Rensselaer

CNGV

Error Resilient Compression and Transmission of Scalable Video

William A. Pearlman

Professor and Director,

Center for Next Generation Video

Rensselaer Polytechnic Institute

E-mail: pearlman@rpi.edu

http://www.cipr.rpi.edu/staff/pearlman.html,
http://www.rpi.edu/web/NGV

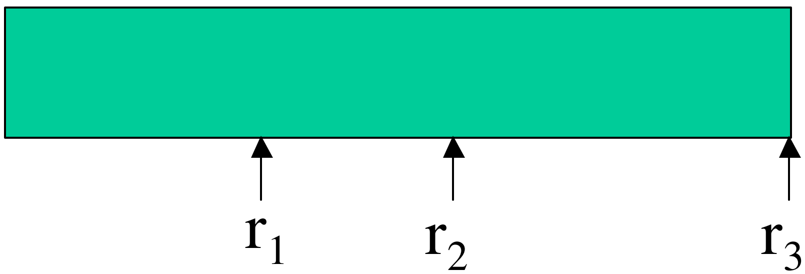
VUB
28 June 2001

Features Requirements

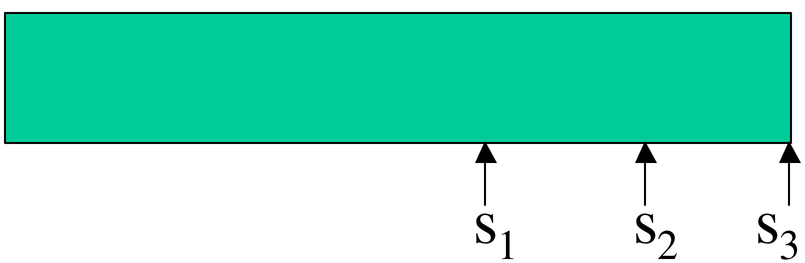
- High performance
  - low MSE (high PSNR) for given rate
- fast, simple encoding and decoding
- low memory usage
  - independent of image/video frame size
- random access

Features Requirements: Scalability

Rate or PSNR scalability: lower rate files embedded in full compressed file



Spatial (resolution) scalability:
lower resolution files
embedded in full compressed file

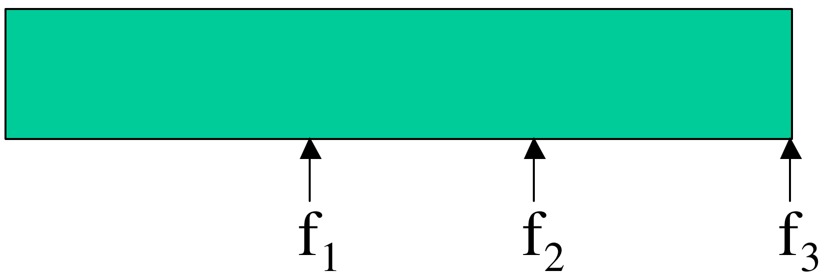


s_2

s_1

Scalability (cont.)

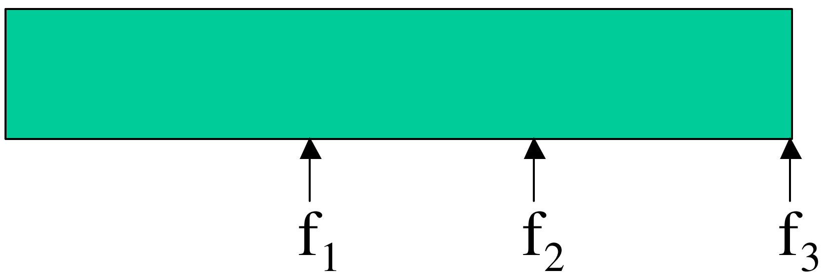
Video frame rate scalability: lower frame rates embedded in full compressed file



f_1 -> every 4th frame
f_2 -> every 2nd frame
f_3 -> every frame

Scalability (cont.)

Video frame rate scalability: lower frame rates embedded in full compressed file



f_1 -> every 4th frame
f_2 -> every 2nd frame
f_3 -> every frame

Transmission Features

- Robust to channel errors/packet loss
  - unprotected bit stream degrades gracefully
  - no overhead required for robustness
- Resilience to channel errors/packet loss
  - redundant bits added for error-correcting channel code
  - video reception of acceptable quality in severe environments

Rensselaer

CNGV

Basis: SPIHT

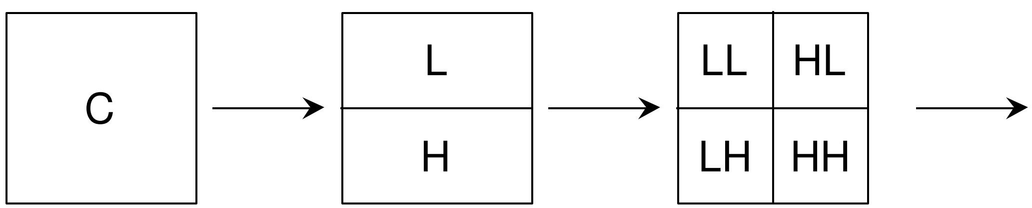
- State of the art performance
  - SPIHT is the modern-day benchmark
- Low in complexity
  - efficient and fast
- Fidelity embedded
  - scalable in rate with precise rate control
- Coding of wavelet subbands
  - capable of progressive resolution

SPIHT (cont.)

