Subband Coding of Video

John W. Woods

Image Processing Laboratory
Rensselaer Polytechnic Institute
Troy, NY 12180–3590

Talk given in
Symposium on
Multiresolution Signal
Decomposition Techniques
at NJIT, Newark
on April 30, 1990

RPI Image Processing Laboratory (IPL):

- 8+ professors
- 23+ students
- recent NSF CISE Infrastructure Grant joint w/Computer Science Department for *parallel image processing*
- DAP and MasPar massively parallel SIMD machines
- with support from NSF, DOJ, CATS, Philips NA, NATO, etc.

This talk largely based on Rensselaer Master's Thesis by T. Naveen:

Subband Encoding of Video

IPL TR-109

January 1990

Work partially supported by David Sarnoff Laboratories and NSF Grant 8519948.

Outline of Talk:

- ☐ History of Subband Coding
- ☐ Subband Encoding of Images
- ☐ Statistics of Image Subbands
- ☐ Motion Compensation
- ☐ QMF Image Pyramids
- ☐ Video Coding
- ☐ Some On-going Work

See results presented at SPIE Symposium last November in Philadelphia, Symposium on Visual Information Processing

Subband Encoding of Digital Waveforms:

- ☐ Subband coding of speech introduced in 1976 by Crochiere, Webber, and Flanagan
- ☐ Quadrature Mirror Filter (QMF) banks introduced by Esteban and Galand in 1977
- ☐ Full frequency band is split into several sub-bands, each of which can be encoded more accurately
- ☐ Coding error is confined to individual frequency bands
- ☐ Noise spectra can be shaped via HVS model

Some Related Early Work:

- ☐ Split-band Quantization — Kretzmer, 1956
- ☐ Synthetic Highs System — Schreiber et al, 1959
- ☐ Gaussian Pyramid Image Coding — Burt & Adelson, 1983

Subband Coding (SBC) of Images:

- ☐ Multidimensional Subband Filtering by Vetterli — April 1984
- ☐ Sean O'Neil's Masters Thesis at RPI (first still image SBC) issued as IPL TR–071 April 1985
- ☐ Achim Brandt of Siemens (first SBC for video conferencing) — June 1985

7

- Subband techniques : Widely used for image (still frame) coding purposes.
- For video :

1. 3–D Separable subband coding (Karlson and Vetterli)
2. Motion compensation (backward) OR no motion compensation, and coding subbands of residue (Westerink)
3. Motion estimation in pyramid using block matching, forward motion compensation, and coding subbands of residue (Kronander)

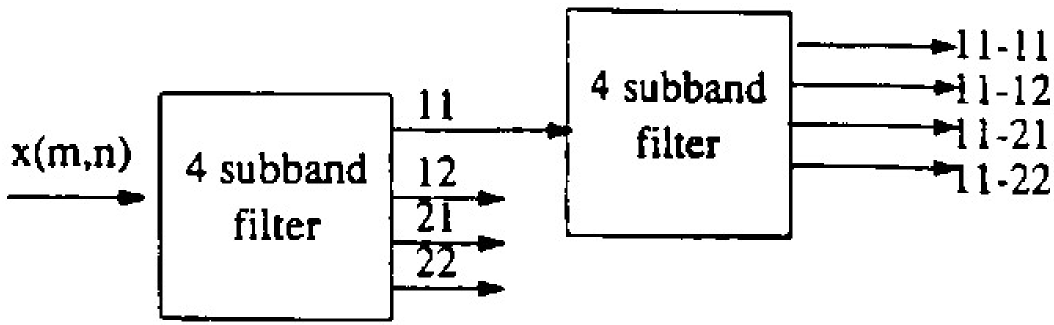


Fig 3.5. Structure for generating seven subbands

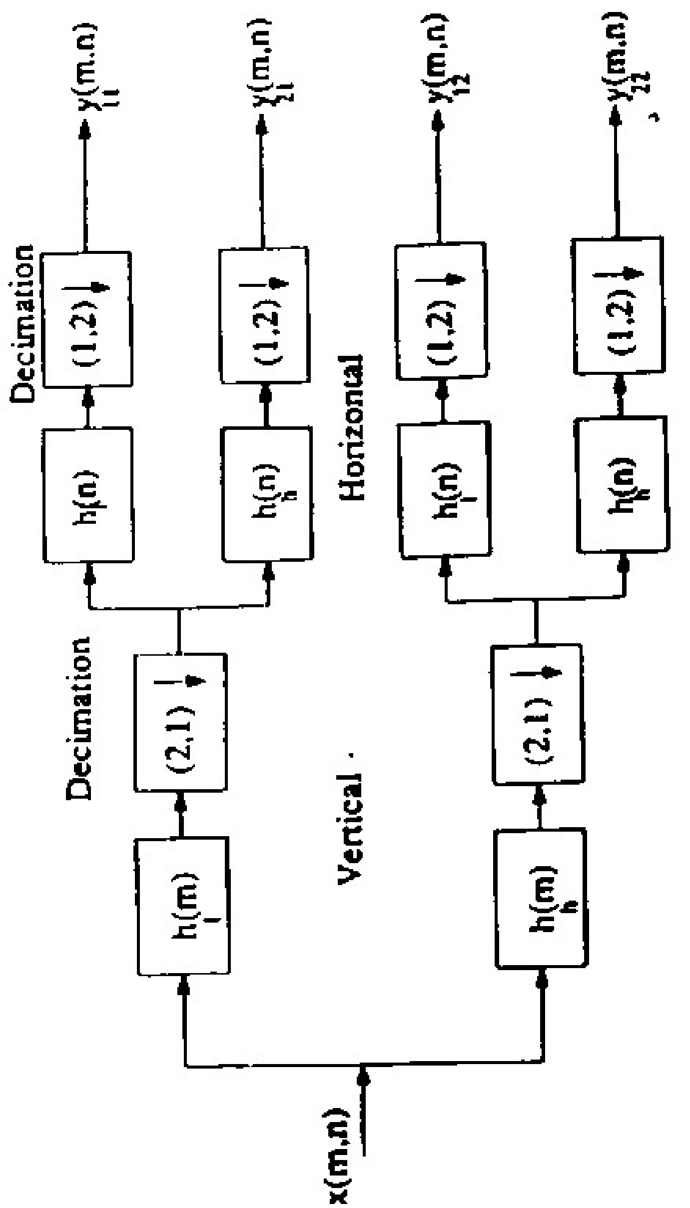


Fig 3.4. Seperable 4 subband filter

Statistics of Image Subbands:

- ☐ Only lowest subband(s) are correlated.
- ☐ Non-separable QP first-order linear predictors for lowest subband(s).
- ☐ Subband pdf's are generalized Gaussian with $0.5 \leq c \leq 1.1$,

$$p(x) = a\,exp\,(-\,|bx|^c)\ .$$

Note $c = 1.0$ for Laplacian pdf.

- ☐ Upper subbands energy is correlated.

Show 35 mm slide here.

