INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

H.263

Annex U

(11/2000)

SERIES H: AUDIOVISUAL AND MULTIMEDIA SYSTEMS

Infrastructure of audiovisual services – Coding of moving video

---

Video coding for low bit rate communication

Annex U: Enhanced reference picture selection mode

ITU-T Recommendation H.263 – Annex U

(Formerly CCITT Recommendation)

Petitioner has added letters to pages (a)-(b) and number 30 to the last page. Otherwise, it leaves the original page numbering.

LENOVO
EXHIBIT 1008

ITU-T H-SERIES RECOMMENDATIONS

AUDIOVISUAL AND MULTIMEDIA SYSTEMS

CHARACTERISTICS OF VISUAL TELEPHONE SYSTEMS	H.100–H.199
INFRASTRUCTURE OF AUDIOVISUAL SERVICES	
General	H.200–H.219
Transmission multiplexing and synchronization	H.220–H.229
Systems aspects	H.230–H.239
Communication procedures	H.240–H.259
Coding of moving video	**H.260–H.279**
Related systems aspects	H.280–H.299
SYSTEMS AND TERMINAL EQUIPMENT FOR AUDIOVISUAL SERVICES	H.300–H.399
SUPPLEMENTARY SERVICES FOR MULTIMEDIA	H.450–H.499

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

ITU-T Recommendation H.263

Video coding for low bit rate communication

ANNEX U

Enhanced reference picture selection mode

Summary

An optional Enhanced Reference Picture Selection (ERPS) mode capable of providing enhanced coding efficiency and enhanced error resilience (particularly against loss of data packets). The ERPS mode operates by managing a multi-picture buffer of stored pictures.

Source

Annex U to ITU-T Recommendation H.263 was prepared by ITU-T Study Group 16 (2001-2004) and approved under the WTSA Resolution 1 procedure on 17 November 2000.

FOREWORD

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

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

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

NOTE

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

INTELLECTUAL PROPERTY RIGHTS

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

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

© ITU 2001

All rights reserved. No part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from ITU.

CONTENTS

		Page
U.1	Introduction	1
U.2	Video source coding algorithm	2
U.3	Forward-channel syntax	3
	U.3.1 Syntax of the picture, GOB, and slice layer	3
	U.3.2 Macroblock layer syntax	15
U.4	Decoder process	21
	U.4.1 Decoder process for short-/long-term picture management	22
	U.4.2 Decoder process for reference picture buffer mapping	22
	U.4.3 Decoder process for sub-picture removal	23
	U.4.4 Decoder process for multi-picture motion compensation	24
	U.4.5 Decoder process for reference picture buffering	25
U.5	Back-channel messages	26
	U.5.1 BCM separate logical channel layer	26
	U.5.2 Back-channel message layer syntax	26

ITU-T Recommendation H.263

Video coding for low bit rate communication

ANNEX U

Enhanced reference picture selection mode

U.1 Introduction

This annex describes the optional Enhanced Reference Picture Selection (ERPS) mode of this Recommendation. The capability to use this optional mode is negotiated by external means (for example, ITU-T H.245). The amount of picture memory accommodated in the decoder for ERPS operation should also be signalled by external means. The use of this mode shall be indicated by setting the formerly reserved bit 16 of the optional part of the PLUSPTYPE (OPPTYPE) to "1". The mode provides benefits for both error resilience and coding efficiency by using a memory buffer of reference pictures.

A sub-mode of the ERPS mode is specified for Sub-Picture Removal. The purpose of Sub-Picture Removal is to reduce the amount of memory required to store multiple reference pictures. The memory reduction is accomplished by specifying the partitioning of each reference picture into smaller rectangular units called sub-pictures. The encoder can then indicate to the decoder that specific sub-picture areas of specific reference pictures will not be used as a reference for the prediction of subsequent pictures, thus allowing the memory allocated in the decoder for storing these areas to be used to store data from other reference pictures. The support for this sub-mode and the allowed fragmentation of the picture memory into minimum picture units (MPUs) for Sub-Picture Removal as defined herein is also negotiated by external means (for example, ITU-T H.245).

A sub-mode of the ERPS mode is specified for enabling two-picture backward prediction in B pictures. This sub-mode can enhance performance by providing encoders for B pictures not only with an ability to use multiple references for forward prediction, but also to use more than one reference picture for backward prediction. The support for this sub-mode is negotiated by external means (for example, ITU-T H.245).

For error resilience, the ERPS mode can use backward channel messages, which are signalled by external means (for example, ITU-T H.245) sent from a decoder to an encoder to inform the encoder which pictures or parts of pictures have been incorrectly decoded. The ERPS mode provides enhanced performance compared to the Reference Picture Selection (RPS) mode defined in Annex N. It shall not be used simultaneously with the RPS mode. (It can be used in such a way as to provide essentially the same functionality as the RPS mode.)

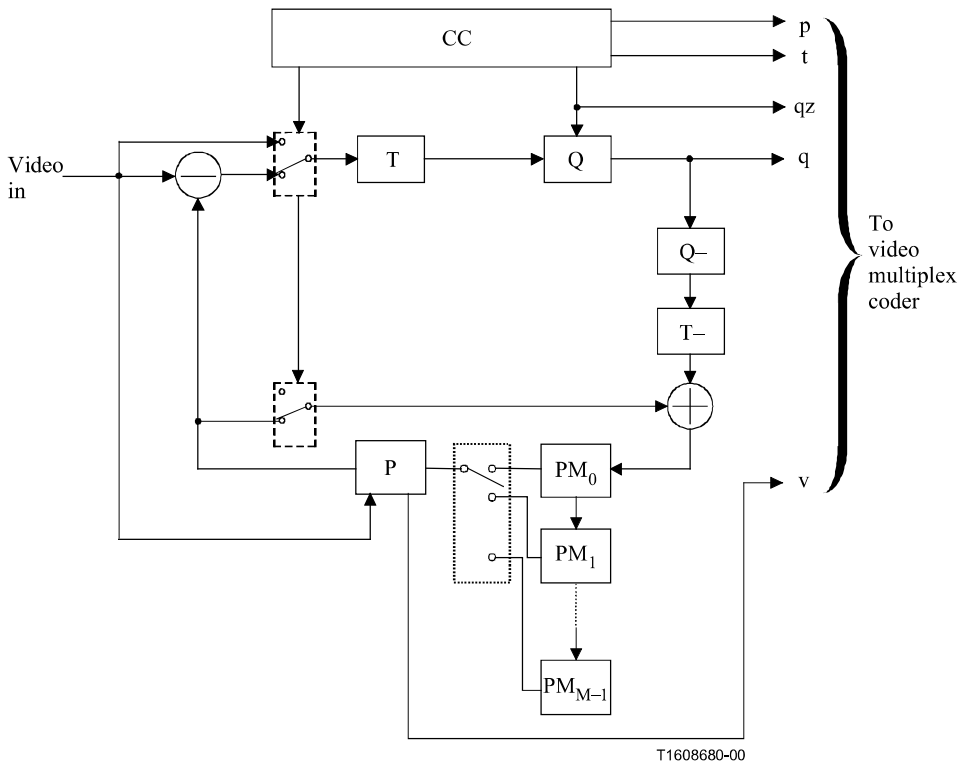
For coding efficiency, motion compensation can be extended to prediction from multiple pictures. The extension of motion compensation to multi-picture prediction is achieved by extending each motion vector by a picture reference parameter that is used to address a macroblock or block prediction region for motion compensation in any of the multiple reference pictures. The picture reference parameter is a variable length code specifying a relative buffer index. The reference pictures are assembled in a buffering scheme that is controlled by the encoder.

The ERPS mode shall not be used with the Syntax-based Arithmetic Coding mode (see Annex E) or the Data Partitioned Slice mode (see Annex V).

Once activated, the ERPS mode shall not be inactivated in subsequent pictures in the bitstream unless the initial inactivation occurs in an I or EI picture and any reactivation is also in an I or EI picture and is accompanied by a buffer reset (RESET equal to "1"). If inactivated, the entire contents of the ERPS multi-picture buffer shall be set to "unused" status.

U.2 Video source coding algorithm

The source coder of this mode is shown in generalized form in Figure U.1. This figure shows a structure that uses a number of picture memories.



T	Transform
Q	Quantizer
P	Picture Memory with motion compensated variable delay
PM	Picture Memory
CC	Coding control
p	Flag for INTRA/INTER
t	Flag for transmitted or not
qz	Quantizer indication
q	Quantizing index for transform coefficients
v	Motion vector

Figure U.1/H.263 – Source coder for Enhanced Reference Picture Selection mode

The video source coding algorithm can be extended to multi-picture motion compensation. Enhanced coding efficiency may be achieved by allowing reference picture selection on the macroblock level. A picture buffering scheme with relative indexing is employed for efficient addressing of pictures in the multi-picture buffer. The multi-picture buffer control may work in two distinct types of operation.

In the first of these two types of operation, a "Sliding Window" over time can be accommodated by the buffer control unit. In such a buffering scheme using M picture memories $PM_0 ... PM_{M-1}$, the most recent preceding (up to M) decoded and reconstructed pictures are stored in the picture memories and can be used as references for decoding. If the number of pictures maximally accommodated by the multi-picture buffer corresponds to M, the motion estimation when coding a picture m, if $0 \le m \le M - 1$, can utilize m pictures. When coding a picture $m \ge M$, the maximum number of pictures M can be used. Alternatively, a second "Adaptive Memory Control" type of operation can be used for a more flexible and specific control of the picture memories than with the simple "Sliding Window" scheme.

The operation of the ERPS mode results in the assignment of "unused" status to some pictures or sub-picture areas of pictures that have been sent to the decoder. Once some picture or area of a picture has been assigned to "unused" status, the bitstream shall not contain any data that causes a reference to any "unused" area for the prediction of subsequent pictures. By managing the assignment of "unused" status to previous pictures, the encoder shall ensure that sufficient memory is available in the decoder to store all data needed for the representation of subsequent pictures. The overall buffer size and structure is conveyed to the decoder in the bitstream, and the encoder shall control the buffer such that the specified total capacity is not exceeded by stored picture data that has not been assigned to "unused" status.

The source coder may select one or several of the picture memories to suppress temporal error propagation caused by inter-picture coding. The Independent Segment Decoding mode (see Annex R), which treats boundaries of GOBs with non-empty headers or slices as picture boundaries, can be used to avoid spatial error propagation due to motion compensation across the boundaries of the GOBs or slices when this mode is applied to a smaller unit than a picture, such as a GOB or slice. The information to signal which picture is selected for prediction is included in the encoded bitstream.

The strategy used by the encoder to select the picture or pictures to be used for prediction is out of the scope of this Recommendation.

U.3 Forward-channel syntax

The syntax is altered in the picture, Group of Blocks (GOB), and slice layers. When indicated by a parameter MRPA being equal to "1", the syntax is also altered in the macroblock layer. In the picture, GOB, and slice layers, an Enhanced Reference Picture Selection layer (ERPS layer) is inserted. In the macroblock layer, picture reference parameters are inserted under certain conditions to enable multi-picture motion compensation.

U.3.1 Syntax of the picture, GOB, and slice layer

The Enhanced Reference Picture Selection syntax for the PLUS header (otherwise as shown in Figure 8) is shown in Figure U.2. The fields of RPSMF, PN, and the ERPS layer are inserted into the PLUS header. The fields of TRPI, TRP, BCI, and BCM are not present (since they are only needed for the RPS mode of Annex N, which is not allowed when the ERPS mode is active).

