#4

International Telegraph and Telephone Consultative Committee (CCITT) | COM XV-No D.2

Period 1985-1988 | English only

Question : 4/XV | Date : November 1984

STUDY GROUP XV - CONTRIBUTION No. D.2

SOURCE : United States of America

TITLE : Description of two of the Video Codecs used in the United States with Primary Rate Transmission Systems

Attached are descriptions of two of the video codecs used in the United States. One is provided by Compression Labs, Incorporated (ANNEX A) and one is provided by NEC America (ANNEX B). These are offered to the Specialists Group on Coding to further its work in addressing Part 3 of Recommendations H.120 and H.130.

ANNEX A

ANNEX A

1.0 INTRODUCTION

This contribution outlines a codec which is capable of transmitting 525 NTSC video, voice and data information over a 1.544 Mbps communications channel.

The codec is suitable for use in visual telephony and is offered for discussion relating to Part 3 of Recommendations H-120 an H-130.

2.0 COMMUNICATIONS

The communications portion of the system is described below. Following a general introduction, the communications frame structure is discussed in detail, covering the frame sync, format, system communications, audio, video field, high speed data, forward error correction, encryption and multipoint conferencing.

2.1 COMMUNICATIONS FRAME STRUCTURE

The communications frame provides a means of separating the various components that make up the teleconferencing signal. The proposed structure is based on the HDLC standard in which the various components are time division multiplexed and referenced to a synchronizing word. Figure 2.1 illustrates the format.

As a feature of the system, an identical structure is used via both a high speed serial data interface (RS449) and a T1/DS1 interface. In the standard form of the T1 bit stream, the data is inserted within the T1 frame so that there is no specific allocation of particular data, such as audio, to particular channel slots. The codec frame is asynchronous with respect to the T1 framing structure. This is illustrated in Figure 2.2.

In variation thereof, the audio is treated differently in the multi-point conferencing mode. This is described in section 2.1.9.

2.1.1 Frame Sync

The Frame Sync consists of a 9-bit pattern (001111101) which has been selected for its superior autocorrelation characteristics. The framing circuitry also ignores one bit errors in the received frame pattern for added error immunity.

2.1.2 Format

The format field consists of 3 bits of status information which has been error protected by encoding it into a (7,3) BCH code. This coding allows a single error correction/double error detection scheme.

The status information specifies whether user data (RS-232 or RS-449) is present in the data fields.

2.1.3 System Communications

Inter-codec commands and system status signaling are handled via a 16 bit field which also serves as a low speed (1200 baud max) RS-232 user data communications channel. A single bit in the format field serves to separate the command mode data from the RS-232 user data.

The command mode serves to coordinate field sizes for the high speed data (video and user channels), graphics frame transfers and handle status information transfers for the user data ports (RS-232, RS-449).

2.1.4 Audio

The audio channel is encoded at a 64KHz rate using a 176 bit field. The coding algorithm is covered in section 3.3.

2.1.5 Video Field

The video field consists of the following major subfields: Sync, header, buffer status, and video block data.

2.1.5.1 Sync

The sync word is a unique 16-bit pattern that signifies the start of a physical video frame or strip. A strip occurs once for every 16 lines of video (see section 3.1). Since physical frames occur asynchronously with respect to transmission frames, a sync word may occur anywhere within the video field. Bit insertion/deletion protects against the accidental occurrence of a sync pattern within the video block data.

2.1.5.2 Header

The header is a five-bit subfield that follows the sync. It helps to differentiate between a frame sync and a strip sync. It also signals various operating modes of the codec, such as scan mode, pixel density, full motion video or graphics mode.

2.1.5.3 Buffer Status

The buffer status subfield consists of a four-bit code that indicates how full the bit rate buffer is (see section 3.1.6).

2.1.5.4 Video Block Data

The video block data consists of a concatenation of variable-length Huffman codes with block separators. The block separators consist of a motion status code at the beginning of each block and an end-of-block (EOB) code at the end. The Huffman codes, if present, indicate the amplitudes and position of any significant DPCM coefficients (see section 3.1.5).

2.1.6 High Speed Data

Two full duplex data ports on each codec communicate via this optional field. The presence and source of the data is indicated by bits in the format field. Field size is determined by the number of channels and their aggregate bit rate. The user data field, when allocated, is at the expense of the video field:

VIDEO + USER DATA = 3840 BITS

In order to facilitate buffer control, every eighth bit in the user data field is a buffer status bit used to indicate an end of message condition.

The user data rates run from 9.6 Kbps up to 448 Kbps in each channel, limited by total available bandwidth.

2.1.7 Forward Error Correction

As the control field contains its own error checking, only the command, audio user data and video data fields are error corrected. The algorithm used is a single burst error correction code (4080, 4098) capable of correcting any burst error up to 9 bits by using a 32 bit syndrome. This syndrome is generated via a Burton code polynomial ($x^{32} + x^{24} + x^{22} + x^{20} + x^8 + x^6 + x^4 + 1$) and forms the last field in the frame.

2.1.8 Encryption

The system communications field and all data fields may be encrypted with a 56 bit user defined key.

The DES algorithm is employed in the Cipher Feed Back mode which links the 64 bit encryption blocks.

2.1.9 Multipoint Conferencing

In a multi-point conference, the audio is treated differently than described above. The audio field is removed from the communications frame as illustrated in Figure 2.3. In the case of communications over a T1 network, the data stream is inserted in channels one through 23 as illustrated in Figure 2.4 . This allows the audio to be separately inserted in channel 24 outside of the codec encryption and forward error correction field. It may then be extracted at a nodal point and combined with audio from many sources before being distributed.

In a satellite network employing earth stations on customer premises, the audio may be carried on a separate SCPC or TDMA channel and combined at a nodal point as above.

The method for coding and controlling audio in a multi-point conferencing environment is left for further study.

3.0 CODING

This section discusses the methods employed to code and decode the video and audio data in terms ot video processing and sampling, input memory and input processing, the cosine transform process, DPCM coding, Hutfman coding, rate buffer control and output processing.

3.1 VIDEO CODING

Figure 3.1 illustrates the major subsystems in the video coding section.

3.1.1 Video Processing and Sampling

In this portion of the system, the composite NTSC signal is separated into three components, one luminance (Y) and two chrominance components (R-Y and B-Y) and then converted into digital form. The sample rate used for luminance is twice the color subcarrier frequency (7.2 MHz) and chrominance is half the subcarrier frequency (1.8 MHz). This provides a luminance bandwidth of 2.75 MHz and a chrominance bandwidth of 680 KHz.

As it is only necessary to code and transmit the active video information, some of the sampled data is discarded. Of the 455 luminance samples per line, 386 are coded. Of the 114 chrominance samples per line, 96 are coded. A similar economy is performed in the vertical dimension. Of the 525 lines in an NTSC signal only 480 luminance lines and 512 chrominance lines are transmitted. Figures 3.2 and 3.3 illustrate this video format. The area of chrominance sampling is greater than that of luminance. After subsampling by 4:1 vertically (resulting in 128 lines), an integer number of a 16 x 16 pixel chrominance block is obtained.

3.1.2 Input Memory and Processing

This subsection performs a scan conversion on the video signal. The luminance frame of 480 lines by 368 pixels is re-scanned as 690 blocks of data, each containing 256 pixels arranged in a 16 X 16 pixel format. The position of each block with respect to a frame is illustrated in Figure 3.3.

The chrominance information is vertically subsampled four to one and scan converted into 48 blocks of R-Y data and 48 blocks of B-Y data, each 16 X 16 pixels in size. To further reduce bandwidth, the video frames are subsampled 2:1. This produces a data stream of 15 frames per second.

To enhance performance, the signal may be temporally filtered and noise reduced. This is optional and does not effect inter-operability.

3.1.3 Cosine Transform

A two-dimensional cosine transform is performed on the 16X16 pixel blocks. Both luminance and chrominance blocks are treated identically. This operation allows the removal of large amounts of redundant spatial information. The one-dimensional cosine transform for N points is given by

$$F(u) = 2C(u)/N \sum_{i=0}^{N-1} f(i)\cos[(2i+1)u\pi/2N]$$

for u = 0,1, ... , N-1, where

$$C(u) = \begin{cases} 1/\sqrt{2} & \text{for } u = 0 \\ 1 & \text{for } u = 0. \end{cases}$$

The inverse transform is given by

$$f(i) = \sum_{u=0}^{N-1} C(u)F(u)\cos[(2i+1)u\pi/2n]$$

When the forward transform is performed both horizontally and vertically on a square block of pixels, a block of coefficients is formed in which the coefficient in the upper left corner represents the d-c energy, or average brightness within the block; coefficients progressing to the right represent increasing horizontal spatial frequencies; and coefficients progressing downwards represent increasing vertical spatial frequencies. The cosine transform has been found to have superior energy compaction properties for typical images, as compared with all known transforms possessing fast computational algorithms.

3.1.4 DPCM Coding

An adaptive DPCM loop is used to remove both spatially redundant and temporally redundant information from the stream of transform coefficients. The DPCM-coded coefficients are quantized to a degree determined by rate buffer fullness. Sixteen degress of quantization severity are used, according to a four-bit buffer fullness code which is updated for every horizontal strip of blocks.

3.1.5 Huffman Coding

The quantized DPCM coefficients are scanned in a diagonal zig-zag within each block, starting from the d-c term in the upper left corner of the block and ending with the highest frequency term in the lower right corner. The amplitude of each nonzero DPCM coefficient is assigned a variable-length Huffman code derived from the statistics for typical teleconferencing images. Runs of zero amplitudes are coded from a separate table of Huffman codes. A run that extends to the end of the block is replaced with a single end-of-block code.

3.1.6 Rate Buffer Control

A rate buffer memory is used to regulate the flow of bits onto the channel. A four-bit word representing buffer fullness is updated for every horizontal strip of blocks and included in the video codestream. This word is used by the forward and inverse DPCM loops to control the degree of quantization used.

3.2 VIDEO DECODING

The video decoding process is complementary to the coding process, i.e., the functions performed by the Forward Transform logic and the Coding subsystems detailed above (sections 3.1.1 through 3.1.6) are basically reversed.

3.2.1 Output Processing

In the Output Memory subsystem, the scan conversion is reversed and the output frames are interpolated to restore a 30 frame per second video signal. The output video processing encodes the Y, R-Y and B-Y into an NTSC signal with conventional syncs and blanking.