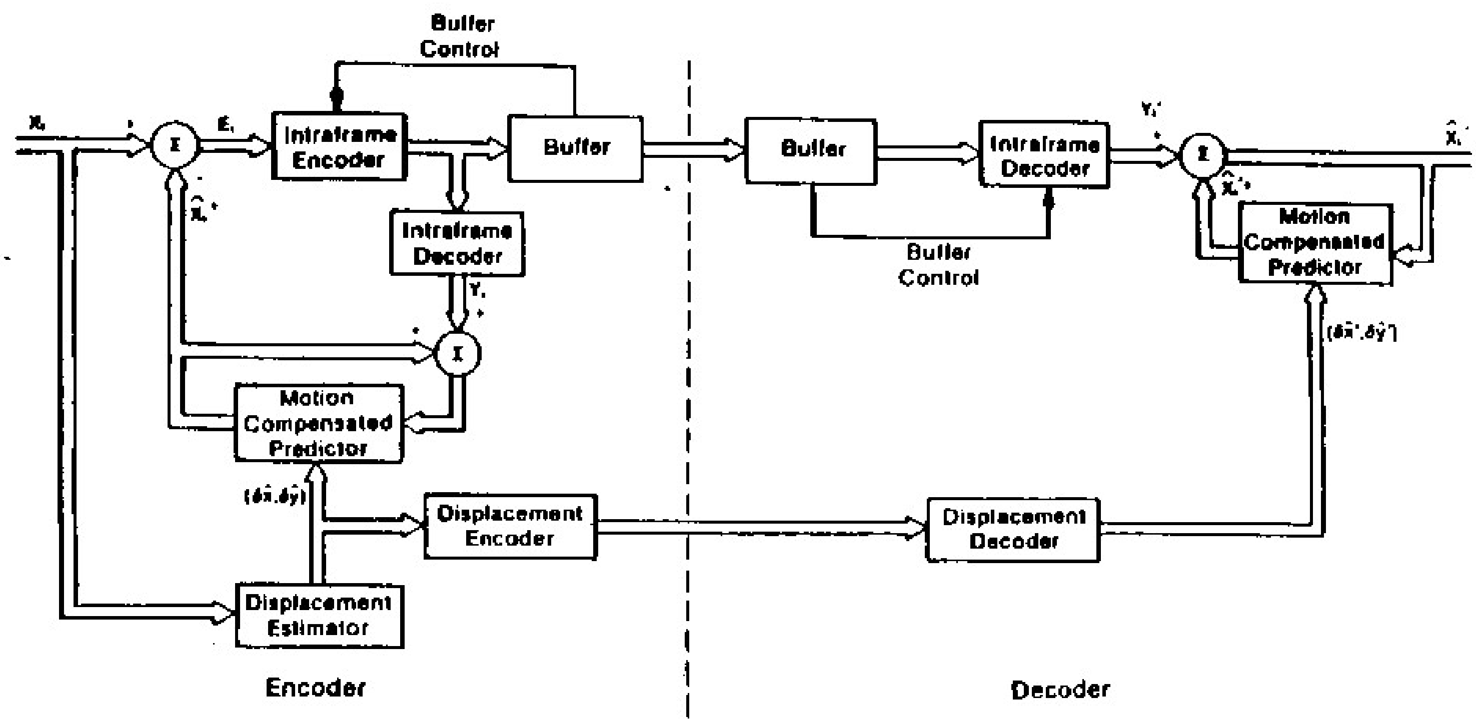


Figure 2.

Generic Video Encoder/Decoder

This Contribution

- Hierarchical motion estimation using pel recursion, and coding subbands of residue.
- QMF lowpass pyramid used for motion estimation.
- Backward motion compensation.
- Motion compensated frame interpolation and extrapolation.

The Subband filtering

- Separable 2–D filtering.
- Johnston's 16C QMF filters.

The Basic Intraframe Encoder

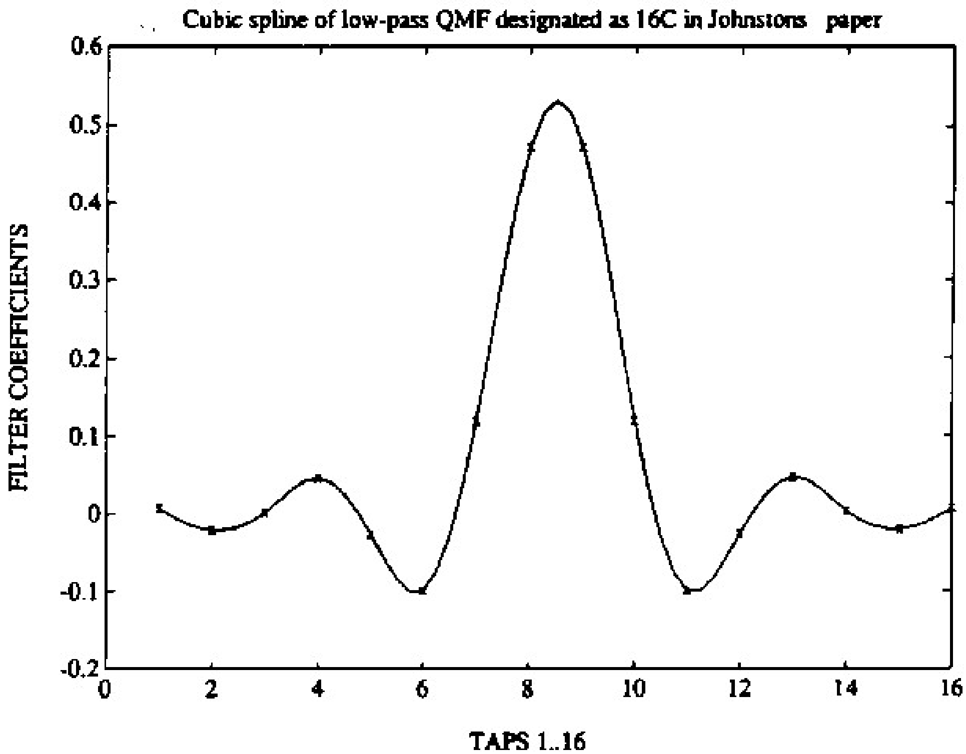
- DPCM with Adaptive Entropy-Coded Quantization.
- Rates reported are the zeroth order entropies.

— all quantizers have 11 uniformly spaced levels

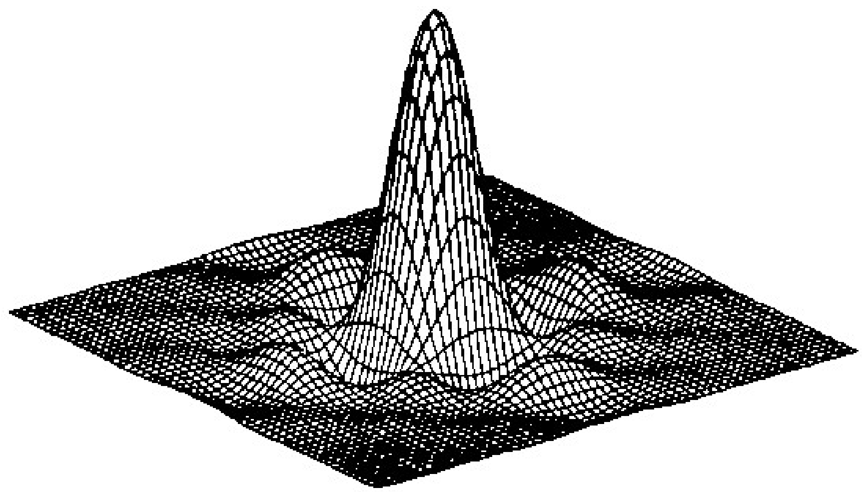
— The quantizers are adaptive per frame + per subband

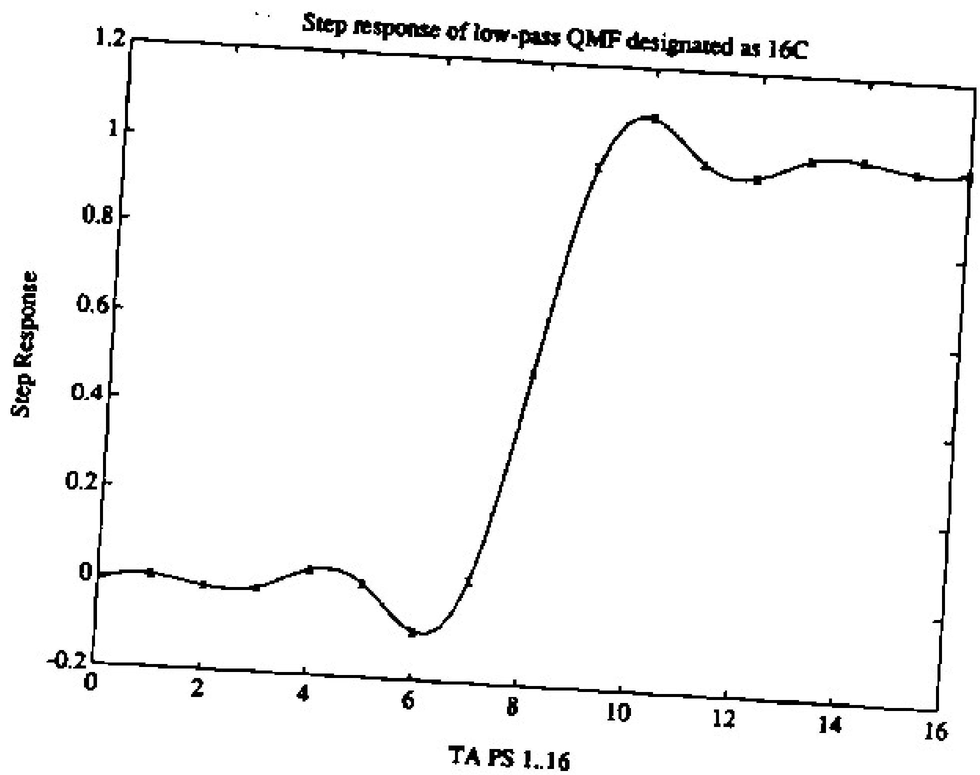
Overhead $11 \times 7 = 77$ (16 bit) words

