FREQUENCY-BASED RECONSTRUCTION OF MULTI-DESCRIPTION CODED JPEG2000 IMAGES

Hang Yu and Benjamin W. Wah

Dept. of Electrical and Computer Engineering
and the Coordinated Science Laboratory
University of Illinois at Urbana-Champaign

Presented at the *IEEE International Conference on Computer Networks and Mobile Computing*

October 22, 2003

Outline

- Introduction
- Internet Traffic Behavior
- Frequency-Based Loss Concealment
  - Approach
  - Correlation Analysis
  - Optimal Linear Reconstruction
- Experimental Results

Motivations

- Quality and delay to assess image transmissions over network
- Two ways of transmitting images over the lossy Internet
  - TCP: high image quality but long delay
  - UDP: high to poor image quality but short delay
- Objective in this paper
  - Design a UDP-based coding scheme with short delay and good quality

An Example

Sending 512×512 *lena* compressed at 0.125 bpp by JPEG2000

- Between UIUC and Thailand$_2$ (`www.kmitnb.ac.th`)
- Two out of the eight packets were lost in UDP

TCP:
PSNR 30.97 dB,
Roundtrip time 4.01 sec.

UDP:
PSNR 20.51 dB,
Roundtrip time 0.71 sec.

Proposed Theme:
PSNR 25.21 dB,
Roundtrip time 0.71 sec.

Previous Work on Loss Concealments

- Receiver Based
  - Post-processing by assuming smoothness in image
    - High computation cost and image dependent
- Sender-Receiver Based
  - Joint source-channel coding: jointly minimize source/channel coding error
    - Need prior information on channel
- Sender Based
  - Layered coding: base layer + enhancement layer
    - Need network support on QoS
  - Multiple description coding: divide source into equally important descriptions, each reproducing acceptable quality

Multiple Description Coding (MDC)

MDC with sample-domain reconstruction

- Decompose image into segments due to packet size restriction
- Interleave samples in each segment into odd/even descriptions
- Problem with segmentation:
  - Coding efficiency is impaired because redundancy among segments cannot be removed

Proposed MDC with frequency-domain reconstruction

- Segmentation and reconstruction in the frequency domain
- Avoid degradation due to segmentation in sample domain

Illustration of Degradations due to Segmentation

Segment, compress, decompress, and reassemble image

Lena					
Sample-level Segment.	PSNR (dB) at Bits Per Pixel (bpp)				
Segment Size	2	1	0.5	0.25	0.125
No segmentation	43.91	40.07	37.16	34.08	30.97
256×256 segments	43.51	39.34	36.02	32.76	29.41
128×128 segments	42.55	37.93	33.92	29.42	24.14
64×64 segments	40.01	33.85	27.01	–	–

PSNR degrades severely even under no loss

Outline

- Introduction
- Internet Traffic Behavior
- Frequency-Based Loss Concealment
  - Approach
  - Correlation Analysis
  - Optimal Linear Reconstruction
- Experimental Results

