CCITT SGXV
Working Party XV/1
Specialists Group on Coding for Visual Telephony

Document # 43
September 13,1985

Source : NTT, KDD, NEC AND FUJITSU

Title : CANDIDATES FOR SUBRATE CODING ALGORITHM

1. Introduction

In the Holmdel meeting (Document #22), we introduced some basic coding algorithms which are applicable to sub-primary rate codec, and a fundamental configuration of coding system. This document presents some candidate algorithms hopeful for 384 kbit/s codec.

2. Common approach

To achieve 384 kbit/s transmission of video teleconferencing signals, highly efficient coding algorithms taking advantage of correlation between frames are indispensable. Therefore, our common approach is to introduce an interframe prediction, especially motion-compensated interframe prediction. A fundamental configuration of motion-compensated interframe prediction codec is shown in Figure 1.

3. Candidates of coding algorithms

Under the same approach to process motion-compensated interframe prediction errors, our possible coding algorithms are:

(1) Interpolative/extrapolative prediction coding
(2) Discrete cosine transform coding
(3) Vector quantization

We have been investigating above algorithms and some of their variations. Annex 1 describes candidates which are under study.

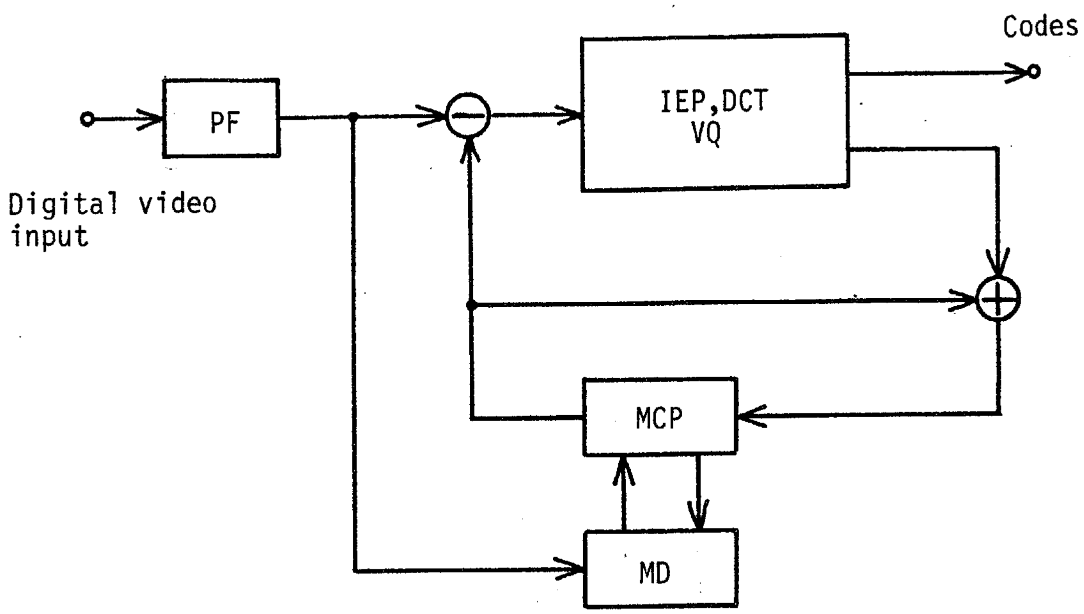
4. Coding characteristics of the candidates

We are now investigating three candidates through comparing coding characteristics in terms of S/N, entropy and decoded picture quality to choose a promising algorithm. Since basic coding characteristics should be clarified in the first stage of study, feedback coding control has not been introduced. Common picture data for investigation are sequence X (split-screen) and sequence Y (Trevor) : these are presented by European countries. Sequence B (3 persons talking) which is one of the common test sequences in Japan is also used.

Annex 2 shows examples of S/N vs. entropy of the candidates, where S/N and entropy are defined as described in the companion document # 44 "Evaluation method of coding algorithm." Annex 2 also presents amount of each information to be transmitted in the candidate algorithms. Examples of decoded pictures are demonstrated in the meeting. Annex 2 implies that every algorithm needs field dropping or some other subsampling method to code at 384 kbit/s rate.