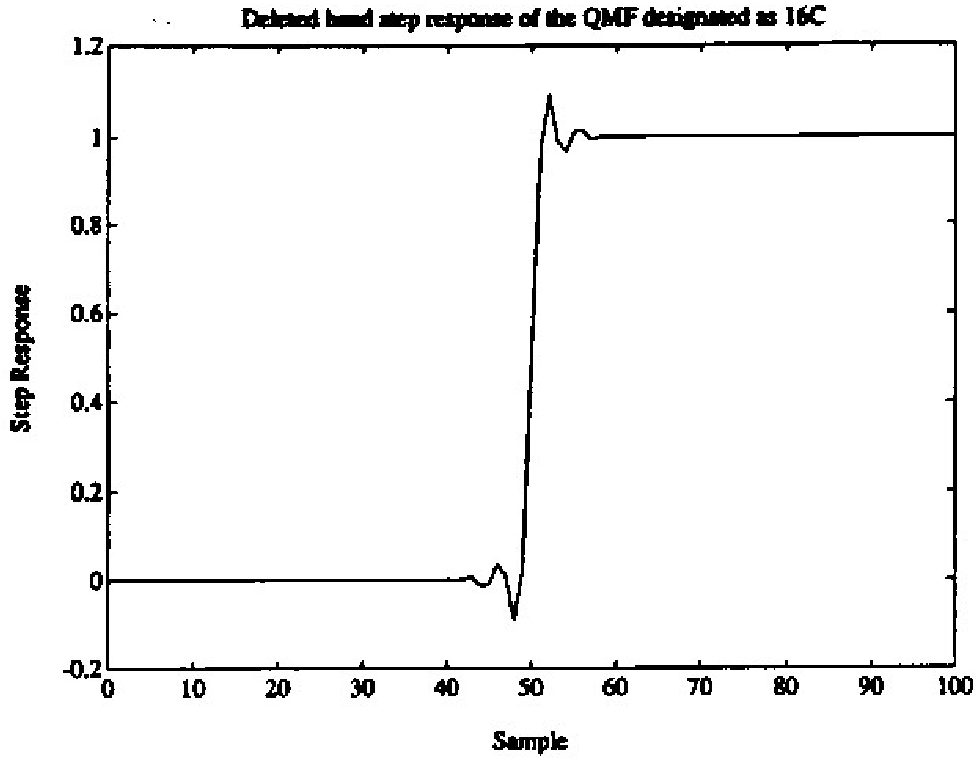
$\Rightarrow$ 0.006 bpp

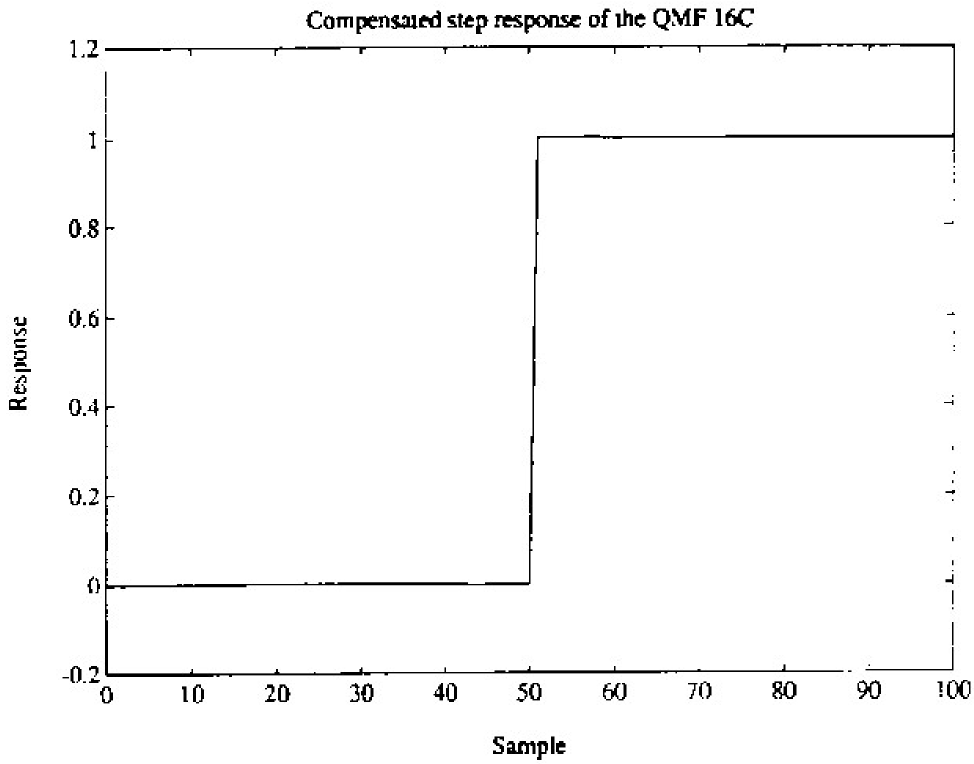

Cubic spline of low-pass QMF designated as 16C in Johnstons paper
FILTER COEFFICIENTS
TAPS 1..16
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
0
2
4
6
8
10
12
14
16

Mesh of the 2-D separable impulse response of 16C QMF




Step response of low-pass QMF designated as 16C
Step Response
TA PS 1..16
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
0
2
4
6
8
10
12
14
16


Delayed band step response of the QMF designated as 16C
Step Response
Sample
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
0
10
20
30
40
50
60
70
80
90
100


Compensated step response of the QMF 16C
Response
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
0
10
20
30
40
50
60
70
80
90
100
Sample

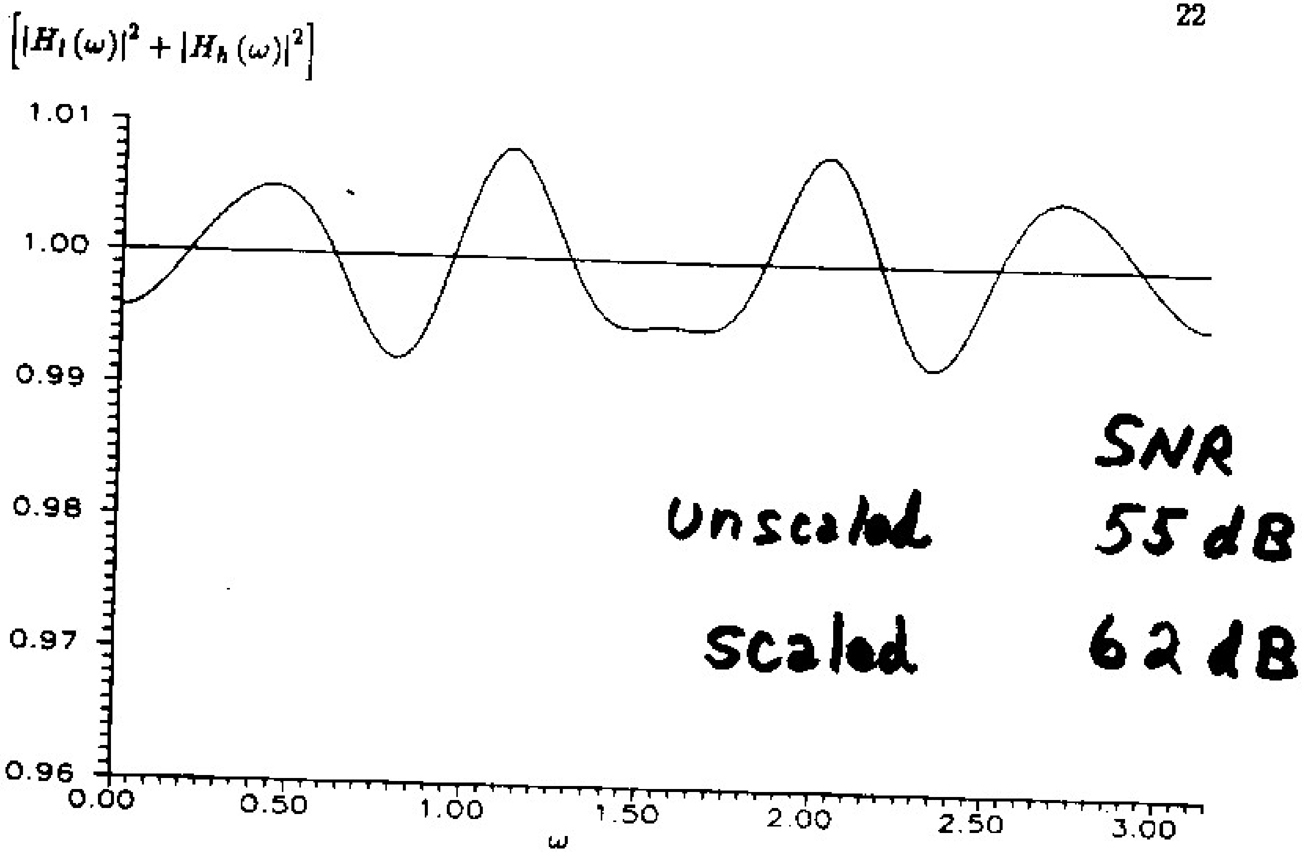


Fig 3.9 $\left[|H_l(\omega)|^2 + |H_h(\omega)|^2\right]$ vs ω for QMF designated as 16C in [John80].

A plot of $[|H(\omega)|^2 + |H(\omega + \pi)|^2]$ for 16C in [John80] is given in Fig. 3.9. This filter is optimal for signals with a uniform spectrum. There is a dip in the transfer

Motion Compensation:

Locally perturb frame N-1 data: $I_{N-1}(\mathbf{x} - \delta\mathbf{x})$
to make it coincide in least-square sense
with current frame N data: $I_N(\mathbf{x})$.

