Implementor's Guide for ITU-T Recommendation H.263

Introduction

This document is the first Implementor’s Guide for Recommendation H.263.

Implementor's Guide for the ITU-T Recommendation H.263 - Video Coding for low bit rate communication

Abstract

This document is a compilation of reported defects identified with the March, 2001 edition of the ITU-T Recommendation H.263. It is intended to be read in conjunction with the Recommendation to serve as an additional authoritative source of information for implementors. The changes, clarifications and corrections defined herein are expected to be included in future versions of Recommendation H.263.

Contact information

ITU-T Study Group 16/Question 6 Rapporteur	Gary Sullivan One Microsoft Way Redmond, WA 98053 USA	Tel: +1 (425) 703-5308 Fax: +1 (425) 936-7329 Email: garysull@microsoft.com
Implementor's Guide Editor	Same as above	

Document history

Revision	Date	Description
1	28 May – 8 June 2001	Initial version approved by the Study Group

1. Introduction

This document is a compilation of reported defects identified with the March, 2001 edition of the ITU-T Recommendation H.263 (containing the main text and Annexes A through X). It is intended to be read in conjunction with the Recommendation to serve as an additional authoritative source of information for implementors. The changes, clarifications and corrections defined herein are expected to be included in future versions of Recommendation H.263.

The first version of the guide was produced at the May-June 2001 ITU-T Study Group 16 meeting. Wide distribution of this document is expected and encouraged.

2. Scope

This guide resolves defects in the following categories:

- editorial errors;
- technical errors such as omissions or inconsistencies;
- ambiguities.

In addition the Guide may include explanatory text found necessary as a result of interpretation difficulties apparent from the defect reports.

This Guide will not address proposed additions, deletions or modifications to the Recommendation that are not strictly related to implementation difficulties in the above categories. Proposals for new features should be made in the normal way through contributions to the ITU-T.

3. Policies for updating this document

This document is managed by the ITU-T Study Group 16 Question 6 Rapporteur's Group. It can be revised at any recognized Q.6/16 Rapporteur's Group meeting provided the proposed revisions are unanimously accepted by the members of the group. A revision history cataloguing the evolution of this document is included.

4. Defect resolution procedure

Upon discovering technical defects with Recommendation H.263, please provide a written description directly to the Q.6/16 Rapporteur. The template for a defect report is enclosed. Contact information is included in this document. Return contact information should also be supplied so a dialogue can be established to resolve the matter and an appropriate reply to the defect report can be conveyed. This defect resolution process is open to anyone interested in Recommendation H.263. Formal membership in the ITU is not required to participate in this process.

5. References

This document refers to the following Recommendation:

– ITU-T Recommendation H.263 (2001), *Video Coding for low bit rate communication*.

6. Nomenclature

In addition to traditional revision marks, the following marks and symbols are used to indicate to the reader how changes to the text of a Recommendation should be applied:

Symbol	Description
[Begin Correction]	Identifies the start of revision marked text based on extractions from the published Recommendation affected by the correction being described.
[End Correction]	Identifies the end of revision marked text based on extractions from the published Recommendatios affected by the correction being described.
...	Indicates that the portion of the Recommendation between the text appearing before and after this symbol has remained unaffected by the correction being described and has been omitted for brevity.
--- SPECIAL INSTRUCTIONS --- ***{instructions}***	Indicates a set of special editing instructions to be followed.

7. Technical and editorial corrections

7.1 Figure 8/H.263

In the published version of H.263, Figure 8 had some lines shifted a little to the left. The same information is shown correctly in Figure 6 (part 2). The incorrect format of Figure 8 seems to indicate that TRP rather than BCI will follow BCM and that RPRP will not follow BCI.

