The 4:2:2 profile does not support scalability. This allows implementation architectures to be similar to those of MP@ML.

This profile can be used for applications requiring multiple generations of encoding and decoding. In the case of multiple generations without picture manipulation or change in picture coding type between generations, the quality remains nearly constant after the first generation. Use of picture manipulation or change in picture coding type between generations causes some degradation in quality. Nevertheless, the resulting quality is acceptable for a broad range of applications.

The 4:2:2 profile permits all I-picture encoding. This enables fast recovery from transmission errors and can simplify editing applications. This profile allows the high bit rates required to maintain high quality while using only I-picture coding. The 4:2:2 profile also allows the use of P- and B-picture coding types which can further improve quality or reduce bit rate for the same quality.

See Annex J for more information on the picture quality of the 4:2:2 profile.

NOTE 2 – On Multi-view Profile: The Multi-view Profile (MVP) is envisioned to be a profile appropriate for applications that require multiple viewpoints within the context of ITU-T Rec. H.262 | ISO/IEC 13818-2. MVP supports stereoscopic pictures as its source images for a wide range of picture resolution and quality as requested by the applications to be used. A base layer of MVP is assigned to a left view and an enhancement layer is assigned to a right view.

**Table 8-4 – Escape profile_and_level_indication identification**

| profile_and_level_indication | Name |
|---|---|
| 10001111 to 11111111 | (Reserved) |
| 10001110 | Multi-view profile @ Low level |
| 10001101 | Multi-view profile @ Main level |
| 10001100 | (Reserved) |
| 10001011 | Multi-view profile @ High1440 level |
| 10001010 | Multi-view profile @ High level |
| 10000110 to 10001001 | (Reserved) |
| 10000101 | 4:2:2 profile @ Main level |
| 10000011 to 10000100 | (Reserved) |
| 10000010 | 4:2:2 profile @ High level |
| 10000000 to 10000001 | (Reserved) |

A monoscopic coding with the same tools as Main Profile (MP), including ISO/IEC 11172-2, is applied to the base layer. An enhancement layer is coded using Temporal Scalability tools and a hybrid prediction of motions and disparity can be utilized in the enhanced layer.

MVP, viewed as one of the scalable profiles in terms of multiple viewpoint layers, is expected to have the same type of compatibility features other scalable profiles have, such as compatibility with MP. For example:

1) decoders compliant to MVP at a certain Level are capable of decoding the bitstreams compliant to MP at the corresponding Level (i.e. forward compatibility);
2) decoders compliant to MP at a certain Level are capable of decoding the bitstream in the base layer of MVP (i.e. backward compatibility).

## 8.1 ISO/IEC 11172-2 compatibility

ISO/IEC 11172-2 "constrained parameter" bitstreams shall be decodable by Simple, Main, SNR Scalable, Spatially Scalable and High profile decoders at all levels. Additionally Simple, Main, SNR Scalable, Spatially Scalable and High profile decoders shall be able to decode D-pictures-only bitstreams of ISO/IEC 11172-2 which are within the level constraints of the decoder.

## 8.2 Relationship between defined profiles

The Simple, Main, SNR Scalable, Spatially Scalable and High profiles have a hierarchical relationship. Therefore the syntax supported by a 'higher' profile includes all the syntactic elements of 'lower' profiles (e.g. for a given level, a Main profile decoder shall be able to decode a bitstream conforming to Simple profile restrictions). For a given profile, the same syntax set is supported regardless of level. The order of hierarchy is given in Table 8-2.

The syntactic differences between constraints of profiles are given in Table 8-5. This table describes the limits which apply to a bitstream. Note that a Simple Profile conformant decoder must be able to fully decode both Simple profile, Main level and Main profile, Low level bitstreams.

**Table 8-5 – Syntactic constraints of profiles**

| Syntactic Element | Profile | | | | | | |
|---|---|---|---|---|---|---|---|
| | Simple | Main | SNR | Spatial | High | 4:2:2 | Multi-view |
| **chroma_format** | 4:2:0 | 4:2:0 | 4:2:0 | 4:2:0 | 4:2:2 or 4:2:0 | 4:2:2 or 4:2:0 | 4:2:0 |
| **frame_rate_extension_n** | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **frame_rate_extension_d** | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **aspect_ratio_information** | 0001, 0010, 0011 | 0001, 0010, 0011 | 0001, 0010, 0011 | 0001, 0010, 0011 | 0001, 0010, 0011 | 0001, 0010, 0011 | 0001, 0010, 0011 |
| **picture_coding_type** | I, P | I, P, B | I, P, B | I, P, B | I, P, B | I, P, B | I, P, B |
| **repeat_first_field** | Constrained | | Unconstrained | | | Constrained | Unconstrained |
| **sequence_scalable_ extension()** | No | No | Yes | Yes | Yes | No | Yes |
| **scalable_mode** | – | – | SNR | SNR or Spatial | SNR or Spatial | – | Temporal |
| **picture_spatial_scalable_ extension()** | No | No | No | Yes | Yes | No | No |
| **picture_temporal_scalable_ extension()** | No | No | No | No | No | No | Yes |
| **intra_dc_precision** | 8, 9, 10 | 8, 9, 10 | 8, 9, 10 | 8, 9, 10 | 8, 9, 10, 11 | 8, 9, 10, 11 | 8, 9, 10 |
| **Slice structure** | Restricted 6.1.2.2 | | | | | | |

For all defined profiles, there is a semantic restriction on the bitstream that all of the data for a macroblock shall be represented with not more than the number of bits indicated by Table 8-6. However, a maximum of two macroblocks in each horizontal row of macroblocks may exceed this limitation.

**Table 8-6 – Maximum number of bits in a macroblock**

| chroma_format | Maximum number of bits |
|---|---|
| 4:2:0 | 4608 |
| 4:2:2 | 6144 |
| 4:2:2 (in 4:2:2 Profile) | Unconstrained |
| 4:4:4 | 9216 |

In this context, a macroblock is deemed to start with the first bit of the macroblock_address_increment (or macroblock_escape, if any) and continues until the last bit of the macroblock() syntactic structure. The bits required to represent any slice() that precedes (or follows) the macroblock are not counted as part of the macroblock.

The High profile is also distinguished by having different constraints on luminance sample rate, maximum bit rate, and VBV buffer size. Refer to Tables 8-12, 8-13 and 8-14.

Decoders that are Simple profile @ Main level compliant shall be capable of decoding Main profile @ Low level bitstreams.

### 8.2.1 Use of repeat_first_field

The use of repeat_first_field in Simple and Main profile bitstreams is constrained as specified in Table 8-7.

**Table 8-7 – Constraints on use of repeat_first_field for Simple and Main Profiles**

| frame_rate_code | frame_rate_value | Repeat_first_field | |
|---|---|---|---|
| | | progressive_ sequence==0 | progressive_ sequence==1 |
| 0000 | Forbidden | | |
| 0001 | 24 000 ÷ 1001 (23,976…) | 0 | 0 |
| 0010 | 24 | 0 | 0 |
| 0011 | 25 | 0 or 1 | 0 |
| 0100 | 30 000 ÷ 1001 (29,97…) | 0 or 1 | 0 |
| 0101 | 30 | 0 or 1 | 0 |
| 0110 | 50 | 0 or 1 | 0 |
| 0111 | 60 000 ÷ 1001 (59,94…) | 0 or 1 | 0 or 1 |
| 1000 | 60 | 0 or 1 | 0 or 1 |
| … | Reserved | | |
| 1111 | Reserved | | |

Additional constraints exist for Main profile @ Main level and Simple profile @ Main level only:

- if (vertical_size > 480 lines) or (frame_rate is "25Hz") then if picture_coding_type == 011 (i.e. B-picture), repeat_first_field shall be 0.
- if vertical_size > 480 lines frame_rate shall be "25Hz"

Additionally, the following constraints exist for 4:2:2 profile @ Main level only:

- if vertical_size > 512 lines,

  then if picture_coding_type=011 (i.e. B-picture), repeat_first_field shall be 0;
- if vertical_size > 512 lines frame_rate shall be "25Hz".

## 8.3 Relationship between defined levels

The Low, Main, High-1440 and High levels have a hierarchical relationship. Therefore the parameter constraints of a 'higher' level equal or exceed the constraints of 'lower' levels (e.g. for a given profile, a Main level decoder shall be able to decode a bitstream conforming to Low level restrictions). The order of hierarchy is given in Table 8-3.

The different parameter constraints for levels are given in Table 8-8.

## 8.4 Scalable layers

The SNR Scalable, Spatial Scalable, High and Multi-view profiles may use more than one bitstream to code the image. These different bitstreams represent layers of coding, which when combined create a higher quality image than that obtainable from one layer alone (see Annex D). The maximum number of layers for a given profile is specified in Table 8-9. The scalable layers are named according to Table 7-31. The syntactic and parameter constraints for these profile/level combinations when coded using the maximum permitted number of layers are given in Tables 8-11, 8-12, 8-13 and 8-14. When the number of layers is less than the maximum permitted, reference should also be made to Tables E-21 to E-46 as appropriate.

It should be noted that a bitstream of the base layer of SNR Scalable and Multi-view profiles can always be decoded by a Main profile decoder of equivalent level. Conversely, a Main profile bitstream shall be decodable by either SNR Scalable or Multi-view profile decoder of equivalent level.

**Table 8-8 – Parameter constraints for levels**

| Syntactic Element | Level | | | |
|---|---|---|---|---|
| | Low | Main | High-1440 | High |
| **f_code[0][0]** (forward horizontal ) | [1:7] | [1:8] | [1:9] | [1:9] |
| **f_code[1][0]**[a] (backward horizontal) | [1:7] | [1:8] | [1:9] | [1:9] |
| Frame picture | | | | |
| **f_code[0][1]** (forward vertical) | [1:4] | [1:5] | [1:5] | [1:5] |
| **f_code[1][1]**[a] (backward vertical ) | [1:4] | [1:5] | [1:5] | [1:5] |
| Vertical vector range[b] | [–64:63,5] | [–128:127,5] | [–128:127,5] | [–128:127,5] |
| Field picture | | | | |
| **f_code[0][1]** (forward vertical) | [1:3] | [1:4] | [1:4] | [1:4] |
| **f_code[1][1]**[a] (backward vertical ) | [1:3] | [1:4] | [1:4] | [1:4] |
| Vertical vector range[b] | [–32:31,5] | [–64:63,5] | [–64:63,5] | [–64:63,5] |
| **frame_rate_code** | [1:5] | [1:5] | [1:8] | [1:8] |
| Sample Density | Table 8-11 | | | |
| Luminance Sample Rate | Table 8-12 | | | |
| Maximum Bit Rate | Table 8-13 | | | |
| Buffer Size | Table 8-14 | | | |

a) For Simple profile bitstreams which do not include B-pictures, f_code[1][0] and f_code[1][1] shall be set to 15 (not used).

b) This restriction applies to the final reconstructed motion vector. In the case of dual prime motion vectors, this restriction applies to all the following values:

vector'[0][0][1]
((vector'[0][0][1] * m[parity_ref][parity_pred])//2)
((vector'[0][0][1] * m[parity_ref][parity_pred])//2")" + e[parity_ref][parity_pred]
((vector'[0][0][1] * m[parity_ref][parity_pred])//2")" + dmvector[1]
((vector'[0][0][1] * m[parity_ref][parity_pred])//2")" + e[parity_ref][parity_pred] + dmvector[1]

**Table 8-9 – Upper bounds for scalable layers in SNR Scalable, Spatial Scalable, High and Multi-view profiles**

| Level | Maximum Number of | Profile | | | |
|---|---|---|---|---|---|
| | | SNR | Spatial | High | Multi-view |
| High | All layers (base + enhancement)<br>Spatial enhancement layers<br>SNR enhancement layers<br>Temporal auxiliary layers | | | 3<br>1<br>1<br>0 | 2<br>0<br>0<br>1 |
| High-1440 | All layers (base + enhancement)<br>Spatial enhancement layers<br>SNR enhancement layers<br>Temporal auxiliary layers | | 3<br>1<br>1<br>0 | 3<br>1<br>1<br>0 | 2<br>0<br>0<br>1 |
| Main | All layers (base + enhancement)<br>Spatial enhancement layers<br>SNR enhancement layers<br>Temporal auxiliary layers | 2<br>0<br>1<br>0 | | 3<br>1<br>1<br>0 | 2<br>0<br>0<br>1 |
| Low | All layers (base + enhancement)<br>Spatial enhancement layers<br>SNR enhancement layers<br>Temporal auxiliary layers | 2<br>0<br>1<br>0 | | | 2<br>0<br>0<br>1 |

### 8.4.1 Permissible layer combinations

Table 8-10 is a summary of the permitted combinations, and is subject to the following rules:

- SNR Scalable and Multi-view profile: maximum of 2 layers; Spatial Scalable and High profile: maximum of 3 layers. (See Table 8-9.)
- Only one SNR and one Spatial scale allowed in 3-layer combinations, either SNR/Spatial or Spatial/SNR order is permitted. (See Table 8-9.)
- Chroma simulcast, which allows to add 4:2:2 chroma information to a 4:2:0 base layer and defined in 7.8 is implemented with SNR scalability.
- A 4:2:0 layer is not permitted if the lower layer is 4:2:2. (See 7.7.3.3.)
- (level – 1) is defined as follows:
  - if level is Main, (level – 1) is Low;
  - if level is High – 1440, (level – 1) is Main;
  - if level is High, (level – 1) is High – 1440.