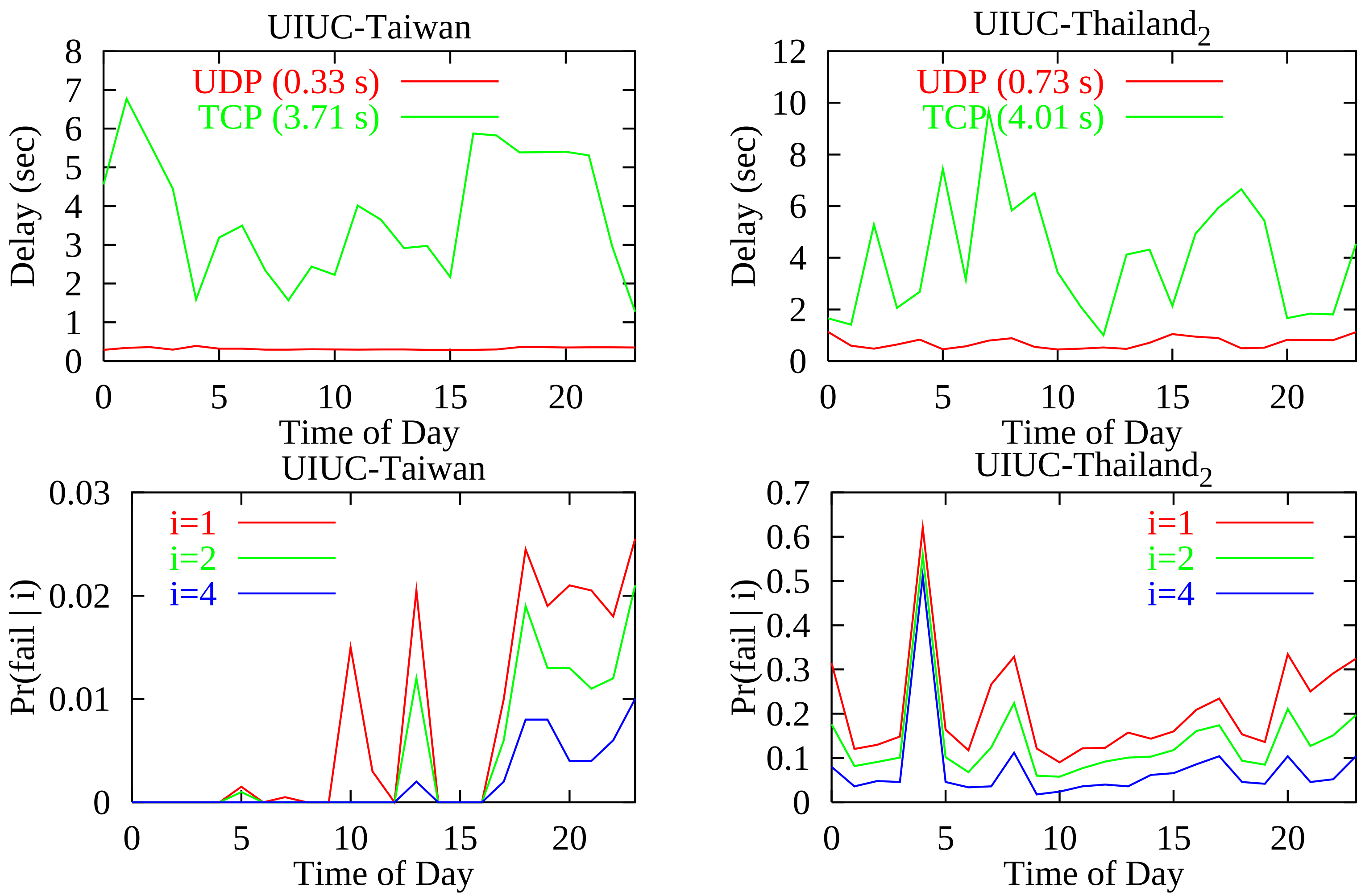
Testbed Setup

- Traces from `cw.crhc.uiuc.edu` to the echo port of three servers

`pager.mit.com.tw`	Low loss	below 5%
`www.iced.moe.go.th`	medium Loss	between 5% to 20%
`www.kmitnb.ac.th`	high Loss	between 20% to 60%

- Statistics on packets bounced back
  - Collected in December 2002
  - UDP sent at 30-ms interval to simulate real-image transmission
- Modified Linux kernel with encapsulation of TCP packets in UDP packets for fair comparison