3.3 AUDIO CODING

(for further study)

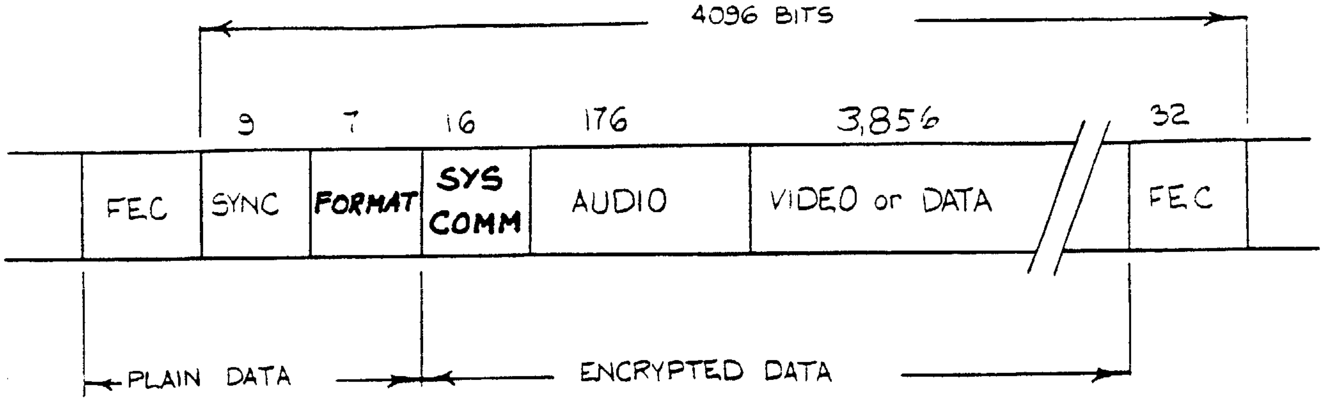


FIGURE 2.1
COMMUNICATIONS FRAME FORMAT

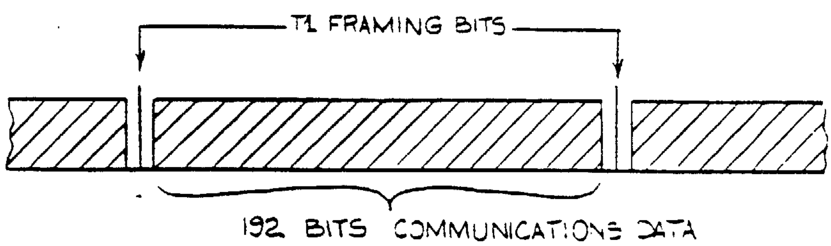


FIGURE 2.2
T1 FRAME FORMAT

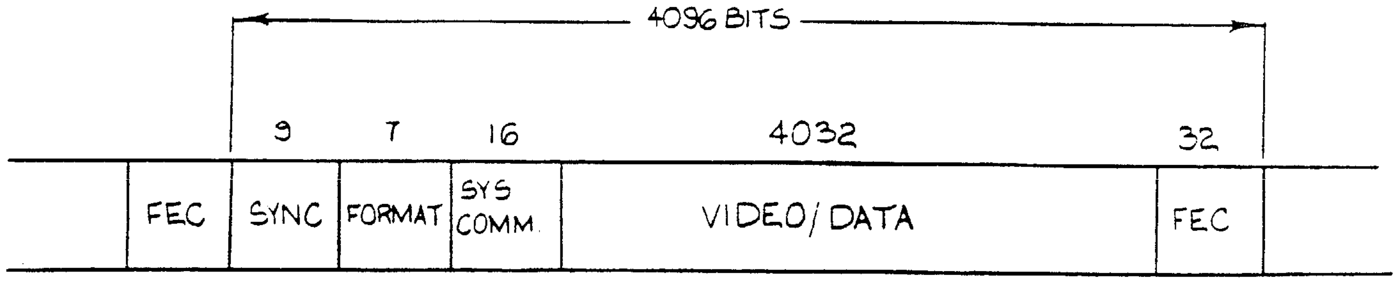


FIGURE 2.3
COMMUNICATIONS FRAME FORMAT - MULTIPOINT CONFERENCING

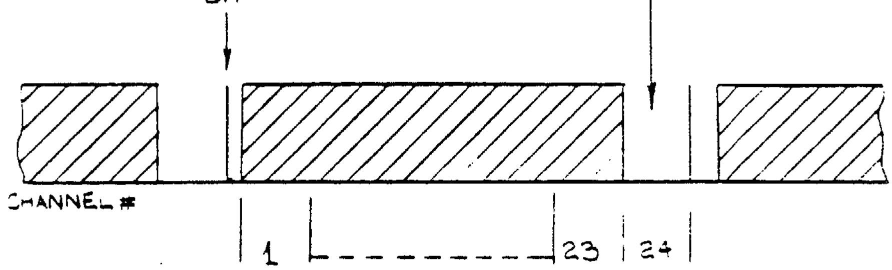
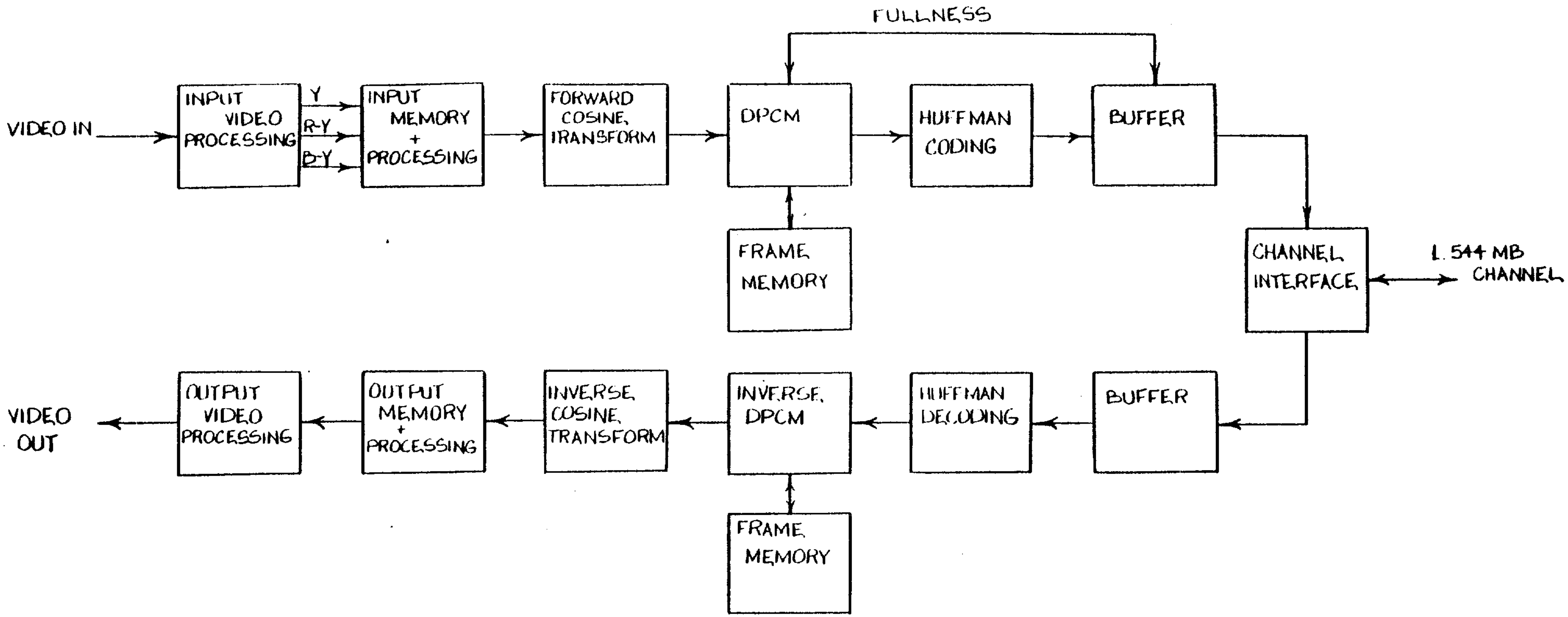


FIGURE 2.4
T1 FRAME FORMAT - MULTIPOINT CONFERENCING



VIDEO CODEC BLOCK DIAGRAM

FIGURE 3.1

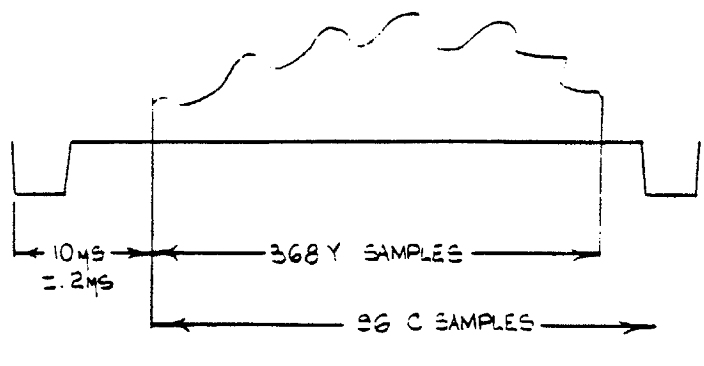


FIGURE 3.2
VIDEO SAMPLING WINDOW

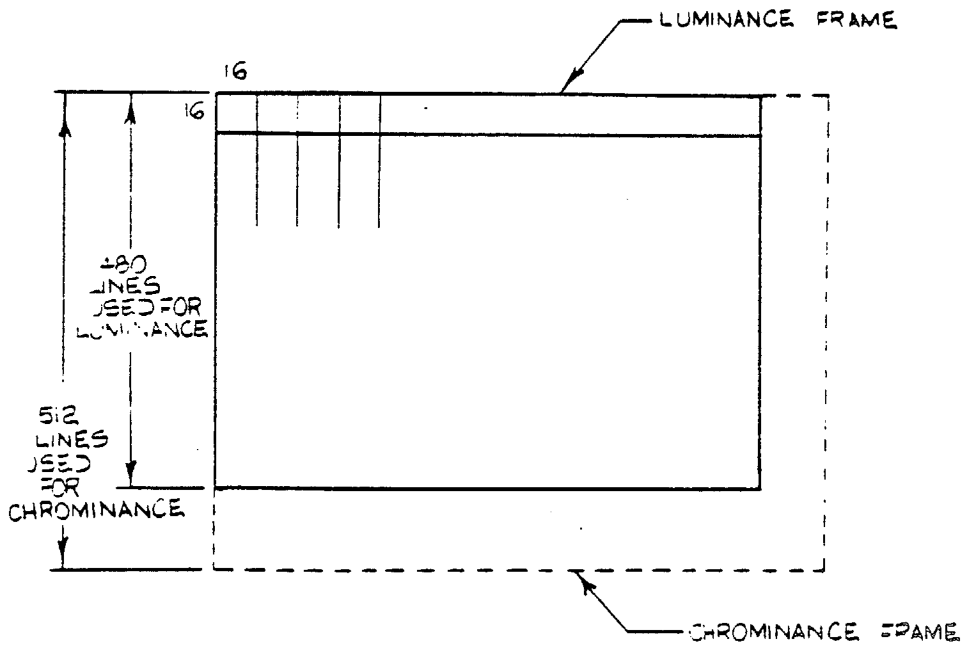


FIGURE 3.3
VIDEO FRAME FORMAT

ANNEX B

ANNEX B

1. Introduction

CCITT Recommendations H.120 and H.130 have been established on the codec and frame structure for videoconferencing for Part 1 and Part 2a. However, Part 3 for intra-regional 1544 kbit/s transmission of videoconferencing signals that employ 525 lines and 60 field/s standard is still under study.