**Table 8-10 – Permissible layer combinations**

| | Scalable mode | | | Profile/level of simplest base layer decoder (level reference top layer) [a] |
|---|---|---|---|---|
| Profile | Base layer | Enhancement layer 1 | Enhancement layer 2 | |
| SNR | 4:2:0 | SNR, 4:2:0 | – | MP@same level |
| Spatial | 4:2:0 | SNR, 4:2:0 | – | MP@same level |
| Spatial | 4:2:0 | Spatial, 4:2:0 | – | MP@(level – 1) |
| Spatial | 4:2:0 | SNR, 4:2:0 | Spatial, 4:2:0 | MP@(level – 1) |
| Spatial | 4:2:0 | Spatial, 4:2:0 | SNR, 4:2:0 | MP@(level – 1) |
| High | 4:2:0 | – | – | HP@same level |
| High | 4:2:2 | – | – | HP@same level |
| High | 4:2:0 | SNR, 4:2:0 | – | HP@same level |
| High | 4:2:0 | SNR, 4:2:2 | – | HP@same level |
| High | 4:2:2 | SNR, 4:2:2 | – | HP@same level |
| High | 4:2:0 | Spatial, 4:2:0 | – | HP@(level – 1) |
| High | 4:2:0 | Spatial, 4:2:2 | – | HP@(level – 1) |
| High | 4:2:2 | Spatial, 4:2:2 | – | HP@(level – 1) [b] |
| High | 4:2:0 | SNR, 4:2:0 | Spatial, 4:2:0 | HP@(level – 1) |
| High | 4:2:0 | SNR, 4:2:0 | Spatial, 4:2:2 | HP@(level – 1) |
| High | 4:2:0 | SNR, 4:2:2 | Spatial, 4:2:2 | HP@(level – 1) [b] |
| High | 4:2:2 | SNR, 4:2:2 | Spatial, 4:2:2 | HP@(level – 1) [b] |
| High | 4:2:0 | Spatial, 4:2:0 | SNR, 4:2:0 | HP@(level – 1) |
| High | 4:2:0 | Spatial, 4:2:0 | SNR, 4:2:2 | HP@(level – 1) |
| High | 4:2:0 | Spatial, 4:2:2 | SNR, 4:2:2 | HP@(level – 1) |
| High | 4:2:2 | Spatial, 4:2:2 | SNR, 4:2:2 | HP@(level – 1) [b] |
| Multi-view | 4:2:0 | Temporal, 4:2:0 | – | MP@same level |

a) The simplest compliant decoder to decode the base layer is specified, assuming that bitstream may contain any syntax and parameter value permitted for the stated profile @ level, except scalability. Note that for High profile @ Main level spatially scaled bitstreams, 'HP @ (level – 1)' becomes 'MP @ (level – 1)'. In the event that a base layer bitstream uses fewer syntactic elements or a reduced parameter range than permitted, profile_and_level_indication may indicate a 'simpler' profile @ level.

b) Note that 4:2:2 chroma format is not supported as a lower spatial layer of High profile @ Main level (see Table 8-12).

Details of the different parameter limits that may be applied in each layer of a bitstream and the corresponding appropriate profile_and_level_indication that should be used are given in Tables E.20 to E.45.

### 8.4.2 Multi-view Profile specific constraints

Both the enhancement and base layers have the same frame rate.

The picture_mux_enable, picture_mux_order and picture_mux_factor are not used in this profile and shall be ignored.

The reference_select_code should be "00" or "01" for the P-frames in the enhancement layer. The reference_select_code should be "01" for B-frames in the enhancement layer.

If the base layer coded frame is the first frame of the Group Of Pictures then the corresponding frame in the enhancement layer should be either I-frame or P-frame with the reference_select_code value of "01".

In a P-field picture with reference_select_code = "01" and which is the first field of a frame, the following restriction applies:

- Dual prime prediction shall not be used.
- Field prediction in which motion_vertical_field_select indicates the second field of the base layer frame shall not be used.
- If base and enhancement layers do not have the same value for top_field_first, there shall be no macroblocks that are coded with macroblock_motion_forward zero and macroblock_intra zero.
- If base and enhancement layer do not have the same value for top_field_first, there shall be no skipped macroblocks.

In a B-field picture which is the first field of a frame, the prediction shall not make reference to the second field of the corresponding base layer frame.

It is inherent in the Multi-view Profile that the two layers are tightly coupled to one another. It is a requirement that the pictures in enhancement layer shall be decoded immediately after their corresponding required reference pictures are decoded unless this requirement makes one to decode the enhancement layer pictures out of display order. In that case, the pictures in the enhancement layer should be decoded in the display order.

## 8.5 Parameter values for defined profiles, levels and layers

See Table 8-11.

**Table 8-11 - Upper bounds for sampling density**

| Level | Spatial resolution layer | | Profile | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Simple | Main | SNR | Spatial | High | 4:2:2 | Multi |
| High | Enhancement | Samples/line<br>Lines/frame<br>Frames/sec | | 1920<br>1088<br>60 | | | 1920<br>1088<br>60 | 1920<br>1088<br>60 | 1920<br>1088<br>60 |
| | Lower | Samples/line<br>Lines/frame<br>Frames/sec | | – | | | 960<br>576<br>30 | – | 1920<br>1088<br>60 |
| High-1440 | Enhancement | Samples/line<br>Lines/frame<br>Frames/sec | | 1440<br>1088<br>60 | | 1440<br>1088<br>60 | 1440<br>1088<br>60 | – | 1440<br>1088<br>60 |
| | Lower | Samples/line<br>Lines/frame<br>Frames/sec | | – | | 720<br>576<br>30 | 720<br>576<br>30 | – | 1440<br>1088<br>60 |
| Main | Enhancement | Samples/line<br>Lines/frame<br>Frames/sec | 720<br>576<br>30 | 720<br>576<br>30 | 720<br>576<br>30 | | 720<br>576<br>30 | 720<br>608 [a]<br>30 | 720<br>576<br>30 |
| | Lower | Samples/line<br>Lines/frame<br>Frames/sec | – | – | – | | 352<br>288<br>30 | – | 720<br>576<br>30 |
| Low | Enhancement | Samples/line<br>Lines/frame<br>Frames/sec | | 352<br>288<br>30 | 352<br>288<br>30 | | | – | 352<br>288<br>30 |
| | Lower | Samples/line<br>Lines/frame<br>Frames/sec | | – | – | | | – | 352<br>288<br>30 |

In the case of single layer or SNR scaled coding, the limits specified by 'Enhancement layer' apply.

[a] 512 lines/frame for 525/60, 608 lines/frame for 625/50

The syntactic elements referenced by this table are as follows:

- samples/line: horizontal_size;
- lines/frame: vertical_size;
- frames/sec: frame_rate.

The upper bound for frame_rate is the same for both progressive_sequence == 0 and progressive_sequence == 1.

**Table 8-12 – Upper bounds for luminance sample rate (samples/s)**

| Level | Spatial resolution layer | Profile | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Simple | Main | SNR | Spatial | High | 4:2:2 | Multi-view |
| High | Enhancement | | 62 668 800 | | | 62 668 800 (4:2:2)<br>83 558 400 (4:2:0) | 62 668 800 | 62 668 800 |
| | Lower | | – | | | 14 745 600 (4:2:2)<br>19 660 800 (4:2:0) | – | 62 668 800 |
| High-1440 | Enhancement | | 47 001 600 | | 47 001 600 | 47 001 600 (4:2:2)<br>62 668 800 (4:2:0) | – | 47 001 600 |
| | Lower | | – | | 10 368 000 | 11 059 200 (4:2:2)<br>14 745 600 (4:2:0) | – | 47 001 600 |
| Main | Enhancement | 10 368 000 | 10 368 000 | 10 368 000 | | 11 059 200 (4:2:2)<br>14 745 600 (4:2:0) | 11 059 200 | 10 368 000 |
| | Lower | – | – | – | | –<br>3 041 280 (4:2:0) | – | 10 368 000 |
| Low | Enhancement | | 3 041 280 | 3 041 280 | | | – | 3 041 280 |
| | Lower | | – | – | | | – | 3 041 280 |
| NOTE – In the case of single layer or SNR scaled coding, the limits specified by "Enhancement layer" apply. | | | | | | | | |

The luminance sample rate P is defined as follows:

- For progressive_sequence == 1:

$$P = (16 * ((horizontal\_size + 15) / 16)) \times (16 * ((vertical\_size + 15) / 16)) \times frame\_rate$$

- For progressive_sequence == 0:

$$P = (16 * ((horizontal\_size + 15) / 16)) \times (32 * ((vertical\_size + 31) / 32)) \times frame\_rate$$

## 8.6 Compatibility requirements on decoders

Table 8-15 defines the requirements on compatibility for decoders. There is a requirement that a decoder of a profile and level represented by a column in Table 8-15 be capable of decoding correctly all bitstreams with profile and level indication marked by an X in the column. In case of scalable hierarchy of bitstreams, the profile and level indication are that of the upper layer.

NOTE – For Profiles and Levels which obey a hierarchical structure, it is recommended that each layer of the bitstream contain the profile_and_level_indication of the "simplest" decoder which is capable of successfully decoding that layer of the bitstream. In the case where the profile_and_level_indication Escape bit == 0, this will be the numerically largest of the possible valid values of profile_and_level_indication.

**Table 8-13 – Upper bounds for bit rates (Mbit/s)**

| Level | Profile | | | | | | |
|---|---|---|---|---|---|---|---|
| | Simple | Main | SNR | Spatial | High | 4:2:2 | Multi-view |
| High | | 80 | | | 100 all layers<br>80 middle + base layer<br>25 base layer | 300 | –<br>130 both layers<br>80 base layer |
| High-1440 | | 60 | | 60 all layers<br>40 middle + base layers<br>15 base layer | 80 all layers<br>60 middle + base layers<br>20 base layer | – | –<br>100 both layers<br>60 base layer |
| Main | 15 | 15 | –<br>15 both layers<br>10 base layer | | 20 all layers<br>15 middle + base layer<br>4 base layer | 50 | –<br>25 both layers<br>15 base layer |
| Low | | 4 | –<br>4 both layers<br>3 base layer | | | – | –<br>8 both layers<br>4 base layer |

NOTE 1 – This table defines the maximum rate of operation of the VBV for a coded bitstream of the given profile and level. This rate is indicated by bit_rate (see 6.3.3).

NOTE 2 – This table defines the maximum permissible data rate for all layers up to and including the stated layer. For multi-layer coding applications, the data rate apportioned between layers is constrained only by the maximum rate permitted for a given layer as stated in this table.