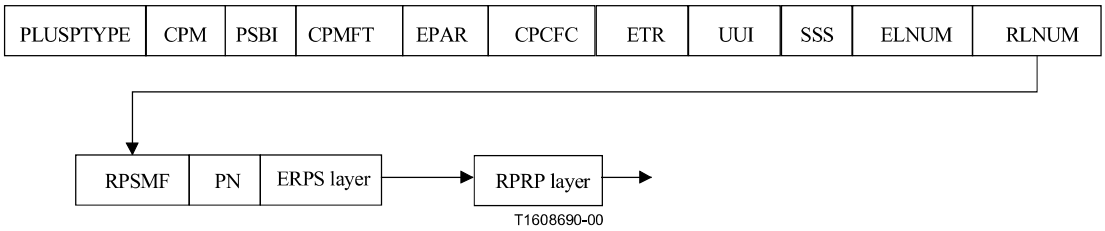


Figure U.2/H.263 – Structure of PLUS header for the ERPS mode

The syntax for the GOB layer is shown in Figure U.3. The fields of PNI, PN, NOERPSL, and the ERPS layer are added to the syntax (otherwise defined as in Figure 9).

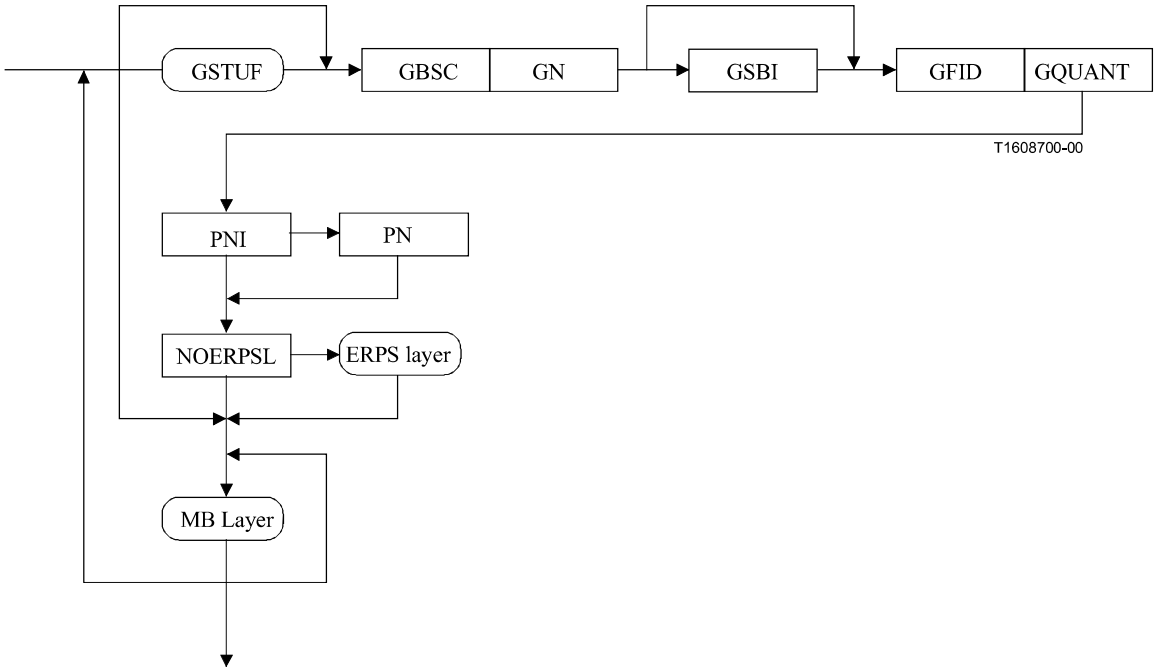


Figure U.3/H.263 – Structure of GOB layer for the ERPS mode

When the optional Slice Structured mode (see Annex K) is in use, the syntax of the slice layer is modified in the same way as the GOB layer. The syntax for the slice layer is shown in Figure U.4. The slice that immediately follows the picture start code in the bitstream also includes all of the added fields PNI, PN, NOERPSL, and the ERPS layer.

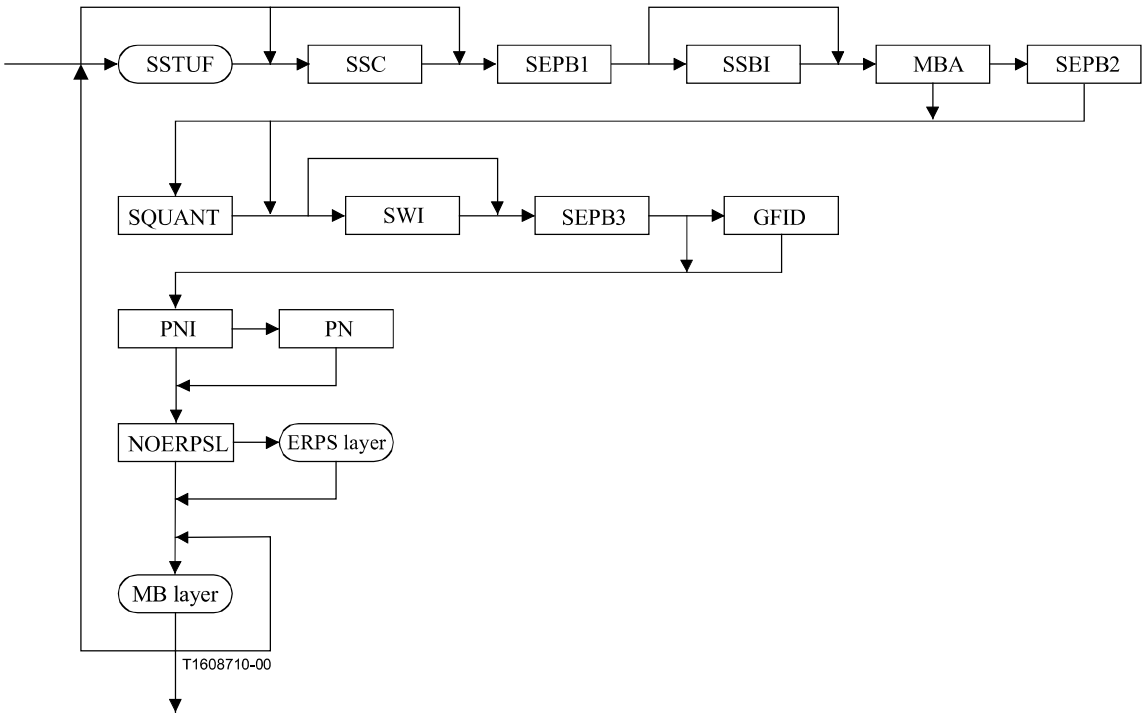


Figure U.4/H.263 – Structure of slice layer for the ERPS mode

The ERPS layer is shown in Figure U.5.

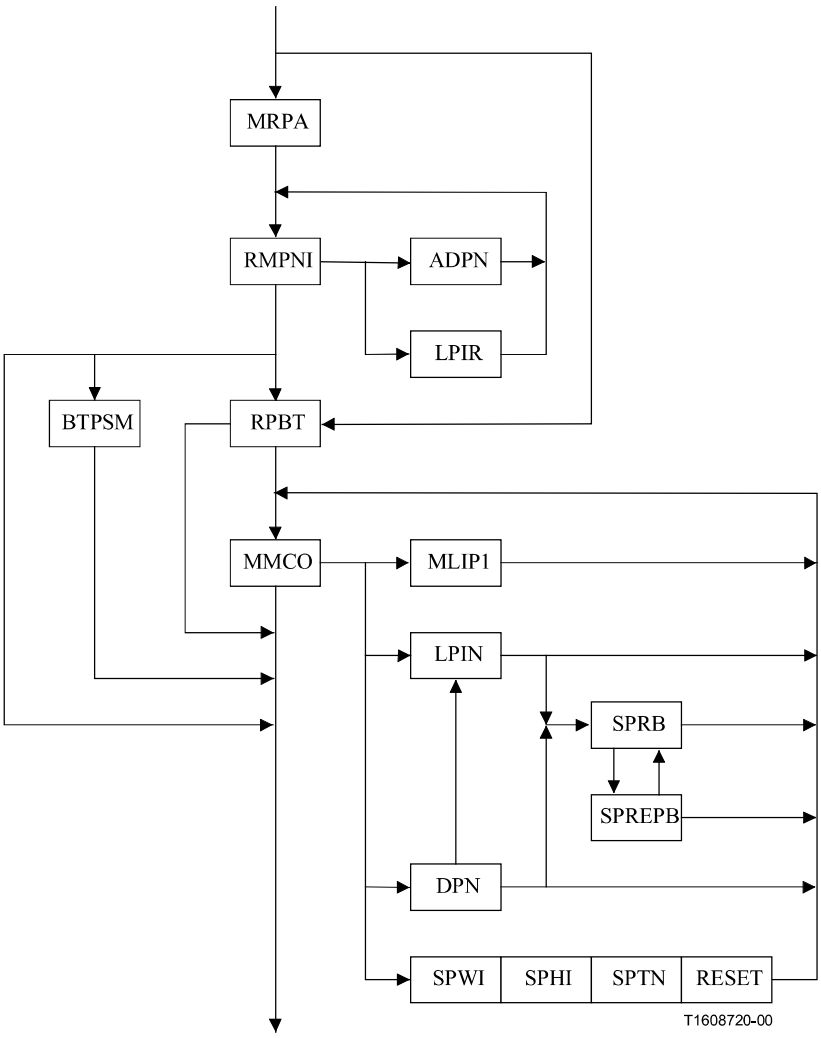


Figure U.5/H.263 – Structure of the ERPS layer

Variable length codes for the ADPN, LPIR, MLIP1, DPN, LPIN, SPTN, PR, PR_0, PR_2, PR_3, PR_4, PRB and PRFW fields are given in Table U.1.

Table U.1/H.263 – Variable length codes for ADPN, LPIR, MLIP1, DPN, LPIN, SPTN, PR, PR_0, PR_2, PR_3, PR_4, PRB, and PRFW

Absolute position	Number of bits	Codes
0	1	1
"x_0"+1 (1:2)	3	$0x_00$
"x_1x_0"+3 (3:6)	5	$0x_11x_00$
"$x_2x_1x_0$"+7 (7:14)	7	$0x_21x_11x_00$
"$x_3x_2x_1x_0$"+15 (15:30)	9	$0x_31x_21x_11x_00$
"$x_4x_3x_2x_1x_0$"+31 (31:62)	11	$0x_41x_31x_21x_11x_00$
"$x_5x_4x_3x_2x_1x_0$"+63 (63:126)	13	$0x_51x_41x_31x_21x_11x_00$
"$x_6x_5x_4x_3x_2x_1x_0$"+127 (127:254)	15	$0x_61x_51x_41x_31x_21x_11x_00$
"$x_7x_6x_5x_4x_3x_2x_1x_0$"+255 (255:510)	17	$0x_71x_61x_51x_41x_31x_21x_11x_00$
"$x_8x_7x_6x_5x_4x_3x_2x_1x_0$"+511 (511:1022)	19	$0x_81x_71x_61x_51x_41x_31x_21x_11x_00$
"$x_9x_8x_7x_6x_5x_4x_3x_2x_1x_0$"+1023 (1023:2046)	21	$0x_91x_81x_71x_61x_51x_41x_31x_21x_11x_00$
"$x_{10}x_9x_8x_7x_6x_5x_4x_3x_2x_1x_0$"+2047 (2047:4094)	23	$0x_{10}1x_91x_81x_71x_61x_51x_41x_31x_21x_11x_00$

U.3.1.1 Reference Picture Selection Mode Flags (RPSMF) (3 bits)

RPSMF is a 3-bit fixed-length codeword that is present in the PLUS header whenever the ERPS mode is in use (regardless of the value of UFEP). RPSMF shall not be present in the GOB or slice layer. When present, RPSMF indicates which type of back-channel messages are needed by the encoder. The values of RPSMF shall be as defined in 5.1.13.

