Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG
Geneva, CH, May, 2003

Document JVT-H036
File: JVT-H036.doc
Generated: 2003-08-30

Title: **Joint Model Reference Encoding Methods and Decoding Concealment Methods**

Status: Approved Output Document

Purpose: Draft of Joint Model

Author(s) or Contact(s):

Gary Sullivan
Microsoft Corporation
One Microsoft Way
Redmond, WA 98052 USA
Tel: +1 (425) 703-5308
Fax: +1 (425) 706-7329
Email: garysull@microsoft.com

Thomas Wiegand
Heinrich Hertz Institute (FhG),
Einsteinufer 37, D-10587 Berlin,
Germany
Tel: +49 - 30 - 31002 617
Fax: +49 - 30 - 392 72 00
Email: wiegand@hhi.de

Keng-Pang Lim
Institute for Infocomm Research
20, Heng Mui Keng Terrace
Singapore 119613
Tel: +65 6874 4303
Fax: +65 6774 4998
Email: kplim@i2r.a-star.edu.sg

Source: Editor

DRAFT INTERNATIONAL STANDARD
DRAFT ISO/IEC 14496-5 / PDAM6: 2003 (E)
DRAFT ITU-T Rec. H.xxx (2003 E)
DRAFT ITU-T RECOMMENDATION

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Foreword

Reference software and descriptions of reference encoding methods and non-normative reference decoding error concealment methods are useful in aiding users of a video coding standard to establish and test conformance and interoperability, and to educate users and demonstrate the capabilities of the standard. In this spirit, the methods described herein and the accompanying software modules are provided for implementers of ITU-T Recommendation H.264 | ISO/IEC International Standard ISO/IEC 14496-10 advanced video coding.

Introduction

ITU-T Recommendation H.264 | ISO/IEC International Standard ISO/IEC 14496-10 advanced video coding was developed as a joint project of the ITU-T SG16 Q.6 Video Coding Experts Group (VCEG) and the ISO/IEC JTC1/SC29/WG11 Moving Picture Experts Group (MPEG), with completion of the text of the video coding standard in 2003. Reference software was also developed jointly for approval soon thereafter, and the reference software accompanies this document. In addition to implementing the normative decoding process specified in the text of ITU-T Rec. H.264 | ISO/IEC 14496-10, software is provided to demonstrate effective (non-normative) encoding techniques and (non-normative) decoding error concealment techniques for use with the standard. These non-normative aspects are described in this document.

JOINT MODEL FOR NON-NORMATIVE ASPECTS OF ADVANCED VIDEO CODING

1 Scope

This document specifies non-normative reference encoding methods and methods of concealing errors and losses in decoders for video data conforming to ITU-T Recommendation H.264 | ISO/IEC International Standard ISO/IEC 14496-10 advanced video coding.

2 Non-normative example encoding methods

2.1 Motion estimation and mode decision

2.1.1 Low-complexity mode

2.1.1.1 Finding optimum prediction mode

Both for intra prediction and motion compensated prediction, a similar loop as indicated in QQ is run through. The different elements will be described below.

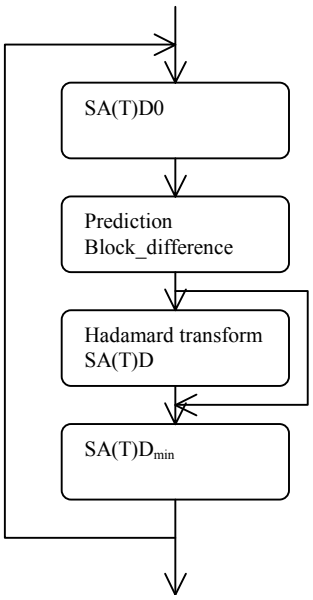

Figure 2-1 – Loop for prediction mode decision

2.1.1.1.1 SA(T)D0

The SA(T)D to be minimised is given a 'bias' value SA(T)D0 initially in order to favour prediction modes that need few bits to be signalled. This bias is basically a parameter representing bit usage times $QP_0(QP)$

Intra mode decision:

$$\mathbf{SA(T)D0 = QP_0(QP)xOrder\_of\_prediction\_mode \quad (see\ above)} \qquad (2\text{-}1)$$

Motion vector search:

$$SA(T)D0 = QP_0(QP)x(Bits\_to\_code\_vector + 2xcode\_number\_of\_ref\_idx\_fwd) \qquad (2\text{-}2)$$

In addition there are two special cases:

- For motion prediction of a 16x16 block with 0 vector components, $16xQP_0(QP)$ is subtracted from SA(T)D to favour the skip mode.
- For the whole intra 4x4 macroblock, $24xQP_0(QP)$ is added to the SA(T)D before comparison with the best SA(T)D for inter prediction. This is an empirical value to prevent using too many intra blocks.

For flat regions having zero motion, B pictures basically fail to make effective use of zero motion and instead are penalized in performance by selecting 16x16 intra mode. Therefore, in order to prevent assigning 16x16 intra mode to a region with little details and zero motion, SA(T)D of direct mode is subtracted by 16xQP0(QP) to bias the decision toward selecting the direct mode.

The calculation of SA(T)D0 at each mode should be as follows.

- Forward prediction mode:

$$SA(T)D0 = QP_0(QP) \times (2 \times code\_number\_of\_ref\_idx\_fwd + Bits\_to\_code\_MVDFW) \quad (2\text{-}3)$$

- Backward prediction mode:

$$SA(T)D0 = QP_0(QP) \times Bits\_to\_code\_MVDBW \quad (2\text{-}4)$$

- Bi-directional prediction mode:

$$SA(T)D0 = QP_0(QP) \times (2 \times code\_number\_of\_ref\_idx\_fwd + Bits\_to\_code\_forward\_Blk\_size + Bits\_to\_code\_backward\_Blk\_size + Bits\_to\_code\_MVDFW + Bits\_to\_code\_MVDBW) \quad (2\text{-}5)$$

- Direct prediction mode:

$$SA(T)D0 = -16 \times QP_0(QP) \quad (2\text{-}6)$$

- Intra 4x4 mode:

$$SA(T)D0 = 24 \times QP0(QP) \quad (2\text{-}7)$$

- Intra 16x16 mode:

$$SA(T)D0 = 0 \quad (2\text{-}8)$$

2.1.1.1.2 Block_difference

For the whole block the difference between the original and prediction is produced

$$Diff(i,j) = Original(i,j) - Prediction(i,j) \quad (2\text{-}9)$$

2.1.1.1.3 Hadamard transform

For integer sample search (see below) we use SAD based on Diff(i,j) for decision. Hence no Hadamard is done and we use SAD instead of SATD.

$$SAD = \sum_{i,j} \left| Diff(i,j) \right| \quad (2\text{-}10)$$

However, since we will do a transform of Diff(i, j) before transmission, we will do a better optimisation if a transform is done before producing SAD. Therefore a two dimensional transform is performed in the decision loop for selecting intra modes and for fractional sample search (see below). To simplify implementation, the Hadamard transform is chosen in this mode decision loop. The relation between samples and basis vectors (BV) in a 4 point Hadamard transform is illustrated below (not normalized):

[Ed. Note: Convert to a figure.]

Samples →

B 1 1 1 1

V 1 1 -1 -1

↓ 1 -1 -1 1

1 -1 1 -1

This transformation is performed horizontally and vertically and result in DiffT(i, j). Finally SATD for the block and for the present prediction mode is produced.

$$SATD = (\sum_{i,j} |DiffT(i,j)|) / 2 \tag{2-11}$$

2.1.1.1.4 Mode decision

Choose the prediction mode that results in the minimum value $SA(T)D_{min} = \min(SA(T)D+SA(T)D0)$.

2.1.1.2 Encoding on macroblock level

2.1.1.2.1 Intra coding and fast intra coding

2.1.1.2.1.1 High-quality intra coding

When starting to code a macroblock, intra mode is checked first. For each 4x4 block, full coding indicated in QQ is performed. At the end of this loop the complete macroblock is intra coded and a SATDintra is calculated.

2.1.1.2.1.2 Fast intra coding

Fast intra mode prediction algorithm is implemented based on the local edge direction information. By determining the edge direction of the current block to be coded, only a few most probable INTRA modes are selected for RDO computations.

Prior to intra prediction, Sobel edge operators are applied to the current image to generate the edge map, where each pixel is associated with an edge vector containing its edge direction and amplitude. For a pixel $p_{i,j}$, in a luminance (or chrominance) picture, the corresponding edge vector, $\vec{D}_{i,j} = \{dx_{i,j}, dy_{i,j}\}$, is defined as,

$$\begin{aligned} dx_{i,j} &= p_{i-1,j+1} + 2\times p_{i,j+1} + p_{i+1,j+1} - p_{i-1,j-1} - 2\times p_{i,j-1} - p_{i+1,j-1} \\ dy_{i,j} &= p_{i+1,j-1} + 2\times p_{i+1,j} + p_{i+1,j+1} - p_{i-1,j-1} - 2\times p_{i-1,j} - p_{i-1,j+1} \end{aligned} \tag{2-12}$$

where *i, = 0, 1, ..., N, j = 0, 1, ..., M*, and *N×M* is the image dimension.

The amplitude of the edge vector is decided by,

$$Amp(\vec{D}_{i,j}) = |dx_{i,j}| + |dy_{i,j}| \tag{2-13}$$

The direction of the edge (in degree) is decided by the hyper-function,

$$Ang(\vec{D}_{i,j}) = \frac{180^0}{\pi} \times \arctan\left(\frac{dy_{i,j}}{dx_{i,j}}\right) \tag{2-14}$$

As there is only limited number of modes that intra prediction could be applied, a thresholding technique is used to substitute Equation (2-14) in the actual implementation of the algorithm. Therefore the edge direction histogram of a 4×4 luma block is decided by the followings,

For all the pixels $p_{i,j}$ in the 4×4 *luma block,*

$$Histo(k) = \sum_{(i,j)\in SET(k)} Amp(\vec{D}_{i,j}),$$

where,

$$SET(k) \in \{(i,j) \mid Ang(\vec{D}_{i,j}) \in a_k\},$$

and

$$\begin{aligned} a_0 &= (-103.3^0, -76.6^0] \\ a_1 &= (-13.3^0, 13.3^0] \\ a_3 &= (35.8^0, 54.2^0] \\ a_4 &= (-54.2^0, -35.8^0] \\ a_5 &= (-76.7^0, -54.2^0] \\ a_6 &= (-35.8^0, -13.3^0] \\ a_7 &= (54.2^0, -76.7^0] \\ a_8 &= (13.3^0, 35.8^0] \end{aligned} \tag{2-15}$$

Similarly, the edge direction histogram of an 8×8 luma block or 16×16 chroma block is decided as follows,

For *all the pixels $p_{i,j}$ in the* 8×8 *chroma block or* 16×16 *luma block,*

$$Histo(k) = \sum_{(i,j)\in SET(k)} Amp(\vec{D}_{i,j}),$$

where,

$$SET(k) \in \{(i,j) \mid Ang(\vec{D}_{i,j}) \in a_k\}, \tag{2-16}$$

and

$$\begin{aligned} a_0 &= (-112.5^0, -67.5^0] \\ a_1 &= (-22.5^0, 22.5^0] \\ a_3 &= (-67.5^0, -22.5^0] \end{aligned}$$

Figure 2-2 shows an example of the edge direction histogram of a 4×4 luma block.