NOTE 3 – 1 Mbit = 1 000 000 bits

**Table 8-14 – VBV buffer size requirements (bits)**

| Level | Layer | Profile | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Simple | Main | SNR | Spatial | High | 4:2:2 | Multi-view |
| High | Enhancement 2<br>Enhancement 1<br>Base | | <br><br>9 781 248 | | | 12 222 464<br>9 781 248<br>3 047 424 | <br><br>47 185 920 | –<br>15 898 480<br>9 787 248 |
| High-1440 | Enhancement 2<br>Enhancement 1<br>Base | | <br><br>7 340 032 | | 7 340 032<br>4 882 432<br>1 835 008 | 9 781 248<br>7 340 032<br>2 441 216 | – | –<br>12 222 464<br>7 340 032 |
| Main | Enhancement 2<br>Enhancement 1<br>Base | <br><br>1 835 008 | <br><br>1 835 008 | –<br>1 835 008<br>1 212 416 | | 2 441 216<br>1 835 008<br>475 136 | <br><br>9 437 184 | –<br>3 047 424<br>1 835 008 |
| Low | Enhancement 2<br>Enhancement 1<br>Base | | <br><br>475 136 | –<br>475 136<br>360 448 | | | – | –<br>950 272<br>475 136 |

NOTE 1 – The buffer size is calculated to be proportional to the maximum allowable bit rate, *rounded down* to the nearest multiple of 16 × 1024 bits. The reference value for scaling is the Main profile, Main level buffer size.

NOTE 2 – This table defines the *total* decoder buffer size required to decode all layers up to and including the stated layer. For multi-layer coding applications, the allocation of buffer memory between layers is constrained only by the maximum size permitted for a given layer as stated in this table.

NOTE 3 – The syntactic element corresponding to this table is vbv_buffer_size (see 6.3.3).

**Table 8-15 – Forward compatibility between different profiles and levels**

| Profile and Level indication in bitstream | Decoder | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | HP @ HL | HP @ H-14 | HP @ ML | Spatial @ H-14 | SNR @ ML | SNR @ LL | MP @ HL | MP @ H-14 | MP @ ML | MP @ LL | SP @ ML | 4:2:2P @ ML | 4:2:2P @ HL | MVP @ HL | MVP @ H-14 | MVP @ ML | MVP @ LL |
| HP@HL | X | | | | | | | | | | | | | | | | |
| HP@H-14 | X | X | | | | | | | | | | | | | | | |
| HP@ML | X | X | X | | | | | | | | | | | | | | |
| Spatial@H-14 | X | X | | X | | | | | | | | | | | | | |
| SNR@ML | X | X | X | X | X | | | | | | | | | | | | |
| SNR@LL | X | X | X | X | X | X | | | | | | | | | | | |
| MP@HL | X | | | | | | X | | | | | | X[c] | X | | | |
| MP@H-14 | X | X | | X | | | X | X | | | | | X[c] | X | X | | |
| MP@ML | X | X | X | X | X | | X | X | X | | | X[b] | X[c] | X | X | X | |
| MP@LL | X | X | X | X | X | X | X | X | X | X | X[a] | X[b] | X[c] | X | X | X | X |
| SP@ML | X | X | X | X | X | | X | X | X | | X | X[b] | X[c] | X | X | X | |
| ISO/IEC 11172-2 | X | X | X | X | X | X | X | X | X | X | X | X[b] | X[c] | X | X | X | X |
| 4:2:2@ML | | | | | | | | | | | | X | X[c] | | | | |
| 4:2:2@HL | | | | | | | | | | | | | X | | | | |
| MVP@HL | | | | | | | | | | | | | | X | | | |
| MVP@H-14 | | | | | | | | | | | | | | X | X | | |
| MVP@ML | | | | | | | | | | | | | | X | X | X | |
| MVP@LL | | | | | | | | | | | | | | X | X | X | X |

X indicates that the decoder shall be able to decode the bit stream including all relevant lower layers.

a) SP@ML decoders are required to decode MP@LL bitstreams.

b) A 4:2:2 profile@Main level decoder shall be able to decode Main profile@Main level, Main profile@Low level and Simple profile@Main level bitstreams, as well as ISO/IEC 11172-2 constrained system parameter bitstreams.

c) A 4:2:2 profile@High level decoder shall be able to decode 4:2:2P@ML, MP@HL, MP@HL-1440, MP@ML, MP@LL and SP@ML, as well as ISO/IEC 11172-2 constrained system parameter bitstreams.

# 9 Registration of Copyright Identifiers

## 9.1 General

Parts 1, 2, and 3 of ISO/IEC 13818 provide support for the management of audiovisual works copyrighting. In ITU-T Rec. H.222.0 | ISO/IEC 13818-1 this is by means of a copyright descriptor, while ITU-T Rec. H.262 | ISO/IEC 13818-2 and ISO/IEC 13818-3 contain fields for identifying copyright holders through syntax fields in the elementary stream syntax. This Recommendation | International Standard presents the method of obtaining and registering copyright identifiers in ITU-T Rec. H.262 | ISO/IEC 13818-2.

ITU-T Rec. H.262 | ISO/IEC 13818-2 specifies a unique 32-bit copyright_identifier which is a work type code identifier (such as ISBN, ISSN, ISRC, etc.) carried in the copyright descriptor. The copyright_identifier enables identification of a wide number of Copyright Registration Authorities. Each Copyright Registration Authority may specify a syntax and semantic for identifying the audiovisual works or other copyrighted works within that particular copyright organization through appropriate use of the variable length additional_copyright_info field which contains the copyright number.

In the following subclause and Annexes G, H and I, the benefits and responsibilities of all parties to the registration of copyright_identifier are outlined.

## 9.2 Implementation of a Registration Authority (RA)

ISO/IEC JTC 1 shall call for nominations for an international organization which will serve as the Registration Authority for the **copyright_identifier** as defined in ITU-T Rec. H.262 | ISO/IEC 13818-2. The selected organization shall serve as the Registration Authority. The so-named Registration Authority shall execute its duties in compliance with Annex H of the JTC 1 Directives. The registered copyright_identifier is hereafter referred to as the Registered Identifier (RID).

Upon selection of the Registration Authority, JTC 1 shall require the creation of a Registration Management Group (RMG) which will review appeals filed by organizations whose request for an RID to be used in conjunction with ITU-T Rec. H.262 | ISO/IEC 13818-2 has been denied by the Registration Authority.

Annexes G, H and I to this Specification provide information on the procedures for registering a unique copyright_identifier.

# Annex A

## Inverse discrete transform

(This annex forms an integral part of this Recommendation | International Standard)

The N × N two dimensional DCT is defined as:

$$F(u, v) = \frac{2}{N} C(u)C(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y)\cos \frac{(2x+1)u\pi}{2N} \cos \frac{(2y+1)v\pi}{2N}$$

with $u, v, x, y = 0, 1, 2, \ldots N-1$

where x, y are spatial coordinates in the sample domain

u, v are coordinates in the transform domain

$$C(u), C(v) = \begin{cases} \frac{1}{\sqrt{2}} & \text{for } u\,v = 0 \\ 1 & \text{otherwise} \end{cases}$$

The definition of the DCT (also called forward DCT) is purely informative. Forward DCT is not used by the decoding process described by this Specification.

The mathematical real-number IDCT is defined as:

$$f(x, y) = \frac{2}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} C(u)C(v)F(u, v)\cos \frac{(2x+1)u\pi}{2N} \cos \frac{(2y+1)v\pi}{2N}$$

f (x, y) is a real number.

The mathematical integer-number IDCT is defined as:

$$f'(x, y) = \text{round}(f(x, y))$$

where round() is the rounding to the nearest integer, with half-integer values rounded away from zero. No clamping or saturation is performed.

The saturated mathematical integer-number IDCT is defined as:

$$f''(x, y) = \text{saturate}(f'(x, y))$$

where saturate() is the saturation in the range [–256, 255], defined as:

$$\text{saturate}(x) = \begin{cases} -256 & x < -256 \\ 255 & x > 255 \\ x & -256 \le x \le 255 \end{cases}$$

The IDCT function f[y][x] used in the decoding process may be any of several approximations of the saturated mathematical integer-number IDCT f ''(x, y), provided that it meets all of the following requirements:

1) The IDCT function f[y][x] used in the decoding process shall have values always in the range [–256, 255].
2) The IDCT function f[y][x] used in the decoding process shall conform to the IEEE Standard Specification for the implementation of 8 by 8 Inverse Discrete Cosine Transform, IEEE Std. 1180-1990, 6 December 1990.
3) This item applies only when input blocks of DCT coefficients cause all the 64 values output of the mathematical integer-number IDCT f '(x, y) to be in the range [–384, 383]. When f '(x, y) > 256, f[y][x] shall be equal to 255 and when f '(x, y) < –257, f[y][x] shall be equal to –256. For all values of f '(x, y) in the range [–257, 256] the absolute difference between f[y][x] and f ''(x, y) shall not be larger than 2.

4) Let F be the set of 4096 blocks Bi[y][x] (i = 0 … 4095) defined as follows:

   a) Bi[0][0] = i – 2048.

   b) Bi[7][7] = 1 if Bi[0][0] is even, Bi[7][7] = 0 if Bi[0][0] is odd.

   c) All other coefficients Bi[y][x] other than Bi[0][0] and Bi[7][7] are equal to 0.

   For each block Bi[y][x] that belongs to set F defined above, an IDCT that conforms to this specification (see Annex A) shall output a block f[y][x] that has a peak error of 1 or less compared to the reference saturated mathematical integer-number IDCT f "(x, y). In other words, | f[y][x] –f "(x, y) | shall be < = 1 for all x and y.

In addition to these requirements, the following is a recommendation on the accuracy of the IDCT function f[y][x].

5) When item 3) does not apply, i.e. for input blocks of DCT coefficients causing the output of the mathematical integer-number IDCT f '(x, y) to contain one or more values out of the range [–384, 383], it is desirable that f[y][x] be as close as possible to f "(x, y) for all bitstreams produced by reasonably well-designed encoders.

NOTE – Clause 2.3 IEEE Standard 1180-1990 "Considerations of Specifying IDCT Mismatch Errors" requires the specification of periodic intra-picture coding in order to control the accumulation of mismatch errors. Every macroblock is required to be refreshed before it is coded 132 times as predictive macroblocks. Macroblocks in B-pictures (and skipped macroblocks in P-pictures) are excluded from the counting because they do not lead to the accumulation of mismatch errors. This requirement is the same as indicated in 1180-1990 for visual telephony according to Recommendation H.261.

# Annex B

## Variable length code tables

(This annex forms an integral part of this Recommendation | International Standard)

### B.1 Macroblock addressing

See Table B.1.

**Table B.1 – Variable length codes for macroblock_address_increment**

| macroblock_address_ increment VLC code | Increment value | macroblock_address_ increment VLC code | Increment value |
|---|---|---|---|
| 1 | 1 | 0000 0101 01 | 18 |
| 011 | 2 | 0000 0101 00 | 19 |
| 010 | 3 | 0000 0100 11 | 20 |
| 0011 | 4 | 0000 0100 10 | 21 |
| 0010 | 5 | 0000 0100 011 | 22 |
| 0001 1 | 6 | 0000 0100 010 | 23 |
| 0001 0 | 7 | 0000 0100 001 | 24 |
| 0000 111 | 8 | 0000 0100 000 | 25 |
| 0000 110 | 9 | 0000 0011 111 | 26 |
| 0000 1011 | 10 | 0000 0011 110 | 27 |
| 0000 1010 | 11 | 0000 0011 101 | 28 |
| 0000 1001 | 12 | 0000 0011 100 | 29 |
| 0000 1000 | 13 | 0000 0011 011 | 30 |
| 0000 0111 | 14 | 0000 0011 010 | 31 |
| 0000 0110 | 15 | 0000 0011 001 | 32 |
| 0000 0101 11 | 16 | 0000 0011 000 | 33 |
| 0000 0101 10 | 17 | 0000 0001 000 | macroblock_escape |

NOTE – The "macroblock stuffing" entry that is available in ISO/IEC 11172-2 is not available in this Specification.

## B.2 Macroblock type

The properties of the macroblock are determined by the macroblock type VLC according to Tables B.2 to B.8.

