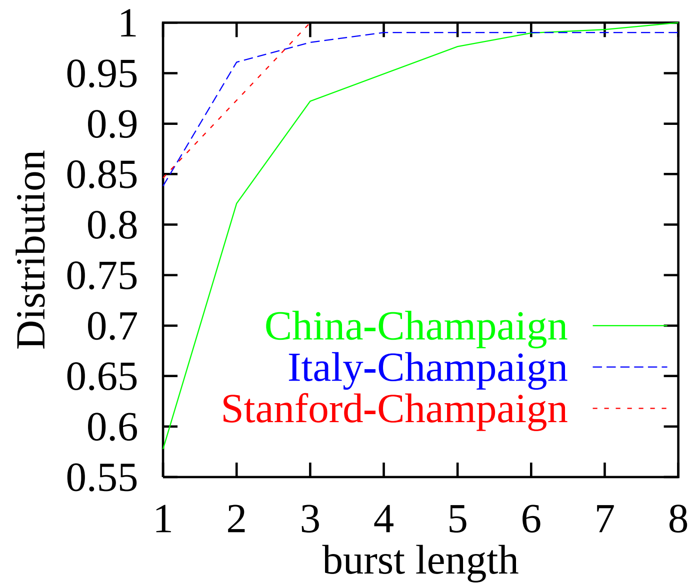
- Compression loss: lossy quantization
- Bitstream loss (addressed in this talk)
  - Due to network – over 50% (network and jitter losses)
  - Due to scaling – dropped by sender in limited bandwidth (video)



- Propagation loss due to dependencies (video)

Statistics on Burst Lengths

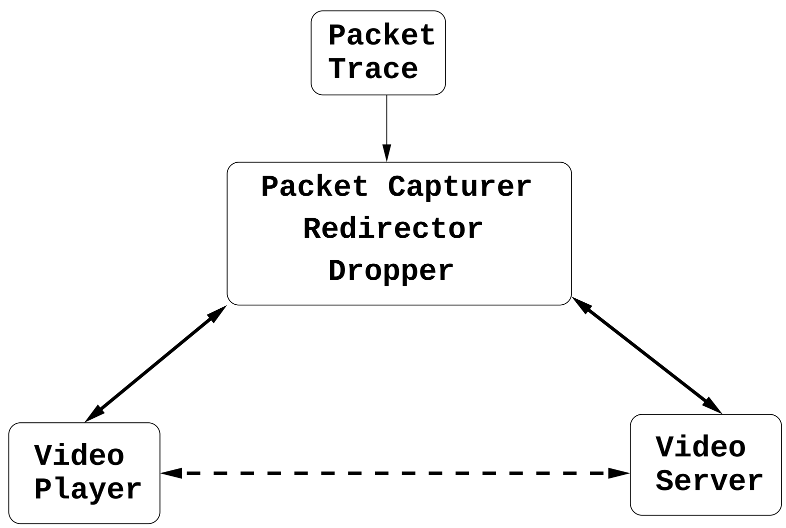




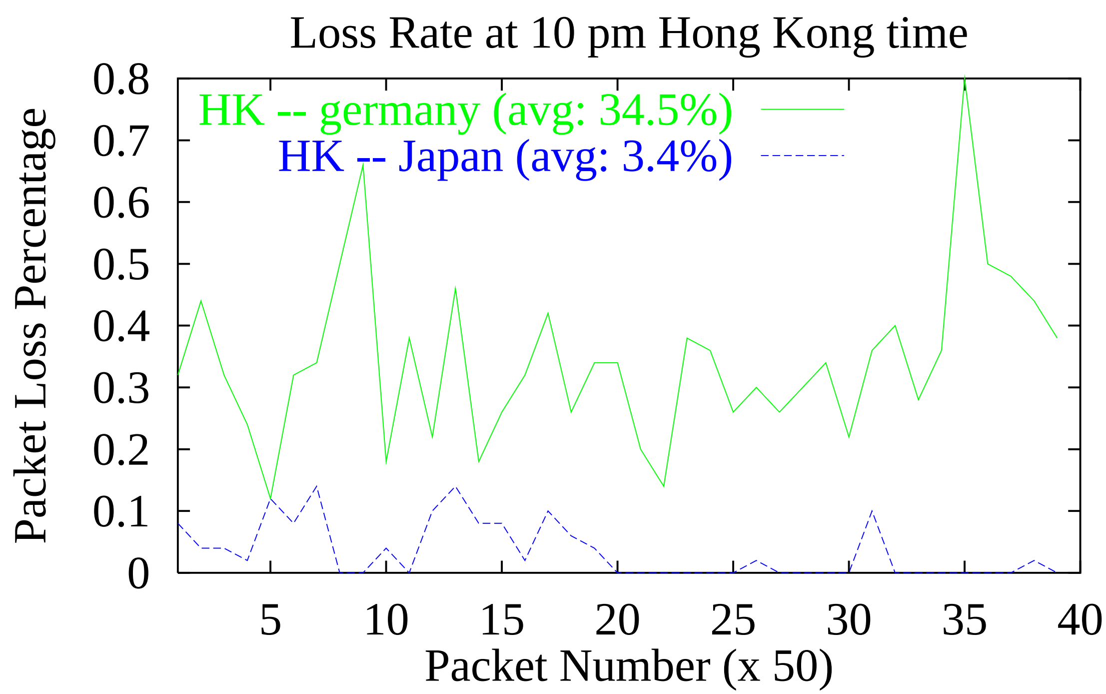
- Probability of burst length > 1 is 2% for connections within US
- Probability of burst length > 3 is 5% for transcontinental connections

