CCITT SGXV
Working Party on Visual Telephony
Specialists Group on Coding for Visual Telephony

Document # 19
Version -
April 5, 1985

SOURCE : JAPAN
TITLE : REVISION OF A DRAFT FOR PARTS 3 OF RECOMMENDATIONS H.120 AND H.130

At the Tokyo meeting of the Working Party on Visual Telephony in December 1984, Japan submitted a contribution titled "A Draft for Parts 3 of Recommendations H.120 and H.130" (COM XV-No.D.1). As a conclusion after discussion, candidates of Parts 3 were requested to "be further studied with the aim of aligning frame structure more closely to that specified in Part 2a of the Recommendation H.130" (Document #14R).

The Annexes to this contribution provide a revised text of draft Recommendation. They include:

- Revision of H.130 to meet the request mentioned above, and

- Some editorial revision improving the clarity of the text.

The new and revised sections of the text are marked with vertical lines in the left-hand margin. All the remainder is as has already been proposed at the previous meeting.

ANNEX 1 : A Draft for PART 3 of Recommendation H.120

A CODEC FOR 525-LINES, 60 FIELD/S AND
1544 KBIT/S TRANSMISSION FOR INTRA-REGIONAL USE

1. Introduction
2. Outline of Codec
3. Brief Description
   3.1 Video Input/Output
   3.2 Digital Output/Input
   3.3 Sampling Frequency
   3.4 TDM (Time Division Multiplexing) Digital Video Format
   3.5 Coding Algorithm
   3.6 Audio Channel
   3.7 Data Channel
   3.8 Mode of Operation
   3.9 Transmission Error Protection
   3.10 Additional Facilities
   3.11 Processing Delay
4. Video Interface
5. Pre- and Post-Processing
   5.1 Analog-to-digital and Digital-to-analog Conversion
   5.2 Color Decoding and Encoding
   5.3 TDM Signal
   5.4 Pre- and Post-filtering
      5.4.1 Pre-filtering
      5.4.2 Post-filtering
6. Source Coding
   6.1 Configuration of Source Coder and Decoder
   6.2 Predictive Coding
      6.2.1 Coding Modes
      6.2.2 Adaptive Prediction
      6.2.3 Background Generation
      6.2.4 Forced Intraframe Prediction
      6.2.5 Definition of Blanking and Block-line and Edge Pels Treatment
      6.2.6 Prediction and Interpolation Functions
      6.2.7 Quantization
      6.2.8 Limiter in Prediction Loop
      6.2.9 Frame Memory Parity Check
      6.2.10 Stop of Coding
   6.3 Motion Vector Transmission
      6.3.1 Block Size
      6.3.2 Maximum Tracking Range
      6.3.3 Definition of Vector Direction
      6.3.4 Motion Detection Method
   6.4 Coding Parameter Control
      6.4.1 Control Method
      6.4.2 Control Timing
      6.4.3 Control Sequence
   6.5 Entropy Coding
      6.5.1 Configuration of Entropy Coding
      6.5.2 Commands for Coding Modes
      6.5.3 Prediction Error Coding (VLC1)
      6.5.4 Motion Vector Coding (VLC2)
      6.5.5 Block-line Skipping

6.6 Buffer Memory
6.6.1 Receiving Buffer Control
6.6.2 Memory Size
6.6.3 Underflow Prevention
6.6.4 Overflow Prevention
7. Audio Coding
8. Transmission Coding
8.1 General
8.2 Encryption
8.3 Error Correction
8.4 Scrambling
8.5 Frame Structure and Stuffing

1. Introduction

This Recommendation describes a 1.5 Mbit/s interframe codec which is capable of transmitting and receiving a single NTSC video signal and audio signal using an adaptive predictive coding technique with motion-compensated prediction, background prediction and intraframe prediction.

The aim of this codec is to effectively transmit video telephone and video conferencing signals which have relatively small movements. The video interface of the codec is a 525-lines, 60 field/s standard analog television signal corresponding to the "class a" standard of CCITT Recommendation H.100.

2. Outline of Codec

The essential parts of the codec blockdiagram are shown in Fig. 1. The coder consists of three basic functional blocks, that is, pre-processing, video source coding and transmission coding.

In the pre-processor, the input analog NTSC video signal is digitized and color decoded into one luminance component and two chrominance components. These three components are time division multiplexed into a digital video form, whose noise and unnecessary signal components are removed by the pre-filter.

In the video source coder, the digital video signal is fed to the predictive coder where interframe and intraframe predictive coding techniques are fully utilized for minimizing prediction errors to be transmitted. The prediction error signal is next entropy-coded using its statistical properties to reduce redundancies. Since the coded error information is generated in irregularly spaced bursts, a buffer is used. If the buffer becomes full, the number of prediction error quantizing levels and/or picture elements to be coded is reduced to prevent any overflow.

In the transmission coder, coded video and audio signals are first encrypted on an optional basis. The coded video signal is then forward error correction coded and scrambled. The three signals, coded video, coded audio and optional data signals are multiplexed into a 1544 kbit/s digital format with a frame structure as defined in H.130.

The decoder carries out a reverse operation.

3. Brief Descriptions

3.1 Video Input/Output

NTSC signals are used for the video input/output signal, with monochrome signals being additionally applicable.

3.2 Digital Output/Input

The interface conditions for the digital output/input signal satisfy CCITT Recommendation G.703 specifications. The signal transmission rate is 1544 kbit/s.

3.3 Sampling Frequency

The video sampling frequency is four times the color subcarrier frequency (f_{SC}) and asynchronous with the 1544 kHz network clock.

3.4 TDM (Time Division Multiplexed) Digital Video Format

An NTSC signal is separated into a luminance component (Y) and two chrominance components(C_1 and C_2). A time-division-multiplexed signal composed of Y and time-compressed C_1 and C_2 is employed in the source coding as the standard digital video format.

3.5 Coding Algorithm

Adaptive predictive coding supplemented by variable-length coding is used to achieve low bit rate transmission. The following three predictions are carried out adaptively on a pel-by-pel basis:

a. motion-compensated interframe prediction for a still or slowly moving area,
b. background prediction for an uncovered background area, and
c. intraframe prediction for a rapidly moving area.

Prediction errors for video signals and motion vectors are both entropy-coded using the following two techniques:

a. variable word-length coding for non-zero errors, and
b. run-length coding for zero errors.

3.6 Audio Channel

An audio signal band-limited to 7 kHz is coded by an adaptive DPCM method and transmitted at a rate of 64 kbit/s.

3.7 Data Channel

An optional 64 kbit/s data channel is available, which is used for video if not required for data.

3.8 Mode of Operation

The normal mode of operation is full duplex, with other modes, e.g. the one-way broadcasting operation mode, also taken into account.

3.9 Transmission Error Protection

A BCH error correcting code is used along with a demand refreshing method to prevent uncorrected errors from degrading the picture quality.

3.10 Additional Facilities

Provision is made in the digital frame structure for the future introduction of such facilities as encryption, graphics transmission and multipoint communication.

3.11 Processing Delay

The coder plus decoder delay is about 165 ms without that of a pre-filter and a post-filter.

4. Video Interface

The video input/output signal of the codec is an analog NTSC signal (System M) in accordance with CCIR Rep.624-2.

5. Pre- and Post-processing

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

An NTSC signal band-limited to 4.5 MHz is sampled at a rate of 14.3 MHz, four times the color subcarrier frequency (f_{SC}), and converted to an 8-bit linear PCM signal. The sampling clock is locked to the horizontal synchronization of the NTSC signal. Since the sampling frequency is asynchronous with the network clock, the justification information is coded and transmitted from the coder to the decoder.

The digital video data is expressed as two's complement form. The input level to the A/D converter is defined as follows;

- Sync tip level (-40 IRE) corresponds to -124 (10000100),
- White level (100 IRE) corresponds to 72 (01001000).

As a national option, a pad can be inserted before the A/D converter if a level fluctuation should be considered at analog transmission lines connecting terminal equipment and codec.

One line deleted

At the decoder, the NTSC signal is reproduced by converting the 8-bit PCM signal to an analog signal.

5.2 Color Decoding and Encoding

The digitized NTSC signal is separated into the luminance component (Y) and the carrier band chrominance component (C) by digital filtering. The two baseband chrominance signals (C_1 and C_2) are obtained by digitally demodulating the separated carrier band chrominance component. The effective sampling frequency after color decoding is converted to 7.2 MHz (2 f_{SC}) and 1.2 MHz (1/3 f_{SC}) for the luminance signal and chrominance signals respectively.

The replica of the NTSC signal is obtained by digitally modulating the C_1 and C_2 signals and adding to the Y signal at the decoder.

Filter characteristics for color decoding and encoding are left to each hardware implementation since they do not affect interworking between different design codecs. Examples of recommended characteristics are described in Annex 1.

5.3 TDM Signal

A time division multiplexing signal is constructed from the separated component signals.

First, the C_1 and C_2 signals are time-compressed to 1/6. Next, each of the time compressed C_1 and C_2 signals, with their horizontal blanking parts removed, is inserted into the Y signal horizontal blanking interval on alternate lines. C_1 is inserted on the first line of the first field and on every other line following throughout the frame, while C_2 is inserted on the second line of the first field and on every other line following throughout the frame.

Active samples for the Y signal are 384 samples/line and 64 samples/line for the C_1 and C_2 signals. The TDM signal is constructed with these active samples and 7 color burst samples(B), which are inserted into the top of the TDM signal.

As shown in Fig. 2, the C_1- and C_2- signal sampling points coincide with that of the Y signal on every sixth sample. The C_1 and C_2 signals of only the odd lines are transmitted to the decoder.

At the decoder, each component signal is again demultiplexed from the TDM signal, and time-expansion processing of 6 times is carried out for the C_1 and C_2 signals.

5.4 Pre- and Post-filtering

In addition to conventional anti-aliasing filtering prior to analog-to-digital conversion, the following two filtering processes should be used as pre-filtering for source coding.

(1) Temporal filtering to reduce random noise included in the input video signal

(2) Spatial filtering to reduce aliasing distortion in subsampling

At the decoder, the following three filtering processes should be used as post-filtering in addition to conventional low pass filtering after digital-to-analog conversion.

(3) Spatial filtering to interpolate the omitted picture elements in subsampling

(4) Spatio-temporal filtering to interpolate the omitted fields in field repetition

(5) Temporal filtering to reduce noise generated in the course of source coding

Though these filtering processes are important for improving reproduced picture quality, their characteristics are independent of interworking between different design codecs. Hence, pre- and post-filtering is left to each hardware implementation.

6. Source Coding

6.1 Configuration of Source Coder and Decoder

The video source coder and decoder configuration of this codec is outlined in Fig. 3.

The Predictive Encoder converts the input video signal x into the prediction error signal e using the motion vector v. This conversion is controlled by the coding mode m.

The VWL Coder codes e and v into the compressed data C using the variable length coding method. The transmission buffer memory BM smoothes out the irregularly spaced data C. The coding mode m is also coded.

The frame memory parity information p is used to check the identity of coder and decoder frame memory contents. If any parity error is detected, frame memories of both coder and decoder are reset by the demand refresh request information DR and the demand refresh confirmation information DDR.

At the decoder, the VWL Decoder decodes e, v, m and p and the Predictive Decoder reproduces the video signal x'.

6.2 Predictive Coding

6.2.1 Coding Modes

Five coding modes as summarized in Table 1 are provided. All of the samples are coded and transmitted in Normal Mode, while half of the samples are omitted in Subsampling Mode. In Field Repetition Mode, one or more consecutive fields are omitted (called multi-field repetition, Note 1). If Field Repetition Mode and Subsampling Mode are used in combination, only a quarter or less of the original picture elements are coded and transmitted.

Subsampling is carried out in a quincunx way, namely by transmitting only odd-numbered pels on odd-numbered lines and even-numbered pels on even-numbered lines in each block-line (Note 2).

In Field Repetition Mode, either the odd or even fields are omitted. For the omitted fields, both the prediction error e and the motion vector v are set to 0.

Note 1 : If odd fields and even fields are omitted in mixture, a severe picture degradation takes place. Hence, 1 out of 2, 3 out of 4 or 5 out of 6 field omission is recommended.

Note 2 : Each block-line consists of 8 lines as defined in 6.2.5.

6.2.2 Adaptive Prediction

Prediction functions are adaptively selected on a pel-by-pel basis as shown in Fig. 4. The selection is carried out so as to minimize probable prediction errors. This is accomplished using the two prediction status signals, which are determined by prediction reference signals, for the preceding pels located on the previous and the present lines.

When subsampling and/or field repetition are operated, omitted pels are interpolated in the prediction loop.

The notations defined for the i-numbered pel are

- X_i : local decoder output,
- Y_i : interpolator output,
- M_i : motion compensated interframe prediction value,
- B_i : background prediction value,
- I_i : intraframe prediction value,
- $*$: logical product, and
- $+$: logical sum.

(1) Motion-compensated Interframe Prediction / Background Prediction

Prediction status signal S_{1i} for pel i is determined as

$$S_{1i} = R_1(i-455)*R_1(i-456) + R_1(i-456)*R_1(i-454) + R_1(i-454)*R_1(i-455), \qquad (1)$$

where prediction reference signal $R_1(i)$ is

$$R_1(i) = \begin{cases} 0 & \text{if } |Y_i - B_i| \geq |Y_i - M_i|, \\ 1 & \text{if otherwise.} \end{cases} \qquad (2)$$

Based on S_{1i}, prediction signal X_{1i} is given as

$$X_{1i} = \begin{cases} M_i & \text{if } S_{1i} = 0, \\ B_i & \text{if } S_{1i} = 1. \end{cases} \quad (3)$$

If pel i is either omitted due to subsampling and/or field repetition or forced intraframe coded or in burst B, its corresponding $R_1(i)$ is set to 0 regardless of Equation (2).

(2) Interframe Prediction / Intraframe Prediction

Prediction status signal S_{2i} for pel i is determined as

$$S_{2i} = R_2(i-1)*R_2(i-455), \quad (4)$$

where prediction reference signal $R_2(i)$ is

$$R_2(i) = \begin{cases} 0 & \text{if } |Y_i-I_i| \geq |Y_i-X_{1i}|, \\ 1 & \text{if otherwise.} \end{cases} \quad (5)$$

Based on S_{2i}, prediction signal X_{2i} is given as

$$X_{2i} = \begin{cases} X_{1i} & \text{if } S_{2i} = 0, \\ I_i & \text{if } S_{2i} = 1. \end{cases} \quad (6)$$

If pel (i-1) is omitted due to subsampling, $R_2(i-2)$ is used instead of $R_2(i-1)$. On the other hand, if pel (i-455) is omitted, $R_2(i-454)*R_2(i-456)$ is used instead of $R_2(i-455)$. If pel i is forced intraframe-coded, its corresponding $R_2(i)$ is set to 1 regardless of Equation (5).