U.3.1.2 Picture Number Indicator (PNI) (1 bit)

PNI is a single-bit fixed-length codeword that is always present at the GOB or slice layer when the ERPS mode is in use, and is not present in the PLUS header. When present, PNI indicates whether or not the following PN field is also present.

"0": PN field is not present.

"1": PN field is present.

U.3.1.3 Picture Number (PN) (10 bits)

PN is a 10-bit fixed-length codeword that is always present in the PLUS header when the ERPS mode is in use, and is present at the GOB or slice layer only when indicated by PNI.

PN shall be incremented by 1 for each coded and transmitted picture, in a 10-bit modulo operation, relative to the PN of the previous stored picture. The term "stored picture" is defined in U.3.1.5.7. For EI and EP pictures, PN shall be incremented be incremented from the value in the last stored EI or EP picture within the same scalability enhancement layer. For B pictures, PN shall be incremented from the value in the most temporally-recent stored non-B picture in the reference layer of the B picture which precedes the B picture in bitstream order (a picture which is temporally subsequent to the B picture). B pictures are not stored in the multi-picture buffer, as they are not used as references for subsequent pictures. Thus, a picture immediately following a B picture in the reference layer of the B picture or another B picture which immediately follows a B picture shall have the same PN as the B picture. Similarly, if a non-B picture is present in the bitstream which is not stored, the picture following this non-B picture (in the same enhancement layer, in the case of Annex O operation) shall have the same PN as the non-stored non-B picture.

In a usage scenario known as "Video Redundancy Coding", the ERPS mode may be used by some encoders in a manner in which more than one representation is sent for the pictured scene at the same temporal instant (usually using different reference pictures). In such a case in which the ERPS mode is in use and in which adjacent pictures in the bitstream have the same temporal reference and the same picture number, the decoder shall regard this occurrence as an indication that redundant copies have been sent of approximately the same pictured scene content, and shall decode and use the first such received picture while discarding the subsequent redundant picture(s).

The PN serves as a unique ID for each picture stored in the picture buffer within 1024 coded and stored pictures. Therefore, a picture cannot be kept in the buffer after more than 1023 subsequent coded and stored pictures (in the same enhancement layer, in the case of Annex O operation) unless it has been assigned a long-term picture index as specified below. The encoder shall ensure that the bitstream shall not specify retaining any short-term picture after more than 1023 subsequent stored pictures. A decoder which encounters a picture number on a current stored picture having a value equal to the picture number of some other short-term picture in the multi-picture buffer should treat this condition as an error.

U.3.1.4 No Enhanced Reference Picture Selection Layer (NOERPSL) (1 bit)

NOERPSL is a single-bit fixed-length codeword that is present at the GOB or slice level whenever the ERPS mode is in use. It is not present in the PLUS header. The values of NOERPSL shall be as follows:

"0": The ERPS layer is sent.

"1": The ERPS layer is not sent.

If NOERPSL is "1", all ERPS settings and re-mappings in effect for the picture shall be applied also for the relevant video picture segment that follows the GOB or slice layer data. ERPS layer information sent at the GOB or slice level governs the decoding process for the video picture segment preceded by the GOB or slice level data, and does not affect the decoding process of any other video picture segment. (See Annex R for the definition of a video picture segment.)

U.3.1.5 Enhanced Reference Picture Selection layer (ERPS) (variable length)

The ERPS layer is always present at the picture level when the ERPS mode is in use, and is present at the GOB or slice level if NOERPSL is "0". It specifies the buffer indexing used to decode the current picture or video picture segment, and manages the contents of the picture buffer.

U.3.1.5.1 Multiple Reference Pictures Active (MRPA) (1 bit)

MRPA is a single-bit fixed-length codeword that is present only if the picture coding type indicates a P picture, an EP picture, an Improved PB frame, or B picture. MRPA is the first element in the ERPS layer if present. MRPA specifies whether the number of active reference pictures for forward-prediction or backward-prediction decoding of the current picture or video picture segment may be larger than one. The value of MRPA shall be as follows:

"1": More than one reference picture may be used for forward or backward motion compensation.

"0": Only one reference picture is used for forward or backward motion compensation. In this case, the extensions of the macroblock layer syntax in U.3.2 do not apply.

MRPA may be changed from video picture segment to video picture segment, so that different video picture segments may address different numbers of reference pictures.

MRPA shall be "0" in any picture which invokes the Reference Picture Resampling mode (see Annex P), and the same picture shall be indicated as the forward reference picture to be used at both the picture and GOB or slice levels for any such current picture. If the current picture is a B picture, the backward reference picture shall have the same size as the current picture, and any reference

picture resampling process shall be applied only to the forward reference picture. Reference picture resampling shall be invoked only if the multi-picture buffer contains sufficient "unused" capacity to store the resampled forward reference picture, but after the resampled reference picture is used for the decoding of the current picture, the resampled forward reference picture shall not be stored in the multi-picture buffer.

U.3.1.5.2 Re-Mapping of Picture Numbers Indicator (RMPNI) (variable length)

RMPNI is a variable-length codeword that is present in the ERPS layer if the picture is a P, EP, Improved PB, or B picture. RMPNI indicates whether any default picture indices are to be re-mapped for motion compensation of the current picture or video picture segment – and how the re-mapping of the relative indices into the multi-picture buffer is to be specified if indicated. RMPNI is transmitted using Table U.2. If RMPNI indicates the presence of an ADPN or LPIR field, an additional RMPNI field immediately follows the ADPN or LPIR field.

Table U.2/H.263 – RMPNI operations for re-mapping of reference pictures

Value	Re-mapping specified
"1"	ADPN field is present and corresponds to a negative difference to add to a picture number prediction value
"010"	ADPN field is present and corresponds to a positive difference to add to a picture number prediction value
"011"	LPIR field is present and specifies the long-term index for a reference picture
"001"	End loop for re-mapping of picture relative indexing default order

A picture reference parameter is a relative index into the ordered set of pictures. The RMPNI, ADPN, and LPIR fields allow the order of that relative indexing into the multi-picture buffer to be temporarily altered from the default index order for the decoding of a particular picture or video picture segment. The default index order is for the short-term pictures (i.e. pictures which have not been given a long-term index) to precede the long-term pictures in the reference indexing order. Within the set of short-term pictures, the default order is for the pictures to be ordered starting with the most recent buffered reference picture and proceeding through to the oldest reference picture (i.e. in decreasing order of picture number in the absence of wrapping of the ten-bit picture number field). Within the set of long-term pictures, the default order is for the pictures to be ordered starting with the picture with the smallest long-term index and proceeding up to the picture with long-term index equal to the most recent value of MLIP1 – 1.

For example, if the buffer contains three short-term pictures with short-term picture numbers 300, 302, and 303 (which were transmitted in increasing picture-number order) and two long-term pictures with long-term picture indices 0 and 3, the default index order is:

- default relative index 0 refers to the short-term picture with picture number 303;
- default relative index 1 refers to the short-term picture with picture number 302;
- default relative index 2 refers to the short-term picture with picture number 300;
- default relative index 3 refers to the long-term picture with long-term picture index 0; and
- default relative index 4 refers to the long-term picture with long-term picture index 3.

The first ADPN or LPIR field that is received (if any) moves a specified picture out of the default order to the relative index of zero. The second such field moves a specified picture to the relative index of one, etc. The set of remaining pictures not moved to the front of the relative indexing order in this manner shall retain their default order amongst themselves and shall follow the pictures that have been moved to the front of the buffer in relative indexing order.

If MRPA is "0", no more than one ADPN or LPIR field shall be present in the same ERPS layer unless the current picture is a B picture. If the current picture is a B picture and MRPA is "0", no more than two ADPN or LPIR fields shall be present in the same ERPS layer.

Any re-mapping of picture numbers specified for some picture shall not affect the decoding process for any other picture. Any re-mapping of picture numbers specified for some video picture segment shall not affect the decoding process for any other video picture segment. A re-mapping of picture numbers specified for a picture shall only affect the decoding process for any video picture segment within that picture in two ways:

- If NOERPSL is "1" at the GOB or slice level, then the re-mapping specified at the picture level is also used for the associated video picture segment.
- If the picture is a B picture, the re-mapping specified at the picture level shall specify the calculation of the value of TR_B and TR_D for direct bidirectional prediction.

An RMPNI "end loop" indication is the last element of the ERPS layer for a B picture if MRPA is "0". In a B picture with MRPA equal to "1", an RMPNI "end loop" indication is followed by BTPSM. In a P or EP picture or Improved PB frame, an RMPNI "end loop" indication is followed by RPBT.

Within one ERPS layer, RMPNI shall not specify the placement of any individual reference picture into more than one re-mapped position in relative index order.

U.3.1.5.3 Absolute Difference of Picture Numbers (ADPN) (variable length)

ADPN is a variable-length codeword that is present only if indicated by RMPNI. ADPN follows RMPNI when present. ADPN is transmitted using Table U.1, where the index into the table corresponds to ADPN – 1. ADPN represents the absolute difference between the picture number of the currently re-mapped picture and the prediction value for that picture number. If no previous ADPN fields have been sent within the current ERPS layer, the prediction value shall be the picture number of the current picture. If some previous ADPN field has been sent, the prediction value shall be the picture number of the last picture that was re-mapped using ADPN.

If the picture number prediction is denoted PNP, and the picture number in question is denoted PNQ, the decoder shall determine PNQ from PNP and ADPN in a manner mathematically equivalent to the following:

```
if (RMPNI == "1") {        // a negative difference
  if (PNP – ADPN < 0)
    PNQ = PNP – ADPN + 1024;
  else
    PNQ = PNP – ADPN;
}else{                     // a positive difference
  if (PNP + ADPN > 1023)
    PNQ  = PNP + ADPN – 1024;
  else
    PNQ  = PNP + ADPN;
}
```

The encoder shall control RMPNI and ADPN such that the decoded value of ADPN shall not be greater than or equal to 1024.

As an example implementation, the encoder may use the following process to determine values of ADPN and RMPNI to specify a re-mapped picture number in question, PNQ:

```
DELTA = PNQ – PNP;
if (DELTA < 0) {
  if (DELTA < –511)
    MDELTA = DELTA + 1024;
  else
```

```
   MDELTA = DELTA;
}else{
 if(DELTA > 512)
   MDELTA = DELTA – 1024;
 else
   MDELTA = DELTA;
}

ADPN = abs(MDELTA);
```

where abs() indicates an absolute value operation. Note that the index into Table U.1 corresponds to the value of ADPN – 1, rather than the value of ADPN itself.