5. Future plan

"Evaluation method of coding algorithm" (Document # 44) suggests a way of choosing a promising algorithm through further investigation. We believe that the number of candidate algorithms can be reduced by this method.



PF : Prefilter
IEP : Interpolative/extrapolative prediction coding
DCT : Discrete cosine transform coding
VQ : Vector quantization
MCP : Motion-compensated predictor
MD : Motion vector detector

Fig.1 Fundamental configuration of 384 kbit/s codec

Annex 1

CANDIDATES FOR 384 KBIT/SEC CODEC WHICH ARE UNDER STUDY

1. Candidate A: Interpolative/extrapolative prediction coding with motion-compensated interframe prediction

1.1 Basic configuration

In this algorithm, motion-compensated interframe prediction errors are coded taking advantage of temporal subsampling and spatial subsampling techniques in which a sort of interpolating filtering function is carried out. Since interpolating functions are utilized as predictive function, differences between the interpolated value and the input signal value are coded.

While conventional codecs are provided with these filters mainly after the coding process, this algorithm performs filtering functions such as interpolation in parallel with prediction function. Therefore, the time margin in the prediction loop is increased. Prediction efficiency and resolution can be improved by interpolative prediction. Figure A1 shown a blockdiagram.

1.2 Algorithm description

(i) Picture element classification

Input picture elements are divided into three categories: pels for extrapolative prediction (called P-pels), pels for interpolative prediction which are in the same field as P-pels (called I-pels) and pels for interpolative prediction which are not in that field (called F-pels).

(ii) Motion compensated interframe prediction

Block matching type of motion compensated prediction is introduced for P-pels prediction. For segmented block of P-pels, an appropriate block in the previous frame is chosen by calculating evaluation functions between them. Significant pels are scalar quantized.

(iii) Intrafield interpolative prediction

The I-pels which are not coded in process (ii) are coded adaptively. In the still picture area the pels are predicted by motion-compensated interframe prediction, while in the moving area the pels are interpolative predicted using adjacent pels in the same field. Prediction errors are quantized with three level quantizer in the basic mode, and quantizer characteristics can be changed block by block.

(iv) Interfield interpolative prediction with motion interpolation

For the fields which are not coded in the (ii) or (iii) ways, all the pels(F-pels) are interpolated by P and I-pels which are in the previous and the following fields with motion-compensation. The pels where the interpolative errors are large, are coded with a coarse quantizer such as three level quantizer.

(v) Information to be transmitted

In the above coding algorithm:
- Moving vector
- Significant/insignificant pel identification address
- Scalar quantizing index for P-pel
- Scalar quantizing index for I-pel
- Scalar quantizing index for F-pel

should be transmitted.

2. Candidate B: Discrete cosine transform coding with motion-compensated interframe prediction

2.1 Basic configuration

In this algorithm, motion-compensated interframe prediction errors are segmented into blocks. Then, blocks are discrete cosine transformed and quantized with some variations. Figure A2 shows a blockdiagram.

2.2 Algorithm description

(i) Motion compensated interframe prediction

Block matching type of motion compensated prediction is introduced. For segmented block of input signals, an appropriate block in the previous frame is chosen by calculating evaluation functions of candidate blocks.

(ii) Significant/insignificant block detection

An evaluation function derived, for example, from prediction errors in a block, classifies the blocks into significant blocks and insignificant blocks. When a block is detected as insignificant, only flag information indicating so is transmitted.

(iii) DCT/ without-DCT block detection

In a variation algorithm, an evaluation function derived from prediction errors in a block classifies the blocks into blocks for DCT and blocks for without-DCT. Prediction errors are directly scalar quantized for the blocks for without-DCT, while the others are discrete cosine transformed. Another evaluation function derived from input signals itself can be used as well as prediction errors in the prediction loop.

In another variation, average value in the block for DCT is calculated in advance, then prediction errors are subtracted by this average and are transformed.

(iv) Orthogonal Transform

Prediction errors are two-dimensional discrete cosine transformed with N X M pel size. N and M are eight or sixteen at present.

(v) Quantization

Simple scalar quantization techniques can be employed for transform coefficient quantization.