If pel i is omitted due to field repetition, its corresponding $R_2(i)$ is set to 0 regardless of Equation (5). When pel i is not forced-intraframe coded, $R_2(i)$ in burst B is set to 0.

6.2.3 <u>Background Generation</u>

The background prediction value is generated scene adaptively as

$$b_i = b_i^{-f} + v(k)\text{sign}(Y_i-b_i^{-f})u(Y_i-Y_i^{-f}) \quad (7)$$

where

$$u(Y_i-Y_i^{-f}) = \begin{cases} 1 & \text{if } |Y_i-Y_i^{-f}| \leq L, \\ 0 & \text{if otherwise,} \end{cases} \quad (8)$$

$$v(k) = \begin{cases} 1 & \text{for one frame period in every block of k frames} \\ 0 & \text{for consecutive (k-1) frames following the frame of } v(k) = 1 \end{cases}$$

and b_i : background prediction value for the present frame,

b_i^{-f} : background prediction value for the previous frame,

Y_i : interpolator output for the present frame,

Y_i^{-f} : interpolator output for the previous frame, u : still area detection function,

k : background update control parameter, and

L : threshold value.

Parameters k and L are set as $k = 8$ and $L = 1$. It is noted that for hardware simplification, b_i^{-f}, instead of b_i, is used as background prediction value B_i.

6.2.4 Forced Intraframe Prediction

This codec usually uses the demand refresh mode to prevent the defected picture due to transmission errors from being left in the decoder frame memory. The demand refresh mode is carried out if BWP (Bit 3.15.4 in codec-to-codec information) = 0, which indicates that backward path from decoder to coder is available. However, the cyclic refresh mode is also provided, considering such applications as broadcasting communication where no backward path (from decoder to coder) is available. This mode is carried out when BWP = 1.

For either of the two refresh modes, the prediction function is forcedly set to intraframe prediction.

In the demand refresh mode, the motion frame memory and the background frame memory are updated block-line by block-line within a frame time by writing the interpolator output simultaneously. Once demand refresh starts by receiving DRR in the coder, the following DRR received is disregarded for one second. (Note)

In the cyclic refresh mode, the two memories are updated simultaneously two lines at a time, by writing the interpolator output. It should be noted that the command for cyclic refresh is neglected in the block-line where updating based on the demand refresh is carried out.

Note : If a transmission error happens on the line from the codec A to the codec B, the decoder of the codec B detects the error occurrence and generates a demand refresh request information DR. This DR is passed to the coder of the codec B and transmitted as a demand refresh request command DRR to the codec A. When the decoder of the codec A receives DRR, a demand refresh confirmation information DDR is passed to the coder of the codec A. Finally, demand refresh mode is operated along with transmission of a demand refresh mode command DRM from the codec A to the codec B.

6.2.5 Definition of Blanking and Block-line and Edge Pels Treatment

(1) The kinds of pels arranged on a horizontal scanning line (see Fig. 2), for which prediction functions are defined, are

Burst B : 7 pels,
Color C : 64 pels and
Luminance Y : 384 pels.

(2) The vertical blanking periods are treated the same as active lines.

(3) Block-line (see Fig. 6)

In the first field, 8 lines consisting of the 8th through the 15th lines form the first block-line, with every 8 lines following this forming each block-line. In the second field, 8 lines consisting of the 7th through the 14th lines form the 33rd block-line. Each field has 32 block-lines.

The last block-line in a frame is defined as the 8 lines which include the last line of the frame, or the line closest to the head line of the frame. The position of the video last line in the last block-line is coded as frame position.

(4) Reset Lines

The lines which are excluded from block-lines are defined as reset lines. These reset lines are clamped to 0 in the predictive coding and decoding loops, or the corresponding prediction values X_2 and prediction errors e in Fig. 4 are set to 0. The reset lines are prediction-coded in Normal Mode with adaptive prediction and setting $v = 0$.

(5) Edge pels suffer from crosstalk due to interpolation between B and C, C and Y, and Y and B. In order to prevent such crosstalk, the first 3 pels in B, the last pel in B and the first pel in Y are clamped to 0 at the source coder input TDM signal as in Fig. 7.

Three lines deleted

(6) Edge pels are not treated specifically in the source coder and decoder. Namely, video signals including reset lines and the 3 clamped pels in B are processed as if they were existing continuously (Note). Consequently, even if a motion vector points pels outside the active picture area, it functions as a delay control for the input time serial video signal.

Note : The right end of each line in the picture is assumed to be connected to the left end of the next line, and the lower end of each frame is assumed to be connected to the upper end of the next frame.

(7) In the forced intraframe prediction mode, the prediction value for the first pel of each line is set to 0.

(8) For the burst signal, no adaptive prediction nor subsampling is applied and no motion vectors are transmitted.

6.2.6 Prediction and Interpolation Functions

Prediction functions and interpolation functions are shown in Table 2 for all of the coding modes. It should be noted that motion vectors for the color signal can be set to 0 without much loss of coding efficiency.

6.2.7 Quantization

Prediction errors for video signals are quantized using one of the four quantizing characteristics indicated in Table 3, that is, Q_0(57 levels), Q_1(57 levels), Q_2(51 levels) and Q_3(37 levels). The same set of quantizing characteristics are applied regardless of prediction functions.

6.2.8 Limiter in Prediction Loop

No limiter is allocated in the prediction loop. Accordingly, input signal x for the prediction loop is limited to $-124 \leq x \leq 123$ so that local decoder output X is maintained in the $-128 \leq X \leq 127$ range.

6.2.9 Frame Memory Parity Check

Parity is counted for each bit plane of the interpolator output during a video frame period from the 1st to the 64th block-line as defined in Fig. 6. If block-lines are omitted in Field Repetition Mode, parity is not counted during these omitted block-lines.

Eight odd parity bits are sent to the decoder, where they are compared with the parity bits of the decoder interpolator output to detect uncorrected errors. If any difference between received and counted parity bits is found, a demand refresh is requested from the decoder to the coder.

6.2.10 Stop of Coding

When the information is generated to the degree that the transmission buffer memory overflows, coding operation is suspended by setting e = 0 and v = 0. This Stop Mode is defined only in the coder. Interpolation and prediction functions for this mode are defined as either of NRM, SBS, FRP or RFS according to the control of the coding parameter controller.

6.3 Motion Vector Transmission

6.3.1 Block Size

A block for motion compensation consists of 8 lines (vertical) x 16 pels (horizontal).

6.3.2 Maximum Tracking Range

Motion vectors are tracked in the range of +7 to -7 lines (vertical) and +15 pels to -15 pels (horizontal) at its maximum. The decoder should be able to reproduce any vector in this maximum range.

6.3.3 Definition of Vector Direction

The motion vector $v(v_x, v_y)$ is defined as

$$v_x = x_a - x_b \quad \text{and} \quad v_y = y_a - y_b, \tag{9}$$

where the block positions in the present frame and in the corresponding previous frame are (x_a, y_a) and (x_b, y_b) respectively. The x and y directions are identical to those of the horizontal and vertical scanning. This definition means that delay in the interframe prediction loop increases for v_x, $v_y > 0$.

6.3.4 Motion Detection Method

A motion vector is detected for each block by the interframe block matching method. Detailed detection methods are left to each hardware implementation (Note).

Note : When multi-field repetition is employed, the detected vector for the previous transmitted frame can be utilized as the initial value for vector detection in the present frame to be omitted, and the detected vector for the present frame can be utilized as the initial value for vector detection in the next frame, and so on.

6.4 Coding Parameter Control

6.4.1 Control Method

Coding control is carried out by selecting quantizing characteristics described in 6.2.7 and coding modes described in 6.2.1.

6.4.2 Control Timing

Coding parameters are controlled according to the timing and commands as shown in Table 4.

6.4.3 Control Sequence

The control sequence is determined based on buffer memory occupancy and other control information. Since this sequence does not affect interworking between different design codecs, it is left to each hardware implementation. However, the codec operating principle is that the coder determines all operating modes, which are transmitted with the coded video data to the decoder as a combination of commands. The decoder dependently reproduces the video signal according to the received commands and data. A control sequence example is shown in Annex 3.

6.5 Entropy Coding

6.5.1 Configuration of Entropy Coding

Configuration of Entropy Coding is shown in Fig. 8. The entropy coder compresses the data of the prediction error e and motion vector v, which are provided by the source coder, using variable length coding. These compressed data are multiplexed with coding mode data m and fed into the transmission buffer memory. The multiplexed data format is outlined in Fig. 9.

6.5.2 Commands for Coding Modes and Data Structure

The commands for coding modes and data structure are defined as follows:

(1) FS : Frame Sync, a unique word to designate the start of a video frame

$$\underbrace{00000000000000}_{14}10$$

(2) FMD : Frame Mode Data

Head	PT	FM1	BC	FM2	FC	FM3
	1 Byte	1B	1B	1B	1B	3B

a. PT : Parity Data
Odd parity for each of the 8 bit planes of the interpolator output during the previous frame period (MSB first).

b. FM1 : Frame Mode 1

SBC	BRC	BUC	DRM	1	FP2	FP1	FP0

SBC : Subsample Control
When SBC = 0, subsampling is carried out throughout the frame excluding burst signals, reset lines and block-lines with FRP=0. See 6.2.1.

BRC : Background Revision Control
When BRC = 0, the contents of the motion frame memory are transferred to the background frame memory during this frame period. See 6.2.4.

BUC : Background Update Control
When BUC = 0, the background frame memory is updated. If BRC is operating, it has priority. See 6.2.3.

DRM : Demand Refresh Mode
When DRM = 0, coding is carried out with Demand Refresh Mode. See 6.2.4.

FP2-FP0 : Frame Position (Note)
This 3-bit word designates the position of the head line of the video frame or the first line in the first field (MSB first). See Fig. 6.

<u>Note</u> : FP bits are employed for preventing degradation in the case that the input signals are asynchronously switched to another signals where those signals have different sync phase or different sync frequency. For this purpose the horizontal sync pulse interval in the codec, namely, the picture element numbers per line should be kept as 455 samples even in the transition period.

c. BC : Buffer Control
Staying time of FS in the transmission buffer memory is coded into an 8 bit word (MSB first). See 6.6.1.

d. FM2 : Frame Mode 2

1	DRR	CMS	CRM1	CRM2	SF1	MAF	1

DRR : Demand Refresh Request
When DRR = 0, the decoder requests a demand refresh to the coder. See 6.2.9.

CMS : Color/Monochrome State
Color(1) / Monochrome(0) where Monochrome is optional and default mode is Color.

CRM1,CRM2 : Cyclic Refresh Mode
This 2 bit word designates the position of the two lines in a block-line which is cyclic refreshed. See 6.2.4.

CRM1	CRM2	line
0	0	1, 2
0	1	3, 4
1	0	5, 6
1	1	7, 8

$\underbrace{1 \ldots 8}_{\text{block-line}}$

SF1 : Spare Frame Mode

MAF : Mode Addition Flag
When MAF = 0, FM3 is added.

e. FC : Sampling Frequency Control
The lower 8 bits for a video frame period counted with the transmission line clock (MSB first). See 5.1.

f. FM3 : Frame Mode 3

1st B	SF2	SF3	SF4	1	SF5	SF6	SF7	SF8
2nd B	SF9	SF10	SF11	1	SF12	SF13	SF14	SF15
3rd B	SF16	SF17	SF18	1	SF19	SF20	SF21	SF22

SF2-SF22 : Spare Frame Mode

(3) LS : Line Sync, a unique word to designate the start of a block-line

$$\underbrace{00000000000000}_{14}11$$

(4) LMD : Line Mode Data

Head	QC1	QC2	IFM	1	LSK	FRP	SL1	SL2	LDN
	1 Byte								1 Byte

QC1,QC2 : Quantizing Characteristics

QC1	QC2	Characteristics (Table 3)
0	0	Q_0
0	1	Q_1
1	0	Q_2
1	1	Q_3

IFM : Forced Intraframe Prediction Mode
When IFM = 0, the prediction function is fixed to intraframe prediction, throughout this block-line if DRM=0, and at the 2 lines designated by CRM1 and CRM2 if DRM=1. See 6.2.4.

LSK : Line Skip
When LSK = 0, the following Byte (LDN, Line Data Number) designates the number of block-lines which are skipped. See 6.5.5. LDN is coded similarly as the the number of vector data, VDN. When LDN = n, consecutive (n+1) block-lines are the same. Therefore $0 \leq n \leq 63$.

FRP : Field Repetition
When FRP = 0, this block-line is omitted because of field repetition. See 6.2.1.

SL1,SL2 : Spare Line Mode

(5) MVD : Motion Vector Data

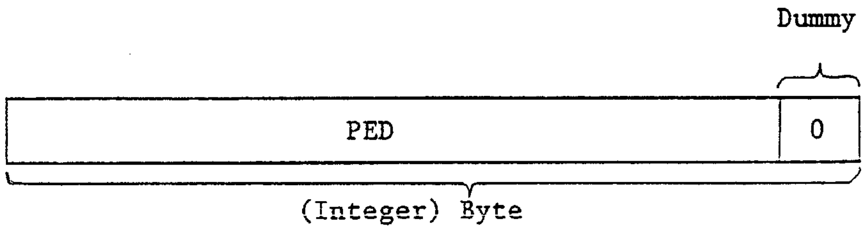
n7	n6	n5	1	n4	n3	n2	n1	VD	*
VDN								n Byte	

VDN : Vector Data Number
Designates the Byte number n of the following VD (MSB first).

VD : Vector Data
Variable length coded motion vector data.

* : Dummy, see 6.5.4(6).

(6) PED : Prediction Error Data (variable word length coded)



6.5.3 <u>Prediction Error Coding (VLC 1)</u>

(1) Coding Method

The quantizing level number corresponding to prediction error e is coded based on its statistical characteristics. For $e \neq 0$, variable word length coding is carried out using code V or F stating the quantizing level number. For $e = 0$, run length code R is used to state a run (RL) of ineffective pels. Note that if RL = 1, a variable word length code V_0 or F_0 to state $e = 0$ is used.