RMPNI would then be determined by the sign of MDELTA.

U.3.1.5.4 Long-term Picture Index for Re-Mapping (LPIR) (variable length)

LPIR is a variable-length codeword that is present only if indicated by RMPNI. LPIR follows RMPNI when present. LPIR is transmitted using Table U.1. It represents the long-term picture index to be re-mapped. The prediction value used by any subsequent ADPN re-mappings is not affected by LPIR.

U.3.1.5.5 B-picture Two-Picture Prediction Sub-Mode (BTPSM) (1 bit)

BTPSM is a single-bit fixed-length codeword that is present only in a B picture (see Annex O) and only when MRPA is "1". It follows an RMPNI "end loop" indication and is the last element of the ERPS layer for the B picture when present. It indicates whether the two-picture backward-prediction sub-mode is in use for the picture as follows:

"0": Single-picture backward prediction.

"1": Two-picture backward prediction.

BTPSM has an implied value of "0" if not present (when MRPA is "0").

The set of pictures available for use as forward-prediction references is the set of pictures in the multi-picture buffer other than the set of backward reference pictures. The set of backward reference pictures is determined by the value of BTPSM. If single-picture backward prediction is specified by BTPSM, the first picture in (possibly re-mapped) relative index order is the only backward reference picture. If two-picture backward prediction is specified by BTPSM, the first two pictures in (possibly re-mapped) relative index order are the two backward reference pictures. The relative index for forward prediction then becomes a relative index into the set of forward reference pictures.

The contents of the multi-picture buffer are not affected by the presence of a B picture. The B picture is not stored in the multi-picture buffer and is not used as a reference for the coding of subsequent pictures.

U.3.1.5.6 Reference Picture Buffering Type (RPBT) (1 bit)

RPBT is a single-bit fixed-length codeword that specifies the buffering type of the currently decoded picture. It follows an RMPNI "end loop" indication when the picture is not an I, EI, or B picture. It is the first element of the ERPS layer if the picture is an I or EI picture. It is not present if the picture is a B picture. The values for RPBT are defined as follows:

"1": Sliding Window.

"0": Adaptive Memory Control.

In the "Sliding Window" buffering type, the current decoded picture shall be added to the buffer with index 0, and any marking of pictures as "unused" in the buffer is performed automatically in a first-in-first-out fashion among the set of short-term pictures. In this case, if the buffer has sufficient "unused" capacity to store the current picture, no additional pictures shall be marked as "unused" in

the buffer. If the buffer does not have sufficient "unused" capacity to store the current picture, the picture (or pictures as necessary to free the needed amount of memory in the case of sub-picture removal) with the largest default index (or indices as necessary in the case of sub-picture removal) among the short-term pictures in the buffer shall be marked as "unused". In the "Sliding Window" buffering type, no additional information is transmitted to control the buffer contents.

In the "Adaptive Memory Control" buffering type, the encoder explicitly specifies any addition to the buffer or marking of data as "unused" in the buffer, and may also assign long-term indices to short-term pictures. The current picture and other pictures may be explicitly marked as "unused" in the buffer, as specified by the encoder. This buffering type requires further information that is controlled by memory management control operation (MMCO) parameters.

RPBT, if present in GOB or slice layers, shall be the same as in the picture layer. Any MMCO command present in GOB or slice layers shall convey the same operation as some MMCO command in the picture layer.

If the picture is a B picture, RPBT shall not be present and the decoded picture shall not be stored in the multi-picture buffer. This ensures that a B picture shall not affect the contents of the multi-picture buffer.

Similarly, the B-picture part of an Improved PB frame shall not be stored in the buffer. All control fields associated with controlling the storage of an Improved PB frame shall be considered to be associated with controlling the storage of only the P-picture part of the Improved PB frame.

U.3.1.5.7 Memory Management Control Operation (MMCO) (variable length)

MMCO is a variable-length codeword that is present only when RPBT indicates "Adaptive Memory Control", and may occur multiple times if present. It specifies a control operation to be applied to manage the multi-picture buffer memory. The MMCO parameter is followed by data necessary for the operation specified by the value of MMCO, and then an additional MMCO parameter follows – until the MMCO value indicates the end of the list of such operations. MMCO commands do not affect the buffer contents or the decoding process for the decoding of the current picture – rather, they specify the necessary buffer status for the decoding of subsequent pictures in the bitstream. The values and control operations associated with MMCO are defined in Table U.3.

Table U.3/H.263 – Memory Management Control Operation (MMCO) Values

Value	Memory Management Control Operation	Associated data fields following
"1"	End MMCO loop	None (end of ERPS layer)
"011"	Mark a short-term picture as "Unused"	DPN
"0100"	Mark a long-term picture as "Unused"	LPIN
"0101"	Assign a long-term index to a picture	DPN and LPIN
"00100"	Mark short-term sub-picture areas as "Unused"	DPN and SPRB
"00101"	Mark long-term sub-picture areas as "Unused"	LPIN and SPRB
"00110"	Specify the maximum long-term picture index	MLIP1
"00111"	Specify the buffer size and structure	SPWI, SPHI, SPTN, and RESET

All memory management control operations specified using MMCO shall be specified in the picture layer. Some or all of the same operations as are specified at the picture layer may also be specified at the GOB or slice layer (with the same associated data). MMCO shall not specify memory operations at the GOB or slice layer that are not also specified with the same associated data at the picture layer.

A buffer size and structure specification MMCO command shall be the first MMCO command if present. No more than one buffer size and structure specification MMCO command shall be present in a given ERPS layer. A buffer size and structure specification MMCO command with RESET equal to "1" shall be present in the first picture in which the ERPS mode is activated in any series of ERPS mode pictures in the bitstream. A buffer size and structure specification MMCO command with RESET equal to "1" shall precede any use of MMCO to indicate marking sub-picture areas of any short-term or long-term pictures as "unused". The sub-picture width and height specified in a buffer size and structure specification MMCO command shall not differ from the value of these parameters in a prior buffer size and structure specification MMCO command unless the current picture is an I or EI picture with RESET equal to "1". The picture height and width shall not change within the bitstream except within a picture containing a buffer size and structure specification MMCO command with RESET equal to "1" (or within a picture in which the ERPS mode is not in use).

If a B picture using single-picture backward prediction is present in the bitstream, exactly one temporally subsequent non-B picture in the reference layer of the B picture shall precede the B picture in bitstream order, as specified in O.2. No memory management control operations shall be present within any ERPS layer of this immediately temporally subsequent non-B picture within the reference layer of the B picture which mark any part of that immediately temporally succeeding non-B picture as "unused", since that reference layer picture is needed for display until after the decoding of the B picture.

The transmission order constraints specified in O.2 are adjusted as necessary for B pictures using two-picture backward prediction. If a B picture using two-picture backward prediction is present in the bitstream, exactly two temporally subsequent non-B pictures in the reference layer of the B picture shall precede the B picture in bitstream order. The other restrictions on the transmission order of the B picture in the bitstream specified in O.2 shall apply, but as adjusted for the use of two temporally subsequent reference layer pictures. No memory management control operations shall be present within any ERPS layer of these two immediately temporally subsequent two non-B pictures within the reference layer of the B picture which mark any part of these two non-B pictures as "unused", since these reference layer pictures are needed for display until after the decoding of the B picture.

A "stored picture" is defined as a non-B picture which does not contain an MMCO command in its ERPS layer which marks that picture as "unused". If the current picture is not a stored picture, its ERPS layer shall not contain any of the following types of MMCO commands:

- an MMCO command to specify the buffer size and structure with RESET equal to "1";
- any MMCO command which marks any other picture as "unused" that has not also been marked as "unused" in the ERPS layer of a prior stored picture;
- any MMCO command which assigns a long-term index to a picture that has not also been assigned the same long-term index in the ERPS layer of a prior stored picture; or
- any MMCO command which marks sub-picture areas of any picture as "unused" that have not also been marked as "unused" in the ERPS layer of a prior stored picture.

U.3.1.5.8 Difference of Picture Numbers (DPN) (variable length)

DPN is present when indicated by MMCO. DPN follows MMCO if present. DPN is transmitted using codewords in Table U.1 and is used to calculate the PN of a picture for a memory control operation. It is used in order to assign a long-term index to a picture, mark a short-term picture as "unused", or mark sub-picture areas of a short-term picture as "unused". If the current decoded picture number is PNC and the decoded value from Table U.1 is DPN, an operation mathematically equivalent to the following equations shall be used for calculation of PNQ, the specified picture number in question:

```
if (PNC – DPN < 0)
  PNQ = PNC – DPN + 1024;
else
  PNQ = PNC – DPN;
```

Similarly, the encoder may compute the DPN value to encode using the following relation:

```
if (PNC – PNQ < 0)
  DPN = PNC – PNQ + 1024;
else
  DPN = PNC – PNQ;
```

For example, if the decoded value of DPN is zero and MMCO indicates marking a short-term picture as "unused", the current decoded picture shall be marked as "unused".

U.3.1.5.9 Long-term Picture Index (LPIN) (variable length)

LPIN is present when indicated by MMCO. LPIN is transmitted using codewords in Table U.1 and specifies the long-term picture index of a picture. It follows DPN if the operation is to assign a long-term index to a picture. It follows MMCO if the operation is to mark a long-term picture as "unused" or to mark sub-picture areas of a long-term picture as "unused".

U.3.1.5.10 Sub-Picture Removal Bit-map (SPRB) (fixed length)

SPRB is a fixed-length codeword that contains one bit for each sub-picture and is present when indicated by MMCO. The number of bits of SPRB data is determined by the most recent values of SPWI and SPHI. SPRB is used to indicate which sub-picture areas of a buffered picture are to be marked as "unused". SPRB follows DPN if the operation is to mark sub-picture areas of a short-term picture as "unused", and follows LPIN if the operation is to mark sub-picture areas of a long-term picture as "unused".

Sub-pictures are numbered in raster scan order starting from the upper-left corner of the picture. For example, consider a case in which a reference picture, specified by DPN, is partitioned into six sub-pictures. Let "s_1 s_2 s_3 s_4 s_5 s_6" represent six bits of SPRB data. If bit s_i is "1", then the decoder should mark the ith sub-picture in the indicated reference picture as "unused". For example, if the SPRB is "000110", then the fourth and fifth sub-pictures areas are marked as "unused".

To prevent start code emulation, all necessary SPREPB emulation prevention bits shall be inserted within or following the SPRB data as specified in U.3.1.5.11.

If SPRB is present and the specified picture has been previously been affected by a prior SPRB bit-map, the bit-map specified by SPRB shall contain a "1" for any sub-picture area that contained a "1" in the previous SPRB bit-map. Every SPRB bit-map shall contain at least one bit having the value "0" and at least one bit having the value "1".

U.3.1.5.11 Sub-Picture Removal Emulation Prevention Bit (SPREPB) (one bit)

SPREPB is a single-bit fixed-length codeword having the value "1" which shall be inserted immediately after any string of 8 consecutive zero bits of SPRB data.