Error = $[I_N(\mathbf{x}) - I_{N-1}(\mathbf{x} - \delta\mathbf{x})]^2$.

Basic Motion Estimator (Netravali and Robbins)

- Goal : Find $\vec{\delta x}$ such that

$$I_N(\vec{x}) = I_{N-1}\left(\vec{x} - \vec{\delta x}\right)$$

- Pel recursion :

$$\vec{\delta x}^{\,i+1} = \vec{\delta x}^{\,i} - \epsilon \cdot DFD\left(\vec{x}, \vec{\delta x}^{i}\right) \nabla I_{N-1}\left(\vec{x} - \vec{\delta x}^{i}\right)$$

where DFD is the displaced frame difference

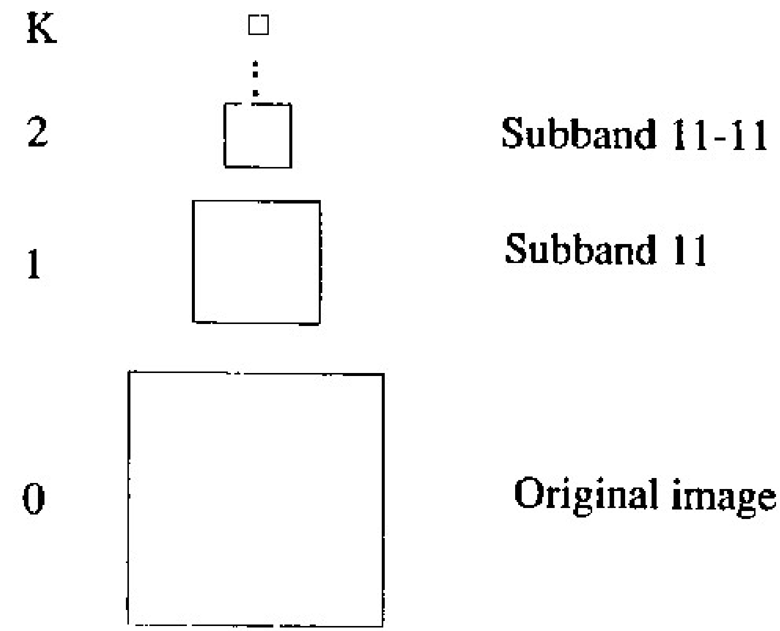
$$DFD\left(\vec{x}, \vec{\delta x}\right) = I_N(\vec{x}) - I_{N-1}\left(\vec{x} - \vec{\delta x}\right)$$

Motion Estimation in a Pyramid (Hierarchical)

Concept :

- Obtain a QMF lowpass pyramid.
- Estimate the motion in the top level, then refine it using lower (higher resolution) levels.
- Refining is done by pel recursion, iterating over the crude estimates obtained from the upper level.
- Crude estimates are smoothed using a lowpass filter (7 tap triangular) before refining.
- Good for large displacements and large object motions.

experimentally optimized over 3 to 13 tap triangular filters.



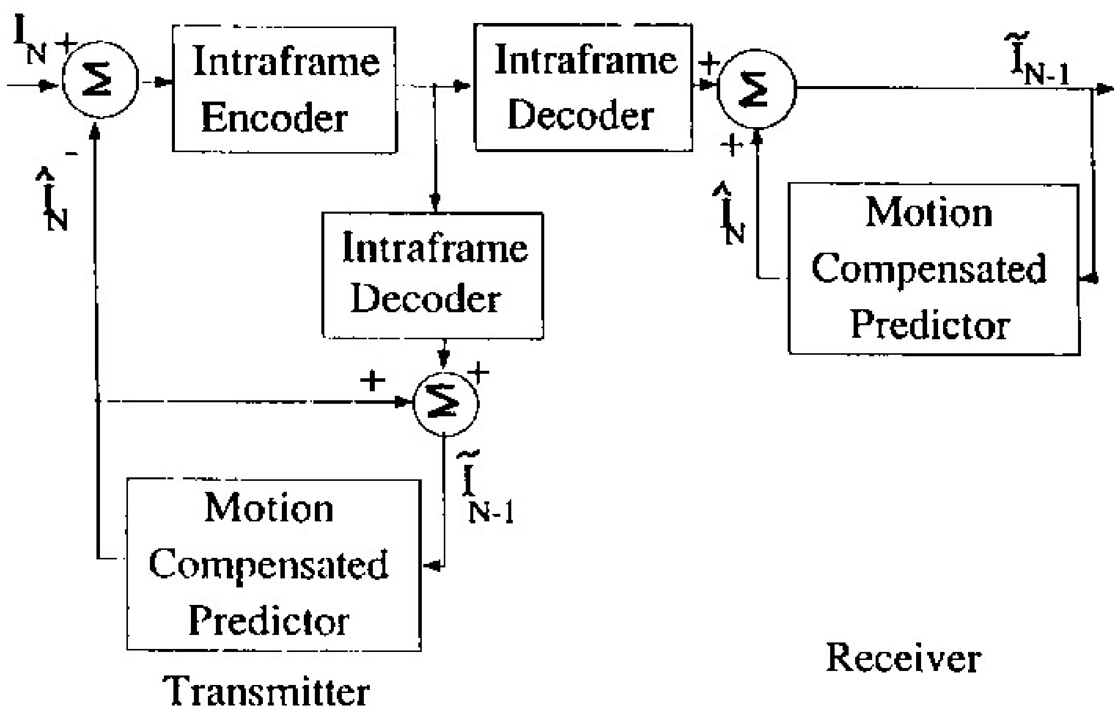
QMF lowpass pyramid

K=3 Miss Amer.

K=4 Caltrain (higher res)

Subband Coding in Pyramid Estimation

One way is :

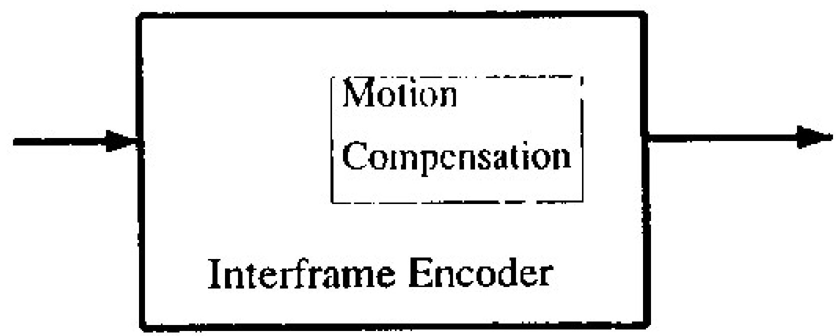


Standard Method

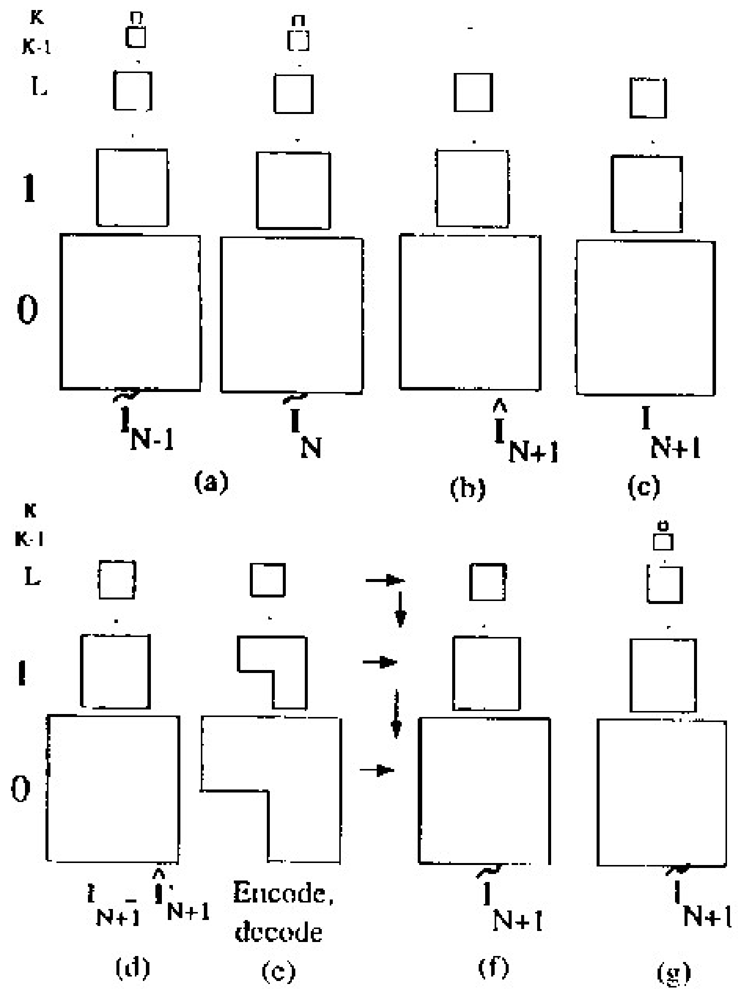
- Do motion estimation and predict the next frame.
- Obtain the error frame.
- Subband encode this residue.
- The coder has to wait for the motion estimation to finish.
- Motion estimator has to wait till the coding is done.