(2) Scanning Sequence

a. Entropy coding for a video frame is carried out from the first through the last block-lines, excluding reset lines.
b. The Frame Sync (FS) and Frame Mode Data (FMD) are coded in the first block-line.
c. When the last line falls on the nth line of the last block-line, the Frame Position is set to FP = mod(n,8). FP is transmitted to the decoder as a part of the Frame Mode Data. (Note 1)
d. Since the first three pels of each line are clamped to 0 in the predictive source coder, and the reset lines are so defined as described in 6.2.5(4), the pels to be entropy coded can be indicated as in Fig. 10. (Note 2)
e. The scanning sequence is a block scan as shown in Fig. 11. The first block after scan conversion consists of 4 pels x 8 lines = 32 pels.

<u>Note</u> 1: Without asynchronous switching of input video signals, the last block-line coincides with the 64th block-line and FP = 0.

<u>Note</u> 2: Entropy coding is not carried out for the reset lines defined in Fig. 6. The number of reset lines in the first field varies according to the FP value.

(3) Group of Codes

	Symbol	Number of Codes	Length of Codes
Amplitude Code No.1	F	57	4,6
Amplitude Code No.2	V	57	2-10
Run Length Code	R	3615	5-19

F: Pseudo fixed length code for stating quantizing level number. This code is introduced to shorten the maximum code length.
V: Variable length code for stating quantizing number.
R: Variable length code for stating run length of ineffective pels for RL $\geq$ 2.

(4) Code Transition Rule

The rule is shown in Fig. 12 and a prediction error coding example is given in Annex 2.

The following points should be noted.

a. The starting code is either the V or R code.
b. The last run in a block-line may not be transmitted since the LS or FS command can be utilized as a termination of the last run.
c. Coding is carried out assuming that omitted pels due to subsampling do not exist.
d. Some 0s are filled at the tail of PED as dummies to make the total number of bits for the block-line data a multiple of 8.

(5) Code Assignments for F and V : Table 5

Code assignments are common to the four quantizing characteristics Q_0, Q_1, Q_2 and Q_3.

(6) Code Assignments for R : Table 6

6.5.4 <u>Motion Vector Coding (VLC 2)</u>

(1) Coding Method

A motion vector v is first coded with predictive coding whose output Δv is variable length coded throughout a block-line.

(2) Predictive Coding

The prediction algorithm is the previous block prediction which is

$$\Delta v = v - v_1, \qquad (10)$$

where v and v_1 represent the present and the previous block vectors. The operation is carried out for each x and y component in two's complement form. The operated results are expressed with 5 bits for the x component and 4 bits for the y component neglecting carries (MSB first). Note that the decoder carries out the inverse operation $v = v_1 + \Delta v$ in two's complement form neglecting carries.

The motion vector for the first block (horizontal blanking) is set to (0,0).

(3) Variable Length Coding

a. For Δv = (0,0), a run length of zero is coded. For $\Delta v \neq (0,0)$, variable length coding is applied with their code lengths shown in Fig. 13.
b. The coding of Δv is carried out for the 28 vectors of the 2nd to the 29th blocks.
c. The last run of Δv = (0,0) may not be transmitted since VDN states the total bits of VD.

(4) Code Assignments

The codes are assigned as shown in Table 7, where the maximum code length is 15. The variable length codes consist of 541 codes, or 512 codes for Δv, 28 codes for run length and one TRANS code for transition of subsampling ON/OFF.

(5) Transition Code for Subsampling (TRANS)

The code TRANS indicates transition between ON and OFF for subsampling (SBS). For the first block in a block-line, SBS is set to OFF. Subsampling is then set to ON at the block just after the first TRANS code is inserted, and returned to OFF at the block just after the second TRANS code is inserted. The same sequence follows on. The TRANS code is expressed as a 6 bit word. When SBS = 0 , transition code is disregarded in the decoder.

(6) Dummy Code Insertion

When a Vector Data for a block-line does not have exactly 8 multiple bits, a dummy code consisting of 1 to 7 bits is inserted at the tail of the Vector Data (VD).

The dummy code has 1 as the head, 0s as the body and 1 as the tail.

Number of Dummy Bits	Dummy Code
1	1
2	11
3	101
4	1001
5	10001
6	100001
7	1000001

(7) Code Transition Rule

The rule is indicated in Fig. 14 with a motion vector coding example given in Annex 2.

6.5.5 Block-line Skipping

If block-lines continue, in which all the prediction error e data and motion vector v data are 0 and whose line mode data (QC1, QC2, IFM, FRP, SL1, SL2) are identical, their number is run length-coded as skipped block-lines. A run ends when it encounters FS or a block-line with new line mode data or some $e \neq 0$ or some $v \neq 0$. A run also ends when variable legnth code V_0 appears due to underflow prevention.

6.6 Buffer Memory

6.6.1 Receiving Buffer Control

The staying period of FS in the transmission buffer is counted with a 1/16 input video line frequency clock and transmitted to the decoder as BC command. The staying period is represented in eight bit binary code. Similarly, the staying period in the receiving buffer is counted and the operation of the receiving buffer is so controlled as to make the total delay time caused by the two buffer memories constant.

Note : This control method is applicable even when the read out speed for the transmission buffer varies.

6.6.2 Memory Size

Transmission buffer memory size B_S is defined as 180 kbits, while receiving buffer memory size B_R should be more than 220 kbits considering the variation of the transmission buffer read out speed.

Note : The delay time due to the transmission and receiving buffer memories becomes about 165 ms for B_S = 180 kbits and B_R = 220 kbits.

6.6.3 Underflow Prevention

If the occupancy of the transmission buffer decreases to a threshold, the run length coding for the prediction error is prohibited and variable length code V_0 is used.

6.6.4 Overflow Prevention

If the occupancy of the transmission buffer increases to another threshold, Stop Mode for setting forcedly all the prediction error and motion vector data to 0 is applied.

7. Audio Coding

A 7 kHz bandwidth analog audio signal is coded into a 64 kbit/s digital stream by subband coding with ADPCM proposed by NTT (Note).

Any recommendation made in the future will replace this part.

Since video coding and decoding introduces a significant delay as described in 3.11, the encoded audio signal should be delayed by the corresponding time in the coder and decoder to obtain the proper synchronization between video and audio at the decoder. The delay inserted in the audio coder should be the sum of a half of the buffer memory delay and other video coding process delay, while the delay inserted in the audio decoder should be the sum of a half of the buffer delay and other video decoding process delay.

Note : SGXVIII Delayed Contribution SE (November 21 - December 2, 1983), Suggestions on the Design Requirements for Wideband Speech Coding at 64 kbit/s and Proposal of an Encoding Scheme.

8. Transmission Coding

8.1 General

The transmission coder assembles the video, audio, optional data and codec-to-codec information channels into a 1544 kbit/s digital stream. With all serialized data, the most significant digit leads.

8.2 Encryption

Video and audio signals can be independently encrypted on an optional basis. Their algorithms are under study. Keys and other control information can be transmitted through the message channel provided in the codec-to-codec information channel.

8.3 Error Correction

An encoded (and encrypted) video signal is forward error corrected by a (255, 239) two-error correcting BCH code. One framing bit is added to each 255 bit error correction frame, and 16 such frames are assembled into one large frame as shown in Fig. 15. The frame alignment pattern is 0001101x (x: for future multiframe alignment signal use). The other 8 bits are used for controlling purpose, whose protocol is under study.

To correct a burst error of up to 32 bits, 16 phase interleaving is employed. The bit allocation rule is also indicated in Fig. 15. Note that framing bits are excluded from interleaving.

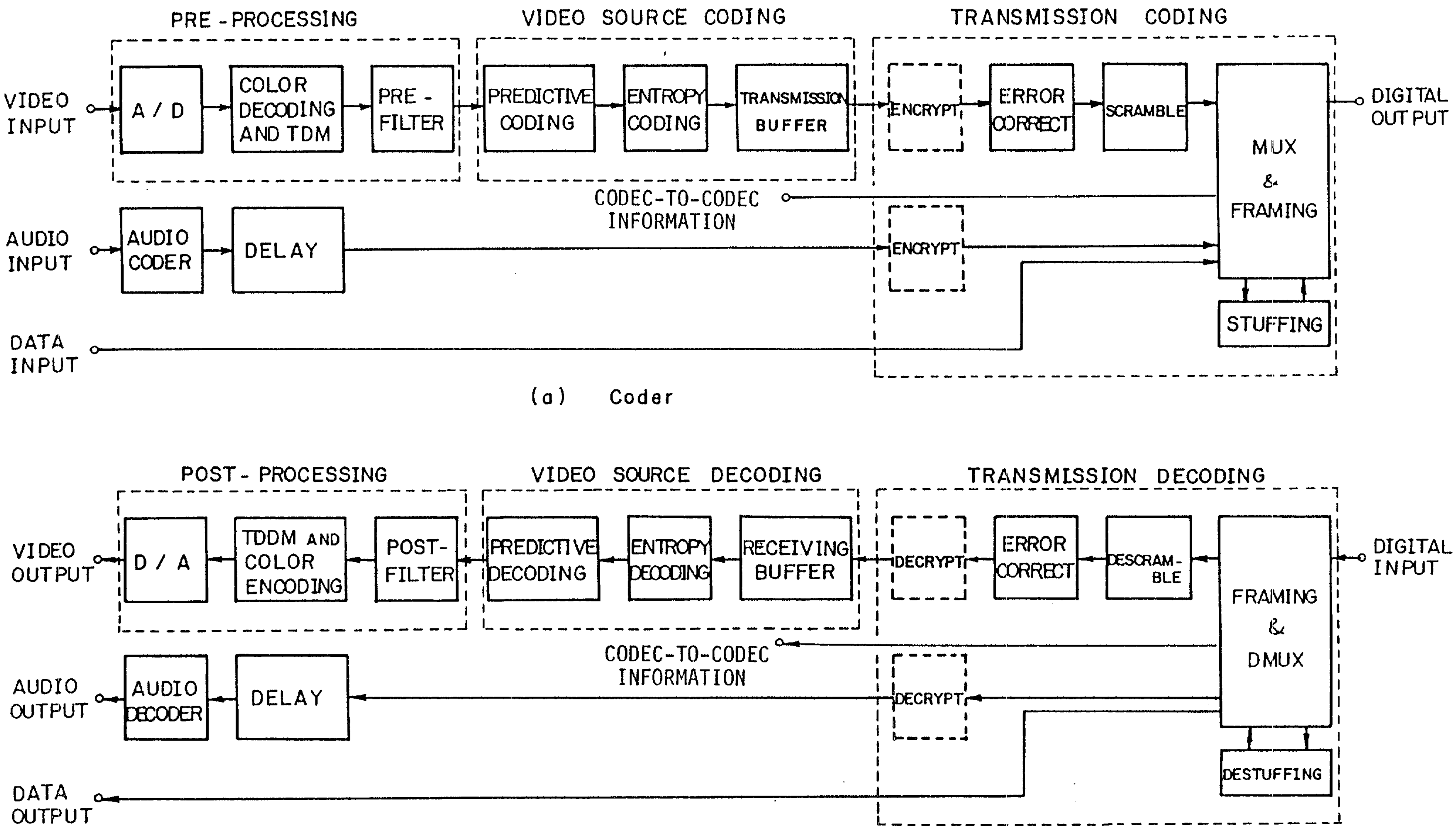
8.4 Scrambling

An error corrected video signal is scrambled with an 8 stage pseudo random pulse generator to reduce stuffing required by the network restrictions. At each error correction frame bit, the scrambler is reset. The polynomial generator and the scrambled output pattern following the reset pulse for the input of all zeros are

$1 + x^4 + x^5 + x^6 + x^8$,
1111010011......1001111011.

8.5 Frame Structure and Stuffing

PART 3 of Recommendation H.130 is applied.



