INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

J.88

(09/99)

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

SERIES J: TRANSMISSION OF TELEVISION, SOUND PROGRAMME AND OTHER MULTIMEDIA SIGNALS

Digital transmission of television signals

Transmission of enhanced definition television signals over digital links

ITU-T Recommendation J.88

(Previously CCITT Recommendation)

ITU-T J-SERIES RECOMMENDATIONS

TRANSMISSION OF TELEVISION, SOUND PROGRAMME AND OTHER MULTIMEDIA SIGNALS

For further details, please refer to ITU-T List of Recommendations.

ITU-T RECOMMENDATION J.88

TRANSMISSION OF ENHANCED DEFINITION TELEVISION SIGNALS OVER DIGITAL LINKS

Summary

This Recommendation considers a bit-rate reduction scheme to be applied for the digital transmission of EDTV-II composite signals compatible with NTSC television signals [1].

Digital transmission thereof involves direct coding of the composite signals.

Source

ITU-T Recommendation J.88 was prepared by ITU-T Study Group 9 (1997-2000) and was approved under the WTSC Resolution No. 1 procedure on 16 September 1999.

FOREWORD

ITU (International Telecommunication Union) is the United Nations Specialized Agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of the ITU. The ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Conference (WTSC), which meets every four years, establishes the topics for study by the ITU-T Study Groups which, in their turn, produce Recommendations on these topics.

The approval of Recommendations by the Members of the ITU-T is covered by the procedure laid down in WTSC Resolution No. 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

INTELLECTUAL PROPERTY RIGHTS

The ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. The ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, the ITU had received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementors are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database.

CONTENTS

TRANSMISSION OF ENHANCED DEFINITION TELEVISION SIGNALS OVER DIGITAL LINKS

(Geneva, 1999)

1 Scope

The introduction of an enhanced television system, namely EDTV-II (compatible with the NTSC system), has led to the need to carry this signal over digital links.

Due to the high cost expected for the enhanced television encoders, a specific coding system has been developed to allow transmission of the EDTV-II signals foreseen for primary distribution and also for contribution where no further post-processing is needed.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; all users of this Recomendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published.

[1] Recommendation ITU-R BT.1298 (1997), *Enhanced wide-screen NTSC TV transmission system*.

[2] ITU-T Recommendation H.222.0 (1995) | ISO/IEC 13818-1:1996, *Information Technology – Generic coding of moving pictures and associated audio informations: Systems*.

3 Definitions

This Recommendation defines the following terms:

3.1 adaptive quantizer: Quantizer step size is controlled by a slice type, buffer occupancy and HVS (human visual system).

3.2 adaptive scanning: Transmission orders of WHT coefficients. Four types of scanning are used to transmit composite TV signals efficiently by a limited number of bits.

3.3 bit rate: The rate at which the compressed bit stream is delivered from the channel to the input of a decoder.

3.4 block: A unit for direct execution of WHT on the composite EDTV-II signal under the 8 pixels × 8 lines size.

3.5 coding mode: Mix A, Mix B and refresh (all intra) mode.

3.6 colour sub-carrier phase shift: Phase shift of 3.58 MHz colour sub-carrier of the coding block from the motion compensated reference block in the previous frame.

3.7 composite motion compensation: Motion Compensation (MC) is conducted on WHT domain not on pixel domain to solve colour sub-carrier phase shift.

3.8 criticality: The reference for the difficulty of the picture judged in macro block units.

3.9 EDTV-II sampling: 14.3 MHz with 8-bit precision.

3.10 EDTV-II signals: NTSC-compatible components transmitted in the centre part of the picture and helper components located in the upper and lower parts of the picture.

3.11 Fukinuki hole: Frequency regions around the colour sub-carrier of the compatible centre part where normal NTSC signals have less spectrum density.

3.12 **helper signals**: Three types of spatio and temporal enhancement signals: Horizontal High frequency helper signal (HH), Vertical High frequency helper signal (VH) and Vertical Temporal helper signal (VT).

3.13 **macro block**: A size of 16 pixels × 16 lines and is composed of four 8 × 8 Walsh-Hadamard Transform.

3.14 **mix mode**: Mixing Inter and Intra mode in the same macro block to enhance coding efficiency.

3.15 **PES packet**: The data structure used to carry elementary stream data. A PES packet consists of a PES packet header followed by a number of continuous bytes from an elementary data stream. It is a layer in the system coding syntax described in 2.4.3.6 of ITU-T Rec. H.222.0 | ISO/IEC 13818-1 [2].

3.16 **picture layer**: A repeated structure composed of a head I-picture and a plurality of subsequent p-pictures.

3.17 **sequence layer**: Uppermost layer of a coding bit stream by which coding and decoding parameters are coordinated.

3.18 **slice layer**: Composed of 768 pixels × 16 lines active pixels, and comprises a horizontal arrangement of macro blocks. The slice is categorized into four types of ID and control signal area, boundary area, block area and active image area.

3.19 **two-dimensional VLC**: Huffman code having the combination of zero-run length and quantization output level as a symbol.

4 Abbreviations

This Recommendation uses the following abbreviations:

Bp Maximum Buffer Capacity

BUFP Buffer Pointer

CF Colour Frame

EDTV Enhanced Definition Television

FSW Video Frame Synchronizing Word

GOP Group of Pictures

HH Horizontal High frequency component

NTSC National Television System Committee

PC Phase Compensation

SSW Sequence Synchronizing Word

VH Vertical High frequency component

VITS Vertical Interval Test Signal

VLC Variable Length Coding

VLD Variable Length Decoding

VT Vertical Temporal frequency component

WHT Walsh-Hadamard Transform

5 Transmission of enhanced definition television signals over digital links

5.1 Introduction

This Recommendation considers a bit-reduction coding scheme for application to digital transmission of EDTV-II composite signals defined for NTSC-compatible enhanced television. The bit-reduction coding system described below is a composite coding system that does not require any decoding/re-encoding process for enhanced signals and colour component signals.

Consequently the system is free from picture quality loss brought about by composite-component conversion required in component coding schemes. The bit rate required for contribution and distribution purposes is approximately 20 Mbit/s.

5.2 System description

The EDTV-II signals consist of NTSC-compatible components transmitted in the centre part of the picture and the helper components located in the upper and lower parts of the picture. Three kinds of helper signals are employed: a horizontal high frequency component (HH) multiplexed in frequency regions around the colour sub-carrier (so-called Fukinuki hole) of the compatible centre part where normal NTSC signals have less spectrum density, the vertical high frequency component (VH), and the vertical temporal helper signal (VT) compressed in the spatial domain and multiplexed in the upper and lower parts of the picture. Thus, EDTV-II composite signals have a complex structure.

Although component bit-reduction coding can be applied to EDTV-II signals, it requires a decoding/re-encoding process for the colour and HH components as well as the VH/VT components at every stage of transmission chain. This decoding/re-encoding process inevitably causes severe damage to picture quality, especially in the case of tandem connections. This degradation is caused by the irreversible nature of the decoding/re-encoding process.

Hardware for the composite coding scheme was implemented based on an interframe motion-compensated Walsh-Hadamard Transform coding scheme. A preliminary picture quality evaluation test was conducted with the hardware and it showed that the system provides picture quality adequate for the contribution and primary distribution of the EDTV-II programmes at a transmission bit rate of approximately 20 Mbit/s. This evaluation test also showed that the system provides coding performance equivalent or superior to the MPEG-2 codec for NTSC composite signals at a transmission bit rate above 15 Mbit/s.

The above considerations and the test results mentioned show that the composite coding system is well suited for transmission of signals such as EDTV-II and NTSC when relatively high quality is required.

Detailed specification for a composite coding system suitable for digital transmission of EDTV-II as well as NTSC signals is given in Annex A.

Annex A

Specification of direct composite coding for EDTV-II and NTSC signals

A.1 Structure of coding bit stream

The coding bit stream is converted to the PES packet in order to be operable on ITU-T Rec. H.222.0 | ISO/IEC 13818-1 (MPEG systems) [2]. Therefore, the bit stream is arranged in a three-layer structure composed of a sequence layer, a picture layer, and a slice layer. The unit of the PES packet is basically a slice in view of the maximum code generation. The outline of the bit coding stream structure is shown in Figure A.1-1.

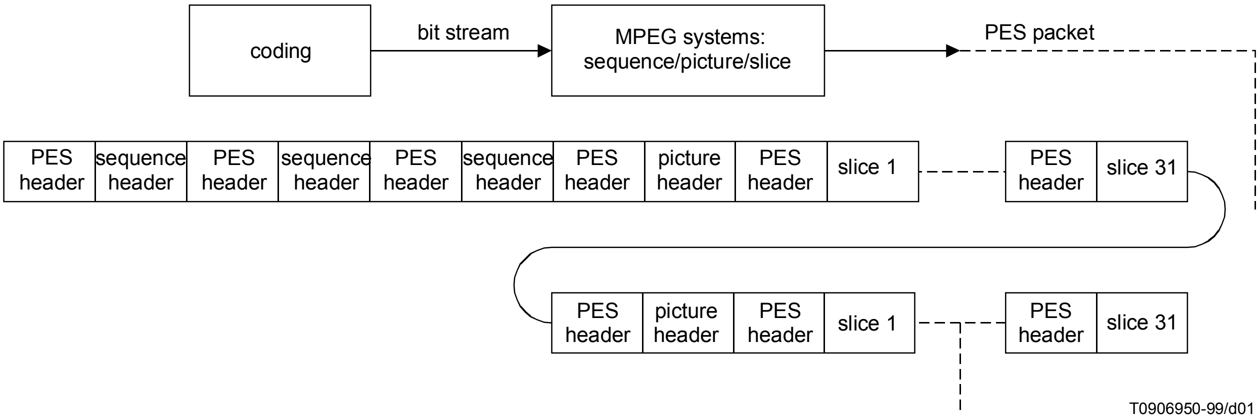
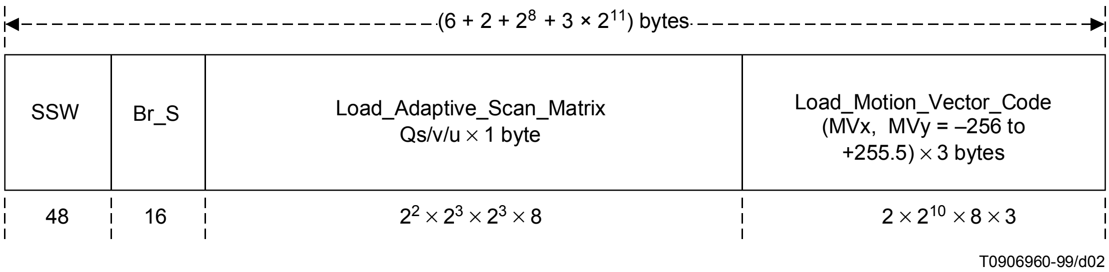


Figure A.1-1/J.88 – Outline of coding bit stream structure

A.1.1 Sequence layer

The sequence layer is the uppermost layer of a coding bit stream. The sequence layer has the purpose of coordinating coding and decoding parameters. In detail, the sequence header transmits required parameters from the coding side to the decoding side with the desired timing in order to coordinate automatically coding and decoding for channel hopping, switching of systems, and problems of systems. A sequence header may be interpolated as desired at any position as long as the position is a blank between pictures shown by the picture header in a picture layer as described later. There are three types of sequence headers from 1 to 3 as described below.

A.1.1.1 Sequence header 1



1) *SSW: Sequence head word*

SSW is a sequence synchronizing word. The pattern is composed of one bit "1" and 47 bits "0". SSW is transmitted once, at the timing when the receiving side becomes receivable, establishes the sequence synchronization and transmits parameters.

2) *Br_S: Bit rate (sequence)*

Under a premise of the change of parameters, bit rate change is executed. The multiple number of 90 kbit/s is represented by 90 × Br_S kbit/s (0 to 5898.15 Mbit/s).

3) *Load_Adaptive_Scan_Matrix*

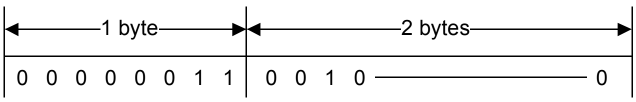
Four types of Adaptive_Scan parameters (represented by Qs 2 bit which will be described in A.2.6 and A.3.2 later) are transmitted.

The scanning pattern is transmitted in the order of Qs = 00, 01, 10, 11, and the scanning order of each coefficient for raster-scanning of 8 × 8 Hadamard transformation. As a result, the capacity of parameters is 8 bits × 64 coefficients × 4 patterns = 2^{11} bits = 2^8 bytes.

4) *Load_Motion_Vector_Code*

A coding pattern (Huffman code) which represents a horizontal motion vector (MVx) and vertical motion vector (MVy) is transmitted. MVx and MVy are composed of half-pixel/half-frame line units, and are retrievable up to the maximum –128.0 to +127.5 pixel/frame line. Therefore, the motion vector requires 1024 coding patterns of –256 to 255.5 pixel/frame line to transmit the differential value. This coding pattern is transmitted in the form as described below composed of 1-byte code length and subsequent 2-byte coding pattern in the order MVx = –256.0 ≥ 0 ≥ 255.5, MVy = –256.0 ≥ 0 ≥ 255.5 as Huffman code with the maximum code length of 16 bits or less. It is assumed that bits other than coding pattern is "0", and bits which have no coding pattern for defining the searching range is also "0".

Case that the coding pattern is "001" (code length of 3 bits).



T0906970-99/d03

A.1.1.2 Sequence header 2

Load_Quantization_Matrix
Bs/M/Es/N/v/u × 1 byte

$2^5 \times 2^1 \times 2^2 \times 2^2 \times 2^3 \times 2^3 \times 8$

This sequence header 2 is data of fixed length 216 bytes subsequent to the sequence header 1. Load_Quantization_Matrix transmits a quantizing parameter described in A.2.6 and A.3.3. The transmission address order of transmission of the quantizing parameter contained in a quantizing parameter table g1 (16-bit input; equivalent to 1M) is represented by 1 byte of Delta (Bs, M, Es, N, v, u) under Bs/M/Es/N/v/u/ = 00 to FF. Therefore, the PES packet length is the maximum length.

A.1.1.3 Sequence header 3

Load_Mixed_Mode_Matrix
Bs/M/Es/N/v/u × 1 byte

$2^5 \times 2^1 \times 2^2 \times 2^2 \times 2^3 \times 2^3 \times 8$

This sequence header 3 is data of fixed length 216 subsequent to the sequence header 2. The Load_Mixed_Mode_Matrix transmits parameters for mixed mode coding described in A.2.5 and A.3.5. The transmission address order of transmission of the mixed mode parameter contained in a quantizing parameter table h1 (16-bit input; equivalent to 1M) is represented by 1 byte of A (Bs, M, Es, N, v, u) under Bs/M/Es/N/v/u/ = 00 to FF. Therefore, the PES packet is the maximum length.

A.1.2 Picture layer

The picture has a repeated structure composed of a head I-picture and a plurality of subsequent p-pictures. The head I-picture is directed by a refresh information (R). This R is included in the picture header shown in Figure A.1-2. The picture header constitutes one PES packet with one picture header.

(PES-packet-length = 000 000 0001 0010)

A.1.2.1 Picture header

The structure of a picture header is shown in Figure A.1-2, and the components are described in 1) to 8).

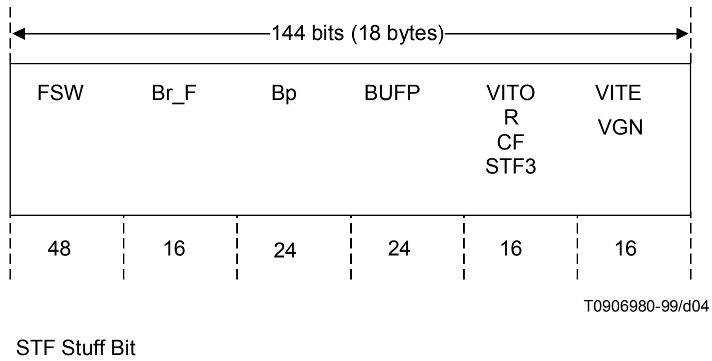


Figure A.1-2/J.88 – Structure of picture header

1) *FSW: Video frame head word*

   FSW is a frame synchronizing word, and composed of 47 bits "0" and subsequent 1 bit "1". FSW is transmitted once, and at the timing when the receiving side becomes receivable, establishes the sequence synchronization and transmits parameters. Once synchronization is established and FSW has not been detected, it is regarded as a step out from the frame synchronization.

2) *Br_F: Bit rate (frame)*

   Parameters are not changed, and bit rate is changed without any breaks of picture. The multiple number of 90 kbit/s is represented by 90 × Br_F kbit/s (0 to 5898.15 Mbit/s).

