ICME2012 Time Machine Session

High Order Entropy Coding – From Conventional Video Coding to Distributed Video Coding

Wenjun (Kevin) Zeng

University of Missouri

http://people.cs.missouri.edu/~zengw/

Acknowledgements: Wei Liu, Lina Dong, Yixuan Zhang, Ce Zhu

An Old Topic!

But a new life!

- Exploiting high order statistics
  - Context modeling and conditional entropy coding
- Coding performance (PSNR) for 512x512 "lena"

Year	1991	1986-1992	1992	1993	1996	1997	1997	1997	2000
Rate (bpp)	JPEG	**Antonini '92 (*)**	ECSBC (ECVQ)	**EZW**	SPIHT	SFQ	EQ	Wu	JPEG 2000
0.21		29.11							
0.25	31.6		33.98 (0.32)	**33.17**	34.13	34.33	34.57	**34.81**	~34
0.37		30.85							
0.5	34.9		35.96 (0.52)	36.28	37.24	37.36	37.68	37.92	
1.00	37.9			39.55	40.45	40.52	40.88	40.85	

***: Close to Woods & O'Neil'86 (SB+DPCM), Westerink'89 (SB+SQ)**

Video Coding: from Hybrid to Distributed

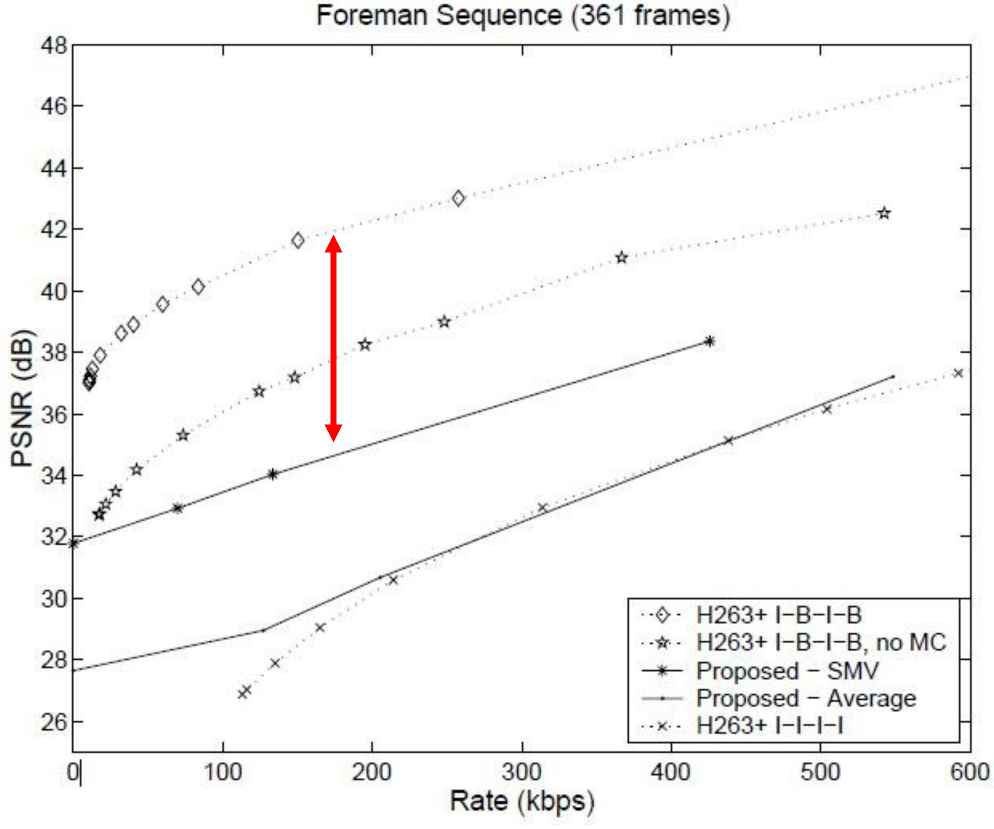
- New applications demand paradigm-shifting approaches
  - Low-complexity wireless video, video surveillance and camera arrays, sensor networks, compression of encrypted data, etc.
  - New structure:
    - **Simple encoding**, but can afford complex decoding
    - **Distributed** encoders/processing
- **Distributed video coding**
  - Theoretical foundations: DSC, laid in 1970s
  - Very good codes for **ideal i.i.d.** sources found (since 1999)
  - How about **real-world** sources such as image/video? (since 2001)

	Coding Efficiency	Encoder Complexity	Decoder Complexity	Error Resilience
Intra	Low	Low	Low	High
Hybrid	Highest	High	Medium	Low
Distributed	**?** ☺	Low	High	High

DVC Performance (2002)

Aaron et al' 2002, **r-d for WZ frames only** (direct quote)