This contribution proposes a codec and frame structure for completing the Part 3 of Recommendations H.120 and H.130.

2. Outline

2.1 Overall Configuration

The codec overall blockdiagram is shown in Fig.1. The coder (Fig.1(a)) is composed of three blocks, video source coding, transmission coding and audio source coding.

In the video source coder, the input analog NTSC signal is first digitized, color decoded, and time division multilplexed. The prefilter removes noise components. The prefilter output is adaptive predictive encoded to remove redundancy. The prediction error signal is then entropy encoded. The entropy-encoded error signal is fed to buffer memory to smooth out the data rate to be transmitted. Buffer occupancy information is fed back to the adaptive predictive coder to prevent buffer overflow.

In the transmission coder, encoded video and audio signals are encrypted on an optional basis, and then forward error corrected and scrambled. The encoded video, encoded audio and optional user data signals are multiplexed into 1544 kbit/s digital data.

In the audio source coder, the input audio signal is encoded into 64 kbit/s data. The encoded audio signal is delayed to compensate for the video source coding delay, and fed to the transmission coder.

The decoder (Fig.1(b)) carries out the reverse operation.

2.2 Video Input/Output

525/30 NTSC composite color/monochrome television signal.

2.3 Digital Output/Input

Conforming to CCITT Recommendation G.703.

2.4 Digital Video Format

Video sampling frequency is 4 fsc (color subcarrier frequency). The sampling clock is locked to the horizontal synchronization of the NTSC signal. An NTSC composite signal is converted to luminance (Y) and chrominance(C_1 and C_2) components, which are then multiplexed to TDM (Time Division Multiplex) format.

2.5 Coding Algorithm

Adaptive predictive coding:

(1) interframe prediction with motion compensation, and

(2) intraframe prediction

Entropy Coding:

(1) variable word-length coding, and

(2) run-length coding

2.6 Mode of Operation

Full duplex. Also one-way broadcasting mode is available.

2.7 Audio Channel

Monaural audio signal (50~7000 Hz).

2.8 Data Channel

Optional 64 kbit/s user data channel.

2.9 Transmission Error Correction

BCH forward error correcting code.

Demand refresh mode available.

2.10 Additional Facilities

Encryption, graphics mode and multipoint communication.

2.11 Codec Processing Delay

Approximately 165 ms.

3. Video Interface

The codec video interface is 525/30 NTSC color/monochrome television signal. Nominal amplitude is 1 V_{p-p} at 75 ohm unbalanced.

4. Video Source Coding

4.1 Analog-to-digital and Digital-to-analog Conversion

(1) Sampling frequency f_s=910f_H

(f_H:Horizontal scanning line frequency),

(2) Precision=8 bits/sample(Linear PCM, two's complement),

(3) Level Definition;

Sync tip level(-40 IRE)=-124(10000100),

White level (100 IRE)= 72(01001000).

4.2 Color Decoding and Encoding

(1) Digital Demodulation and Separation at the Encoder.

Luminance component : Y

Chrominance components : C_1/C_2.

(2) Digital Incorporation and Modulation at the Decoder.

4.3 TDM Signal (see Fig.2)

(1) Time compression of C_1 and C_2

Compression Ratio = 6:1

(2) Horizontal blanking interval of Y is substituted line-sequentially for either of the C_1 and C_2 signals.

(3) TDM signal consists of Burst (7 samples), Color signal (64 samples of C_1 or C_2), and Luminance signal (384 samples).

4.4 Pre- and Post-filtering

(1) At the Encoder;

- Temporal filtering for noise reduction,
- Spatial filtering for folding-over suppression in subsampling.

(2) At the Decoder;

- Spatial filtering for interpolating eliminated samples in Subsampling,
- Spatio-temporal filtering for interpolating eliminated fields in Field Repetition.
- Temporal filtering for reduction of noise due to video source coding.

4.5 Predictive Coding

4.5.1 Source Coder and Decoder (See Fig.3)

(1) At the Coder;

- Motion Vector Detection,
- Predictive Encoder based upon adaptive intra-interframe prediction with motion compensation,
- Entropy Coder for encoding prediction error signals, motion vector information, and coding mode commands.

(2) At the Decoder;

- Entropy Decoder,
- Predictive Decoder.

(3) Demand Refresh Capability for restoring identity of frame memory contents at the coder and the Decoder.

4.5.2 Adaptive Prediction (See Fig.4 and Table-1)

(1) Adaptive intra-interframe prediction with motion compensation

- sample-by-sample prediction selection

(2) Prediction Selection Rule

Selection of Intraframe Prediction ($\hat{x}_I = Z^{-1}$)

when s=1

- Selection of Motion-compensated Interframe Prediction ($\hat{x}_M = Z^{-F+V}$) , when s=0.

- $s = R_a \cdot R_b$,

where R_a is a reference signal for a previous sample, and R_b for a previous line sample.

Definition of R;

$$R_x = \begin{cases} 0 \text{ if } |\hat{x}_I - \hat{x}| \geqq |\hat{x}_M - \hat{x}|, \\ 1 \text{ otherwise,} \end{cases}$$

where $\hat{x}$ is local decoder output signal for input x.

(3) For Subsampling mode,

$\hat{x}_I = Z^{-2}$,

$S = R_c \cdot R_d \cdot R_e$, when samples a and b are eliminated.

(4) For Refresh mode,

$R_x = 1$ for all samples in lines under Refresh mode.

4.5.3 Motion Compensation (See Fig.5)

- Block size: 8 lines x 16 samples
- Tracking range: ±7lines (vertical); ±15samples(horizontal)
- Motion Vector Direction

Positive for delay increasing direction in Interframe prediction loop.

4.5.4 Quantization (See Table-2)

(1) Four quantizing characteristics

Q_0 (57 levels), Q_1 (57), Q_2 (51) and Q_3 (37).

4.5.5 Subsampling

(1) Subsampling using 2:1 line quincunx pattern.

(2) Interpolation of eliminated samples

- Four neighboring samples (two-dimensional) for luminance signal interpolation.
- Two horizontally adjacent samples for color signal interpolation.

(See Table 1)

4.5.6. Field Repetition (See Table 1)

(1) 2:1 Field Omission

(2) Interpolation of omitted fields by vertical average using two lines in the previous field.
Same component (C_1 or C_2) in the previous field is used for color signal interpolation.

4.5.7. Stop of Coding

(1) Coding operation halt for excessive generated information rate in order to prevent buffer memory from overflowing.

(2) Both prediction error signal and motion vectors are forcibly set zero.

4.5.8 Refresh

(1) Refresh for renewal of frame memory contents.

(2) Intraframe prediction for Refresh

(3) Two types of Refresh

- Demand Refresh for communication with backward channel
- Cyclic Refresh for broadcasting type communication without backward channel.

4.6 Coding Parameter Control

(1) Normal mode

full sampling

(2) Subsampling

2:1 sample elimination

(3) Field Repetition

2:1 field elimination

(4) Stop

Coding halt

(5) Refresh

renewal of frame memory contents

4.7 Entropy Coding

4.7.1 Entropy Coder (See Fig. 6)

(1) Coding of Mode Commands

(2) Compression of prediction error data

(3) Compression of motion vector data

(4) Multiplexing

4.7.2 Coding Mode Commands (See Fig. 7)

(1) Data Structure

- Frame Sync
- Frame Mode Data
- Line Sync
- Line Mode Data
- Motion Vector Data
- Prediction Error Data

(2) Frame Mode Data

- Frame Position to indicate the head line position in a video frame
- Delay time in coder buffer memory
- Demand Refresh (Execution and Request)
- Color/Monochrome State
- Cyclic Refresh to indicate lines to be refreshed
- Mode Flag for extension of mode information

(3) Line Mode Data

- Quantizing characteristics $Q_0 \sim Q_3$
- Refresh
- Field Repetition
- Line Skip to indicate that there is a code representing the number of lines which can be skipped.

(4) Motion Vector Data

- Compressed Motion Vector Data preceded by its number (in Bytes)
- Entropy Coding

(5) Prediction Error Data

Entropy Coding

4.7.3 Prediction Error Coding

(1) Variable Word-length coding for non-zero prediction error signal. (See Table 3)
The same code set is used for all quantizing characteristics.

(2) Adaptive code switchig between Variable Word-length (V) code and Fixed Word-length (F) code on a sample-by-sample basis

·Long V code is followed by F code

·F code representing small amplitude error is followed by V code

·Switching threshold between Level Numbers 4 and 5. (See Fig. 8)

(3) Run-length coding for zero prediction error signal (See Table 4)

·Short run (Run-length RL<64) is expressed using a single code (Terminate code).

·Long run (RL>65) is expressed using two codes, Make-up (MK) and Terminate codes.

(4) Scanning method conversion from TV scan to Block scan, where scanning is completed within a two-dimensional block, thus resulting in longer runs.

(5) Bridging of isolated zero amplitude error signal.

V_0 or F_0.

(6) Eliminated samples in Subsampling are not coded.