- Sensitive to channel errors
  - error in significance decision bit causes catastrophic degradation for rest of bit stream
  - error in refinement bit causes graceful degradation only

Rensselaer

Image Wavelet/Subband Transform



SPIHT Algorithm

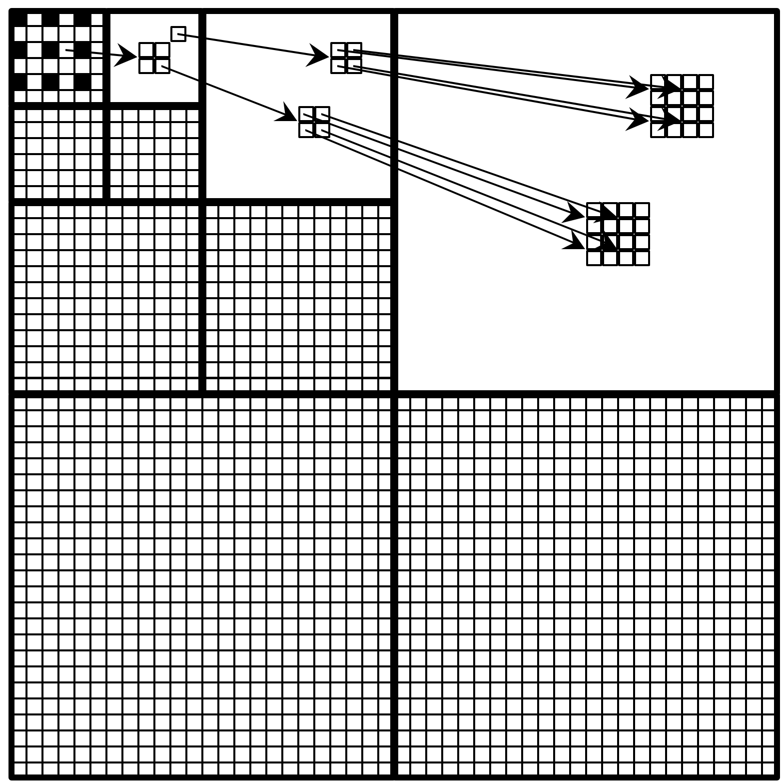
- Discovered by Said and Pearlman (IEEE CSVT 1996)
- Partitions spatial orientation tree sets
- The algorithm maintains three lists of sets
  - LIP: list of coordinates of insignificant pixels
    - initialized by highest level low-pass (DC) subband
  - LIS: list of coordinates of insignificant sets and their type (D or G)
    - initialized by coordinates in DC subband with descendants as D type
  - LSP: list of coordinates of significant pixels
    - initially empty

Rensselaer

SPIHT Tree Structure

3 level dyadic subband transform shown
Arrows depict parent-child relationships in the SPIHT tree structure
Coefficients are grouped to exploit magnitude dependence
Each subband coefficient has four children
Some coefficients in the DC subband have no children

Each coefficient is denoted by it coordinates (i,j)



Set Types

(i,j): Single coefficient
C(i,j): Children of (i,j)
D(i,j): Descendants of (i,j)
G(i,j): D(i,j) - C(i,j)

Rensselaer

Set Partitioning

Significance test for pixels $c_{i,j}$ in set B

$$S_n(B) = \begin{cases} 1, & \text{if } \left(\max_{(i,j)\in B} \left|c_{ij}\right|\right) \geq 2^n, \\ 0, & \text{otherwise.} \end{cases}$$

Coding Sequence

- SPIHT uses three lists to minimize the number of tests for a given bit-plane (n).
  - LIP (List of Insignificant Pixels) – visit pixels that are insignificant and do not belong to insignificant sets.
  - LIS (List of Insignificant Sets) – visit all the sets (with more than one pixel) that are insignificant but do not belong to a larger insignificant set.
  - LSP (List of Significant Pixels) – send n-th bit of all pixels found to be significant in previous passes.

Rensselaer

Sorting by Magnitude and Bit-Plane Transmission

Transmission of magnitude-sorted coefficients

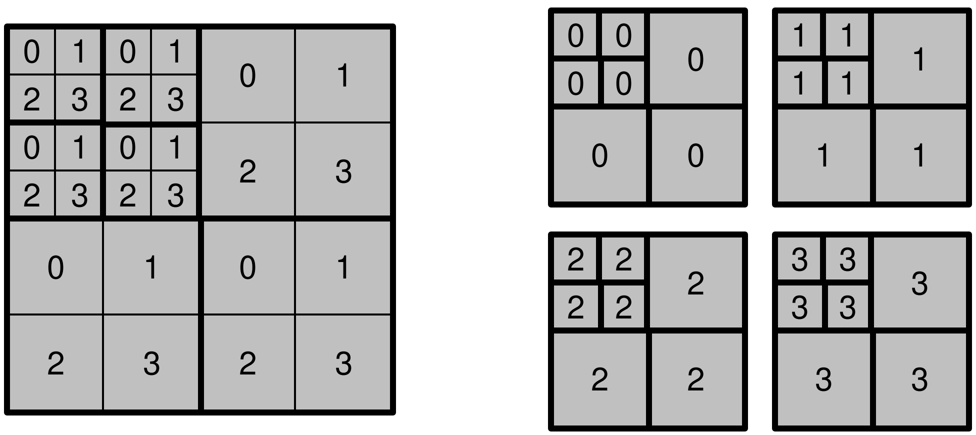
	sign	s	s	s	s	s	s	s	s	s	s	s	s	s
msb	5	1	1	0	0	0	0	0	0	0	0	0	0	0
	4	→	→	1	1	0	0	0	0	0	0	0	0	0
	3	→	→	→	→	1	1	1	1	0	0	0	0	0
	2	→	→	→	→	→	→	→	→	1	1	1	1	1
	1	→	→	→	→	→	→	→	→	→	→	→	→	→
lsb	0	→	→	→	→	→	→	→	→	→	→	→	→	→

Spatial (Tree) Blocks

- SPIHT coding applied to small transform tree-blocks independently
  - these *spatial blocks* in transform domain correspond to spatial regions plus overlap area in image domain
- Embedded bit streams from independent tree-block encoders are packetized and multiplexed to generate a progressive main bit stream