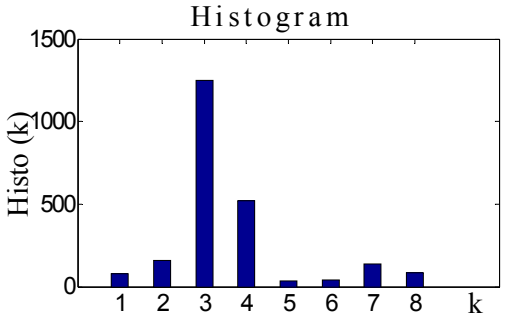

Figure 2-2 Edge direction histogram

According to (G-4) and (G-5), each cell in the edge direction histogram sums up the amplitudes of the edge with the similar direction in that block. Since the pixels along edge direction are likely to have similar values, the best prediction mode is probably in the edge direction whose cell has the maximum amplitude, or the directions close to the maximum amplitude cell.

2.1.1.2.1.2.1 4×4 luma block prediction modes

The histogram cell with the maximum amplitude and the two adjacent cells are considered as candidates for the best prediction mode. In addition, DC mode is chosen as the fourth candidate. Thus, for each 4×4 luma block, only 4 prediction modes are performed.

2.1.1.2.1.2.2 16×16 luma block prediction modes

Only the histogram cell with maximum amplitude is considered as a candidate of best prediction mode. In addition, DC mode is chosen as the next candidate. Thus, for each 16×16 luma block, only 2 prediction modes are performed.

2.1.1.2.1.2.3 8×8 chroma prediction modes

In the case of chroma blocks, there are two different histograms, one from component U and the other from V. The histogram cells with maximum amplitude from the two components are both considered as candidate modes. In addition, DC mode is chosen as another candidate. Note that if the candidates from the two components are the same, there is 2 candidate prediction modes; otherwise, it is 3.

Table 2-1 summarizes the number of candidates selected for RDO calculation based on edge direction histogram.

Table 2-1Number of selected modes

	Block size	Total No. of modes	No. of modes selected
Luma (Y)	4×4	9	4
Luma (Y)	16×16	4	2
Chroma (U, V)	8×8	4	3 or 2

2.1.1.2.2 Table for intra prediction modes to be used at the encoder side

QQ gives the table of intra prediction modes according to probability of each mode to be used on the decoder side. On the encoder side we need a sort of inverse table. Prediction modes for A and B are known as in QQ. For the encoder we have found a Mode that we want to signal with an ordering number in the bitstream (whereas on the decoder we receive the order in the bitstream and want to convert this to a mode). QQ is therefore the relevant table for the encoder. Example: Prediction mode for A and B is 2. The string in QQ is 2 1 0 3 4 5. This indicates that prediction mode 0 has order 2 (third most probable). Prediction mode 1 is second most probable and prediction mode 2 has order 0 (most probable) etc. As in QQ '–' indicates that this instance can not occur because A or B or both are outside the picture.

[Ed. Note: Isn't this table already in the normative section?]

Table 16

Prediction ordering to be used in the bitstream as a function of prediction mode (see text).

B\A	outside	0	1	2	3	4	5
outside	0-----	021---	102---	120---	012---	012---	012---
0	0---12	025314	104325	240135	143025	035214	045213
1	0---12	014325	102435	130245	032145	024315	015324
2	0---12	012345	102345	210345	132045	032415	013245
3	0---12	135024	214035	320154	143025	145203	145032
4	1---02	145203	125403	250314	245103	145203	145302
5	1---20	245310	015432	120534	245130	245301	135420

2.1.1.2.3 Inter mode selection

Next motion vector search is performed for motion compensated prediction. A search is made for all 7 possible block structures for prediction as well as from the 5 past decoded pictures. This result in 35 combinations of blocksizes and reference frames. For B frames, the motion search is also conducted for the temporal following reference picture to obtain backward motion vectors.

2.1.1.2.4 Integer sample search

Motion search range decision plays as a role of a pre-processing step of a motion estimation to reduce encoder complexity. Search range is differently determined at each macro-block by using its neighboring information.

Let assume that E is a macro-block under motion estimation, and A, B, C are its neighboring 4×4 blocks, as shown in Fig. 1. Also, assume that motion vectors of A, B, and C are (MV_A_x, MV_A_y), (MV_B_x, MV_B_y), and (MV_C_x, MV_C_y), and input search range is input_search_range. The first step of the motion search range is to estimate the local maximum vector of E. It is defined as

$$\begin{aligned} \text{max\_MV\_E\_x} &= \max(\text{abs}(\text{MV\_A\_x}), \max(\text{abs}(\text{MV\_B\_x}), \text{abs}(\text{MV\_C\_x})) \\ \text{max\_MV\_E\_y} &= \max(\text{abs}(\text{MV\_A\_y}), \max(\text{abs}(\text{MV\_B\_y}), \text{abs}(\text{MV\_C\_y})) \end{aligned} \tag{2-17}$$

In the above equation, if a neighboring block is outside of image, the motion vector is replaced by input search range. The second step is to determine temporal search range by the local statistics as

$$\begin{aligned} \text{local\_search\_range\_x} &= \max(\text{k\_x},\ 2 \times \text{max\_MV\_E\_x}) \\ \text{local\_search\_range\_y} &= \max(\text{k\_y},\ 2 \times \text{max\_MV\_E\_y}) \end{aligned} \tag{2-18}$$

where k_x and k_y are determined as

$$\begin{aligned} \text{k\_i} &= \begin{cases} (\text{input\_search\_range} + 2)/4 & \text{if } \alpha_i \geq 2 \\ (\text{input\_search\_range} + 4)/8 & \text{otherwise} \end{cases} \\ \alpha_{\text{i}} &= \text{abs}(\text{MV\_A\_i}) + \text{abs}(\text{MV\_B\_i}) + \text{abs}(\text{MV\_C\_i}) \qquad \text{for } \text{i} = \text{x}, \text{y} \end{aligned} \tag{2-19}$$

The maximum search range is subject to input search range, and therefore, the motion search range is constrained as

$$\begin{aligned} \text{new\_search\_range\_x} &= \min(\text{input\_search\_range}, \text{local\_search\_range\_x}) \\ \text{new\_search\_range\_y} &= \min(\text{input\_search\_range}, \text{local\_search\_range\_y}) \end{aligned} \tag{2-20}$$

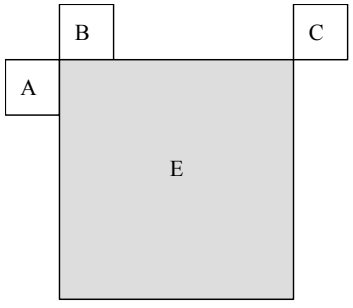

Figure 2-3 – Reference block location for motion search range

2.1.1.2.5 Fractional sample search

Fractional sample search is performed in two steps. This is illustrated below where capital letters represent integer positions, numbers represent ½ sample positions and lower case letters represent ¼ sample positions.

[Ed. Note: Convert to a figure]

A B C

1 2 3

D 4 E 5 F

a b c

6 d 7 e 8

f g h

G H I

Assume that the integer search points to position E. Then ½ sample positions 1, 2, 3, 4, 5, 6, 7, 8 are searched. Assume that 7 is the best position. Then the ¼ sample positions a, b, c, d, e, f, g, h are searched. (Notice that by this procedure a position with more low pass filtering – see subclause QQ – is automatically checked). If motion compensation with 1/8 sample accuracy is used, an additional sub-sample refinement step is performed in the described way. After fractional sample search has been performed for the complete macroblock, the SATD for the whole macroblock is computed: SATDinter.

2.1.1.2.6 Decision between intra and inter

If SATDintra < SATDinter intra coding is used. Otherwise inter coding is used.

2.1.2 High-complexity mode

2.1.2.1 Full-search motion Estimation

For each block or macroblock the motion vector is determined by full search on integer-sample positions followed by sub-sample refinement.

2.1.2.1.1 Integer-sample search

As in low-complexity mode, the search positions are organized in a spiral structure around a prediction vector. The full search range given by MC_range is used for all INTER-modes and reference frames. To speed up the search process , the prediction vector of the 16x16 block is used as center of the spiral search for all INTER-modes. Thus the SAD values for 4x4 blocks can be pre-calculated for all motion vectors of the search range and then used for fast SAD calculation of all larger blocks. The search range is not forced to contain the (0, 0)-vector.

2.1.2.1.2 Fractional sample search

The fractional sample search is performed as in the low-complexity case.

2.1.2.1.3 Finding the best motion vector

The integer-sample motion search as well as the sub-sample refinement returns the motion vector that minimizes

$$J(\mathbf{m}, \lambda_{MOTION}) = SA(T)D(s, c(\mathbf{m})) + \lambda_{MOTION} \cdot R(\mathbf{m} - \mathbf{p}) \tag{2-21}$$

with $\mathbf{m} = (m_x, m_y)^T$ being the motion vector, $\mathbf{p} = (p_x, p_y)^T$ being the prediction for the motion vector, and λ_{MOTION} being the Lagrange multiplier. The rate term $R(\mathbf{m} - \mathbf{p})$ represents the motion information only and is computed by a table-lookup. The rate is estimated by using the universal variable length code (UVLC) table, even if CABAC is used as entropy coding method. For integer-sample search, is used as distortion measure. It is computed as

$$SAD(s,\ c(\mathbf{m})) = \sum_{x=1,\ y=1}^{B,\ B} \left| s[x, y] - c[x - m_x, y - m_y] \right|, \tag{2-22}$$

with s being the original video signal and c being the coded video signal. In the sub-sample refinement search, the distortion measure SATD is calculated after a Hadamard transform (see section QQ). The Lagrangian multiplier λ_{MOTION} is given by

$$\lambda_{MOTION,P} = \sqrt{0.85 \times 2^{QP/3}} \tag{2-23}$$

for I and P frames and

$$\lambda_{MOTION,B} = \sqrt{\max\left(2, \min(4, \frac{QP}{6})\right) \times \lambda_{I,P}} \tag{2-24}$$

for B frames, where *QP* is the macroblock quantisation parameter.

2.1.2.1.4 Finding the best reference frame

The determination of the reference frame REF and the associated motion vectors for the NxM inter modes is done after motion estimation by minimizing

$$J(REF \mid \lambda_{MOTION}) = SATD(s, c(REF, \mathbf{m}(REF))) + \lambda_{MOTION} \cdot (R(\mathbf{m}(REF) - \mathbf{p}(REF)) + R(REF)) \tag{2-25}$$

The rate term $R(REF)$ represents the number of bits associated with choosing *REF* and is computed by table-lookup using UVLC.

2.1.2.1.5 Finding the best prediction direction for B frames

The determination of the prediction direction PDIR for the NxM inter modes in B frames is done after motion estimation and reference frame decision by minimizing

$$\begin{aligned} J(PDIR \mid \lambda_{MOTION}) = {} & SATD(s, c(PDIR, \mathbf{m}(PDIR))) + \\ & \lambda_{MOTION} \times (R(\mathbf{m}(PDIR) - \mathbf{p}(PDIR)) + R(REF(PDIR))) \end{aligned} \tag{2-26}$$

2.1.2.2 Fast search motion estimation

A fast integer sample motion estimation and sub-sample motion estimation are adopted to speed up the motion estimation progress.