4.7.4 Motion Vector Coding

(1) Coding of vector difference

·Difference between two successive vectors

(2) Variable Word-length coding for non-zero difference vector (See Fig. 9, Table 5)

(3) Run-length coding for zero difference vector (See Table 5)

(4) Transient code (TRANS) is inserted when Subsampling is turned on or off. The first TRANS code indicates start of Subsampling and the second end of it, and so on. (See Fig. 10)

4.8 Buffer Control

(1) Delay time at the coder buffer memory is transmitted as a part of Frame Mode Data. Buffer memory operation at the decoder is controlled so that the total delay time at the coder and decoder is kept constant.

(2) Buffer memory size

180kbit/s for the coder buffer,

more than 220kbit/s for the decoder buffer memory.

(Total delay time is about 165ms.)

(3) Underflow/Overflow Prevention

For Underflow prevention, each zero amplitude prediction error signal is represented with variable word-length code V_0 instead of run length coding.

For Overflow prevention, Stop Mode is applied, resulting in zero amptitude for all prediction error signal and motion vector.

5. Audio Coding

Audio signal bandlimited to 7kHz is encoded into 64 kbit/s. The encoding scheme is, at present, subband coding with ADPCM (SG. XVIII Delayed Contribution SE). However, provision is made for future use of more efficient coding.

6. Transmission Coding

6.1 Encryption

Video and audio signals can be independently encrypted.

6.2 Error Correction

(255, 239) double error correcting BCH code, along with interleaving.

6.3 Scrambling

8 stage scrambler.

6.4 Frame Structure

Details are shown in Annex.

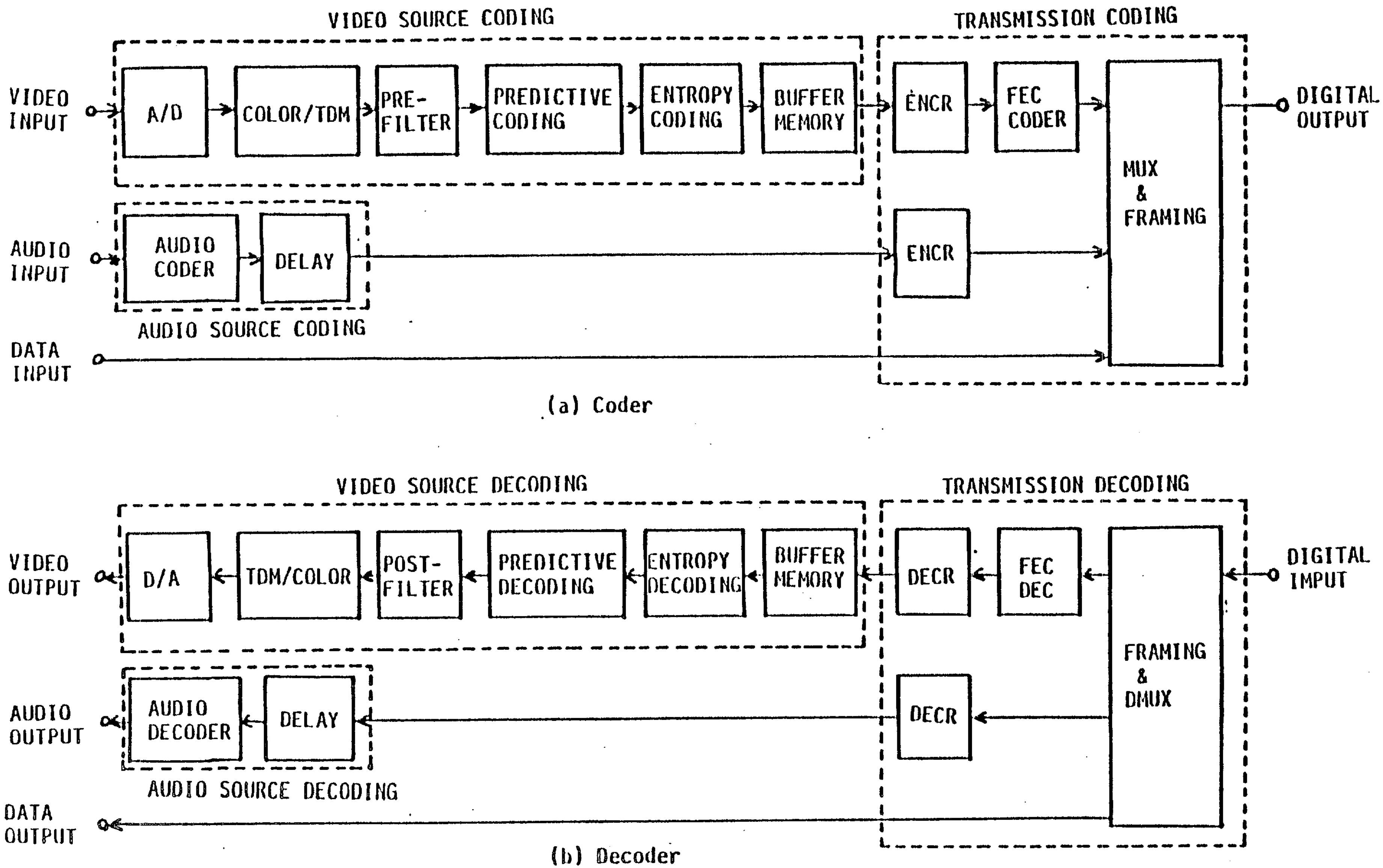


FIGURE 1. Codec blockdiagram

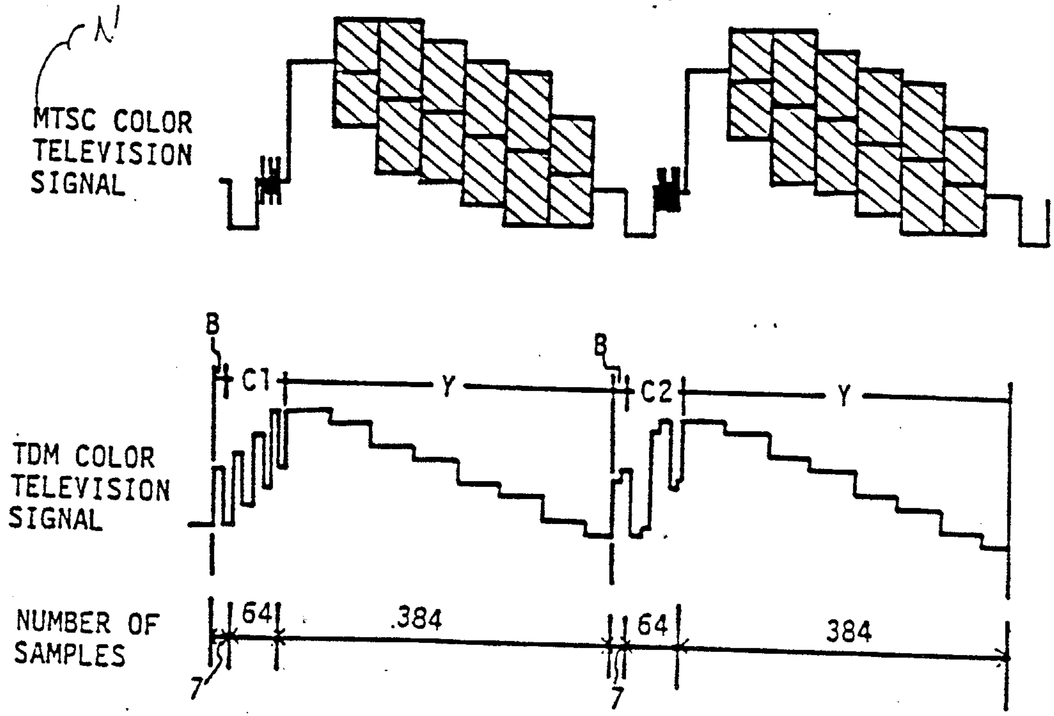
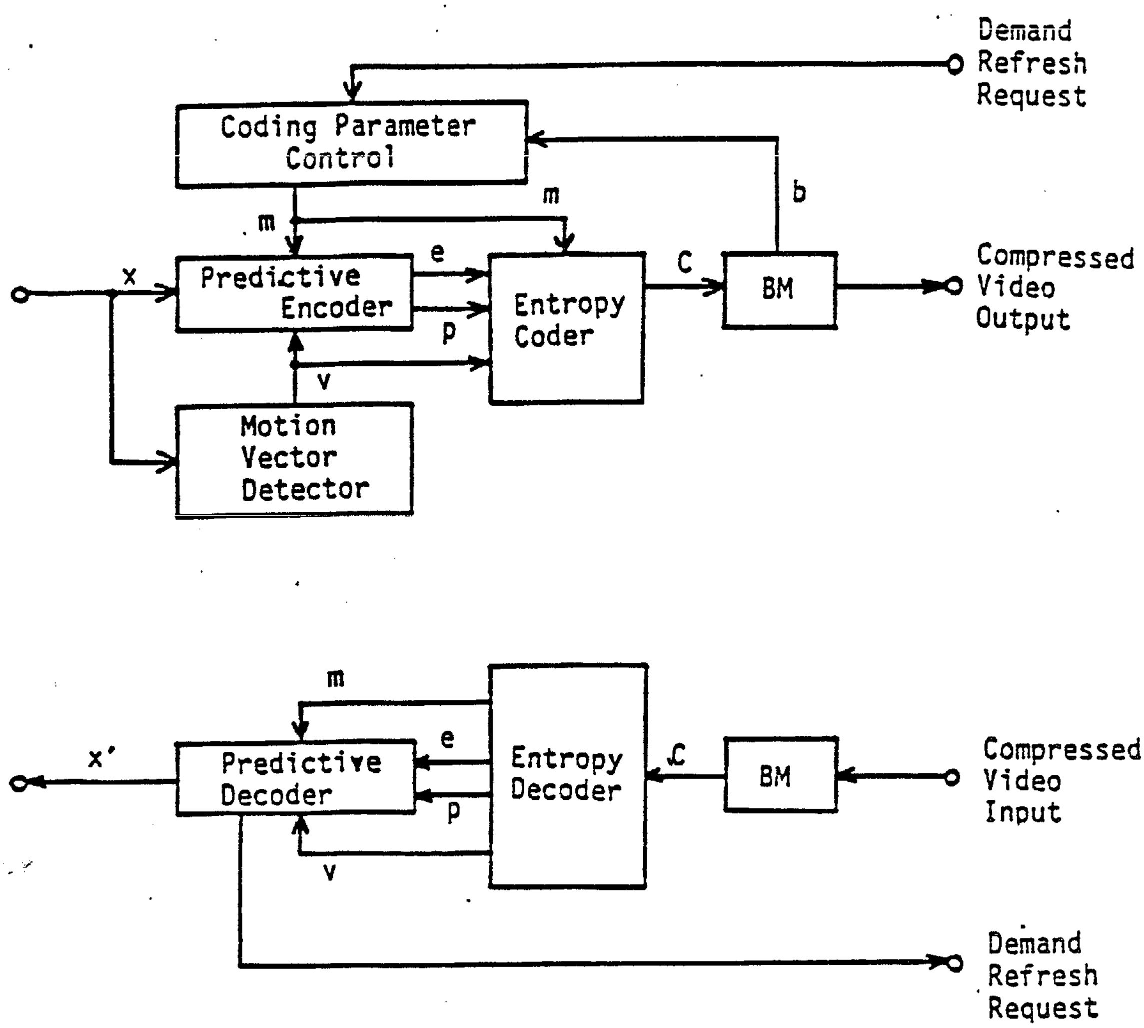