5 to 7 dB gap from H.263+ **inter** coding



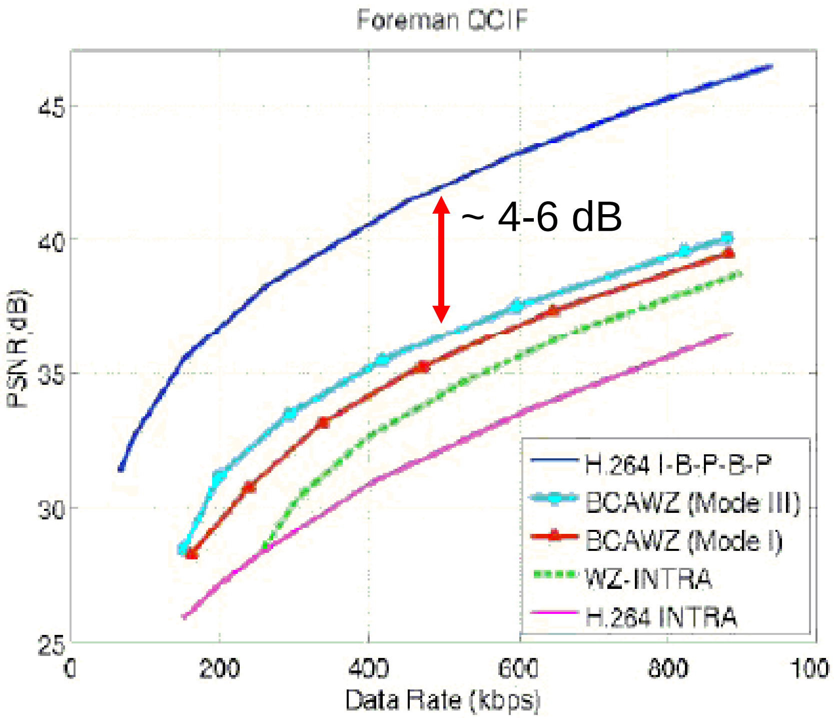
DVC Performance (2007)

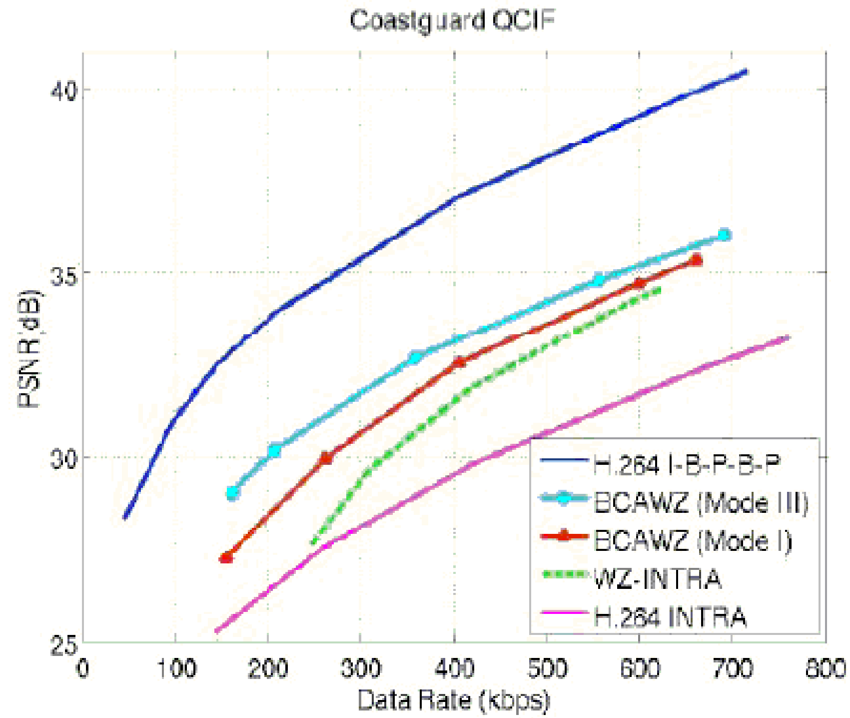
Ed. Delp, Nov. 2007, Discover Workshop (direct quote)

(odd frame: Decoder ME, MV sent back to encoder through backward channel
even frame: WZ coded)

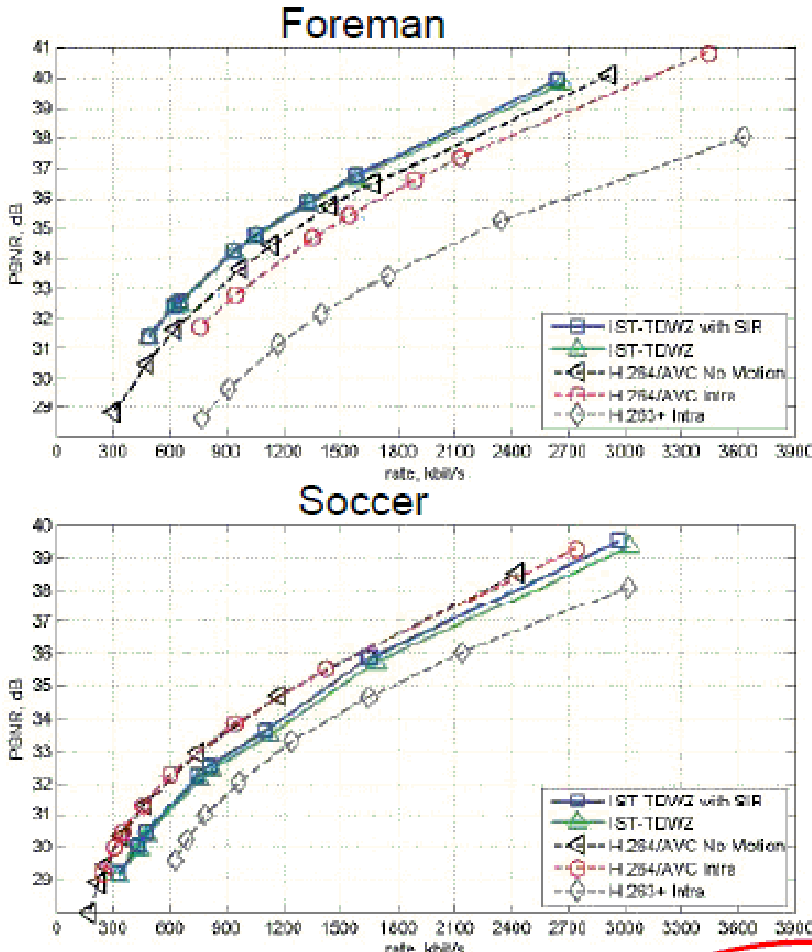
Real gap could be 6 - 8 dB

4 to 6 dB gap (**diluted!**) from H.264 **inter** coding





DVC Performance (2009)



RD performance , CIF, 30 Hz, GOP size 2

F. Pereira, ICME'09 (direct quote)

"it can be concluded that the SIR based IST-TDWZ codec performs better than H.264/AVC **Intra** for the sequences with **low motion** content, notably for the **lower GOP sizes**."