2.1.2.2.1 Integer-sample search

The search range given by MC_range is used for all INTER-modes and reference frames. To speed up the search process, a hexagon based search strategy is used together with an early termination technique.

2.1.2.2.1.1 Fast integer-sample search strategy

There are four steps in the proposed fast integer sample search algorithm:

Step_1: Initial search point prediction

A motion prediction set $\mathcal{S}$ is defined including all the candidate predictors described in section 2.1.2.2.1.2, then:

1) $\mathbf{m}_{\min} = \arg[\min_{\mathbf{m}_i} J(\mathbf{m}_i, \lambda_{MOTION})], s.t. \mathbf{m}_i \in \mathcal{S}$ (*s.t means" subject to"*)

2) Early_Termination(refer 2.1.2.2.1.3.2)

Step_2: Unsymmetrical-cross search

An unsymmetrical-cross search pattern is defined by set Ω_1, which is:

$\Omega_1 = \{\mathbf{m} = (m_x, m_y)^T \mid \mathbf{m} = (cm_x \pm 2i, cm_y)^T, i = 0,1,2,....,\frac{W}{2}; \mathbf{m} = (cm_x, cm_y \pm 2j)^T, j = 1,2,....,\frac{W}{4}\}$, where $\mathbf{cm} = \mathbf{m}_{\min}$, and W is the search range. Then:

1) $\mathbf{m}_{\min 2} = \arg[\min_{\mathbf{m}_i} J(\mathbf{m}_i, \lambda_{MOTION})], s.t. \mathbf{m}_i \in \Omega_1$

2) $\mathbf{m}_{\min} = \arg[\min(J(\mathbf{m}_{\min}, \lambda_{MOTION}), J(\mathbf{m}_{\min 2}, \lambda_{MOTION}))]$

3) Early_Termination(refer 2.1.2.2.1.3.2)

Step_3: Uneven Multi-Hexagon-grid Search

Two sub-steps in this search step:

Step_3-1:

A square search pattern with search range equals 2 is defined by Ω_2, which is:

$\Omega_2 = \{\mathbf{m} = (m_x, m_y)^T \,||\, m_x - cm_x \,|\leq 2, |\, m_y - cm_y \,|\leq 2\}$, where $\mathbf{cm} = \mathbf{m_{min}}$. Then:

1) $\mathbf{m}_{\min 3} = \arg[\min_{\mathbf{m}_i} J(\mathbf{m}_i, \lambda_{MOTION})], s.t. \mathbf{m}_i \in \Omega_2$

2) $\mathbf{m}_{\min} = \arg[\min(J(\mathbf{m}_{\min}, \lambda_{MOTION}), J(\mathbf{m}_{\min 3}, \lambda_{MOTION}))]$

3) Early_Termination(refer 2.1.2.2.1.3.2)

Step_3-2:

A 16 points hexagon search pattern is defined by Ω_{16-HP}, which is:

$\Omega_{16-HP} = \{\mathbf{m} = (x, y)^T \,|\mathbf{m} = (\pm 4, \pm 2)^T, (\pm 4, \pm 1)^T, (\pm 4, 0)^T, (\pm 2, \pm 3)^T, (0, \pm 4)^T\}$. By extending the search pattern with a scale factor *k*, a multi-hexagon-grid search scheme is taken.

for k= 1 k< W/4 k++

{

Search pattern after scaling with factor k is Π_k, which is:

$\Pi_k = \{\mathbf{m} = (m_x, m_y) \,|\, m_x = cm_x + k \bullet x', m_y = cm_y + k \bullet y', (x', y') \in \Omega_{16-HP}\}$, then:

1) $\mathbf{m}_{\min k} = \arg[\min_{\mathbf{m}_i} J(\mathbf{m}_i, \lambda_{MOTION})], s.t. \mathbf{m}_i \in \Pi_k$

2) $\mathbf{m}_{\min} = \arg[\min(J(\mathbf{m}_{\min}, \lambda_{MOTION}), J(\mathbf{m}_{\min k}, \lambda_{MOTION}))]$

3) Early_Termination(refer 2.1.2.2.1.3.2)

}

Step_4: Extended Hexagon-based Search

Two sub-steps in this search step:

Step_4-1:

A hexagon search pattern is defined by Ω_3, which is:

$\Omega_3 = \{\mathbf{m} = (m_x, m_y)^T \,|\mathbf{m} = (cm_x \pm 2, cm_y)^T, (cm_x \pm 1, cm_y \pm 2)^T\}$, where $\mathbf{cm} = \mathbf{m_{min}}$. Then:

1) $\mathbf{m}_{\min 4} = \arg[\min_{\mathbf{m}_i} J(\mathbf{m}_i, \lambda_{MOTION})], s.t. \mathbf{m}_i \in \Omega_3$

2) $\mathbf{m}_{\min} = \arg[\min(J(\mathbf{m}_{\min}, \lambda_{MOTION}), J(\mathbf{m}_{\min 4}, \lambda_{MOTION}))]$

3) if $\mathbf{m}_{\min} == \mathbf{cm}$, goto Step_4-2

else goto Step_4-1

Step_4-2:

A diamond search pattern is defined by Ω_4, which is:

$\Omega_4 = \{\mathbf{m} = (m_x, m_y)^T \,|\mathbf{m} = (cm_x \pm 1, cm_y)^T, (cm_x, cm_y \pm 1)^T\}$, where $\mathbf{cm} = \mathbf{m_{min}}$. Then:

1) $\mathbf{m}_{\min 5} = \arg[\min_{\mathbf{m}_i} J(\mathbf{m}_i, \lambda_{MOTION})], s.t. \mathbf{m}_i \in \Omega_4$

2) $\mathbf{m}_{\min} = \arg[\min(J(\mathbf{m}_{\min}, \lambda_{MOTION}), J(\mathbf{m}_{\min 5}, \lambda_{MOTION}))]$

3) if $\mathbf{m}_{\min} == \mathbf{cm}$, stop

else goto Step_4-2

2.1.2.2.1.2 Motion prediction set for fast search

Four prediction modes are used to predict current block's motion vector, they are:

1) Median Prediction:

mvLXN is the motion vector (with N being either the neighboring partition A, B, or C, see Figure 2-3) of the neighboring partitions.

refIdxLXN is the reference indices (with N being either A, B, or C) of the neighboring partitions, and refIdxLX is the reference index of the current partition.

mvMpLX is the median motion prediction of current partition.

The following rules are applied in sequential order to determine the motion prediction mvMpLX:

- If both partitions B and C are not available and A is available, then mvLXB = mvLXA and mvLXC = mvLXA. As well as refIdxLXB = refIdxLXA and refIdxLXC = refIdxLXA.

- If one and only one of the reference indices refIdxLXA, refIdxLXB and refIdxLXC is equal to the reference index refIdxLX of the current partition, the following applies. Let refIdxLXN be the reference index that is equal to refIdxLX, then the motion vector mvLXN is assigned to the motion prediction mvMpLX: mvMpLX = mvLXN.

- Otherwise, each component of the motion prediction mvMpLX is given by the median of the corresponding vector components of the motion vector mvLXA, mvLXB, and mvLXC:

mvMpLX [0] = Median(mvLXA[0], mvLXB[0], mvLXC[0])

mvMpLX[1] = Median(mvLXA[1], mvLXB[1], mvLXC[1])

2) Uplayer prediction

The recommendation supports motion compensation block size ranging from 16x16 to 4x4 luminance samples, and these block partition modes can be named from Mode1 to Mode7.

Assuming the prediction mode of current partition is $Mode_{curr}$, the uplayer mode $Mode_{Up}$ is:

$$Mode_{Up} = \begin{cases} unavailable, & if \quad Mode_{curr} == Mode1 \\ Mode1, & if \quad Mode_{curr} == Mode2, Mode3 \\ Mode2, & if \quad Mode_{curr} == Mode4 \\ Mode4, & if \quad Mode_{curr} == Mode5, Mode6 \\ Mode5, & if \quad Mode_{curr} == Mode7 \end{cases}$$

Then if refIdxLX == refIdxLXU, the uplayer predicted vector is:

mvUpLX =mvLXU

where mvLXU and refIdxLXU are the motion vector and the reference indice of the uplayer partition respectively, and refIdxLX is the reference index of the current partition.

3) Corresponding Prediction

Where corresponding partition is the partition in the previous reconstructed frame and this partition has the same address as that of the current partition.

Then is, if refIdxLX ==refIdxLXCP, the corresponding predicted vector of current partion is: mvCpLX = mvLXCP, where mvLXCP is the motion vector of the corresponding partition; refIdxLXCP is the reference indices of the corresponding partition; and refIdxLX is the reference index of the current partition.

4) Neighboring Ref-frame Prediction

mvNrpLX is defined as the motion pediction of current partition, and mvNrpLX is not available while refIdxLX<1, otherwise mvNrpLX = mvLXNRP*(refIdxLX+1)/(refIdxLXNRP+1), where, refIdxLX is the reference index of the current partition; mvLXNRP is the motion vector of the current partition when reference index refIdxLXNRP==refIdxLX-1.

The final motion prediciton set is based on the above four prediction 1), 2), 3), and 4)

Let Ψ be a set:

$$\Psi(CMV) = \{\mathbf{MV} = (MV[0], MV[1])^T \mid \mathbf{MV} = (CMV[0] \pm 1, CMV[1])^T, (CMV[0], CMV[1] \pm 1)^T\},$$

Then set:

$$\mathcal{S}_1 = \{\text{mvMpLX}, \text{mvUpLX}, \text{mvCpLX}, \text{mvNrpLX}, \overrightarrow{(0,0)}, \Psi(\text{mvMpLX}), \Psi(\overrightarrow{(0,0)})\}$$

Let $\mathbf{m}_{S\min} = \arg[\min_{\mathbf{m}_i} J(\mathbf{m}_i, \lambda_{MOTION})], s.t. \mathbf{m}_i \in \mathcal{S}_1$, the motion prediction set is:

$$\mathcal{S} = \mathcal{S}_1 + \Psi(\mathbf{m}_{S\min})$$

2.1.2.2.1.3 Early_termination

2.1.2.2.1.3.1 Cost prediction of the current paritition

Three cost prediction modes are used to predict current partition's cost, they are:

1) Median Prediction:

- If both partitions B and C are not available and A is available, then csLXB = csLXA and csLXC = csLXA, as well as refIdxLXB = refIdxLXA and refIdxLXC = refIdxLXA.
  - If one and only one of the reference indices refIdxLXA, refIdxLXB and refIdxLXC is equal to the reference index refIdxLX of the current partition, the following applies. Let refIdxLXN, with N being A, B, or C, be the reference index that is equal to refIdxLX. Then the cost csLXN, with N being the correspondent A, B or C, is assigned to the cost prediction: J_{pred_MP} = csLXN.
  - Otherwise, each component of the motion prediction is given by the median of the corresponding vector components of the motion vector mvLXA, mvLXB, and mvLXC:

mvMpLX[0] = Median(mvLXA[0], mvLXB[0], mvLXC[0])

mvMpLX[1] = Median(mvLXA[1], mvLXB[1], mvLXC[1])

Let mvLXN (with N being either A, B, or C) be the motion vector that satisfies mvLXN[0]== mvMpLX[0], and let mvLXM(with M being either A, B, or C) be the motion vector that satisfies mvLXM[1]== mvMpLX[1]. Then the median predicted cost is: J_{pred_MP} = min(csLXN, csLXM)

2) Uplayer Prediction

If refIdxLX == refIdxLXU, the uplayer cost predictor is: J_{pred_UP} = csLXU/2, where, csLXU is the cost of the uplayer partition, refIdxLXU is the reference indices of the uplayer partition, index refIdxLX is the reference of the current partition.