Vector quantization technique is introduced as a variation. In such variation, an L-dimensional vector set consisting of K vectors are generated, in which L lower coefficients of orthogonal transform coefficients in the significant blocks are utilized. Training sequences make a converged vector set. A scalar quantizer tables is also generated for each vector. In the case of L coefficients, residuum values which are subtracted representative vector value from input coefficients are referenced. Above are offline processing.

Input transformed blocks are classified into K classes by the vector set matching. Then, each classified coefficient block is first subtracted by vector quantized representative in the L coefficients, then the residuum and the other coefficients are quantized according to the quantization table. (See Fig.A3)

(vi) Filtering

In a variation, spatial adaptive filtering and temporal adaptive filtering are applied to decoded signals to improve picture quality and coding efficiency.

In another variation, spatial adaptive filtering is applied to outputs of the inverse DCT. Only non-edged areas, which is detected using decoded prediction errors in the previous frame, is low pass filtered so that block-like noise is reduced without degradation of resolution.

(vii) Information to be transmitted

In the above coding algorithm:

- Significant/insignificant block identification flag
- DCT/ without-DCT block identification flag *
- Average value in a block *
- Moving vector
- Vector quantizing set index *
- Scalar quantizing index

should be transmitted. Mark * presents information transmitted as variation algorithms.

3. Candidate C: Vector quantization with motion compensated interframe prediction

3.1 Basic configuration

In this algorithm, motion-compensated interframe prediction errors are segmented into blocks. Blocks are vector quantized by detecting the most appropriate vector and some blocks with large amount of resiual error are further processed by a scalar quantizer. Vector quantizing index and scalar quantizing index are transmitted. Figure A4 shows a blockdiagram.

3.2 Algorithm description

(i) Motion compensated interframe prediction

Block matching type of motion compensated prediction is introduced as described in 2.2.

(ii) Significant/insignificant block detection

Each block is classified into significant blocks and insignificant blocks. Only flag information is transmitted for the insignificant block.

(iii) Vector quantization

Vector sets are generated using training sequences in the offline processing.

The most suitable vector index in the vector set is chosen by minimum square error criterion.

(iv) Scalar quantization

Some blocks which are determinded by moving vector value and vector set index are scalar quantized. Namely, residuum of interframe difference and vector quantization representative value is quantized.