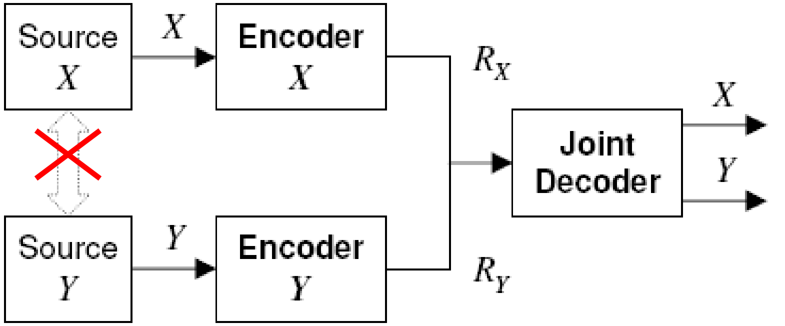
What is wrong with the field???

Conventional Video Coding

- Major tools
  - Motion compensated prediction
  - Spatial transform
  - High order entropy coding
    - run-length coding, end-of-block symbol
    - zero-tree coding,
    - context based coding

Distributed Source Coding (Lossless)

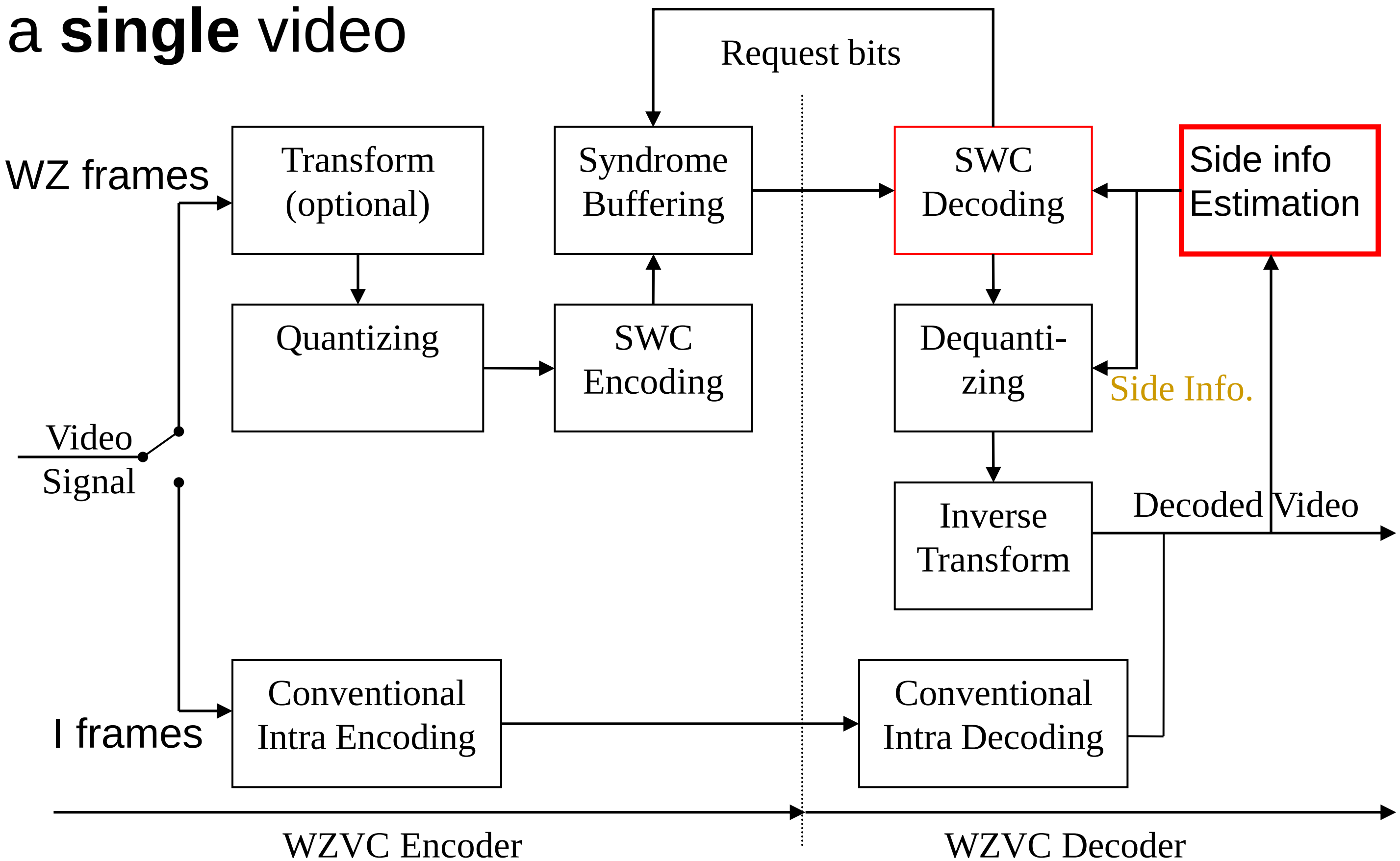
- Slepian and Wolf, 1973 (SWC)
  - No rate loss even if encode separately
- A practical approach: using **Forward Error Correcting** Code (Wyner, 1974)
  - $R_Y = H(Y)$, $R_X = H(X|Y)$ (treat Y as noisy version of X)
- Recent advances in channel codes produces practical SWC with capacity-approaching performance
  - Turbo codes, LDPC codes