3) *Bp: Maximum buffer capacity*

   Transmission is implemented once at the head of each frame. Only when different patterns are received is the maximum buffer capacity change regarded as effective, and the change is implemented on the decoder side.

   This is the parameter which represents the capacity of the buffer memory. One unit has the weight of 32 bits, maximum 32×2^{24} = 512 Mbits.

4) *BUFP: Buffer pointer*

   The buffer pointer has 32-bit units and represents buffer occupancy quantity for each picture. It is possible to represent up to the maximum 32×2^{24} = 512 Mbits.

5) *VITO/VITE: VITS Position (ODD/EVEN)*

   This shows whether the odd lines 10 to 20 (11 lines) and even lines 273 to 282 (10 lines) are selected by VITO-11 bits and VITE-10 bits as VITS line. For example, when the third bit of VITO is "1", 12-th line is selected as VITS line, and this line is addressed for transmission.

6) *R: Refresh information*

   The existence of I-picture at the head of GOP is directed to a decoder with transmission of 1 bit. "1" represents the refresh mode and "0" represents the adaptive coding mode.

7) *CF: Colour frame*

   Whether the picture has 1,2 field or 3,4 field is directed to a decoder with one transmission of 1 bit.

8) *VGN: Picture gain information*

   The input signal is subjected to a signal compression in the range of 1/2 to 1 with 6 bits (1/64) precision.

A.1.2.2 EDTV-II composite picture format

Sampling frequency must be 4 times of colour sub-carrier (3.58 MHz). To obtain active pixel area to be coded out of these sampled data, the vertical blanking period signal (Figure A.1-3) and horizontal blanking period signal (Figure A.1-4) are removed.

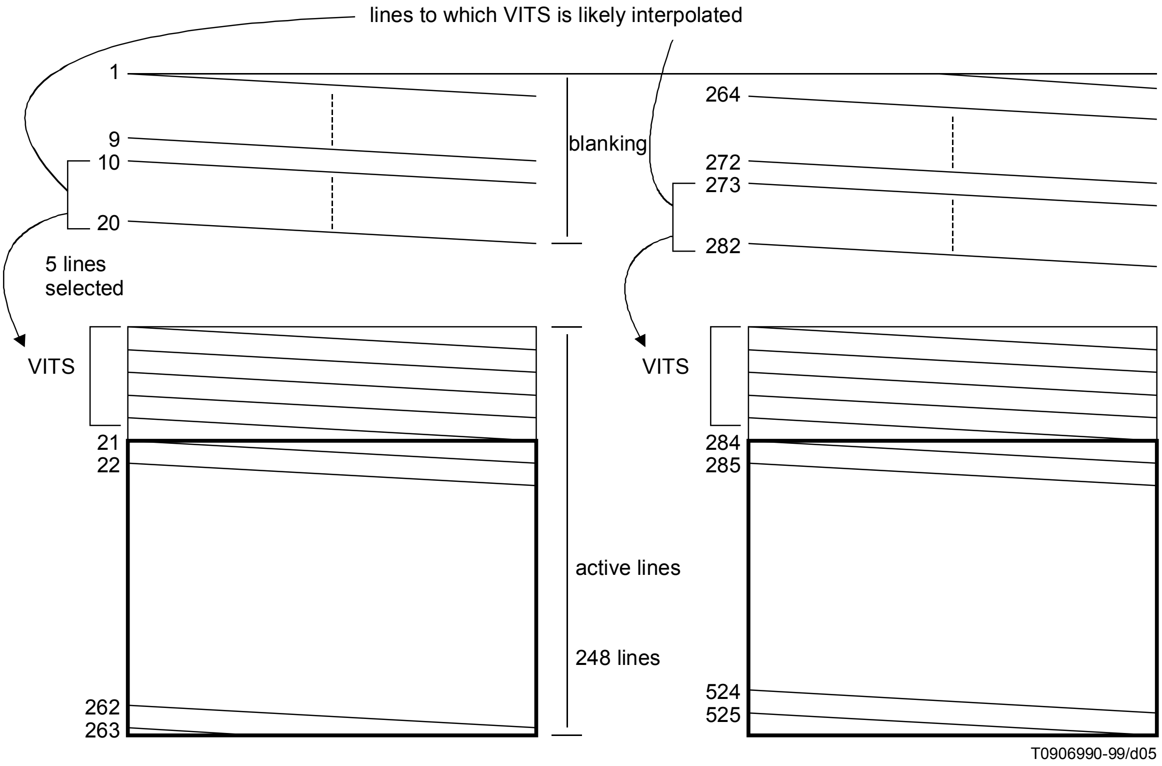


Figure A.1-3/J.88 – Removal of vertical blanking period signal

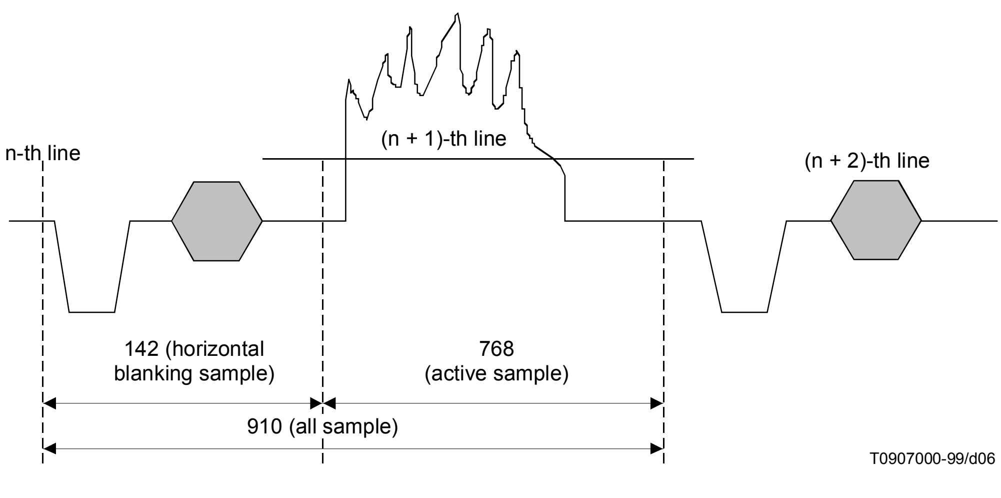


Figure A.1-4/J.88 – Removal of horizontal blanking period signal

A.1.3 Slice layer

A.1.3.1 Structure of slice

A slice is composed of 768 pixels × 16 lines active pixels, and comprises a horizontal arrangement of macroblocks having a size of 16 × 16 lines which is the transmission unit of various control signals and motion vectors. Therefore, 1 slice is composed of 48 macroblocks taken in the horizontal direction from the active 768 × 496 pixels of the picture after having removed the vertical and horizontal blanking signals as described above, and 1 picture is composed of 31 slices.

The relationship between the slice and line in the EDTV-II signal is shown in Figure A.1-5, and the slice is categorized into four types (Es) of ID and control signal area, boundary area, black area, and active image area.

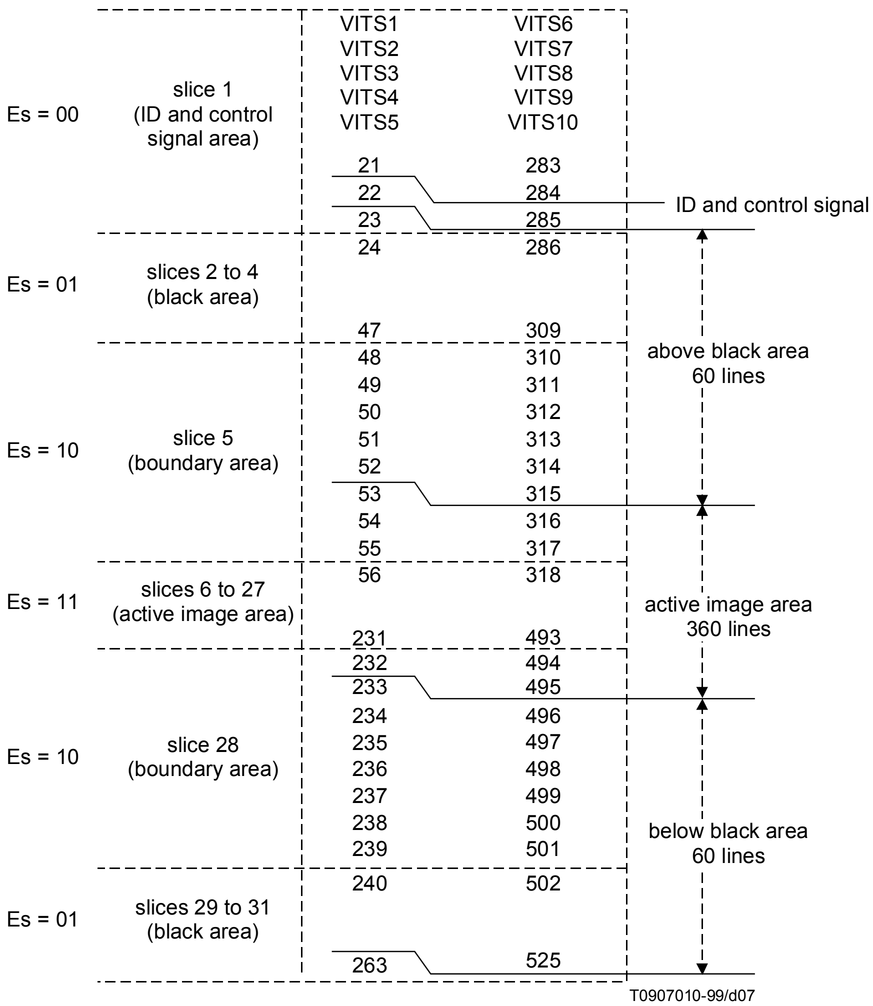


Figure A.1-5/J.88 – Four types of slices in EDTV-II signal

The structure of each slice is shown in Figure A.1-6 and described in 1) to 7) below:

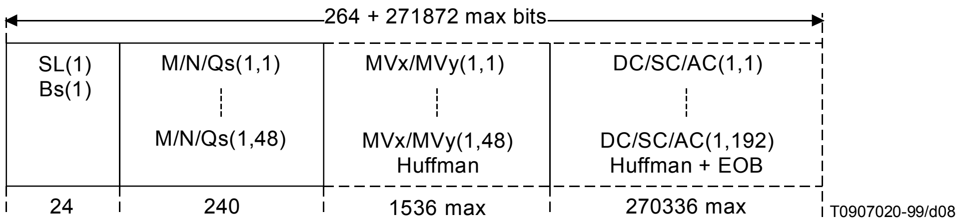


Figure A.1-6/J.88 – Internal structure of a slice

1) *SL(n)*

SL(n) positioned at the head of the n-th slice represents the quantity of information to be generated from the slice. This SL is represented using 19 bits. Because the maximum information quantity generated from variable length code of 1 slice is 272 136 bits < 2^{19} bits = 2^{16} bytes, which is the total of the assumable longest codes of all coefficients in a slice (48 × 64 × 4) and motion vectors (48), namely 22 bits and 32 bits respectively (described in A.2.7 and A.3.1 respectively), SL can express it and it can be contained in a PES packet. A counter decodes up to the bit length indicated by SL(n) and, when finished, starts to decode the next slice.

2) *Bs(n): Buffer information*

Buffer occupancy quantity for coding the n-th slice is represented using 5 bits (32 steps).

3) *M(n, m): Coding mode selection information*

The coding mode of the m-th macroblock (m = 1 to 48) of the n-th slice is represented using 1 bit. "0" is for inter (mixed) mode and "1" is for intra mode. The mode decision is described in A.2.5.

4) *N(n, m): Criticality information*

The activity of the m-th macroblock (m = 1 to 48) of the n-th slice is represented using 2 bits (four types). "0" is for flat area, "1" is for gradient area, "2" is for edge area, and "3" is for detail area. The determination of the criticality information will be described later in A.2.4.

5) *Qs(n, m): Scanning selection information*

The scanning selection for transmission order of Hadamard transform coefficient of the m-th macroblock (m = 1 to 48) of the n-th slice is represented using 2 bits (four types). The determination of the scanning selection information is described in A.3.6.

6) *MV(n, m): Motion vector information*

The motion vector information of the m-th macroblock of the n-th slice is represented by Huffman code. The motion vector detection is described in A.2.3.

7) *DC/SC/AC(n, l): Transform coefficient data*

The transform coefficient data is two-dimensionally Huffman-coded DS/SC/AC component of the 1st block (l = 1 to 192) of the n-th slice (described later).

In the above, slice data of 1) to 5) are in the form of fixed length, but 6) MV and 7) DC/SC/AC are in the form of variable length data. 2) Bs is control information with slice unit, and 3) to 6) are control information with macroblock unit.

A.1.3.2 Macroblock

The macroblock has a size of 16 pixels × 16 lines always, and is composed of four 8 × 8 WHT (Hadamard transform). The counting of the macroblocks in a slice proceeds from left to right, namely raster scanning; the counting of blocks in a macroblock is carried out as shown in Figure A.1-7.

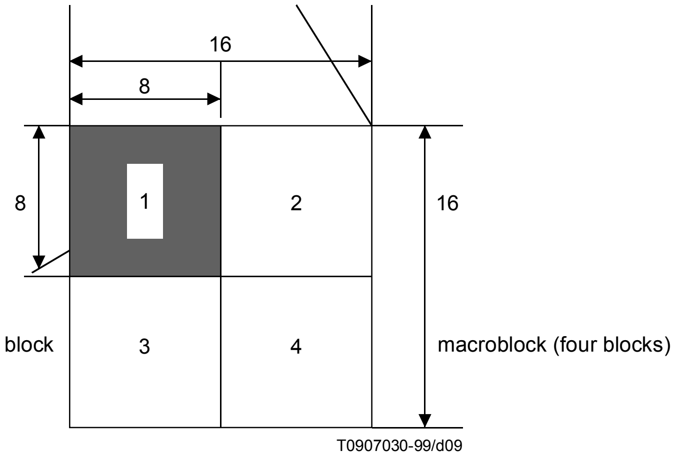


Figure A.1-7/J.88 – Structure of macroblock

Because the macroblock is addressed only on frame-based Hadamard transform coding, the macroblock is structured as shown in Figure A.1-8.

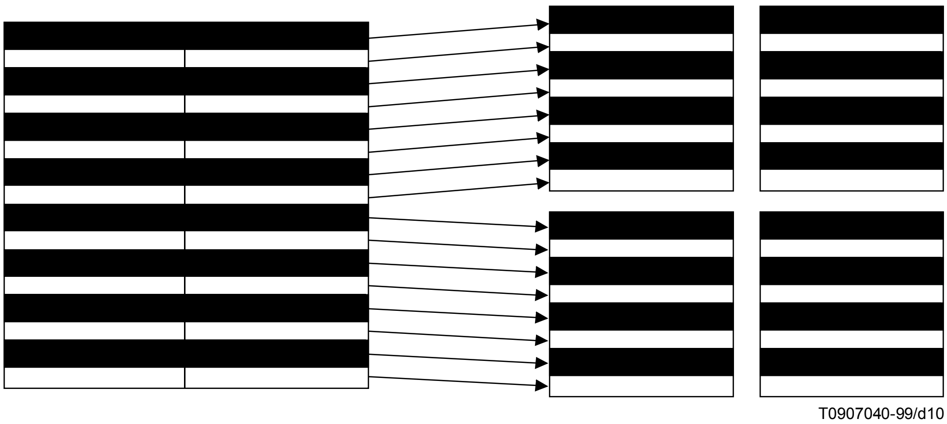


Figure A.1-8/J.88 – The relationship between frame and macroblock

A.1.3.3 Block

The block is a unit for direct execution of WHT on the composite EDTV-II signal under the 8 pixels × 8 lines size. It is assumed that Hadamard transform coding is performed under the frame base generated from field merging (Figure A.1-9).

A.2 Coding process

The whole structure of the coding process is shown in Figure A.2-1.

EDTV-II signal must be sampled at 14.3 MHz and with 8-bit precision. The coding process of this data involves basically the removal of time space redundant component by inter/intra composite motion compensated adaptive Hadamard transform coding. Elements associated with the coding process are described in the following.