Tree-Block Partitioning

- Works with line- or block-based wavelet transform
- Transform will deliver subband partition corresponding to image block
- Subband partition is organized into tree-blocks, if necessary
- SPIHT encoder is applied to trees
- Example shows a 2 level subband partition organized into 4 trees forming a tree-block



In actual experiments a 5 level transform is used and partitioning generates 128 by 128 or 64 by 64 tree-blocks

Image Sequence Coding

- 3D Wavelet Subbands of Groups of Frames
  - 4, 8, 16 frames
- Temporal (axial) subbanding decorrelates and compacts energy into low frequency bands
  - *no motion estimation/compensation*

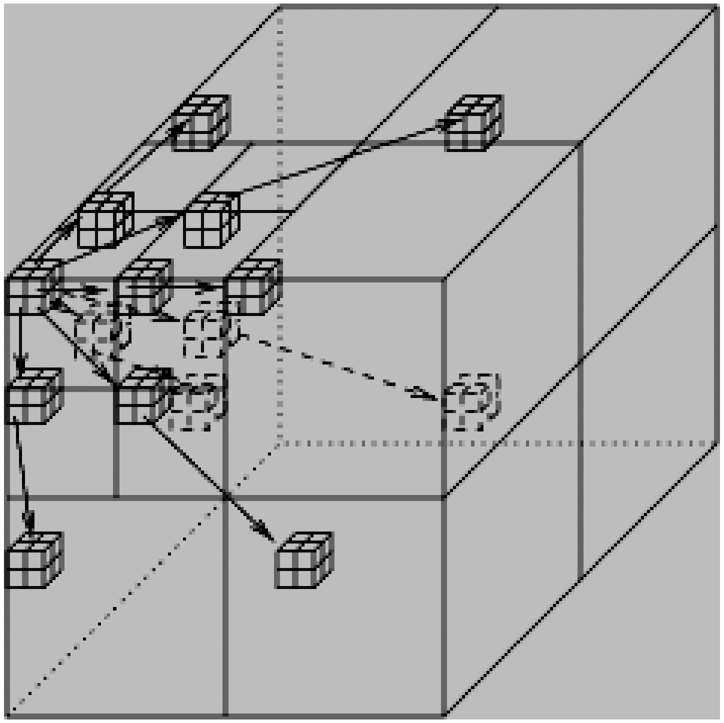
Image Sequence Coding (cont.)

- SPIHT coding on 3D s-t orientation trees
- Independent coding of s-t blocks for robust bitstream transmission
- Packet interleaving for fidelity embedding
- Channel coding of packets for error resilient reception
- Selective decoding in received bit stream for desired rate, spatial resolution, and frame rate