X: *i.i.d.; Y: i.i.d.*

A Generalized WZVC Architecture

For a **single** video

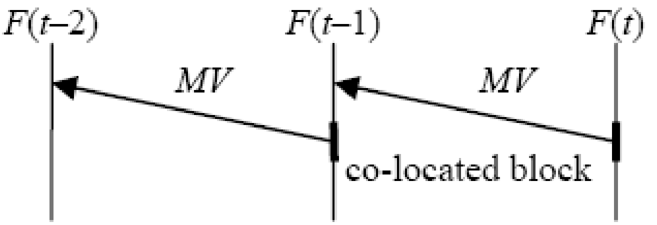


Wyner-Ziv **Video** Coding

- Wyner-Ziv Video Coding (WZVC)
  - Shift the computational burden (exploiting redundancy) to the decoder side
  - The decoder usually does
    - Generate side information Y
    - Estimate the (noisy) channel statistics between Y and the frame to be decoded (X)
    - Wyner-Ziv (channel) decoding
  - Better performance of WZVC can be achieved if any of the above aspects is improved

Common Assumptions and Models

- Stationary motion field (?)
  - interpolation/extrapolation of motion vectors from neighboring decoded frames

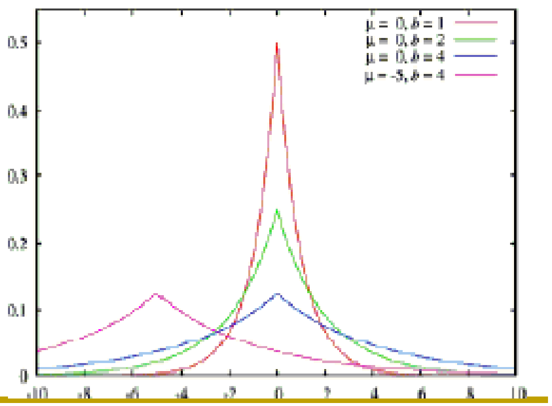


Example side info frame

Common Assumptions and Models

- Correlation between source and side information (i.e., virtual channel model)
  - Virtual channel noise: **independently** distributed (?)
  - Laplacian distribution
    - Variance estimated from previous decoded frames

$$\frac{1}{2b}\exp\left(-\frac{|x-\mu|}{b}\right)$$



Progressive-learning-based DVC

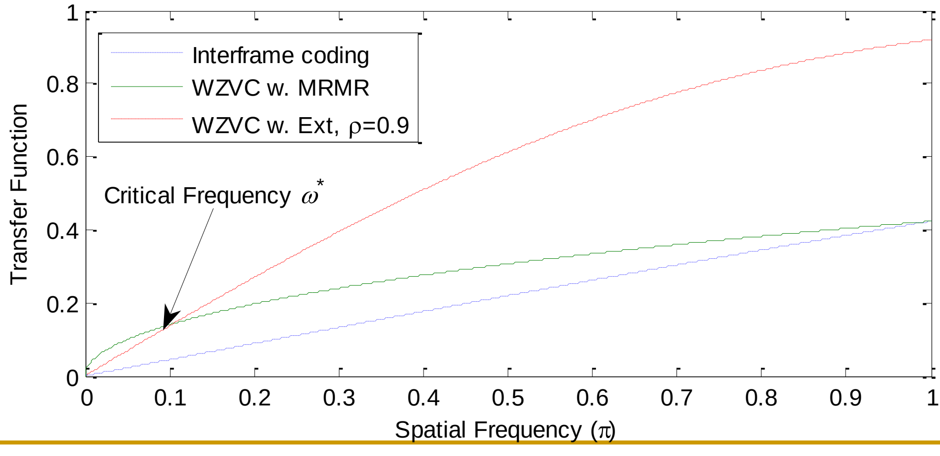
- Allow decoder learning
  - Learn from lower/earlier layers and use for higher/later layers
- Allow better side info estimation
- Better exploitation of **local** redundancy at the decoder
- Allow employment of **symmetric** DSC
- Scalability

WZVC with Multi-Resolution Motion Refinement (MRMR) (Liu et al'09,10)

- Basic idea
  - Multi-resolution:
    - Resolution-progressive packetization and decoding
    - From low-frequency to high-frequency components.
  - **Motion Refinement**:
    - Decode a low-resolution frame using current motion field
    - Refine the motion field based on partially decoded frame
    - Go to the next resolution level, repeat the process
- **Comprehensive R-D analysis**
- A wavelet-domain codec

Performance Comparison

- Recall $\Delta R = R_{MCP} - R_{intra} \approx \frac{1}{8\pi^2} \iint \log_2 \left[1 - \exp\left(-\omega^T \omega \sigma_{\Delta d}^2 \right) \right] d\omega$
  - An equivalent transfer function $\sqrt{1 - \exp\left(\omega^T \omega \sigma_{\Delta d}^2 (\omega) \right)}$
  - **Smaller value → better performance**
- Comparison
  - Encoder inter-frame coding is always the best.
  - MRMR is the worst in very low frequency, but overall much better than motion extrapolation (even for a large ρ).