FIGURE 2. TDM Color Signal Format



x : input video signal
x': output video signal
m : coding mode
e : prediction error
v : motion vector
p : parity check information
C : compressed data
b : buffer memory occupancy imformation
BM : Buffer Memory

FIGURE 3. Source coder and decoder

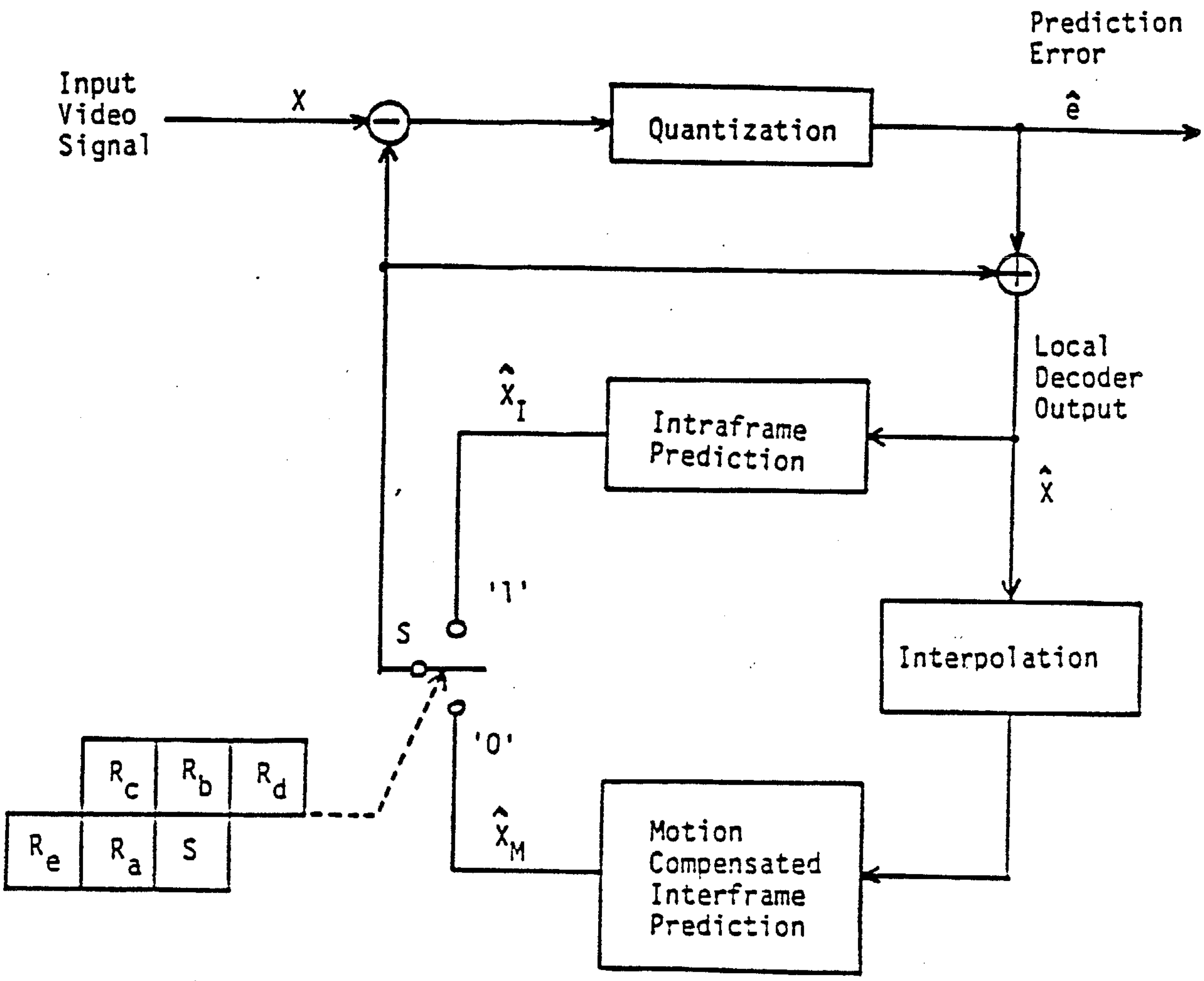


FIGURE 4. Adaptive Prediction

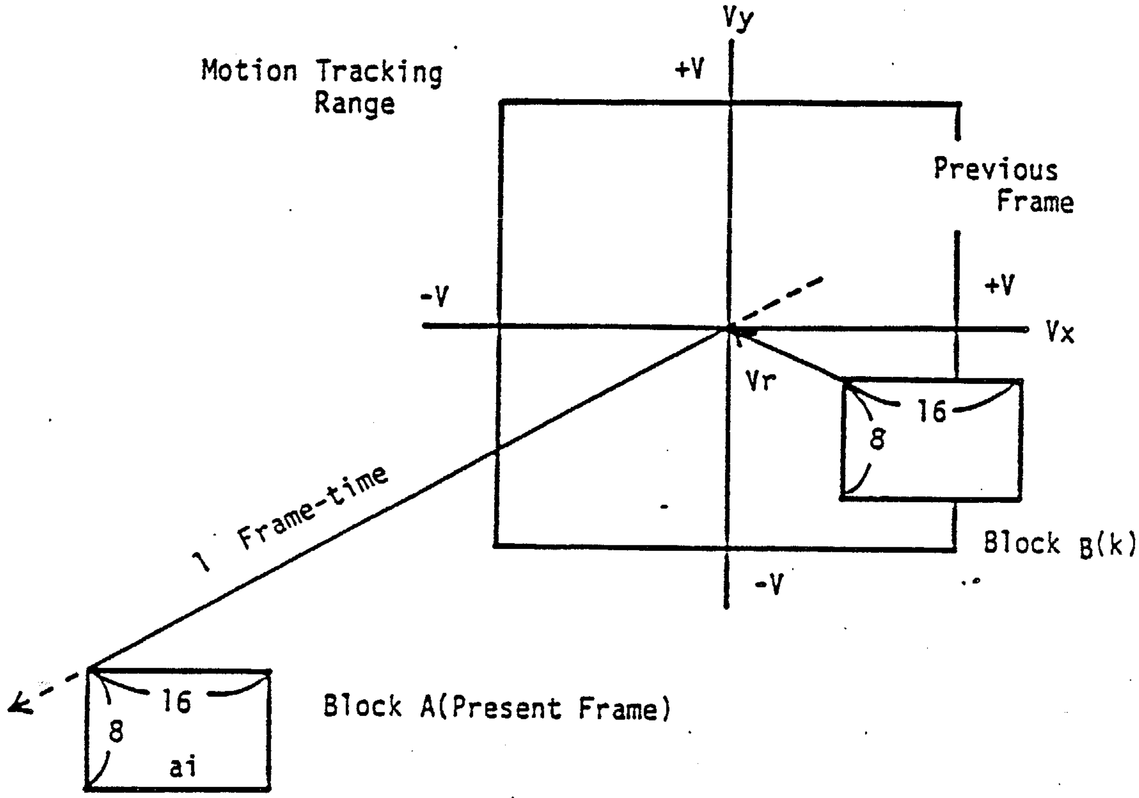
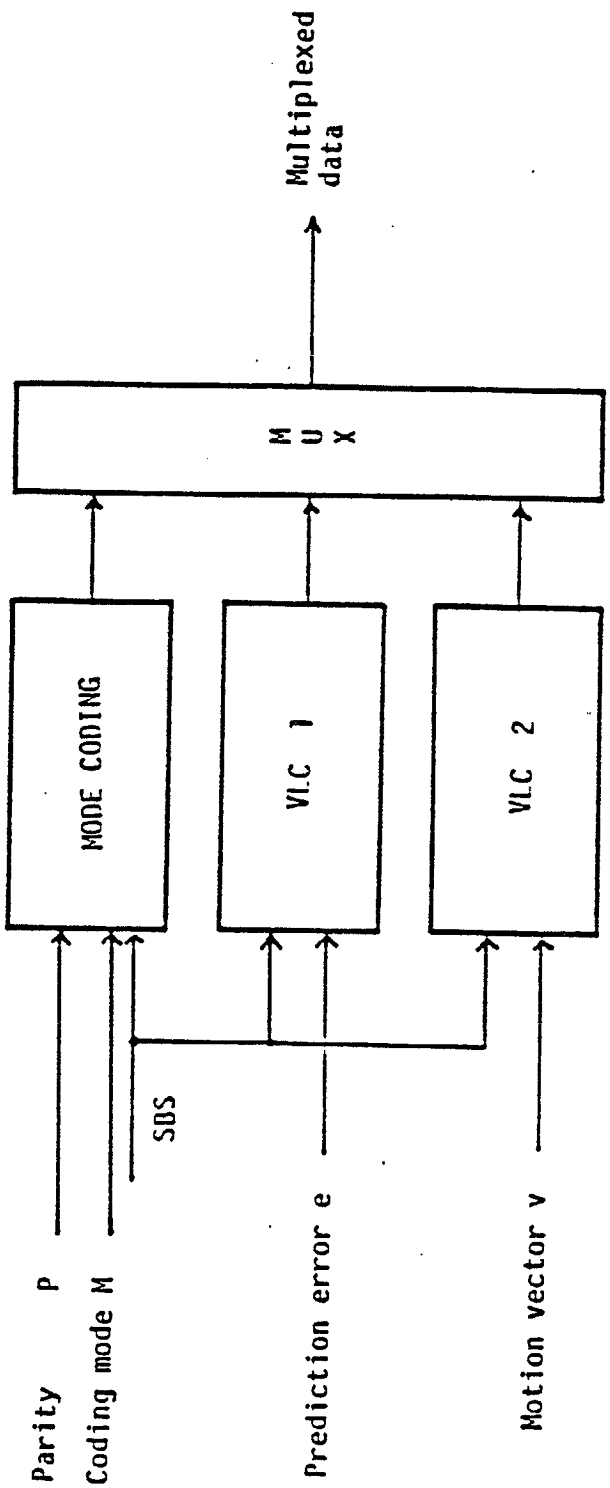
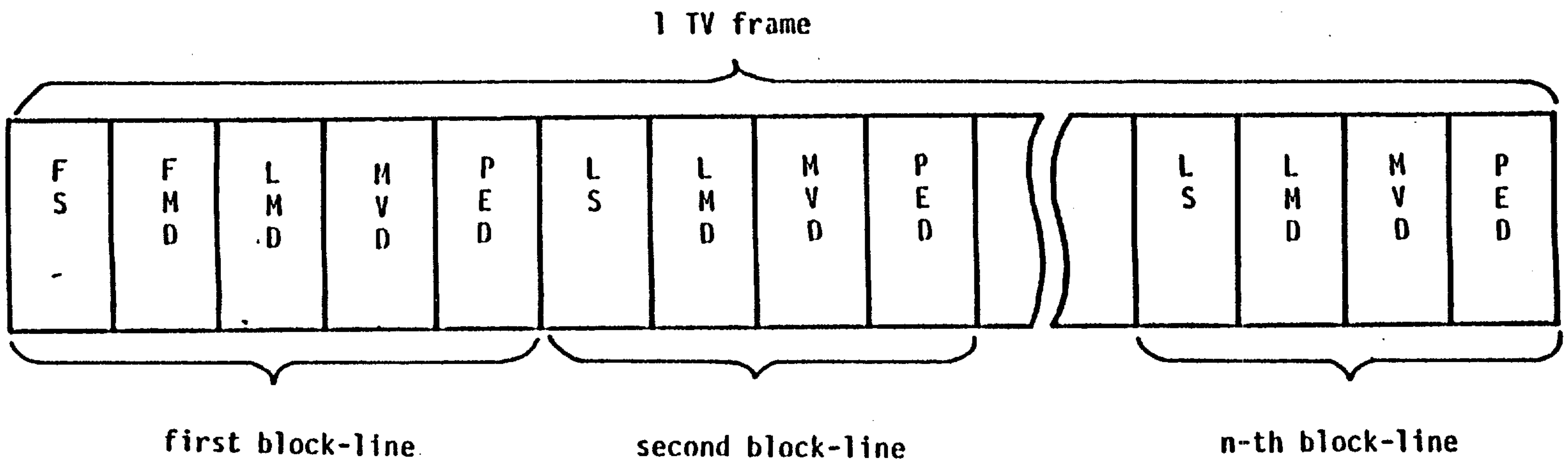