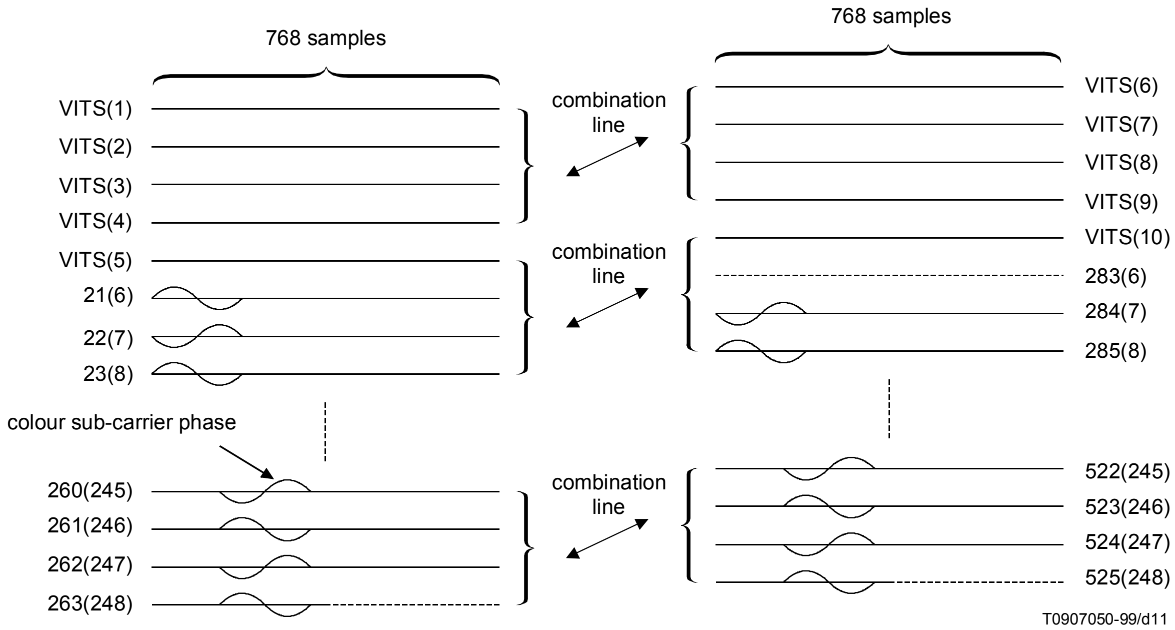


Figure A.1-9/J.88 – Frame-based WHT coding based on field merge

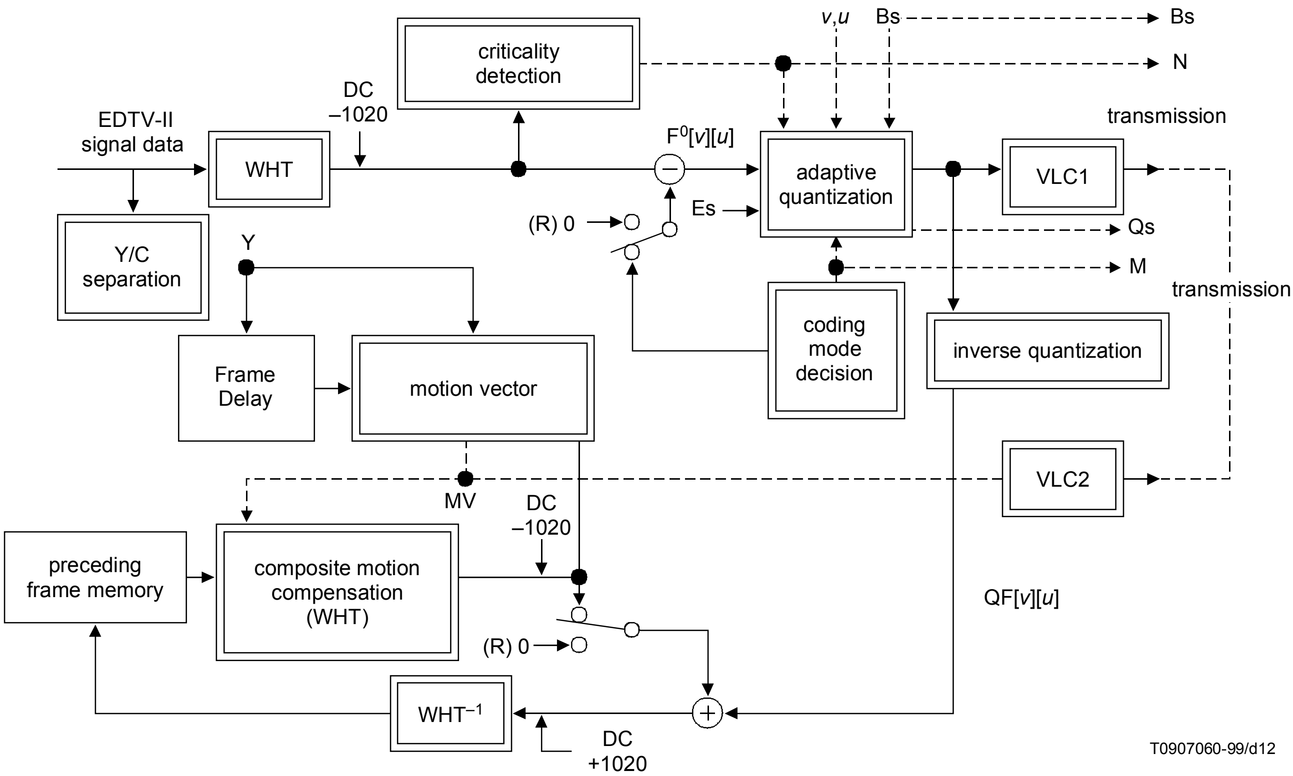


Figure A.2-1/J.88 – Whole structure of coding process

A.2.1 Intra-frame Hadamard transform (WHT)

A block of 8 (pixels) × 8 (lines) shown in the figure is formed by combining alternately 4 lines (4n + 1), (4n + 2), (4n + 3), and (4n + 4) in an odd field and 4 lines (4n + 1), (4n + 2), (4n + 3), and (4n + 4) in an even field with 768 (pixels) × 496 (lines) data inputted as the coding active pixel (refer to A.1.2.2).

$$[V](n, m) = \begin{bmatrix} A(4n+1, 8m+1) \; A(4n+1, 8m+2) \; ... \; A(4n+1, 8m+8) \\ B(4n+1, 8m+1) \; B(4n+1, 8m+2) \; ... \; B(4n+1, 8m+8) \\ A(4n+2, 8m+1) \; A(4n+2, 8m+2) \; ... \; A(4n+2, 8m+8) \\ B(4n+2, 8m+1) \; B(4n+2, 8m+2) \; ... \; B(4n+2, 8m+8) \\ A(4n+3, 8m+1) \; A(4n+3, 8m+2) \; ... \; A(4n+3, 8m+8) \\ B(4n+3, 8m+1) \; B(4n+3, 8m+2) \; ... \; B(4n+3, 8m+8) \\ A(4n+4, 8m+1) \; A(4n+4, 8m+2) \; ... \; A(4n+4, 8m+8) \\ B(4n+4, 8m+1) \; B(4n+4, 8m+2) \; ... \; B(4n+4, 8m+8) \end{bmatrix}$$

In the equation, A(j, i) represents the value of the i-th pixel on the j-th line in an odd field, and B(j, i) represents the value of the i-th pixel on the j-th line in an even field, m is 0 to 95, and n is 0 to 61.

The block [V] of 8 (pixels) × 8 (lines) obtained by the block forming is subjected to Walsh-Hadamard transform.

A.2.2 Y/C separation

The input composite signal is subjected first to Y/C separation in the field base. It is recommended to use a function for pre-filtering independently Y signal and C signal in field base. The filtering function is shown in Figure A.2-2. Y signal before filtering is used for motion detection.

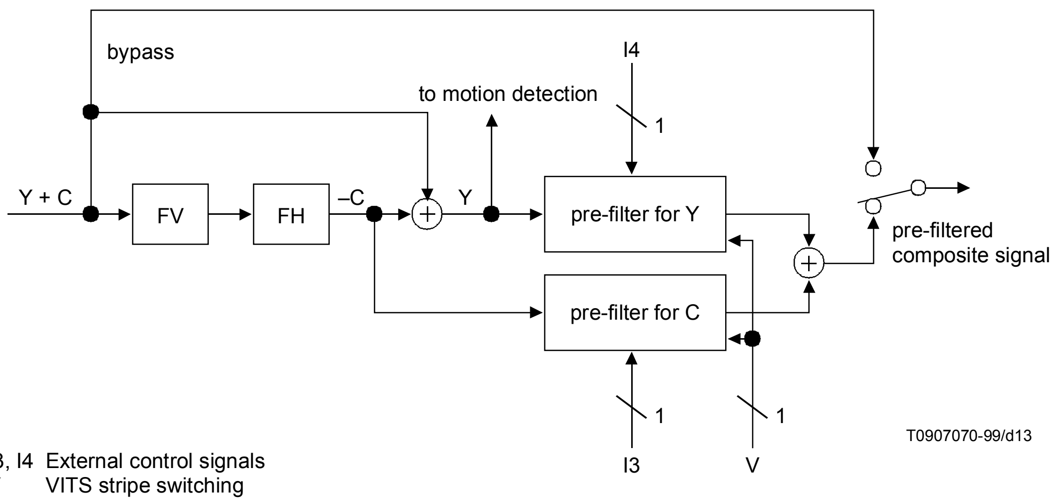


Figure A.2-2/J.88 – Structure of Y/C separation type pre-filter

Filter coefficients in the following are recommended:

– Vertical filter (FV):

(–0.25 0.0 0.5 0.0 –0.25)

– Horizontal filter (FH):

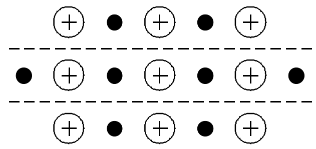
(0.03125 0.0 –0.0625 0.0 0.09375 0.0 –0.125

0.0 0.125 0.0 $\boxed{-0.125}$ 0.0 0.125 0.0

–0.125 0.0 0.09375 0.0 –0.0625 0.0 0.03125)

pre-filter for Y (two-dimensional)

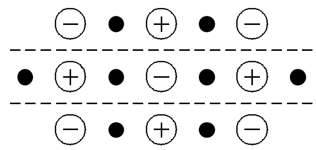
$$\frac{1}{(KY + 2)^2} \begin{bmatrix} 1 & KY & 1 \\ KY & KY^2 & KY \\ 1 & KY & 1 \end{bmatrix}$$



T0907350-99/d14

pre-filter for C (two-dimensional)

$$\frac{1}{-(KC - 2)^2} \begin{bmatrix} -1 & KC & -1 \\ KC & -KC^2 & KC \\ -1 & KC & -1 \end{bmatrix}$$



T0907080-99/d15

The value of KY and KC is variable in a range from 1 to 64. It is desirable that KY and KC are independently variable on the pre-filter coefficient table and controlled by the bit rate Br and buffer information Bs. KY and KC are turned ON/OFF in response to the external control signals I3/I4. When the pre-filter is OFF, the input is bypassed.

A.2.3 Motion vector detection

The motion detection is operated in macroblock unit by frame base on the original signal (before prefilter) of Y obtained by Y/C separation and Y signal of the original picture of the preceding frame. The range of the motion detection covers at most horizontal –128.0 to +127.5 pixels and vertical –128.0 to +127.5 frame lines. The precision is a half pixel for horizontal direction and 1/2 frame line for vertical direction. As for a method for forming the half-pixel and 1/2 frame line, the half-pixel on field base and 1/4 field line interpolation are recommended as described in Figure A.2-3.

The recommended filter coefficients for generating the decimal pixel/line is shown in Figure A.2-4. The right side and lower side of the motion vector (MVx, MVy) are referred to as MVx and MVy respectively in the view from coding block (current frame) to reference block (preceding block).

A.2.4 Criticality detection

The criticality which is the reference for the difficulty of the picture is judged in macroblock units. The absolute value of 64 transform coefficients are compared with a certain threshold value on Hadamard transform, and values larger than that threshold value are counted as significant coefficient, and the total significant coefficients are classified into 4 criticalities.

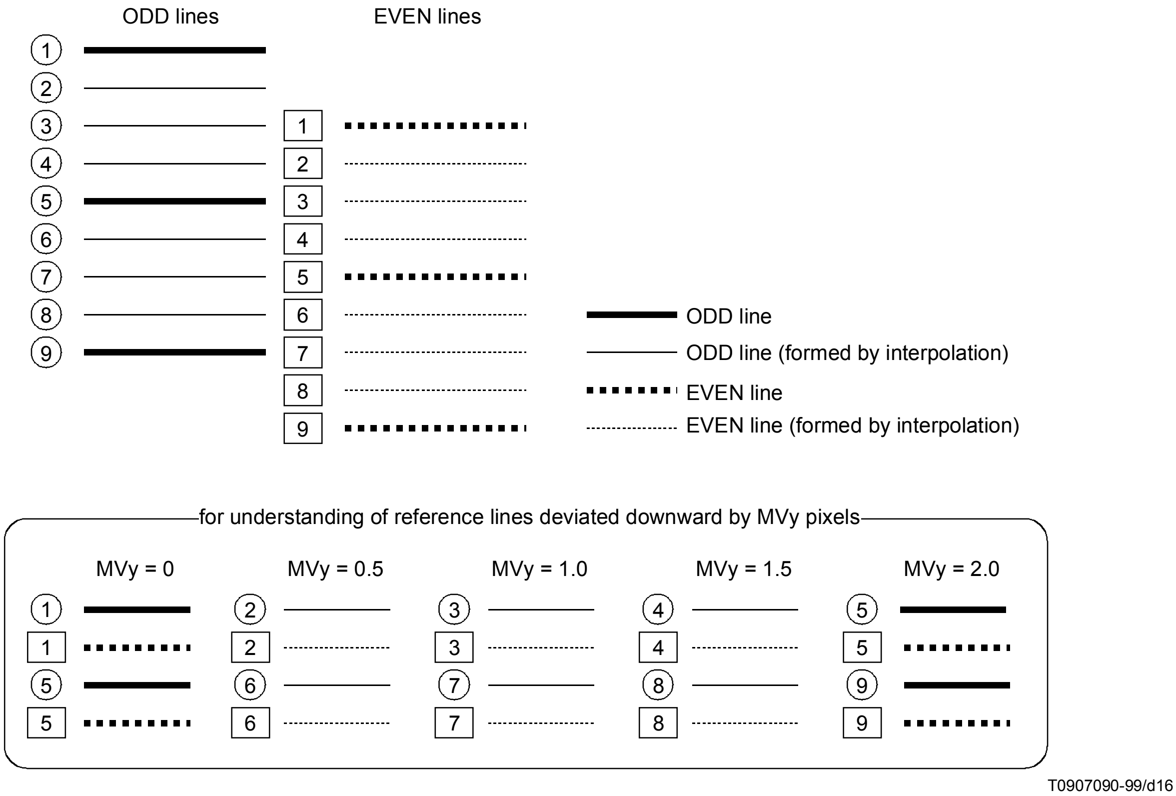


Figure A.2-3/J.88 – Half-pixel, 1/4 field line interpolation

The threshold values in the following are recommended.

The threshold value for comparison of coefficients: 7

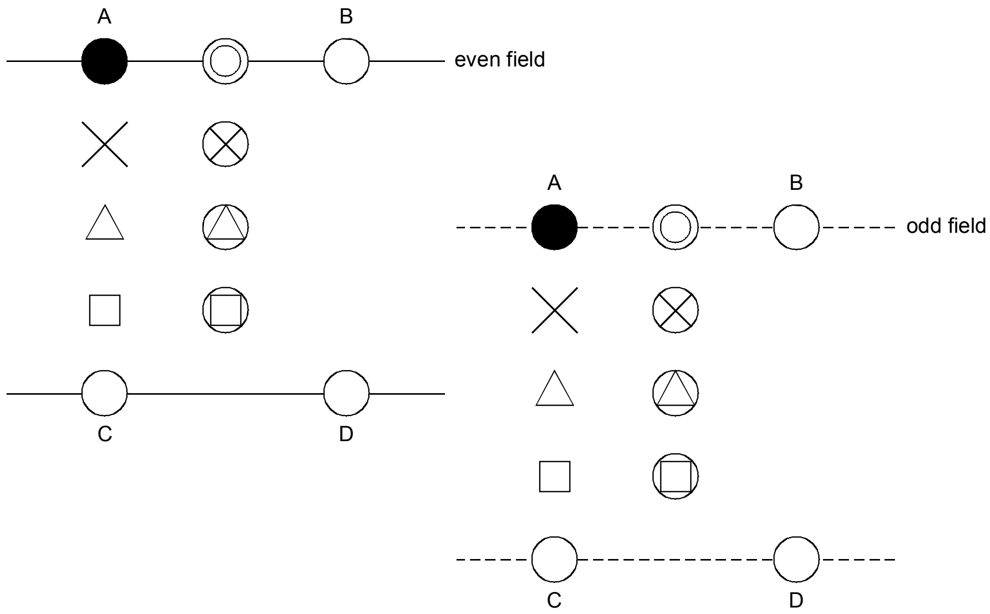
- N = 00 (flat area): 0 to 20
- N = 01 (gradient area): 21 to 50
- N = 10 (edge area): 51 to 85
- N = 11 (detail area): 86 or subsequent values

A.2.5 Coding mode decision section

Two types of mode are provided as adaptive coding modes. If there is a refresh indication signal (R), the coding mode is refresh mode = all intra mode. If there is no refresh indication signal (R), whether it is Mix A mode or Mix B mode is judged according to the procedure shown in Figure A.2-5. The prediction frame signal is inputted to both prediction value conversion tables A and B, and by calculating the difference from the current frame signal, any one mode is selected.