[Begin Correction]

Replace figure 8 which appears as:

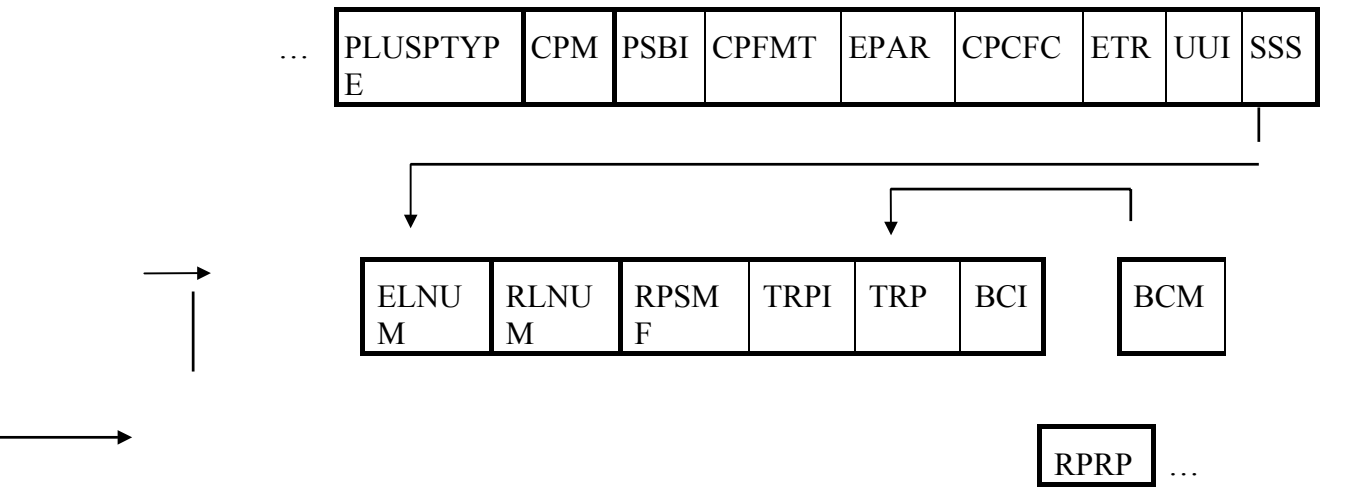


FIGURE 8/H.263

Structure of optional PLUSPTYPE-related fields (located immediately after PTYPE when present)

with the following figure:

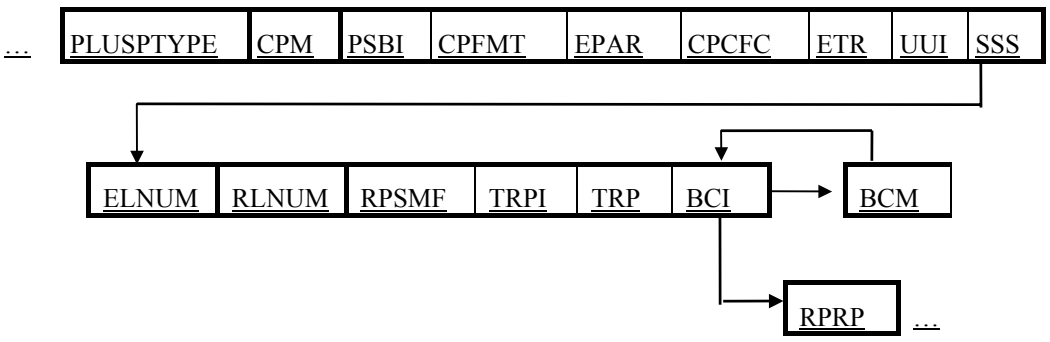


FIGURE 8/H.263

Structure of optional PLUSPTYPE-related fields (located immediately after PTYPE when present)