Rensselaer

CNGV

3-D S-T Orientation Tree



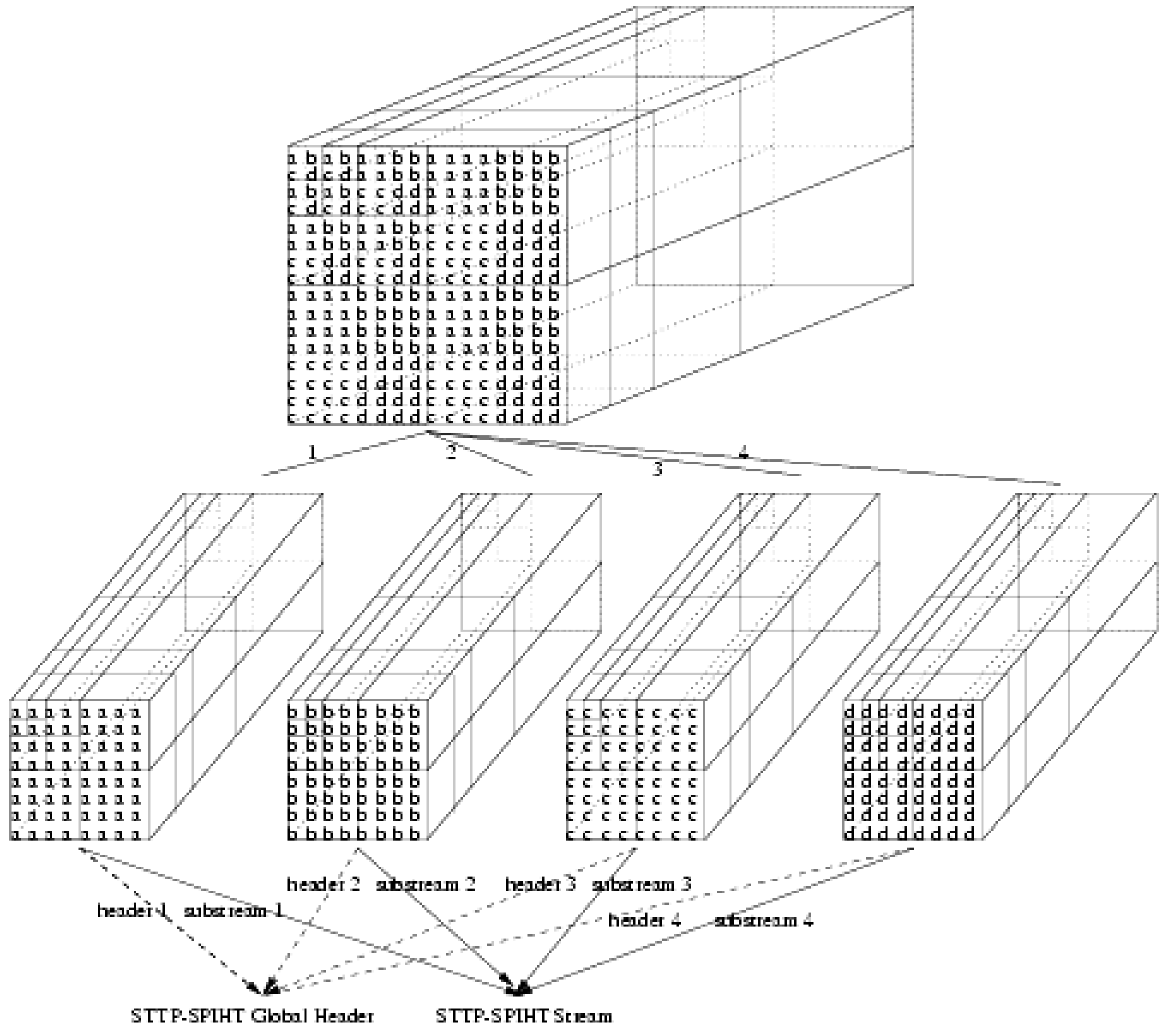
STTP-SPIHT

- Break transform into n groups of spatio-temporal trees, called spatio-temporal (s-t) blocks (corresponds to s-t region in sequence
- Encode (decode) each s-t block independently
- Channel error or packet loss affects only one block and hence one s-t region

Rensselaer

CNGV

Spatio-Temporal Tree Preserving SPIHT



VUB
28 June 2001

Progressive Bit Stream

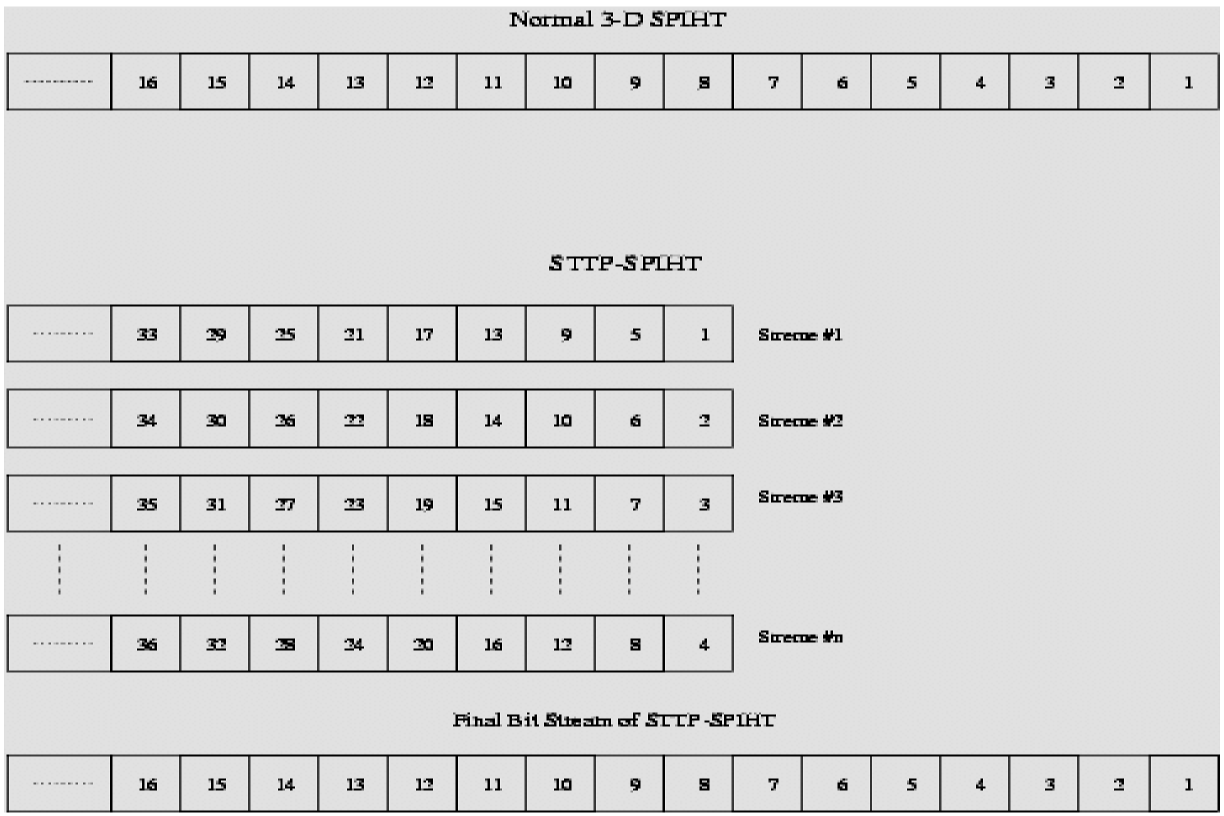
- The SPIHT algorithm generates a progressive bit stream for each tree-block
- As the SPIHT algorithm progresses it periodically records the current bit plane and list (LIP, LIS or LSP) being processed
- Once a sub-bit stream has been generated for each sub-image they are packetized and multiplexed to generate the primary bit stream
- Priority is given to packets with data for more significant bit planes
- Within the same bit plane, priority is given to packets with data for the LIP, then LIS then LSP
- The packetized bit stream is progressive

Example main bit stream structure:

hdr	0	data for 0	1	data for 1	2	data for 2	3	data for 3

2	data for 2	3	data for 3	3	data for 3	1	data for 1	2	data for 2

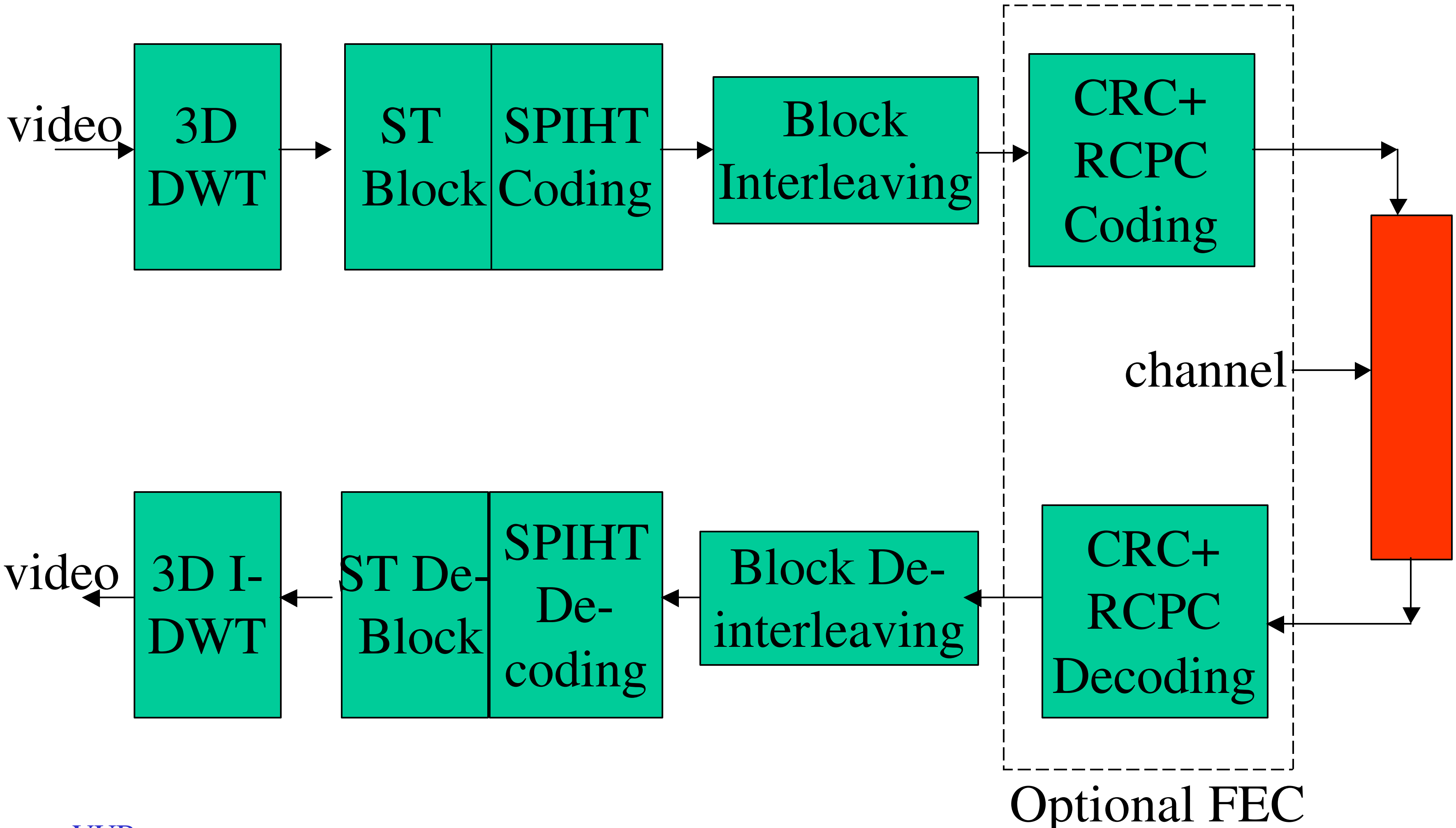
Packet Interleaving



STTP-SPIHT

- Requires overhead of header for each s-t block
- Fidelity embedding retained
  - bit stream packetized and re-organized