**Table B.2 – Variable length codes for macroblock_type in I-pictures**

| macroblock_type VLC code | macroblock_quant | macroblock_motion_forward | macroblock_motion_backward | macroblock_pattern | macroblock_intra | spatial_temporal_weight_code_flag | permitted spatial_temporal_weight_classes | Description | |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 1 | 0 | | Intra | 0 |
| 01 | 1 | 0 | 0 | 0 | 1 | 0 | | Intra, Quant | 0 |

**Table B.3 – Variable length codes for macroblock_type in P-pictures**

| macroblock_type VLC code | macroblock_quant | macroblock_motion_forward | macroblock_motion_backward | macroblock_pattern | macroblock_intra | spatial_temporal_weight_code_flag | permitted spatial_temporal_weight_classes | Description | |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 1 | 0 | 1 | 0 | 0 | | MC, Coded | 0 |
| 01 | 0 | 0 | 0 | 1 | 0 | 0 | | No MC, Coded | 0 |
| 001 | 0 | 1 | 0 | 0 | 0 | 0 | | MC, Not Coded | 0 |
| 0001 1 | 0 | 0 | 0 | 0 | 1 | 0 | | Intra | 0 |
| 0001 0 | 1 | 1 | 0 | 1 | 0 | 0 | | MC, Coded, Quant | 0 |
| 0000 1 | 1 | 0 | 0 | 1 | 0 | 0 | | No MC, Coded, Quant | 0 |
| 0000 01 | 1 | 0 | 0 | 0 | 1 | 0 | | Intra, Quant | 0 |

**Table B.4 – Variable length codes for macroblock_type in B-pictures**

| macroblock_type VLC code | macroblock_quant | macroblock_motion_forward | macroblock_motion_backward | macroblock_pattern | macroblock_intra | spatial_temporal_weight_code_flag | permitted spatial_temporal_weight_classes Description | |
|---|---|---|---|---|---|---|---|---|
| 10 | 0 | 1 | 1 | 0 | 0 | 0 | Interp, Not Coded | 0 |
| 11 | 0 | 1 | 1 | 1 | 0 | 0 | Interp, Coded | 0 |
| 010 | 0 | 0 | 1 | 0 | 0 | 0 | Bwd, Not Coded | 0 |
| 011 | 0 | 0 | 1 | 1 | 0 | 0 | Bwd, Coded | 0 |
| 0010 | 0 | 1 | 0 | 0 | 0 | 0 | Fwd, Not Coded | 0 |
| 0011 | 0 | 1 | 0 | 1 | 0 | 0 | Fwd, Coded | 0 |
| 0001 1 | 0 | 0 | 0 | 0 | 1 | 0 | Intra | 0 |
| 0001 0 | 1 | 1 | 1 | 1 | 0 | 0 | Interp, Coded, Quant | 0 |
| 0000 11 | 1 | 1 | 0 | 1 | 0 | 0 | Fwd, Coded, Quant | 0 |
| 0000 10 | 1 | 0 | 1 | 1 | 0 | 0 | Bwd, Coded, Quant | 0 |
| 0000 01 | 1 | 0 | 0 | 0 | 1 | 0 | Intra, Quant | 0 |

**Table B.5 – Variable length codes for macroblock_type in I-pictures with spatial scalability**

| macroblock_type VLC code | macroblock_quant | macroblock_motion_forward | macroblock_motion_backward | macroblock_pattern | macroblock_intra | spatial_temporal_weight_code_flag | permitted spatial_temporal_weight_classes Description | |
|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 1 | 0 | 0 | Coded, Compatible | 4 |
| 01 | 1 | 0 | 0 | 1 | 0 | 0 | Coded, Compatible, Quant | 4 |
| 0011 | 0 | 0 | 0 | 0 | 1 | 0 | Intra | 0 |
| 0010 | 1 | 0 | 0 | 0 | 1 | 0 | Intra, Quant | 0 |
| 0001 | 0 | 0 | 0 | 0 | 0 | 0 | Not Coded, Compatible | 4 |

**Table B.6 – Variable length codes for macroblock_type in P-pictures with spatial scalability**

| macroblock_type VLC code | macroblock_quant | macroblock_motion_forward | macroblock_motion_backward | macroblock_pattern | macroblock_intra | spatial_temporal_weight_code_flag | permitted spatial_temporal_weight_classes | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | Description | |
| 10 | 0 | 1 | 0 | 1 | 0 | 0 | MC, Coded | 0 |
| 011 | 0 | 1 | 0 | 1 | 0 | 1 | MC, Coded, Compatible | 1,2,3 |
| 0000 100 | 0 | 0 | 0 | 1 | 0 | 0 | No MC, Coded | 0 |
| 0001 11 | 0 | 0 | 0 | 1 | 0 | 1 | No MC, Coded, Compatible | 1,2,3 |
| 0010 | 0 | 1 | 0 | 0 | 0 | 0 | MC, Not Coded | 0 |
| 0000 111 | 0 | 0 | 0 | 0 | 1 | 0 | Intra | 0 |
| 0011 | 0 | 1 | 0 | 0 | 0 | 1 | MC, Not coded, Compatible | 1,2,3 |
| 010 | 1 | 1 | 0 | 1 | 0 | 0 | MC, Coded, Quant | 0 |
| 0001 00 | 1 | 0 | 0 | 1 | 0 | 0 | No MC, Coded, Quant | 0 |
| 0000 110 | 1 | 0 | 0 | 0 | 1 | 0 | Intra, Quant | 0 |
| 11 | 1 | 1 | 0 | 1 | 0 | 1 | MC, Coded, Compatible, Quant | 1,2,3 |
| 0001 01 | 1 | 0 | 0 | 1 | 0 | 1 | No MC, Coded, Compatible,Quant | 1,2,3 |
| 0001 10 | 0 | 0 | 0 | 0 | 0 | 1 | No MC, Not Coded, Compatible | 1,2,3 |
| 0000 101 | 0 | 0 | 0 | 1 | 0 | 0 | Coded, Compatible | 4 |
| 0000 010 | 1 | 0 | 0 | 1 | 0 | 0 | Coded, Compatible, Quant | 4 |
| 0000 011 | 0 | 0 | 0 | 0 | 0 | 0 | Not Coded, Compatible | 4 |

**Table B.7 – Variable length codes for macroblock_type in B-pictures with spatial scalability**

| macroblock_type VLC code | macroblock_quant | macroblock_motion_forward | macroblock_motion_backward | macroblock_pattern | macroblock_intra | spatial_temporal_weight_code_flag | permitted spatial_temporal_weight_classes Description | |
|---|---|---|---|---|---|---|---|---|
| 10 | 0 | 1 | 1 | 0 | 0 | 0 | Interp, Not coded | 0 |
| 11 | 0 | 1 | 1 | 1 | 0 | 0 | Interp, Coded | 0 |
| 010 | 0 | 0 | 1 | 0 | 0 | 0 | Back, Not coded | 0 |
| 011 | 0 | 0 | 1 | 1 | 0 | 0 | Back, Coded | 0 |
| 0010 | 0 | 1 | 0 | 0 | 0 | 0 | For, Not coded | 0 |
| 0011 | 0 | 1 | 0 | 1 | 0 | 0 | For, Coded | 0 |
| 0001 10 | 0 | 0 | 1 | 0 | 0 | 1 | Back, Not Coded, Compatible | 1,2,3 |
| 0001 11 | 0 | 0 | 1 | 1 | 0 | 1 | Back, Coded, Compatible | 1,2,3 |
| 0001 00 | 0 | 1 | 0 | 0 | 0 | 1 | For, Not Coded, Compatible | 1,2,3 |
| 0001 01 | 0 | 1 | 0 | 1 | 0 | 1 | For, Coded, Compatible | 1,2,3 |
| 0000 110 | 0 | 0 | 0 | 0 | 1 | 0 | Intra | 0 |
| 0000 111 | 1 | 1 | 1 | 1 | 0 | 0 | Interp, Coded, Quant | 0 |
| 0000 100 | 1 | 1 | 0 | 1 | 0 | 0 | For, Coded, Quant | 0 |
| 0000 101 | 1 | 0 | 1 | 1 | 0 | 0 | Back, Coded, Quant | 0 |
| 0000 0100 | 1 | 0 | 0 | 0 | 1 | 0 | Intra, Quant | 0 |
| 0000 0101 | 1 | 1 | 0 | 1 | 0 | 1 | For, Coded, Compatible, Quant | 1,2,3 |
| 0000 0110 0 | 1 | 0 | 1 | 1 | 0 | 1 | Back, Coded, Compatible, Quant | 1,2,3 |
| 0000 0111 0 | 0 | 0 | 0 | 0 | 0 | 0 | Not Coded, Compatible | 4 |
| 0000 0110 1 | 1 | 0 | 0 | 1 | 0 | 0 | Coded, Compatible, Quant | 4 |
| 0000 0111 1 | 0 | 0 | 0 | 1 | 0 | 0 | Coded, Compatible | 4 |

**Table B.8 – Variable length codes for macroblock_type in I-pictures, P-pictures and B-pictures with SNR scalability**

| macroblock_type VLC code | macroblock_quant | macroblock_motion_forward | macroblock_motion_backward | macroblock_pattern | macroblock_intra | spatial_temporal_weight_code_flag | permitted spatial_temporal_weight_classes | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | Description | |
| 1 | 0 | 0 | 0 | 1 | 0 | 0 | Coded | 0 |
| 01 | 1 | 0 | 0 | 1 | 0 | 0 | Coded, Quant | 0 |
| 001 | 0 | 0 | 0 | 0 | 0 | 0 | Not Coded | 0 |

NOTE – There is no differentiation between picture types, since macroblocks are processed identically in I-, P- and B-pictures. The "Not coded" type is needed, since skipped macroblocks are not allowed at beginning and end of a slice.

## B.3 Macroblock pattern

See Table B.9.

**Table B.9 – Variable length codes for coded_block_pattern**

| coded_block_pattern VLC code | cbp | coded_block_pattern VLC code | cbp |
|---|---|---|---|
| 111 | 60 | 0001 1100 | 35 |
| 1101 | 4 | 0001 1011 | 13 |
| 1100 | 8 | 0001 1010 | 49 |
| 1011 | 16 | 0001 1001 | 21 |
| 1010 | 32 | 0001 1000 | 41 |
| 1001 1 | 12 | 0001 0111 | 14 |
| 1001 0 | 48 | 0001 0110 | 50 |
| 1000 1 | 20 | 0001 0101 | 22 |
| 1000 0 | 40 | 0001 0100 | 42 |
| 0111 1 | 28 | 0001 0011 | 15 |
| 0111 0 | 44 | 0001 0010 | 51 |
| 0110 1 | 52 | 0001 0001 | 23 |
| 0110 0 | 56 | 0001 0000 | 43 |
| 0101 1 | 1 | 0000 1111 | 25 |
| 0101 0 | 61 | 0000 1110 | 37 |
| 0100 1 | 2 | 0000 1101 | 26 |
| 0100 0 | 62 | 0000 1100 | 38 |
| 0011 11 | 24 | 0000 1011 | 29 |
| 0011 10 | 36 | 0000 1010 | 45 |
| 0011 01 | 3 | 0000 1001 | 53 |
| 0011 00 | 63 | 0000 1000 | 57 |
| 0010 111 | 5 | 0000 0111 | 30 |
| 0010 110 | 9 | 0000 0110 | 46 |
| 0010 101 | 17 | 0000 0101 | 54 |
| 0010 100 | 33 | 0000 0100 | 58 |
| 0010 011 | 6 | 0000 0011 1 | 31 |
| 0010 010 | 10 | 0000 0011 0 | 47 |
| 0010 001 | 18 | 0000 0010 1 | 55 |
| 0010 000 | 34 | 0000 0010 0 | 59 |
| 0001 1111 | 7 | 0000 0001 1 | 27 |
| 0001 1110 | 11 | 0000 0001 0 | 39 |
| 0001 1101 | 19 | 0000 0000 1 | 0 (Note) |
| NOTE – This entry shall not be used with 4:2:0 chrominance structure. | | | |

## B.4 Motion vectors

See Tables B.10 and B.11.