FIGURE 5. Motion Vector Detection



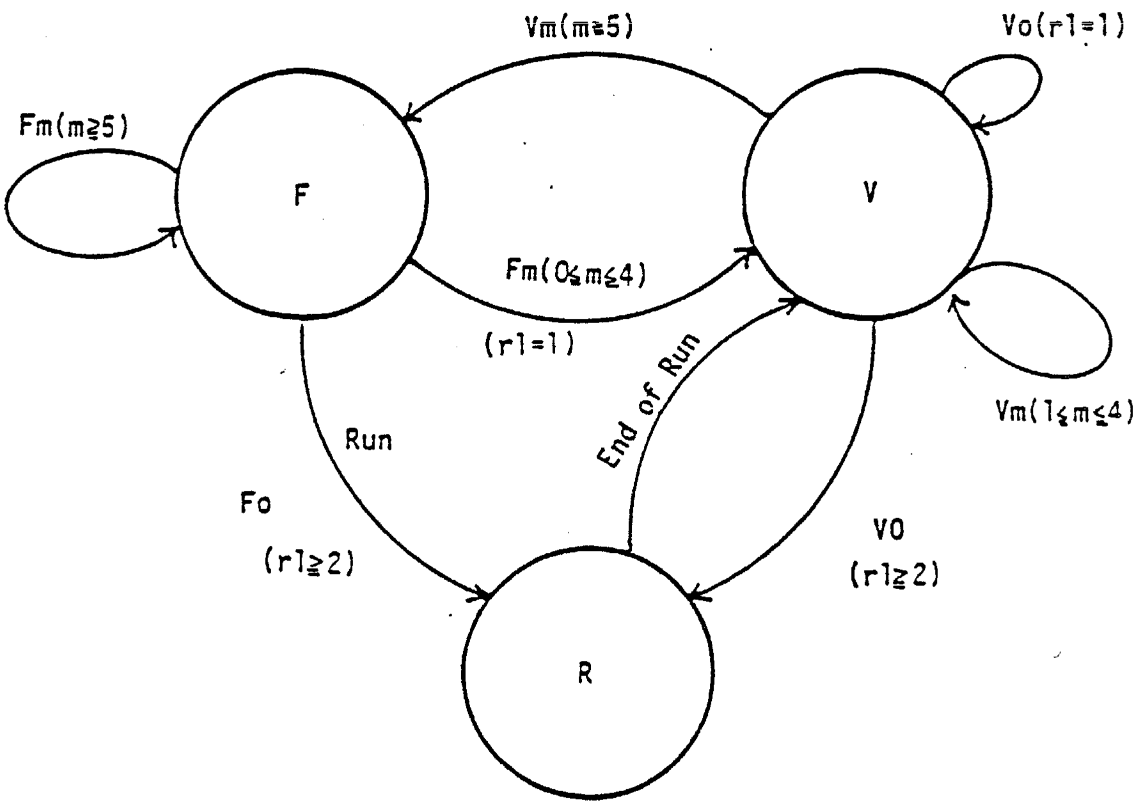
SBS : Subsampling
VLC : Variable Wordlength coder

FIGURE 6. Entropy coder



FS : Frame Sync
FMD : Frame Mode DATA
LS : Line Sync
LMD : Line Mode Data
MVD : Motion Vector Data
PED : Prediction Error Data

FIGURE 7. Multiplexed data format



Note : Run-length is represented by "rl".

Figure 8. Code Transition Rule for Prediction Error

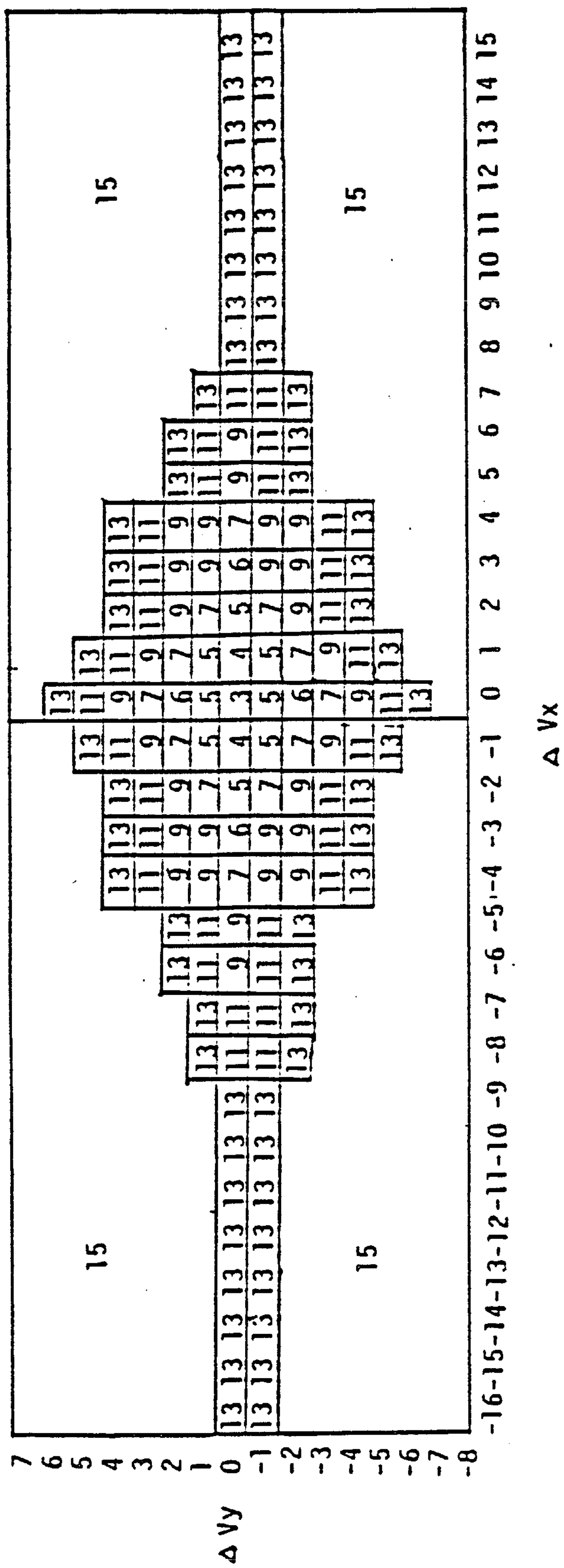
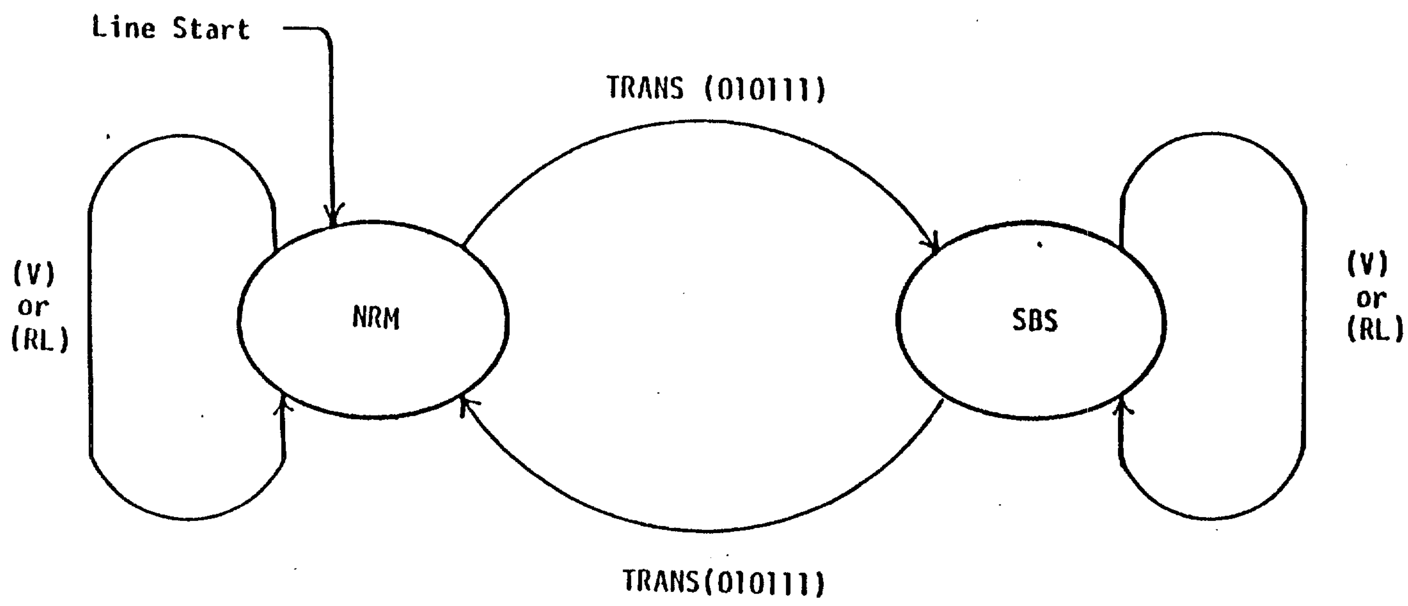