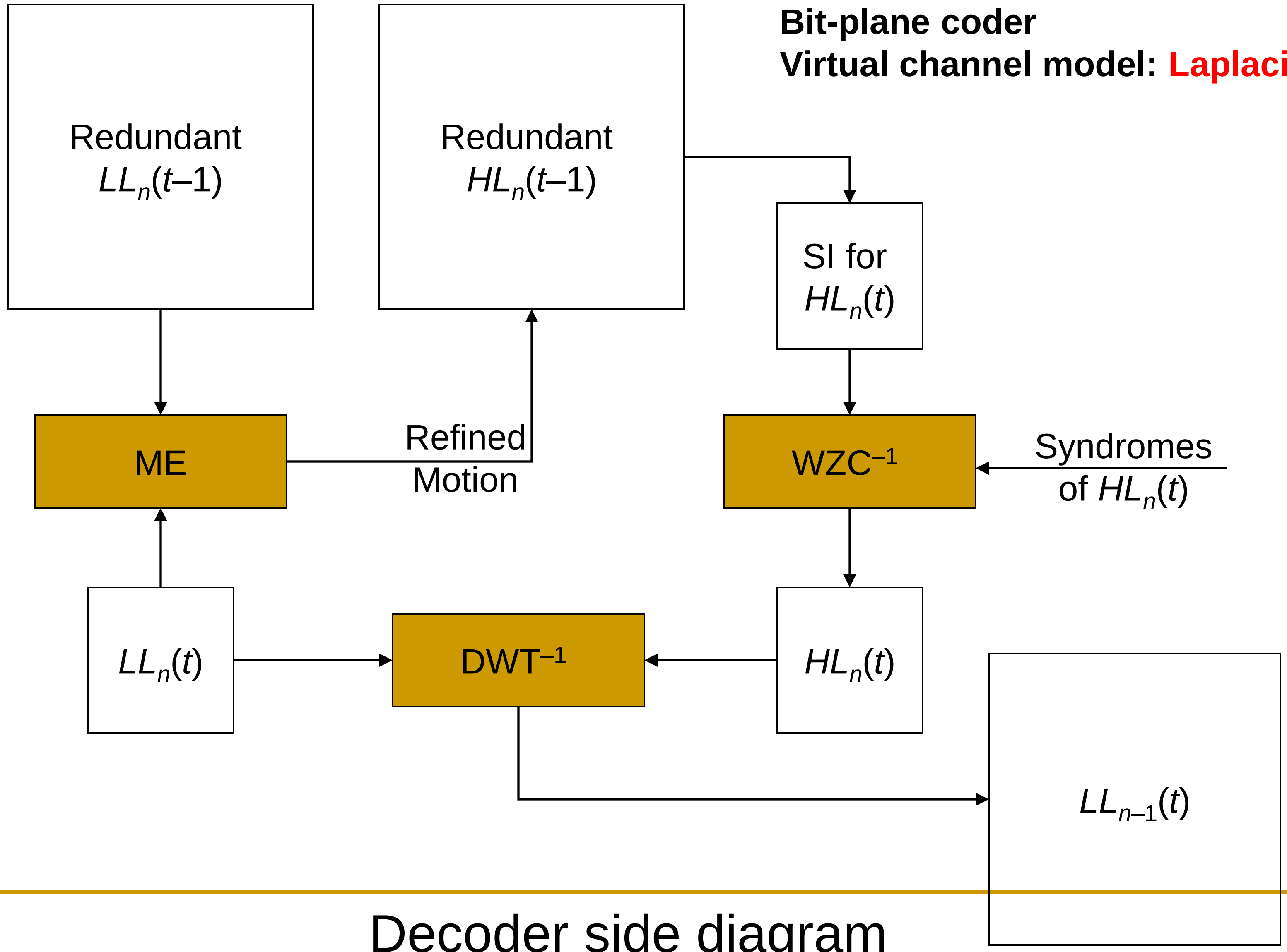
Result: Sample SI Estimation (Residue)

Motion Extrapolation

MRMR

Wavelet Domain WZVC with MRMR

Bit-plane coder
Virtual channel model: Laplacian



Decoder side diagram

Result: MRMR vs. H.264/AVC Predictor

GOP pattern: first frame I, all rest P/WZ; only one reference frame, QCIF

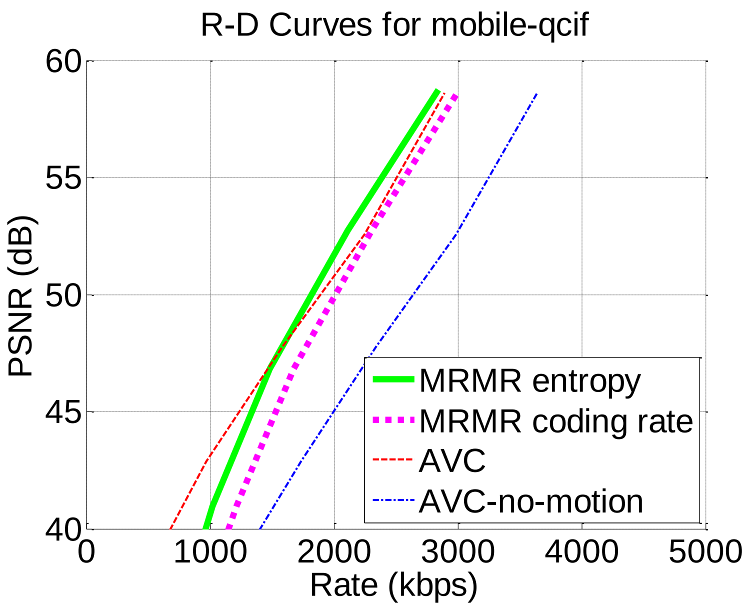
Saved rates	Akiyo	Car phone	Con-tainer	Foot-ball	Fore-man	Miss-Am	M&D	News	Suzie	Sales-man	**Avg.**
MRMR	–2.89	–1.70	–3.48	–0.42	–1.88	–1.46	–2.09	–2.80	–1.10	–2.44	**–2.03**
AVC	–2.58	–1.88	–3.04	–0.75	–2.03	–1.53	–2.12	–2.82	–1.31	–2.54	**–2.06**