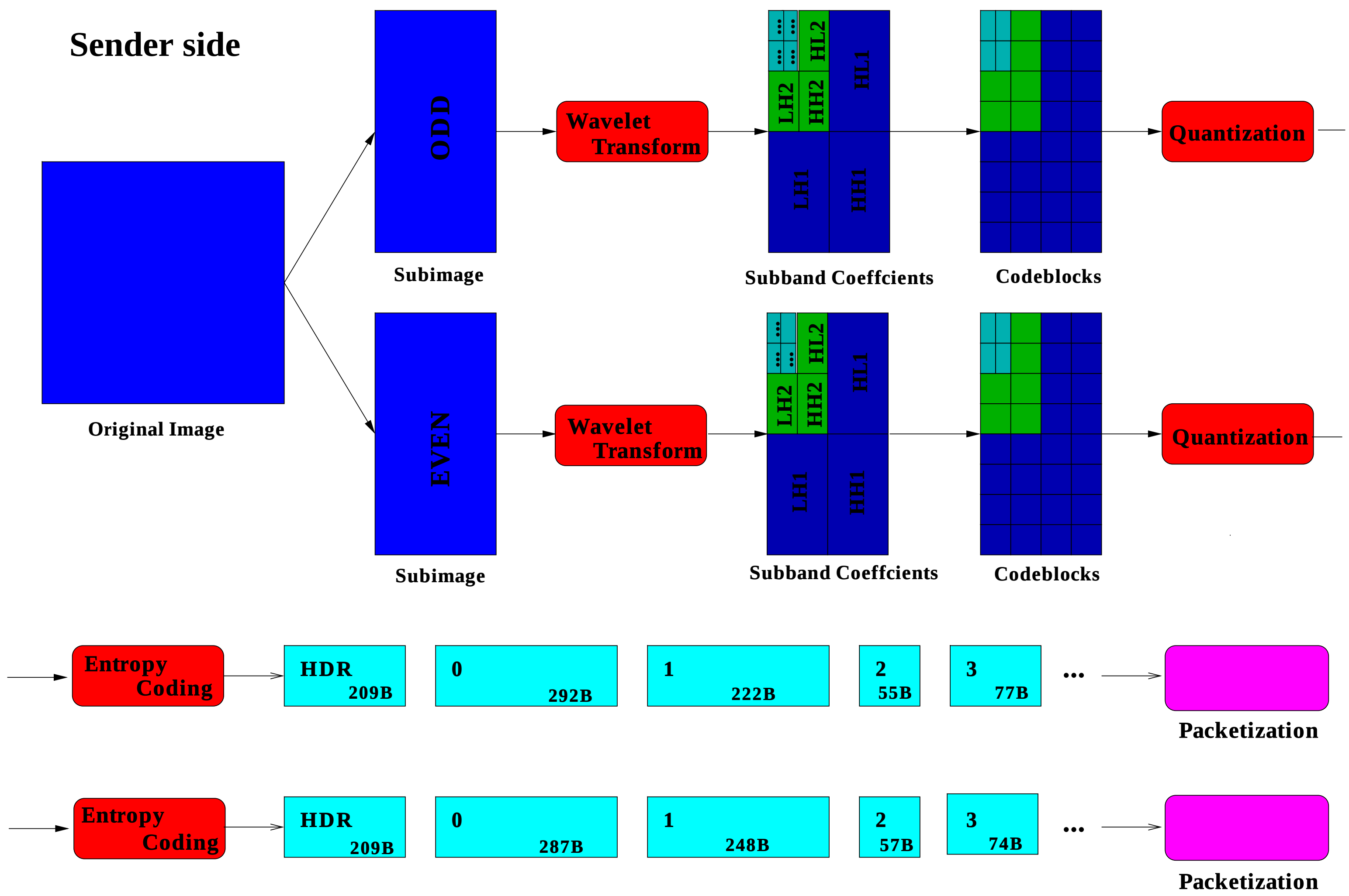
Comparison of TCP and UDP Delays and Losses