**Table B.10 – Variable length codes for motion_code**

| Variable length code | motion_code[r][s][t] |
|---|---|
| 0000 0011 001 | –16 |
| 0000 0011 011 | –15 |
| 0000 0011 101 | –14 |
| 0000 0011 111 | –13 |
| 0000 0100 001 | –12 |
| 0000 0100 011 | –11 |
| 0000 0100 11 | –10 |
| 0000 0101 01 | –9 |
| 0000 0101 11 | –8 |
| 0000 0111 | –7 |
| 0000 1001 | –6 |
| 0000 1011 | –5 |
| 0000 111 | –4 |
| 0001 1 | –3 |
| 0011 | –2 |
| 011 | –1 |
| 1 | 0 |
| 010 | 1 |
| 0010 | 2 |
| 0001 0 | 3 |
| 0000 110 | 4 |
| 0000 1010 | 5 |
| 0000 1000 | 6 |
| 0000 0110 | 7 |
| 0000 0101 10 | 8 |
| 0000 0101 00 | 9 |
| 0000 0100 10 | 10 |
| 0000 0100 010 | 11 |
| 0000 0100 000 | 12 |
| 0000 0011 110 | 13 |
| 0000 0011 100 | 14 |
| 0000 0011 010 | 15 |
| 0000 0011 000 | 16 |

**Table B.11 – Variable length codes for dmvector[t]**

| Code | Value |
|---|---|
| 11 | –1 |
| 0 | 0 |
| 10 | 1 |

## B.5 DCT coefficients

See Tables B.12 to B.16.

**Table B.12 – Variable length codes for dct_dc_size_luminance**

| Variable length code | dct_dc_size_luminance |
|---|---|
| 100 | 0 |
| 00 | 1 |
| 01 | 2 |
| 101 | 3 |
| 110 | 4 |
| 1110 | 5 |
| 1111 0 | 6 |
| 1111 10 | 7 |
| 1111 110 | 8 |
| 1111 1110 | 9 |
| 1111 1111 0 | 10 |
| 1111 1111 1 | 11 |

**Table B.13 – Variable length codes for dct_dc_size_chrominance**

| Variable length code | dct_dc_size_chrominance |
|---|---|
| 00 | 0 |
| 01 | 1 |
| 10 | 2 |
| 110 | 3 |
| 1110 | 4 |
| 1111 0 | 5 |
| 1111 10 | 6 |
| 1111 110 | 7 |
| 1111 1110 | 8 |
| 1111 1111 0 | 9 |
| 1111 1111 10 | 10 |
| 1111 1111 11 | 11 |

**Table B.14 – DCT coefficients Table zero**

| Variable length code (Note 1) | Run | Level |
|---|---|---|
| 10 (Note 2) | End of Block | |
| 1 s (Note 3) | 0 | 1 |
| 11 s (Note 4) | 0 | 1 |
| 011 s | 1 | 1 |
| 0100 s | 0 | 2 |
| 0101 s | 2 | 1 |
| 0010 1 s | 0 | 3 |
| 0011 1 s | 3 | 1 |
| 0011 0 s | 4 | 1 |
| 0001 10 s | 1 | 2 |
| 0001 11 s | 5 | 1 |
| 0001 01 s | 6 | 1 |
| 0001 00 s | 7 | 1 |
| 0000 110 s | 0 | 4 |
| 0000 100 s | 2 | 2 |
| 0000 111 s | 8 | 1 |
| 0000 101 s | 9 | 1 |
| 0000 01 | Escape | |
| 0010 0110 s | 0 | 5 |
| 0010 0001 s | 0 | 6 |
| 0010 0101 s | 1 | 3 |
| 0010 0100 s | 3 | 2 |
| 0010 0111 s | 10 | 1 |
| 0010 0011 s | 11 | 1 |
| 0010 0010 s | 12 | 1 |
| 0010 0000 s | 13 | 1 |
| 0000 0010 10 s | 0 | 7 |
| 0000 0011 00 s | 1 | 4 |
| 0000 0010 11 s | 2 | 3 |
| 0000 0011 11 s | 4 | 2 |
| 0000 0010 01 s | 5 | 2 |
| 0000 0011 10 s | 14 | 1 |
| 0000 0011 01 s | 15 | 1 |
| 0000 0010 00 s | 16 | 1 |

**Table B.14 – DCT coefficients Table zero** *(continued)*

| Variable length code (Note 1) | Run | Level |
|---|---|---|
| 0000 0001 1101 s | 0 | 8 |
| 0000 0001 1000 s | 0 | 9 |
| 0000 0001 0011 s | 0 | 10 |
| 0000 0001 0000 s | 0 | 11 |
| 0000 0001 1011 s | 1 | 5 |
| 0000 0001 0100 s | 2 | 4 |
| 0000 0001 1100 s | 3 | 3 |
| 0000 0001 0010 s | 4 | 3 |
| 0000 0001 1110 s | 6 | 2 |
| 0000 0001 0101 s | 7 | 2 |
| 0000 0001 0001 s | 8 | 2 |
| 0000 0001 1111 s | 17 | 1 |
| 0000 0001 1010 s | 18 | 1 |
| 0000 0001 1001 s | 19 | 1 |
| 0000 0001 0111 s | 20 | 1 |
| 0000 0001 0110 s | 21 | 1 |
| 0000 0000 1101 0 s | 0 | 12 |
| 0000 0000 1100 1 s | 0 | 13 |
| 0000 0000 1100 0 s | 0 | 14 |
| 0000 0000 1011 1 s | 0 | 15 |
| 0000 0000 1011 0 s | 1 | 6 |
| 0000 0000 1010 1 s | 1 | 7 |
| 0000 0000 1010 0 s | 2 | 5 |
| 0000 0000 1001 1 s | 3 | 4 |
| 0000 0000 1001 0 s | 5 | 3 |
| 0000 0000 1000 1 s | 9 | 2 |
| 0000 0000 1000 0 s | 10 | 2 |
| 0000 0000 1111 1 s | 22 | 1 |
| 0000 0000 1111 0 s | 23 | 1 |
| 0000 0000 1110 1 s | 24 | 1 |
| 0000 0000 1110 0 s | 25 | 1 |
| 0000 0000 1101 1 s | 26 | 1 |

**Table B.14 – DCT coefficients Table zero** *(continued)*

| Variable length code (Note 1) | Run | Level |
|---|---|---|
| 0000 0000 0111 11 s | 0 | 16 |
| 0000 0000 0111 10 s | 0 | 17 |
| 0000 0000 0111 01 s | 0 | 18 |
| 0000 0000 0111 00 s | 0 | 19 |
| 0000 0000 0110 11 s | 0 | 20 |
| 0000 0000 0110 10 s | 0 | 21 |
| 0000 0000 0110 01 s | 0 | 22 |
| 0000 0000 0110 00 s | 0 | 23 |
| 0000 0000 0101 11 s | 0 | 24 |
| 0000 0000 0101 10 s | 0 | 25 |
| 0000 0000 0101 01 s | 0 | 26 |
| 0000 0000 0101 00 s | 0 | 27 |
| 0000 0000 0100 11 s | 0 | 28 |
| 0000 0000 0100 10 s | 0 | 29 |
| 0000 0000 0100 01 s | 0 | 30 |
| 0000 0000 0100 00 s | 0 | 31 |
| 0000 0000 0011 000 s | 0 | 32 |
| 0000 0000 0010 111 s | 0 | 33 |
| 0000 0000 0010 110 s | 0 | 34 |
| 0000 0000 0010 101 s | 0 | 35 |
| 0000 0000 0010 100 s | 0 | 36 |
| 0000 0000 0010 011 s | 0 | 37 |
| 0000 0000 0010 010 s | 0 | 38 |
| 0000 0000 0010 001 s | 0 | 39 |
| 0000 0000 0010 000 s | 0 | 40 |
| 0000 0000 0011 111 s | 1 | 8 |
| 0000 0000 0011 110 s | 1 | 9 |
| 0000 0000 0011 101 s | 1 | 10 |
| 0000 0000 0011 100 s | 1 | 11 |
| 0000 0000 0011 011 s | 1 | 12 |
| 0000 0000 0011 010 s | 1 | 13 |
| 0000 0000 0011 001 s | 1 | 14 |

**Table B.14 – DCT coefficients Table zero** *(concluded)*

| Variable length code (Note 1) | Run | Level |
|---|---|---|
| 0000 0000 0001 0011 s | 1 | 15 |
| 0000 0000 0001 0010 s | 1 | 16 |
| 0000 0000 0001 0001 s | 1 | 17 |
| 0000 0000 0001 0000 s | 1 | 18 |
| 0000 0000 0001 0100 s | 6 | 3 |
| 0000 0000 0001 1010 s | 11 | 2 |
| 0000 0000 0001 1001 s | 12 | 2 |
| 0000 0000 0001 1000 s | 13 | 2 |
| 0000 0000 0001 0111 s | 14 | 2 |
| 0000 0000 0001 0110 s | 15 | 2 |
| 0000 0000 0001 0101 s | 16 | 2 |
| 0000 0000 0001 1111 s | 27 | 1 |
| 0000 0000 0001 1110 s | 28 | 1 |
| 0000 0000 0001 1101 s | 29 | 1 |
| 0000 0000 0001 1100 s | 30 | 1 |
| 0000 0000 0001 1011 s | 31 | 1 |

NOTE 1 – The last bit 's' denotes the sign of the level: '0' for positive, '1' for negative.

NOTE 2 – "End of Block" shall not be the only code of the block.

NOTE 3 – This code shall be used for the first (DC) coefficient in the block.

NOTE 4 – This code shall be used for all other coefficients.

**Table B.15 – DCT coefficients Table one**

| Variable length code (Note 1) | Run | Level |
|---|---|---|
| 0110 (Note 2) | End of Block | |
| 10s | 0 | 1 |
| 010 s | 1 | 1 |
| 110 s | 0 | 2 |
| 0010 1 s | 2 | 1 |
| 0111 s | 0 | 3 |
| 0011 1 s | 3 | 1 |
| 0001 10 s | 4 | 1 |
| 0011 0 s | 1 | 2 |
| 0001 11 s | 5 | 1 |
| 0000 110 s | 6 | 1 |
| 0000 100 s | 7 | 1 |
| 1110 0 s | 0 | 4 |
| 0000 111 s | 2 | 2 |
| 0000 101 s | 8 | 1 |
| 1111 000 s | 9 | 1 |
| 0000 01 | Escape | |
| 1110 1 s | 0 | 5 |
| 0001 01 s | 0 | 6 |
| 1111 001 s | 1 | 3 |
| 0010 0110 s | 3 | 2 |
| 1111 010 s | 10 | 1 |
| 0010 0001 s | 11 | 1 |
| 0010 0101 s | 12 | 1 |
| 0010 0100 s | 13 | 1 |
| 0001 00 s | 0 | 7 |
| 0010 0111 s | 1 | 4 |
| 1111 1100 s | 2 | 3 |
| 1111 1101 s | 4 | 2 |
| 0000 0010 0 s | 5 | 2 |
| 0000 0010 1 s | 14 | 1 |
| 0000 0011 1 s | 15 | 1 |
| 0000 0011 01 s | 16 | 1 |

**Table B.15 – DCT coefficients Table one** *(continued)*

| Variable length code (Note 1) | Run | Level |
|---|---|---|
| 1111 011 s | 0 | 8 |
| 1111 100 s | 0 | 9 |
| 0010 0011 s | 0 | 10 |
| 0010 0010 s | 0 | 11 |
| 0010 0000 s | 1 | 5 |
| 0000 0011 00 s | 2 | 4 |
| 0000 0001 1100 s | 3 | 3 |
| 0000 0001 0010 s | 4 | 3 |
| 0000 0001 1110 s | 6 | 2 |
| 0000 0001 0101 s | 7 | 2 |
| 0000 0001 0001 s | 8 | 2 |
| 0000 0001 1111 s | 17 | 1 |
| 0000 0001 1010 s | 18 | 1 |
| 0000 0001 1001 s | 19 | 1 |
| 0000 0001 0111 s | 20 | 1 |
| 0000 0001 0110 s | 21 | 1 |
| 1111 1010 s | 0 | 12 |
| 1111 1011 s | 0 | 13 |
| 1111 1110 s | 0 | 14 |
| 1111 1111 s | 0 | 15 |
| 0000 0000 1011 0 s | 1 | 6 |
| 0000 0000 1010 1 s | 1 | 7 |
| 0000 0000 1010 0 s | 2 | 5 |
| 0000 0000 1001 1 s | 3 | 4 |
| 0000 0000 1001 0 s | 5 | 3 |
| 0000 0000 1000 1 s | 9 | 2 |
| 0000 0000 1000 0 s | 10 | 2 |
| 0000 0000 1111 1 s | 22 | 1 |
| 0000 0000 1111 0 s | 23 | 1 |
| 0000 0000 1110 1 s | 24 | 1 |
| 0000 0000 1110 0 s | 25 | 1 |
| 0000 0000 1101 1 s | 26 | 1 |