3) Neighboring Ref-frame Prediction

J_{pred_NRP} is cost predictor of current partition. If refIdxLX<1, J_{pred_NRP} is not available, otherwise J_{pred_NRP} = csLXNRP, Where, refIdxLX is the reference index of the current partition, csLXNRP is the cost of the current partition when reference indices is refIdxLXNRP==refIdxLX-1.

2.1.2.2.1.3.2 Early termination

Based on above three cost predictors, the final cost predictor of current partition is:

$$J_{pred} = \begin{cases} J_{pred\_NRP}, & if \quad refIdxLX > 0 \\ J_{pred\_MP}, & if \quad refIdxLX == 0 \,\&\&\, Mode_{curr} == 1 \\ J_{pred\_UP}, & if \quad refIdxLX == 0 \,\&\&\, Mode_{curr} > 1 \end{cases}$$

The Early_Termination is described as following:

if $J(\mathbf{m}, \lambda_{MOTION})_{curr} - J_{pred} < J_{pred} * \beta_1$, then jump to a motion search strategy with small diamond search pattern to do a minor refinement, see section 2.1.2.2.1.1 step_4-2. And else if $J(\mathbf{m}, \lambda_{MOTION})_{curr} - J_{pred} < J_{pred} * \beta_2$, then jump to a motion search strategy with hexagon search pattern to do a preferable refinement, see section 2.1.2.2.1.1 step_4-1.

Where two thresholds β_1 and β_2 are set for each partition mode according to the error judgement possibility, and their values are chosen from the experiments. β_1 = {0.01, 0.01, 0.01, 0.02, 0.03, 0.03, 0.04} from Mode1 to Mode7 and β_2 = {0.06, 0.07, 0.07, 0.08, 0.12, 0.11, 0.15} from Mode1 to Mode7 are adopted in fast integer-sample search.

Assuming $J_{th1} = J_{pred} * (1 + \beta_1)$ and $J_{th2} = J_{pred} * (1 + \beta_2)$, the macro definition of Early_Termination in section 2.1.2.2.1.1 can be derived:

#define Early_Termination

if $J(\mathbf{m}_{\min}, \lambda_{MOTION}) < J_{th1}$ goto Step_4-2

else if $J(\mathbf{m}_{\min}, \lambda_{MOTION}) < J_{th2}$, goto Step_4-1

else continue

#endif

2.1.2.2.2 Sub-sample search

Assume $\mathbf{m}$ is the motion vector of the current partition after integer sample search, then $\mathbf{fm} = \mathbf{m} \times 4$, which is the motion vector in fractional sample unit after integer sample motion estimation.

The whole fast fractional sample motion estimation algorithm can be described in the following two steps:

Step_1. Initial search point prediction

Set the predicted fractional sample motion vector $\mathbf{fm}_{pred} = (\mathbf{fm}_{pred\_MP} - \mathbf{fm})\%4$, where $\mathbf{fm}_{pred\_MP} = \mathbf{m}_{pred\_MP} \times 4$ (see section 0) and % is the mod operation. Then:

$\mathbf{fm}_{\min} = \arg[\min_{\mathbf{m}_i} J(\mathbf{fm}_i, \lambda_{MOTION})], s.t. \mathbf{fm}_i \in T$, where $T = \{\mathbf{fm}_i \,|\, \mathbf{fm}_{pred}, (0,0)\}$.

Step_2. Diamond search strategy

For (i=0; i<7; i++)

{

Let diamond search pattern is: $\Omega_5 = \{\mathbf{fm}_i = (fm_{i,x}, fm_{i,y}) \,|\, \mathbf{fm}_i = (fcm_x \pm 1, fcm_y)^T, \mathbf{fm}_i = (fcm_x, fcm_y \pm 1)^T\}$,

where $\mathbf{fcm} = \mathbf{fm}_{\min}$, then

1) $\mathbf{fm}_{\min f} = \arg[\min_{\mathbf{m}_i} J(\mathbf{fm}_i, \lambda_{MOTION})], s.t. \mathbf{fm}_i \in \Omega_5$

2) $\mathbf{fm}_{\min} = \arg[\min(J(\mathbf{fm}_{\min f}, \lambda_{MOTION}), J(\mathbf{fm}_{\min}, \lambda_{MOTION}))]$

3) if ($\mathbf{fm}_{\min} == \mathbf{fcm}$) stop

}

2.1.2.2.3 For interlace frame and B frame cases

The fast search motion estimation can be done for interlace and B frame cases. For interlace case, motion estimation of current partition is done on the corresponding field of reference frames. For B frame case, when choosing motion prediction set and cost prediction set for backward prediction , Median Prediction mode and Uplayer Prediction mode is the same as that in forward prediction and corresponding partition defined in Corresponding Prediction mode is the partition in the previous reconstructed B frame for backward prediction case. Neighboring Ref-frame Prediction mode is not used for backward prediction.

2.1.2.2.4 Finding the best motion vector

This part is the same as Full Search Motion Estimation

2.1.2.2.5 Finding the best reference frame

This part is the same as Full Search Motion Estimation

2.1.2.2.6 Finding the best prediction direction for B frames

This part is the same as Full Search Motion Estimation

2.1.2.3 Mode decision

2.1.2.3.1 Macroblock mode decision

The macroblock mode decision is done by minimizing the Lagrangian functional

$$J(s,c,MODE \mid QP,\lambda_{MODE}) = SSD(s,c,MODE \mid QP) + \lambda_{MODE} \cdot R(s,c,MODE \mid QP) \tag{2-27}$$

where *QP* is the macroblock quantiser, λ_{MODE} is the Lagrange multiplier for mode decision, and *MODE* indicates a mode chosen from the set of potential prediction modes:

I frame: $$MODE \in \{INTRA\,4x4,\ INTRA16x16\}, \tag{2-28}$$

P frame: $$MODE \in \begin{Bmatrix} INTRA\,4x4,\ INTRA16x16,\ SKIP, \\ 16x16,\ 16x8,\ 8x16,\ 8x8 \end{Bmatrix}, \tag{2-29}$$

B frame: $$MODE \in \begin{Bmatrix} INTRA\,4x4,\ INTRA16x16,\ \ DIRECT, \\ 16x16,\ 16x8,\ 8x16,\ 8x8 \end{Bmatrix} \tag{2-30}$$

Note that the *SKIP* mode refers to the 16x16 mode where no motion and residual information is encoded. *SSD* is the sum of the squared differences between the original block *s* and its reconstruction *c* given as

$$\begin{aligned} SSD(s,c,MODE \mid QP) = & \sum_{x=1,y=1}^{16,16} \left(s_Y[x,y] - c_Y[x,y,MODE \mid QP]\right)^2 \\ & + \sum_{x=1,y=1}^{8,8} \left(s_U[x,y] - c_U[x,y,MODE \mid QP]\right)^2 + \sum_{x=1,y=1}^{8,8} \left(s_V[x,y] - c_V[x,y,MODE \mid QP]\right)^2, \end{aligned} \tag{2-31}$$

and $R(s,c,MODE \mid QP)$ is the number of bits associated with choosing *MODE* and *QP*, including the bits for the macroblock header, the motion, and all DCT blocks. $c_Y[x,y,MODE \mid QP]$ and $s_Y[x,y]$ represent the reconstructed and original luminance values; c_u, c_v, and s_u, s_v the corresponding chrominance values.

The Lagrangian multiplier λ_{MODE} is given by

$$\lambda_{MODE,P} = 0.85 \times 2^{QP/3} \tag{2-32}$$

for I and P frames and

$$\lambda_{MODE,B} = \max\left(2, \min(4, \frac{QP}{6})\right) \times \lambda_{MODE,P} \tag{2-33}$$

for B frames, where *QP* is the macroblock quantisation parameter.

2.1.2.3.2 8x8 mode decision

The mode decision for 8x8 sub-partitions is done similar to the macroblock mode decision by minimizing the Lagrangian functional

$$J(s,c,MODE \mid QP,\lambda_{MODE}) = SSD(s,c,MODE \mid QP) + \lambda_{MODE} \cdot R(s,c,MODE \mid QP) \tag{2-34}$$

where *QP* is the macroblock quantisation parameter, λ_{MODE} is the Lagrange multiplier for mode decision, and *MODE* indicates a mode chosen from the set of potential prediction modes:

P frame: $$MODE \in \begin{Bmatrix} INTRA\,4x4, \\ 8x8,\ 8x4,\ 4x8,\ 4x4 \end{Bmatrix} \tag{2-35}$$

B frame: $$MODE \in \begin{Bmatrix} INTRA\,4x4,\ DIRECT, \\ 8x8,\ 8x4,\ 4x8,\ 4x4 \end{Bmatrix} \tag{2-36}$$

2.1.2.3.3 INTER 16x16 mode decision

The INTER16x16 mode decision is performed by choosing the INTER16x16 mode which results in the minimum SATD value.

2.1.2.3.4 INTER 4x4 mode decision

For the Intra4x4 prediction, the mode decision for each 4x4 block is performed similar to the macroblock mode decision by minimizing

$$J(s,c,IMODE \mid QP,\lambda_{MODE}) = SSD(s,c,IMODE \mid QP) + \lambda_{MODE} \cdot R(s,c,IMODE \mid QP) \tag{2-37}$$

where *QP* is the macroblock quantiser, λ_{MODE} is the Lagrange multiplier for mode decision, and *IMODE* indicates an intra prediction mode:

$$IMODE \in \{DC,\ HOR,\ VERT,\ DIAG,\ DIAG\_RL,\ DIAG\_LR\} \tag{2-38}$$

SSD is the sum of the squared differences between the original 4x4 block luminance signal *s* and its reconstruction *c*, and $R(s,c,IMODE \mid QP)$ represents the number of bits associated with choosing *IMODE*. It includes the bits for the

intra prediction mode and the DCT-coefficients for the 4x4 luminance block. The rate term is computed using the UVLC entropy coding, even if CABAC is used for entropy coding.