FIGURE 9. Word length for motion vector prediction errors



Note : The same motion vector coding method is applied for both NRM and SBS modes.

FIGURE 10. Code Transition rule for motion vector data and Normal/Subsampling mode change

TABLE 1. Prediction and interpolation functions.

Coding Mode	Kind of Pel	Prediction Function P(Z): $P_Y(Z)$	$P_C(Z)$	$P_B(Z)$	Interpolation Functions I(Z): $I_Y(Z)$	$I_C(Z)$	$I_B(Z)$
Normal	Coded	Z^{-1} ; S=1 Z^{-F+V} ; S=0		Z^{-F}	1		
Subsampling	Coded	Z^{-2} ; S=1 Z^{-F+V} ; S=0			1		
Subsampling	Omitted	(not defined)			$\frac{1}{2}\left\{\frac{1}{2}(Z^{-1}+Z^{+1}) + \frac{1}{2}(Z^{-H}+Z^{+H})\right\}$	$\frac{1}{2}(Z^{-1}+Z^{+1})$	
Field Repetition	Omitted	(not defined)			$\frac{1}{2}(Z^{-262H}+Z^{-263H})$	Z^{-283H};first field Z^{-282H};second field	
Refresh NRM	Coded	Z^{-1}			1		
Refresh SBS	Coded	Z^{-2}			1		
Refresh SBS	Omitted	(not defined)			$\frac{1}{2}(Z^{-1}+Z^{+1})$		

TABLE 2. Quantizing characteristics

Q0

Input Range	Output Level
0 ~ 1	0
2	1
3	2
4 ~ 5	3
6 ~ 7	5
8 ~ 9	7
10 ~ 11	10
12 ~ 14	13
15 ~ 17	16
18 ~ 20	19
21 ~ 23	22
24 ~ 26	25
27 ~ 29	28
30 ~ 32	31
33 ~ 37	35
38 ~ 42	40
43 ~ 48	45
49 ~ 54	51
55 ~ 60	57
61 ~ 67	64
68 ~ 74	71
75 ~ 81	78
82 ~ 88	85
89 ~ 95	92
96 ~ 102	99
103 ~ 109	106
110 ~ 116	113
117 ~ 123	120
124 ~ 255	127

Q1

Input Range	Output Level
0 ~ 3	0
4 ~ 6	3
7 ~ 8	6
9 ~ 10	9
11 ~ 13	12
14 ~ 16	15
17 ~ 19	18
20 ~ 22	21
23 ~ 26	24
27 ~ 30	28
31 ~ 34	32
35 ~ 39	37
40 ~ 44	42
45 ~ 49	47
50 ~ 54	52
55 ~ 59	57
60 ~ 64	62
65 ~ 69	67
70 ~ 74	72
75 ~ 79	77
80 ~ 84	82
85 ~ 89	87
90 ~ 94	92
95 ~ 99	97
100 ~ 104	102
105 ~ 109	107
110 ~ 116	113
117 ~ 123	120
124 ~ 255	127

Q2

Input Range	Output Level
0 ~ 4	0
5 ~ 8	5
9 ~ 12	10
13 ~ 17	15
18 ~ 22	20
23 ~ 27	25
28 ~ 32	30
33 ~ 37	35
38 ~ 42	40
43 ~ 47	45
48 ~ 52	50
53 ~ 57	55
58 ~ 62	60
63 ~ 67	65
68 ~ 72	70
73 ~ 77	75
78 ~ 82	80
83 ~ 87	85
88 ~ 92	90
93 ~ 97	95
98 ~ 102	100
103 ~ 107	105
108 ~ 112	110
113 ~ 118	115
119 ~ 124	121
125 ~ 255	127

Q3

Input Range	Outp Lev
0 ~ 6	
7 ~ 11	
12 ~ 17	1
18 ~ 24	2
25 ~ 31	2
32 ~ 38	3
39 ~ 45	4
46 ~ 52	4
53 ~ 59	5
60 ~ 66	6
67 ~ 73	7
74 ~ 80	7
81 ~ 87	8
88 ~ 94	9
95 ~ 101	9
102 ~ 108	10
109 ~ 115	11
116 ~ 123	11
124 ~ 255	12

Note :Characteristics are symmeterical with respect to zero

TABLE 3. Variable length code for non-zero amplitude prediction errors

Level Number	Code Length	V Code	
V 0	4	0 1 1 1	
1	2	1 S	
2	5	0 1 1 0	S
3	7	0 1 0 1	1 1 S
4	7	0 1 0 1	1 0 S
5	8	0 1 0 1	0 1 1 S
6	8	0 1 0 1	0 1 0 S
7	8	0 1 0 1	0 0 1 S
8	8	0 1 0 1	0 0 0 S
9	9	0 1 0 0	1 1 1 1 S
10	9	0 1 0 0	1 1 1 0 S
11	9	0 1 0 0	1 1 0 1 S
12	9	0 1 0 0	1 1 0 0 S
13	9	0 1 0 0	1 0 1 1 S
14	9	0 1 0 0	1 0 1 0 S
15	9	0 1 0 0	1 0 0 1 S
16	9	0 1 0 0	1 0 0 0 S
17	10	0 1 0 0	0 1 1 1 1 S
18	10	0 1 0 0	0 1 1 1 0 S
19	10	0 1 0 0	0 1 1 0 1 S
20	10	0 1 0 0	0 1 1 0 0 S
21	10	0 1 0 0	0 1 0 1 1 S
22	10	0 1 0 0	0 1 1 1 0 S
23	10	0 1 0 0	0 1 0 0 1 S
24	10	0 1 0 0	0 1 0 0 0 S
25	10	0 1 0 0	0 0 1 1 1 S
26	10	0 1 0 0	0 0 1 1 0 S
27	10	0 1 0 0	0 0 1 0 1 S
28	10	0 1 0 0	0 0 1 0 0 S

Level Number	Code Length	F Code
F 0	4	0 0 0 1
1	6	1 1 1 1 1 S
2	6	1 1 1 1 0 S
3	6	1 1 1 0 1 S
4	6	1 1 1 0 0 S
5	6	1 1 0 1 1 S
6	6	1 1 0 1 0 S
7	6	1 1 0 0 1 S
8	6	1 1 0 0 0 S
9	6	1 0 1 1 1 S
10	6	1 0 1 1 0 S
11	6	1 0 1 0 1 S
12	6	1 0 1 0 0 S
13	6	1 0 0 1 1 S
14	6	1 0 0 1 0 S
15	6	1 0 0 0 1 S
16	6	1 0 0 0 0 S
17	6	0 1 1 1 1 S
18	6	0 1 1 1 0 S
19	6	0 1 1 0 1 S
20	6	0 1 1 0 0 S
21	6	0 1 0 1 1 S
22	6	0 1 0 1 0 S
23	6	0 1 0 0 1 S
24	6	0 1 0 0 0 S
25	6	0 0 1 1 1 S
26	6	0 0 1 1 0 S
27	6	0 0 1 0 1 S
28	6	0 0 1 0 0 S

Note:S denotes Sign.S=0 for positive, S=1 for negative

TABLE 4. Run length code for zero amplitude prediction errors

RL	Code Length	Code Word	Remark
2	5	00 001	
3	5	00 000	
4	6	00 1010	
5	6	00 1001	
6	6	00 1000	
7	7	00 10111	
8～11	7	00 110XX	X=11-RL
12	8	00 111101	
13	8	00 111100	
14～17	8	00 1110XX	X=17-RL
18～25	9	00 0111XX X	X=25-RL
26～33	10	00 01100X XX	X=33-RL
34～37	10	00 010100 XX	X=37-RL
38～64	12	00 01001X XXXX	X=69-RL
MK1	13	00 10110Y YYYYY	Y= 0～63
MK2	14	00 111111 YYYYYY	
MK3	14	00 111110 YYYYYY	
MK 4～7	15	00 01101X XYYYYYY	X= 7-MK
MK 8～15	16	00 01011X XXYYYYYY	X=15-MK
MK16～19	16	00 010101 XXYYYYYY	X=19-MK
MK20～34	18	00 010001 XXXXYYYY YY	X=35-MK
MK35～49	19	00 010000 1XXXXYYY YYY	X=50-MK
MK50～57	19	00 010000 01XXXYYY YYY	X=57-MK

Note : RL=64×(MK number)+1+Y, $0 \leqq Y \leqq 63$

TABLE5. Variable length code and run length code for motion vector data

ΔV_x	ΔV_y	Code Length	Code Word	Number of Codes
±1	0	4	001 Sx	2
±1	±1	5	11 1 SxSy	8
0	±1	5	11 01 Sy	
±2	0	5	11 00 Sx	
0	±2	6	1011 1 Sy	4
±3	0	6	1011 0 Sx	
±1	±2	7	100 11 SxSy	12
±2	±1	7	100 10 SxSy	
0	±3	7	100 011 Sy	
±4	0	7	100 010 Sx	
±3	±1	9	1010 111 SxSy	30
±1	±3	9	1010 110 SxSy	
±2	±2	9	1010 101 SxSy	
±3	±2	9	1010 100 SxSy	
±4	±1	9	1010 011 SxSy	
±4	±2	9	1010 010 SxSy	
±5	0	9	1010 0011 Sx	
±6	0	9	1010 0010 Sx	
0	±4	9	1010 0001 Sx	
-8～7	-5～+5 (See Fig.13)	11	100001 XXXX Sy [X]=ΔV_x	32
-16～15	-6～+6 (See Fig.13)	13	0100001 XXXXX Sy [X]=ΔV_x	64
-16～15	-8～+7 (See Fig.13)	15	100000 XXXXXYYYY [X]=ΔV_x [Y]=ΔV_y	360

RL	Code Length	Code Word	Number of Codes
1	3	000	1
2	4	0111	1
3～6	6	0110XX XX= 6-RL	4
7～12	7	0101XXX XXX=12-RL	6
13～20	8	01001XXX XXX=20-RL	8
21～28	9	010001XX XXX=28-RL	8
TRANS	6	010111	1

Note: Sx and Sy denote signs. Si=0 for positive, Si=1 for negative.