**Table B.15 – DCT coefficients Table one** *(continued)*

| Variable length code (Note 1) | Run | Level |
|---|---|---|
| 0000 0000 0111 11 s | 0 | 16 |
| 0000 0000 0111 10 s | 0 | 17 |
| 0000 0000 0111 01 s | 0 | 18 |
| 0000 0000 0111 00 s | 0 | 19 |
| 0000 0000 0110 11 s | 0 | 20 |
| 0000 0000 0110 10 s | 0 | 21 |
| 0000 0000 0110 01 s | 0 | 22 |
| 0000 0000 0110 00 s | 0 | 23 |
| 0000 0000 0101 11 s | 0 | 24 |
| 0000 0000 0101 10 s | 0 | 25 |
| 0000 0000 0101 01 s | 0 | 26 |
| 0000 0000 0101 00 s | 0 | 27 |
| 0000 0000 0100 11 s | 0 | 28 |
| 0000 0000 0100 10 s | 0 | 29 |
| 0000 0000 0100 01 s | 0 | 30 |
| 0000 0000 0100 00 s | 0 | 31 |
| 0000 0000 0011 000 s | 0 | 32 |
| 0000 0000 0010 111 s | 0 | 33 |
| 0000 0000 0010 110 s | 0 | 34 |
| 0000 0000 0010 101 s | 0 | 35 |
| 0000 0000 0010 100 s | 0 | 36 |
| 0000 0000 0010 011 s | 0 | 37 |
| 0000 0000 0010 010 s | 0 | 38 |
| 0000 0000 0010 001 s | 0 | 39 |
| 0000 0000 0010 000 s | 0 | 40 |
| 0000 0000 0011 111 s | 1 | 8 |
| 0000 0000 0011 110 s | 1 | 9 |
| 0000 0000 0011 101 s | 1 | 10 |
| 0000 0000 0011 100 s | 1 | 11 |
| 0000 0000 0011 011 s | 1 | 12 |
| 0000 0000 0011 010 s | 1 | 13 |
| 0000 0000 0011 001 s | 1 | 14 |

**Table B.15 – DCT coefficients Table one** *(concluded)*

| Variable length code (Note 1) | Run | Level |
|---|---|---|
| 0000 0000 0001 0011 s | 1 | 15 |
| 0000 0000 0001 0010 s | 1 | 16 |
| 0000 0000 0001 0001 s | 1 | 17 |
| 0000 0000 0001 0000 s | 1 | 18 |
| 0000 0000 0001 0100 s | 6 | 3 |
| 0000 0000 0001 1010 s | 11 | 2 |
| 0000 0000 0001 1001 s | 12 | 2 |
| 0000 0000 0001 1000 s | 13 | 2 |
| 0000 0000 0001 0111 s | 14 | 2 |
| 0000 0000 0001 0110 s | 15 | 2 |
| 0000 0000 0001 0101 s | 16 | 2 |
| 0000 0000 0001 1111 s | 27 | 1 |
| 0000 0000 0001 1110 s | 28 | 1 |
| 0000 0000 0001 1101 s | 29 | 1 |
| 0000 0000 0001 1100 s | 30 | 1 |
| 0000 0000 0001 1011 s | 31 | 1 |
| NOTE 1 –The last bit 's' denotes the sign of the level: '0' for positive, '1' for negative.<br>NOTE 2 – "End of Block" shall not be the only code of the block. | | |

**Table B.16 – Encoding of run and level following an ESCAPE code**

| Fixed length code | Run |
|---|---|
| 0000 00 | 0 |
| 0000 01 | 1 |
| 0000 10 | 2 |
| … | … |
| | |
| … | … |
| … | … |
| … | … |
| 1111 11 | 63 |

| Fixed length code | signed_level |
|---|---|
| 1000 0000 0000 | reserved |
| 1000 0000 0001 | –2047 |
| 1000 0000 0010 | –2046 |
| … | … |
| 1111 1111 1111 | –1 |
| 0000 0000 0000 | Forbidden |
| 0000 0000 0001 | +1 |
| … | … |
| 0111 1111 1111 | +2047 |

# Annex C

## Video buffering verifier

(This annex forms an integral part of this Recommendation | International Standard)

Coded video bitstreams shall meet constraints imposed through a Video Buffering Verifier (VBV) defined in this annex. Each bitstream in a scalable hierarchy shall not violate the VBV constraints defined in this annex.

The VBV is a hypothetical decoder, which is conceptually connected to the output of an encoder. It has an input buffer known as the VBV buffer. Coded data is placed in the buffer as defined below in C.3 and is removed from the buffer as defined in C.5, C.6, and C.7. It is required that a bitstream that conforms to this Specification shall not cause the VBV buffer to overflow. When low_delay equals zero, the bitstream shall not cause the VBV buffer to underflow. When low_delay equals one, decoding a picture at the normally expected time might cause the VBV buffer to underflow. If this is the case, the picture is not decoded and the VBV buffer is re-examined at a sequence of later times specified in C.7 and C.8 until it is all present in the VBV buffer.

All the arithmetic in this annex is done with real-values, so that no rounding errors can propagate. For example, the number of bits in the VBV buffer is not necessarily an integer.

**C.1** The VBV and the video encoder have the same clock frequency as well as the same frame rate, and are operated synchronously.

**C.2** The VBV buffer is of size B, where B is the vbv_buffer_size coded in the sequence header and sequence extension if present.

**C.3** This subclause defines the input of data to the VBV buffer. Two mutually exclusive cases are defined in C.3.1 and C.3.2. In both cases the VBV buffer is initially empty. Let $R_{max}$ be the bitrate specified in the bit_rate field.

**C.3.1** In the case where vbv_delay is coded with a value not equal to hexadecimal FFFF, the picture data of the n-th coded picture enters the buffer at a rate R(n) where:

$$R(n) = d^*_n / (\tau(n) - \tau(n+1) + t(n+1) - t(n))$$

Where

- R(n) is the rate, in bits/s, that the picture data for the n-th coded picture enters the VBV.
- $d^*_n$ is the number of bits after the final bit of the n-th picture start code and before and including the final bit of the (n + 1)-th picture start code.
- $\tau$ (n) is the decoding delay coded in vbv_delay for the n-th coded picture, measured in seconds.
- t(n) Is the time, measured in seconds, when the n-th coded picture is removed from VBV buffer. t(n) is defined in C.9, C.10, C.11, and C.12.

Ambiguity at the beginning of a sequence:

The interval of time $t_{n+1} - t_n$ between removal of two consecutive pictures can normally be derived from the bitstream as described in C.9, C.10, C.11 and C.12.

When random access is made in a sequence, $t_{n+1} - t_n$ cannot be determined from the video bitstream alone for the first picture(s) after the sequence header since the previous coded P- or I-frame does not exist in the decoded sequence. If the bitstream is multiplexed as part of a systems bitstream according to ITU-T Rec. H.222.0 | ISO/IEC 13818-1, then it is possible (but not certain) that information in the systems bitstream may be used to determine unambiguously this interval of time. This information is available if Decoding Time Stamps (DTS) are transmitted for pictures n and n+1.

If the rate R(n) cannot be determined unambiguously, it is not possible for the VBV to precisely determine the fullness in trajectories in the VBV buffer during a limited period (always less than the maximum value for vbv_delay, which is approximately 0.73 seconds), therefore strict VBV verification of the entire bitstream is not always possible. Note that an encoder always knows the values of $t_{n+1} - t_n$ after each repeated sequence headers and therefore knows how to generate a bitstream that does not violate the VBV constraints at those points.

The ambiguity may become a problem when the video bitstream is remultiplexed and delivered at a rate different from the intended piecewise constant rate R(n).

It should also be noted that the input rate for the bits preceding the first picture header cannot be determined from the bitstream.

Ambiguity at the end of a sequence:

The input of all the bits following the picture start code of the picture preceding an end of sequence code cannot be determined from the bitstream. There shall exist an input rate for these bits that does not lead to an overflow or, if low_delay is equal to 1, an underflow of the VBV buffer. This rate shall be less than the maximum rate specified in the sequence header.

After filling the VBV buffer with all the data that precedes the first picture start code of the sequence and the picture start code itself, the VBV buffer is filled from the bitstream for the time specified by the vbv_delay field in the picture header. At this time decoding begins. The data input continues at the rates specified in this subclause.

For all bitstreams R(n) <= $R_{max}$ for all picture data.

NOTE – For constant rate video the sequence of values R(n) are constant throughout the sequence to within the accuracy permitted by the quantisation of vbv_delay.

**C.3.2** In the case where vbv_delay is coded with the value hexadecimal FFFF, data enters the VBV buffer as specified in this subclause.

If the VBV buffer is not full, data enters the buffer at $R_{max}$.

If the VBV buffer becomes full after filling at $R_{max}$ for some time, no more data enters the buffer until some data is removed from the buffer.

After filling the VBV buffer with all the data that precedes the first picture start code of the sequence and the picture start code itself, the VBV buffer is filled from the bitstream until it is full. At this time decoding begins. The data input continues at the rate specified in this subclause.

**C.4** Starting at the time defined in C.3, the VBV buffer is examined at successive times defined in C.9 to C.12. C.5 to C.8 defines the actions to be taken at each time the VBV buffer is examined.

**C.5** This subclause defines a requirement on all video bitstreams.

At the time the VBV buffer is examined *before* removing any picture data, the number of bits in the buffer shall lie between zero bits and B bits where B is the size of the VBV buffer indicated by vbv_buffer_size.

For the purpose of this annex, picture data is defined as all the bits of the coded picture, all the header(s) and user data immediately preceding it if any (including any stuffing between them) and all the stuffing following it, up to (but not including) the next start code, except in the case where the next start code is an end of sequence code, in which case it is included in the picture data.

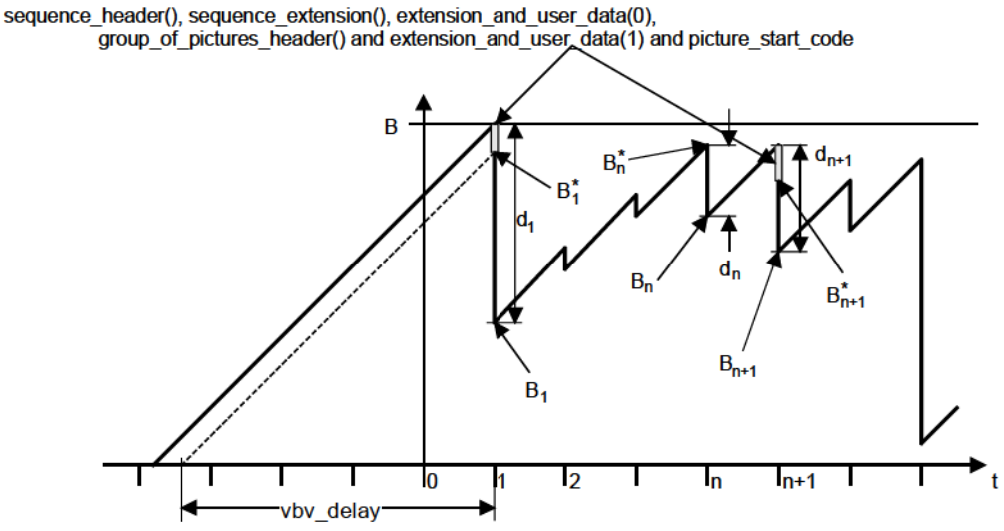



**Figure C.1 – VBV Buffer Occupancy – Constant bit-rate operation**

**C.6** This subclause defines a requirement on the video bitstreams when the low_delay flag is equal to zero.

At each time the VBV buffer is examined and before any bits are removed, all of the data for the picture which (at that time) has been in the buffer longest shall be present in the VBV buffer. This picture data shall be removed instantaneously at this time.