2.1.2.4 Algorithm for motion estimation and mode decision

The procedure to encode one macroblock *s* in an I, P or B frame in the high-complexity mode is summarized as follows.

a) Given the last decoded frames, λ_{MODE} , λ_{MOTION} , and the macroblock quantisation parameter *QP*

b) Choose intra prediction modes for the *Intra 4x4* macroblock mode by minimizing

e [Ed. Note: Something wrong here]

with

$$IMODE \in \{DC, HOR, VERT, DIAG, DIAG\_RL, DIAG\_LR\} \tag{2-39}$$

c) Determine the best *Intra16x16* prediction mode by choosing the mode that results in the minimum SATD.

d) For each 8x8 sub-partition

Perform motion estimation and reference frame selection by minimizing

$$SSD + \lambda\ Rate(MV, REF) \tag{2-40}$$

B frames: Choose prediction direction by minimizing

$$SSD + \lambda\ Rate(MV(PDIR), REF(PDIR)) \tag{2-41}$$

Determine the coding mode of the 8x8 sub-partition using the rate-constrained mode decision, i.e. minimize

$$SSD + \lambda\ Rate(MV, REF, Luma\text{-}Coeff, block\ 8x8\ mode) \tag{2-42}$$

Here the *SSD* calculation is based on the reconstructed signal after DCT, quantisation, and IDCT.

e) Perform motion estimation and reference frame selection for 16x16, 16x8, and 8x16 modes by minimizing

$$J(REF, \mathbf{m}(REF) \mid \lambda_{MOTION}) = SA(T)D(s, c(REF, \mathbf{m}(REF))) + \lambda_{MOTION} \cdot (R(\mathbf{m}(REF) - \mathbf{p}(REF)) + R(REF)) \tag{2-43}$$

for each reference frame and motion vector of a possible macroblock mode.

f) B frames: Determine prediction direction by minimizing

$$J(PDIR \mid \lambda_{MOTION}) = SATD(s, c(PDIR, \mathbf{m}(PDIR))) + \lambda_{MOTION} \cdot (R(\mathbf{m}(PDIR) - \mathbf{p}(PDIR)) + R(REF(PDIR))) \tag{2-44}$$

g) Choose the macroblock prediction mode by minimizing

$$J(s, c, MODE \mid QP, \lambda_{MODE}) = SSD(s, c, MODE \mid QP) + \lambda_{MODE} \cdot R(s, c, MODE \mid QP), \tag{2-45}$$

given *QP* and λ_{MODE} when varying *MODE*. *MODE* indicates a mode out of the set of potential macroblock modes:

I frame: $$MODE \in \{INTRA\,4x4, INTRA16x16\}, \tag{2-46}$$

$$\text{P frame:} \quad MODE \in \begin{Bmatrix} INTRA4x4,\ INTRA16x16,\ SKIP, \\ 16x16,\ 16x8,\ 8x16,\ 8x8 \end{Bmatrix}, \tag{2-47}$$

$$\text{B frame:} \quad MODE \in \begin{Bmatrix} INTRA4x4,\ INTRA16x16,\ DIRECT, \\ 16x16,\ 16x8,\ 8x16,\ 8x8 \end{Bmatrix}. \tag{2-48}$$

The computation of $J(s,c,SKIP \mid QP,\lambda_{MODE})$ and $J(s,c,DIRECT \mid QP,\lambda_{MODE})$ is simple. The costs for the other macroblock modes are computed using the intra prediction modes or motion vectors and reference frames, which have been estimated in the previous steps.

2.2 Transform and quantisation

2.2.1 Quantisation tables

The coefficients Q(m,i,j), used in the formulas below, are defined in pseudo code by:

Q(m,i,j) = $M_{m,0}$ for (i,j) = {(0,0),(0,2),(2,0),(2,2)},

Q(m,i,j) = $M_{m,1}$ for (i,j) = {(1,1),(1,3),(3,1),(3,3)},

Q(m,i,j) = $M_{m,2}$ otherwise;

where

$$M = \begin{bmatrix} 13107 & 5243 & 8066 \\ 11916 & 4660 & 7490 \\ 10082 & 4194 & 6554 \\ 9362 & 3647 & 5825 \\ 8192 & 3355 & 5243 \\ 7282 & 2893 & 4559 \end{bmatrix} \tag{2-49}$$

2.2.2 4x4 spatial block processing

Instead of a discrete cosine transform (DCT), an integer transform with a similar coding gain as a 4x4 DCT is used. The transformation of input samples X={x_{00} … x_{33}} to output coefficients Y is defined by

$$Y = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} \begin{bmatrix} x_{00} & x_{01} & x_{02} & x_{03} \\ x_{10} & x_{11} & x_{12} & x_{13} \\ x_{20} & x_{21} & x_{22} & x_{23} \\ x_{30} & x_{31} & x_{32} & x_{33} \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 1 & -1 & -2 \\ 1 & -1 & -1 & 2 \\ 1 & -2 & 1 & -1 \end{bmatrix} \tag{2-50}$$

Multiplication by two can be performed either through additions or through left shifts, so that no actual multiplication operations are necessary. Thus, we say that the transform is multiplier-free.

For input samples with 9-bit dynamic range (because they are residuals from 8-bit sample data), the transform coefficients are guaranteed to fit within 16 bits, even when the second transform for DC coefficients is used. Thus, all transform operations can be computed in 16-bit arithmetic.

The transform and inverse transform matrices above have orthogonal basis functions. Unlike the DCT, though, the basis functions don't have the same norm. Therefore, for the inverse transform to recover the original samples, appropriate normalization factors must be applied to the transform coefficients before quantisation and after inverse quantisation. Such factors are absorbed by the quantisation and inverse quantisation scaling factors.

Quantisation is performed according to the following equation:

$$Y_Q(i,j) = \left[Y(i,j) \cdot Q((QP+12)\%6, i, j) + f \right] / 2^{15+(QP+12)/6}, \quad i,j = 0,\ldots,3 \tag{2-51}$$

where Y are the transformed coefficients, Y_Q are the corresponding quantised values, $Q(m,i,j)$ are the quantisation coefficients listed below, with f having the same sign as $Y(i,j)$, and |f| is in the range 0 to $2^{14+(QP+12)/6/2}$. Specifically,

$$|f| = \begin{cases} (2 \cdot 2^{15+(QP+12)/6}) / 6 & \text{for intra coefficient quantization} \\ (1 \cdot 2^{15+(QP+12)/6}) / 6 & \text{for inter coefficient quantization} \end{cases} \tag{2-52}$$

NOTE – while the intermediate value inside square brackets in Equation 2-43 has a 32-bit range, the final value Y_Q is guaranteed to fit in 16 bits.

2.2.3 DC luminance coefficients in 16x16 intra mode

To minimize the dynamic range expansion due to transformations (and thus minimize rounding errors in the quantisation scale factors), a simple scaled Hadamard transform is used. The direct transform is defined by:

$$Y_D = \left(\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_{D00} & x_{D01} & x_{D02} & x_{D03} \\ x_{D10} & x_{D11} & x_{D12} & x_{D13} \\ x_{D20} & x_{D21} & x_{D22} & x_{D23} \\ x_{D30} & x_{D31} & x_{D32} & x_{D33} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \right) // 2 \tag{2-53}$$

where the symbol // denotes division with rounding to the nearest integer:

$$a // 2^b = \text{sign}(a) \times \left[\left(\text{abs}(a) + 2^{b-1} \right) >> b \right] \tag{2-54}$$

Quantisation is performed according to the following equation:

$$Y_Q(i,j) = [Y(i,j) \cdot Q((QP+12)\%6, i, j) + 2f] / 2^{16+(QP+12)/6}, \quad i,j = 0,...,3 \tag{2-55}$$

where f has the same sign as $Y(i,j)$ and |f| is given by Equation 2-44.

2.2.4 DC chrominance coefficients

The 2 dimensional 2x2 transform procedure is:

$$Y = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} DC_{00} & DC_{01} \\ DC_{10} & DC_{11} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{2-56}$$

Quantisation is performed according to the following equation:

$$Y_Q(i,j) = \left[Y(i,j) \cdot Q((QP+12)\%6, i, j) + 2f \right] / 2^{16+(QP+12)/6}, \quad i,j = 0,\ldots,3 \tag{2-57}$$

where f has the same sign as $Y(i,j)$ and |f| is given by Equation 2-44.

2.3 Elimination of single coefficients in inter macroblocks [Ed. Note: Do we still use this?]

2.3.1 Luminance

With the small 4x4 blocks, it may happen that for instance a macroblock has only one nonzero coefficient with |Level| =1. This will probably be a very "expensive" coefficient in terms of bit usage and it could have been better to set it to zero. For that reason a procedure to check single coefficients has been used for inter luma blocks. During the quantisation process, a parameter Single_ctr is accumulated depending on Run and Level according to the following rule:

- If Level = 0 or (|Level| = 1 and Run > 5) nothing is added to Single_ctr.
- If |Level| > 1, 9 is added to Single_ctr.
- If |Level| = 1 and Run < 6, a value T(Run) is added to Single_ctr. where T(0:5) =(3,2,2,1,1,1)
- If the accumulated Single_ctr for a 8x8 block is less than 4, all coefficients of that luma block are set to zero. Similarly, if the accumulated Single_ctr for the whole macroblock is less than 6, all coefficients of that luma macroblock are set to zero.

2.3.2 Chrominance

A similar method to the one for luma is used. Single_ctr is calculated similarly for each chroma component, but for AC coefficients only and for the whole macroblock.

If the accumulated Single_ctr for each chroma component of a macroblock is less than 7, all the AC chroma coefficients of that component for the whole macroblock are set to zero.

2.4 S-Pictures

2.4.1 Encoding of secondary SP-pictures

This section suggests an algorithm that can be used to create a secondary SP-picture with identical sample values to another SP-picture (called the target SP-picture). The secondary SP-picture is typically used for switching from one bitstream to another and thus has different reference frames for motion compensation than the target SP-picture.

Intra-type macroblocks from the target SP-picture are copied into the secondary SP-picture without alteration. SP-macroblocks from the target SP-pictures will be replaced by the following procedure: Let Lrec be the quantised coefficients, see subclause 8.2, that are used to reconstruct the SP-block. Define the difference levels Lerr by:

$$Lerr = Lrec - L_P \tag{2-58}$$

where L_P is the quantised transform coefficients of the predicted block from any of the SP coding modes. Then a search over all the possible SP modes is performed and the mode resulting in the minimum number of bits is selected.

2.4.2 Encoding of SI-pictures

The following algorithm is suggested to encode an SI-picture such that the reconstruction of it is identical to that of an SP-picture. SP-pictures consist of intra and SP-macroblocks. Intra-type macroblocks from SP-pictures are copied into SI-pictures with minor alteration. Specifically, the MB_Type is altered to reflect Intra Modes in SI-pictures, i.e., MB_Type is incremented by one, See Table 15. The SP-macroblocks from SP-pictures will be replaced by SIntra 4x4 modes. Let Lrec be the quantised coefficients, see subclause 8.2, that are used to reconstruct the SP-block. Define the difference levels Lerr by:

$$Lerr = Lrec - L_I \tag{2-59}$$

where L_I is the quantised transform coefficients of the predicted block from any of the intra prediction modes. Then a search over the possible intra prediction modes is performed and the mode resulting in the minimum number of bits is selected.

2.5 Encoding with anticipation of slice losses

Low delay video transmission may lead to losses of slices. The decoder may then stop decoding until the next I picture or P picture may conduct a concealment, for example as explained in Appendix IV, and continue decoding. In the latter case, spatio-temporal error propagation occurs if the concealed picture content is referenced for motion compensation. There are various means to stop spatio-temporal error propagation including the usage of multiple reference pictures and Intra coding of macroblocks. For the latter case, a Lagrangian mode selection algorithm is suggested as follows.