U.3.1.5.12 Maximum Long-term Picture Index Plus 1 (MLIP1) (variable length)

MLIP1 is a variable-length codeword that is present if indicated by MMCO. MLIP1 follows MMCO if present. MLIP1 is transmitted using codewords in Table U.1. If present, MLIP1 is used to determine the maximum index allowed for long-term reference pictures (until receipt of another value of MLIP1). The decoder shall initially assume MLIP1 is "0" until some other value has been received. Upon receiving an MLIP1 parameter, the decoder shall consider all long-term pictures having indices greater than the decoded value of MLIP1 – 1 as "unused" for referencing by the decoding process for subsequent pictures. For all other pictures in the multi-picture buffer, no change of status shall be indicated by MLIP1.

U.3.1.5.13 Sub-Picture Width Indication (SPWI) (7 bits)

SPWI is a fixed-length codeword of 7 bits that is present if indicated by MMCO. SPWI follows MMCO when indicated. SPWI specifies the width of a sub-picture in units of 16 luminance samples, such that the indicated sub-picture width is 16 · (SPWI + 1) luminance samples. The current picture has a width in sub-picture units of ceil(ceil(pw/16)/(SPWI + 1)) sub-pictures, where pw is the width of the picture and "/" indicates floating-point division. For positive numbers, the ceiling function, ceil(x), equals x if x is an integer and otherwise ceil(x) equals one plus the integer part of x. If a minimum picture unit (MPU) size defining the minimum width and height of a sub-picture has been negotiated by external means (for example, ITU-T H.245), the sub-picture width specified by SPWI shall be an integer multiple of the width of the MPU; otherwise, the sub-picture width specified by SPWI shall be such that SPWI is equal to ceil(pw/16) – 1.

U.3.1.5.14 Sub-Picture Height Indication (SPHI) (7 bits)

SPHI is a fixed-length codeword of 7 bits that is present if SPWI is present (as indicated by MMCO). SPHI follows SPWI if present. SPHI specifies the height of a sub-picture in units of 16 luminance samples, such that the indicated sub-picture height is 16 · SPHI. The allowed range of values of SPHI is from 1 to 72. The current picture has a height of ceil(ceil(ph/16)/SPHI) sub-pictures, where ph is the height of the picture and "/" indicates floating-point division. If a minimum picture unit (MPU) size defining the minimum width and height of a sub-picture has been negotiated by external means (for example, ITU-T H.245), the sub-picture height specified by SPHI shall be an integer multiple of the height of the MPU; otherwise, the sub-picture height specified by SPHI shall be such that SPHI is equal to ceil(ph/16).

U.3.1.5.15 Sub-Picture Total Number (SPTN) (variable length)

SPTN is a variable-length codeword that is present if SPWI and SPHI are present (as indicated by MMCO). SPTN follows SPHI if present. SPTN is coded using Table U.1, where the index into Table U.1 corresponds to the decoded value of SPTN – 1. The decoded value of STPN is the total operational size capacity of the multi-picture buffer in units of sub-pictures as specified by SPWI and SPHI. The memory capacity needed for the decoding of current pictures is not included in SPTN – only the memory capacity needed for storing the reference pictures to use for the prediction of other pictures. When sub-picture removal is not in use (i.e. when SPWI and SPHI have whole-picture dimensions), the maximum number of active short-term reference pictures (for example, for sliding window operation) is thus given by SPTN minus the number of pictures that have been assigned to long-term indices and have not been subsequently marked as "unused".

U.3.1.5.16 Buffer Reset Indicator (RESET) (1 bit)

RESET is a single-bit fixed-length codeword that is present if SPWI, SPHI, and SPTN are present (as indicated by MMCO). RESET follows SPTN if present. The values of RESET shall be as follows:

"0": The buffer contents are not reset.

"1": The buffer contents are reset.

If RESET is "1", all pictures in the multi-picture buffer (but not the current picture unless specified separately) shall be marked "unused" (including both short-term and long-term pictures).

U.3.2 Macroblock layer syntax

U.3.2.1 P-picture and Improved PB frames macroblock syntax

The macroblock layer syntax is modified if the ERPS layer is present for P pictures and Improved PB frames when the number of selected forward reference pictures may be greater than one, as indicated by MRPA. The field MRPA is signalled in the ERPS layer. The macroblock layer syntax is

shown in Figure U.6 when MRPA is "1". Otherwise, the macroblock syntax format in a P picture or Improved PB frame is not altered from that shown in Figure 10.

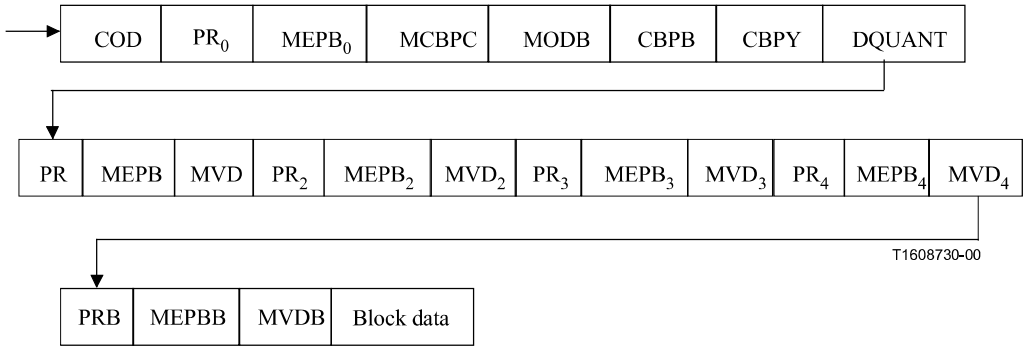


Figure U.6/H.263 – Structure of P-picture and Improved PB frame macroblock layer for the ERPS mode

U.3.2.1.1 Interpretation of COD

If the COD bit is "1", no further information is transmitted for the macroblock. In that case, the decoder shall treat the macroblock as an INTER macroblock with the motion vector for the entire macroblock equal to zero, picture reference parameter equal to zero, and with no coefficient data. If the COD bit is "0", indicating that the macroblock is coded, the syntax of the macroblock layer is depicted in Figure U.6 with the fields PR_0, PR, PR_2, PR_3, PR_4, and PRB being included in the syntax. PR_0, PR, PR_2, PR_3, PR_4, and PRB each consist of a variable length codeword as given in Table U.1.

U.3.2.1.2 Picture Reference Parameter 0 (PR_0) (variable length)

PR_0 is a variable length codeword as specified in Table U.1. It is present whenever COD is "0". If PR_0 has a decoded value of zero (codeword "1"), it indicates that further information will follow for the macroblock. If decoded as non-zero, it indicates the coding of the macroblock using only a picture reference parameter.

If the field PR_0 does not have a decoded value of zero (codeword "1"), no further information is transmitted for this macroblock. In that case the decoder shall treat the macroblock as an INTER macroblock with the motion vector for the whole block equal to zero, the picture reference parameter equal to PR_0, and with no coefficient data.

If the field PR_0 has a decoded value of zero (codeword "1"), the macroblock is coded. The meaning and usage of the fields MCBPC, CBPB, CBPY, and DQUANT remains unaltered. The field PR is included together with the field MVD for all INTER macroblocks (and in Improved PB frames mode also for INTRA macroblocks). The use of MODB in Improved PB frames is described in U.3.2.1.4.

U.3.2.1.3 Macroblock Emulation Prevention Bit 0 ($MEPB_0$) (1 bit)

$MEPB_0$ is a single-bit fixed-length codeword having the value "1" that follows PR_0 if and only if PR_0 is present and has a decoded value of "1" (codeword "000"), and either of the following two conditions are satisfied:

1) the slice structured mode (see Annex K) is in use; or

2) the COD for the current macroblock immediately follows after another macroblock which also has COD = "0" and PR_0 = "1" (codeword "000"), and the PR_0 of the previous macroblock is not followed by an $MEPB_0$ bit.

The purpose of $MEPB_0$ is to prevent start-code emulation and, in the slice structured mode, to aid in determining the number of macroblocks in a slice.

U.3.2.1.4 Macroblock Picture Reference parameters (PR, $PR_{2–4}$, and PRB) (variable length)

PR is the primary picture reference parameter. PR is present whenever MVD is present. The three codewords $PR_{2–4}$ are included together with $MVD_{2–4}$ if indicated by PTYPE and if MCBPC specifies an INTER4V or INTER4V+Q macroblock (a macroblock of type 2 or 5 in Tables 8 and 9). $PR_{2–4}$ and $MVD_{2–4}$ are only present when in Advanced Prediction mode (see Annex F) or Deblocking Filter mode (see Annex J). PRB is only present in an Improved PB frame when MODB indicates that MVDB is present. PR, $PR_{2–4}$, and PRB each specify a picture reference relative index into the multi-picture buffer.

PR is used as the picture reference parameter for motion compensation of the entire macroblock if the macroblock is not an INTER4V or INTER4V+Q macroblock. If the macroblock is an INTER4V or INTER4V+Q macroblock, PR is used for motion compensated prediction of the first of the four 8×8 luminance blocks in the macroblock and for the two chrominance blocks of the macroblock (with the motion compensation process otherwise as specified in 6.1). $PR_{2–4}$ are used for motion compensation of the remaining three 8×8 blocks of luminance data in the macroblock. If MODB indicates that MVDB is present, PRB is the picture reference parameter for forward prediction of the B part of the Improved PB frame.

In Improved PB frames when MODB indicates B_{PB} bidirectional prediction, the values of TR_D and TR_B shall be computed as the temporal reference increments based on the temporal reference data of the current picture and that of the most recent previous reference picture, regardless of whether or not the most recent previous reference picture has been re-mapped to a difference relative index order, marked as "unused", or assigned to a long-term index. The picture used as the forward reference picture for B_{PB} bidirectional prediction in Improved PB frames shall be the picture specified by PR.

U.3.2.1.5 Macroblock Emulation Prevention Bits (MEPB, $MEPB_{2–4}$ and MEPBB) (1 bit each)

MEPB, $MEPB_{2–4}$, and MEPBB are each a single bit having the value "1" if present. Each shall be present if and only if the Unrestricted Motion Vector mode (see Annex D) is not in use and the associated PR, $PR_{2–4}$, or PRB field is present and has the decoded value "1" (codeword "000"). Their purpose is to prevent start-code emulation.

U.3.2.2 B-picture and EP-picture macroblock syntax

The macroblock layer syntax for B and EP pictures (see Annex O) is modified in a similar fashion as in P pictures. The COD bit, if equal to "1", indicates a skipped macroblock as defined in Annex O, using a picture reference parameter of zero for the forward (skipped) prediction in an EP picture and for the forward part of direct (skipped) bidirectional prediction in a B picture and using the first backward prediction picture for the backward part of direct (skipped) bidirectional prediction in a B picture (in the case of two-picture backward prediction, as when BSBBW is present and equal to "0"). If COD is "0", a PR_0 parameter is inserted into the syntax and is used in a similar manner as described in U.3.2.1.2. If PR_0 is present and does not have a decoded value of zero (codeword "1"), it indicates that the macroblock is to be predicted with forward INTER prediction using a zero-valued motion vector and a picture reference parameter of PR_0. If PR_0 has a decoded value of

zero, MBTYPE follows and specifies the macroblock type. The format of the CBPC, CBPY, and DQUANT fields is unchanged. The MVDFW and MVDBW fields are encoded in the same manner as when the ERPS mode is not in use, but are each used in conjunction with a picture reference, and possibly an emulation prevention bit.