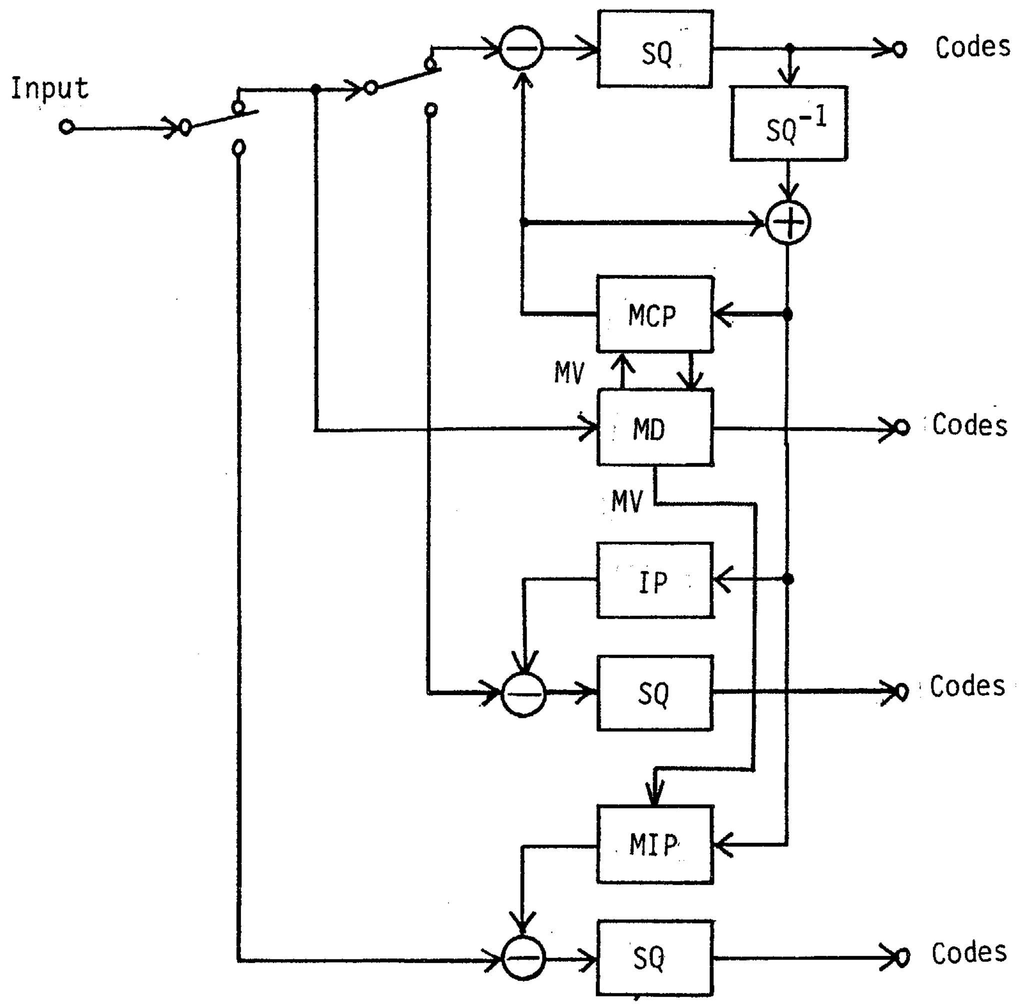
(v) Filtering

In a variation, spatial adaptive filtering and temporal adaptive filtering are applied to decoded signals to improve picture quality and coding efficiency.

(vi) Information to be transmitted

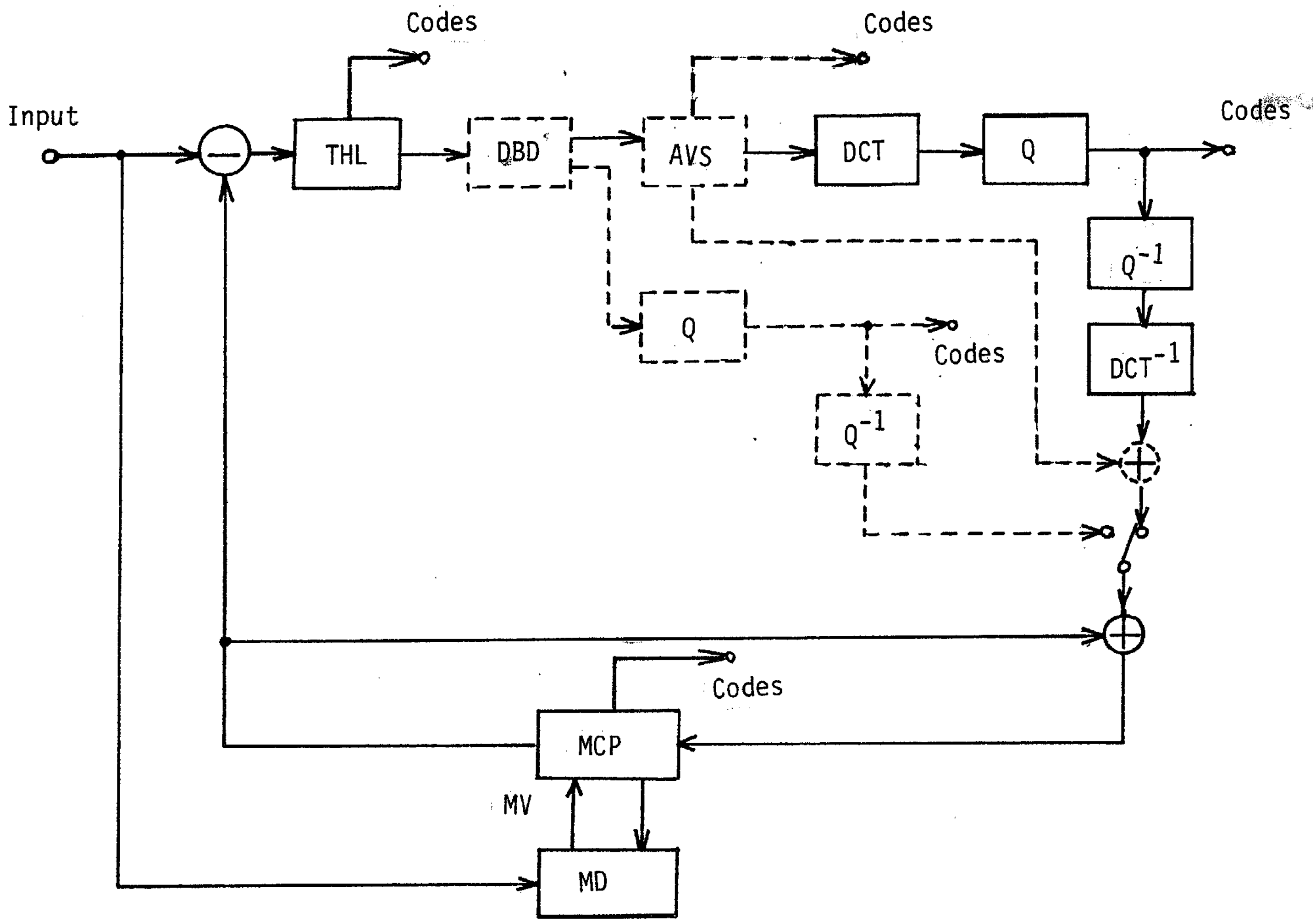
In the above coding algorithm:
- Significant/insignificant block identification flag
- Moving vector
- Vector quantizing set index
- Scalar quantizing index