Since transmission errors occur randomly, the decoding result is also a random process. Therefore, the average decoder distortion is estimated to control the encoder for a specified probability of packet losses p. The average decoding result is obtained by running *N* complete decoders at the encoder in parallel. The statistical process of loosing a slice is assumed to be independent for each of the *N* decoders. The slice loss process for each decoder is also assumed to be independent and identically distributed, and a certain slice loss probability *p* is assumed to be known at the encoder. Obviously for large *N* the decoder gets a very good estimate of the average decoder distortion. However, with increasing *N* a linear increase of storage and decoder complexity in the encoder is incurred. Therefore, this method might not be practical in real-time encoding processes and complexity and memory efficient algorithms are currently under investigation.

To encode a macroblock in a P picture, the set of possible macroblock types is given as

$$\mathcal{S}_{MB} = \{ \text{MBskip, MBinter16x16, Binter16x8, MBinter8x16, MBinter8x8, MBintra16x16, MBintra4x4} \} \quad (2\text{-}60)$$

For each macroblock the coding mode m' is selected according to

$$m' = \min_{m \in \mathcal{S}_{MB}} \{ D_m + \lambda R_m \} \quad (2\text{-}61)$$

with D_m being the distortion in the current macroblock when selecting macroblock mode m and R_m being the corresponding rate, i.e. the number of bits. For the COPY_MB and all INTER_MxN types, the distortion D_m is computed as

$$D_m = \frac{1}{N} \sum_{n=1}^{N} \sum_{i} (f_i - \hat{f}_{i,n,m}(p))^2 \quad (2\text{-}62)$$

with f_i being the original sample value at position i within the macroblock and $\hat{f}_{i,n,m}$ being the reconstructed sample value at position i for coding macroblock mode m in the simulated decoder n. The distortion for the intra macroblocks remains unchanged. Since the various reconstructed decoders also contain transmission errors, the Lagrangian cost function for the COPY_MB and all INTER_MxN types increases making Intra_NxN types more popular.

The λ parameter for mode decision depends on the quantisation parameter q as follows

[Ed. Note: Change "l" to Lambda below.]

$$\lambda = (1 - p) 0.85 \cdot 2^{QP/3} \quad (2\text{-}63)$$

A reference frames restriction is introduced to avoid the prediction from areas which have been intra-updated in later frames. This restriction is only effective for P-frames where multiple reference frames are used, rate-distortion optimized mode selection is applied and the reference restriction flag (RRF) is set to 1 or 2. The RRF has the following effects on the encoded video:

RRF	Effect on Encoded Video
0	No restriction on reference frames for any MB is applied
1	Disallows combinations of macroblock modes and reference frames which reference pixels previous to a macroblock intra-updated due to error resilience reasons.
2	Disallows combinations of macroblock modes and reference frames which reference pixels previous to any intra-updated macroblock

The detailed algorithm on the reference frame restriction is as follows. Let us define a intra pixel map *ip(i,j), i=1,..., img_width, j=1,..., img_height*, which specifies for each pixel the lowest past frame number in which the reference chain referenced an intra-updated pixel. The generation of *ip(i,j)* will be discussed later. In addition we define for each 8x8 block a intra-refresh flag *r(i,j), i=1,..., img_width/8, j=1,..., img_height/8*, which signals if the block at position *(i,j)* restricts further prediction. The generation of *r(i,j)* will also be discussed later.

The reliability of a certain combination of MB coding mode *mode*, reference frame *ref* and 8x8block position *block* is checked and only if the combination is reliable, it is valid to be selected in the encoding process. A combination is not valid if any referenced pixel at position *(i,j)* or any corresponding quarter-pel position has been intra-updated in a frame later than the selected reference frame, i.e. it is not valid if *ip(i,j)* < *ref*.

The intra-refresh flag *r(i,j)* is updated for each 8x8 block after encoding a macroblock depending of the selected reference restriction flag. If the RRF is set to 2, then *r(i,j)* is set to 1 if this block is intra-coded (MB mode = INTRA16 or block mode = IBLOCK). If the RRF is set to 1, then *r(i,j)* is set to 1 if this block is not intra-coded due to compression reasons, but intra-coded (MB mode = INTRA16 or block mode = IBLOCK) due to error-resilience reasons (forced intra-update by regular or random intra-update, optimized intra-updates).

After encoding one frame the intra pixel map is updated. If the current pixel position is intra updated, the intra pixel map is set to 1 at this position. Otherwise, the intra pixel map is increased by one as the last intra update is moved further to the past. The maximum number is the number of possible reference frames specified. The pseudo-code is as follows:

```
// for all blocks
for (my=0; my<img->height/8; my++)
    for (mx=0; mx<img->width/8;    mx++)    {
            j = my*8 + 8;
            i = mx*8 + 8;
            // for all pixels
            for (y=my*8; y<j; y++)
                  for (x=mx*8; x<i; x++)
                        if (r[my][mx])
                              ip[y][x] = 1;
                        else
                              ip[y][x] = min(ip[y][x]+1, input->no_multpred+1);
```

2.6 Rate Control

This section presents the rate control algorithm with an assumption that there exists a predefined pattern for the order of the coded pictures. The algorithm is used to create the stream satisfying the available bandwidth provided by a channel and is also compliant to hypothetical reference decoder (HRD). It consists of three tightly consecutive components: GOP level rate control, picture level rate control and the optional basic unit level rate control. The basic unit is defined as a group of successive macroblocks in the same frame. When the basic units arise, each shall contain at least a macroblock.

2.6.1 GOP level rate control

GOP level rate control calculates the total bits for the rest pictures in this GOP and the initial quantization parameter of instantaneous decoding refresh (IDR) picture and that of the first stored picture.

When the j[th] picture in the i[th] GOP is coded, the total bits for the rest pictures in this GOP are computed as follows

$$B_i(j)=\begin{cases}\dfrac{R_i(j)}{f}\times N_i - V_i(j) & j=1\\ B_i(j-1)+\dfrac{R_i(j)-R_i(j-1)}{f}\times(N_i-j+1)-b_i(j-1) & j=2,3,...,N_i\end{cases}. \tag{2-64}$$

(1) For the first picture in a GOP (i.e. j = 1), the total bits are calculated from the upper formula in (2-64). f is the predefined coding frame rate. N_i is the total number of pictures in the i[th] GOP. $R_i(j)$ and $V_i(j)$ are the instant available bit rate and the occupancy of the virtual buffer, respectively, when the j[th] picture in the i[th] GOP is coded.

(2) For other pictures, the total bits are calculated from the bottom formula in (2-64). $b_i(j\text{-}1)$ is the actual generated bits in the (j-1)[th] picture. Considering the case of the dynamic channels, $R_i(j)$ may vary at different frames and GOPs. But, in the case of constant bit rate, $R_i(j)$ is always equal to $R_i(j\text{-}1)$. The formula can be simplified as

$$B_i(j)=B_i(j-1)-b_i(j-1) \tag{2-65}$$

(3) $V_i(j)$ is updated after coding each picture as

$$\begin{aligned}V_i(1)&=\begin{cases}0 & i=1\\ V_{i-1}(N_{i-1}) & other\end{cases}\\ V_i(j)&=V_i(j-1)+b_i(j-1)-\frac{R_i(j-1)}{f} \qquad j=2,3,...,N_i\end{aligned} \tag{2-66}$$

An initial quantization parameter $QP_i(1)$ is set for the IDR picture and the first stored picture of the i[th] GOP.

(1) For the first GOP, $QP_1(1)$ is predefined based on the available channel bandwidth as follows,

$$QP_1(1) = \begin{cases} 40 & bpp \le l1 \\ 30 & l1 < bpp \le l2 \\ 20 & l2 < bpp \le l3 \\ 10 & bpp > l3 \end{cases} \quad (2\text{-}67)$$

where $bpp = \dfrac{R_1(1)}{f \times N_{pixel}}$

N_{pixel} is the number of pixel in a picture. $l1$=0.15, $l2$=0.45, $l3$=0.9 is recommended for QCIF/CIF, and $l1$=0.6, $l2$=1.4, $l3$=2.4 is recommended for the picture size larger than CIF.

(2) For the other GOPs,

$$QP_i(1) = \max\left\{QP_{i-1}(1) - 2, \min\left\{QP_{i-1}(1) + 2, \frac{SumPQP(i-1)}{N_p(i-1)} - \min\left\{2, \frac{N_{i-1}}{15}\right\}\right\}\right\} \quad (2\text{-}68)$$

$N_P(i\text{-}1)$ is the total number of stored pictures in the (i-1)th GOP, and $SumPQP(i\text{-}1)$ is the sum of average picture quantization parameters for all stored pictures in the (i-1)th GOP. It's further adjusted by

$$QP_i(1) = QP_i(1) - 1 \quad if \quad QP_i(1) > QP_{i-1}(N_{i-1} - L) - 2 \quad (2\text{-}69)$$

$QP_{i-1}(N_{i-1} - L)$ is the quantization parameter of the last stored picture in the previous GOP, and L is the number of successive non-stored pictures between two stored pictures.

2.6.2 Picture level rate control

The picture level rate control consists of two stages: pre-encoding and post-encoding.

2.6.2.1 Pre-encoding stage

The objective of this stage is to compute a quantization parameter of each picture. Different methods are proposed for stored and non-stored pictures

2.6.2.1.1 Non-stored pictures

The quantization parameters for non-stored pictures are computed by a simple interpolation method as follows:

Suppose that the jth and (j+L+1)th pictures are stored pictures and the quantization parameters for these stored pictures are $QP_i(j)$ and $QP_i(j+L+1)$, respectively. The quantization parameter of the ith non-stored pictures is calculated according to the following two cases:

Case 1: When L=1, there is only one non-stored picture between two stored pictures. The quantization parameter is computed by

$$QP_i(j+1) = \begin{cases} \dfrac{QP_i(j) + QP_i(j+2) + 2}{2} & \text{if } QP_i(j) \ne QP_i(j+2) \\ QP_i(j) + 2 & \text{Otherwise} \end{cases} \quad (2\text{-}70)$$

Case 2: when L>1, there are more than one non-stored picture between two stored pictures. The quantization parameters are computed by

$$QP_i(j+k) = QP_i(j) + \alpha + \max\{\min\{\frac{(QP_i(j+L+1) - QP_i(j)}{L-1}, 2\times(k-1)\}, -2\times(k-1)\} \quad (2\text{-}71)$$

where k =1, ..., L, and α is given by

$$\alpha = \begin{cases} -3 & QP_i(j+L+1) - QP_i(j) \le -2\times L - 3 \\ -2 & QP_i(j+L+1) - QP_i(j) = -2\times L - 2 \\ -1 & QP_i(j+L+1) - QP_i(j) = -2\times L - 1 \\ 0 & QP_i(j+L+1) - QP_i(j) = -2\times L \\ 1 & \mathrm{QP_i}(j+L+1) - QP_i(j) = -2\times L + 1 \\ 2 & \text{Otherwise} \end{cases} \tag{2-72}$$

The quantization parameter $QP_i(j+k)$ is further bounded by 0 and 51.

2.6.2.1.2 Stored pictures

The quantization parameter of stored picture is computed via the following two steps.

Step 1: Determine target bits for each stored picture.

(1) Determine the target buffer level for each stored picture in the current GOP.