For a B picture, the backward reference pictures in the multi-picture buffer are defined as follows:

- in the case of single-picture backward prediction, there is only one backward reference picture, which is the first picture in (possibly re-mapped) relative index order; and
- in the case of two-picture backward prediction, there are two backward reference pictures, which are the first two pictures in (possibly re-mapped) relative index order.

The forward reference pictures in the multi-picture buffer are defined as the pictures in the multi-picture buffer other than the backward reference pictures. The relative indexing for forward prediction is a relative index into the forward reference picture set, and the relative indexing for backward prediction is a relative index into the backward reference picture set.

For example, if the buffer contains three short-term pictures with short-term picture numbers 300, 302, and 303 (which were transmitted in increasing picture-number order) and two long-term pictures with long-term picture indices 0 and 3, the default index order in the case of two-picture backward prediction is:

- default backward relative index 0 refers to the short-term picture with picture number 303;
- default backward relative index 1 refers to the short-term picture with picture number 302;
- default forward relative index 0 refers to the short-term picture with picture number 300;
- default forward relative index 1 refers to the long-term picture with long-term picture index 0; and
- default forward relative index 2 refers to the long-term picture with long-term picture index 3;

and in the case of single-picture backward prediction:

- the single default backward reference picture is the short-term picture with picture number 303;
- default forward relative index 0 refers to the short-term picture with picture number 302;
- default forward relative index 1 refers to the short-term picture with picture number 300;
- default forward relative index 2 refers to the long-term picture with long-term picture index 0; and
- default forward relative index 3 refers to the long-term picture with long-term picture index 3;

and if these pictures have been re-mapped to a new relative indexing order of short-term picture 302, followed by short-term picture 303, followed by long-term picture 0, followed by short-term picture 300, followed by long-term picture 3, the new relative index order in the case of two-picture backward prediction is:

- re-mapped backward relative index 0 refers to the short-term picture with picture number 302;
- re-mapped backward relative index 1 refers to the short-term picture with picture number 303;
- re-mapped forward relative index 0 refers to the long-term picture with long-term picture index 0;
- re-mapped forward relative index 1 refers to the short-term picture with picture number 300; and

- re-mapped forward relative index 2 refers to the long-term picture with long-term picture index 3;

and in the case of single-picture backward prediction:

- the single re-mapped backward reference picture is the short-term picture with picture number 302;
- re-mapped forward relative index 0 refers to the short-term picture with picture number 303;
- re-mapped forward relative index 1 refers to the long-term picture with long-term picture index 0;
- re-mapped forward relative index 2 refers to the short-term picture with picture number 300; and
- re-mapped forward relative index 3 refers to the long-term picture with long-term picture index 3.

The TR_D used for direct bidirectional prediction in a B picture shall be computed as the temporal reference increment between the first forward reference picture in (possibly re-mapped) relative index order and the first backward reference picture in (possibly re-mapped) relative index order (i.e. if two-picture backward prediction is in use, this would be the picture referenced when BSBBW is "0" as described in U.3.2.2.3). The TR_B used for direct bidirectional prediction in a B picture shall be computed as the temporal reference increment between the B picture and the first forward reference picture in (possibly re-mapped) relative index order. The relative index order used in the computation of TR_D and TR_B shall be that specified by the ERPS layer at the picture level of the B picture syntax (i.e. re-mappings at the GOB or slice level shall not affect the values of TR_D and TR_B). (See Figure U.7.)

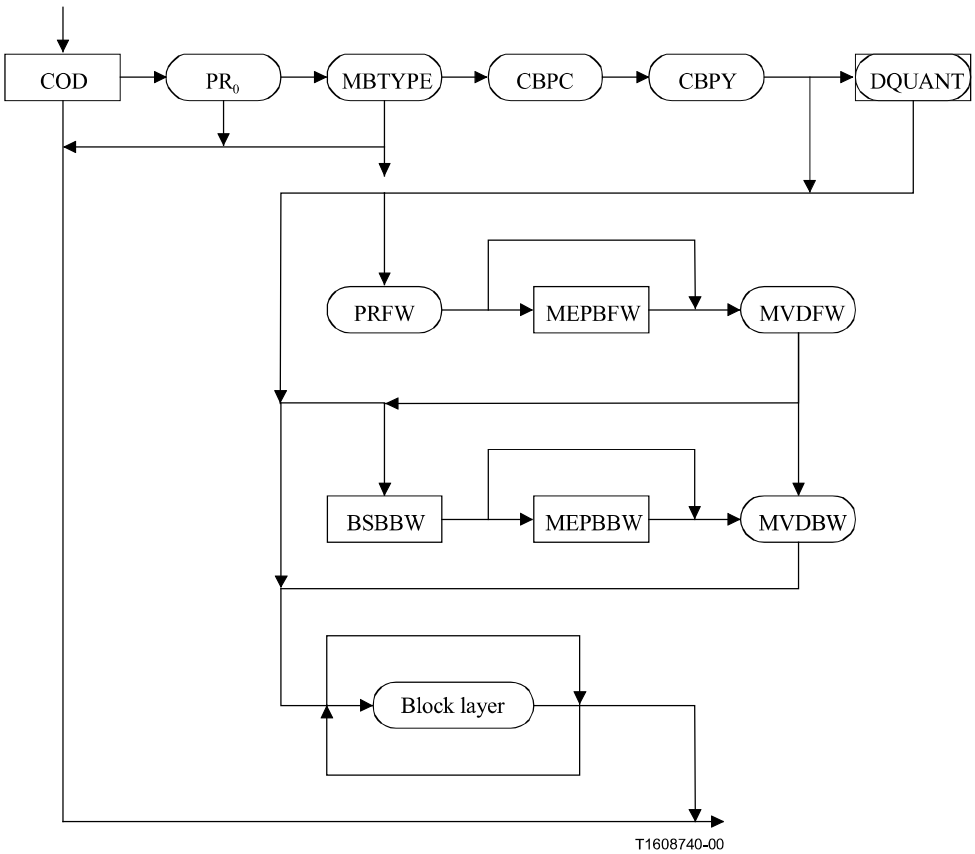


Figure U.7/H.263 – Structure of EP- and B-picture macroblock layer for the ERPS mode

U.3.2.2.1 Picture Reference for Forward Prediction (PRFW) (variable length)

PRFW is a variable-length picture reference parameter that is present whenever forward motion vector data is present, and is encoded using Table U.1. PRFW is a relative index into the set of forward reference pictures.

U.3.2.2.2 Emulation Prevention Bit for Forward Prediction (MEPBFW) (1 bit)

MEPBFW is a single-bit fixed-length codeword having the value "1" which shall be inserted after PRFW if and only if PRFW is present and has a decoded value of "1" (codeword "000") and the unrestricted motion vector mode (see Annex D) is not in use.

U.3.2.2.3 B-picture Selection Bit for Backward Prediction (BSBBW) (1 bit)

BSBBW is a single-bit fixed-length codeword that is present only for B pictures when MVDBW is present and only when two-picture backward prediction is specified for the B-picture operation. The meaning of this bit shall be defined as:

"0": Prediction from the first backward reference picture in relative index order (in default order, this would be the most recent short-term reference picture if that picture has not been assigned a long-term index or marked as "unused").

"1": Prediction from the second backward reference picture in relative index order (in default order, this would be the second-most recent short-term reference picture if neither of the last two reference pictures has been assigned a long-term index or marked as "unused").

U.3.2.2.4 Emulation Prevention Bit for Backward Prediction (MEPBBW) (1 bit)

MEPBBW is a single-bit fixed-length codeword having the value "1" that is present only under the following conditions:

- BSBBW is present and equal to "0"; and
- the unrestricted motion vector mode (see Annex D) is not in use; and
- BSBBW is preceded by five bits having the value "00000".

U.4 Decoder process

The decoder for the ERPS mode stores the reference pictures for inter-picture decoding in a multi-picture buffer. The decoder may need additional memory capacity to store the multiple decoded pictures (relative to the memory capacity needed without support of the ERPS mode). The decoder replicates the multi-picture buffer of the encoder according to the reference picture buffering type and any memory management control operations specified in the bitstream. The buffering scheme may also be operated when partially erroneous pictures are decoded.

Each transmitted and stored picture is assigned a Picture Number (PN) which is stored with the picture in the multi-picture buffer. PN represents a sequential picture counting identifier for stored pictures. PN is constrained, using modulo 1024 arithmetic operation. For the first transmitted picture, PN should be "0". For each and every other transmitted and stored picture, PN shall be increased by 1 (within a given scalability layer, if Annex O is in use). If the difference (modulo 1024) of the PNs of two consecutively received and stored pictures is not 1, the decoder should infer a loss of pictures or corruption of data. In such a case, a back-channel message indicating the loss of pictures may be sent to the encoder.

Besides the PN, each picture stored in the multi-picture buffer has an associated index, called the default index. When a picture is first added to the multi-picture buffer it is given default index 0 – unless it is assigned to a long-term index. The indices of pictures in the multi-picture buffer are modified when pictures are added to or removed from the multi-picture buffer.

The pictures stored in the multi-picture buffers can also be divided into two categories: long-term pictures and short-term pictures. A long-term picture can stay in the multi-picture buffer for a long time (more than 1023 coded and stored picture intervals). The current picture is initially considered a short-term picture. Any short-term picture can be changed to a long-term picture by assigning it a long-term index according to information in the bitstream. The PN is the unique ID for all short-term pictures in the multi-picture buffer. When a short-term picture is changed to a long-term picture, it is also assigned a long-term picture index (LPIN). A long-term picture index is assigned to a picture by associating its PN to an LPIN. Once a long-term picture index has been assigned to a picture, the only potential subsequent use of the long-term picture's PN within the bitstream shall be in a repetition of the long-term index assignment. The PNs of the long-term pictures are unique within 1024 transmitted and stored pictures. Therefore, the PN of a long-term picture cannot be used for assignment of a long-term index after 1023 transmitted subsequent stored pictures. LPIN becomes the unique ID for the life of a long-term picture.

PN (for a short-term picture) or LPIN (for a long-term picture) can be used to re-map the pictures into re-mapped indices for efficient reference picture addressing.

U.4.1 Decoder process for short-/long-term picture management

The decoder may have both long-term pictures and short-term pictures in its multi-picture buffer. The MLIP1 field is used to indicate the maximum long-term picture index allowed in the buffer. If no prior value of MLIP1 has been sent, no long-term pictures shall be in use, i.e. MLIP1 shall initially have an implied value of "0" upon invocation of the ERPS mode. Upon receiving an MLIP1 parameter, a new MLIP1 shall take effect until another value of MLIP1 is received. Upon receiving a new MLIP1 parameter in the bitstream, all long-term pictures with associated long-term indices greater than or equal to MLIP1 shall be considered marked "unused". The frequency of transmitting MLIP1 is out of the scope of this Recommendation. However, the encoder should send an MLIP1 parameter upon receiving an error message, such as an INTRA request message.