(a) Coder

(b) Decoder

FIGURE 1. Codec blockdiagram

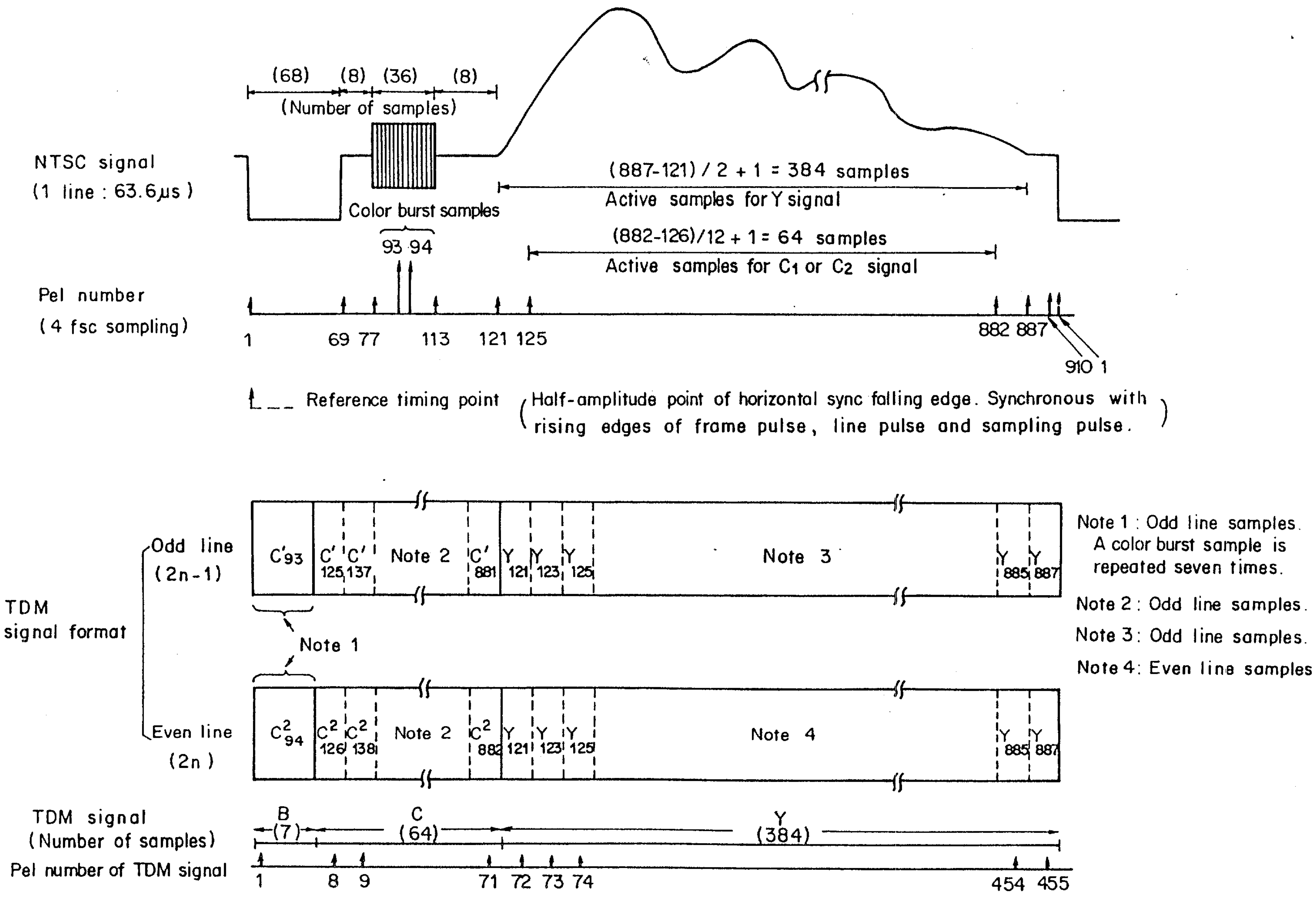
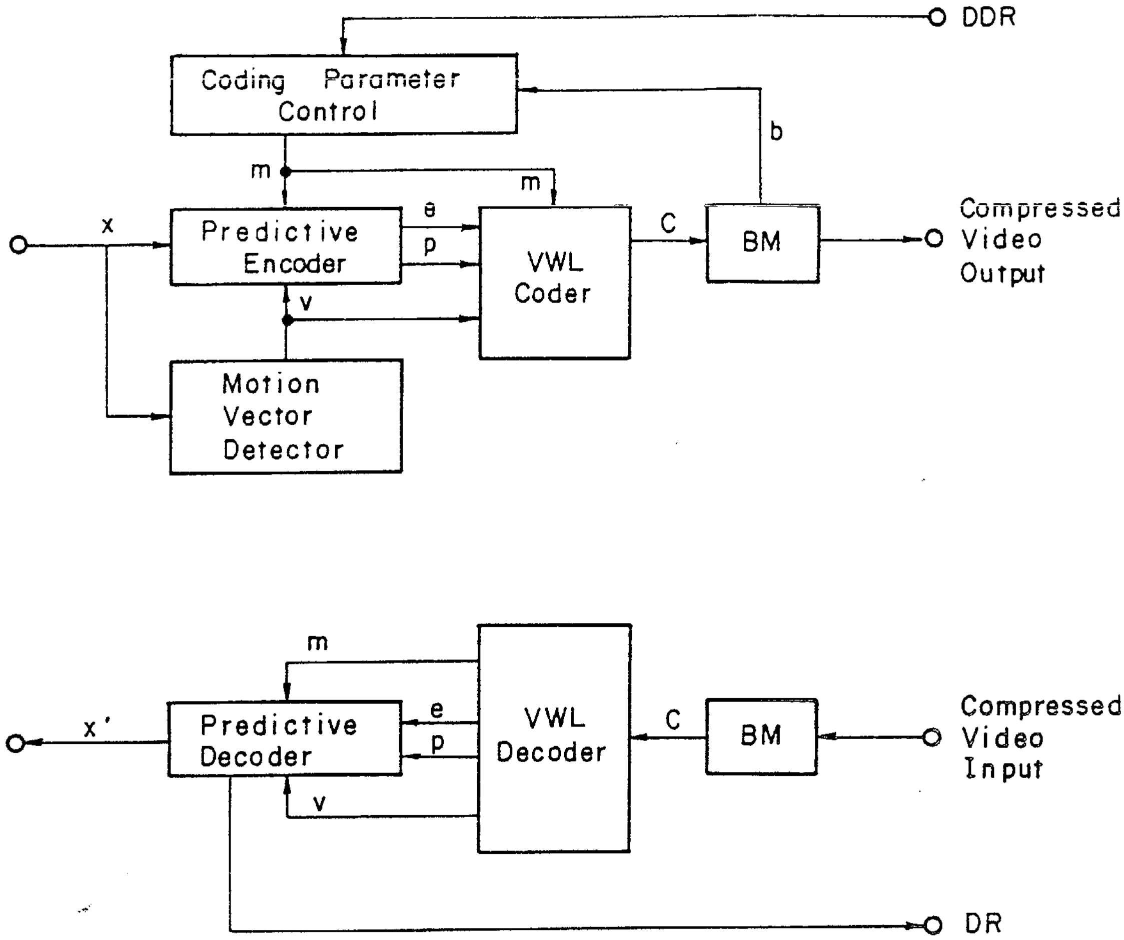


FIGURE 2. TDM signal format



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 information

VWL : Variable Word Length

BM : Buffer Memory

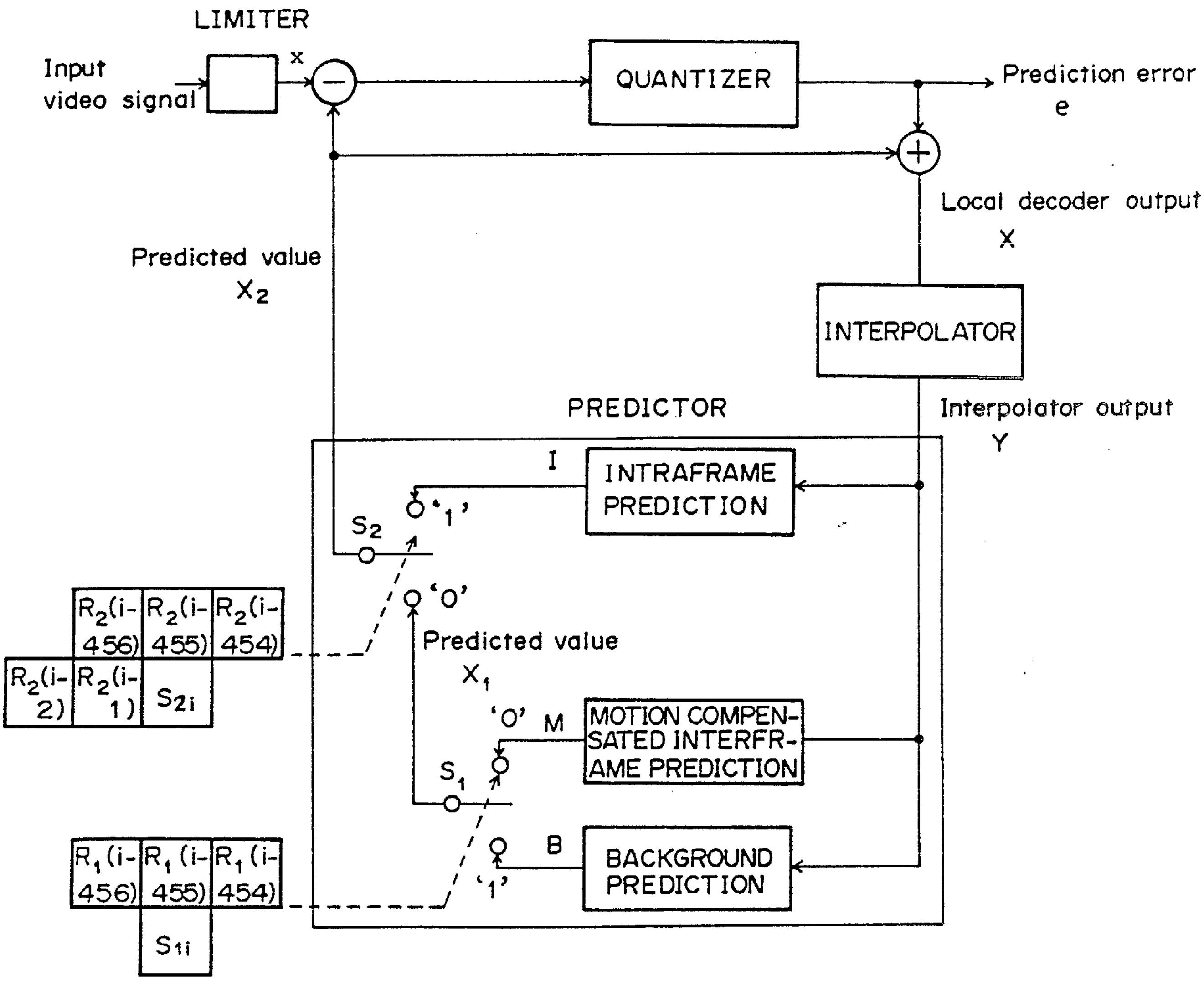
DR : Demand Refresh Request information

DDR : Demand Refresh Confirmation information

FIGURE 3 Source coder and decoder configuration

TABLE 1. Coding modes

Coding Modes		Abbreviation	Operation
1	Normal	NRM	Full sampling
2	Field Repetition	FRP	One or more fields omission
3	Subsampling	SBS	2 : 1 pel omission
4	Stop	STP	Suspension of coding
5	Refresh	RFS	Renewal of frame memory



S : prediction status signal
R : prediction reference signal

FIGURE 4. Adaptive prediction

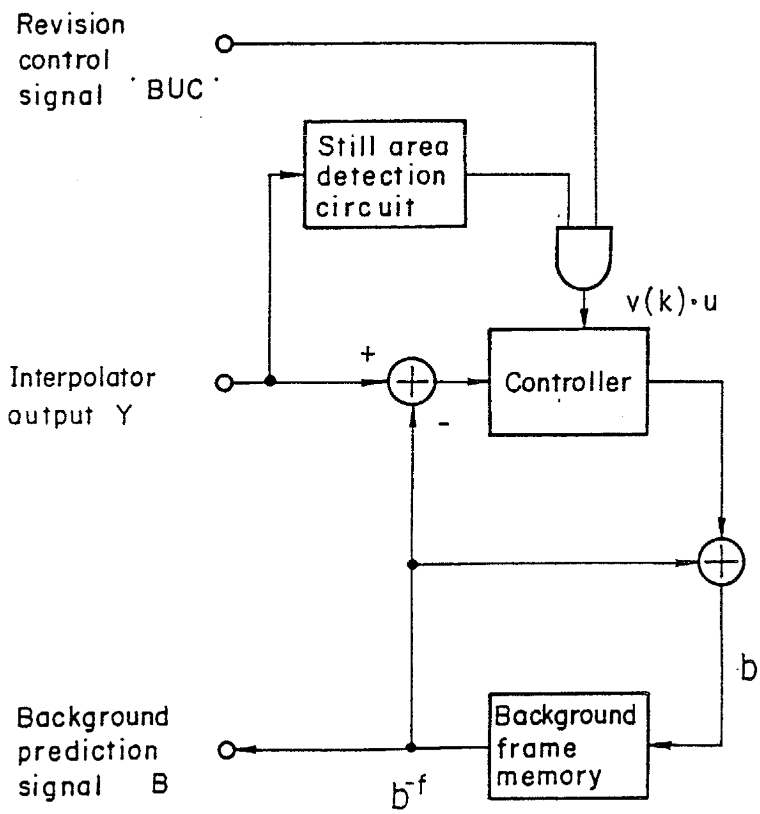
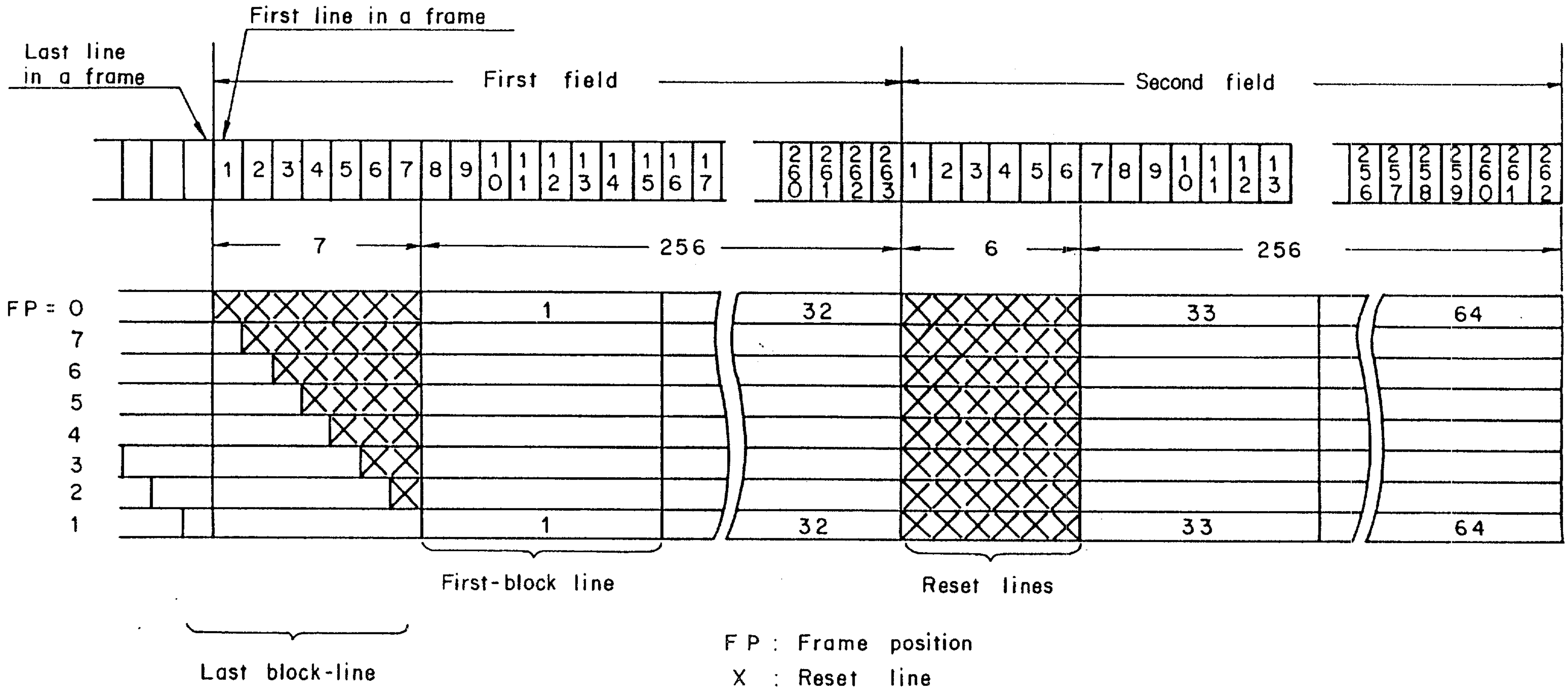
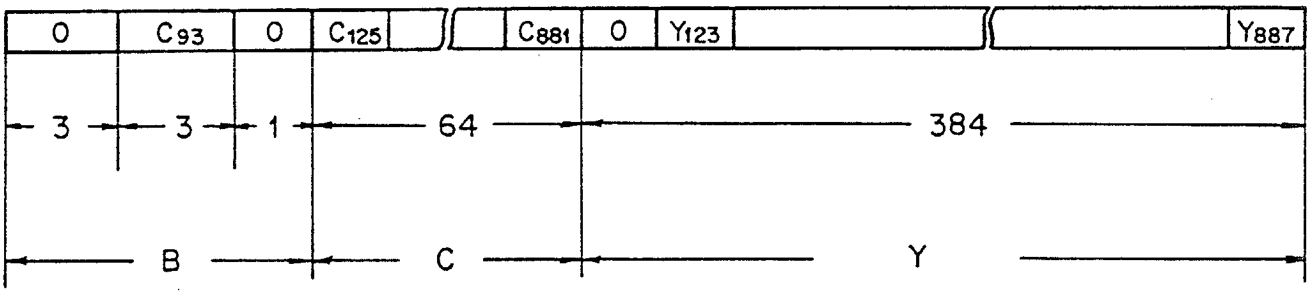


FIGURE 5. Background generation



Note : When the last line in a frame falls on the n-th line of the last block-line, FP = mod(n,8).