VBV buffer underflow shall not occur when the low_delay flag is equal to 0. This requires that all picture data for the n-th picture shall be present in the VBV buffer at the decoding time, $t_n$.

**C.7** This subclause only applies when the low_delay flag is equal to one.

When low_delay is equal to one, there may be situations where the VBV buffer shall be re-examined several times before removing a coded picture from the VBV buffer. It is possible to know if the VBV buffer has to be re-examined and how many times by looking at the temporal_reference of the next picture (the one that follows the picture currently to be decoded), see 6.3.10. If the VBV buffer has to be re-examined, the picture currently to be decoded is referred to as a big picture.

If picture currently to be decoded is a big picture, the VBV buffer is re-examined at intervals of 2 field-periods before removing the big picture, and no picture data is removed until the final re-examination.

At this time, the number of bits the VBV buffer immediately before removing the big picture shall be less than B, all the picture data for the picture that has been in the buffer longest (the big picture) shall be present in the buffer and shall be removed instantaneously. Then, normal operation of the VBV resumes, and C.5 applies.

The last coded picture of a sequence shall not be a big picture.

**C.8** This subclause is informative only.

The situation where the VBV buffer would underflow (see C.7) can happen when low-delay applications transmit occasionally large pictures, for example in case of scene-cuts.

Decoding such bitstreams will cause the display process associated with a decoder to repeat a previously decoded field or frame until normal operation of the VBV can resume. This process is sometimes referred to as the occurrence of "skipped pictures". Note that this situation should normally not occur except occasionally. It shall not occur when low_delay is equal to 0.

**C.9** This subclause defines the time intervals between successive examination of the VBV buffer in the case where progressive_sequence equals to 1 and low_delay equals to 0. In this case, the frame re-ordering delay always exists and B-pictures can occur.

The time interval $t_{n+1} - t_n$ between two successive examinations of the VBV buffer is a multiple of T, where T is the inverse of the frame rate.

If the n-th picture is a B-picture with repeat_first_field equals to 0, then $t_{n+1} - t_n$ is equal to T.

If the n-th picture is a B-Picture with repeat_first_field equal to 1 and top_field_first equals 0, then $t_{n+1} - t_n$ is equal to 2*T.

If the n-th picture is a B-Picture with repeat_first_field equal to 1 and top_field_first equals 1, then $t_{n+1} - t_n$ is equal to 3*T.

If the n-th picture is a P-Picture or I-Picture and if the previous P-Picture or I-Picture has repeat_first_field equal to 0, then $t_{n+1} - t_n$ is equal to T.

If the n-th picture is a P-Picture or I-Picture and if the previous P-Picture or I-Picture has repeat_first_field equal to 1 and top_field_first equal to 0, then $t_{n+1} - t_n$ is equal to 2*T.

If the n-th picture is a P-Picture or I-Picture and if the previous P-Picture or I-Picture has repeat_first_field equal to 1 and top_field_first equal to 1, then $t_{n+1} - t_n$ is equal to 3*T.

**C.10** This subclause defines the time intervals between successive examination of the VBV buffer in the case where progressive_sequence equals to 1 and low_delay equals to 1. In this case the sequence contains no B-Pictures and there is no frame re-ordering delay.

The time interval $t_{n+1} - t_n$ between two successive examinations of the VBV buffer is a multiple of T, where T is the inverse of the frame rate.

If the n-th picture is a P-Picture or I-Picture with repeat_first_field equal to 0, then $t_{n+1} - t_n$ is equal to T.

If the n-th picture is a P-Picture or I-Picture with repeat_first_field equal to 1 and top_field_first equal to 0, then $t_{n+1} - t_n$ is equal to 2*T.

If the n-th picture is a P-Picture or I-Picture with repeat_first_field equal to 1 and top_field_first equal to 1, then $t_{n+1} - t_n$ is equal to 3*T.

**C.11** This subclause defines the time intervals between successive examination of the VBV buffer in the case where progressive_sequence equals to 0 and low_delay equals to 0. In this case, the frame re-ordering delay always exists and B-pictures can occur.

The time interval $t_{n+1} - t_n$ between two successive examinations of the VBV input buffer is a multiple of T, where T is the inverse of two times the frame rate.

If the n-th picture is a *frame-structure* coded B-frame with repeat_first_field equal to 0, then $t_{n+1} - t_n$ is equal to 2*T.

If the n-th picture is a *frame-structure* coded B-frame with repeat_first_field equal to 1, then $t_{n+1} - t_n$ is equal to 3*T.

If the n-th picture is a *field-structure* B-picture (B-field picture), then $t_{n+1} - t_n$ is equal to T.

If the n-th picture is a *frame-structure* coded P-frame or coded I-Frame and if the previous coded P-Frame or coded I-Frame has repeat_first_field equal to 0, then $t_{n+1} - t_n$ is equal to 2*T.

If the n-th picture is a *frame-structure* coded P-Frame or coded I-Frame and if the previous coded P-Frame or coded I-Frame has repeat_first_field equal to 1, then $t_{n+1} - t_n$ is equal to 3*T.

If the n-th picture is the *first* field of a *field-structure* coded P-frame or coded I-Frame, then $t_{n+1} - t_n$ is equal to T.

If the n-th picture is the *second* field of a *field-structure* coded P-Frame or coded I-Frame and if the previous coded P-Frame or coded I-Frame is using field-structure or has *repeat_first_field* equal to 0, then $t_{n+1} - t_n$ is equal to (2*T – T).

If the n-th picture is the *second* field of a *field-structure* coded P-Frame or coded I-Frame and if the previous coded P-Frame or coded I-Frame is using frame-structure and has *repeat_first_field* equal to 1, then $t_{n+1} - t_n$ is equal to (3*T – T).

Figure C.2 shows the VBV in a simple case with only frame-pictures. Frames $P_0$, $B_2$ and $B_4$ have a display duration of 3 fields.

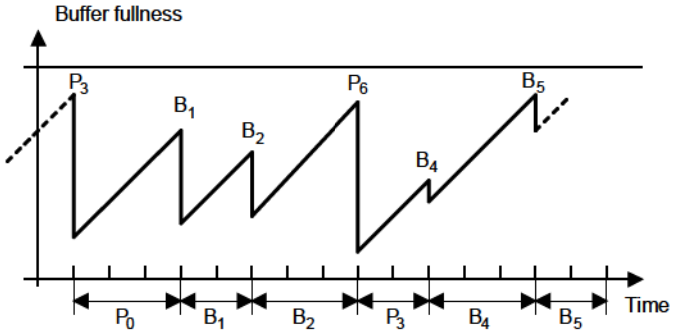



**Figure C.2**

**C.12** This clause defines the time intervals between successive examination of the VBV buffer in the case where progressive_sequence equals to 0 and low_delay equals to 1. In this case the sequence contains no B-Pictures and there is no frame re-ordering delay.

The time interval $t_{n+1} - t_n$ between two successive examinations of the VBV input buffer is a multiple of T, where T is the inverse of two times the frame rate.

If the n-th picture is a *frame-structure* coded P-Frame or coded I-Frame with repeat_first_field equal to 0, then $t_{n+1} - t_n$ is equal to 2*T.

If the n-th picture is a *frame-structure* coded P-Frame or coded I-Frame with repeat_first_field equal to 1, then $t_{n+1} - t_n$ is equal to 3*T.

If the n-th picture is a *field-structure* coded P-Frame or coded I-Frame, then $t_{n+1} - t_n$ is equal to T.

Figure C.3 shows the VBV in a simple case with only frame-pictures. Frames $I_0$, $P_2$ and $P_4$ have repeat_first_field equal to 1.

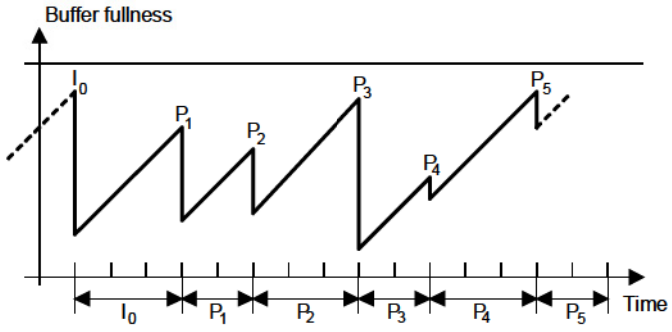



**Figure C.3**

# Annex D

## Features supported by the algorithm

(This annex does not form an integral part of this Recommendation | International Standard)

## D.1 Overview

The following non-exhaustive list of features is included in this Specification:

1) Different chrominance sampling formats (i.e. 4:2:0, 4:2:2 and 4:4:4) can be represented.
2) Video in both the progressive and interlaced scan formats can be encoded.
3) The decoder can use 3:2 pull down to represent a ~24 fps film as ~30 fps video.
4) The displayed video can be selected by a movable pan-scan window within a larger raster.
5) A wide range of picture qualities can be used.
6) Both constant and variable bitrate channels are supported.
7) A low delay mode for face-to-face applications is available.
8) Random access (for DSM, channel acquisition, and channel hopping) is available.
9) ISO/IEC 11172-2 constrained parameter bitstreams are decodable.
10) Bitstreams for high and low (hardware) complexity decoders can be generated.
11) Editing of encoded video is supported.
12) Fast-forward and fast-reverse playback recorded bitstreams can be implemented.
13) The encoded bitstream is resilient to errors.

## D.2 Video formats

### D.2.1 Sampling formats and colour

This Specification video coding supports both interlaced and progressive video. The respective indication is provided with a progressive_sequence flag transmitted in the Sequence Extension code.

Allowed raster sizes are between 1 and (2^14 – 1) luminance samples each of the horizontal and vertical directions. The video is represented in a luminance/chrominance colour space with selectable colour primaries. The chrominance can be sampled in either the 4:2:0 (half as many samples in the horizontal and vertical directions), 4:2:2 (half as many samples in the horizontal direction only). Furthermore, application specific sample aspect ratios and image aspect ratios are flexibly supported. A chroma_format parameter is contained in the Sequence Extension code.

Sample aspect ratio information is provided by means of aspect_ratio_information and (optional) display_horizontal_size and display_vertical_size in the sequence_display_extension(). Examples of appropriate values for signals sampled in accordance with Recommendation ITU-R BT. 601 are given in Table D.1.

This Specification implements tools to support 4:4:4 chrominance, for possible future use. However, this is currently not supported in any profile.

**Table D.1 – Example display size values**

| Signal Format | display_horizontal_size | display_vertical_size |
|---|---|---|
| 525-line | 711 | 483 |
| 625-line | 702 | 575 |

### D.2.2 Movie timing

A decoder can implement 3:2 pull down when a sequence of progressive pictures is encoded. Each encoded movie picture can independently specify whether it is displayed for two or three video field periods, so "irregular" 3:2 pull down source material can be transmitted as progressive video. Two flags, top_field_first and repeat_first_field, are transmitted with the Picture Coding Extensions and adequately describe the necessary display timing.

### D.2.3 Display format control

The display process converts a sequence of digital frames (in the case of progressive video) or fields (in the case of interlaced video) to output video. It is not a normative part of this Recommendation | International Standard. The video syntax of this Specification does communicate certain display parameters for use in reconstructing the video. Optional information (in the sequence display extension) specifies the chromaticities, the display primaries, the opto-electronic transfer characteristics (e.g. the value of gamma) and the RGB-to-luminance/chrominance conversion matrix.

Moreover, a display window within the encoded raster may be defined as, e.g. in the case of pan and scan. Alternatively the encoded raster may be defined as a window on a large area display device. In the case of pan-scan the position of the window representing the displayed region of a larger picture can be specified on a field-by-field basis. It is specified in the Picture display extension described in 6.3.12. A typical use for the pan-scan window is to describe the "important" 4:3 aspect ratio rectangle within a 16:9 video sequence. Similarly, in the case of small encoded pictures on a large display, the size of the display and the position of the window within that display may be specified.