The prediction value conversion of each mode is shown in Figure A.2-6 (refer to A.3.6).

M = 0 is given to Mix A mode and M = 1 is given to Mix B mode. As for the content of the table h1, the sequence header 3 is used for decoding process. The function h2 is used for multiplying A[*v*][*u*] by prediction frame signal, wherein *v* represents the vertical coefficient position (0 to 7), and *u* represents the horizontal coefficient position (0 to 7).



	A	B	C	D
●	1	0	0	0
×	3/4	0	1/4	0
△	1/2	0	1/2	0
□	1/4	0	3/4	0
⊗	5/16	5/16	3/16	3/16
△⃝	1/4	1/4	1/4	1/4
□⃝	3/16	3/16	5/16	5/16
◎	1/2	1/2	0	0

T0907100-99/d17

Figure A.2-4/J.88 – Motion compensation of the decimal pixel/line precision

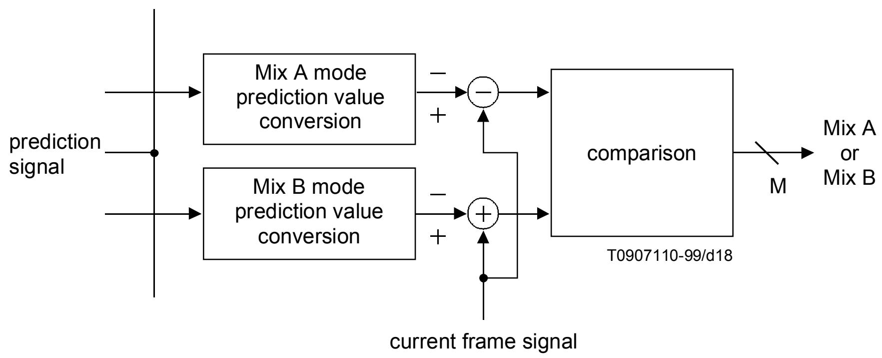


Figure A.2-5/J.88 – Mix mode decision

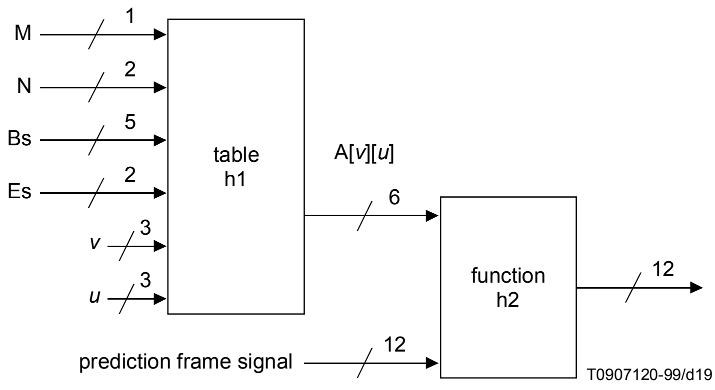


Figure A.2-6/J.88 – Structure of prediction value conversion

A.2.6 Adaptive quantization/adaptive scanning

The structure of adaptive quantization and adaptive scanning is shown in Figure A.2-7 including an inverse operation.

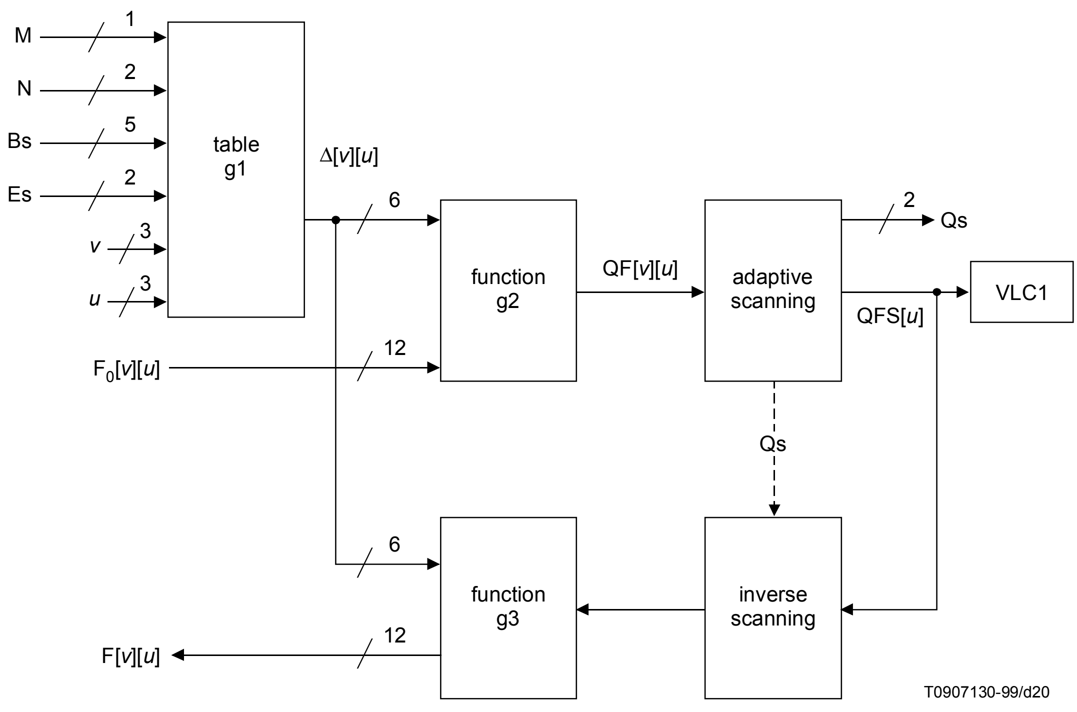


Figure A.2-7/J.88 – Adaptive quantization/adaptive scanning

In the adaptive quantization, the quantizer step size Delta[v][u] is determined using the function g1 by the coding mode M (1 bit), criticality N (2 bits), buffer occupancy Bs (5 bits), and slice type Es (2 bits). Subsequently, QF[v][u] is obtained using the function g2.

As for the content of the Table g1, the sequence header 2 is served for decoding process. The function g2 is basically the round-off quotient of F_0[v][u] divided by Delta[v][u].

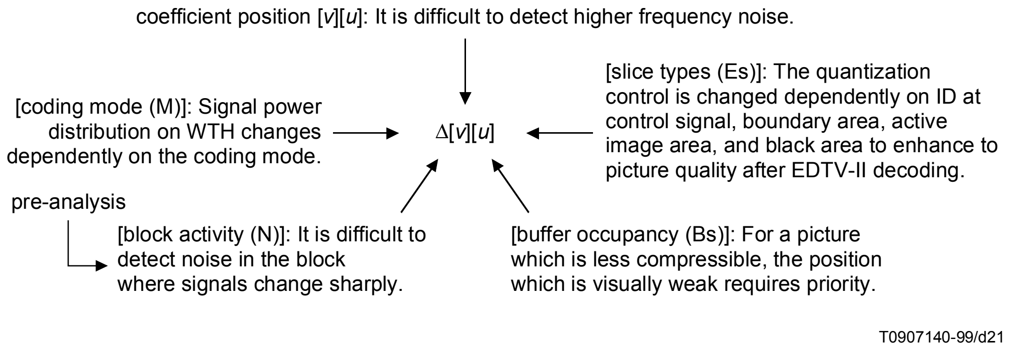


Figure A.2-8/J.88 – Design guideline for adaptive quantization table

Next, QF[*v*][*u*] which is a two-dimensional coefficient array is converted to QFS[*u*] which is a one-dimensional data by selecting the optimal pattern out of 4 types of scanning pattern which are pre-determined. For the selection, a method for selecting the pattern which minimizes the number of coefficients up to EOB (end of block) is recommended. The selected pattern is transmitted in response to the control information of Qs (2 bits).

QFS[*u*] is coded by VLC1 and transmitted, and returned simultaneously to QF[*v*][*u*] by inverse scanning as a local decoding process (refer to A.3.3), and returned then to F[*v*][*u*] by the function g3, and this value is used as a prediction value for the next frame (refer to A.3.4).

A.2.7 Variable length coding

A.2.7.1 Coding (VLC1) of WHT coefficient

For a variable length code corresponding to coded data, a two-dimensional Huffman code, having the combination of zero-run length and quantization output level as a symbol, is used. A Huffman code having a code length of 3 to 14 bits is used for 126 specified symbols in the code assignment.

The code in the form described in the following is assigned to the residual symbols. Code language (22 bits) = escape code (6 bits) + run length (6 bits) + quantization level number (10 bits).

A different combination of a symbol and Huffman code pattern is used for a different mode.

The Huffman code is required to satisfy the following conditions (Figure A.2-9).

1) The quantization level corresponding to 126 specified symbols should be in a range of –32 to +31.

2) Huffman codes having a code length of exceeding 12 bits out of Huffman code patterns corresponding to 126 specified symbols should have a fixed pattern at two low-order bits.

Variable length code pattern (VLC1) for WHT coefficient coding is shown in Appendix V.

A.2.7.2 Coding of motion vector (VLC2)

The motion vector information is transmitted in the form described in the following for each macroblock.

1) In the case of refresh mode:

   Because the motion vector information is not required, it is not transmitted.

2) In the case of the left end macroblock of each stripe:

   The head macroblock of each stripe transmits the motion vector itself instead of differential value.

3) Other cases:

   The differential value from the motion vector of the preceding macroblock is coded and transmitted.

Zero-run length	Quantization level number	Code word
0	1	11s
0	2	0100 s
0	3	0010 1s
0	4	0000 110s
0	5	0010 0110 s
0	6	0010 0001 s
0	7	0000 0010 10s
0	8	0000 0001 1101 s
0	9	0000 0001 1000 s
0	10	0000 0001 0011 s
0	11	0000 0001 0000 s
0	12	0000 0000 1101 0s
0	13	0000 0000 1100 1s
0	14	0000 0000 1100 0s
0	15	0000 0000 1011 1s
1	1	011s
1	2	0001 10s
1	3	0010 0101 s
1	4	0000 0011 00s
1	5	0000 0001 1011 s
1	6	0000 0000 1011 0s
1	7	0000 0000 1011 1s

Quantization level number assigned to one zero-run is restricted to a range of –32 to +31

T0907150-99/d22

NOTE – Lowest order 2 bits of a code word exceeding 12 bits is formed of a fixed pattern.

Figure A.2-9/J.88 – Restriction on Huffman code

A.2.8 Composite motion compensation

A.2.8.1 Concept

The concept of applying the motion compensated interframe coding to NTSC composite signal is shown in Figure A.2-10.

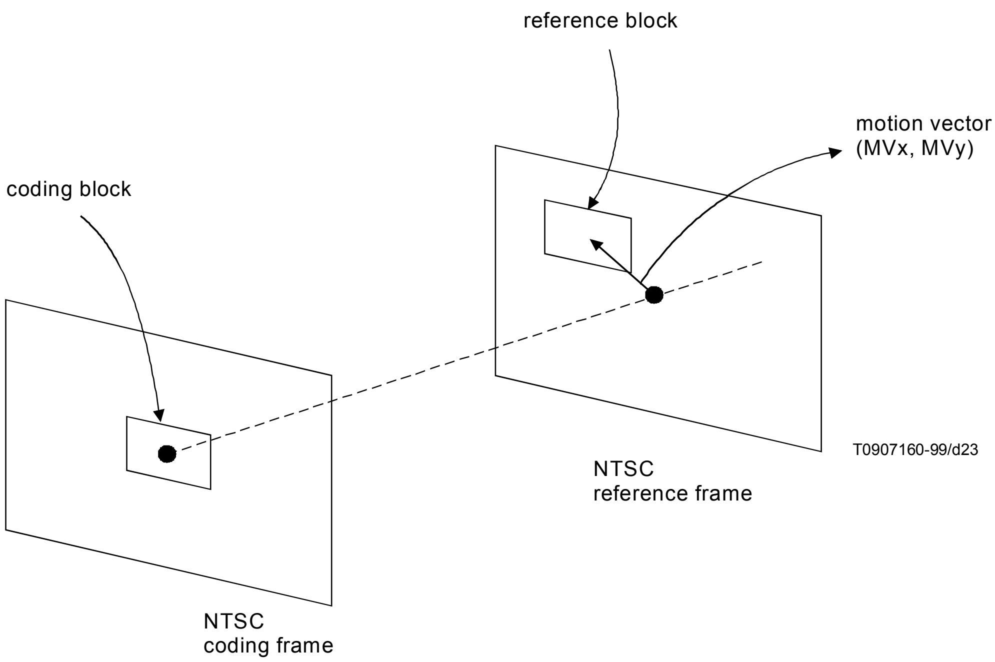


Figure A.2-10/J.88 – Motion compensation interframe coding of NTSC

There is a phase shift of colour sub-carrier dependent on the value of motion vector (MVx, MVy) between the coding block and reference block as shown in Figure A.2-10. Considering that NTSC is a signal, the precision of (MVx, MVy) is prescribed to the precision of horizontal 1 pixel and vertical 1 field line (2 frame lines) under 4 fsc (fsc = 3.58 MHz) sampling. There are 4 types of phase shift – 0 degree (4), 90 degrees (3), 180 degrees (1), and 270 degrees (2) – of colour sub-carrier. The right side is defined as MVx+ and the lower side is defined as MVy+ in view of the direction from the coding frame to the reference frame; the phase shift, depending on the value of (MVx, MVy), is shown in Figure A.2-11.

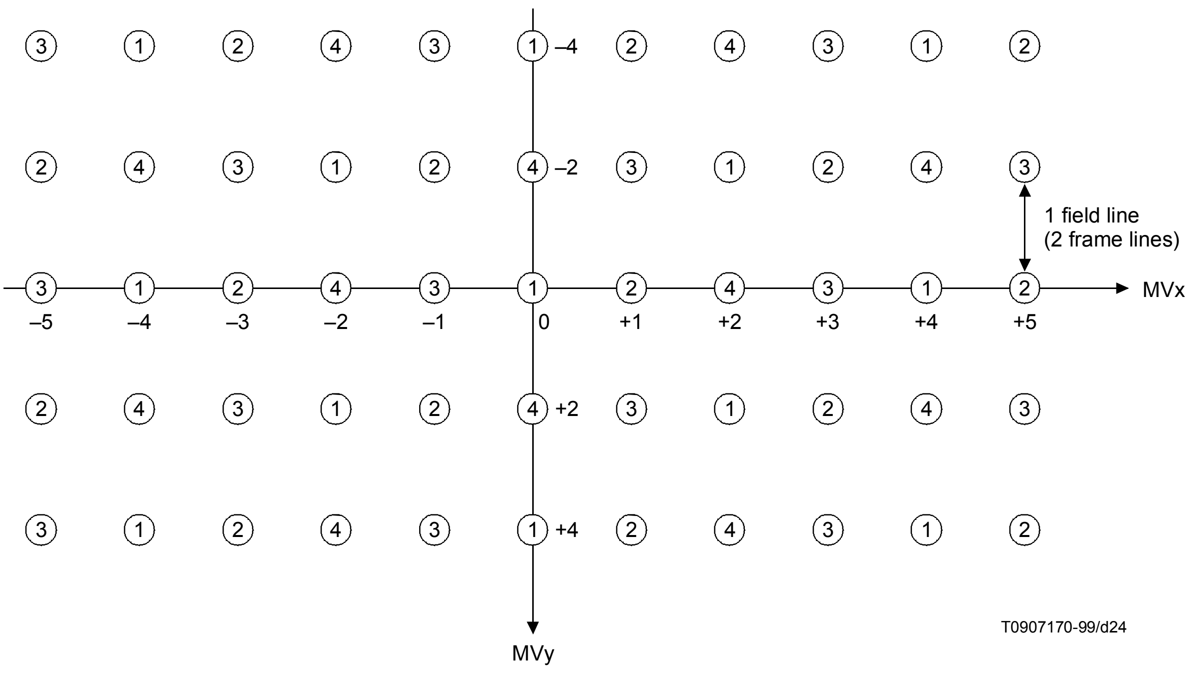


Figure A.2-11/J.88 – Phase shift of colour sub-carrier depending on the motion vector (MVx, MVy)

A.2.8.2 Colour sub-carrier phase shift depending on the motion vector

The relationship between the motion vector (MVx, MVy) (MVx, MVy integer) and phase shift is formulated to form the following equation:

$$\{MVx + MVy\} = \begin{cases} 4n + 2 \text{ (0 degree shift: 4)} \\ 4n \text{ (180 degrees shift: 1)} \\ 4n + 1 \text{ (270 degrees shift: 2)} \\ 4n + 3 \text{ (90 degrees shift: 3)} \end{cases}$$

To compensate the phase shift, the coding block and reference block shown in Figure A.2-10 is 8 × 8 Hadamard transformed.