should be coded.



SQ : Scalar quantizer
SQ^{-1} : Scalar quantizer decoder
MCP : Motion-compensated predictor for P-pels
MD : Motion vector detector
IP : Interpolative predictor for I-pels
MIP : Motion-compensated interpolative predictor for F-pels
MV : Motion vector

Fig. A1 Blockdiagram of interpolative/extrapolative prediction coding

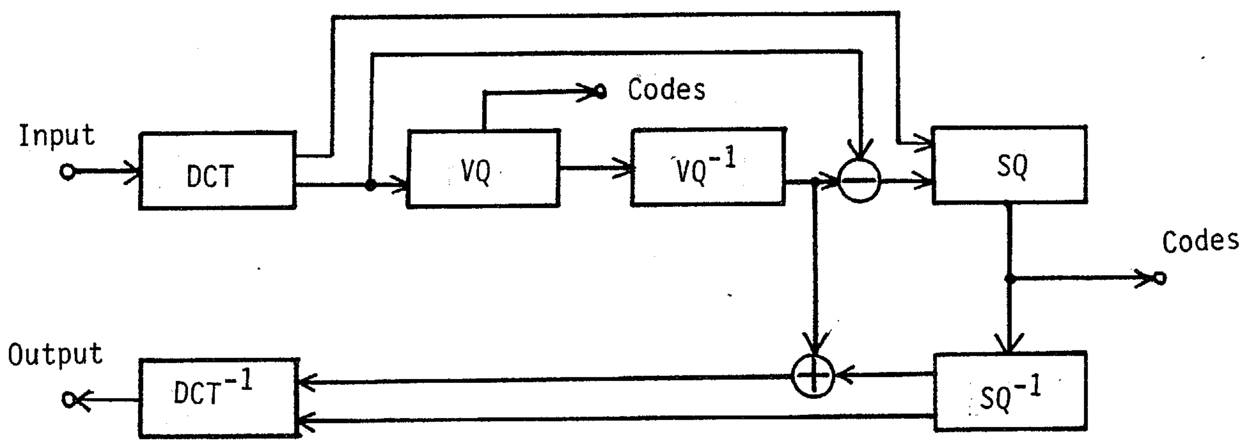


THL : Significant/insignificant block detector
DBD : DCT / without-DCT block detector
AVS : Average separator
DCT : Discrete cosine transform
Q : Quantizer
Q^{-1} : Quantization decoder
DCT^{-1} : Inverse discrete cosine transform
MD : Motion vector detector
MCP : Motion-compensated interframe predictor
MV : Motion vector

Note 1. "[dashed box]" represents variations of algorithms.