New Method

- Coding and motion estimation in parallel: could save time.



Basic idea: Put motion compensation inside the coder



(a) Previous reconstructed pyramids

(b) Predicted pyramid

(c) Original pyramid

L = 2 = coding

K = 3 or 4

(d) Error frame motion est.

(e) Encoded subbands

(f) Reconstructed pyramid of L levels

(g) Reconstructed pyramid of K levels

Bit allocation

- Seven subbands used for coding.
- Equidistortion criterion : May not be optimal.
- Bit rates can be changed with the step sizes of the uniform quantizer in DPCM.
- Step sizes are set once for a particular (the first) error frame, then kept constant for the entire sequence.

Bit allocations when coding seven subbands of the residue after pyramid estimation.

S. No	Subband	Step Size of the quantizer *	Av. zeroth order entropy
1	11-11	1.3	0.9959
2	11-12	4.0	0.2019
3	11-21	2.2	0.5287
4	11-22	4.0	0.2107
5	12	6.5	0.0748
6	21	3.0	0.3058
7	22	6.5	0.0723

* Normalized with respect to the standard deviation of the innovations sequence in the DPCM loop.

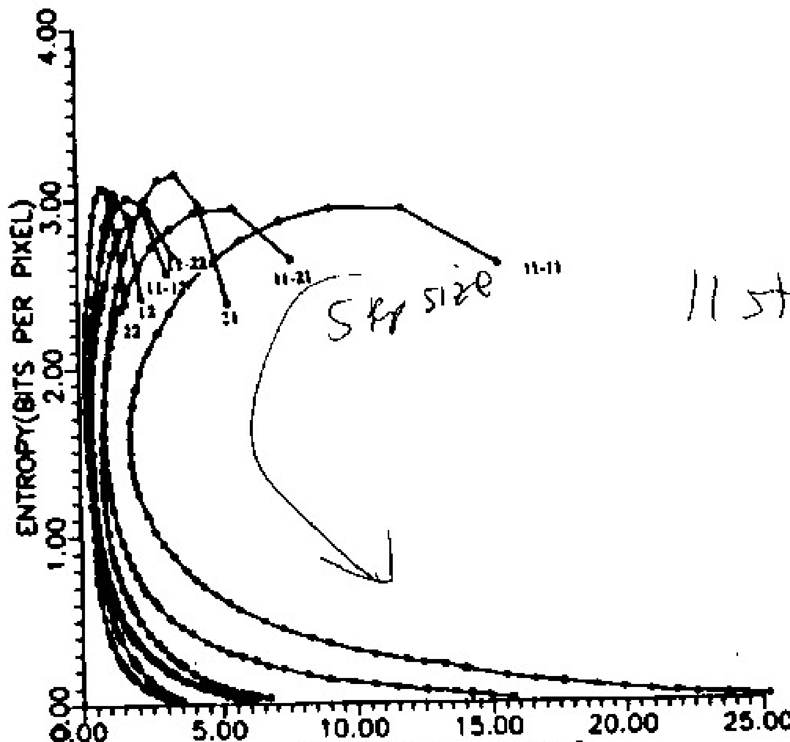


Fig. 5.2 TYPICAL R-D CURVES FOR DPCM ON SEVEN SUBBANDS.

Video Coding Experiments:

- ☐ Frame Difference Coding
- ☐ Motion Compensated Coding
- ☐ Motion Compensation within a QMF Pyramid

Frame Difference Coding — no motion compensation:

- ☐ monochrome *Salesman* sequence (from AT&T Bell Labs)
- ☐ 70 frames, 8 bits/pixel, 288x360
- ☐ DPCM — 40.6 dB @ 1.24 bits/pixel
- ☐ SBC/DPCM — 41.4 dB @ 1.04 bits/pixel
- ☐ Result: 0.8 dB better at 16% lower bit rate

Motion Compensated Coding:

- ☐ monochrome *CALTRAIN* sequence (from Sarnoff Labs)
- ☐ 33 frames, 8 bits/pixel, 400x512 size
- ☐ coded using 3 methods

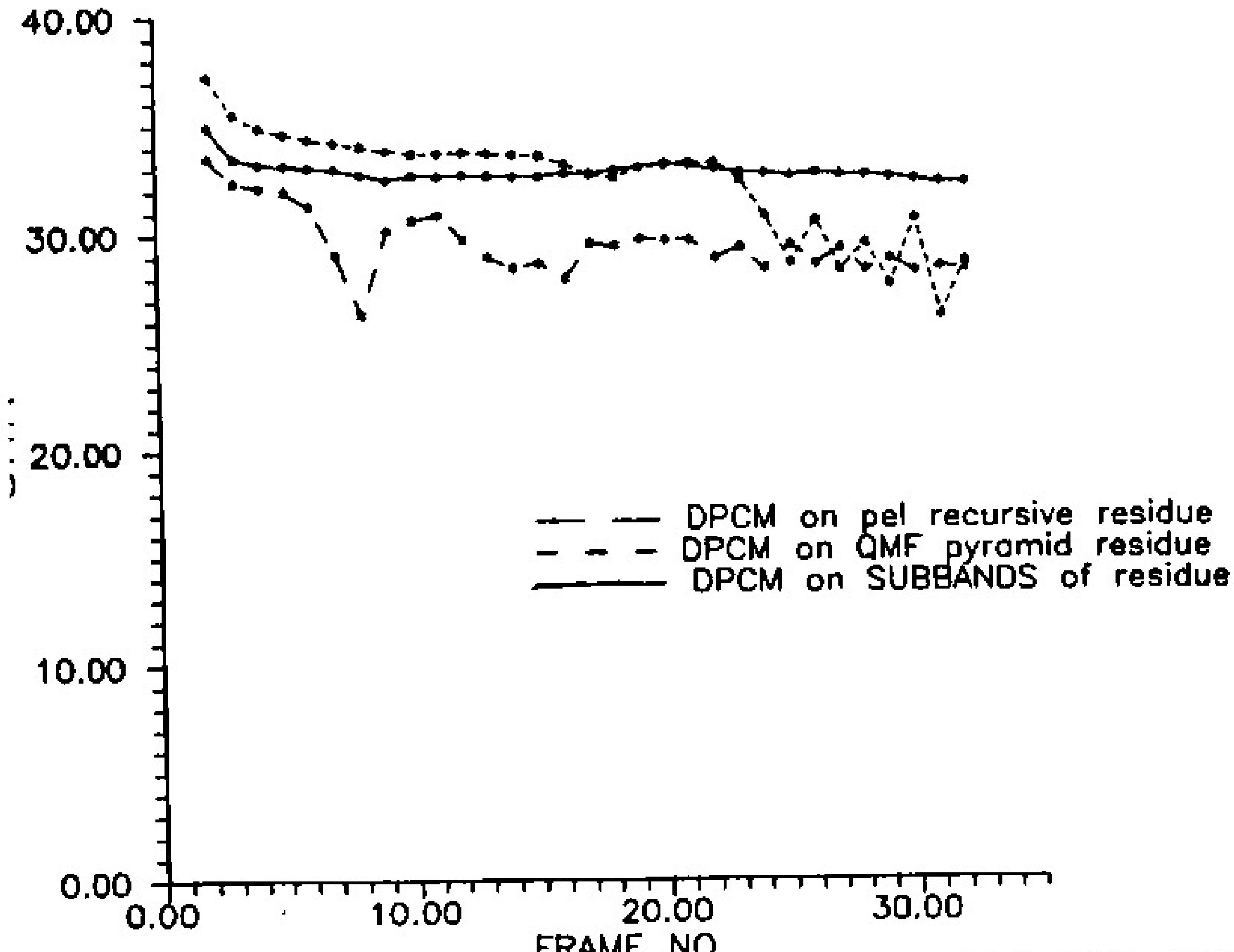
12

Schemes tried :

1. DPCM on pel recursive residue.
2. DPCM on QMF pyramid (motion estimation) residue.
3. DPCM on seven subbands of the QMF pyramid (motion estimation) residue.