8 × 8 coding block matrix in the frame base is designated as Sframe and the reference block matrix of the same size is designated as S′frame; the block forming is shown in Figure A.2-12. Next, the colour sub-carrier on the two dimensional Hadamard transformation based on Figure A.2-12 is shown in Figure A.2-13.

A.2.8.3 Overlap WHT for half-pixel precision motion compensation

As it is obvious from Figure A.2-12, phase compensation is performed by simple processing of polarity inversion and transposition of C1, C2, C3, and C4.

As for the motion vector of decimal pixel precision, peripheral pixels required for motion compensation of half-pixel and half-line precision is subjected to phase compensation on a plurality of blocks (4 blocks), subjected subsequently to interpolation processing of Hadamard transform coefficient with the same order, inverse Hadamard transform (same as H) is executed, and final prediction matrix is formed. This concept is shown in Figure A.2-14.

The structure for realizing the function described above is described. A block diagram is shown in Figure A.2-15.

A.2.8.4 Signal processing from input to output

The motion compensation of half-pixel precision/half-line precision is the premise. In this case, because interpolation processing is required, it is necessary that all reference pixels in the vicinity are inputted. This concept is shown in Figure A.2-16. Therefore, 9 pixels × 10 lines = 90 pixels are required to input per 1 block. 8 pixels × 8 lines = 64 pixels overlapped between A, B, C, and D shown in Figure A.2-16 are subjected individually to Hadamard transform (WHT). The transform coefficient is subjected to PC (phase compensation) with matching to phase shift 1, 2, 3, and 4 by the register as serial data from motion vector of each A, B, C, and D block. The expression of Figure A.2-11 on serial data is Figure A.2-17.

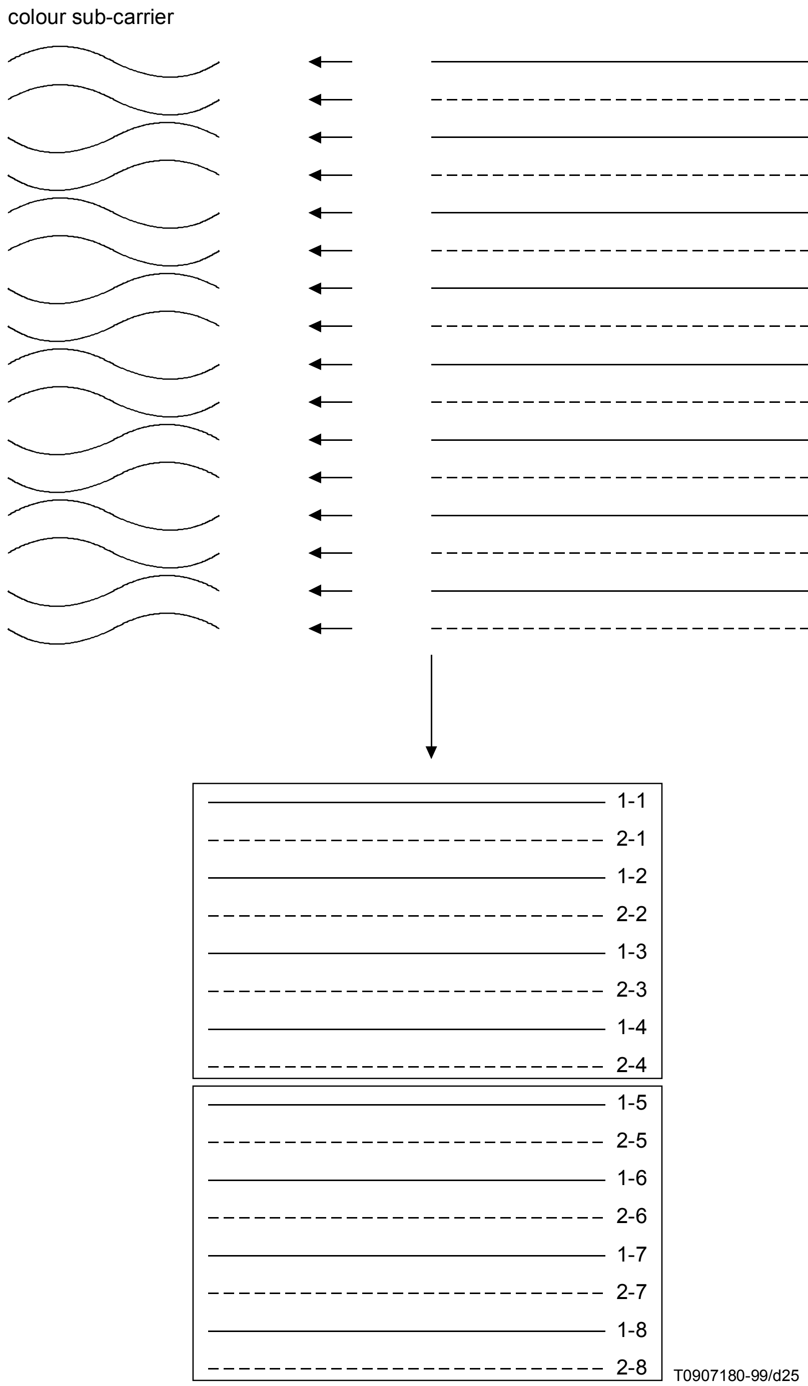


Figure A.2-12/J.88 – Matrix in the frame base

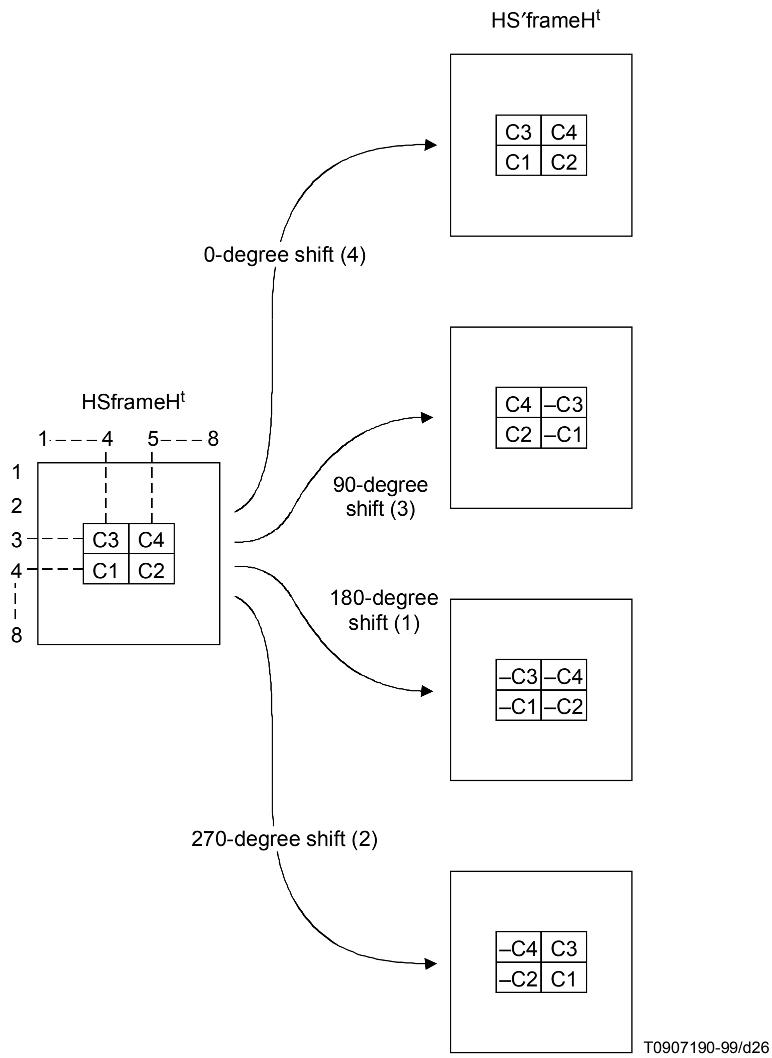


Figure A.2-13/J.88 – Colour sub-carrier on two-dimensional Hadamard transform

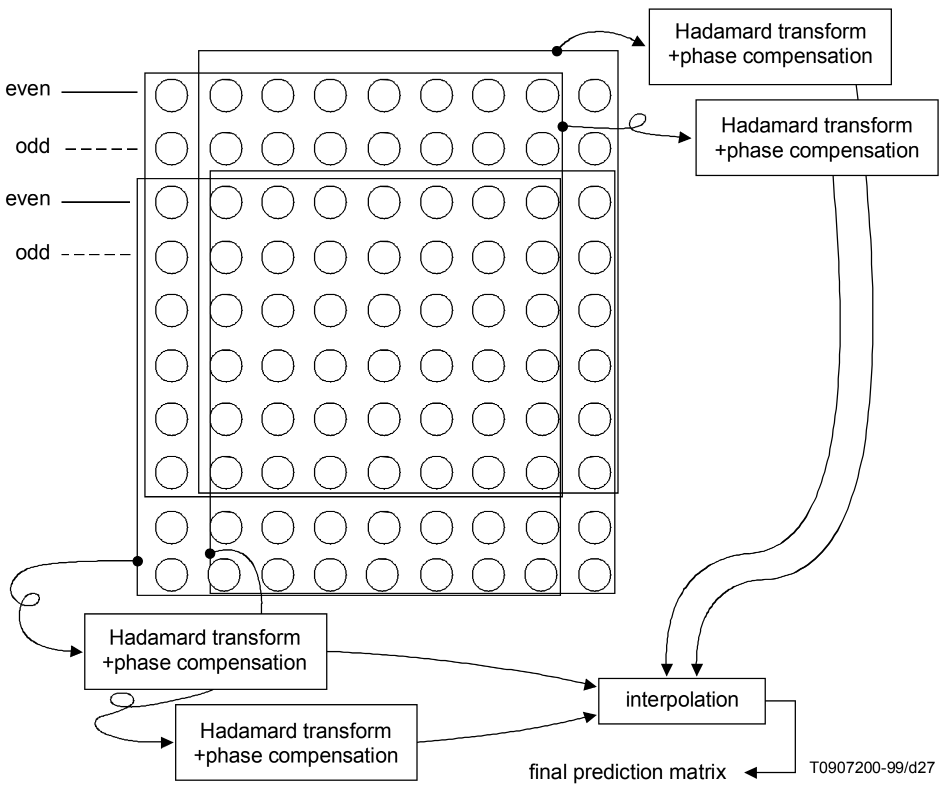


Figure A.2-14/J.88 – Concept of interpolation process on Hadamard transformation structure

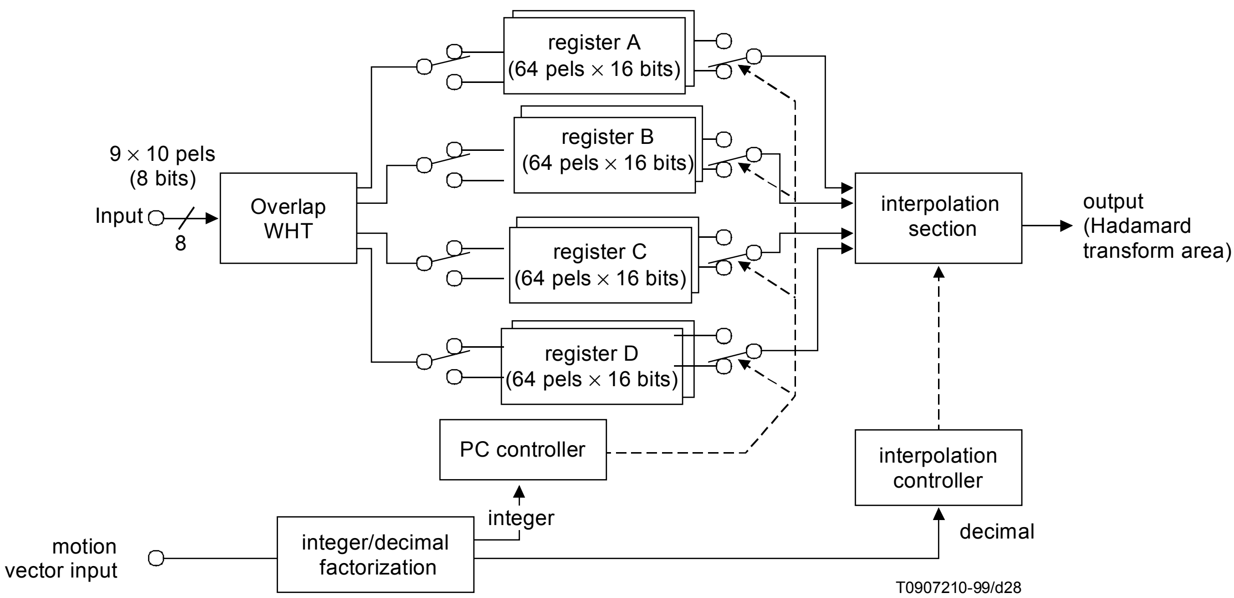


Figure A.2-15/J.88 – Structure of composite motion compensation

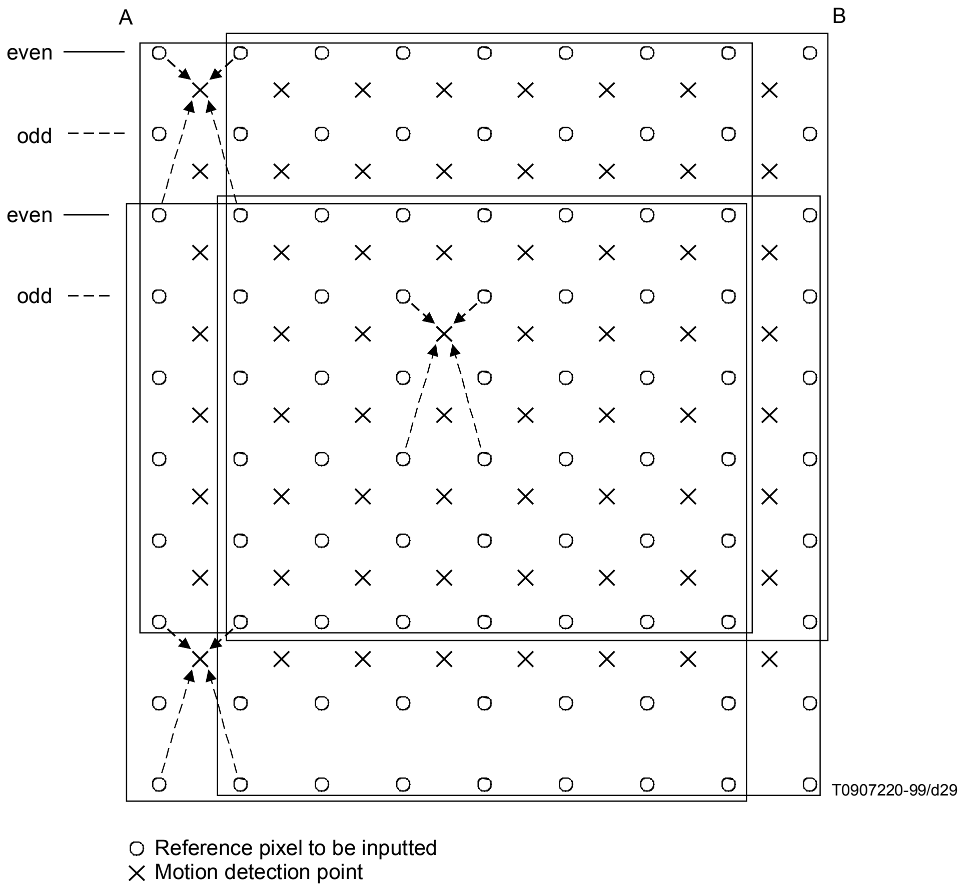


Figure A.2-16/J.88 – Reference pixel to be inputted and motion detection point

Writing and reading order	1, ..., 19, 20, ..., 51, 52, ..., 64
Writing	______, C_1, C_2, ______, C_3, C_4, ______
0-degree shift ④	______, C_1, C_2 ______, C_3, C_4, ______
90-degree shift ③	______, C_2, $-C_1$ ______, C_4, $-C_3$, ______
180-degree shift ①	______, $-C_1$, $-C_2$ ______, $-C_3$, $-C_4$, ______
270-degree shift ②	______, $-C_2$, C_1 ______, $-C_4$, $-C_3$, ______

Figure A.2-17/J.88 – Phase compensation (PC) expression on serial data

The phase shift in each register A, B, C, and D is calculated as described in the following.