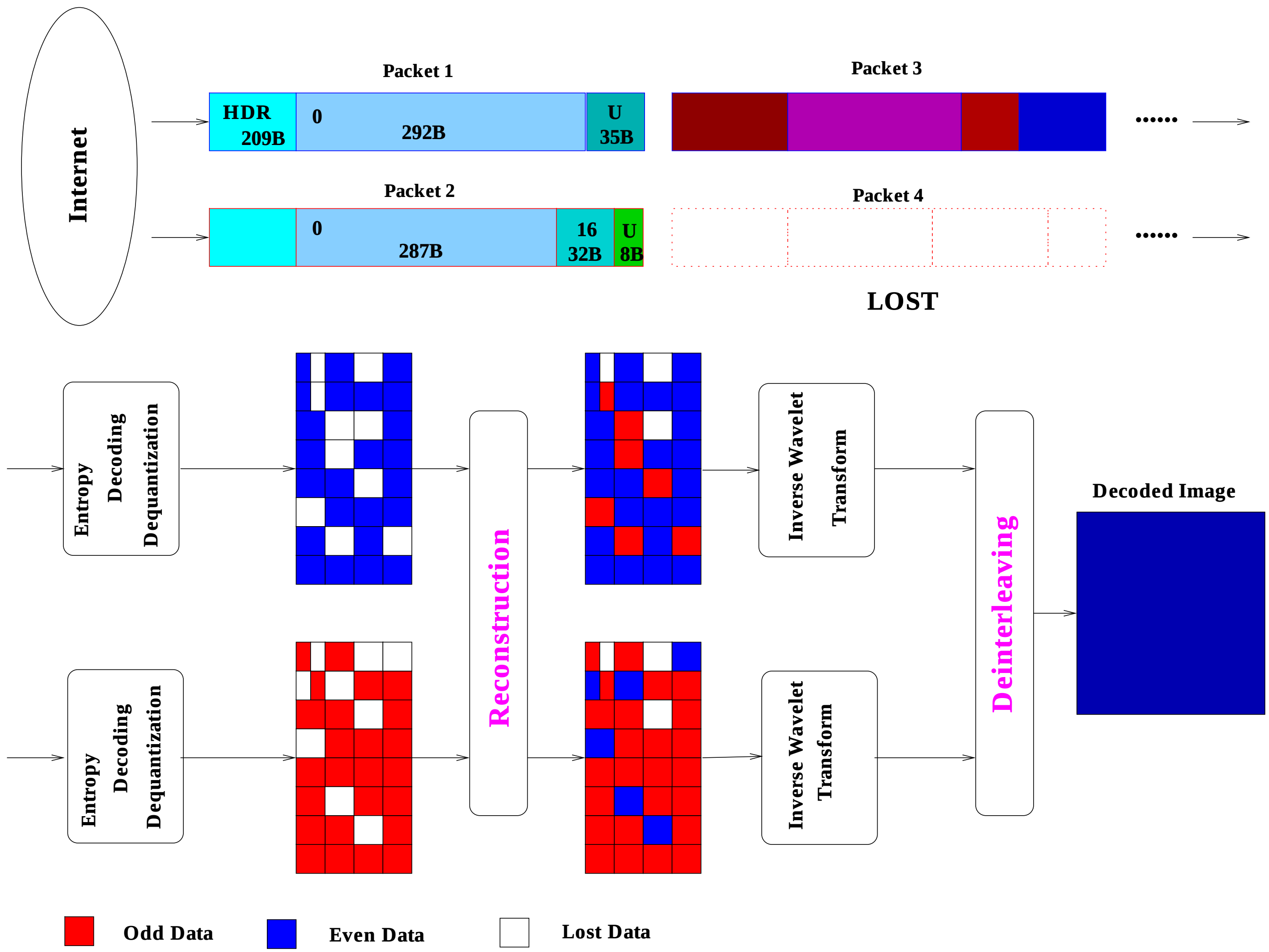
Outline

- Introduction
- Internet Traffic Behavior
- Frequency-Based Loss Concealment
  - Approach
  - Correlation Analysis
  - Optimal Linear Reconstruction
- Experimental Results

Sender-Side Encoding and Packetization



Receiver-Side Reconstruction and Decoding



Correlation Analysis/Reconstruction Performance

Subband	Image *lena*			
	ρ	Average Distortion Per Pixel d_0^2		
		Duplication	Padding-0	Interpolation
Unfiltered	0.972	64.60	1151.86	22.20
LH1	0.369	31.44	25.07	20.54
HL1	0.370	2.12	1.68	1.96
HH1	0.119	3.44	1.96	2.08
LH2	0.820	12.76	42.83	14.64
HL2	0.851	1.06	2.97	1.11
HH2	0.674	2.43	3.73	2.41
LH3	0.954	6.50	72.78	8.43
HL3	0.954	0.41	4.50	0.46
HH3	0.914	1.07	6.28	1.38
LH4	0.992	2.41	160.93	1.65
HL4	0.981	0.21	5.79	0.23
HH4	0.977	0.47	10.80	0.63
LL4	0.999	2.07	811.08	4.01