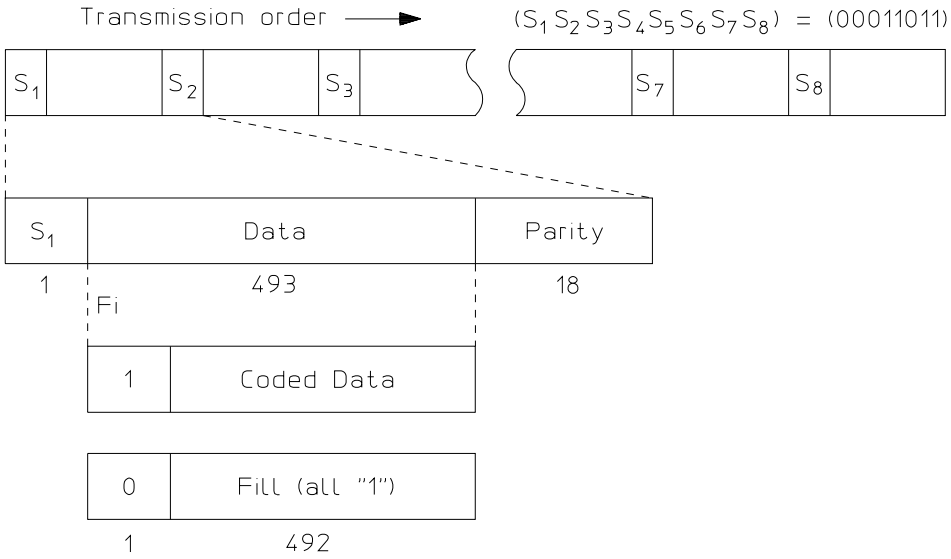
[End Correction]

7.2 Figure H.1/H.263

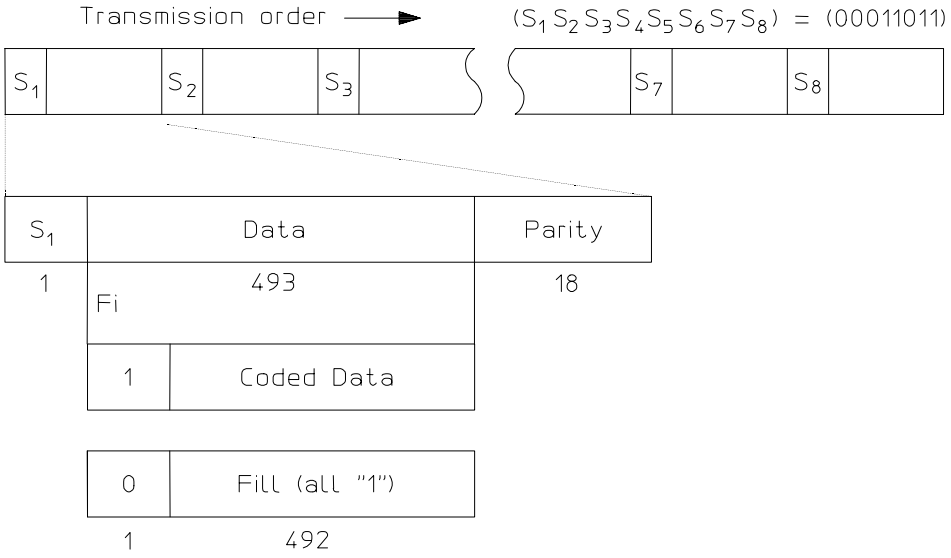
Figure H.1/H.263 has a misleading dotted line, which appears to show the S2 frame alignment bit being within the first 512-bit data frame. Actually, the right dotted line should properly be shown a little further to the left on top, to show that S2 falls into the next data frame.

[Begin Correction]

Replace the following figure



with the following figure



[End Correction]

7.3 BPPmaxKb table and Picture Padding

Table 1/H.263 lists Minimum BPPmaxKb as a function of "Y Picture size in pixels". It is not stated whether this is the size of the picture specified in the picture header (width x height) or whether this is the size of the picture after the width and height have been padded out to the nearest multiple of 16.

The correct interpretation is that the size specified in the table is the size of the picture specified in the picture header (i.e., the unpadded size).

This interpretation is supported by noting the 16-pixel gaps in size found in the table, which signal the intention for this to specify the unpadded size (a multiple of 16 since the width and height are each multiples of 4). Also, this is the most conservative answer as it ensures interoperability regardless of which size is assumed in a decoder design.