Evaluation of Commercial Video Streaming Systems

- Leading Internet video streaming players – RealPlayer, MediaPlayer
  - Proprietary codecs
  - Unknown error concealment schemes
  - Initial buffering delay of several seconds
- Trace-based evaluations to compare quality under same traffic conditions and bit rate



Playback Quality of Commercial Players under Loss



- Test sequence: boxing at 320×224, 5 fps and 80 Kbps
- Observations: like a slide show
  - Video freezes on packet losses, 5-15 seconds to recover
  - Visible corrupted blocks

Previous Work: Coder Independent

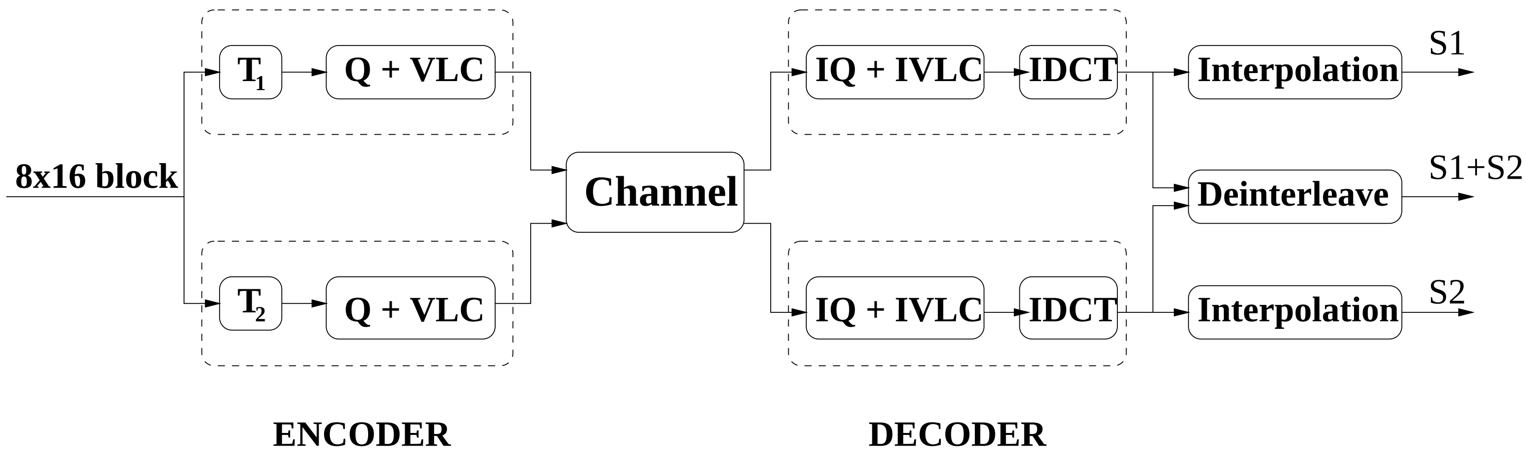
- Sender-based:
  - Intelligent packetization to prevent loss propagation among packets
  - Only handling error propagation
- Receiver-based:
  - Recovery in spatial, temporal, or frequency domains
  - No interaction with senders
- Sender receiver-based:
  - Forward error correction (FEC) – not for bursty packet losses
  - Retransmission-based recovery – increased bandwidth
  - Interleaving or scrambling
    - Does not work well when pixel values change rapidly