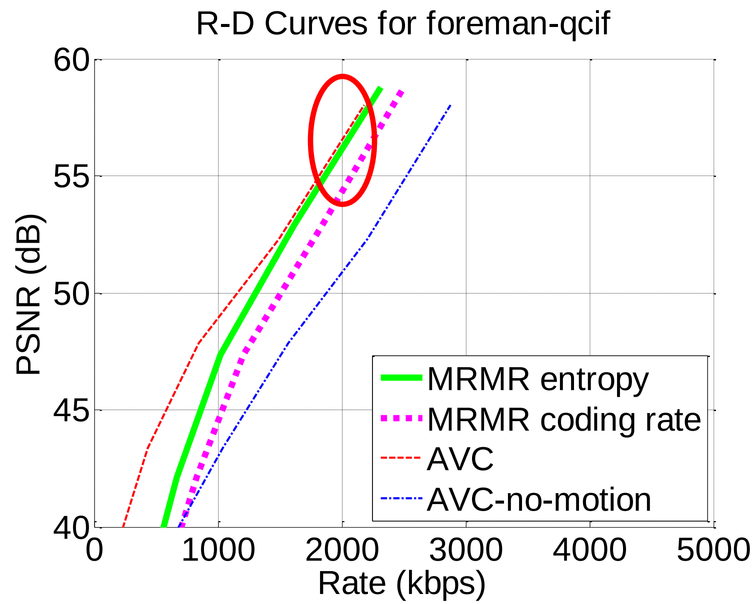
- For MRMR, $B = 2$, $q = 1/4$, $M = 8$ (**extensive motion exploration**).
- For AVC baseline, $B = 4$, $q = 1/4$, $M = 1$.

- Estimate rate saving over intra-coding
  - Calculate $\Delta R = \frac{1}{2}\log_2 \frac{MSE_1}{MSE_0}$, for each subband, then do a weighted average.

- Very small gap is observed.

High Rate Coding Comparison





Y components of the **P/WZ-frames only**
I followed by all P @30fps

Distributed Video Coding (DVC)

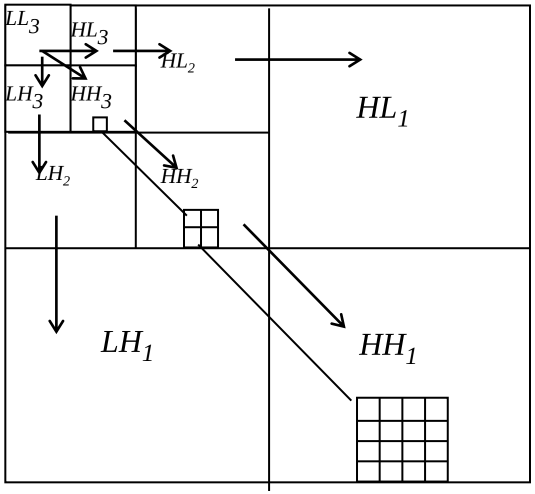
- Simple encoder, complex decoder
  - burden of removing redundancy shifted to decoder
    - **Hard!**
- Major steps
  - Side information estimation (**at decoder**)
    - ill posed problem when the source is non-stationary
    - Major obstacle
    - NOT solved until recently
  - Entropy coding
    - Channel coding approach
      - Much less flexible
    - Little effort so far for high order entropy

High Order Entropy Coding **for DVC**

- Now is the time !
  - Side info estimation well understood/tackled
- How?
  - Design specific channel code for virtual channel
    - Complicated (Hard!)
  - Context modeling and **bit-level** conditional probability
    - Fit channel coding paradigm well
    - Borrow ideas from conventional image/video coding!!!

Embedded Zerotree Wavelet Coding

(Shapiro'93)