[Begin Correction]

TABLE 1/H.263

Minimum BPPmaxKb for different source picture formats

Y Picture Size in Pixels (= width x height as given in picture header)	Minimum BPPmaxKb
up to 25,344 (or QCIF)	64
25,360 to 101,376 (or CIF)	256
101,392 to 405,504 (or 4CIF)	512
405,520 and above	1024

[End Correction]

7.4 Direct bi-dir predicted MBs and picture extrapolation

Section O.4 states: "In B and EP pictures, motion vectors over picture boundaries may be used as described in section D.1 [..]"

Section D.1.1 states: "If PLUSPTYPE is present in the picture header, the motion vector values are restricted such that no element of the 16x16 (or 8x8) region that is selected shall have a horizontal or vertical distance more than 15 pixels outside the coded picture area".

This is a problem for DIRECT predicted macroblocks in B frames for the following reason.

If UMV is on and the new extended motion vectors are used, consider the case where the forward MV is 40 pixels and the B frame is in the middle of the previous and future frames. For a DIRECT

predicted macroblock at the top edge of the frame, the forward motion vector would be 20 pixels and the backward motion vector would be -20 pixels. Unless the DIRECT mode is not used for this macroblock, or unless the backward motion vector is clipped (to –15 pixels), pixels outside the 15 pixels mentioned in section D.1.1 may be accessed.

Two solutions to this problem have been considered: should the encoder make sure no such macroblocks occur, or should the decoder be able to handle this situation gracefully?

The agreed solution is that the encoder shall make sure that no such macroblocks occur, by manipulating mode decisions and/or motion vector values as necessary, as this ensures interoperability regardless of which interpretation is assumed in a decoder.

[Begin Correction]

Modify the text in section O.4/H.263 as follows.

In B and EP pictures, motion vectors over picture boundaries may be used as described in section D.1 of Annex D (although extension of the motion vector range as described in section D.2 is only active if the Unrestricted Motion Vector mode is also in use).

Encoders shall ensure compliance with clause D.1.1 for all macroblocks, including those predicted with DIRECT prediction; that is, if PLUSPTYPE is present in the picture header, the motion vector values shall be restricted such that no element of the 16x16 (or 8x8) region that is selected shall have a horizontal or vertical distance more than 15 pixels outside the coded picture area

[End Correction]

7.5 Parsability of GN/MBA field in BCM

Under the "videomux" mode of Annex N (Reference Picture Selection), a BCM received within a video bitstream refers to the contents of a transmitted video bitstream.

Parsing of the variable length GN/MBA field within the BCM, however, is dependent upon some bitstream's image resolution, use of slices, and use of reduced resolution update. The length of the GN/MBA field is determined by these characteristics. It is considered a little unclear whether the relevant video bitstream is the transmitted or received video bitstream: does Annex N assume that the receiver of the BCM has an awareness of the characteristics of the video bitstream it has transmitted? or does Annex N assume video bitstream symmetry with respect to these characteristics? If neither of these assumptions is made, BCM cannot be parsed correctly in the videomux mode.

The agreed interpretation is that a receiver of BCM data shall be aware of the characteristics of the video bitstream it has transmitted. No symmetry is assumed.

[Begin Correction]

Modify section N.4.2.9 as follows.

N.4.2.9 GOB Number/Macroblock Address (GN/MBA) (5/6/7/9/11/12/13/14 bits):

A GOB number or Macroblock address is present in this field. If the optional Slice Structured mode (see Annex K) is not in use, this field contains the GOB number of the beginning of the video picture segment for which the NACK/ACK message is indicated in the back-channel message. If the optional Slice Structured mode is in use, this field contains the macroblock address of the beginning of the slice for which the NACK/ACK message is indicated in the back-channel message. The length of this field is the length specified elsewhere in this Recommendation for GN or MBA.