Previous Work: Coder Dependent

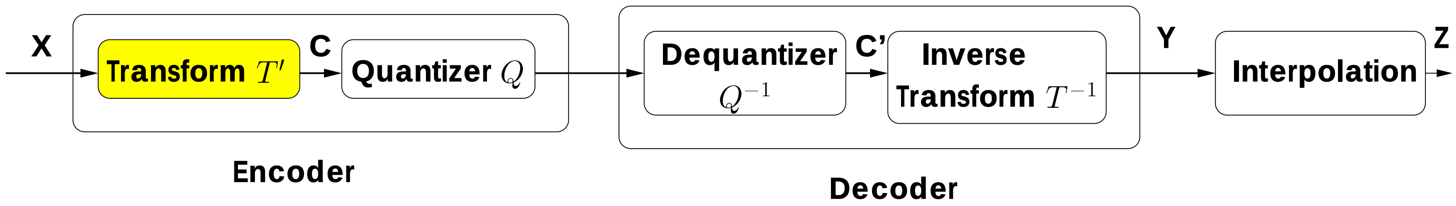
- Source coding that exploits redundancy
  - Robust entropy coding
  - Restricting prediction domain
  - Layered coding
    - Assume prioritized transmissions
  - Multiple-description coding
    - Computationally expensive
    - Mediocre reconstruction quality from a side description
- Source coding followed by channel coding
  - Protection is added according to the importance of a bit stream
- Joint source-channel coding
  - Need a generalizable channel error model

Overview of Proposed MDC Based on Block Transform Coders

- Use block as a basic unit
- T_1 and T_2 are of dimension 128×64



Approach: Deriving Transformations T_1 and T_2



- Objective: assume T^{-1} is fixed, find quantized transformed coefficients $C'_{n\times n}$ to minimize reconstruction error $\mathcal{E}_r$, after inverse DCT transform T^{-1} and average interpolation.

$$\mathcal{E}_r = \| \underbrace{Interpolate(T^{-1}(IQ(Q(\mathbf{c}))))}_{decompression + reconstruction.} - \mathbf{x} \|^2 .$$

- Features:
  - Standard-compliant decoder

Derivations of ORB-DCT (I)

- Derive Y after inverse DCT transform

$$Y = \sum_{i=1}^{n} \sum_{j=1}^{n} C'(i,j) b_i b_j^T$$

b_i – basis vector (8×1) of DCT transform, $i = 1, 2, \ldots, n$
$b_i b_j^T$ – the $(i,j)_{th}$ basis image (8×8), $i, j = 1, 2, \ldots, n$

- Derive Z after average interpolation

$$Z = \sum_{i=1}^{n} \sum_{j=1}^{n} C'(i,j) b_i e_j^T$$

$$e_j = (b_{j1}, \frac{b_{j1} + b_{j2}}{2}, b_{j2}, \frac{b_{j2} + b_{j3}}{2}, b_{j3}, \ldots, b_{jn}, b_{jn})^T$$

e_j – extended basis vector (16×1) of DCT transform, $j = 1, 2, \ldots, n$

Derivations of ORB-DCT (II)

- Problem formulation:

$$\min \mathcal{E}_r(C') \;=\; \| \sum_{i=1}^{n} \sum_{j=1}^{n} C'(i,j) b_i e_j^T - X \|^2$$

- Constrained integer optimization problem
  - $C'(i,j)$ are constrained to values in a subset of integers
- Approximate solution can be obtained by minimizing

$$\mathcal{E}_r \;=\; \| \underbrace{Interpolate(T^{-1}(\mathbf{c}))}_{decompression + reconstruction.} - \mathbf{x} \|^2 \; .$$

- In the approximation, $C(i,j) = C'(i,j)$ is assumed.

Optimized Reconstruction-Based DCT

- Intra-coded blocks:

$$\min \mathcal{E}_r(C) = \left\| \sum_{i=1}^{n} \sum_{j=1}^{n} C(i,j) b_i e_j^T - X \right\|^2$$

$$\Longrightarrow \vec{C}_{64\times1} = \vec{T'}_{64\times128} \ \vec{X}_{128\times1}$$

- Inter-coded blocks:

$$\min \mathcal{E}_r(C) = \left\| \sum_{i=1}^{n} \sum_{j=1}^{n} C(i,j) b_i e_j^T - (X - P) \right\|^2$$

$$\Longrightarrow \vec{C}_{64\times1} = \vec{T'}_{64\times128} \ (\vec{X} - \vec{P})_{128\times1}$$

where P denotes its interpolated reference block

One Description (out of 2) Consistently Lost

Video Sequence	Odd received			Even received		
	DCT	ORB-DCT	Gain	DCT	ORB-DCT	Gain
	No quantization effects					
Missa	39.44	41.31	**1.87**	39.51	41.45	**1.94**
Football	36.05	37.48	**1.43**	36.01	37.47	**1.46**
	With quantization effects					
Missa	36.20	36.61	**0.41**	36.14	36.59	**0.45**
Football	29.43	29.82	**0.39**	29.40	29.83	**0.43**