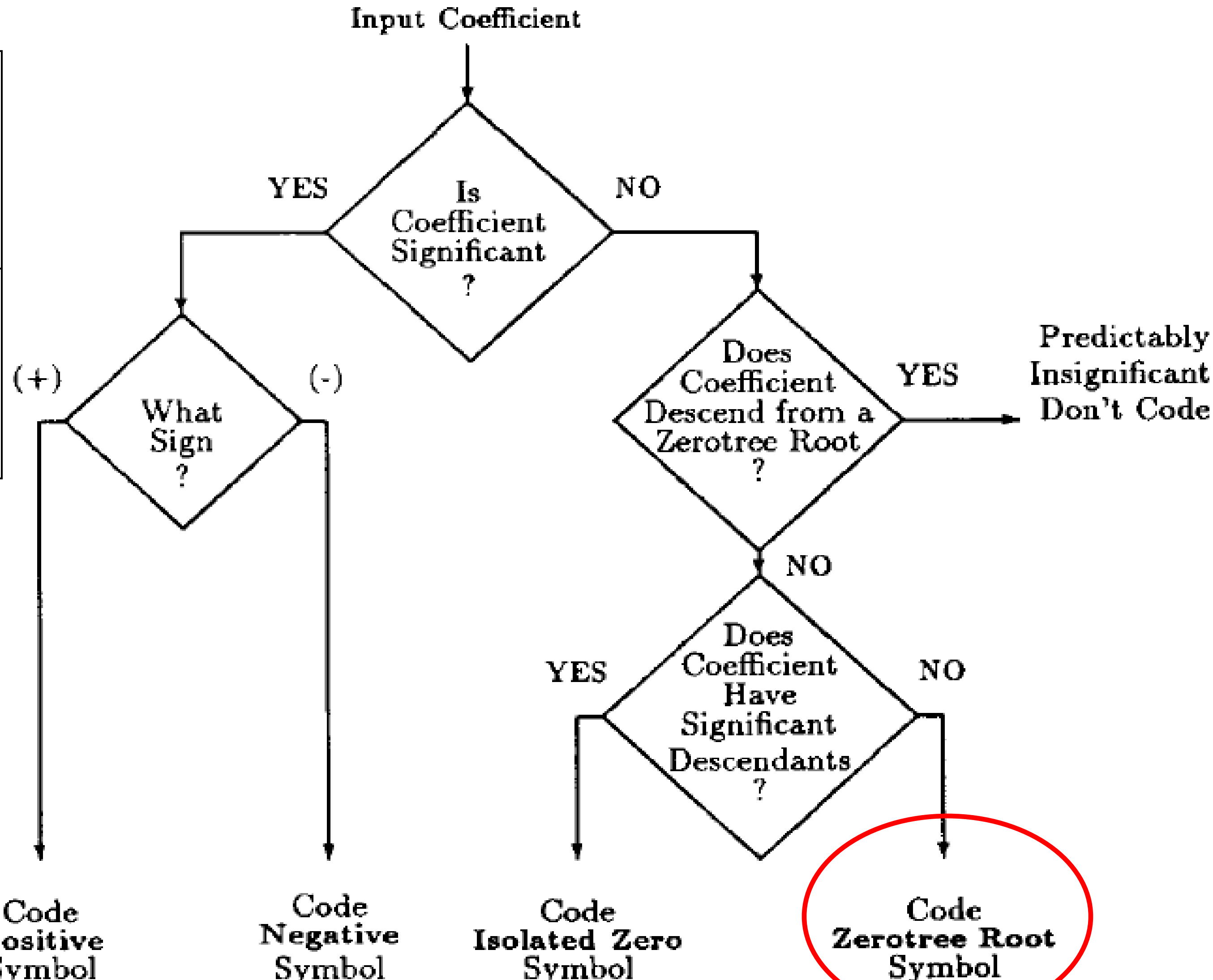
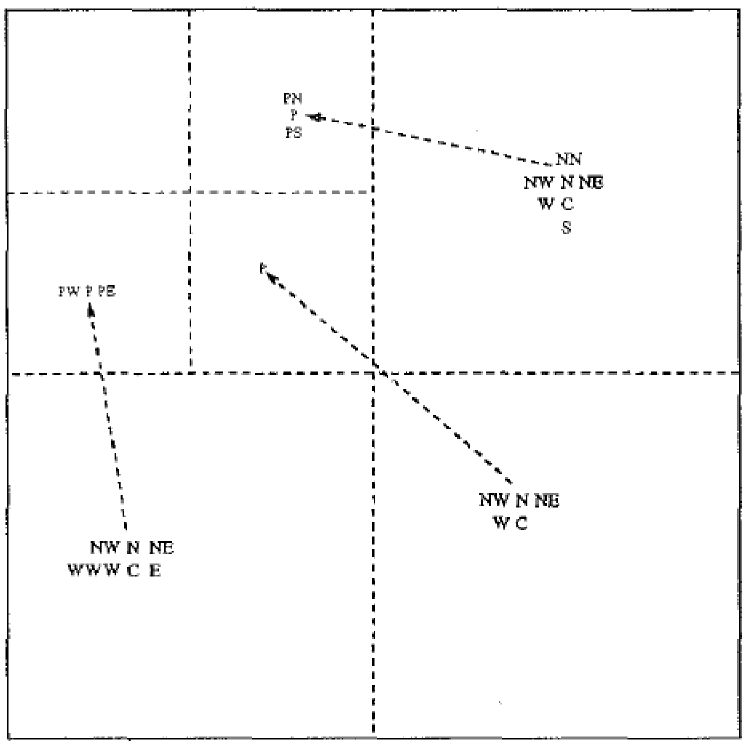


Fig. 6. Flow chart for encoding a coefficient of the significance map.

- Can we use zero-tree in DVC???
  - It relies on **direct** coding of the tree.
  - Not straightforwardly!
  - But concept could be borrowed.

Embedded Conditional Entropy Coding (Wu'97)

- Modeling contexts in different sub-bands



- High order statistical model at bit-level:

$$\hat{P}(C_b \mid N_{m..b}, W_{m..b}, S_{m..b+1}, E_{m..b+1}, NW_{m..b}, NE_{m..b}, P_{m..b} \cdots).$$

- Learn on the fly

JPEG2000: Context-based Arithmetic Coding

Neighboring States used to Form Context

D_0	V_0	D_1
H_0	X	H_1
D_2	V_1	D_3

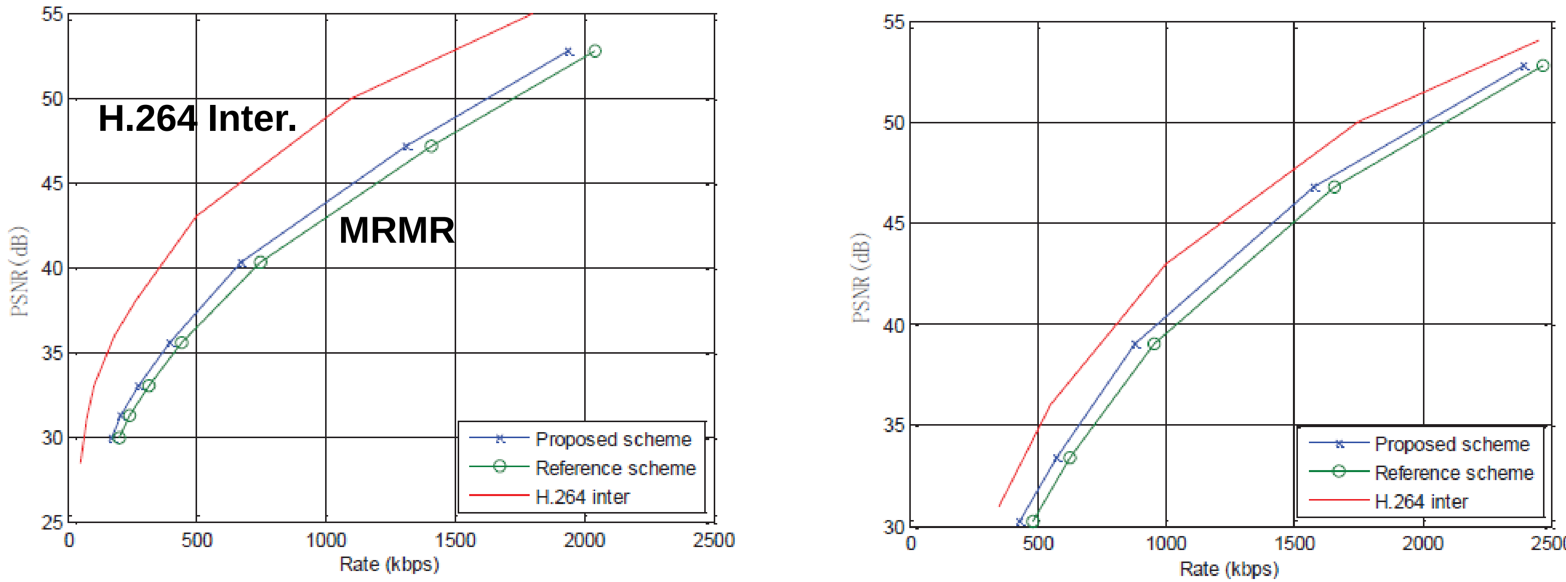
- Each coefficient in a code-block has an associated binary state variable called its ***significance state***.