Integer portions (integer portion as complement expression of 2) of MVx and MVy are referred to X and Y.

Register A:

$$\{X + Y\} = \begin{cases} 4n + 2 & \rightarrow ④ \\ 4n & \rightarrow ① \\ 4n + 1 & \rightarrow ② \\ 4n + 3 & \rightarrow ③ \end{cases}$$

Register B, Register C:

$$\{X + Y\} = \begin{cases} 4n + 1 & \rightarrow ④ \\ 4n + 3 & \rightarrow ① \\ 4n & \rightarrow ② \\ 4n + 2 & \rightarrow ③ \end{cases}$$

Register D:

$$\{X + Y\} = \begin{cases} 4n & \rightarrow ④ \\ 4n + 2 & \rightarrow ① \\ 4n + 3 & \rightarrow ② \\ 4n + 1 & \rightarrow ③ \end{cases}$$

Half pixel/half line is interpolated from the coefficient of the same order of the 4 blocks A, B, C, and D having the same phase resultant from the phase compensation to form the final prediction block. The even line is distant 1/60 seconds from the odd line in Figure A.2-16; in the case of a rapidly moving scene, the correlation between even line and odd line becomes significantly poor. To avoid this disadvantage, interpolation is performed for the half pixel/half line in the same field as shown in Figure A.2-18.

A.2.8.5 Calculation precision

The calculation precision in composite motion compensation is shown in Figure A.2-19.

In WHT calculation, the full bit calculation is supported by widening the input bit width by 3 bits or more to form the wider output bit. The bit width is the same before and after phase compensation. All interpolation coefficients for interpolation are integer multiples of 1/16 (0 to 16); the interpolation is performed by 4-bit shift calculation to widen the input bit width by 4 bits to form the wider output bit width, and the full bit calculation is supported. In the final rounding, real number expression with 19-bit precision is rounded to convert it to integer expression with 12-bit precision.

A.3 Decoding process

The decoding process for frame regenerating from coded data is specified. The whole structure of this process is shown in Figure A.3-1.

Four types of slices are reflected on the control of inverse quantization and inverse mixed mode. Because Es is formed by the slice counter in the decoder, Es is not transmitted as control information.

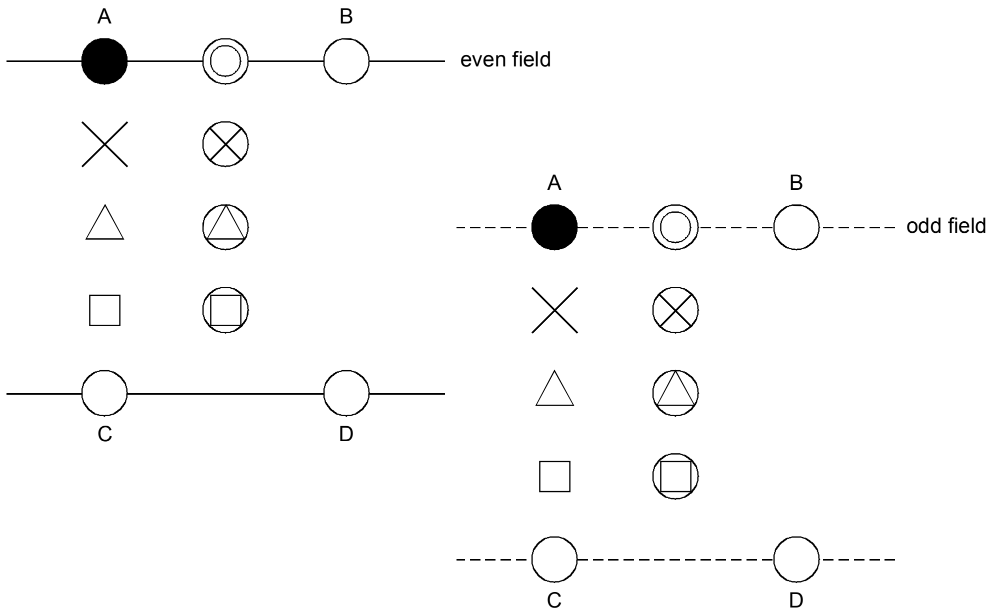
A.3.1 Variable length decoding (VLD1, VLD2)

For variable length coding of coded data, a two-dimensional Huffman code having the combination of zero-run length and quantization output level as a symbol is used (refer to A.2.7).

A Huffman code having code length of 3 to 14 bits is used for 126 specified symbols in assignment of code. Codes in the form described in the following are assigned to the residual symbols: Code word (22 bits) = escape code (6 bits) + run length (6 bits) + quantization level number (10 bits).

Different combinations of symbols and Huffman code patterns are utilized for different coding modes (M). Variable length decoding is performed under the condition described above (VLD1).

The motion vector is subjected to DPCM between macroblocks in the raster scanning direction of each MVx and MVy; the difference value transmitted in the form of Huffman code is decoded (VLD2) (see Appendix II).



	A	B	C	D
●	1	0	0	0
×	3/4	0	1/4	0
△	1/2	0	1/2	0
□	1/4	0	3/4	0
⊗	5/16	5/16	3/16	3/16
△⃝	1/4	1/4	1/4	1/4
□⃝	3/16	3/16	5/16	5/16
◎	1/2	1/2	0	0

T0907230-99/d30

Figure A.2-18/J.88 – Interpolation process for half-pixel and half-line precision motion compensation

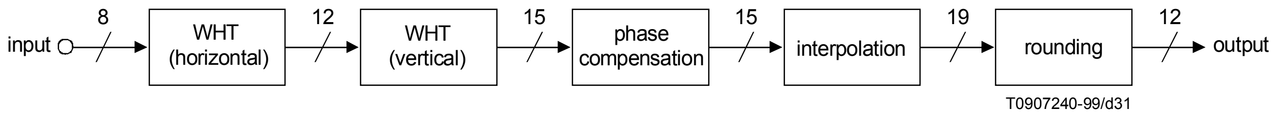


Figure A.2-19/J.88 – Calculation precision in composite motion compensation

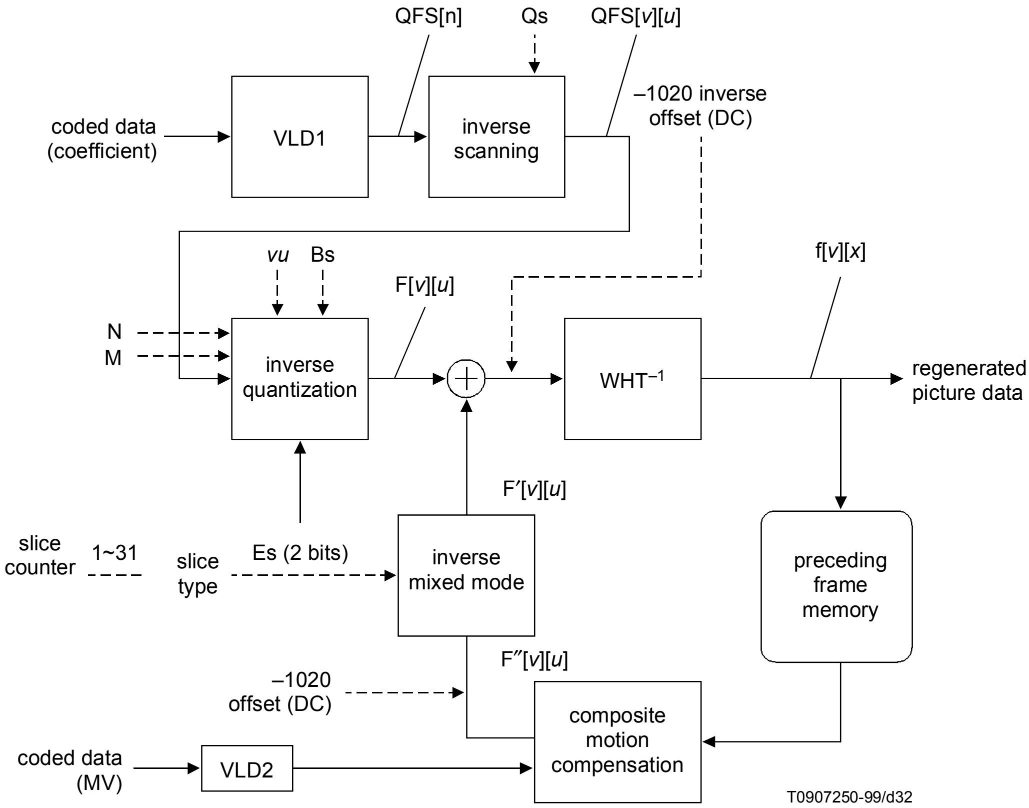


Figure A.3-1/J.88 – Whole structure of decoding process

A.3.2 Inverse scanning

QFS[n] in the form of one-dimensional data is converted to QF[v][u] in the form of a two-dimensional coefficient array, wherein u and v are in the range of 0 to 7.

The inverse scanning is performed according to the following process:

for (v = 0; v < 8; v++)

for (u = 0; u < 8; u++)

QF[v][u] = QFS[Scan[alternate-scan[v][u]]]

A.3.3 Inverse quantization

In the inverse quantization processing, the quantization step size Delta[v][u] is determined using the function g1 by coding mode M (1 bit), macroblock activity N (2 bits), buffer occupancy Bs (5 bits), and slice type Es (2 bits). Next, F[v][u] is obtained using the function g3 (refer to Figure A.3-2).

Delta[v][u] = g1(M, N, Bs, Es)

F[v][u] = g3(Delta[v][u], QF[v][u])

wherein g3:

$$\mathrm{F}[v][u] = \begin{cases} 1023 \left(\mathrm{F}[v][u] > 1023\right) \\ \mathrm{QF}[v][u]\mathrm{Delta}[v][u] \left(-1024 \le \mathrm{F}[v][u] \le 1023\right) \\ -1024 \ \left(\mathrm{F}[v][u] < -1024\right) \end{cases}$$

The basic operation of the function g3 is multiplication of QF[v][u] by Delta[v][u]. As for the content of g1, parameters transmitted in the sequence header 2 is used for matching to coding.

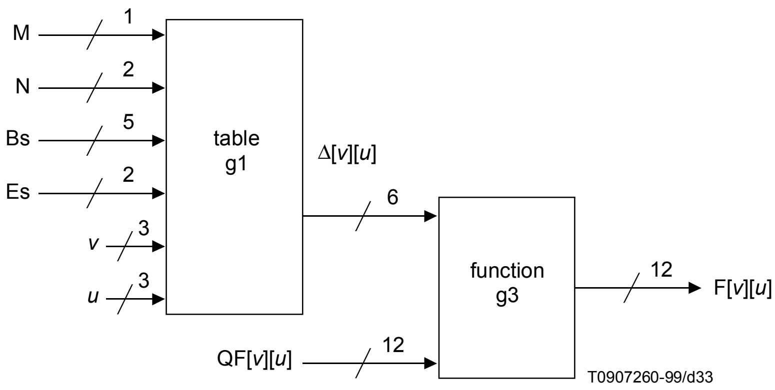


Figure A.3-2/J.88 – Structure of inverse quantization

A.3.4 Inverse Hadamard transformation (WHT^{-1})

WHT^{-1} is given as follows:

$$[\mathrm{C}] = \frac{1}{8}\ [\mathrm{H}] \otimes [\mathrm{V}] \otimes [\mathrm{H}]$$

$$\text{where } [\mathrm{H}] = \begin{bmatrix} + & + & + & + & + & + & + & + \\ + & + & + & + & - & - & - & - \\ + & + & - & - & - & - & + & + \\ + & + & - & - & + & + & - & - \\ + & - & - & + & + & - & - & + \\ + & - & - & + & + & + & + & - \\ + & - & + & - & - & + & - & + \\ + & - & + & - & + & - & + & - \end{bmatrix}$$

A.3.5 Inverse mixed mode

In the inverse mixed mode, the prediction coefficient A[v][u] is determined using the function h1 by coding mode M (1 bit), macroblock activity N (2 bits), buffer occupancy Bs (5 bits), and slice type (2 bits). Next, F′[v][u] is obtained using the function h2.

$$\mathrm{A}[v][u] = \mathrm{h1}(\mathrm{M, N, Bs, Es})$$

$$\mathrm{F}'[v][u] = \mathrm{h2}(\mathrm{A}[v][u], \mathrm{F}''[v][u])$$

wherein h2:

$$\mathrm{F}'[v][u] = \mathrm{F}''[v][u] \cdot \mathrm{A}[v][u]$$

The function of the function h2 is multiplication of F″[v][u] by $0 \le \mathrm{A}[v][u] \le 1$. The mode is intramode for A = 0 and intermode for A = 1. A[v][u] is controlled so that the prediction efficiency is maximized from the viewpoint of motion compensation frame correlation. To maximize it, the control is performed with linking to Delta[v][u] which determines the extent of feedback quantization noise with impact on the correlation. Accordingly, the method of control is shown in Figure A.3-3.

As for the content of h1, parameters transmitted in the sequence header 3 are used for matching to coding.

A.3.6 Composite motion compensation

Refer to A.2.8.

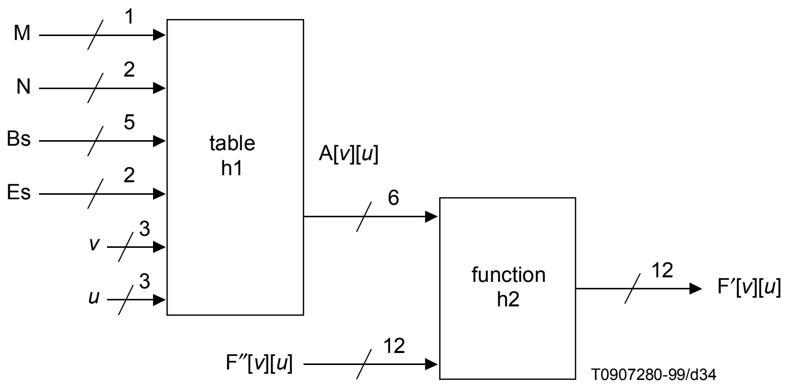


Figure A.3-3/J.88 – Structure of inverse mixed mode

Appendix I

Examples for adaptive scanning patterns

For adaptive scannings (refer to A.2.6) and inverse scannings (refer to A.3.2), the following examples are recommended.

Scanning pattern QS = 00							
0	3	6	15	22	24	26	29
4	8	17	27	31	33	36	40
7	18	30	14	13	38	42	45
16	28	11	1	2	10	47	53
23	32	35	39	44	48	49	55
25	34	37	43	46	50	51	58
12	21	41	52	54	57	60	61
5	9	19	20	56	59	62	63

Scanning pattern QS = 01							
0	3	6	11	18	20	22	25
4	8	13	23	51	49	45	43
7	14	26	27	28	47	41	39
12	24	29	1	2	30	48	53
19	31	32	34	50	44	42	55
21	33	35	36	46	40	38	58
10	17	37	52	54	57	60	61
5	9	15	16	56	59	62	63

Scanning pattern QS = 00							
0	3	6	13	17	19	21	24
4	8	15	22	31	32	34	35
7	16	25	12	11	33	36	37
14	23	10	1	2	9	48	53
18	51	49	45	50	44	42	55
20	47	43	41	46	40	38	58
29	26	39	52	54	57	60	61
5	30	28	27	56	59	62	63

Scanning pattern QS = 01							
0	3	6	9	13	15	17	20
4	8	11	18	50	48	44	42
7	12	21	25	26	46	40	38
10	19	28	1	2	29	32	53
14	51	49	45	31	34	35	55
16	47	43	41	33	36	37	58
27	22	39	52	54	57	60	61
5	30	24	23	56	59	62	63

Appendix II

Examples for VLC of motion vectors

The following examples are recommended as a variable length code (VLC) of motion vectors (–16.0 to 15.5 pixels, frame lines). Refer to A.2.7 and A.3.1.