NOTE – when this field is received in the videomux mode, the use of the optional Slice Structured mode refers to the use of this mode in the video bitstream to which the BCM applies and not to the video bitstream that transports the BCM data.

[End Correction]

7.6 Macroblock Stuffing Preceding a Start Code

It is not clear whether it is allowed to have a macroblock type coded as "stuffing" immediately followed by a picture, GOB, or slice start code in the bitstream.

The tentative answer is that encoders should not do this, as we believe it was not the intent of the original specification to allow this (with the intent being that after the last macroblock in a GOB, slice, or picture, the data flow is returned to the higher level rather than allowing it to remain at the macroblock level). However, there are verbal reports that some encoders actually do this, so decoders should be designed to allow this. Moreover, we acknowledge that allowing macroblock-level stuffing just prior to a start code would add flexibility to the manner in which encoders control their bit rate, and thus might (in an ideal world) be a desirable feature to support.

[Begin Correction]

Modify the text in section 5.3.2 as follows.

An extra codeword is available in the tables for bit stuffing. This codeword should be discarded by decoders. If an Improved PB-frame is indicated by MPPTYPE bits 1-3 and Custom Source Format is indicated in OPPTYPE bits 1-3, then MBA shall not indicate stuffing before the first macroblock of the picture (in order to prevent start code emulation).

NOTE – Decoders should be designed to allow the macroblock type to indicate bit stuffing immediately prior to the location of a picture, GOB, or slice start code in the bitstream. However, encoders should not use macroblock-layer stuffing in this manner (for interoperability with decoders that may have been designed before the need for decoders to support this was clarified).

[End Correction]

7.7 Interaction Between H.263 Annex J and IDCT Rounding Error

The problem of IDCT mismatch can be aggravated when using H.263 Annex J with a small quantization step size and not using the Annex W IDCT on both sides of the communication link. In some cases the amount of rounding error mismatch may even be amplified by the Annex J filtering process.

[Begin Correction]

Add the following note to the end of section J.1 'Introduction'.

NOTE – The problem of IDCT mismatch can be aggravated when using Annex J with a small quantization step size and not using the Annex W IDCT in both the encoder and decoder. In some cases the amount of rounding error mismatch may even be amplified by the Annex J filtering process. Encoders should avoid this problem, by, for example, not using Annex J when the quantization step size is very small and the Annex W IDCT is not used in both the encoder and decoder.

[End Correction]

7.8 Clarification on the Use of Annex P with Annex N

The use of Annex P together with Annex N is not prohibited in section 5.1.4.6, and a note in section Q.1 seems to indicate that the use of Annex P with Annex N is allowed. However, when Annex P and Annex N are used together it is unclear how the decoder would operate. For example, if Annex N and Annex P are used together, some reference pictures may have different picture resolutions than others, so the same resampling process may not be appropriate for all reference pictures. A suitable clarification seems to be to specify that Annex N operates in a similar manner to Annex U, such that only one reference picture should be active during any invocation of Annex P.

[Begin Correction]

Add a paragraph to section P.1 as follows.

If the Reference Picture Resampling mode is invoked for an Improved PB-frame, one set of warping parameters is sent and the resampled reference picture is used as the reference for both the B- and P-parts of the Improved PB-frame.

The Reference Picture Resampling mode shall not be invoked when the Reference Picture Selection mode (see Annex N) is in use unless the values of TRPI and TRP in all picture, GOB, and slice headers of the current picture specify the use of the same reference picture – in which case the indicated reference picture is the picture that determines whether the reference picture resampling process is to be invoked implicitly and is the picture to which the resampling process shall be applied.

[End Correction]

7.9 Ambiguity in Figure U.8

Figure U.8 is ambiguous. It is not clear whether a backward motion vector (MVDBW) can occur without a forward motion vector (MVDFW) and whether a backward motion vector (MVDBW) can

occur without a B-Picture Selection Bit for Backward Prediction (BSBBW). This can actually happen.

The figure should be corrected to clarify this issue.