FIGURE 6. Definition of block-line and reset line



<u>Note</u> : This figure shows an odd line in the first field. See Fig. 2 as for pel numbers.

FIGURE 7. Zero insertion for preventing crosstalks due to interpolation

TABLE 2 Prediction and interpolation functions

Coding Mode		Kind of Pel	Prediction Functions $P(Z)$ (Note 1)			Interpolation Functions $I(Z)$ (Note 2)		
			$P_Y(Z)$	$P_C(Z)$	$P_B(Z)$	$I_Y(Z)$	$I_C(Z)$	$I_B(Z)$
Normal		Coded	Z^{-1} ; $S_2 = 1$ Z^{-F+V} ; $S_2 = 0,\ S_1 = 0$ $P_b(Z)$ (Note 3) ; $S_2 = 0,\ S_1 = 1$		Z^{-F}	1		
Subsampling		Coded	Z^{-2} (Note 4) ; $S_2 = 1$ Z^{-F+V} ; $S_2 = 0,\ S_1 = 0$ $P_b(Z)$; $S_2 = 0,\ S_1 = 1$			1		
		Omitted	(not defined)			$\frac{1}{2}\{\frac{1}{2}(Z^{-1}+Z^{+1}) + \frac{1}{2}(Z^{-H}+Z^{+H})\}$	$\frac{1}{2}(Z^{-1}+Z^{+1})$	
Field Repetilion		Omitted	(not defined)			$\frac{1}{2}(Z^{-262H}+Z^{-263H})$	Z^{-263H} ; first field Z^{-262H} ; second field	
Refresh	NRM	Coded	Z^{-1}			1		
	SBS	Coded	Z^{-2} (Note 4)			1		
		Omitted	(not defined)			$\frac{1}{2}(Z^{-1}+Z^{+1})$		

Note 1 : S_1 and S_2 are prediction status signals defined in 6.2.2.

Note 2 : To deal with fractions generated by operation of (A+B)/2, (A+B+1)/2 is executed and the 8 bits from MSB are used.

Note 3 : Background is generated as described in 6.2.3.

Note 4 : Z^{-1}, if the previous pel is coded.

TABLE 3. Quantizing characteristics

Q_0

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

Q_1

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

Q_2

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

Q_3

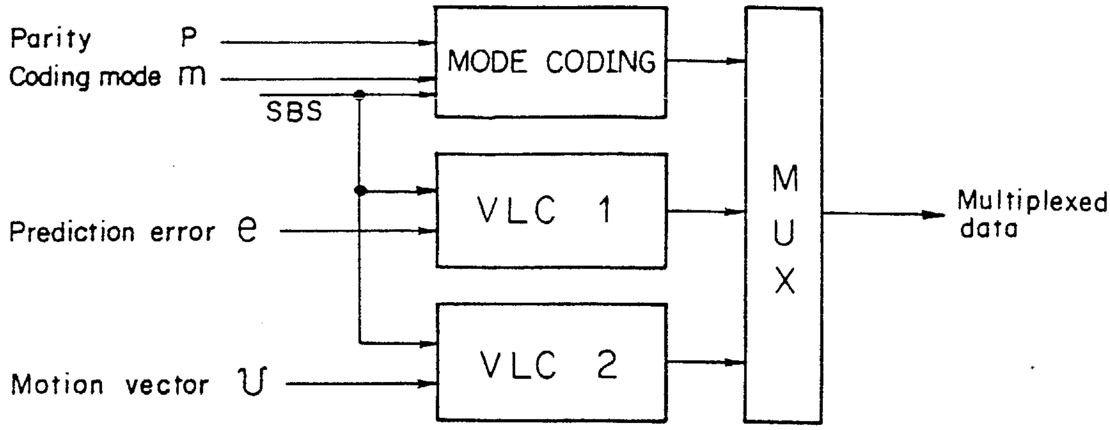
Input Range	Output Level
0 ~ 6	0
7 ~ 11	7
12 ~ 17	14
18 ~ 24	21
25 ~ 31	28
32 ~ 38	35
39 ~ 45	42
46 ~ 52	49
53 ~ 59	56
60 ~ 66	63
67 ~ 73	70
74 ~ 80	77
81 ~ 87	84
88 ~ 94	91
95 ~ 101	98
102 ~ 108	105
109 ~ 115	112
116 ~ 123	119
124 ~ 255	127

Note : Characteristics are symmetrical with respect to zero

TABLE 4 Coding parameter control unit and commands

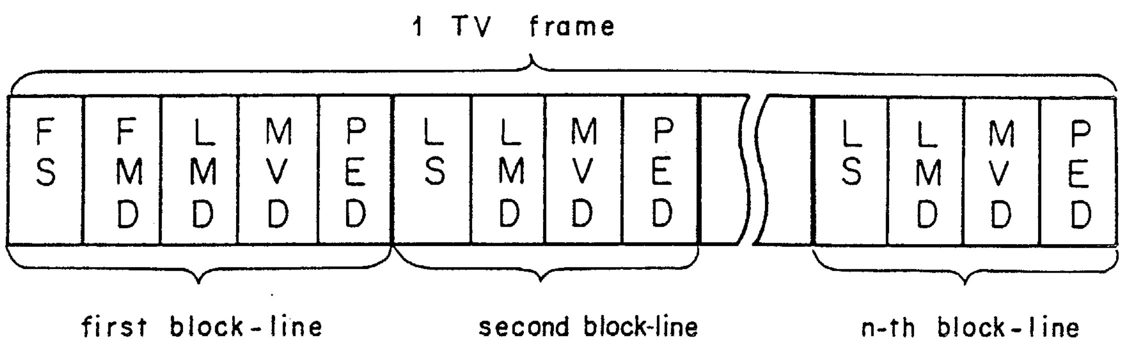
Coding Parameter	Control Unit	Commands
Normal	Frame Block-line (8 lines) Block (8 x 16 pels)	SBC = 1, IFM = 1, FRP = 1 and TRANS (SBS : off)
Quantization	Block-line	QC 1 and QC 2
Field Repetition	Block-line (Note)	FRP = 0
Sub Sampling	Frame	SBC = 0 and FRP = 1
	Block	TRANS (SBS:ON) and FRP=1
Stop	arbitrary	Prediction error e = 0 , Motion vector V = 0
Demand Refresh	Block - line	DRM = 0 and IFM = 0
Cyclic Refresh	Two lines	DRM = 1 , IFM = 0 and CRM 1 , 2

Note : Consecutive 32 block-lines from the first through the 32nd block-lines or from the 33rd through the 64th block-lines are omitted for ordinary field repetition. Other methods are also possible using the FRP command controlled in unit of block-line.



SBS : Subsampling
VLC : Variable Wordlength Coder

FIGURE 8. Configuration of entropy coding



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

FIGURE 9. Multiplexed data format

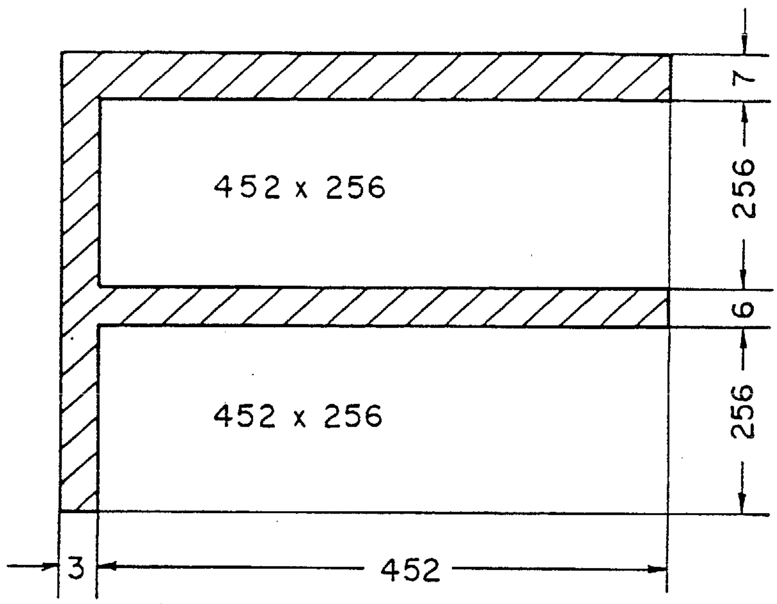


FIGURE 10. Entropy - coded pels

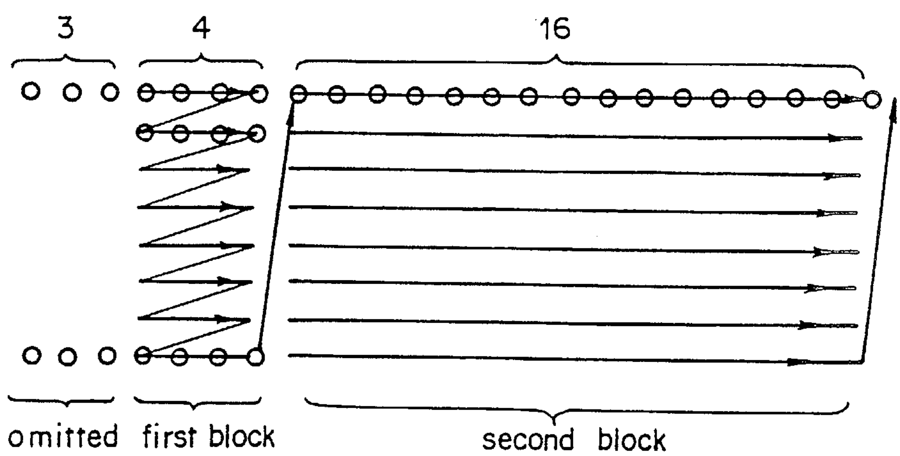
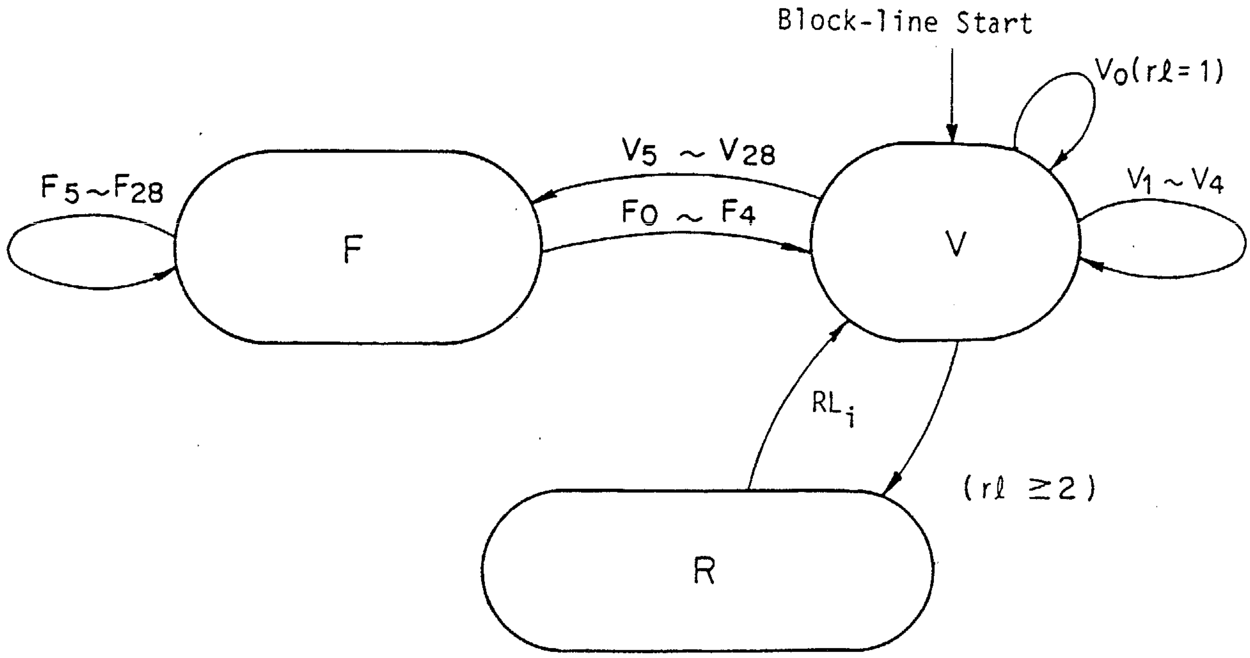


FIGURE 11. Scanning sequence



Note 1 : RL is a length of the run to be coded, while $r\ell$ is a number of continued pels whose $e = 0$.

Note 2 : Prediction error data start with an R or V code. R code is used if $r\ell \geq 2$, while V code is used if otherwise.

Note 3 : The code can shift to V even if $RL \geq 2$ to prevent a buffer memory underflow.