Video Coding/Transmission System

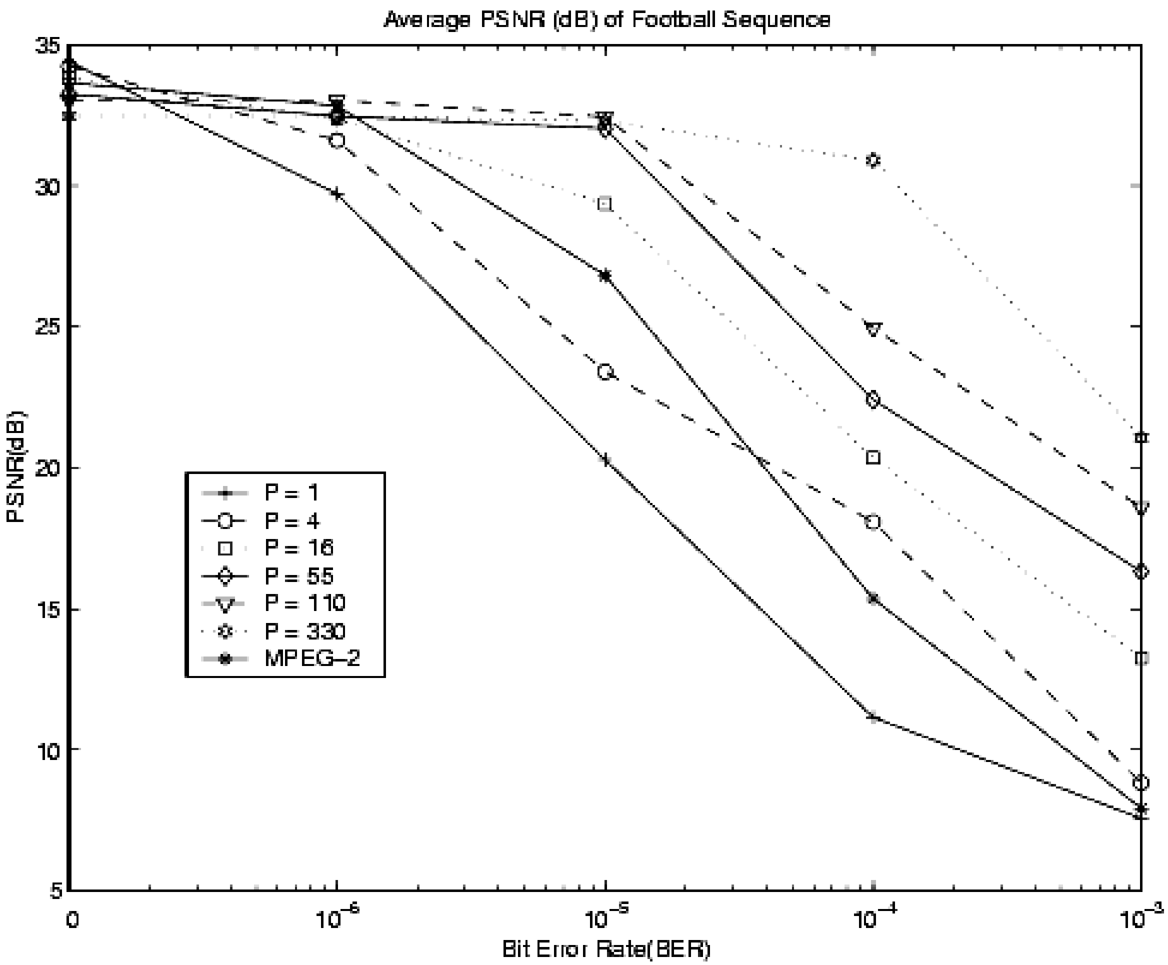


VUB
28 June 2001

Rensselaer

PSNR vs. BER for Different *P*

No FEC

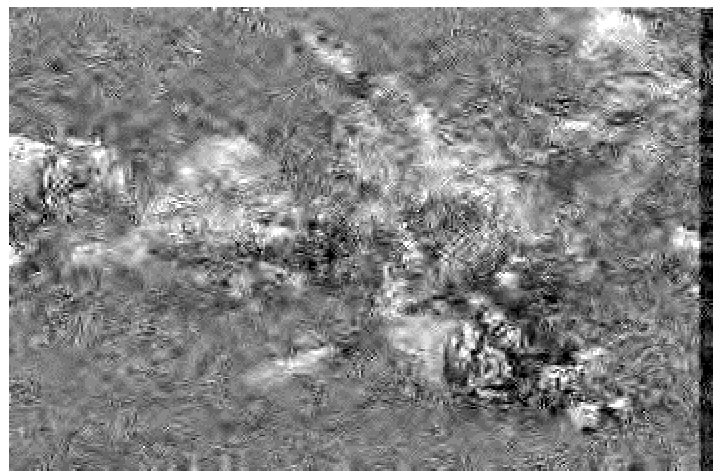


VUB
28 June 2001

Rensselaer Error Sensitivity of Normal SPIHT

Original Frame 15

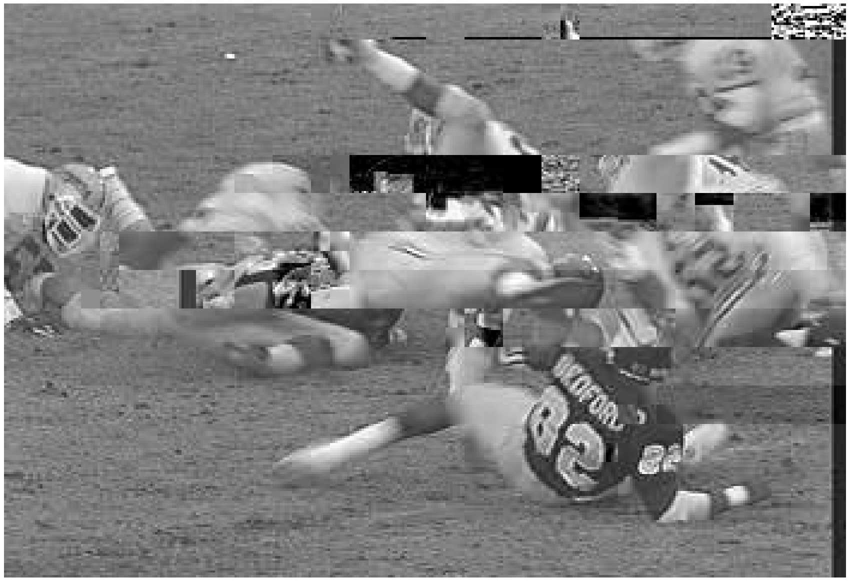
BER=0, 33.26 dB



BER=10^{-5}, P=16, 27.00 dB   BER=10^{-5}, PSNR=18.25 dB

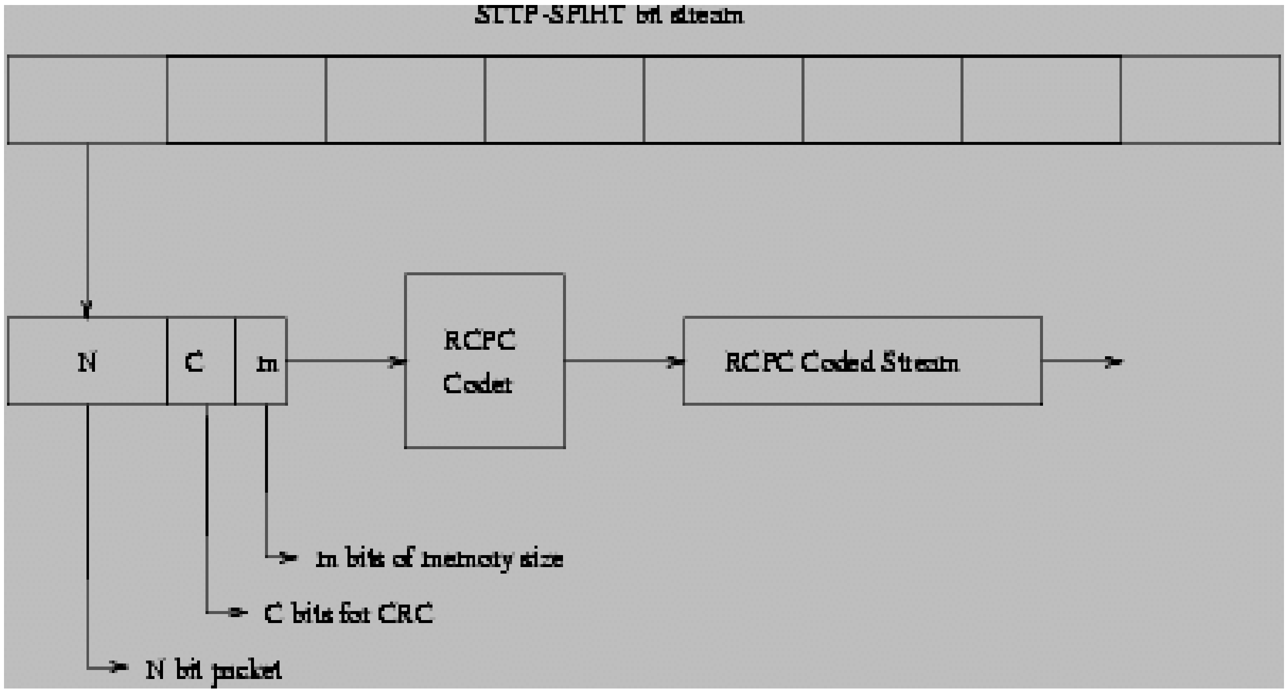
STTP-SPIHT vs. MPEG: BER=10^{-4}, No FEC