[Begin Correction]

Replace the following figure

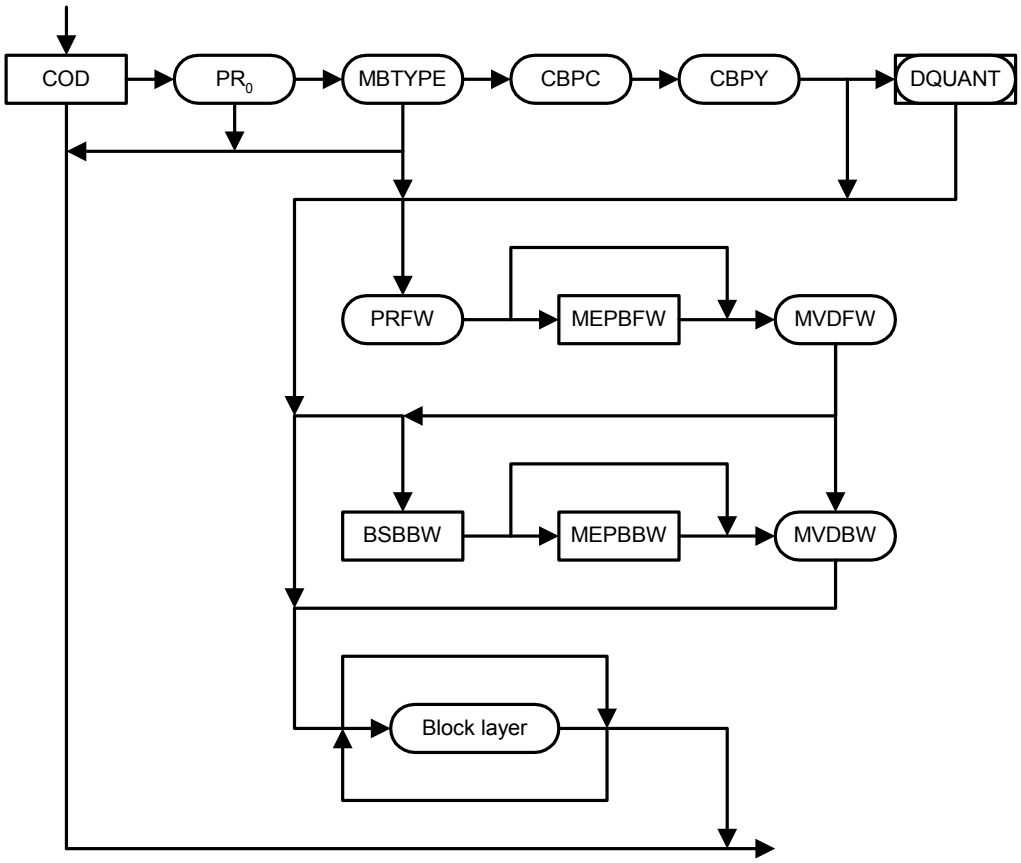


FIGURE U.8/H.263

Structure of EP and B picture Macroblock layer for the ERPS mode

with the following figure:

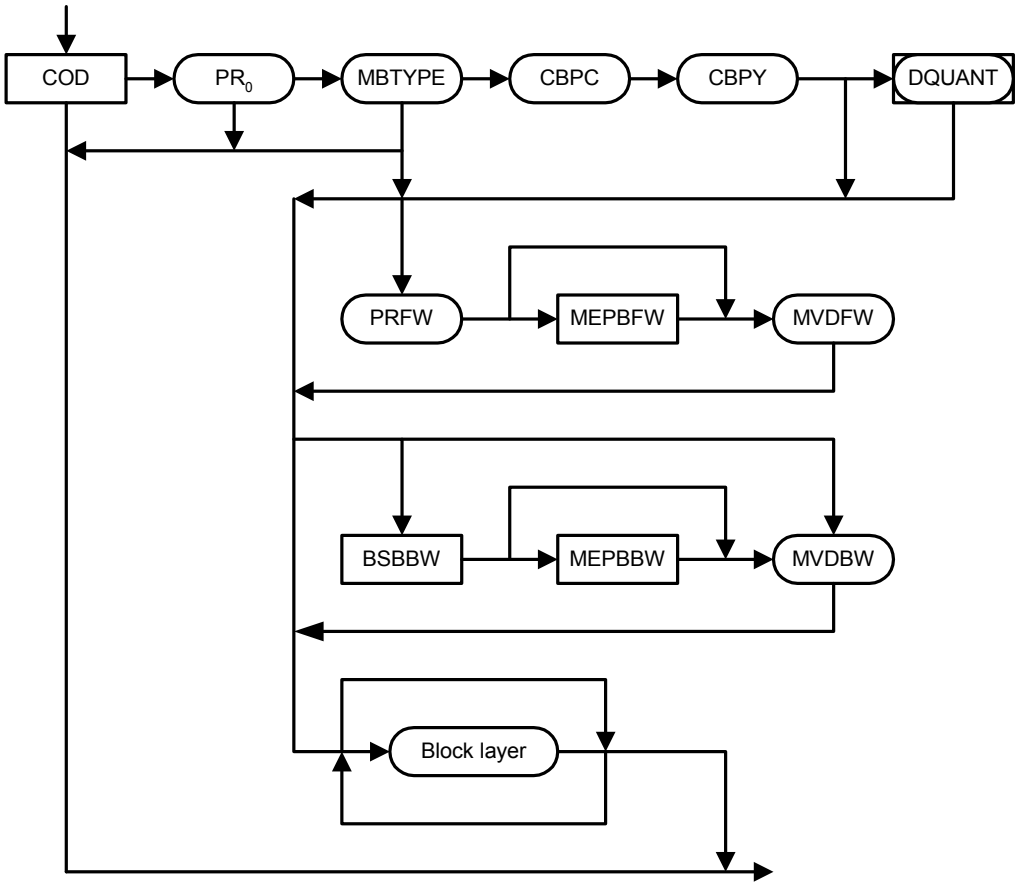


FIGURE U.8/H.263

Structure of EP and B picture Macroblock layer for the ERPS mode

---

[End Correction]

7.10 Clarification on the Use of B pictures with Intra Reference Pictures

Section O.5.2 refers only to section G.4 to specify the "direct" prediction process for B pictures. However, the temporally subsequent reference in Annex G is always the P part of a PB frame, and is thus never an intra picture. Also, there is always a motion vector predictor and MV_D in a PB frames macroblock, even for a macroblock of type intra. In the case of Annex O, the temporally subsequent reference may be an intra picture. No unambiguous statement is present in section G.4 regarding what value to use as the motion vector predictor in this case and the fact that nonzero motion vectors can be inferred for macroblocks of type intra in section G.4 seems to confuse matters further. In this case, it may be unclear what value to use for the motion vector of a macroblock using "direct" prediction in a B picture when the temporally subsequent reference

picture is an intra picture. The obvious choice is that all motion vectors should be assumed to be zero in this case.

[Begin Correction]

Add a statement to section O.5.2 as follows.

O.5.2 Motion vectors in direct mode

For macroblocks coded in direct mode, no vector differences are transmitted. Instead, the forward and backward motion vectors are directly computed from the temporally consecutive P-vector as described in G.4 with the restriction that MV_D is always zero. These derived vectors are not used for prediction of other motion vectors. If the temporally subsequent reference picture is an INTRA picture, all forward and backward motion vectors for use in the prediction process for the direct mode shall have the value zero.

[End Correction]

RECOMMENDATION H.263 DEFECT REPORT FORM

DATE:	
CONTACT INFORMATION **NAME:** **COMPANY:** **ADDRESS:** **TEL:** **FAX:** **EMAIL:**	
DESCRIPTION OF PROBLEM:	
SUGGESTIONS FOR RESOLUTION:	

NOTE - Attach additional pages if more space is required than is provided above.