FIGURE 12. Code transition rule for prediction error data

TABLE 5 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 6 Run length code for zero amplitude prediction errors

RL (Note 1)	Code Length	Code Word				Remark
2	5	00	0 0 1			
3	5	00	0 0 0			
4	6	00	1 0 1 0			
5	6	00	1 0 0 1			
6	6	00	1 0 0 0			
7	7	00	1 0 1 1 1			
8 ~ 11	7	00	1 1 0 X X			X = 11 - RL
12	8	00	1 1 1 1 0 1			
13	8	00	1 1 1 1 0 0			
14 ~ 17	8	00	1 1 1 0 X X			X = 17 - RL
18 ~ 25	9	00	0 1 1 1 X X	X		X = 25 - RL
26 ~ 33	10	00	0 1 1 0 0 X	X X		X = 33 - RL
34 ~ 37	10	00	0 1 0 1 0 0	X X		X = 37 - RL
38 ~ 64	12	00	0 1 0 0 1 X	X X X X		X = 69 - RL
MK1	13	00	1 0 1 1 0 Y	Y Y Y Y Y		Y = 0 ~ 63
MK2	14	00	1 1 1 1 1 1	Y Y Y Y Y Y		
MK3	14	00	1 1 1 1 1 0	Y Y Y Y Y Y		
MK4 ~ 7	15	00	0 1 1 0 1 X	X Y Y Y Y Y Y		X = 7 - MK
MK8 ~ 15	16	00	0 1 0 1 1 X	X X Y Y Y Y Y Y		X = 15 - MK
MK16 ~ 19	16	00	0 1 0 1 0 1	X X Y Y Y Y Y Y		X = 19 - MK
MK20 ~ 34	18	00	0 1 0 0 0 1	X X X X Y Y Y Y	Y Y	X = 35 - MK
MK35 ~ 49	19	00	0 1 0 0 0 0	1 X X X X Y Y Y	Y Y Y	X = 50 - MK
MK50 ~ 56	19	00	0 1 0 0 0 0	0 1 X X X Y Y Y (Note 2)	Y Y Y	X = 57 - MK

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

Note 2 : The maximum run length is (455-3) x 8 - 1 = 3615. Corresponding MK and Y turn out to be 56 and 30, respectively. Hence, $0 \leqq Y \leqq 30$ for MK = 56.

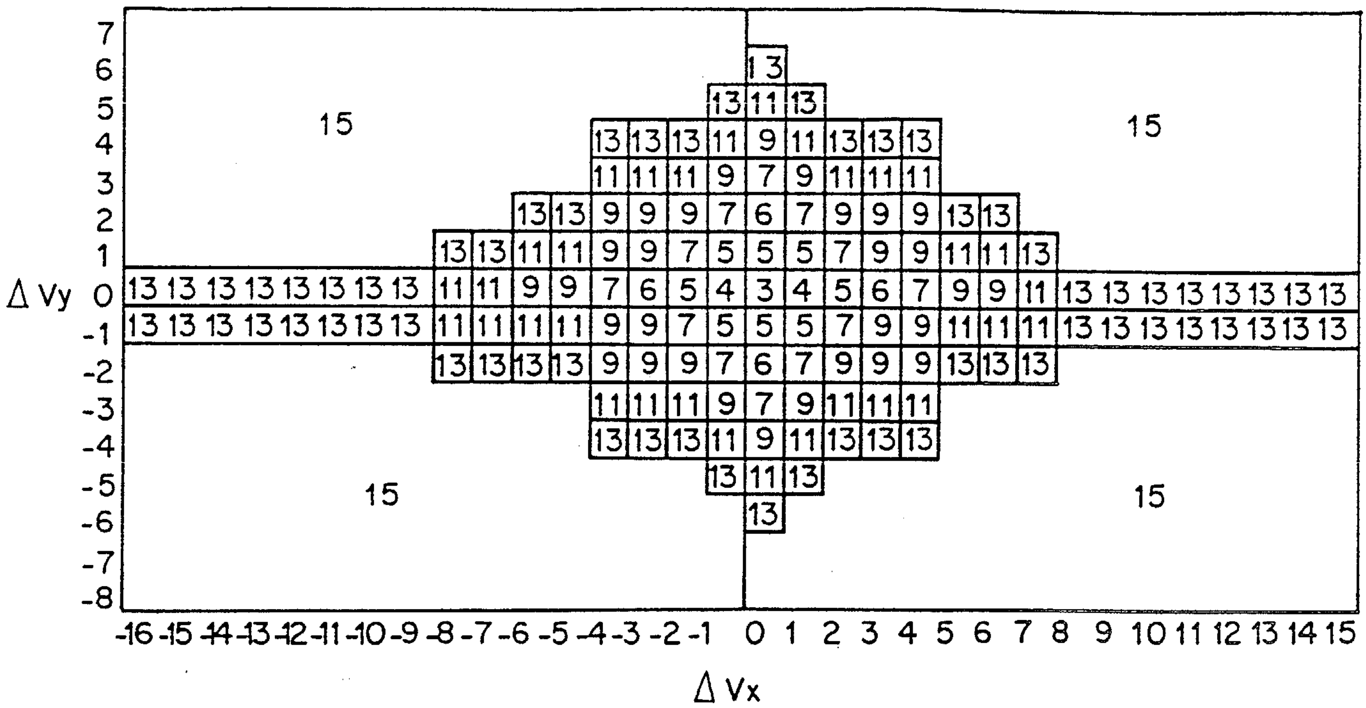
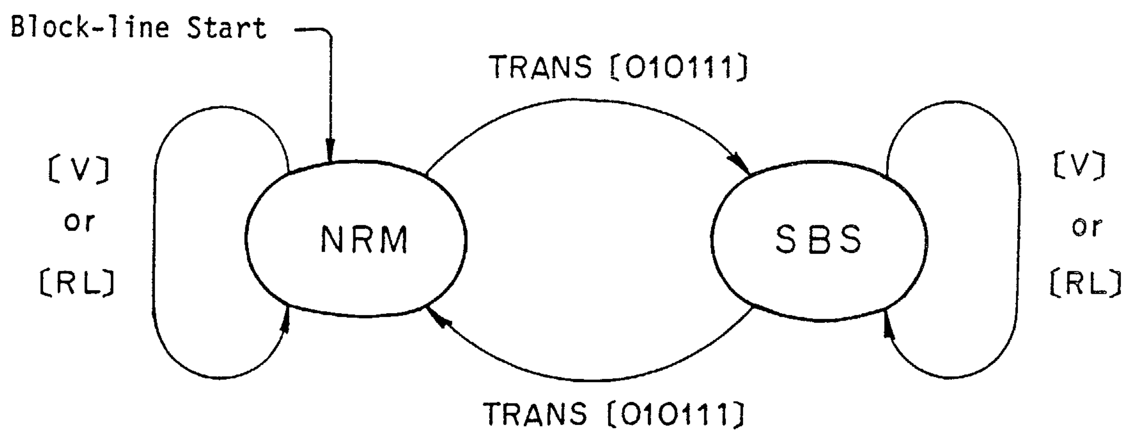


FIGURE 13. Word length for motion vector prediction errors



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

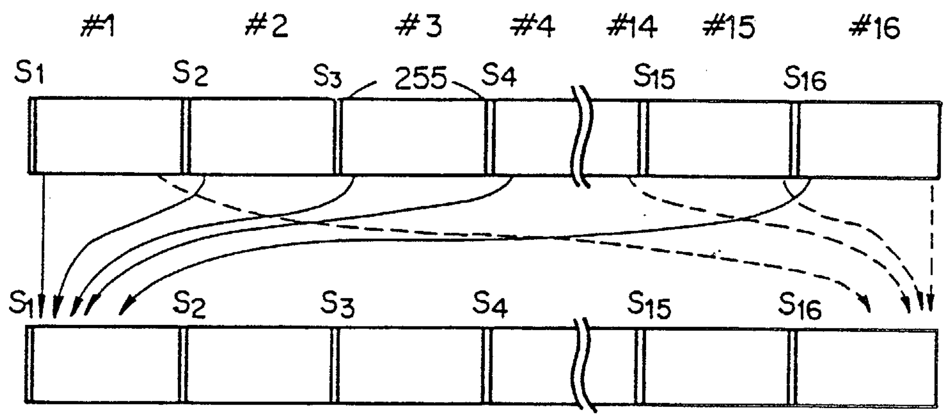
FIGURE 14. Code transition rule for motion vector data and Normal / Subsampling mode change

TABLE 7 Variable length code and run length code for motion vector data

ΔVx	ΔVy	Code Length	Code Word			Number of Codes
±1	0	4	0 0 1 Sx			2
±1	±1	5	1 1 1 Sx Sy			8
0	±1	5	1 1 0 1 Sy			
±2	0	5	1 1 0 0 Sx			
0	±2	6	1 0 1 1 1 Sy			4
±3	0	6	1 0 1 1 0 Sx			
±1	±2	7	1 0 0 1 1 Sx Sy			12
±2	±1	7	1 0 0 1 0 Sx Sy			
0	±3	7	1 0 0 0 1 1 Sy			
±4	0	7	1 0 0 0 1 0 Sx			
±3	±1	9	1 0 1 0 1 1 1 Sx Sy			30
±1	±3	9	1 0 1 0 1 1 0 Sx Sy			
±2	±2	9	1 0 1 0 1 0 1 Sx Sy			
±3	±2	9	1 0 1 0 1 0 0 Sx Sy			
±4	±1	9	1 0 1 0 0 1 1 Sx Sy			
±4	±2	9	1 0 1 0 0 1 0 Sx Sy			
±5	0	9	1 0 1 0 0 0 1 1 Sx			
±6	0	9	1 0 1 0 0 0 1 0 Sx			
0	±4	9	1 0 1 0 0 0 0 1 Sy			
-8 ~ 7	-5 ~ +5 (see Fig.13)	11	1 0 0 0 0 1	X X X X Sy	[X] = ΔVx	32
-16~15	-6 ~ +6 (see Fig.13)	13	0 1 0 0 0 0 1	X X X X X Sy	[X] = ΔVx	64
-16~15	-8 ~ +7 (see Fig.13)	15	1 0 0 0 0 0	X X X X X Y Y Y Y	[X] = ΔVx [Y] = ΔVy	359

R L	Code Length	Code Word		Number of Codes
1	3	0 0 0		1
2	4	0 1 1 1		1
3 ~ 6	6	0 1 1 0 X X	X X = 6 - RL	4
7 ~ 12	7	0 1 0 1 X X X	X X X = 12 - RL	6
13 ~ 20	8	0 1 0 0 1 X X X	X X X = 20 - RL	8
21 ~ 28	9	0 1 0 0 0 1 X X X	X X X = 28 - RL	8
TRANS	6	0 1 0 1 1 1		1

Note 1 : Sx and Sy denote signs. Si = 0 for positive, Si = 1 for negative.
Note 2 : XX..X and YY..Y are expressed in two's complement form (MSB first).

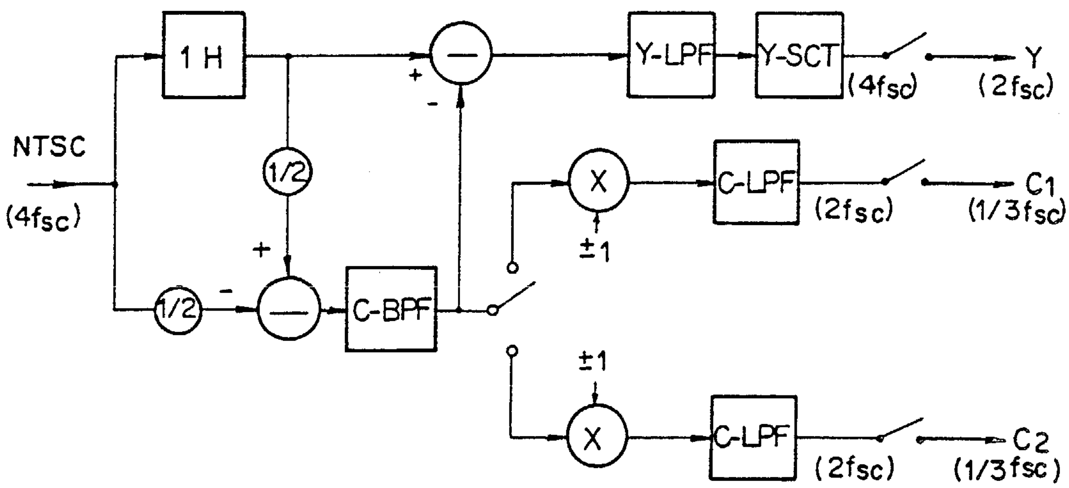


$S_1 S_3 S_5 S_7 S_9 S_{11} S_{13}$ = 0001101
S_{15} : multiframe alignment signal
$S_2 S_4 S_6$ ... S_{16} : control information

FIGURE 15. Error correction frames and interleaving

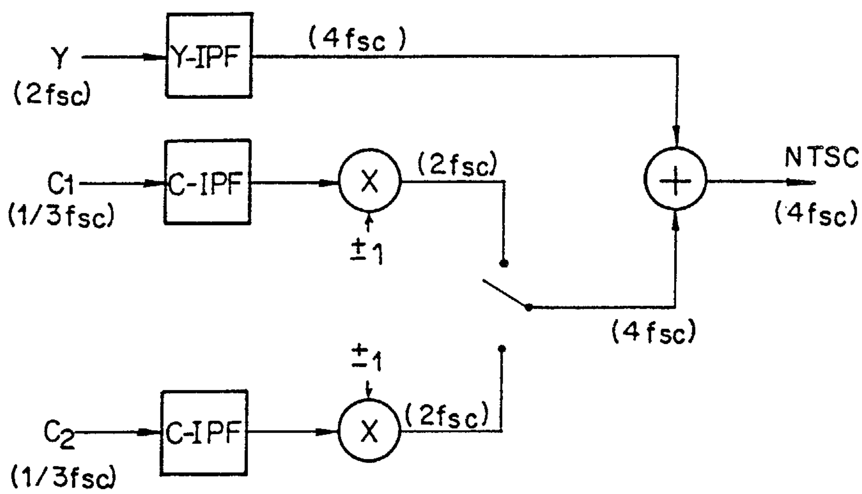
ANNEX 1 Color Decoding and Coding Filters

1. Configuration



H : Line delay
Y-LPF: Low Pass Filter for Y signal
Y-SCT: SubCarrier Trapper
C-BPF: Band Pass Filter for C signal
C-LPF: Low Pass Filter for C signal

(a) Digital color separation circuit



Y-IPF: Interpolation Filter for Y signal
C-IPF: Interpolation Filter for C signal