ANNEX

Frame Structure for Use with the Proposed Codec

1. Outline

This annex summarizes the frame structure for use with the proposed codec operating at 1544 kbit/s.

A frame consists of 193 bits. A multiframe and a supermultiframe consist of 12 frames and 4 multiframes, respectively. To satisfy the format restrictions specified in Recommendation G.703 (one's density), provision is made by means of a stuffing system. Three frames are observed, and when needed, stuffing is carried out to control one's density.

2. Frame Structure

Figure A1 shows the line data frame format. The upper figure depicts the normal mode, where no stuffing is carried out. On the other hand, the lower figure shows the stuffing mode, where 1's are stuffed at the bit positions indicated by dashed lines.

Figure A2 shows the multiframe structure.

The abbreviated names for the bits in Figures A1 and A2 are shown in Table A1.

Table A1 Bit Allocations

Abbr.	Information
F_T	Transmission Line Frame Sync.
F_D	Codec Frame Sync.
X_0-X_4	Stuffing Flags Majority Decision (3 out of 5) 1 : Normal Mode 2 : Stuffing Mode
C	Codec-to-Codec Information
A_1-A_8	Audio Data
D	Audio Frame Information
V	Video Data

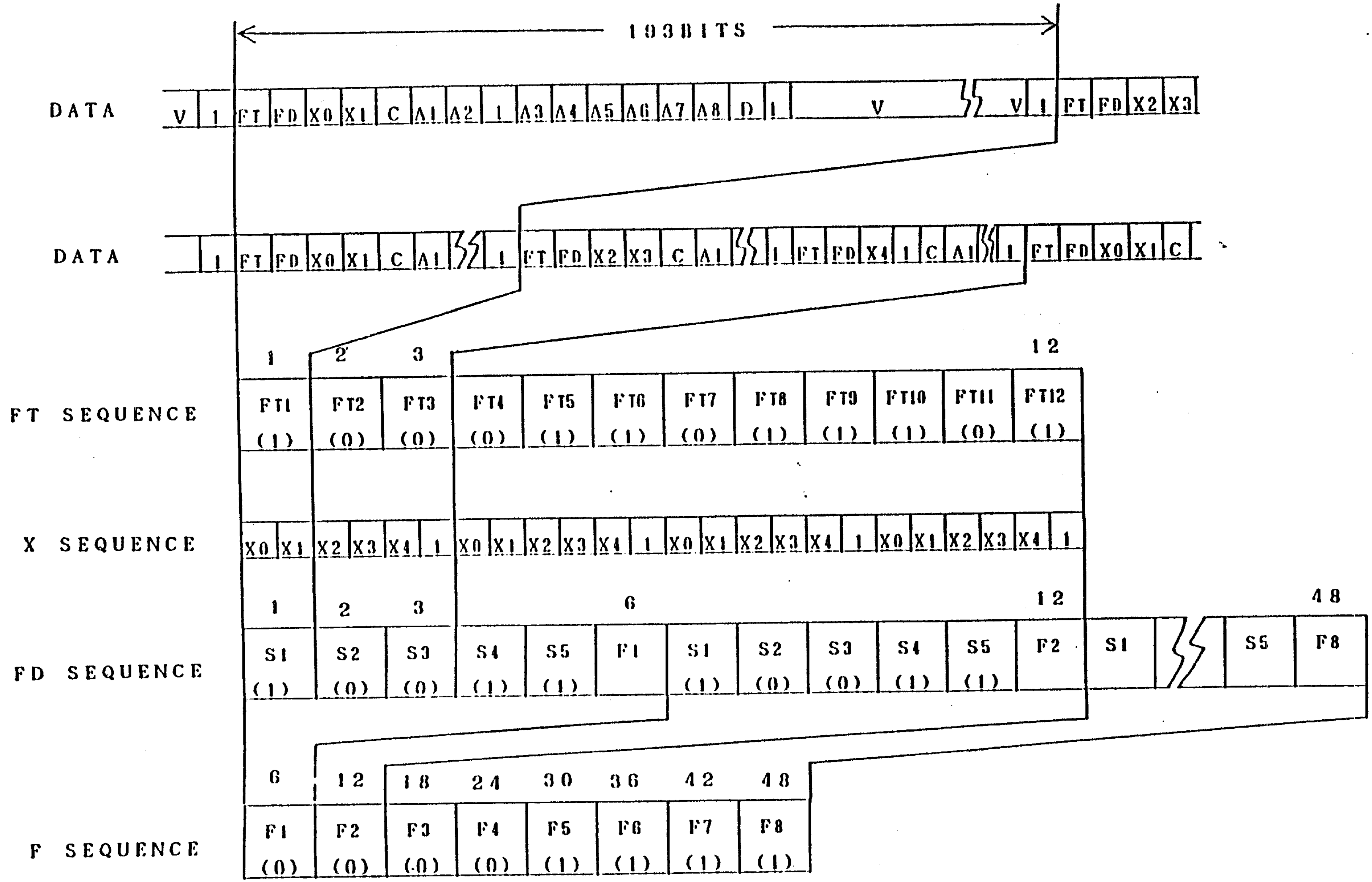
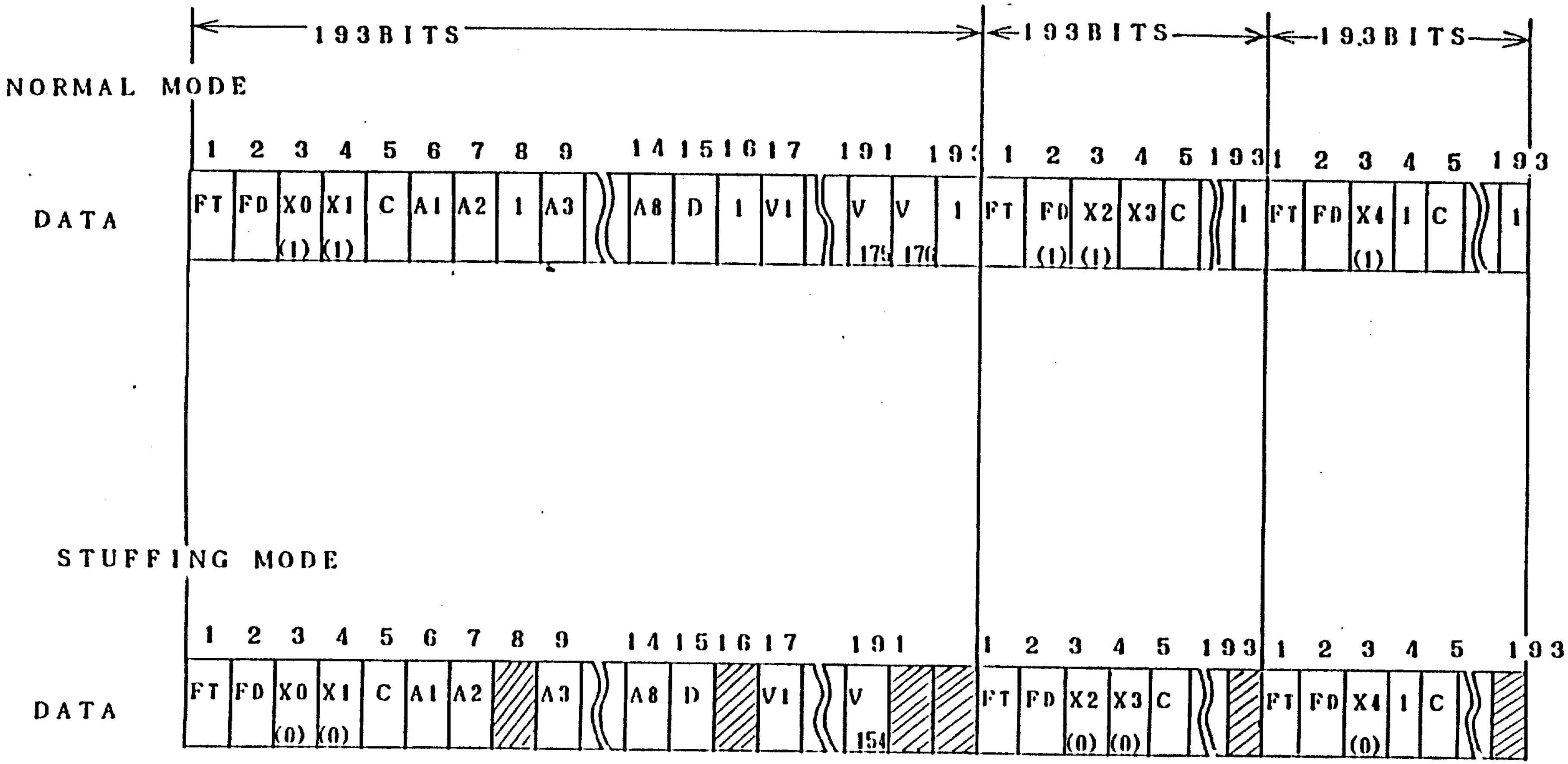


Fig. A2 Multiframe Structure



STUFFING BIT POSITIONS (Indicated as ▨)

8 16 24 32 40 48 56 64 72 80 88 96

104 112 123 128 136 144 152 160 168 176 184 192 193

Fig. A1 Frame Format