A short-term picture can be changed to a long-term picture by using an MMCO command with an associated DPN and LPIN. The short-term picture number is derived from DPN and the long-term picture index is LPIN. Upon receiving such an MMCO command, the decoder shall change the short-term picture with PN indicated by DPN to a long-term picture and shall assign it to the long-term index indicated by LPIN. If a long-term picture with the same long-term index already exists in the buffer, the previously-existing long-term picture shall be marked "unused". An encoder shall not assign a long-term index greater than MLIP1 – 1 to any picture. If LPIN is greater than MLIP1 – 1, this condition should be treated by the decoder as an error. For error resilience, the encoder may send the same long-term index assignment operation or MLIP1 specification message repeatedly. If the picture specified in a long-term assignment operation is already associated with the required LPIN, no action shall be taken by the decoder. An encoder shall not assign the same picture to more than one long-term index value. If the picture specified in a long-term index assignment operation is already associated with a different long-term index, this condition should be treated as an error. An encoder shall only change a short-term picture to a long-term picture within 1024 transmitted consecutive stored pictures. In other words, a short-term picture shall not stay in the short-term buffer after more than 1023 subsequent stored pictures have been transmitted. An encoder shall not assign a long-term index to a short-term picture that has been marked as "unused" by the decoding process prior to the first such assignment message in the bitstream. An encoder shall not assign a long-term index to a picture number that has not been sent.

U.4.2 Decoder process for reference picture buffer mapping

The decoder employs indices when referencing a picture for motion compensation on the macroblock layer using the fields PR_0, PR, PR_2, PR_3, PR_4, PRB, PRFW, and BSBBW. In pictures other than B pictures, these indices are the default relative indices of pictures in the multi-picture buffer when the fields ADPN and LPIR are not present in the current picture, GOB, or slice layer as applicable, and are re-mapped indices when these fields are present. In B pictures, the first one or two pictures (depending on BTPSM) in relative index order are used for backward prediction, and the forward picture reference parameters specify a relative index into the remaining pictures for use in forward prediction.

The indices of pictures in the multi-picture buffer can be re-mapped onto newly specified indices by transmitting the RMPNI, ADPN, and LPIR fields. RMPNI indicates whether ADPN or LPIR is present. If ADPN is present, RMPNI specifies the sign of the difference to be added to a picture number prediction value. The ADPN value corresponds to the absolute difference between the PN of the picture to be re-mapped and a prediction of that PN value. The first transmitted ADPN is computed as the absolute difference between the PN of the current picture and the PN of the picture to be re-mapped. The next transmitted ADPN field represents the difference between the PN of the previous picture that was re-mapped using ADPN and that of another picture to be re-mapped. The process continues until all necessary re-mapping is complete. The presence of re-mappings specified using LPIR does not affect the prediction value for subsequent re-mappings using ADPN. If RMPNI indicates the presence of an LPIR field, the re-mapped picture corresponds to a long-term picture with a long-term index of LPIR. If any pictures are not re-mapped to a specific order by RMPNI,

these remaining pictures shall follow after any pictures having a re-mapped order in the indexing scheme, following the default order amongst these non-re-mapped pictures.

If the decoder detects a missing picture, it may invoke some concealment process, and may insert an error-concealed picture into the multi-picture buffer. Missing pictures can be identified if one or several picture numbers are missing or if a picture not stored in the multi-picture buffer is indicated in a transmitted ADPN or LPIR. Concealment may be conducted by copying the closest temporally preceding picture that is available in the multi-picture buffer into the position of the missing picture. The temporal order of the short-term pictures in the multi-picture buffer can be inferred from their default index order and PN fields. In addition or instead, the decoder may send a forced INTRA update signal to the encoder by external means (for example, ITU-T H.245), or the decoder may use external means or back-channel messages (for example, ITU-T H.245) to indicate the loss of pictures to the encoder. A concealed picture may be inserted into the multi-picture buffer when using the "Sliding Window" buffering type. If a missing picture is detected when decoding a GOB or Slice layer, the concealment may be applied to the picture as if the missing picture had been detected at the picture layer.

U.4.3 Decoder process for sub-picture removal

Sub-picture removal may be used to reduce the amount of memory required to save multiple reference pictures. In sub-picture removal, each reference picture is partitioned into smaller equal-sized sub-pictures. The memory reduction is accomplished by marking undesired sub-pictures as "unused". The strategy used by the encoder to decide which of the sub-pictures to mark as "unused" is outside the scope of this Recommendation. The encoder signals to the decoder the size of the sub-pictures and which of the sub-pictures to mark as "unused" using MMCO commands in the enhanced reference picture selection (ERPS) layer. The encoder shall not send information in the bitstream that causes any samples in reference pictures or sub-pictures that it has caused to be marked as "unused" to be indicated for use in the prediction of subsequent pictures.

The sub-picture removal capability is negotiated by external means (for example, ITU-T H.245). In addition, the decoder signals, also by external means, the minimum partition unit (MPU) which is described in terms of a minimum width and height (in units of 16 luminance samples) of a sub-picture and the total amount of memory it has available for its multi-picture buffer. Memory management is facilitated by the partition rules described below.

Each reference picture is partitioned into rectangular sub-pictures of equal size. The encoder specifies the sub-picture size which shall be an integer multiple of the MPU. The width and height of the sub-picture shall be integer multiples of the minimum width and height negotiated externally as the MPU. The upper-left-hand corner of the first sub-picture is coincident with the upper-left-hand corner of the reference picture. Consequently, the entire partition may be described by specifying the width and height of a sub-picture. If the picture size is not an integer multiple of the sub-picture size, some sub-pictures may extend beyond the right and bottom boundaries of the reference picture. When a sub-picture that extends past the reference picture boundary is saved, a convenient memory management strategy is to set aside enough memory to save the entire sub-picture, rather than just the memory necessary to save the portion of the reference picture that lies within that sub-picture. This is the convention which shall be followed in any calculation of buffer spare capacity for the purpose of determining buffer fullness (e.g. in order to determine whether to automatically mark buffered pictures as "unused" in "sliding window" operation). A decoder designed such that each sub-picture occupies the same amount of memory will prevent the possibility of memory fragmentation.

An example method designed to access referenced picture samples when sub-picture removal is in use is described briefly as follows. One important element in any reference picture access technique is a mechanism to identify where the samples in each sub-picture are stored in memory. If there are R reference pictures and each picture is partitioned into S sub-pictures then there are a total of $K = R \cdot S$ sub-pictures. For example, the sub-picture in the upper-left hand corner of the first

reference picture number can be considered sub-picture number 0, and the sub-picture to the right of it can be considered sub-picture number 1 and so on in raster scan order progressing from reference picture 1 to R until all K sub-pictures have a label. The total buffer capacity is SPTN sub-picture memory buffers, and SPTN is typically less than K. A K-element array can be defined, subPicMem[K], such that t = subPicMem[k] corresponds to the sub-picture memory area that contains the samples in the kth sub-picture. For example, a case can be considered in which R = 5 reference pictures each have S = 12 sub-pictures. Then the samples for the 6th sub-picture in reference picture 3 would be found in sub-picture memory area t = subPicMem[k] where $k = 3 \cdot S + 6 = 42$.

For example, when referencing samples for motion-compensated prediction of one block of luminance or chrominance data when the Advanced Prediction and Reduced Resolution Update optional modes are not in use, it is necessary to acquire $n \times m$ samples, where n and m may take values of 8 or 9 to accommodate half-integer motion compensation. Since the samples in one block may lie in up to four different sub-pictures, four separate cases must be considered. In all cases, the first step is to find the location in memory that contains the upper-left hand sample (U) of the block to be referenced. The sub-picture containing U can be identified by dividing the horizontal or vertical location of U by the sub-picture width or height. If U lies in sub-picture k, then that sample will be located in the subPicMem[k] sub-picture memory area. Next, if both the sample $m - 1$ samples to the right of U (i.e. the upper-right-hand corner of the block) and the sample $n - 1$ samples down from U (i.e. the lower-left-hand corner of the block) lie in sub-picture k, this can be considered case number one. If the sample $n - 1$ samples down from U lies within k, but the sample $m - 1$ samples to the right of U does not, this can be considered case two. If the sample $m - 1$ samples to the right of U lies within k, but the sample $n - 1$ down does not, this can be considered case three. Otherwise, when both the sample $m - 1$ samples to the right of U and the one $n - 1$ samples down lie outside of sub-picture k, this can be considered case four.

In case number one, all samples in the reference block are contained within the kth sub-picture. In this case, all relevant $n \times m$ samples may be found in sub-picture memory area subPicMem[k] and it is a simple task to access them. In case two, the samples that lie in the kth sub-picture can be obtained from sub-picture memory area subPicMem[k] and the remaining samples can be obtained from subPicMem[k_r] where k_r is the sub-picture to the right of k. In case three, the samples that lie in the kth sub-picture can be obtained from memory area subPicMem[k], and the remaining samples can be obtained from subPicMem[k_d] where k_d is the sub-picture below k. In case four, the samples that lie in the kth sub-picture can be obtained from sub-picture memory area subPicMem[k] and the remaining samples can be obtained from memory areas subPicMem[k_r], subPicMem[k_d] and subPicMem[k_{rd}] where k_r and k_d are defined above and k_{rd} is the sub-picture to the right and below k.

U.4.4 Decoder process for multi-picture motion compensation

Multi-picture motion compensation is applied if the MRPA field indicates the use of more than one reference picture. For multi-picture motion compensation, the decoder chooses a reference picture as indicated using the fields PR_0, PR, PR_2, PR_3, PR_4, PRB, PRFW and BSBBW on the macroblock layer. Once the reference picture is specified, the decoding process for motion compensation proceeds as described in 6.1.

In case four motion vectors per macroblock are used and the MRPA field indicates the use of more than one reference picture, the picture reference index for both chrominance blocks is that associated with the first of the four motion vectors (with the motion compensation process otherwise as specified by 6.1).

U.4.5 Decoder process for reference picture buffering

The buffering of the currently decoded picture can be specified using the reference picture buffering type (RPBT) for non-B pictures. The buffering may follow a first-in, first-out ("Sliding Window") mode. Alternatively, the buffering may follow a customized adaptive buffering ("Adaptive Memory Control") operation that is specified by the encoder in the forward channel. B pictures do not affect buffer contents.

The "Sliding Window" buffering type operates as follows. First, the decoder determine whether the picture can be stored into "unused" buffer capacity. If there is insufficient "unused" buffer capacity, the short-term picture with the largest default index (i.e. the oldest short-term picture in the buffer) shall be marked as "unused". This process is repeated if necessary (in the case of sub-picture removal) until sufficient memory capacity is freed to hold the current decoded picture. The current picture is stored in the buffer and assigned a default relative buffer index of zero. The default relative index of all other short-term pictures is incremented by one. The default relative index of all long-term pictures is incremented by one minus the number of short-term pictures removed.