VLC for DPCM	Motion vector
0000 0000 000	–32.0
0000 0000 010	–31.5
0000 0000 100	–31.0
0000 0000 110	–30.5
0000 0001 000	–30.0
0000 0001 010	–29.5
0000 0001 100	–29.0
0000 0001 110	–28.5
0000 0010 000	–28.0
0000 0010 010	–27.5
0000 0010 100	–27.0
0000 0010 110	–26.5
0000 0011 000	–26.0
0000 0011 010	–25.5
0000 0011 100	–25.0
0000 0011 110	–24.5
0000 0100 000	–24.0
0000 0100 010	–23.5
0000 0100 100	–23.0
0000 0100 110	–22.5
0000 0101 000	–22.0
0000 0101 010	–21.5
0000 0101 100	–21.0
0000 0101 110	–20.5
0000 0110 000	–20.0
0000 0110 010	–19.5
0000 0110 100	–19.0
0000 0110 110	–18.5
0000 0111 000	–18.0
0000 0111 010	–17.5
0000 0111 100	–17.0
0000 0111 110	–16.5
0000 1000 000	–16.0

VLC for DPCM	Motion vector
0000 1000 010	–15.5
0000 1000 100	–15.0
0000 1000 110	–14.5
0000 1001 000	–14.0
0000 1001 010	–13.5
0000 1001 100	–13.0
0000 1001 110	–12.5
0000 1010 000	–12.0
0000 1010 010	–11.5
0000 1010 100	–11.0
0000 1010 110	–10.5
0000 1011 000	–10.0
0000 1011 010	–9.5
0000 1011 100	–9.0
0000 1011 110	–8.5
0000 1100 000	–8.0
0000 1100 010	–7.5
0000 1100 100	–7.0
0000 1100 110	–6.5
0000 1101 000	–6.0
0000 1101 010	–5.5
0000 1101 100	–5.0
0000 1101 110	–4.5
0000 1110 00	–4.0
0000 1110 10	–3.5
0000 1111 00	–3.0
0000 1111 10	–2.5
0001 10	–2.0
0001 0	–1.5
0010	–1.0
010	–0.5
1	0.0

011	0.5
0011	1.0
0001 11	1.5
0000 1111 11	2.0
0000 1111 01	2.5
0000 1110 11	3.0
0000 1110 01	3.5
0000 1101 111	4.0
0000 1101 101	4.5
0000 1101 011	5.0
0000 1101 001	5.5
0000 1100 111	6.0
0000 1100 101	6.5
0000 1100 011	7.0
0000 1100 001	7.5
0000 1011 111	8.0
0000 1011 101	8.5
0000 1011 011	9.0
0000 1011 001	9.5
0000 1010 111	10.0
0000 1010 101	10.5
0000 1010 011	11.0
0000 1010 001	11.5
0000 1001 111	12.0
0000 1001 101	12.5
0000 1001 011	13.0
0000 1001 001	13.5
0000 1000 111	14.0
0000 1000 101	14.5
0000 1000 011	15.0
0000 1000 001	15.5
0000 0111 111	16.0

0000 0111 101	16.5
0000 0111 011	17.0
0000 0111 001	17.5
0000 0110 111	18.0
0000 0110 101	18.5
0000 0110 011	19.0
0000 0110 001	19.5
0000 0101 111	20.0
0000 0101 101	20.5
0000 0101 011	21.0
0000 0101 001	21.5
0000 0100 111	22.0
0000 0100 101	22.5
0000 0100 011	23.0
0000 0100 001	23.5
0000 0011 111	24.0
0000 0011 101	24.5
0000 0011 011	25.0
0000 0011 001	25.5
0000 0010 111	26.0
0000 0010 101	26.5
0000 0010 011	27.0
0000 0010 001	27.5
0000 0001 111	28.0
0000 0001 101	28.5
0000 0001 011	29.0
0000 0001 001	29.5
0000 0000 111	30.0
0000 0000 101	30.5
0000 0000 011	31.0
0000 0000 001	31.5

Appendix III

Examples for quantizer and mix mode parameters design

The following design scheme is recommended for quantizer parameters (A.2.6, A.3.3) and mix mode parameters (A.2.5, A.3.5).

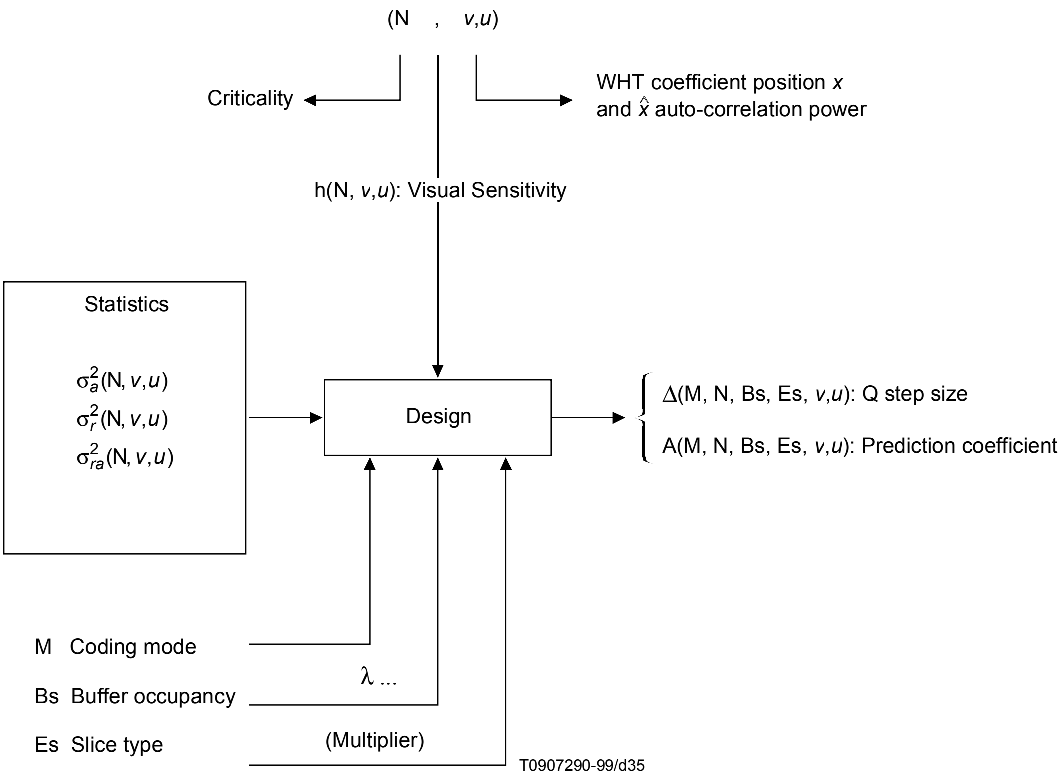
Suppose that x is a current frame signal to be coded, $\hat{x}$ is a reference frame signal and e is a difference value of these two values,

$$e = x - \hat{x}$$

$$E\left[e^2\right] = E\left[(x - \hat{x})^2\right]$$

$$\sigma_r^2 = 2\sigma_a^2(1-\rho) = 2(\sigma_a^2 - \sigma_{ra}^2),$$

where ρ is a correlation between x and $\hat{x}$, σ_a^2 is an original signal power, σ_r^2 is a difference signal power and σ_{ra}^2 is a x and $\hat{x}$ auto-correlation power.



1) Es = 00 (Control)

   A(0, N, Bs, 00, v,u) = 1 at any N, Bs, v,u

   A(1, N, Bs, 00, v,u) = 0 at any N, Bs, v,u

   Δ(M, N, Bs, 00, v,u) = 3 at any M, N, Bs, v,u

2) Es = 01 (Helper signal)

Let N = 01 and h (01, v,u) = 1.0 for statistics and for each λ such that $\lambda = 1.5 \times Bs - 26$ (Bs = 0, 1, ..., 31),

when M = 0,

A is a solution for (III-1), and the optimum A is searched at $0 \leq A \leq 1$ with 1/64 accuracy,

Δ is a solution for (III-2) with the obtained A and the optimum Δ is searched at $1 \leq \Delta \leq 127$ with 2.0 accuracy with a clipping by a triple value of Δ (N, 0, 0) (= DC).

When M = 1,

A = 0,

Δ is a solution for (III-3) with the obtained A and the optimum Δ is searched at $1 \leq \Delta \leq 127$ with 2.0 accuracy with a clipping by an 8 times value of Δ (N, 0, 0) (= DC).

3) Es = 10 (Boundary)

Let N = 01 and h (01, v,u) = 1.0 for statistics and for each λ such that $\lambda = 1.5 \times Bs - 21$ (Bs = 0, 1, ..., 31),

when M = 0,

A is a solution for (III-1) and the optimum A is searched at $0 \leq A \leq 1$ with 1/64 accuracy,

Δ is a solution for (III-2) with the obtained A and the optimum Δ is searched at $1 \leq \Delta \leq 127$ with 2.0 accuracy with a clipping by a triple value of Δ (N, 0, 0) (= DC).

When M = 1,

A = 0,

Δ is a solution for (III-3) with the obtained A and the optimum Δ is searched at $1 \leq \Delta \leq 127$ with 2.0 accuracy with a clipping by an 8 times value of Δ (N, 0, 0) (= DC).

4) Es = 11 (Main)

Let N = 00, 01, 10, 11 and h (01, v,u) = 1.0 for statistics and for each λ such that $\lambda = 1.5 \times Bs - 16$ (Bs = 0, 1, ..., 31),

when M = 0,

A is a solution for (III-1) and the optimum A is searched at $0 \leq A \leq 1$ with 1/64 accuracy,

Δ is a solution for (III-2) with the obtained A and the optimum Δ is searched at $1 \leq \Delta \leq 127$ with 2.0 accuracy with a clipping by a triple value of Δ (N, 0, 0) (= DC).

When M = 1,

A = 0,

Δ is a solution for (III-2) with the obtained A and the optimum Δ is searched at $1 \leq \Delta \leq 127$ with 2.0 accuracy with a clipping by a triple value of Δ (N, 0, 0) (= DC).

$$\alpha A^3 + \beta A^2 + \gamma A + \delta = 0 \qquad \text{(III-1)}$$

where:

$$\alpha = \frac{2\lambda(\sigma_a^2 + \sigma_r^2 - 2\sigma_{ra}^2)}{(\ell n2)h}$$

$$\beta = -4(\sigma_a^2 + 2\sigma_{ra}^4 - 3\sigma_a^2\sigma_{ra}^2 - \sigma_r^2\sigma_{ra}^2 + \sigma_a^2\sigma_r^2) - \frac{4\lambda(\sigma_a^2 - \sigma_{ra}^2)}{(\ell n2)h}$$

$$\gamma = 4(2\sigma_a^2 + 2\sigma_{ra}^4 + \sigma_a^2\sigma_r^2 - 4\sigma_a^2\sigma_{ra}^2) + \frac{2\lambda\sigma_a^2}{(\ell n2)h}$$

$$\frac{4p\left\{1 - \exp(-q\Delta) - q\Delta\exp(-q\Delta/2)\right\}}{q\left\{1 - \exp(-q\Delta)\right\}} - n = 0 \qquad \text{(III-2)}$$

where:

$$p = \frac{1}{\sqrt{2(n+\sigma^2)}}$$

$$q = \frac{\sqrt{2}}{\sqrt{(n+\sigma^2)}}$$

$$\sigma^2 = (1-A)^2\sigma_a^2 + 2(1-A)A\sigma_{ra}^2 + A^2\sigma_r^2$$

$$n = \frac{\sigma_a^2 - \sigma_{ra}^2}{A} - (\sigma_a^2 + \sigma_r^2 - 2\sigma_{ra}^2)$$

$$\frac{4p\left\{1-\exp(-q\Delta) - q\Delta\exp(-q\Delta/2)\right\}}{q\left\{1-\exp(-q\Delta)\right\}} - n = 0 \qquad \text{(III-3)}$$

where:

$$p = \frac{1}{\sqrt{2\sigma_a^2}}$$

$$q = \frac{\sqrt{2}}{\sqrt{\sigma_a^2}}$$

Statistics $(\sigma_a^2, \sigma_r^2, \sigma_{ra}^2)$ examples:

σ_a^2 **(N = 00, v, u)**							
141685.49	273.01	63.13	67.89	379.26	27.60	20.69	14.21
737.55	6.19	3.89	4.07	353.54	3.55	3.94	1.91
161.59	4.49	3.56	4.25	362.82	4.81	3.10	1.66
166.20	5.18	4.38	8188.33	26680.72	8.94	3.71	2.42
72.66	4.33	2.73	2.64	3.54	2.50	2.28	1.12
69.87	3.14	2.49	2.35	2.93	2.38	2.55	1.11
67.33	2.89	2.27	2.58	3.26	3.35	2.48	1.13
72.26	2.77	2.88	2.81	4.13	4.13	3.27	1.30

σ_a^2 (N = 01, v, u)							
122636.61	3270.23	943.68	729.11	340.11	204.31	195.06	138.47
1459.28	187.32	107.13	85.24	154.59	40.48	26.99	14.84
561.93	145.75	84.57	128.55	428.81	46.73	20.74	12.34
528.26	128.55	287.97	10943.77	25130.32	217.79	59.70	39.63
178.78	47.00	32.05	30.97	32.59	18.28	10.53	5.07
136.13	43.28	29.52	28.79	24.51	15.27	10.97	3.91
112.17	29.03	25.81	26.30	23.04	15.52	11.36	4.25
102.63	82.79	42.70	57.01	34.00	52.90	19.99	6.84

σ_a^2 (N = 10, v, u)							
91710.61	4079.17	1380.48	1201.00	671.84	395.47	304.16	181.04
5504.97	802.79	402.07	269.16	365.04	165.73	94.93	35.79
2333.62	556.00	331.89	431.75	1297.27	154.90	89.27	35.24
1507.20	363.44	313.90	8257.84	15439.49	440.50	95.77	88.00
694.17	196.79	130.81	109.29	109.62	54.43	31.76	11.52
548.41	157.65	102.73	91.96	91.22	50.95	31.61	10.32
510.61	154.15	95.59	76.43	74.06	48.90	27.67	7.31
306.90	113.40	119.63	107.36	103.86	60.08	36.30	10.37

σ_a^2 (N = 11, v, u)							
46033.57	5377.94	2709.41	2210.38	1572.38	1014.42	576.27	258.89
7304.38	2353.84	1477.81	1008.19	847.24	510.26	309.17	109.03
3610.62	1459.47	1000.93	925.10	1568.21	433.22	240.33	89.94
2188.28	818.27	719.30	6862.91	10760.50	662.72	208.54	133.15
1222.92	714.33	556.14	445.76	391.61	247.58	134.79	43.52
1100.34	574.40	457.66	375.33	328.17	225.20	125.40	34.86
887.81	452.25	360.80	319.69	292.12	199.91	114.09	30.96
501.80	295.10	290.79	300.74	308.23	201.16	98.57	25.80

σ_r^2 (N = 00, v, u)							
9.96	5.23	5.20	6.74	9.91	6.86	6.01	3.04
9.07	4.13	3.35	4.76	7.41	5.32	4.83	2.08
9.14	5.72	4.30	5.94	9.11	4.79	3.95	1.75
10.04	5.43	4.37	5.93	9.71	6.06	4.11	2.01
6.82	4.48	3.88	3.92	4.47	4.11	3.17	1.48
6.34	4.00	3.76	3.80	4.35	4.28	3.44	1.37
4.80	3.56	3.85	3.69	4.50	4.28	3.66	1.63
5.76	3.27	3.91	3.76	4.52	4.53	3.74	1.50

σ_r^2 (N = 01, v, u)							
80.86	44.84	36.40	35.39	41.92	39.17	20.84	8.60
69.41	37.05	25.37	22.53	43.79	23.48	14.63	6.43
76.87	41.49	31.09	93.08	78.55	27.86	13.41	7.17
69.35	38.12	32.58	55.45	85.83	36.17	19.79	9.87
73.73	31.66	25.11	18.57	16.78	13.68	9.29	3.43
51.05	18.77	22.24	15.86	16.66	12.97	10.12	3.19
23.02	11.81	14.16	13.88	17.67	14.37	10.37	3.76
22.16	11.67	15.91	13.54	19.38	17.31	11.23	3.79

σ_r^2 (N = 10, v, u)							
236.74	122.15	96.27	144.44	166.83	84.20	46.98	19.49
281.53	136.44	105.64	139.43	163.16	63.00	35.86	13.16
343.59	128.70	99.14	269.84	279.28	82.56	38.81	16.25
362.95	123.03	133.80	181.57	238.55	143.04	49.94	29.94
353.45	140.79	91.49	86.56	57.23	45.52	29.24	9.61
186.16	69.48	60.56	66.62	53.76	35.21	20.89	6.72
108.12	54.54	45.86	53.53	51.61	35.80	22.05	6.32
55.47	38.63	36.63	50.18	51.90	34.01	21.30	6.17