(b) Digital color composition circuit

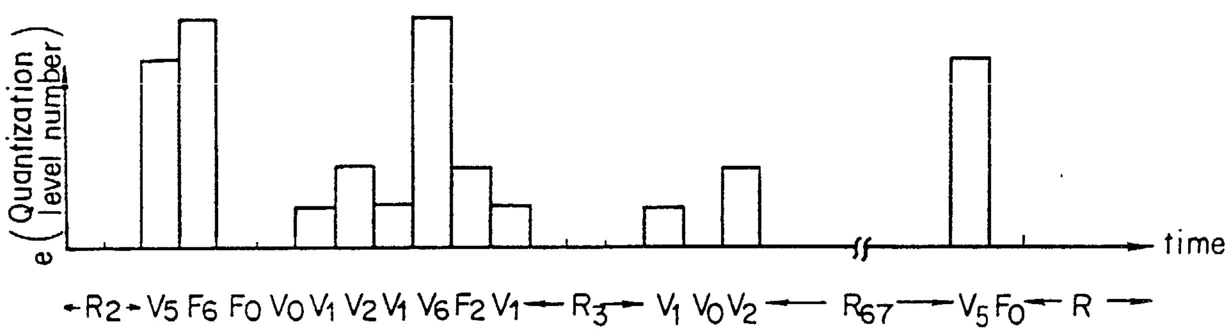
2. Basic Filter Characteristics

Filter	Transfer Function H (z)
C-BPF	$(-Z^{-2}+2-Z^{2})/4$
Y-LPF	$(-3Z^{-3}+19Z^{-1}+32+19Z-3Z^{3})/64$
Y-SCT	$(Z^{-5}-3Z^{-3}+10Z^{-1}+10Z-3Z^{3}+Z^{5})/16$
C-LPF	$(Z^{-4}+3Z^{-2}+4+3Z^{2}+Z^{4})/12$
Y-IPF	$(-3Z^{-3}+19Z^{-1}+32+19Z-3Z^{3})/64$
C-IPF	$(Z^{-2}+1+Z^{2})(Z^{-1}+2+Z)(-Z^{-8}-2Z^{-6}+2Z^{-4}+6Z^{-2}+6+6Z^{2}+2Z^{4}-2Z^{6}-Z^{8})/192$

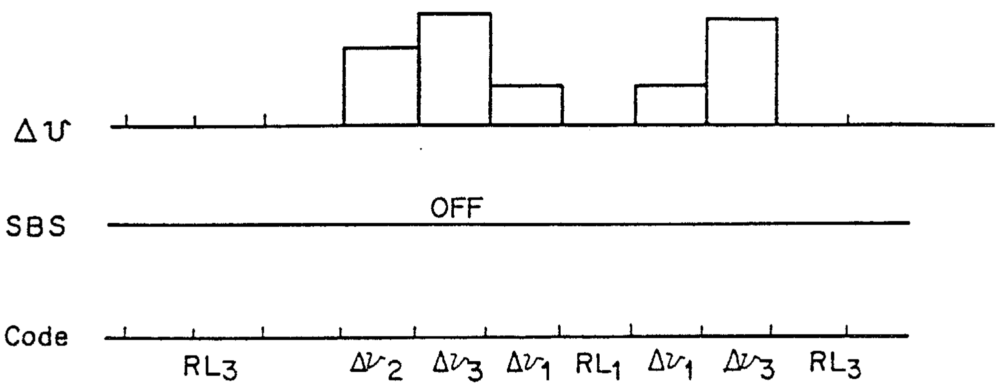
3. Advanced Filter Characteristics

Filter	Transfer Function H(z)
C-BPF	$(Z^{-8}-9Z^{-6}+17Z^{-4}-23Z^{-2}+28-23Z^{2}+17Z^{4}-9Z^{6}+Z^{8})/128$
Y-LPF	$(-Z^{-7}+4Z^{-5}-10Z^{-3}+39Z^{-1}+64+39Z-10Z^{3}+4Z^{5}-Z^{7})/128$
Y-SCT	$(Z^{-5}-3Z^{-3}+10Z^{-1}+10Z-3Z^{3}+Z^{5})/16$
C-LPF	$(Z^{-4}+3Z^{-2}+4+3Z^{2}+Z^{4})/12$
Y-IPF	$(-Z^{-7}+4Z^{-5}-10Z^{-3}+39Z^{-1}+64+39Z-10Z^{3}+4Z^{5}-Z^{7})/128$
C-IPF	$(Z^{-2}+1+Z^{2})(Z^{-1}+2+Z)(-Z^{-8}-2Z^{-6}+2Z^{-4}+6Z^{-2}+6+6Z^{2}+2Z^{4}-2Z^{6}-Z^{8})/192$

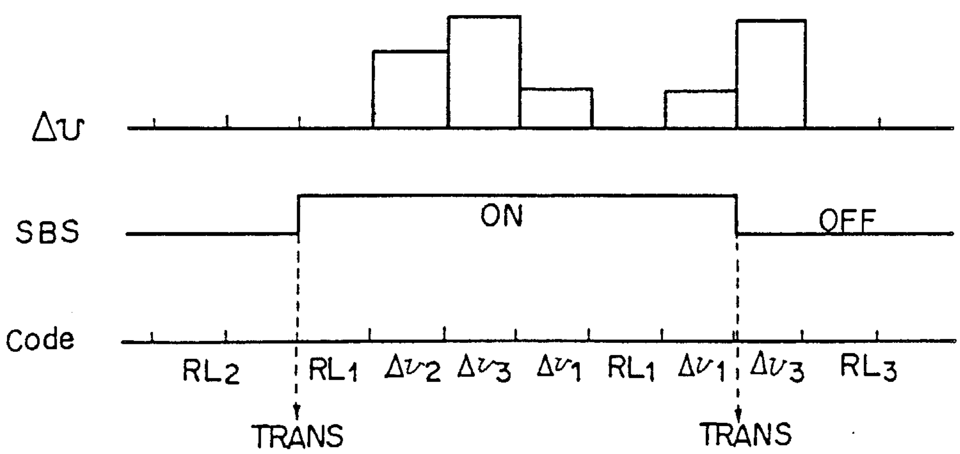
ANNEX 2 Examples of Entropy Coding



(a) Coding of prediction error e



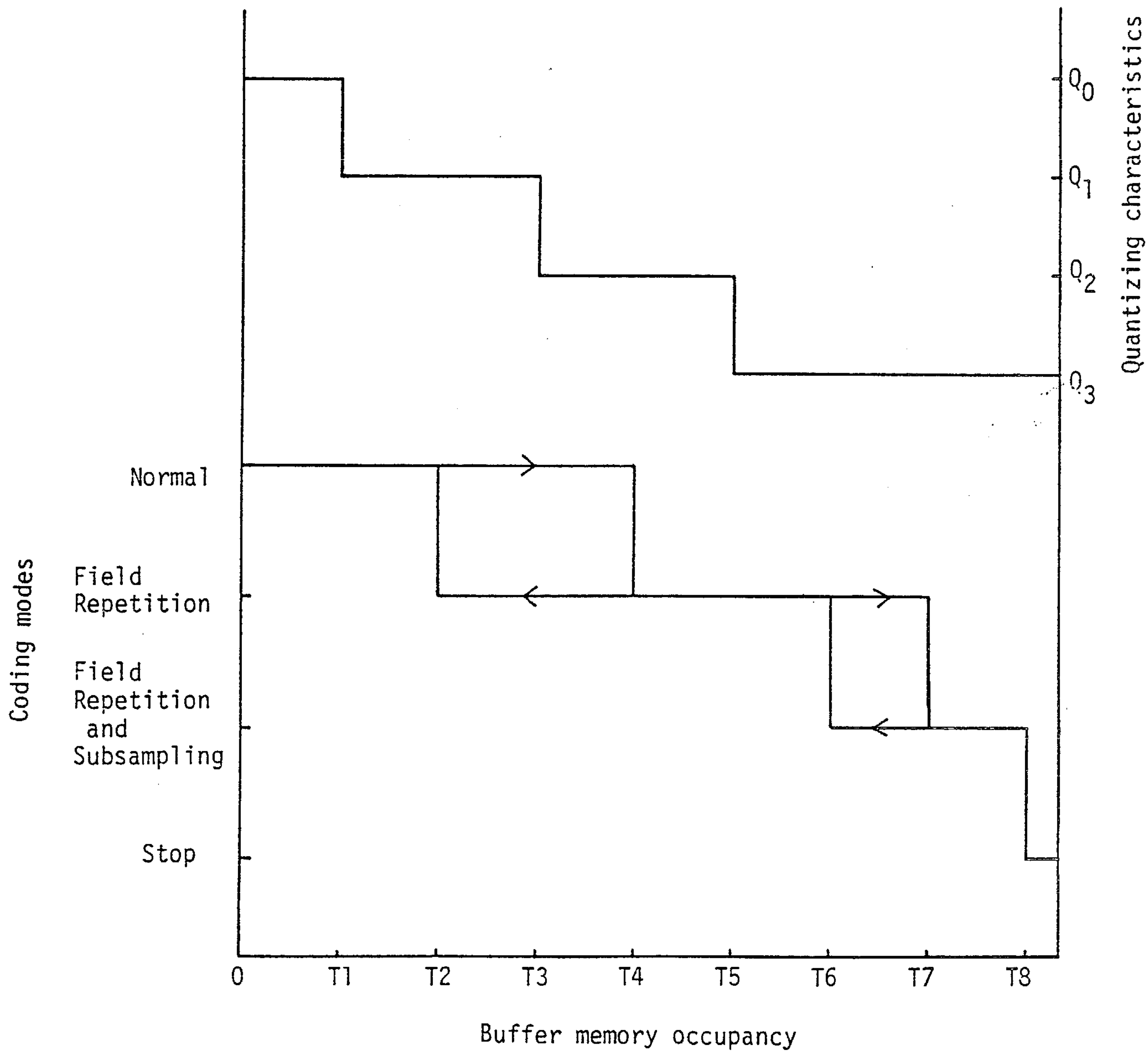
When SBS is off



When SBS is on

(b) Coding of motion vector v

ANNEX 3 An Example of Coding Control Sequence



ANNEX 2 : A Draft for PART 3 of Recommendation H.130

CHARACTERISTICS OF A 1544 KBIT/S (n=4) FRAME STRUCTURE FOR USE WITH CODECS DESCRIBED IN PART 3 OF RECOMMENDATION H.120

1. General Characteristics
   1.1 Fundamental Characteristics
   1.2 Bit Rate
   1.3 Timing Signal
   1.4 Interfaces
   1.5 Format Restrictions Enforced by the Network
2. Frame Structure and Bit Allocation
   2.1 Frame Alignment
   2.2 Audio Signal
   2.3 Codec-to-codec Information
   2.4 Auxiliary Data Information
   2.5 Encoded Video
3. Codec-to-codec Information Channel
   3.1 C_1 Bit
   3.2 C_2 Bit : Stuffing flag
   3.3 C_3 Bit : Codec facility / coding mode
   3.4 C_4 Bit : Channel assignment flag
   3.5 C_5 Bit : Message channel 1
   3.6 C_6 Bit : Message channel 2
   3.7 C_7 Bit : Stuffing flag
   3.8 C_8 Bit : Multiframe alignment
4. Stuffing
   4.1 General
   4.2 Details of Stuffing
   4.3 Stuffing Mode Operation

1. General Characteristics

The multiplex structure described in this Recommendation is suitable for use on digital paths and connections which interconnect video codecs for videoconferencing or visual telephony using 1544 kbit/s transmission. The connection may either be directly via ISDN defined in Recommendation I.431 or via higher order digital multiplex equipment compatible with the primary PCM multiplex equipment defined in Recommendation G.733.

The main features of the multiplex structure are that it provides

- one 8 kbit/s channel for frame alignment, alarm signals and other signals as required,
- one 64 kbit/s channel for the audio signal,
- one 32 kbit/s channel for codec-to-codec information,
- one optional 64 kbit/s channel for auxiliary data service, and
- the use of the remaining capacity (between 1376 and 1440 kbit/s) for the encoded video signal.

1.1 Fundamental Characteristics

The multiplex structure contains 192 bits per frame plus one bit per frame for frame alignment and others. The nominal frame repetition rate is 8000 Hz.

1.2 Bit Rate

The nominal bit rate is 1544 kbit/s with a tolerance of ±50 parts per million (ppm).

1.3 Timing Signal

The timing signal is a 1544 kHz signal from which the bit rate is derived. It should be possible to derive the timing signal either from an internal source or from the network.

1.4 Interfaces

The interfaces should comply with Recommendation G.703. The interface code should be either of AMI/B8ZS described in Recommendation G.703, in addition to which CMI (Coded Mark Inversion) code is also applicable when the codec is installed as a part of terminal equipment. Which of the three codes is used should be determined by bilateral agreement.

1.5 Format Restrictions Enforced by the Network

As indicated in Recommendation G.703, runs of more than 15 'zeros' are forbidden in some networks. Additionally, on the average, there must be at least three 'ones' in every 24 digits. Provision is made by means of a stuffing system to ensure that forbidden patterns do not occur.

2. Frame Structure and Bit Allocation

The basic frame structure follows Recommendation G.704 with changes in bit allocations. The bits in a frame are numbered from 0 to 192, with a transmission frame bit as numbered zero. Remained 192 bits are divided into 24 time slots (TS) in which each time slot has 64 kbit/s rate. Time slot number is assigned to each TS in the way that the first slot is TS1 and the last slot is TS24. Bit allocation in a frame is shown in Fig. 1.

2.1 Frame Alignment

The basic frame alignment is obtained at bit No. 0 as in Recommendation G.704.

2.2 Audio Signal

Audio signal is transmitted at 64 kbit/s in TS1.

2.3 Codec-to-codec Information

This information is transmitted in odd numbered TS2 at the 32 kbit/s channel. Identification of codec-to-codec information is made by detecting multiframe alignment which is inserted in the eighth bit of the odd numbered TS2.

The channel is structured in multiframes of 16 frames each (numbered from 1 to 16) and supermultiframes of 8 multiframes each (numbered from 1 to 8). Multiframe and supermultiframe alignment is obtained from Bit No. 8 in TS2.

The multiframe of the codec-to-codec information channel is quite independent of the multiframe of the transmission frame generated by Bit No. 0.

2.4 Auxiliary Data Information

When required, this information is transmitted basically in TS16, which is used for the encoded video signal when no optional auxiliary equipment is connected. If stuffing is performed due to some channel restrictions, data alignment are represented in 4.2

2.5 Encoded Video

A minimum of 64 x 21.5 kbit/s capacity is basically reserved for encoded video in even numbered TS2, TS3 through TS15 and TS17 through TS24. When the auxiliary data information channel is not set up, the capacity is increased to 64 x 22.5 kbit/s with TS16 added. The available bit rate for the encoded video signal therefore lies between

1376 and 1440 kbit/s. If stuffing is performed, data alignment is represented in 4.2.

3. Codec-to-codec Information Channel

The use of the bits in the codec-to-codec information channel is as follows (see Table 1). The notation, 'm.n.l', is used for a bit position expressing the n-th multiframe and the l-th supermultiframe of Bit No. m.