STTP-SPIHT, P=110, 25.80 dB



MPEG-2, PSNR=16.88 dB

Channel Coding Sub-system



FEC SPIHT Framework

- Cascade 3D SPIHT with FEC scheme
  - c = 16 bit CRC parity check code + RCPC code: same decoder for different rates r
    - $m = 6$ bits to flush memory
- For length N source segment, effective rate

$$R\_eff = [Nr/(N+c+m)]*R\_trans$$

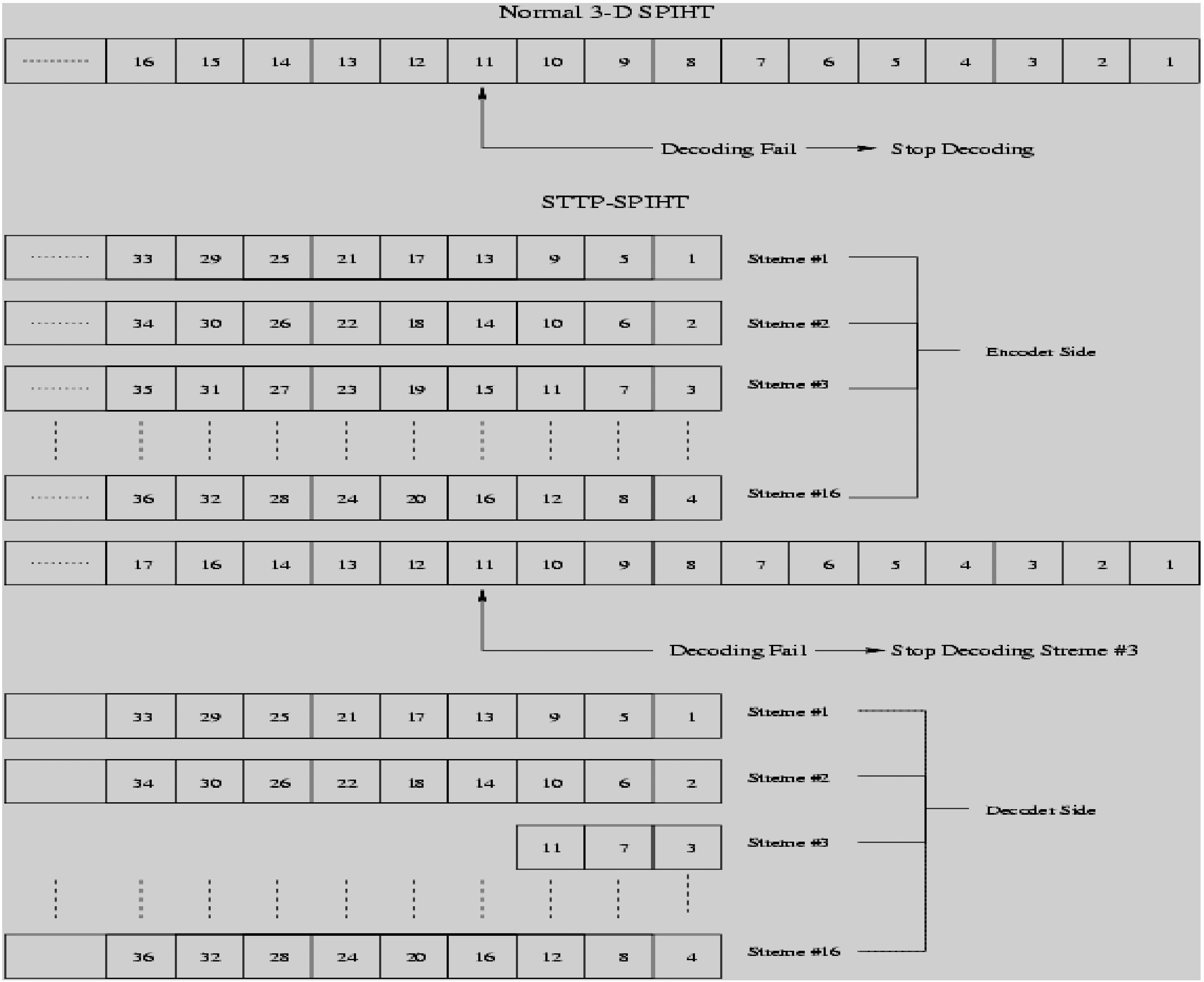
$$R\_trans = 2.53\ Mbps$$

r	*2/7*	*2/3*	*8/9*
BER	*0.1*	*0.01*	*0.001*
R_eff (bps)	*651,223*	*1,519,519*	*2,026,026*