Scheme	Bit Rate (bits/pixel)	SNR (dB)
DPCM on pel recursive residue	0.22	29.6
DPCM on QMF pyramid residue	0.22	32.5
DPCM on subbands of QMF pyramid residue	0.23	32.9

1. no pyramid no subband
2. motion est using pyramid but coding w/o subbands
3. use both



3. PERFORMANCES OF IMAGE SEQUENCE CODING SCHEMES (0.22 BITS/PIX)

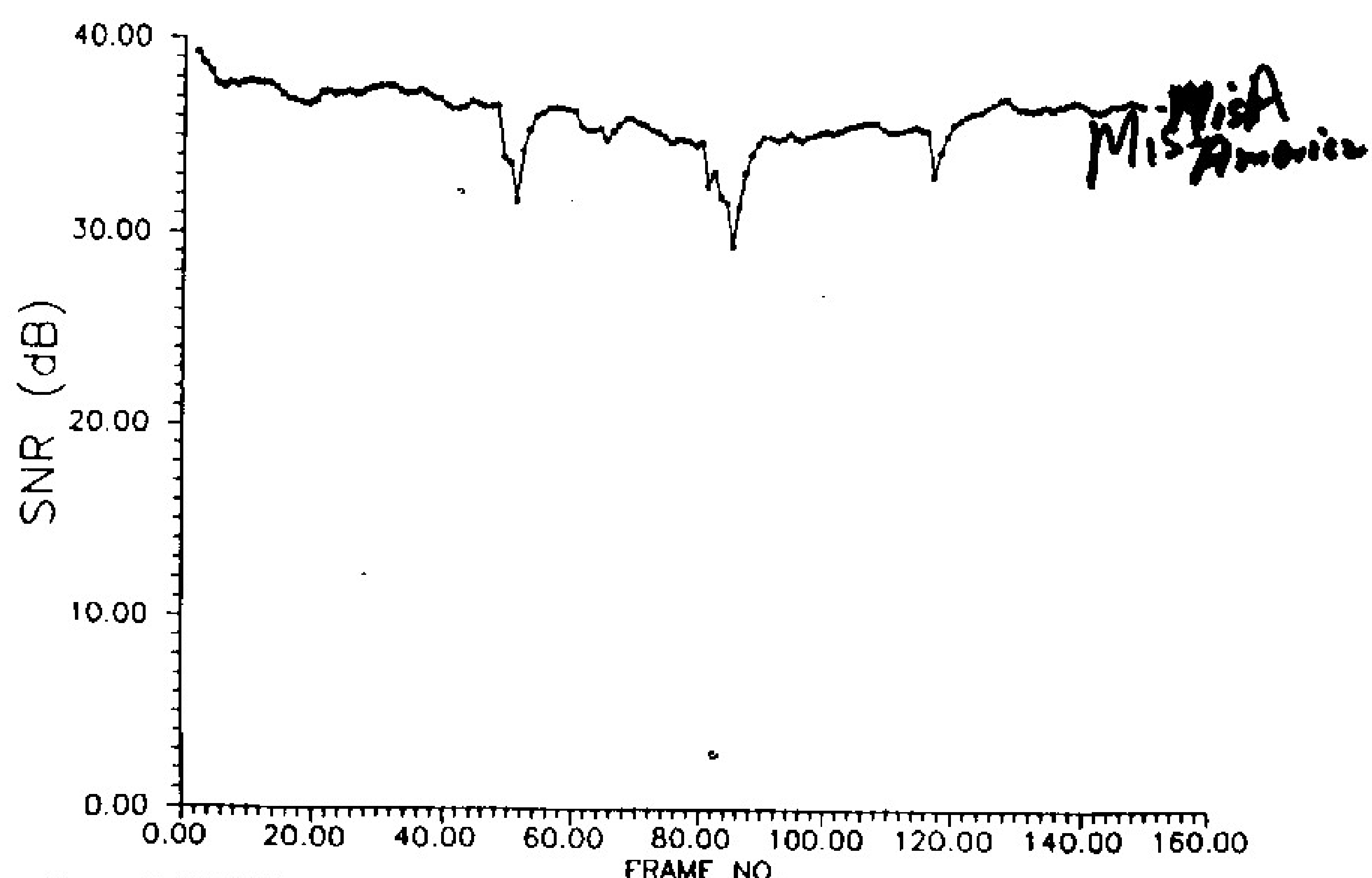


Fig. 6.3 PERFORMANCE OF UNIFIED PYRAMID SCHEME–MISS AMERICA (0.13 bpp)

S. No	Subband	Step Size of the quantizer	Av. zeroth order entropy
1	11-11	4.0	0.1859
2	11-12	9.0	0.0422
3	11-21	9.0	0.0335
4	11-22	9.0	0.0353
5	12	8.0	0.0368
6	21	4.0	0.1748
7	22	3.5	0.2435

Table 6.5: Bit allocations when coding seven subbands of the residue using the unified pyramid structure on the Miss America sequence

Caltrain

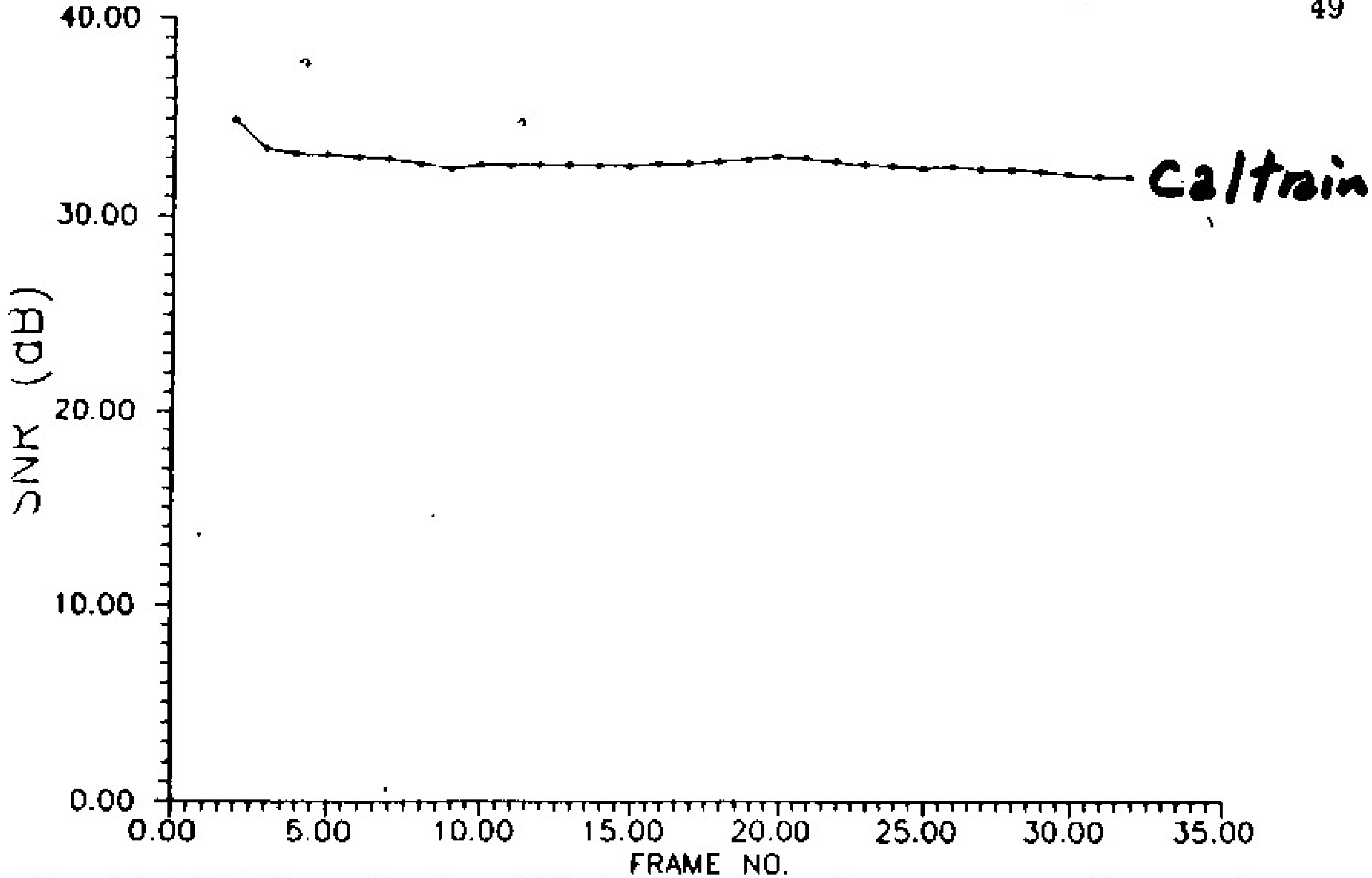


Fig. 6.2 PERFORMANCE OF UNIFIED PYRAMID SCHEME-CALTRAIN SEQ.(0.23 bpp)

Sequence	Bit Rate (bits/pixel)	SNR (dB)
Caltrain	0.23	32.9
Miss America	0.13	36.2

Table 6.6: Performance of the Unified Pyramid Structure on various sequences

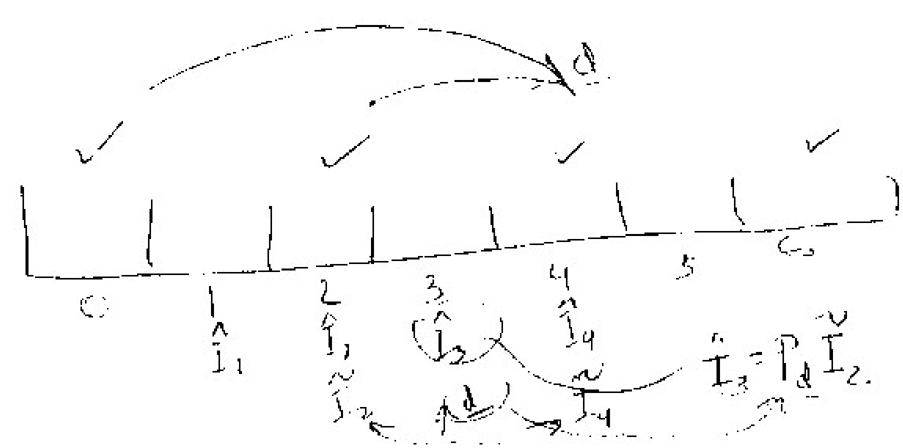
<u>Interpolation and extrapolation</u>

- To reduce frame rate and computation.
- Motion estimation done hierarchically.
- Interpolation

☐ Motion analysis done on even numbered frames but predict all the frames.