Correlation and distortion measured after inverse transform of subband

Correlation Analysis/Reconstruction Performance (cont'd)

Subband	Image *teeth*			
	ρ	Average Distortion Per Pixel d_0^2		
		Duplication	Padding-0	Interpolation
Unfiltered	0.993	46.79	3609.21	18.36
LH1	0.367	19.75	15.43	13.28
HL1	0.528	3.16	3.39	2.97
HH1	0.174	4.30	2.61	2.67
LH2	0.827	8.77	25.52	9.37
HL2	0.901	1.57	7.97	1.59
HH2	0.728	2.37	4.37	2.36
LH3	0.956	3.55	40.92	4.67
HL3	0.984	0.61	19.34	0.65
HH3	0.935	1.32	10.06	1.65
LH4	0.989	1.29	61.83	1.77
HL4	0.996	0.27	38.59	0.31
HH4	0.984	0.59	18.79	0.81
LL4	0.999	0.87	3356.90	0.71

Correlation and distortion measured after inverse transform of subband

Sample-Domain Linear Optimal Reconstruction

Given two zero-mean random variables X and Y, the optimal linear reconstruction of X using Y is to find a that minimizes:

$$\min \quad e = E[(\hat{X} - aY)^2]$$

- If X and Y have the same variance, then $a = \rho_{XY}$
- If ρ_{XY} is near to 1, then duplication ($\hat{X} = Y$) is a good approximation

Subband-Domain Optimal Linear Reconstruction

- Decompose each description into two bands

$$X_{even} = X_{even}^{L} + X_{even}^{H}, \quad X_{odd} = X_{odd}^{L} + X_{odd}^{H}$$

  – With assumptions

$$E[X_{even}^{L}] = E[X_{even}^{H}] = E[X_{odd}^{L}] = E[X_{odd}^{H}] = 0$$

$$E[(X_{even}^{L})^2] = E[(X_{odd}^{L})^2] = \sigma_L^2; \quad E[(X_{even}^{H})^2] = E[(X_{odd}^{H})^2] = \sigma_H^2$$

- Two methods for reconstructing X_{odd} from X_{even}
  – min $e_1 = (X_{odd} - aX_{even})^2$
  – min $e_2 = (X_{odd}^{L} - bX_{even}^{L})^2 + (X_{odd}^{H} - cX_{even}^{H})^2$
- $e_2 \leq e_1 \Longrightarrow$ Reconstruction in separate subbands is better

Model to Estimate ρ

- Assuming $d_0 = X_{odd} - X_{even}$, for the i^{th} subband:

$$\rho_i = \frac{E(\mathbf{X}^i_{odd}\mathbf{X}^i_{even})}{\sqrt{E({\mathbf{X}^i_{odd}}^2)E({\mathbf{X}^i_{even}}^2)}} = \frac{E({\mathbf{X}^i_{even}}^2) + E({\mathbf{X}^i_{odd}}^2) - d_i^2}{2\sigma_i^2} = 1 - \frac{d_i^2}{2\sigma_i^2}$$

- Model d_i by geometric relationship:

$$d^2_{LHi} = c_1 b k_1^{(4-i)}; \qquad d^2_{HLi} = c_2 b k_2^{(4-i)}; \qquad d^2_{HHi} = c_3 b k_3^{(4-i)}$$

Subject to

$$d_0^2 = b + \sum_{i=1}^{3} c_i b(1 + k_i + k_i^2 + k_i^3) \qquad \text{where } b = d^2_{LL4}$$

  - 4 image-dependent parameters c_1, c_2, c_3, d_0
  - 3 other parameters k_1, k_2 and k_3 that can be generated beforehand