In the "Adaptive Memory Control" buffering type, specified pictures or sub-picture areas may be removed from the multi-picture buffer explicitly. The currently decoded picture, which is initially considered a short-term picture, may be inserted into the buffer with default relative index 0, may be assigned to a long-term index, or may be marked as "unused" by the encoder. Other short-term pictures may also be assigned to long-term indices. The buffering process shall operate in a manner functionally equivalent to the following: First, the current picture is added to the multi-picture buffer with default relative index 0, and the default relative indices of all other pictures is incremented by one. Then, the MMCO commands are processed:

- If MMCO indicates a reset of the buffer contents by using RESET equal to "1", all pictures in the buffer are marked as "unused" except the current picture (which will be the picture with default relative index 0 since a buffer reset must be the first MMCO command as required by U.3.1.5.7).
- If MMCO indicates a maximum long-term index using MLIP1, all long-term pictures having long-term indices greater than or equal to MLIP1 are marked as "unused" and the default relative index order of the remaining pictures are not affected.
- If MMCO indicates that a picture is to be marked as "unused" in the multi-picture buffer and if that picture has not already been marked as "unused", the specified picture is marked as "unused" in the multi-picture buffer and the default relative index of all subsequent pictures in default order is decremented by one.
- If MMCO indicates that sub-picture areas of some picture are to be marked as "unused" in the multi-picture buffer, the specified sub-picture areas are marked as "unused" and the default relative index order of the pictures is not affected. As required by U.3.1.5.10, not all sub-picture areas of any given picture will be marked "unused" by a sub-picture removal MMCO command (instead, the encoder should send an MMCO command marking the picture as a whole as "unused").
- If MMCO indicates the assignment of a long-term index to a specified short-term picture and if the specified long-term index has not already been assigned to the specified short-term picture, the specified short-term picture is marked in the buffer as a long-term picture with the specified long-term index. If another picture is already present in the buffer with the same long-term index as the specified long-term index, the other picture is marked as "unused". All short-term pictures that were subsequent to the specified short-term picture in default relative index order and all long-term pictures having a long-term index less than the specified long-term index have their associated default relative indices decremented by one. The specified picture is assigned to a default relative index of one plus the highest of the decremented default relative indices, or zero if there are no such decremented indices.

The resulting buffered quantity of pictures or sub-picture regions not marked as "unused" shall not exceed the buffer capacity indicated by the most recent value of SPTN. If the decoder detects this condition, it should be treated as an error.

U.5 Back-channel messages

An out-of-band channel, which need not necessarily be reliable, can be used to convey back-channel messages. The syntax of this out-of-band channel (which could be a separate logical channel, for example using ITU-T H.223 or ITU-T H.225.0) should be the one defined herein. The "videomux" operation of back-channel messages as defined in Annex N is not supported in the ERPS mode.

U.5.1 BCM separate logical channel layer

The BCM layer as specified in U.5.2 should be carried by a BCM separate logical channel layer as shown in Figure U.8.

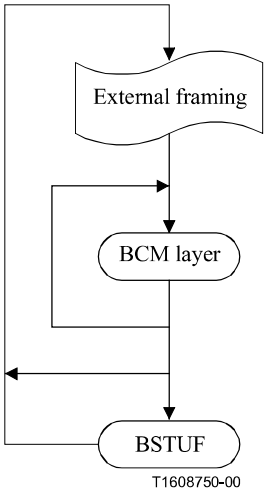


Figure U.8/H.263 – Structure of BCM separate logical channel layer for ERPS mode

U.5.1.1 External framing

External framing of back-channel messages should be provided as shown in Figure U.8. The external framing is used to determine the starting point for the back-channel messages and the amount of back-channel message data to follow.

U.5.1.2 Back-channel Stuffing (BSTUF) (variable length)

BSTUF is a variable-length codeword that may be present only after the last back-channel message in an external frame. BSTUF consists of a codeword of variable length consisting of one or more bits of value "0".

U.5.2 Back-channel message layer syntax

The syntax for the back-channel message (BCM) layer defined herein shall be as shown in Figure U.9.

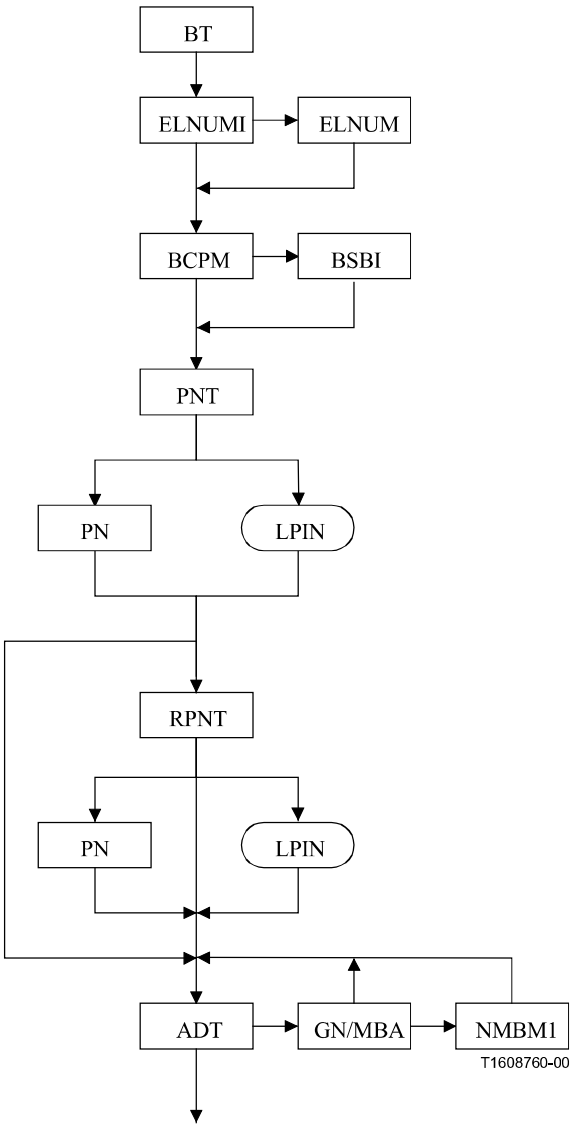


Figure U.9/H.263 – Structure of Back-Channel Message (BCM) layer for ERPS mode

U.5.2.1 Back-channel message Type (BT) (2 bits)

BT is a 2-bit fixed-length codeword which indicates the type of back-channel message. BT is the first codeword present in each back-channel message. Which type or types of message are requested by the encoder is indicated in the RPSMF field of the forward-channel syntax. The values of BT shall be defined as:

"00": Reserved for future use.

"01": Reserved for future use.

"10": NACK. This indicates the loss or erroneous decoding of the corresponding part of the forward channel data.

"11": ACK. This indicates the correct decoding of the corresponding part of the forward channel data.

U.5.2.2 Enhancement Layer Number Indication (ELNUMI) (1 bit)

ELNUMI is a single-bit fixed-length codeword that follows BT in the back-channel message. ELNUMI shall be "0" unless the optional Temporal, SNR, and Spatial Scalability mode (see Annex O) is used in the forward channel and some enhancement layers of the forward channel are combined in one logical channel and the back-channel message refers to an enhancement layer (rather than the base layer), in which case ELNUMI shall be "1".

U.5.2.3 Enhancement Layer Number (ELNUM) (4 bits)

ELNUM is a 4-bit fixed-length codeword that is present only if ELNUMI is "1". It follows ELNUMI if present. When present, ELNUM contains the layer number of the enhancement layer referred to in the back-channel message.

U.5.2.4 Back-channel CPM Indicator (BCPM) (1 bit)

BCPM is a single-bit fixed-length codeword that follows ELNUMI or ELNUM in the back-channel message. BCPM shall be "0" unless the CPM mode (see 5.2.4 and Annex C) is used in the forward channel data, in which case BCPM shall be "1". If BCPM is "1", this indicates that BSBI is present.

U.5.2.5 Back-channel Sub-Bitstream Indicator (BSBI) (2 bits)

BSBI is a 2-bit fixed-length codeword that follows BCPM when present. BSBI is present only if BCPM is "1". BSBI is the natural binary representation of the Sub-Bitstream number in the forward channel data to which the back-channel message refers (see 5.2.4 and Annex C).

U.5.2.6 Picture Number Type (PNT) (1 bit)

PNT is a single-bit fixed-length codeword that is always present and follows BCPM or BSBI in the back-channel message. The values of PNT shall be defined as:

"0": The message concerns a picture specified by a short-term picture number (PN).

"1": The message concerns a picture specified by a long-term picture index (LPIN).

PNT is followed by PN or LPIN, depending on the value of PNT. PN and LPIN shall be represented as specified for use in forward channel data in U.4.1.3 and U.4.1.5.9, respectively.

U.5.2.7 Requested Picture Number Type (RPNT) (2 bits)

RPNT is a 2-bit fixed-length codeword that is present only if BT indicates a NACK message. It follows PN or LPIN when present. It determines how to identify a picture in the multi-picture buffer which may be used as a reference for the coding of subsequent pictures. The values of RPNT shall be defined as:

"00": No valid pictures in buffer – Buffer should be reset by an I or EI picture with RESET equal to "1".

"01": No particular picture is identified to be used as a reference.

"10": A picture which may be used as a reference is identified by a short-term picture number (PN).

"11": A picture which may be used as a reference is identified by a long-term picture index (LPIN).

If RPNT is "10" or "11", RPNT is followed by PN or LPIN, depending on the value of RPNT. PN and LPIN shall be represented as specified for use in forward channel data in U.4.1.3 and U.4.1.5.9, respectively. Typically the PN or LPIN specified using RPNT identifies the last correctly decoded spatially-corresponding picture area for the picture or region identified in the back-channel message.

U.5.2.8 Additional Data Type (ADT) (2 bits)

ADT is a 2-bit fixed-length codeword that is present after PN, LPIN, or RPNT, as determined by PNT (in an ACK message) or RPNT (in a NACK message). It may occur multiple times if present. It specifies the type of additional data used to identify a region of the picture of concern to which the back-channel message applies. The values of ADT shall be defined as:

"00": End of additional data.

"01": A region is identified by only a GN/MBA field.

"10": A region is identified as a raster-scan area within a picture by GN/MBA and NMBM1.

"11": A region is identified as a raster-scan area within a rectangular slice by GN/MBA and NMBM1.

If ADT is "00", no more data follows in the back-channel message. If ADT is "01", ADT is followed by GN/MBA and then by another ADT. If ADT is "10" or "11", ADT is followed by GN/MBA and NMBM1 and then by another ADT.

If ADT is "10", the region is identified as a region starting at a particular spatial location specified by GN/MBA and containing a specified number of macroblocks in raster-scan order within the picture. If ADT is "11", the region is identified as a region starting at a particular spatial location specified by GN/MBA and containing a specified number of macroblocks in raster-scan order within a rectangular slice. If ADT is present only once and is "00", the region identified is the picture as a whole. If ADT is present more than once, the value "00" is used only to end the loop rather than to identify a region.

U.5.2.9 GOB Number/Macroblock Address (GN/MBA) (5/6/7/9/11/12/13/14 bits)

GN/MBA is a fixed-length codeword which specifies a GOB number or macroblock address. GN/MBA follows ADT when present. GN/MBA is present when indicated by ADT. If the optional Slice Structured mode (see Annex K) is not in use, GN/MBA contains the GOB number of the beginning of an area to which the back-channel message refers. If the optional Slice Structured mode is in use, GN/MBA contains the macroblock address of the beginning of the area to which the back-channel message refers. The length of this field shall be as specified elsewhere in this Recommendation for GN or MBA.

U.5.2.10 Number of Macroblocks Minus 1 (NMBM1) (5/6/7/9/11/12/13/14 bits)

NMBM1 is a fixed-length codeword which specifies a number of macroblocks. NMBM1 is present when indicated by ADT. It follows GN/MBA when present. It contains the natural representation of the number of specified macroblocks minus 1. The length of this field shall be the length defined for a macroblock address in K.2.5 and in Table K.2.

SERIES OF ITU-T RECOMMENDATIONS

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