Reconstruction Quality with 4 Descriptions

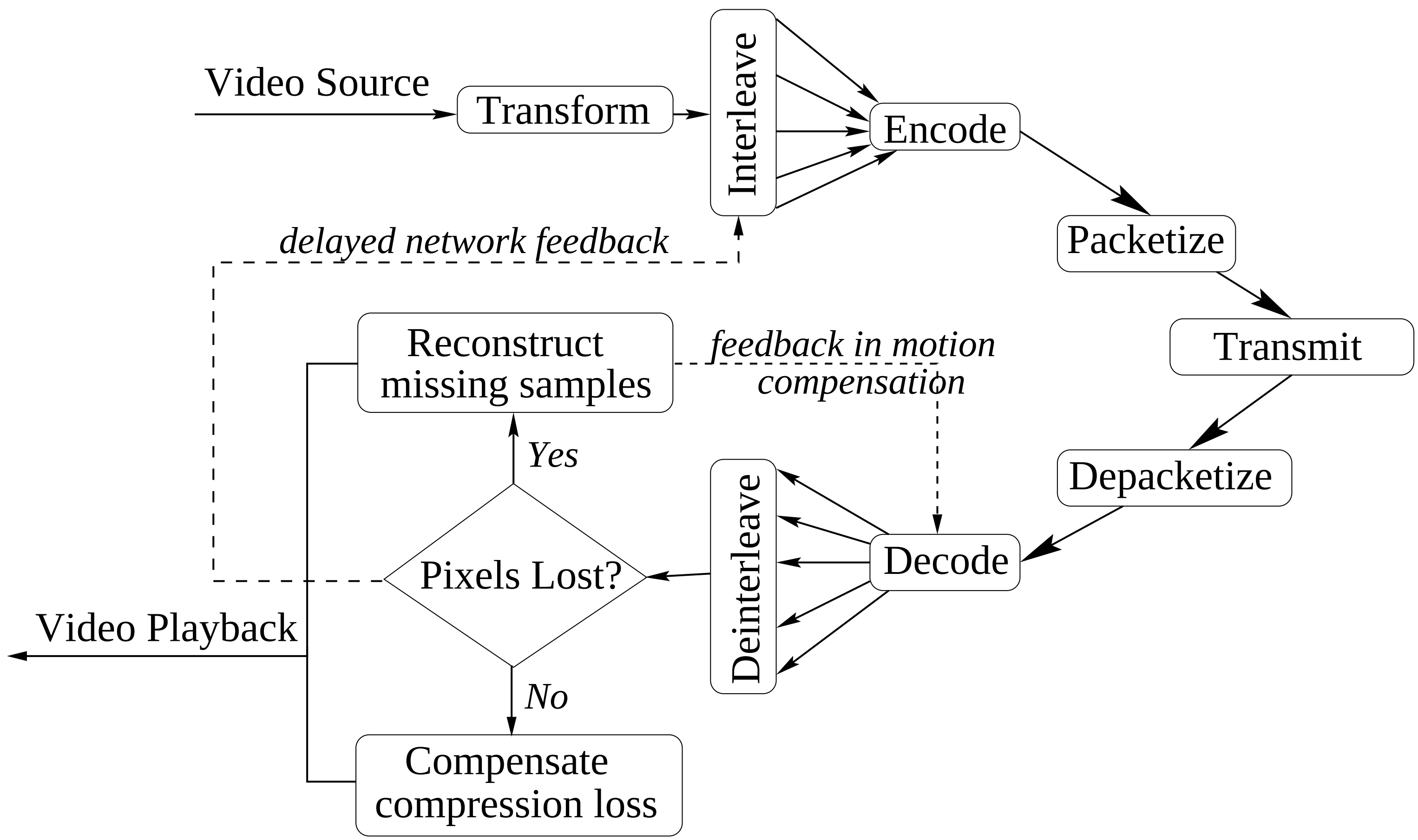
Video Sequence	Quant. Effects	Case I DCT	Case I ORBDCT	Case I gain	Case II DCT	Case II ORBDCT	Case II gain
Missa	No	35.84	37.27	**1.43**	39.35	39.88	**0.53**
football	No	34.92	35.97	**1.05**	35.72	36.15	**0.43**
Missa	Yes	33.58	33.93	**0.35**	34.01	34.23	**0.22**
Football	Yes	24.32	24.68	**0.36**	27.76	27.96	**0.20**
		Case III			Case IV		
Missa	No	39.38	41.25	**1.87**	42.82	42.74	**-0.08**
Football	No	35.99	37.35	**1.36**	40.15	40.00	**-0.15**
Missa	Yes	34.47	34.89	**0.42**	35.07	35.16	**0.09**
Football	Yes	28.43	28.83	**0.40**	29.24	29.37	**0.13**

I: three out of the four interleaved descriptions were lost;
II: two descriptions, both from the same horizontal group, were lost;
III: two descriptions, each from a different horizontal group, were lost;
IV: one out of the four interleaved descriptions was lost.