Experimental Values of k_i's

- Linear regression (with R-square measure) to extract k_i's from each image

Image	HL Band		LH Band		HH Band		Image	HL Band		LH Band		HH Band	
	k_1	R_1^2	k_2	R_2^2	k_3	R_3^2		k_1	R_1^2	k_2	R_2^2	k_3	R_3^2
barbara	3.76	0.97	3.86	0.98	4.61	0.85	*cloth*	4.13	0.80	1.31	0.52	0.97	0.23
boat	3.95	0.99	2.28	0.99	2.54	0.88	*grape*	2.49	0.98	1.97	0.99	1.56	0.80
goldhill	2.88	0.99	2.80	0.99	2.70	0.99	*pines*	3.95	0.97	2.46	0.95	2.50	0.90
lena	2.36	0.99	2.16	0.99	2.04	0.99	*smoke*	2.54	0.99	2.31	0.99	1.97	0.99
peppers	1.82	0.98	2.57	0.96	2.08	0.99	*teeth*	3.79	0.95	2.53	0.87	2.03	0.54
zelda	2.07	0.98	2.27	0.98	1.92	0.82	*thumb*	3.59	0.68	1.75	0.19	1.40	0.02
Group 1	2.69	0.79	2.60	0.80	2.53	0.68	*trick*	3.79	0.99	1.42	0.86	2.05	0.96
Groups 1 & 2	3.05	0.80	2.20	0.68	2.05	0.40	**Group 2**	3.40	0.82	1.91	0.55	1.71	0.22

Parameters not very sensitive to image

Synthetic Experiments (1 out of 2 Lost)