A target buffer level is predefined for each stored picture according to the coded bits of the first IDR picture and the first stored picture, and the average picture complexity. After coding the first stored picture in the i[th] GOP, the initial value of target buffer level is set to,

$$S_i(2) = V_i(2) \tag{2-73}$$

The target buffer level for the subsequent stored picture is determined by

$$S_i(j+1) = S_i(j) - \frac{S_i(2)}{N_p(i)-1} + \frac{\overline{W}_{p,i}(j)\times(L+1)\times R_i(j)}{f\times(\overline{W}_{p,i}(j)+\overline{W}_{b,i}(j)\times L)} - \frac{R_i(j)}{f} \tag{2-74}$$

where $\overline{W}_{p,i}(j)$ is the average complexity weight of stored pictures, $\overline{W}_{b,i}(j)$ is the average complexity weight of non-stored pictures. They are computed by

$$\begin{aligned} \overline{W}_{p,i}(j) &= \frac{W_{p,i}(j)}{8} + \frac{7\times\overline{W}_{p,i}(j-1)}{8} \\ \overline{W}_{b,i}(j) &= \frac{W_{b,i}(j)}{8} + \frac{7\times\overline{W}_{b,i}(j-1)}{8} \\ W_{p,i}(j) &= b_i(j)\times QP_{p,i}(j) \\ W_{b,i}(j) &= \frac{b_i(j)\times QP_{b,i}(j)}{1.3636} \end{aligned} \tag{2-75}$$

When there is no non-stored picture between two stored pictures, Equation (2-74) is simplified as

(2-73)

$$S_i(j+1) = S_i(j) - \frac{S_i(2)}{N_p(i)-1} \tag{2-76}$$

(2) Compute the target bits for the current stored picture.

The target bits allocated for the j[th] stored picture in the i[th] GOP are determined based on the target buffer level (2-74), the frame rate, the available channel bandwidth, and the actual buffer occupancy of (2-66) as follows:

$$\widetilde{T}_i(j) = \frac{R_i(j)}{f} + \gamma\times(S_i(j) - V_i(j)) \tag{2-77}$$

where γ is a constant and its typical value is 0.5 when there is no non-stored picture and 0.25 otherwise.

Meanwhile, the number of remaining bits should also be considered when the target bit is computed.

$$\hat{T}_i(j) = \frac{W_{p,i}(j-1) \times B_i(j)}{W_{p,i}(j-1) \times N_{p,r} + W_{b,i}(j-1) \times N_{b,r}} \tag{2-78}$$

where $N_{p,r}$ and $N_{b,r}$ are the number of the remaining stored pictures and the number of the remaining non-stored pictures, respectively.

The target bits are a weighted combination of $\widetilde{T}_i(j)$ and $\hat{T}_i(j)$:

$$T_i(j) = \beta \times \hat{T}_i(j) + (1-\beta) \times \widetilde{T}_i(j) \tag{2-79}$$

where β is a constant and its typical value is 0.5 when there is no non-stored picture and is 0.9 otherwise.

To conform with the HRD requirement, the target bits are bounded by

$$\begin{aligned} T_i(j) &= \max\{Z_i(j), \mathrm{T_i}(j)\} \\ T_i(j) &= \min\{U_i(j), \mathrm{T_i}(j)\} \end{aligned} \tag{2-80}$$

where $Z_i(j)$ and $U_i(j)$ are computed by

$$Z_i(j) = \begin{cases} B_{i-1}(N_{i-1}) + \dfrac{R_i(j)}{f} & j = 1 \\ Z_i(j-1) + \dfrac{R_i(j)}{f} - b_i(j) & other \end{cases} \tag{2-81}$$

$$U_i(j) = \begin{cases} (B_{i-1}(N_{i-1}) + t_{r,1}(1)) \times \varpi & j = 1 \\ U_i(j-1) + (\dfrac{R_i(j)}{f} - b_i(j)) \times \varpi & other \end{cases} \tag{2-82}$$

$t_{r,1}(1)$ is the removal time of the first picture from the coded picture buffer. ϖ is a constant with typical value of 0.9.

Step 2: Compute the quantization parameter and perform RDO.

The MAD of the current stored picture, $\widetilde{\sigma}_i(j)$, is predicted by a linear model (2-83) using the actual MAD of the previous stored picture, $\sigma_i(j-1-L)$.

$$\widetilde{\sigma}_i(j) = a_1 \times \sigma_i(j-1-L) + a_2 \tag{2-83}$$

where a_1 and a_2 are two coefficients. The initial value of a_1 and a_2 are set to 1 and 0, respectively. They are updated by a linear regression method similar to that of MPEG-4 Q2 after coding each picture or each basic unit.

The quantization step corresponding to the target bits is then computed by using the following quadratic :

$$T_i(j) = c_1 \times \frac{\widetilde{\sigma}_i(j)}{Q_{step,i}(j)} + c_2 \times \frac{\widetilde{\sigma}_i(j)}{Q^2_{step,i}(j)} - m_{h,i}(j) \tag{2-84}$$

where $m_{h,i}(j)$ is the total number of header bits and motion vector bits, c_1 and c_2 are two coefficients.

The corresponding quantization parameter $QP_i(j)$ is computed by using the relationship between the quantization step and the quantization parameter of AVC. To maintain the smoothness of visual quality among successive frames, the quantization parameter $QP_i(j)$ is adjusted by

$$QP_i(j) = \min\{QP_i(j-L-1)+2, \max\{QP_i(j-L-1)-2, QP_i(j)\}\} \tag{2-85}$$

The final quantization parameter is further bounded by 51 and 0. The quantization parameter is then used to perform RDO for each MB in the current frame.

2.6.2.2 Post-encoding stage

After encoding a picture, the parameters *a1* and *a2* of linear prediction model (2-83), as well as c_1 and c_2 of quadratic R-D model (2-84) are updated. A linear regression method similar to MPEG-4 Q2 is used to update these parameters.

Meanwhile, the actual bits generated are added to the buffer. To ensure that the updated buffer occupancy is not too high, a number of pictures may be skipped by using the method similar to MPEG-4 Q2.

2.6.3 Basic unit level rate control

Suppose that a picture is composed of N_{mbpic} MBs. A basic unit is defined to be a group of continuous MBs, and consists of N_{mbunit} MBs, where N_{mbunit} is a fraction of N_{mbpic}. If N_{mbunit} equals to N_{mbpic}, it would be a picture level rate control, and if N_{mbunit} equals to 1, it falls back to a macroblock level rate control,

The total number of basic units in a frame, N_{unit}, is computed by

$$N_{unit} = \frac{N_{mbpic}}{N_{mbunit}} \tag{2-86}$$

If the basic unit is not selected as a frame, an additional basic unit layer rate control for the stored picture should be added.

Same as the frame layer rate control, the quantization parameters for IDR picture and non-stored pictures are the same for all basic unit in the same picture. It is computed the similar way as that at frame layer provided that $QP_i(j)$ and $QP_i(j+L+1)$ are replaced by the average values of quantization parameters of all basic units in the corresponding picture.

The basic unit layer rate control selects the values of quantization parameters of all basic units in a frame, so that the sum of generated bits is close to the frame target $T_i(j)$.

The following is a step-by-step description of this method.

Step 1 Predict the MADs, $\tilde{\sigma}_{l,i}(j)$, of the remaining basic units in the current stored picture by model (2-83) using the actual MADs of the co-located basic units in previous stored picture.

Step 2 Compute the number of texture bits $\hat{b}_l$ for the l^{th} basic unit. This step is composed of the following three sub-steps:

Step 2.1 Compute the target bits for the l^{th} basic unit.

Let T_r denote the number of remaining bits for the current frame and its initial value is set to $T_i(j)$. The target bits for the l^{th} basic unit are given by

$$\tilde{b}_l = T_r \times \frac{\tilde{\sigma}_{l,i}^2(l)}{\sum_{k=l}^{N_{unit}} \tilde{\sigma}_{k,i}^2(j)} \tag{2-87}$$

Step 2.2 Compute the average number of header bits generated by all coded basic units:

$$\begin{aligned}\tilde{m}_{hdr,l} &= \tilde{m}_{hdr,l-1} \times (1-\frac{1}{l}) + \frac{\hat{m}_{hdr,l}}{l} \\ m_{hdr,l} &= \tilde{m}_{hdr,l} \times \frac{l}{N_{unit}} + m_{hdr,1} \times (1-\frac{l}{N_{unit}}); 1 \le l \le N_{unit}\end{aligned} \tag{2-88}$$

where $\hat{m}_{hdr,l}$ is the actual number of header bits generated by the l^{th} basic unit in the current stored picture. $m_{hdr,1}$ is the estimation from all basic units in the previous stored picture.

Step 2.3 Compute the number of texture bits $\hat{b}_l$ for the l^{th} basic unit.

$$\hat{b}_l = \tilde{b}_l - m_{hdr,l} \tag{2-89}$$

Step 3 Compute the quantization step for the l^{th} basic unit of j^{th} picture in i^{th} GOP by using the quadratic R-D model (2-84) and it is converted to the corresponding quantization parameter $QP_{l,i}(j)$ by using the method provided by AVC. We need to consider the following three cases:

Case 1 The first basic unit in the current frame.

$$QP_{1.i}(j) = \overline{QP}_i(j-L-1) \tag{2-90}$$

where $\overline{QP}_i(j-L-1)$ is the average value of quantization parameters for all basic units in the previous stored picture.

Case 2 When the number of remaining bits is less than 0, the quantization parameter should be greater than that of previous basic unit such that the sum of generated bits is close to the target bits, i.e.

$$QP_{l,i}(j) = QP_{l-1,i}(j) + \Delta_{Bu} \tag{2-91}$$

where Δ_{Bu} is the varying range of quantization parameter along basic units, the initial value of Δ_{Bu} is 1 if N_{unit} is greater than 8, and 2 otherwise. It is updated after coding each basic unit as follows:

$$\Delta_{Bu} = \begin{cases} 1; & \text{if } QP_{l-1,i}(j) > 25 \\ 2; & \text{otherwise} \end{cases} \tag{2-92}$$

To maintain the smoothness of perceptual quality, the quantization parameter is further bounded by

$$QP_{l,i}(j) = \max\{0, \overline{QP}_i(j-L-1) - \Delta_{Fr}, \min\{51, \overline{QP}_i(j-L-1) + \Delta_{Fr}, QP_{l,i}(j)\}\} \tag{2-93}$$

where Δ_{Fr} is the varying range of quantization parameter along frames, and is defined by

$$\Delta_{Fr} = \begin{cases} 2; & \text{if } N_{unit} > 18 \\ 4; & \text{if } 18 \ge N_{unit} > 9 \\ 6; & \text{otherwise} \end{cases} \tag{2-94}$$

Case 3 Otherwise, we shall first compute a quantization step by using the quadratic model (2-84), and convert it into the corresponding quantization parameter $QP_{l,i}(j)$. Similar to case 2, it is bounded by

$$QP_{l,i}(j) = \max\{QP_{l-1,i}(j) - \Delta_{Bu}, \min\{QP_{l.i}(j), QP_{l-1,i}(j) + \Delta_{Bu}\}\} \tag{2-95}$$

Meanwhile, in order to maintain the smoothness of visual quality, it is further bounded by (2-93).