Exploiting High Order Correlation **for DVC**

(Yixuan Zhang, Ce Zhu, Wenjun Zeng' 2012)

- Exploit the inter-coefficient correlation across scales in the wavelet domain
- Modeling of the distribution of $p(x_n|x_{n+1})$ for current symbol x_n given its parent x_{n+1}

1 dB gain, still 4 - 5 dB from H.264 Inter.



(a) "Foreman"

I followed by all P/WZ.

(b) "Mobile"

An Old Topic!

But a new life!

~ 3 dB

- Coding performance (PSNR) for 512x512 "lena"

Year	1991	1986-1992	1992	1993	1996	1997	1997	1997	2000
Rate (bpp)	JPEG	**Antonini '92 (*)**	ECSBC (ECVQ)	**EZW**	SPIHT	SFQ	EQ	Wu	JPEG 2000
0.21		29.11							
0.25	31.6		33.98 (0.32)	**33.17**	34.13	34.33	34.57	**34.81**	~34
0.37		30.85							
0.5	34.9		35.96 (0.52)	36.28	37.24	37.36	37.68	37.92	
1.00	37.9			39.55	40.45	40.52	40.88	40.85	

***: Close to Woods & O'Neil'86 (SB+DPCM), Westerink'89 (SB+SQ)**

High Order Entropy Coding for DVC

- Now is the time !
  - Side info estimation well understood/tackled
- From Conventional to Distributed Video Coding
  - **Leverage the past for the future !**

Thanks

Q & A

References

- J. W. Woods and *S .* D. O'Neil, "Subband coding of images." *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-34. pp. 1278-1288, Oct. 1986.
- P. H. Westerink, J . Biemond, and D. E. Boekee, "Progressive transmission of images using subband coding," in *Proc. IEEE ICASSP,* 1989 , pp. 1811 -1814.
- M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image Coding Using Wavelet Transform," *IEEE Trans. Image Processing*, vol. I , no *2.* April 1992.
- Y. H. Kim and J. W. Modestino, "Adaptive entropy coded subband coding of images," *IEEE Trans. Image Processing,* vol. I , pp. 31-48, Jan. 1992.
- J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients", *IEEE Trans. Signal Processing,* vol. 41, no. 12, pp. 3445-3462,Dec. 1993.
- A. Said and W. A. Pearlman, "New, fast, and efficient image codec based on set partitioning in hierarchical trees", *IEEE Trans. Circ.* & *Sys. Video Tech.,* vol. 6, no. 3, pp. 243-249, June 1996.
- Z. Xiong, K. Ramchandran, and M. T. Orchard, "Space frequency quantization for wavelet image coding", *IEEE Trans. Image Processing*, 1997.
- S. M. LoPresto, K. Ramchandran, and M. T. Orchard, "Image coding based on mixture modeling of wavelet coefficients and a fast estimation-quantization framework', *Proc.* of *I997 Data Compression Conference,* pp. 241-250, Mar. 1997.
- X. Wu, "High-Order Context Modeling and Embedded Conditional Entropy Coding of Wavelet Coefficients for Image Compression," Thirty-First Asilomar Conference on Signals, Systems & Computers, 1997.

References

- S. S. Pradhan and K. Ramchandran, “Distributed Source Coding Using Syndromes (DISCUS): Design and Construction,” *Proc. of the Data Compression Conference* (DCC), March, Snowbird, UT, 1999.
- A. Aaron, R. Zhang, and B. Girod, “Wyner-Ziv Coding of Motion Video,” *Asilomar Conference on Signals, Systems & Computers*, 2002.
- B. Girod, A. Aaron, S. Rane and D. Rebollo-Monedero **,** “Distributed video coding,” *Proceedings of the IEEE*, Special Issue on Video Coding and Delivery*,* vol. 93, no. 1, pp. 71–83, January 2005.
- E. J. Delp, “Distributed Video Coding: A Technique or a Toolbox?,” Discover Workshop, “Recent Advances in Distributed Video Coding,” November 6, 2007, Lisbon.
- F. Pereira, “Distributed video coding: Basics, main solutions and trends,” *IEEE Inter. Conf. Multimedia & Expo,* 2009.
- W. Liu, Lina Dong and W. Zeng*, “*Motion Refinement Based Progressive Side-Information Estimation for Wyner-Ziv Video Coding,” *IEEE Trans. on Cir. and Sys. for Video Technology*, Dec. 2010.
- Y. Zhang, C. Zhu, and W. Zeng, “Exploiting High-Order Correlation in Distributed Video Coding,” submitted, 2012.