☐ Send error frames for even numbered frames only.

☐ Receiver assumes zero error for odd numbered frames.



- Extrapolation

Leave out 1 frame

transmit

n, n+1, skip, n+3, n+4, skip

Estimate motion between	N	N+1			
Predict	~~N~~	~~N+1~~	N+2	N+3	N+4
Send error				N+3	N+4
Reconstruct			N+2	N+3	N+4
Estimate motion between				N+3	N+4

$\hat{I}_{n+2} \equiv \tilde{I}_{n+2}$ $\quad \tilde{I}_{n+3} \quad \tilde{I}_{n+4}$

Results of Interpolation and Extrapolation

Scheme	Coder	Bit Rate (bits/pixel)	SNR (dB)
Interpolation	DPCM	0.13	29.34
		0.18	31.48
	DPCM on Subbands	0.11	29.54
		0.22	31.62
Extrapolat-ion	DPCM	0.13	27.67
	DPCM on Subbands	0.14	30.1

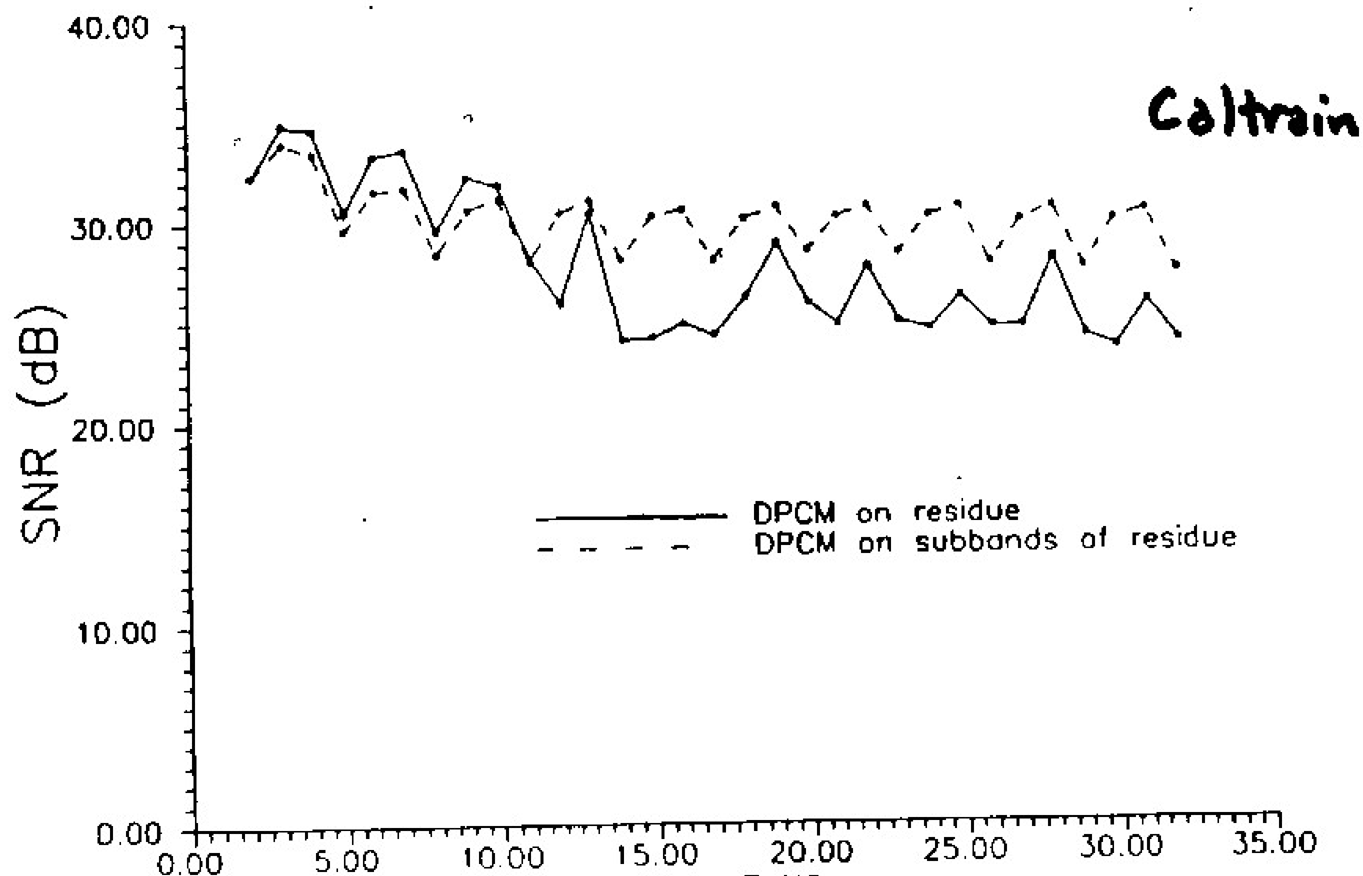


Fig. 6.6 PERFORMANCE OF MOT. COMP. EXTRAPOLATION SCHEME (0.13 bpp)

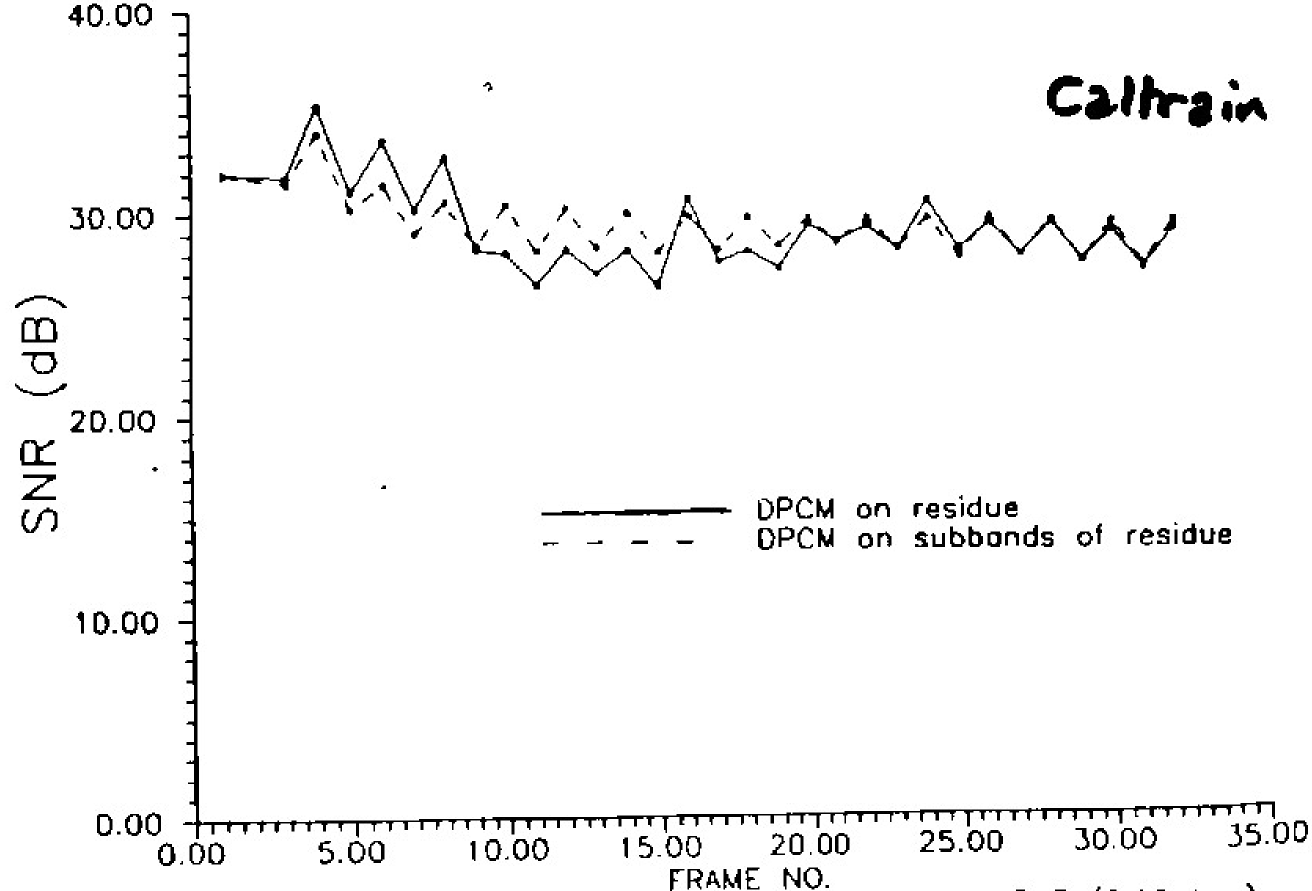


Fig. 6.4 PERFORMANCE OF MOT. COMP. INTERPOLATION SCHEME (0.12 bpp)