All Descriptions Correctly Received

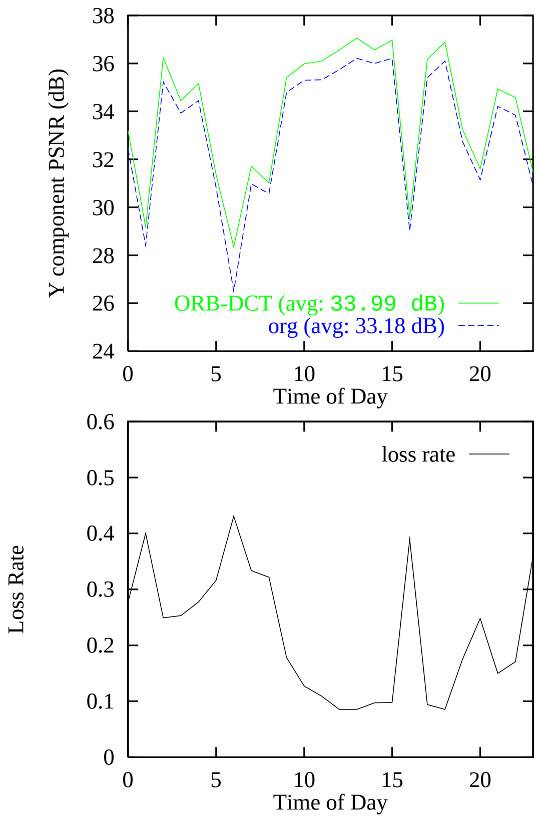
Video Sequence	Interleaving Degree	DCT	DCT & NN	ORB -DCT	ORB-DCT & NN	Gain in PSNR (dB)
Missa	2	36.74	37.06	36.70	37.05	**0.31**
Football		30.16	30.69	30.09	30.67	**0.51**
Missa	4	35.53	36.09	35.43	36.02	**0.49**
Football		29.73	30.35	29.72	30.31	**0.58**

Video Streaming Prototype

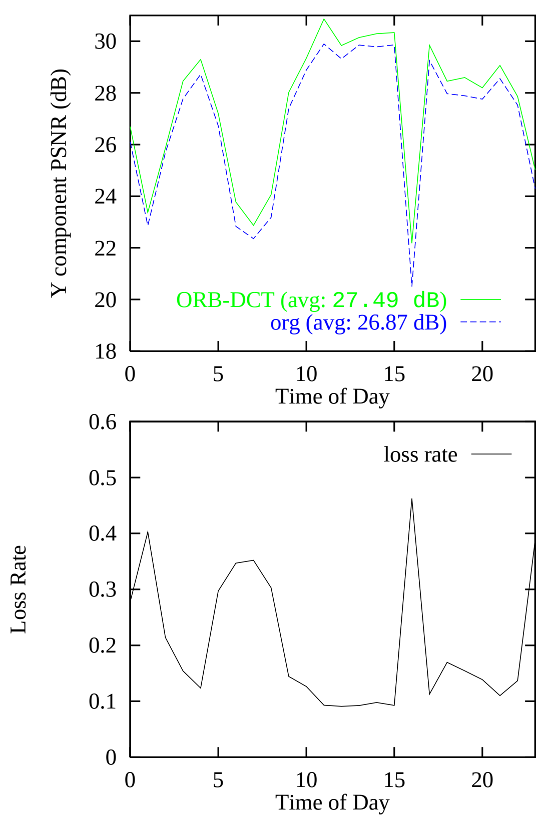


Performance: Tests over the Internet

- Traces between Champaign and Qing Dao, China



a) *missa* sequence



b) *football* sequence

Conclusions and Future Work

Conclusions:

- ORB-DCT is designed in joint sender-receiver fashion
- Improved over conventional DCT under loss scenarios

Future work:

- Fast implementations of ORB-DCT
- Bandwidth-restricted concealment schemes
- Extended to object-based coding schemes