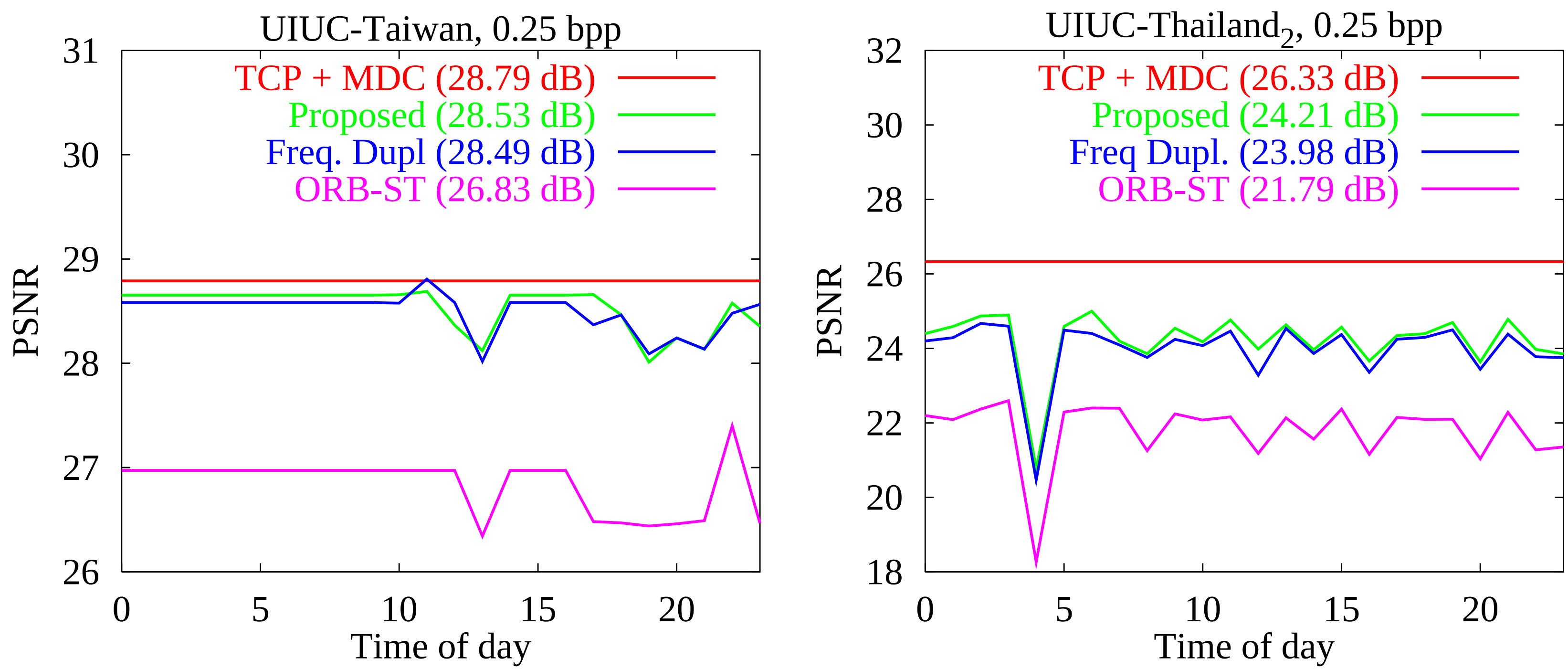
Image	dB Improvement over Full Duplication of Subband Parameters											
	No Quantization				0.5 bpp				0.125 bpp			
	G_u	G_t	G_i	G_0	G_u	G_t	G_i	G_0	G_u	G_t	G_i	G_0
barbara	2.28	2.39	2.57	2.81	2.02	2.02	2.15	2.35	0.77	0.78	0.78	0.84
boat	1.44	1.32	1.42	1.46	1.09	1.00	1.08	1.11	0.28	0.28	0.28	0.29
goldhill	1.02	1.14	1.13	1.15	0.50	0.50	0.49	0.59	0.17	0.17	0.17	0.18
lena	0.90	0.98	0.98	0.98	0.64	0.69	0.70	0.70	0.16	0.17	0.17	0.18
peppers	0.86	0.96	1.09	1.15	0.64	0.64	0.66	0.76	0.25	0.25	0.25	0.28
zelda	1.13	1.25	1.25	1.28	0.70	0.76	0.77	0.78	0.20	0.20	0.20	0.22
cloth	0.26	0.21	0.09	0.49	0.17	0.08	0.00	0.38	0.06	0.06	0.06	0.11
grape	0.68	0.60	0.73	0.75	0.41	0.36	0.45	0.52	0.06	0.06	0.07	0.08
pines	1.06	1.05	1.05	1.08	0.48	0.42	0.41	0.50	0.07	0.07	0.07	0.07
smoke	0.92	0.71	1.02	1.02	0.19	0.19	0.19	0.26	0.03	0.03	0.03	0.03
teeth	0.73	0.69	0.67	0.77	0.48	0.44	0.43	0.51	0.04	0.04	0.04	0.04
thumb	0.05	-0.03	-0.03	0.49	0.08	0.11	0.11	0.39	0.05	0.06	0.06	0.10
trick	1.48	1.61	1.63	1.65	0.79	0.76	0.76	0.83	0.06	0.06	0.07	0.08

G_u: unified k_i's across all images; G_t: image-type dependent k_i's;
G_i: image-dependent k_i's G_0, actual ρ_i's

Outline

- Introduction
- Internet Traffic Behavior
- Frequency-Based Loss Concealment
  - Correlation Analysis
  - Optimal Linear Reconstruction
  - System Architecture Diagram
- Experimental Results

Trace-Driven Results for *Smoke*



An Illustration on *Smoke*

Sending 512×512 *smoke* compressed at 0.25 bpp by JPEG2000

- Between UIUC and Thailand$_1$ (`www.iced.moe.go.th`)
- Five out of the sixteen packets were lost in UDP

SDC and TCP:
PSNR 30.96 dB,
Roundtrip time 13.03 s.

SDC and UDP:
PSNR 22.03 dB,
Roundtrip time 0.46 sec.

Proposed MDC and UDP:
PSNR 28.72 dB,
Roundtrip time 0.46 sec.

Conclusions

- Image transmission involves delay-quality trade-offs
- Proposed frequency-based MDC has good image quality and acceptable delay
- Future Work
  - Develop a coding scheme that can adapt to network conditions
  - Develop more TCP-friendly transmission scheme