Step 4 Perform RDO for all MBs in the current basic unit and code them by AVC.

Step 5 Update the number of remaining bits, the coefficients of the linear prediction model (2-83), and those of the quadratic R-D model (2-84).

To obtain a good trade-off between average PSNR and bit fluctuation, N_{mbunit} is recommended to be the number of MBs in a row for field coding, adaptive field/frame coding, or MB-AFF coding, and N_{unit} is recommended to be 9 for other cases. (2-85)

3 Non-normative decoder error concealment description

3.1 Introduction

It is assumed that no erroneous or incomplete slices are decoded. When all received slices of a picture have been decoded, skipped slices are concealed according to the presented algorithms. In practice, record is kept in a macroblock (MB) based status map of the frame. The status of an MB in the status map is "Correctly received" whenever the slice that the MB is included in was available for decoding, "Lost" otherwise. After the frame is decoded if the status map contains "Lost" MBs, concealment is started.

Given the slice structure and MB-based status map of a frame, the concealment algorithms were designed to work MB-based. The missing frame area (samples) covered by MBs marked as "Lost" in the status map are concealed MB-by-MB (16x16 Y samples, 8x8 U, V samples). After an MB has been concealed it is marked in the status map as "Concealed". The order in which "Lost" MBs are concealed is important as also the "Concealed", and not only the "Correctly received" MBs are treated as reliable neighbours in the concealment process whenever no "Correctly received" immediate neighbour of a "Lost" MB exists. In such cases a wrong concealment can result in propagation of this concealment mistake to several neighbour concealed MBs. The processing order chosen is to take the MB columns at the edge of the frame first and then move inwards column-by-column so to avoid a concealment mistake made in the usually "difficult" (discontinuous motion areas, large coded prediction error) center part of the frame propagate to the "easy" (continuous motion area, similar motion over several frames) side parts of the frame.

Figure 3-1 shows a snapshot of the status map during the concealment phase where already concealed MBs have the status of "Concealed", and the currently processed (concealed) MB is marked as "Current MB".

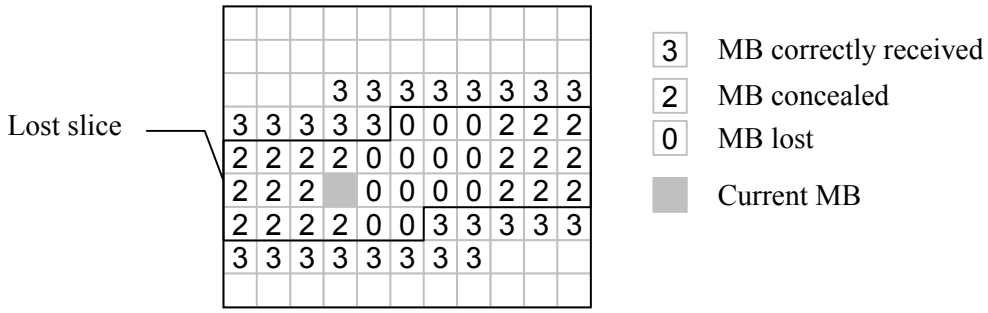

Figure 3-1– MB status map at the decoder

3.2 Intra frame concealment

Lost areas in intra frames have to be concealed spatially as no prior frame may resemble the intra frame. The selected spatial concealment algorithm is based on weighted sample averaging presented in A. K. Katsaggelos and N. P. Galatsanos (editors), "Signal Recovery Techniques for Image and Video Compression and Transmission", Chapter 7, P. Salama, N. B. Shroff, and E. J. Delp, "Error Concealment in Encoded Video Streams", Kluwer Academic Publishers, 1998.

Each sample value in a macroblock to be concealed is formed as a weighted sum of the closest boundary samples of the selected adjacent macroblocks. The weight associated with each boundary sample is relative to the inverse distance between the sample to be concealed and the boundary sample. The following formula is used:

$$\text{Sample value} = (\sum a_i B\text{-}d_i)) / \sum (B\text{-}d_i) \qquad (3\text{-}1)$$

where a_i is the sample value of a boundary sample in an adjacent macroblock, B is the horizontal or vertical block size in samples, and d_i is the distance between the destination sample and the corresponding boundary sample in the adjacent macroblock.

In [Ed. Note: Something missing]

Figure 3-2, the shown destination sample is calculated as follows

Sample value = (15x(16-3) + 21x(16-12) + 32x(16-7) + 7x(16-8)) / (13 + 4 + 9 + 8) □ 18

Only "Correctly received" neighbouring MBs are used for concealment if at least two such MBs are available. Otherwise, neighbouring "Concealed" MBs are also used in the averaging operation.

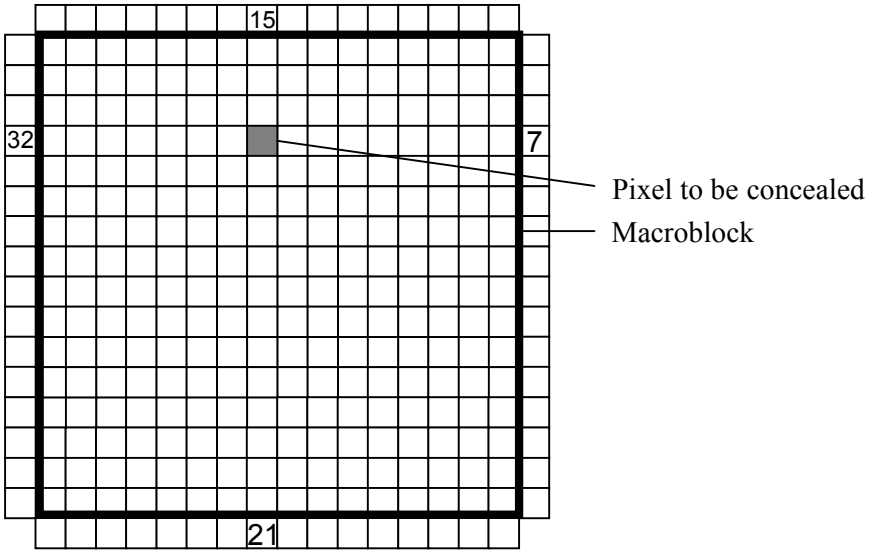

Figure 3-2 – Spatial concealment based on weighted sample averaging

3.3 Inter and SP frame concealment

3.3.1 General

Instead of directly operating in the sample domain a more efficient approach is to try to "guess" the motion in the missing sample area (MB) by some kind of prediction from available motion information of spatial or temporal neighbours. This "guessed" motion vector is then used for motion compensation using the reference frame. The copied sample values give the final reconstructed sample values for concealment, and no additional sample domain operations are used. The presented algorithm is based on W.-M. Lam, A. R. Reibman, and B. Liu, “Recovery of lost or erroneously received motion vectors,” in Proc. ICASSP’93, Minneapolis, Apr. 1993, pp. V417–V420.

3.3.2 Concealment using motion vector prediction

The motion activity of the correctly received slices of the current picture is investigated first. If the average motion vector is smaller than a pre-defined threshold (currently ¼ samples for each motion vector component), all the lost slices are concealed by copying from co-located positions in the reference frame. Otherwise, motion-compensated error concealment is used, and the motion vectors of the lost macroblocks are predicted as described in the following paragraphs.

The motion of a "Lost" MB is predicted from a spatial neighbour MB's motion relying on the statistical observation, that the motion of spatially neighbour frame areas is highly correlated. For example, in a frame area covered by a moving foreground scene object the motion vector field is continuous, which means that it is easy to predict.

The motion vector of the "Lost" MB is predicted from one of the neighbour MBs (or blocks). This approach assumes, that the motion vector of one of the neighbour MBs (or blocks) models the motion in the current MB well. It was found in previous experiments, that median or averaging over all neighbours' motion vectors does not give better results. For simplicity, in the current implementation the smallest neighbour block size that is considered separately as predictor is

set to 8x8 Y samples. The motion of any 8x8 block is calculated as the average of the motion of the spatially corresponding 4x4 or other shaped (e.g. 4x8) blocks.

The decision of which neighbour's motion vectors to use as prediction for the current MB is made based on the smoothness of the concealed (reconstructed) image. During this trial procedure the concealment sample values are calculated using the motion vector of each candidate (motion compensated sample values). The motion vector, which results in the smallest luminance change across block boundaries when the block is inserted into its place in the frame is selected. (see Figure 3-3). The zero motion vector case is always considered and this copy concealment (copy sample values from the co-located MB in the reference frame) is evaluated similarly as the other motion vector candidates.

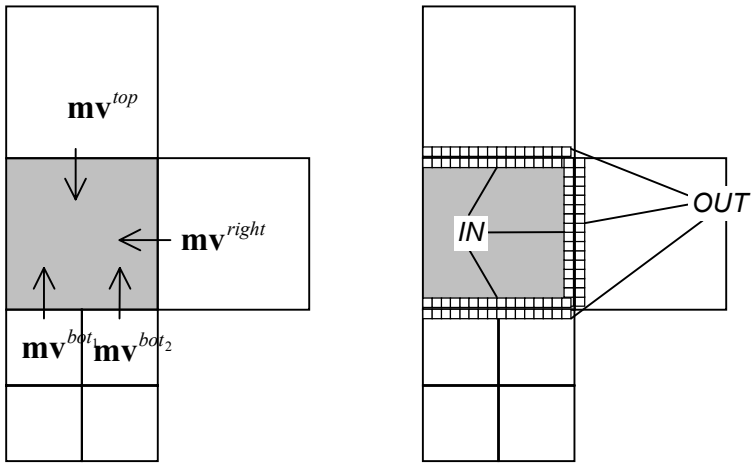

Figure 3-3 – Selecting the motion vector for prediction

The winning predictor motion vector is the one which minimizes the side match distortion , which is the sum of absolute Y sample value difference of the IN-block and neighbouring OUT-block samples at the boundaries of the current block:

[Ed. Note: Missing equation]

When "Correctly received" neighbour MBs exist the side match distortion is calculated only for them. Otherwise all the "Concealed" neighbour MBs are included in the calculation.

3.3.3 Handling of multiple reference frames

When multiple references are used, the reference frame of the candidate motion vector is used as the reference frame for the current MB. That is, when calculating the side match distortion , the IN-block samples are from the reference frame of the candidate motion vector.

3.4 B frame concealment

A simple motion vector prediction scheme according to the prediction mode of the candidate MB is used as follows:

If the prediction mode of the candidate MB is

- forward prediction mode, use the forward MV as the prediction the same way as for P frames.
- backward prediction mode, use the backward MV as the prediction.
- bi-directional prediction mode, use the forward MV as the prediction, and discard the backward MV.
- direct prediction mode, use the backward MV as the prediction.

Note that 1) Each MV, whether forward or backward, has its own reference frame. 2) An Intra coded block is not used as a motion prediction candidate.

3.5 Handling of entire frame losses

TML currently lacks H.263 Annex U type of reference picture buffering. Instead, a simple sliding window buffer model is used, and a picture is referred using its index in the buffer. Consequently, when entire frames are lost, the reference buffer needs to be adjusted. Otherwise, the following received frames would use wrong reference frames. To solve this

problem, the reference picture ID is used to infer how many frames are lost, and the picture indices in the sliding window buffer.