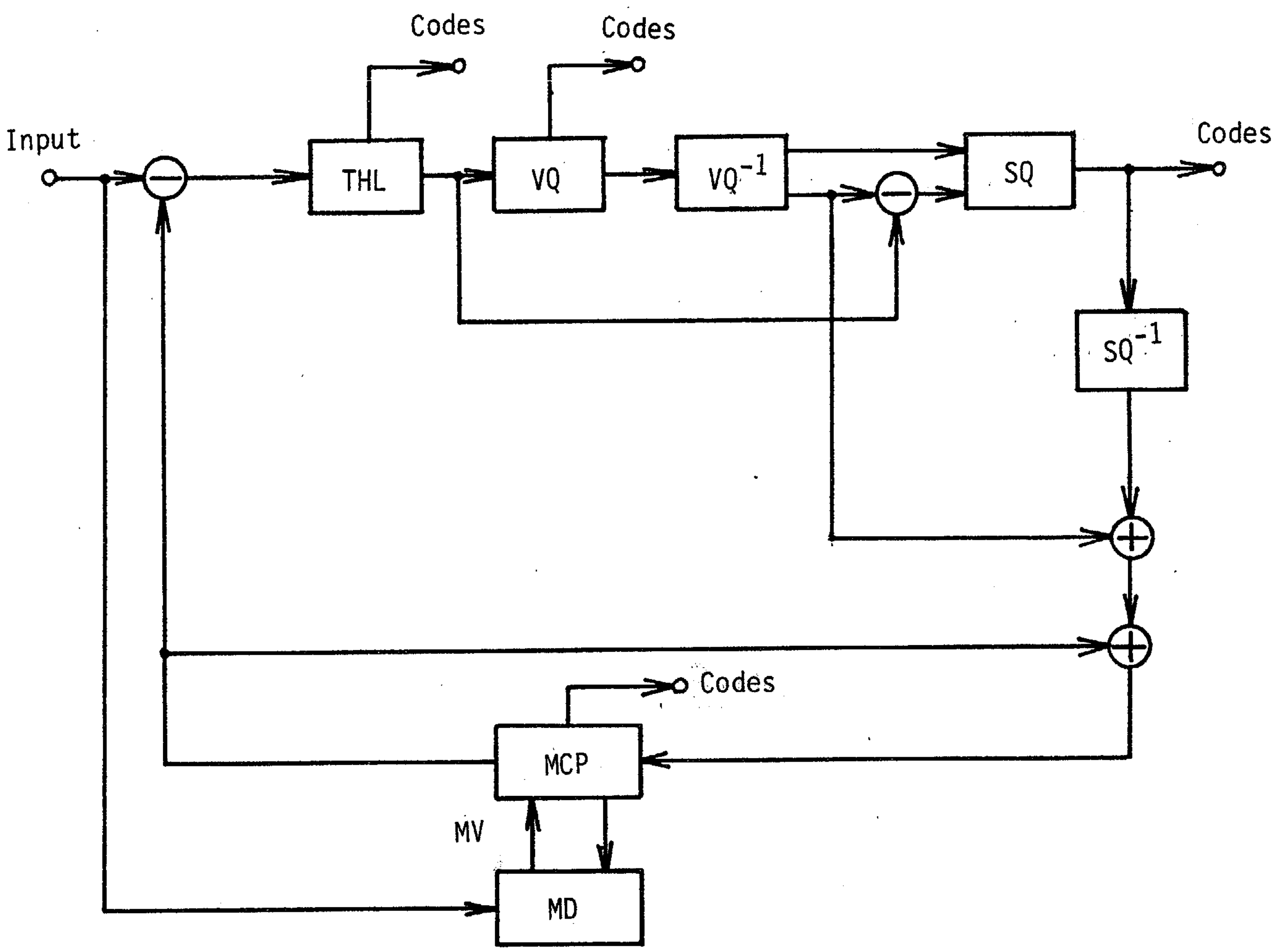
Note 2. For quantizer Q, vector quantization technique is able to be utilized as a variation as well as simple scalar quantization. See Fig.A3.