3.1 C_1 Bit

Bit 1.1, 1.5, 1.9, 1.13 : Permanently set to 1

Bit 1.3, 1.7, 1.11, 1.15 : Spare (Note 1)

3.2 C_2 Bit : Stuffing flag

Bit 2.1-2.15 (odd number) : 0 if not stuffed

Stuffing flag consists of four bits including C_1 and C_7 in each violation detection block (four frame length) which is defined in 4.2. The first three bits are used for majority decision logic at the decoder. When the result indicates 'stuffing', the decoder undergoes destuffing.

3.3 C_3 Bit : Codec facility/coding mode

Bit 3.1 : Codec Facilities
- 3.1.1 : Graphics Mode 1 (high resolution) (0 if provided)
- 3.1.2 : Bit Sequence Independency (0 if secured)
- 3.1.3 : Monochrome Mode (0 if provided)
- 3.1.4 : Video Encryption (0 if provided)
- 3.1.5 : Audio Encryption (0 if provided)
- 3.1.6 : Pointing Function (0 if provided)
- 3.1.7 : Graphics Mode 2 (standard resolution) (0 if provided)
- 3.1.8 : Spare (Note 1)

Bit 3.3 : Spare (Note 1)
Bit 3.5 : Spare (Note 1)
Bit 3.7 : Spare (Note 1)
Bit 3.9 : Spare (Note 1)
Bit 3.11 : Spare (Note 1)
Bit 3.13 : Spare (Note 1)
Bit 3.15 : Coding Mode
- 3.15.1 : Video Encryption (0 if used)
- 3.15.2 : Audio Encryption (0 if used)
- 3.15.3 : Frame Memory Refresh Request (0 if requested)
- 3.15.4 : Backward Path (0 if available)
- 3.15.5-3.15.8 : Spare (Note 1)

3.4 C_4 Bit : Channel assignment flag

Bit 4.1, 4.3, 4.5, 4.7 : Auxiliary Data Channel Flag (0 if used)

Bit 4.9, 4.11, 4.13, 4.15 : Graphics Mode Flag (0 if used)

In Graphics Mode, video data are inhibited and their bit positions are used for graphics data transmission.

These two flags consist of four bits as Stuffing Flag. Both auxiliary data and graphics data can be inserted or removed in a unit of multiframe (16 frames). Flags should precede the data by a multiframe.

3.5 C_5 Bit : Message channel 1

Bit 5.1-5.15 (odd number) : Message Channel 1 (Note 2)

3.6 C_6 Bit : Message channel 2

Bit 6.1-6.15 (odd number) : Message Channel 2 (Note 2)

3.7 C_7 Bit : Stuffing flag

Bit 7.1-7.15 (odd number) : 0 if stuffed

3.8 C_8 Bit : Multiframe alignment

Bit 8.1, 8.3, 8.5, 8.7, 8.9, 8.11, 8.13 : Multiframe Alignment Signal (1110010)

Bit 8.15 : Supermultiframe Alignment Signal (1110010*) (Note 3)

Note 1 : Spare bits are set to 1.
Note 2 : Protocols for these message channels are under study.
Note 3 : The bit * is used for future higher order multiframing.

4. Stuffing

4.1 General

The bit sequence produced by a videoconferencing codec is not subject to any limitation on the bit patterns that are generated. Therefore, reversible processing has to be carried out at the output and input ports to ensure that the format restrictions specified for some 1544 kbit/s networks (described in 1.5 above) are not violated.

As the surest way, the stuffing method should be employed in which necessary 'ones' are inserted, or stuffed, if any violations are found in a block of bitstreams to be transmitted. A flag is attached to the block to identify whether or not the block is stuffed.

4.2 Details of Stuffing

(1) Each block, which consists of 4 transmission frame length, 4 x 193 = 772 bits starting from C_1 bit of codec-to-codec information in the (4n-3)th frame, is checked. If any violations occur with respect to

- no more than 15 consecutive zeros, and
- at least 3 ones in any 24 bits,

ones are stuffed as follows:

TS1 not stuffed
TS2 not stuffed in odd numbered frames, stuffed at the top bit of TS in even numbered frames
TS3-23 stuffed at the top bit of each TS
TS24 stuffed at the top bit and bottom bit of TS

The stuffing position is shown in Fig. 1.

Note : When stuffing pulses are inserted, the transmission bit rate for encoded video is reduced to 1252 kbit/s without auxiliary data transmission and to 1188 kbit/s with auxiliary data transmission.

(2) In order to ease processing at block boundaries, the C_1 Bit at the start of any block is so assigned to be always one as described in 3.1 above and as shown in Table 1.

(3) To prevent 8 consecutive zeros in the codec-to-codec information when stuffing is carried out, the stuffing flag transmitted in the (C_2, C_7) Bits are assigned as (1,0) for stuffing and (0,1) for no stuffing.

(4) Violations are checked assuming that all Transmission Framing bits in Bit No. 0 and Stuffing Flag bits in C_2 and C_7 as zero.

Note : If audio data are processed in the network, corresponding bits should also be assumed as zero for violation checking. However, as this may increase the probability of stuffing, measures are necessary to prevent such stuffing from becoming excessive.

4.3 Stuffing Mode Operation

Stuffing should be operated only when necessary. To identify the network restrictions, the Bit Sequence Independency (BSI) bit in the codec-to-codec information channel is used. A coder usually operates without stuffing, but shifts to the stuffing mode if the received BSI is 'one'.

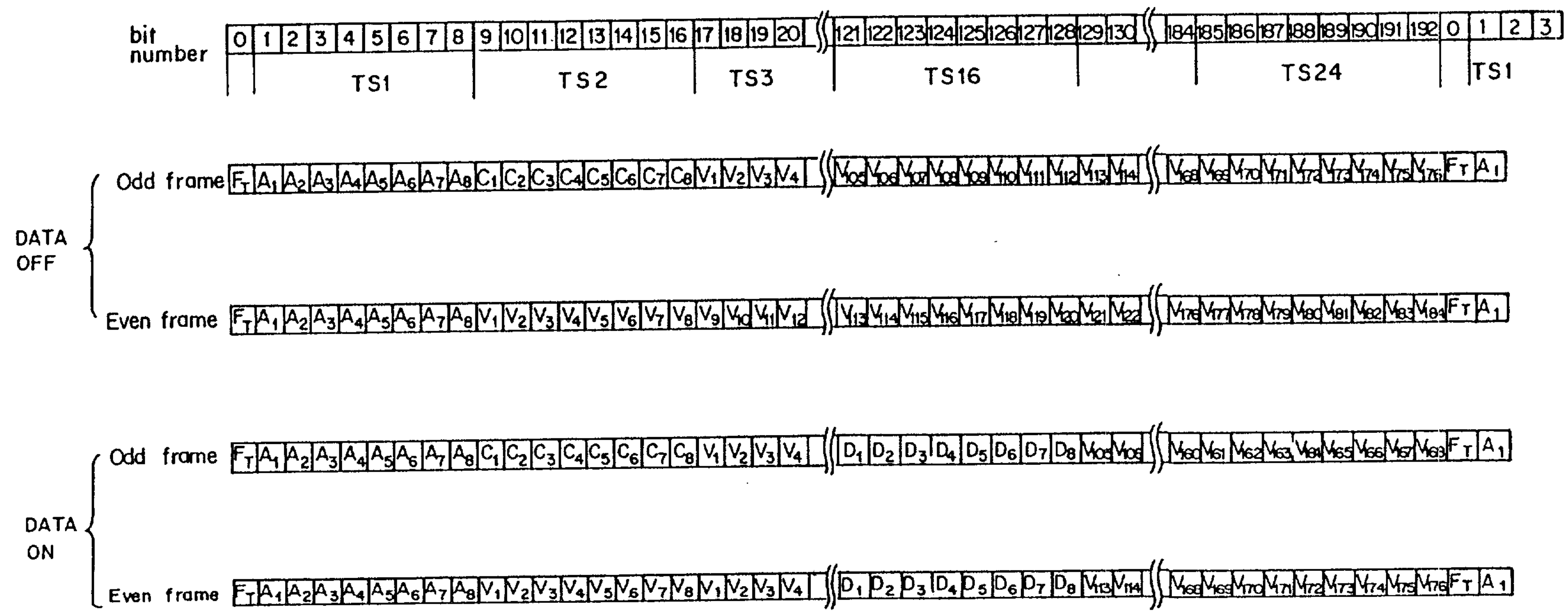


Figure 1 (a)   Frame structure and bit allocation (without stuffing)

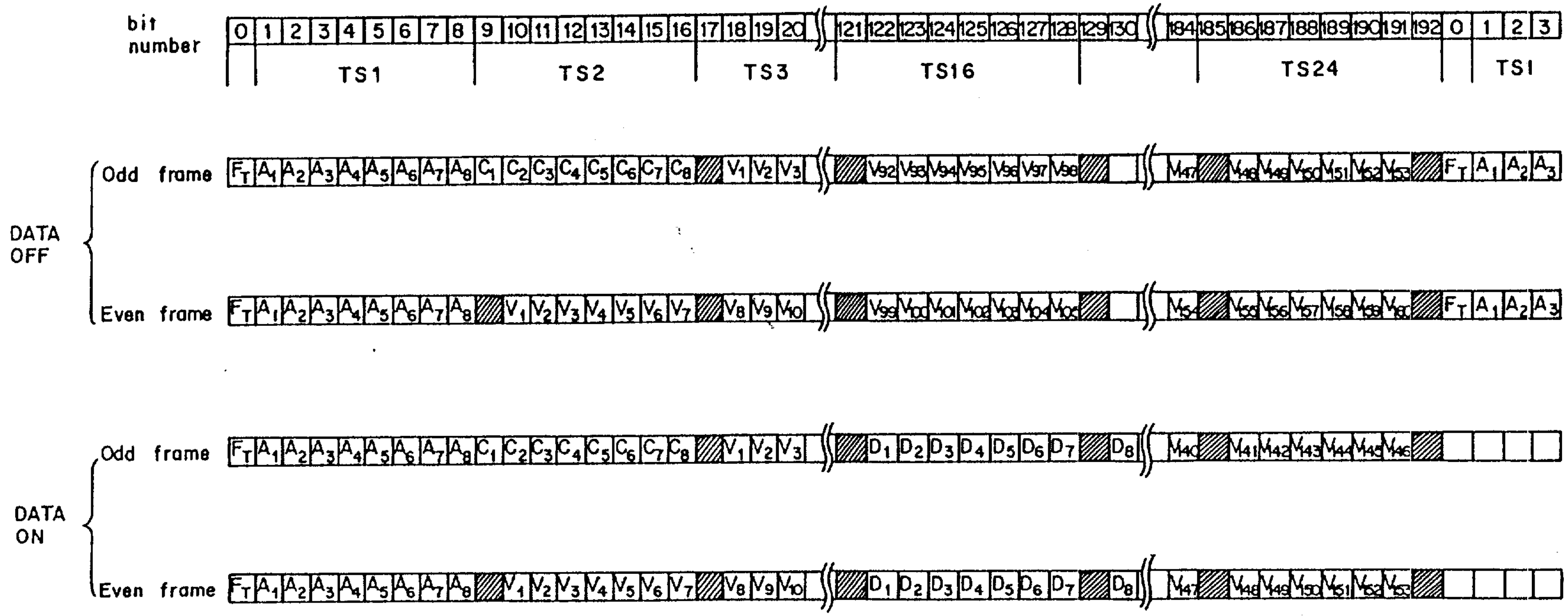


Figure 2 (b)   Frame structure and bit allocation (with stuffing)

Table 1 Codec-to-codec information

<table>
<tr><th>Multiframe Number</th><th>C_1</th><th>C_2</th><th>C_3</th><th>C_4</th><th>C_5</th><th>C_6</th><th>C_7</th><th>C_8</th></tr>
<tr><td>1</td><td>1</td><td rowspan="2">Stuffing Flag</td><td>Codec Facility</td><td rowspan="4">Data Channel Flag</td><td rowspan="8">Message Channel 1</td><td rowspan="8">Message Channel 2</td><td rowspan="2">Stuffing Flag</td><td>MAS (1)</td></tr>
<tr><td>3</td><td>Spare</td><td rowspan="6">Spare</td><td>MAS (1)</td></tr>
<tr><td>5</td><td>1</td><td rowspan="2">Stuffing Flag</td><td rowspan="2">Stuffing Flag</td><td>MAS (1)</td></tr>
<tr><td>7</td><td>Spare</td><td>MAS (0)</td></tr>
<tr><td>9</td><td>1</td><td rowspan="2">Stuffing Flag</td><td rowspan="4">Graphics Mode Flag</td><td rowspan="2">Stuffing Flag</td><td>MAS (0)</td></tr>
<tr><td>11</td><td>Spare</td><td>MAS (1)</td></tr>
<tr><td>13</td><td>1</td><td rowspan="2">Stuffing Flag</td><td rowspan="2">Stuffing Flag</td><td>MAS (0)</td></tr>
<tr><td>15</td><td>Spare</td><td>Coding Mode</td><td>SAS</td></tr>
</table>

MAS : Multiframe Alignment Signal
SAS : Supermultiframe Alignment Signal (1110010 * : * is for future use)

LIST OF ERRATA

SOURCE : JAPAN

TITLE : A DRAFT FOR PARTS 3 OF RECOMMENDATIONS H.120 AND H.130

Page		Error	Correction
1	line 4	Period 1981-1984	Period 1985-1988

ANNEX 1 : A Draft for PART 3 of Recommendation H.120

A CODEC FOR 525-LINES, 60 FIELD/S AND
1544 KBIT/S TRANSMISSION FOR INTRA-REGIONAL USE

Page		Error	Correction
2	line 11	8.4 Stuffing	8.4 Scrambling
30	Table 4	Quantization--- Block-line (Note) Field Repetition---- Field (32 block-line)	See an annexed paper

ANNEX

TABLE 4 Coding parameter control unit and commands

Coding Parameter	Control Unit	Commands
Normal	Frame Block-line (8 lines) Block (8 x 16 pels)	SBC = 1, IFM = 1, FRP = 1 and TRANS (SBS : off)
Quantization	Block-line	QC 1 and QC 2
Field Repetition	Block-line (Note)	FRP = 0
Sub Sampling	Frame	SBC = 0 and FRP = 1
	Block	TRANS (SBS : on)
Stop	arbitrary	Prediction error e = 0 , Motion vector = 0
Demand Refresh	Block - line	DRM = 0 and IFM = 0
Cyclic Refresh	Two lines	DRM = 1 , IFM = 0 and CRM 1 , 2

Note : Consecutive 32 block-lines from the first through the 32nd block-lines or from the 33rd through the 64th block-lines are omitted for ordinary field repetition. Other methods are also possible using the FRP command controlled in unit of block-line.