σ_r^2 (N = 11, v, u)							
393.62	327.97	291.72	408.88	417.00	259.75	143.09	35.04
498.70	322.73	269.81	310.88	348.16	189.06	102.18	27.70
673.23	338.39	310.78	1011.83	889.25	209.59	96.42	35.03
695.33	339.92	392.65	784.14	685.86	386.52	121.12	72.22
781.48	418.44	377.25	326.32	298.77	187.51	101.71	30.44
410.21	241.45	241.96	239.77	224.62	155.19	90.04	23.87
257.98	152.58	157.80	189.69	180.27	131.09	67.84	18.30
115.65	86.40	110.88	153.93	166.93	114.31	70.14	15.00

σ_{ra}^2 (N = 00, v, u)							
–147.34	2.42	2.75	2.74	1.06	3.47	3.81	1.94
7.41	1.85	1.42	2.45	2.71	2.47	2.92	1.11
4.33	2.44	2.30	2.84	4.10	3.01	2.12	0.91
2.02	2.30	2.28	6.19	14.68	4.07	2.54	1.22
1.47	1.65	1.91	1.95	2.59	2.14	1.76	0.72
0.91	1.25	1.69	1.87	2.31	2.01	1.97	0.64
0.24	1.49	1.73	1.79	2.17	2.49	1.97	0.74
0.56	1.19	1.67	1.57	2.03	2.51	1.86	0.63

σ_{ra}^2 (N = 01, v, u)							
–98.65	7.23	12.28	16.74	28.48	21.16	14.01	7.51
68.38	26.17	30.12	17.02	24.05	11.17	11.50	4.64
52.67	31.47	20.33	55.92	46.40	15.94	7.18	4.13
55.88	22.52	21.74	20.49	–26.08	18.16	11.99	5.17
57.78	22.65	17.36	12.42	9.41	7.57	5.46	2.04
38.16	11.66	14.39	10.14	7.94	7.16	6.41	1.92
17.30	6.30	9.06	7.86	8.11	7.63	5.71	2.06
19.09	11.18	14.18	9.33	5.91	8.90	4.90	1.51

σ_{ra}^2 **(N = 10, v, u)**							
270.44	95.98	23.19	37.17	130.06	43.82	28.90	11.65
322.86	117.69	79.12	58.00	103.12	43.58	23.89	8.23
403.99	105.67	67.83	161.16	159.87	51.20	27.70	10.10
318.06	90.48	63.68	84.56	79.09	94.84	32.74	21.90
308.31	89.61	62.32	55.25	37.86	30.66	19.44	5.72
157.51	42.95	40.61	41.18	35.88	23.41	14.10	4.26
100.88	38.99	29.09	33.19	31.78	20.95	12.49	3.17
28.56	22.89	22.84	28.38	28.02	16.53	11.88	3.20

σ_{ra}^2 **(N = 11, v, u)**							
215.63	239.86	150.60	233.20	229.63	138.83	66.65	15.96
462.41	267.85	211.23	200.69	238.45	114.55	61.91	18.08
759.47	295.49	235.77	590.67	452.90	144.21	71.63	26.67
631.80	248.12	226.37	447.87	356.88	278.19	88.28	53.05
648.37	324.00	282.16	233.76	211.48	130.67	71.81	21.87
350.83	181.33	166.19	155.59	149.44	101.90	59.85	15.73
215.02	103.80	100.77	111.81	122.75	84.34	43.20	10.98
94.04	52.48	69.36	80.46	96.70	66.93	37.45	7.58

Visual sensitivity (h) examples:

h(N = 00, v, u)							
1.0000	1.0000	0.7719	0.6314	0.5359	0.5359	0.4384	0.4384
1.0000	0.7719	0.6314	0.4384	0.3384	0.3384	0.3384	0.1813
0.7719	0.6314	0.4384	0.6314	0.7719	0.3384	0.1813	0.1813
0.6314	0.4384	0.6314	1.0000	1.0000	0.7719	0.1813	0.1813
0.5359	0.3384	0.3384	0.3384	0.1813	0.1813	0.1813	0.1813
0.5359	0.3384	0.3384	0.1813	0.1813	0.1813	0.1813	0.1813
0.7719	0.5359	0.1813	0.1813	0.1813	0.1813	0.1813	0.1813
0.7719	0.7719	0.5359	0.1813	0.1813	0.1813	0.1813	0.1813

h(N = 01, v, u)							
0.7268	0.7268	0.6581	0.6094	0.5722	0.5722	0.5298	0.5298
0.7268	0.6581	0.6094	0.5298	0.4798	0.4798	0.4798	0.3777
0.6581	0.6094	0.5298	0.6094	0.6581	0.4798	0.3777	0.3777
0.6094	0.5298	0.6094	0.7268	0.7268	0.6581	0.3777	0.3777
0.5722	0.4798	0.4798	0.4798	0.3777	0.3777	0.3777	0.3777
0.5722	0.4798	0.4798	0.3777	0.3777	0.3777	0.3777	0.3777
0.6581	0.5722	0.3777	0.3777	0.3777	0.3777	0.3777	0.3777
0.6581	0.6581	0.5722	0.3777	0.3777	0.3777	0.3777	0.3777

h(N = 10, v, u)							
0.6072	0.6072	0.5670	0.5377	0.5149	0.5149	0.4882	0.4882
0.6072	0.5670	0.5377	0.4882	0.4559	0.4559	0.4559	0.3866
0.5670	0.5377	0.4882	0.5377	0.5670	0.4559	0.3866	0.3866
0.5377	0.4882	0.5377	0.6072	0.6072	0.5670	0.3866	0.3866
0.5149	0.4559	0.4559	0.4559	0.3866	0.3866	0.3866	0.3866
0.5149	0.4559	0.4559	0.3866	0.3866	0.3866	0.3866	0.3866
0.5670	0.5149	0.3866	0.3866	0.3866	0.3866	0.3866	0.3866
0.5670	0.5670	0.5149	0.3866	0.3866	0.3866	0.3866	0.3866

h(N = 11, v, u)							
0.4125	0.4125	0.3395	0.2919	0.2580	0.2580	0.2218	0.2218
0.4125	0.3395	0.2919	0.2218	0.1826	0.1826	0.1826	0.1142
0.3395	0.2919	0.2218	0.2919	0.3395	0.1826	0.1142	0.1142
0.2919	0.2218	0.2919	0.4125	0.4125	0.3395	0.1142	0.1142
0.2580	0.1826	0.1826	0.1826	0.1142	0.1142	0.1142	0.1142
0.2580	0.1826	0.1826	0.1142	0.1142	0.1142	0.1142	0.1142
0.3395	0.2580	0.1142	0.1142	0.1142	0.1142	0.1142	0.1142
0.3395	0.3395	0.2580	0.1142	0.1142	0.1142	0.1142	0.1142

Appendix IV

Fast Hadamard transform

The following method is recommended as FHT.

8×8 WHT matrix H(8) is as follows (Constant value is omitted):

$$H(8) = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{bmatrix}$$

This matrix can be rewritten as follows:

$$h(8) = [A] \times [A] \times [A]$$

where:

$$[A] = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$

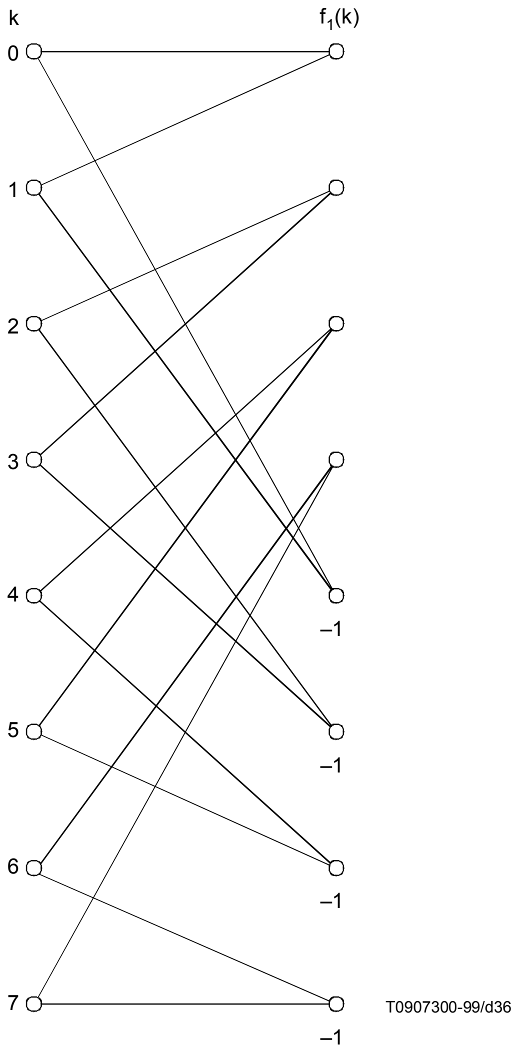


Figure IV.1/J.88 – Flow diagram of [A]

Therefore, the flow diagram α of H(8) can be obtained by Figure IV.2.

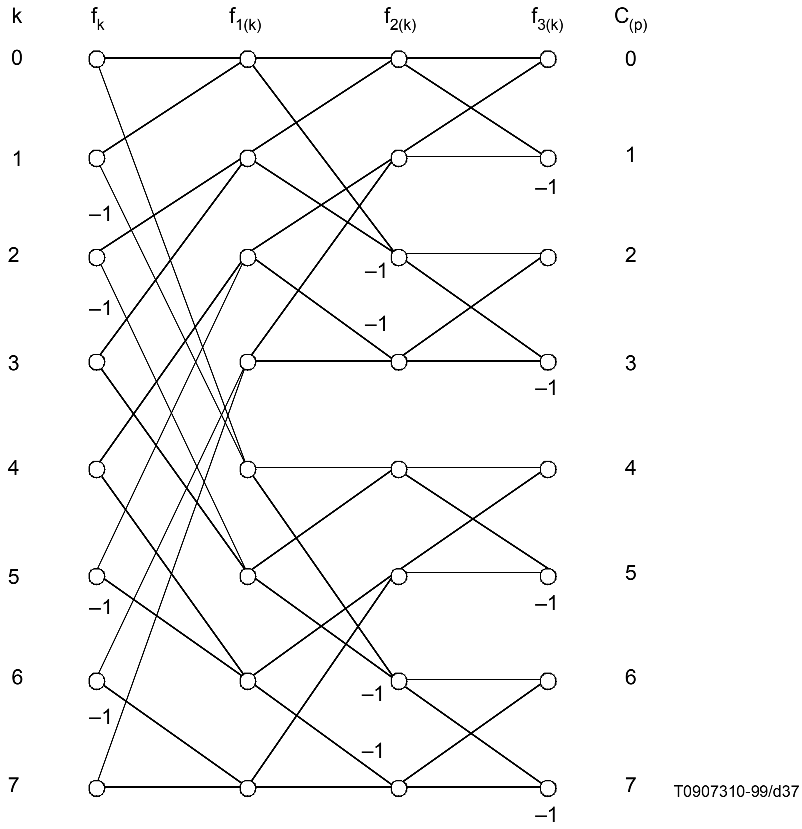


Figure IV.2/J.88 – Flow diagram α of H(8)

Accordingly, horizontal and vertical calculations of 2D WHT are:

$$C = H(8)\,SH^{t}\,(8)$$

$$= ([A]\,[A]\,[A]\,([A]\,[A]\,[A]\,S)^{t})^{t}$$

These can be implemented using α and a transpose operation β as shown in Figure IV.3.

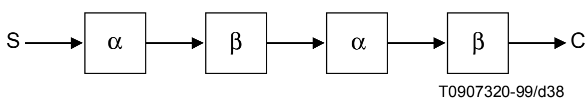


Figure IV.3/J.88 – Signal processing of FHT

Appendix V

VLC patterns for WHT coefficients

Regarding VLC (refer to A.2.7) and VLD (refer to A.3.1) for WHT coefficients, the following patterns must be used.

VLC1 for WHT coefficients (Mix B)

Variable length code	run	level
10	EOB	
0000 01	ESC	
11s	0	1
010s	0	2
0010 0s	0	3
0001 01s	0	4
0001 000s	0	5
0000 1101 0s	0	6
0000 1100 1s	0	7
0000 1100 0s	0	8
0000 1010 01s	0	9
0000 1001 01s	0	10
0000 0011 00s	0	11
0000 0010 11s	0	12
0000 0010 10s	0	13
0000 0010 011s	0	14
0000 0010 010s	0	15
0000 0010 001s	0	16
0000 0001 101s	0	17
0000 0000 1010 s	0	18
0000 0000 0100 s	0	19
0000 0000 0000 1s	0	20
0000 0000 0000 0s	0	21
011s	1	1
0001 11s	1	2
0000 1011 1s	1	3
0000 1000 11s	1	4
0000 1000 01s	1	5
0000 0001 111s	1	6
0000 0000 111s	1	7
0000 0000 0011 s	1	8

VLC1 for WHT coefficients (Mix B) *(cont.)*

Variable length code	run	level
0011 s	2	1
0000 1101 1s	2	2
0000 1001 10s	2	3
0000 1000 00s	2	4
0000 0001 011s	2	5
0000 0001 001s	2	6
0010 1s	3	1
0000 1010 00s	3	2
0000 0010 000s	3	3
0000 0000 1000 s	3	4
0000 0000 0010 s	3	5
0001 10s	4	1
0000 1001 00s	4	2
0000 0000 0111 s	4	3
0000 0000 0110 s	4	4
0001 001s	5	1
0000 0011 11s	5	2
0000 0000 0101 s	5	3
0000 111s	6	1
0000 0011 01s	6	2
0000 0000 0001 s	6	3
0000 1011 0s	7	1
0000 0000 1100 s	7	2
0000 1010 1s	8	1
0000 0000 1011 s	8	2
0000 1001 11s	9	1
0000 1000 10s	10	1
0000 0011 10s	11	1
0000 0001 110s	12	1
0000 0001 100s	13	1
0000 0001 010s	14	1
0000 0001 000s	15	1
0000 0000 1101 s	16	1
0000 0000 1001 s	17	1
10	EOB	
0000 01	ESC	
11s	0	1
0011 s	0	2

VLC1 for WHT coefficients *(cont.)*

Variable length code	run	level
0001 000s	0	3
0000 0011 1s	0	4
0000 0001 110s	0	5
0000 0000 1011 s	0	6
0000 0000 0011 0s	0	7
011s	1	1
0001 11s	1	2
0000 1010 s	1	3
0000 0010 010s	1	4
0000 0000 1100 s	1	5
0000 0000 0001 1s	1	6
0101 s	2	1
0001 010s	2	2
0000 0010 101s	2	3
0000 0001 0001 s	2	4
0000 0000 0010 0s	2	5
0100 s	3	1
0000 110s	3	2
0000 0001 101s	3	3
0000 0000 1010 s	3	4
0010 1s	4	1
0000 1000 0s	4	2
0000 0001 0010 s	4	3
0000 0000 0000 0s	4	4
0010 0s	5	1
0000 0010 110s	5	2
0000 0000 1001 s	5	3
0001 10s	6	1
0000 0010 100s	6	2
0000 0000 1000 s	6	3
0001 011s	7	1
0000 0010 000s	7	2
0000 0000 0101 1s	7	3
0001 001s	8	1
0000 0001 111s	8	2
0000 0000 0101 0s	8	3
0000 111s	9	1
0000 0000 1110 s	9	2

VLC1 for WHT coefficients *(concluded)*

Variable length code	run	level
0000 1011 s	10	1
0000 0000 1101 s	10	2
0000 1001 s	11	1
0000 0000 0111 s	11	2
0000 1000 1s	12	1
0000 0000 0011 1s	12	2
0000 0011 0s	13	1
0000 0000 0010 1s	13	2
0000 0010 111s	14	1
0000 0010 011s	15	1
0000 0010 001s	16	1
0000 0001 100s	17	1
0000 0001 011s	18	1
0000 0001 010s	19	1
0000 0001 0011 s	20	1
0000 0001 0000 s	21	1
0000 0000 1111 s	22	1
0000 0000 0110 1s	23	1
0000 0000 0110 0s	24	1
0000 0000 0100 1s	25	1
0000 0000 0100 0s	26	1
0000 0000 0001 0s	27	1
0000 0000 0000 1s	28	1

ITU-T RECOMMENDATIONS SERIES

Series A	Organization of the work of the ITU-T
Series B	Means of expression: definitions, symbols, classification
Series C	General telecommunication statistics
Series D	General tariff principles
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	**Transmission of television, sound programme and other multimedia signals**
Series K	Protection against interference
Series L	Construction, installation and protection of cables and other elements of outside plant
Series M	TMN and network maintenance: international transmission systems, telephone circuits, telegraphy, facsimile and leased circuits
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks and open system communications
Series Y	Global information infrastructure
Series Z	Languages and general software aspects for telecommunication systems