Fig. A2 Blockdiagram of DCT coding



DCT : Discrete cosine transform
VQ : Vector quantizer
VQ^{-1} : Vector quantizer decoder
SQ : Scalar quantizer
SQ^{-1} : Scalar quantizer decoder
DCT^{-1} : Inverse discrete cosine transform

Fig. A3 A variation of quantizer in DCT coding



THL : Significant/insignificant block detector
VQ : Vector quantizer
VQ^{-1} : Vector quantizer decoder
SQ : Scalar quantizer
SQ^{-1} : Scalar quantizer decoder
MD : Motion vector detector
MCP : Motion-compensated interframe predictor
MV : Motion vector

Fig. A4 Blockdiagram of VQ coding

Annex 2 of Document #43

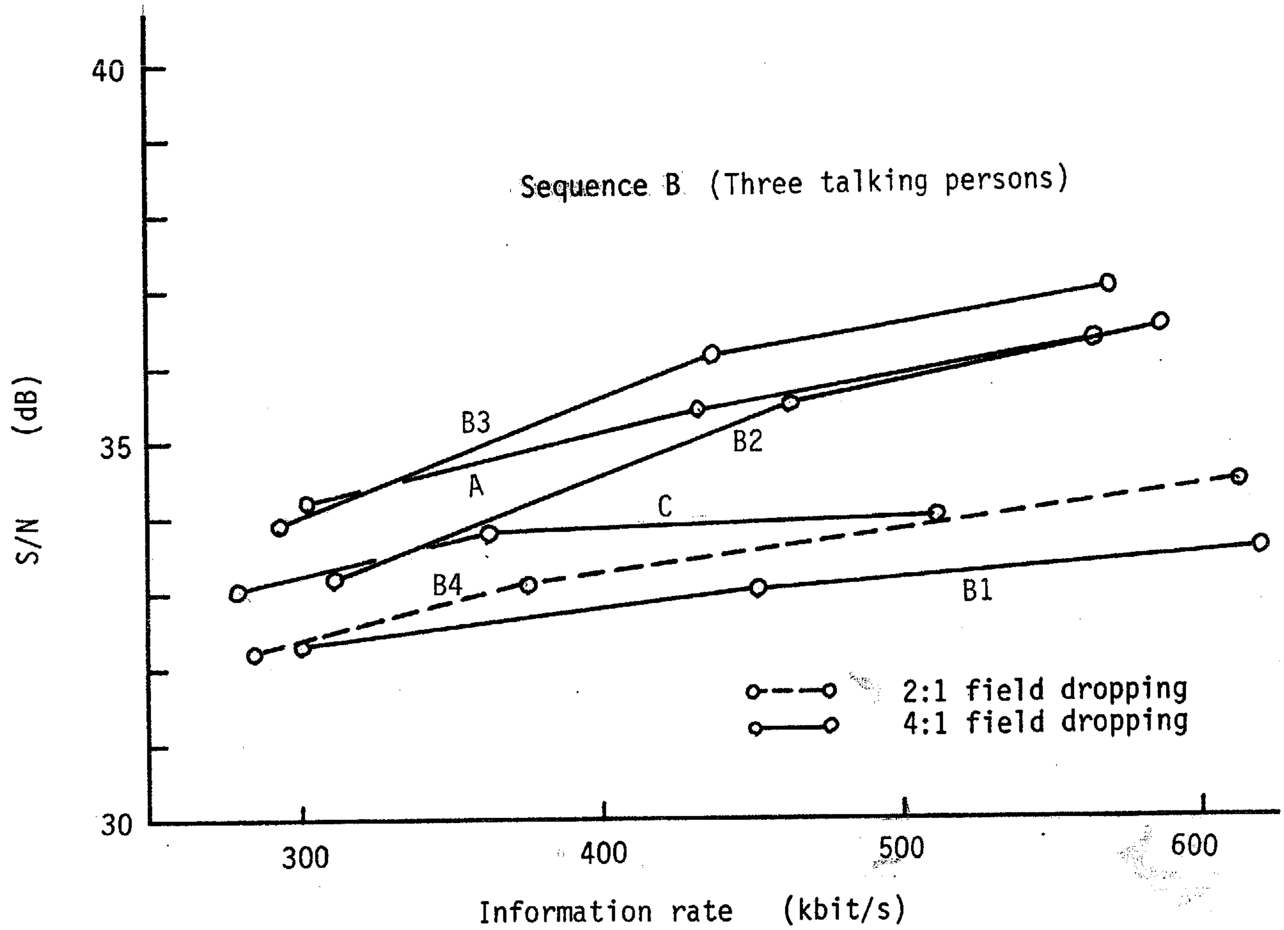
EXPERIMENTAL RESULTS OF THE CANDIDATE ALGORITHMS

This annex describes some experimental results of the candidate algorithms. Sequence B (Three talking persons), X (Split-screen) and Y (Trevor) are utilized for processing. Each sequence includes 26 frames, but the last 20 frames are displayed, because mid-gray level picture is used as the first frame prediction and first several frames are degraded by transient effect. S/N and entropy of the 25th frame data are shown in the following figure and tables.