Rensselaer

Bit Stream Organization

CNGV



Rensselaer FEC, BER=10^{-2} Comparisons

Original Frame 15

SPIHT, P=1, 24.41 dB

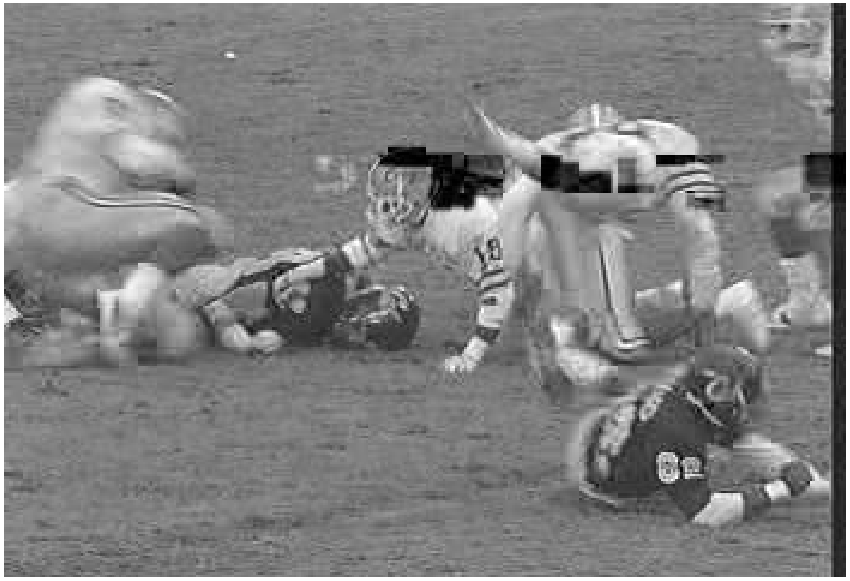
SPIHT, P=16, 29.35 dB

MPEG-2, 27.45 dB

STTP-SPIHT vs. MPEG-2: BER=10^{-2}, FEC

Frame 10

STTP-SPIHT, *P*=16, 30.49 dB

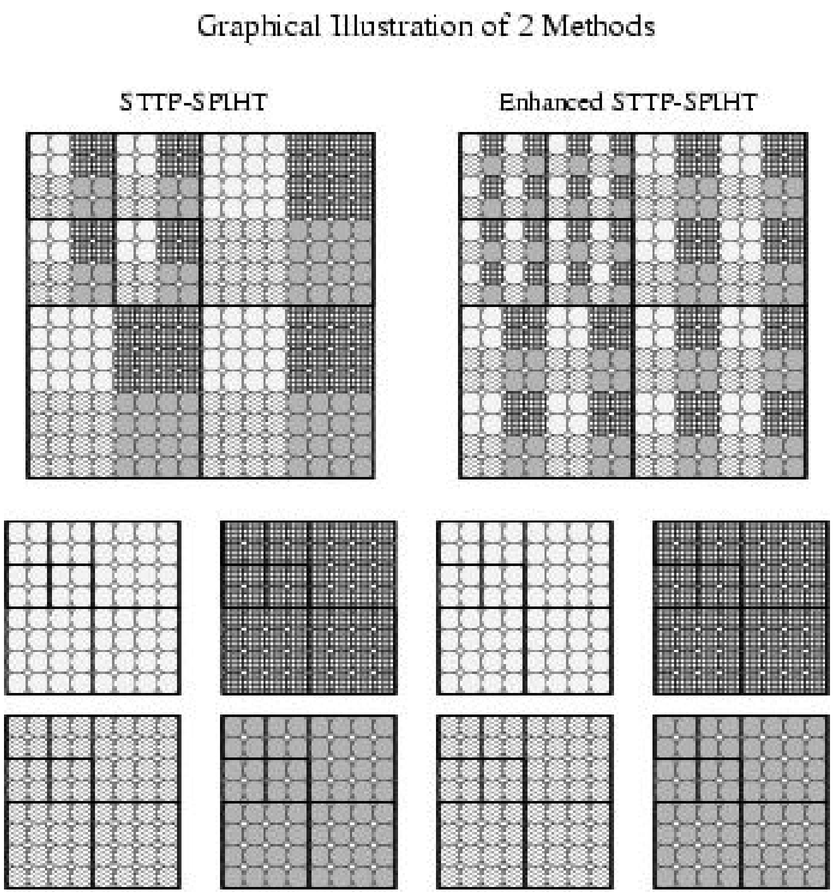


MPEG-2, 21.48 dB

Rensselaer

CNGV

New Tree Block Formation



Rensselaer

Video Frames for Different Methods

Comparison in Noisy Channel

Original Image

Normal 3D-SPIHT

PSNR = 24.41 dB

ESTTP-SPIHT

PSNR = 31.47 dB

STTP-SPIHT

PSNR = 28.68 dB

(352x240 Football Sequence, en/decoded with STTP/ESTTP-SPIHT (n = 16)
(Early decoding error occurred in the same position (substream #11), Frame #1

Rensselaer

CNGV

SPIHT FEC Average Results*

BER	ENER-SPIHT	STTP-SPIHT	3D-SPIHT	3D SPIHT/ ARQ
0.01	30.23	29.75	24.5	32.1
0.001	31.75	31.75	28.2	32.8

*Football Sequence, PSNR numbers

Rensselaer

CNGV

Conclusions

- Even w/o MC, better than MPEG-2 in noiseless and noisy channels
- SPIHT video bit stream can be robust or resilient to channel errors and packet loss
- Acceptable video quality can be delivered over very noisy channels
- Improvement possible via unequal error protection
- Promising for real-time, hardware applications