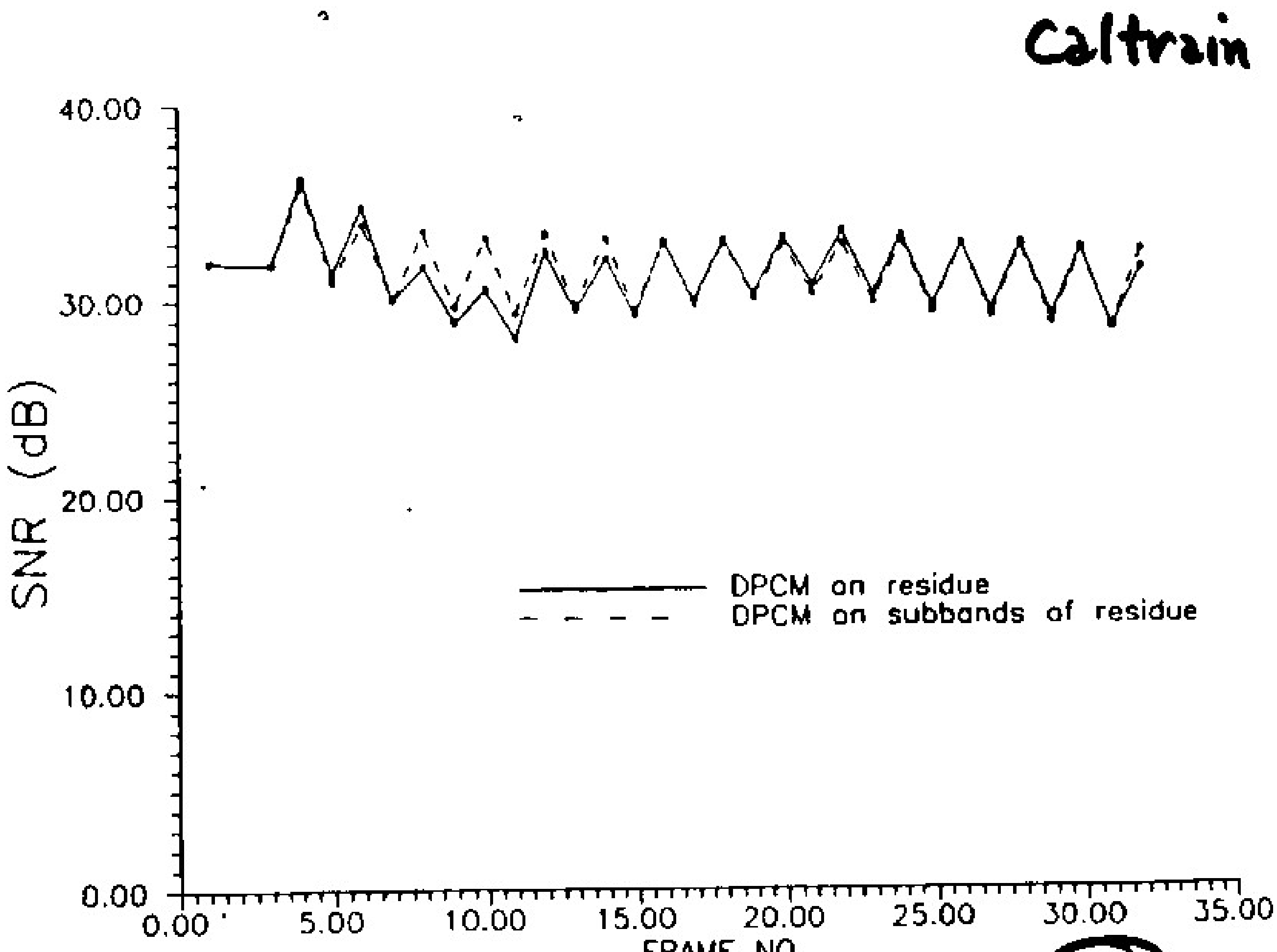


Fig. 6.5 PERFORMANCE OF MOT. COMP. INTERPOLATION SCHEME (0.2 bpp)

Conclusions

- Pel recursion on QMF pyramid structure gives better results than pel recursion in non-pyramid environment.
- Subband coding gave a stable performance.
- Motion compensated interpolation, extrapolation performed.

Future work

- Bit allocation : Robust and dynamic bit allocation system required.
- Conditional replenishment to reduce bit rate further.
- No actual coding done. Incorporate Arithmetic coding for practical system.
- Extend to code color video.

Some Future Work:

- ☐ Adaptive SBC for Motion Compensated Video
- ☐ 3–D Subband Coding with IIR temporal filters
- ☐ Investigation of 2–D Run-length Coding for Upper Subbands — (after Gharavi and Tabatabai)
- ☐ Incorporation of Conditional Replenishment Concept

Some On-going Work:

- ☐ Incorporation of Arithmetic Coding
- ☐ The Compositing of Subband Coded Images w/o Decoding
- ☐ SBC for Digital Zooming — after A. Lippman
- ☐ Subband Coding for the HVS
- ☐ 3–D Subband Filtering with IIR temporal filters

VIDEO
TAPE

SUBBAND ENCODING OF VIDEO

JOHN W. WOODS and T. NAVEEN

ECSE Department
Rensselaer Polytechnic Institute
Troy, NY 12180–3590.

CALTRAIN SEQUENCE
(ORIGINAL)

Length of sequence : 33 frames
Speed : 30 frames/sec
Proscan
400 x 512 pixels, 8 bits/pixel.

49 [redacted] M bits /sec.

7-4 Caltrain

1st on tape

2

CALTRAIN SEQUENCE

- DPCM on seven subbands of the QMF pyramid residue.
- Average bit rate : 0.23 bits/pixel.
- Average SNR : 32.9 dB.

1.4 Mbits/sec

CR = 34.8

2nd on tape

CALTRAIN SEQUENCE

- DPCM on the QMF pyramid residue.
- Average bit rate : 0.22 bits/pixel.
- Average SNR : 32.5 dB.

1.35 Mbits/sec

C.R. = 36.4

3rd on tape

CALTRAIN SEQUENCE

- DPCM on the pel recursive residue.
- Average bit rate : 0.21 bits/pixel.
- Average SNR : 29.6 dB.

1.29 Mbits/sec

CR = 38.1

4th on tape

MISS AMERICA SEQUENCE (ORIGINAL)

Length of sequence : 150 frames
Speed : 30 frames/sec
Proscan
288 x 360 pixels, 8 bits/pixel.

24.88 M bits/sec.

5 of 6 2nd Miss Ame.

5th on tape

MISS AMERICA SEQUENCE

- DPCM on seven subbands of the QMF pyramid residue.
- Average bit rate : 0.13 bits/pixel.
- Average SNR : 36.2 dB.

404 Kbits/sec.

Compression Ratio : 61

6th on tape

CALTRAIN SEQUENCE

- Motion compensated interpolation.
- DPCM on seven subbands of the QMF pyramid residue.
- Average bit rate : 0.11 bits/pixel.
- Average SNR : 29.54 dB.

675 Kbits/sec.

CR = 72 both

7 & 8 interpolation

7th on tape

CALTRAIN SEQUENCE

- Motion compensated interpolation.
- DPCM on the QMF pyramid residue.
- Average bit rate : 0.13 bits/pixel.
- Average SNR : 29.34 dB.

799 Kbits/sec.

CR=61

8th on tape

CALTRAIN SEQUENCE

- Motion compensated extrapolation.
- DPCM on seven subbands of the QMF pyramid residue.
- Average bit rate : 0.14 bits/pixel.
- Average SNR : 30.1 dB.

860 Kbits/sec

CR = 57

9th on tape

CALTRAIN SEQUENCE

- Motion compensated extrapolation.
- DPCM on the QMF pyramid residue.
- Average bit rate : 0.13 bits/pixel.
- Average SNR : 27.67 dB.

799 K bits/sec.

CR = 61

10 ft on Tape