Figure 1 shows the relationship between signal to noise ratio and entropy. Only coded fields are used for S/N calculation. From this figure, the differences of characteristics are not so large as far as S/N is concerned.

To clarify the ability of the candidates, we will demonstrate some decoded pictures in this meeting. Table 1 is a list for demonstration, in which signal to noise ratio, entropy and field dropping ratio are also described.

Entropy of each information described in Fig. 1 and Table 1 is derived from its probability function p(x), namely, -p(x)logp(x). In some variations, where run length coded entropy is introduced, run length probability function is calculated. Table 2 shows examples of information rate assigned for the candidate algorithms. Mark " * " represents run length coded entropy.



Note : Only coded field is used for S/N calculation.

Figure. 1 Experimental results

Table 1 Demonstration list

#	Candidates	Sequence	S/N* (dB)	Entropy (kbit/s)	Field dropping ratio	Comments
1	A	B	34.2	303	4:1	
2		Y	33.7	296	4:1	
3		X	32.6	634	4:1	
4	B1	B	32.3	299	4:1	Fundamental algorithm
5		Y	33.8	301	4:1	
6		X	33.0	496	4:1	
7	B2	B	33.2	310	4:1	DCT/without-DCT block detection and average separation
8		Y	34.8	275	4:1	
9		X	33.6	568	4:1	
10	B3	B	33.4	293	4:1	Adaptive filtering
11		Y	33.8	299	4:1	
12		X	31.7	482	4:1	
13	B4	B	32.2	286	2:1	Vector quantization and scalar quantization
14		Y	30.9	290	2:1	
15		X	32.2	421	2:1	
16	C	B	33.1	281	4:1	
17		Y	33.4	307	4:1	
18		X	33.2	484	4:1	

Note : Only decoded field is used for S/N calculation.

Table 2 **Assigned information rate for the candidates**

<table>
<tr><th colspan="2" rowspan="2">Items</th><th colspan="6">Percentage of information (%)</th></tr>
<tr><th>A</th><th>B1</th><th>B2</th><th>B3</th><th>B4</th><th>C</th></tr>
<tr><td colspan="2">Significant/insignificant pel and block identification information</td><td>11.0*</td><td>7.9</td><td>-</td><td>7.9</td><td>3.7</td><td>8.2</td></tr>
<tr><td colspan="2">Moving vector</td><td>16.9</td><td>11.8</td><td>1.8*</td><td>7.1</td><td>1.7*</td><td>5.6</td></tr>
<tr><td colspan="2">DCT / without-DCT block identification flag</td><td>-</td><td>-</td><td>5.5</td><td>2.7</td><td>-</td><td>-</td></tr>
<tr><td colspan="2">SQ / VQ block identification flag</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td><td>7.1</td></tr>
<tr><td colspan="2">Average value in a block</td><td>-</td><td>-</td><td>8.6</td><td>-</td><td>-</td><td>-</td></tr>
<tr><td colspan="2">Vector quantizing set index</td><td>-</td><td>-</td><td>-</td><td>-</td><td>5.7</td><td>70.9</td></tr>
<tr><td rowspan="6">Scalar quantizing index</td><td>for without DCT block</td><td>-</td><td>-</td><td>-</td><td>4.8</td><td>-</td><td>8.2</td></tr>
<tr><td>for DC block</td><td>-</td><td>6.8</td><td>9.1*</td><td>3.2</td><td rowspan="2">85.5</td><td>-</td></tr>
<tr><td>for AC block</td><td>-</td><td>73.6</td><td>75.0*</td><td>74.3</td><td>-</td></tr>
<tr><td>for P-pels</td><td>19.7*</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td></tr>
<tr><td>for I-pels</td><td>48.8</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td></tr>
<tr><td>for F-pels</td><td>3.6</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td></tr>
</table>

Note : Mark " * " represents run-length coded entropy.