### D.2.4 Transparent coding of composite video

Decoding from PAL/NTSC before transmission and re-coding to PAL/NTSC after transmission of composite source signals in non-low quality applications, such as contribution and distribution, requires a precise reconstruction of the carrier amplitude and phase reference signal (and v-axis switch for PAL).

The input format can be indicated in the sequence header using the video_format bits. Possible source formats are: PAL, NTSC, SECAM and MAC. Reconstruction of the carrier signal is possible by using the carrier parameters: v_axis, field_sequence, sub_carrier, burst_amplitude and sub_carrier_phase that are enabled by setting the composite_display_flag in the picture_coding_extension().

## D.3 Picture quality

High picture quality is provided according to the bitrate used. Provision for very high picture quality is made by sufficiently high bitrate limits relating to a certain level in a particular profile. High chrominance band quality can be achieved by using 4:2:2 chrominance

Quantiser matrices can be downloaded and used with a small a quantiser_scale_code to achieve near lossless coding.

Moreover, scalable coding with flexible bitrate allows for service or quality hierarchy and graceful degradation. E.g. decoding a subset of the bitstream carrying a lower resolution picture allows for decoding this signal in a low-cost receiver with related quality; decoding the complete bitstream allows to obtain the high overall quality.

Furthermore, operation at low bitrates can be accommodated by using low frame rates (by either pre-processing before coding or frame skipping indicated by the temporal_reference in the picture header) and low spatial resolution.

## D.4 Data rate control

The number of transmitted bits per unit time, which is selectable in a wide range, may be controlled in two ways, which are both supported by this Specification. A bit_rate description is transmitted with the Sequence Header Code.

For Constant Bitrate (CBR) coding, the number of transmitted bits per unit time is constant on the channel. Since the encoder output rate generally varies depending on the picture content, it shall regulate the rate constant by buffering, etc. In CBR, picture quality may vary depending on its content.

The other mode is the Variable Bitrate (VBR) coding, in which case the number of transmitted bits per unit time may vary on the channel under some constriction. VBR is meant to provide constant quality coding. A model for VBR application is near-constant-quality coding over B-ISDN channels subject to Usage Parameter Control (UPC).

## D.5 Low delay mode

A low encoding and decoding delay mode is accommodated for real-time video communications such as visual telephony, video-conferencing, monitoring. Total encoding and decoding delay of less than 150 milliseconds can be achieved for low delay mode operation of this Specification. Setting the low_delay flag in the Sequence Header code defines a low delay bitstream.

The total encoding and decoding delay can be kept low by generating a bitstream which does not contain B-pictures. This prevents frame re-ordering delay. By using dual-prime prediction for coded P-frames, the picture quality can still be high.

A low buffer occupancy for both encoder and decoder is needed for low delay. Large coded pictures should be avoided by the encoder. By using intra update on the basis of one or more slices per frame (intra slices) instead of intra frames this can be accommodated.

In case of exceeding, for low delay operation, the desired number of bits per frame the encoder can skip one or more frames. This action is indicated by a discontinuity in the value of temporal_reference for the next picture (see the semantic definition in 6.3.9) and may cause C.7 of the VBV to apply, i.e. the decoder buffer would underflow if some frames are not repeated by the decoder.

## D.6 Random access/channel hopping

The syntax of this Specification supports random access and channel hopping. Sufficient random access/channel hopping functionality is possible by encoding suitable random access points into the bitstream without significant loss of image quality.

Random access is an essential feature for video on a storage medium. It requires that any picture can be accessed and decoded in a limited amount of time. It implies the existence of access points in the bitstream – that is segments of information that are identifiable and can be decoded without reference to other segments of data. In this Specification access points are provided by sequence_header() and this is then followed by intra information (picture data that can be decoded without access to previously decoded pictures). A spacing of two random access points per second can be achieved without significant loss of picture quality.

Channel hopping is the similar situation in transmission applications such as broadcasting. As soon as a new channel has been selected and the bitstream of the selected channel is available to the decoder, the next data entry, i.e. random access point has to be found to start decoding the new program in the manner outlined in the previous paragraph.

## D.7 Scalability

The syntax of this Specification supports bitstream scalability. To accommodate the diverse functionality requirements of the applications envisaged by this Specification, a number of bitstream scalability tools have been developed:

- **SNR scalability** mainly targets for applications which require graceful degradation.
- **Chroma simulcast** targets at applications with high chrominance quality requirements.
- **Data partitioning** is primarily targeted for cell loss resilience in ATM networks.
- **Temporal scalability** is a method suitable for interworking of services using high temporal resolution progressive video formats. Also suitable for high quality graceful degradation in the presence of channel errors.
- **Spatial scalability** allows multiresolution coding technique suitable for video service interworking applications. This tool can also provide coding modes to achieve compatibility with existing coding standards, i.e. ISO/IEC 11172-2, at the lower layer.

### D.7.1 Use of SNR scalability at a single spatial resolution

The aim of SNR scalability is primarily to provide a mechanism for transmission of a two layer service, these two layers providing the same picture resolution but different quality level. For example, the transmission of service with two different quality levels is expected to become useful in the future for some TV broadcast applications, especially when very good picture quality is needed for large size display receivers. The sequence is encoded into two bitstreams called lower and enhancement layer bitstreams. The lower layer bitstream can be decoded independently from the enhancement layer bitstream. The lower layer, at 3 to 4 Mbit/s, would provide a picture quality equivalent to the current NTSC/PAL/SECAM quality. Then, by using both the lower and the enhancement layer bitstreams, an enhanced decoder can deliver a picture quality subjectively close to the studio quality, with a total bitrate of 7 to 12 Mbit/s.

#### D.7.1.1 Additional features

##### D.7.1.1.1 Error resilience

As described in D.13, the SNR scalable scheme can be used as a mechanism for error resilience. If the two layer bitstreams are received with different error rate, the lower layer, better protected, stands as a good substitute to fall back on, if the enhancement layer is damaged.

##### D.7.1.1.2 Chroma simulcast

The SNR scalable syntax can be used in a chroma simulcast system. The goal of such a scheme would be to provide a mechanism for simultaneous distribution of services with the same luminance resolution but different chrominance sampling format (e.g. 4:2:0 in the lower layer and 4:2:2, when adding the enhancement layer and the simulcast chrominance components) for applications which would require such a feature. The SNR scalable enhancement layer contains some luminance refinement. The 4:2:2 chrominance is sent in simulcast. Only chrominance DC is predicted from the lower layer. The combination of both layer luminance and of the 4:2:2 chrominance constitutes the high quality level.

#### D.7.1.2 SNR scalable encoding process

##### D.7.1.2.1 Description

In the lower layer, the encoding is similar to the non scalable situation in terms of decisions, adaptive quantisation, buffer regulation. The intra or error prediction macroblocks are DCT transformed. The coefficients are then quantised using a first rather coarse quantiser. The quantised coefficients are then VLC coded and sent together with the required side information (macroblock_type, motion vectors, coded_block_pattern()).

In parallel, the quantised DCT coefficients coming from the lower layer, are dequantised. The residual error between the coefficients and the dequantised coefficients is then re-quantised, using a second finer quantiser. The resulting refinement coefficients are VLC coded and form the additional enhancement layer, together with a marginal amount of side information (quantiser_scale_code, coded_block_pattern() …). The non-intra VLC table is used for all the coefficients in the enhancement layer, since it is of differential nature.

##### D.7.1.2.2 A few important remarks

Since the prediction is the same for both layers, it is recommended to use the refined images in the motion estimation loop (e.g. the images obtained by the conjunction of the lower and the enhancement layer). Thus, there is a drift between the prediction used at the encoder side and what the low level decoder can get as a prediction. This drift does accumulate from P-picture to P-picture and is reset to zero at each I-Picture. However, the drift has been found to have little visual effect when there is an I- picture every 15 pictures or so.

Since the enhancement layer only contains refinement coefficients, the needed overhead is quite reduced: most of the information about the macroblocks (macroblock types, motion vectors …) are included in the lower layer. Therefore the syntax of this stream is very much simplified:

- the macroblock type table only indicates if the quantiser_scale_code in the enhancement layer has changed and if the macroblock is NOT-CODED (for first and last macroblock of the slices), which amounts to three VLC words.
- quantiser_scale_code in the enhancement layer is sent if the value has changed.
- coded_block_pattern() is transmitted for all coded macroblocks.

All NON-CODED macroblocks that are not at the beginning or end of a slice are skipped, since the overhead information can be deduced from the lower layer.

It is recommended to use different weighting matrices for the lower and the enhancement layer. Some better results are obtained when the first quantisation is steeper than the second one. However, it is recommended not to quantise too coarsely the DCT coefficient that corresponds to the interlace motion, to avoid juddering effects.

### D.7.2 Multiple resolution scalability bitstreams using SNR scalability

The aim of resolution scalability is to decode the base layer video suitable for display at reduced spatial resolution. In addition, it is desirable to implement a decoder with reduced complexity for this purpose. This functionality is useful for applications where the receiver display is either not capable or willing to display the full spatial resolution supported by both layers and for applications where software decoding is targeted. The method described in this subclause uses the SNR Scalability syntax outlined in clause 7 to transmit the video in two layers. Note that none of the options suggested in this subclause changes the structure of the highest resolution decoder, which remains identical to the one outlined in Figure 7-14. The bitstream generated on both layers is compatible with the HIGH profile. However, the base layer decoder could be implemented differently with reduced implementation complexity suitable to software decoding.

#### D.7.2.1 Decoder implementation

In decoding to a smaller spatial resolution, an inverse DCT of reduced size could be used when decoding the base layer. The frame memory requirement in the decoder MC loop would also be reduced accordingly.

If the bitstream of the two SNR Scalability layers was generated with only one MC loop at the encoder the base video will be subject to drift. This drift may or may not be acceptable depending on the application. Image quality will, to a large extent, depend on the sub-sample accuracy used for motion compensation in the decoder. It is possible to use the full precision motion vector as transmitted in the base layer for motion compensation with a sub-sample accuracy comparable to that of the higher layer. Drift can be minimised by using advanced sub-sample interpolation filters (see [12], [13] and [16] in the Bibliography of Annex F).

#### D.7.2.2 Encoder implementation

It is possible to tailor the base layer SNR Scalability bitstream to the particular requirements of the resolution scaled decoder. A smaller DCT size can be more easily supported by only transmitting the appropriate DCT-coefficients belonging to the appropriate subset in the base layer bitstream.

Finally it is possible to support a drift-free decoding at lower resolution scale by incorporating more than one MC loop in the encoder scheme. An identical reconstruction process is used in the encoder and decoder.

### D.7.3 Bitrate allocation in data partitioning

Data partitioning allows splitting a bitstream for increased error resilience when two channels with different error performance are available. It is often required to constrain the bitrate of each partition. This can be achieved at the encoder by adaptively changing priority breakpoint at each slice.

The encoder can use two virtual buffers for the two bitstreams, and implement feedback rate control by picking a priority breakpoint that approximately meets the target rate for each channel. Difference between target and actual rates is used to revise the target for the next frame in a feedback loop.

It is desirable to vary the bitrate split from frame to frame for higher error resilience. Typically, I-pictures benefit from having more of the data in partition 0 than the P-pictures while B-pictures could be placed entirely in partition 1.

### D.7.4 Temporal scalability

A two layer temporally scalable coding structure consisting of a base and an enhancement layer is shown in Figure D.1. Consider video input at full temporal rate to temporal demultiplexer; in our example it is temporally demultiplexed to form two video sequences, one input to the base layer encoder and the other input to the enhancement layer encoder. The base layer encoder is a non-hierarchical encoder operating at half temporal rate, the enhancement layer encoder is like a MAIN profile encoder and also operates at half temporal rate except that it uses base layer decoded pictures for motion compensated prediction. The encoded bitstreams of base and enhancement layers are multiplexed as a single stream in the systems multiplexer. The systems demultiplexer extracts two bitstreams and inputs corresponding bitstreams to base and enhancement layer decoders. The output of the base layer decoder can be shown standalone at half temporal rate or after multiplexing with enhancement layer decoded frames and shown at full temporal rate.

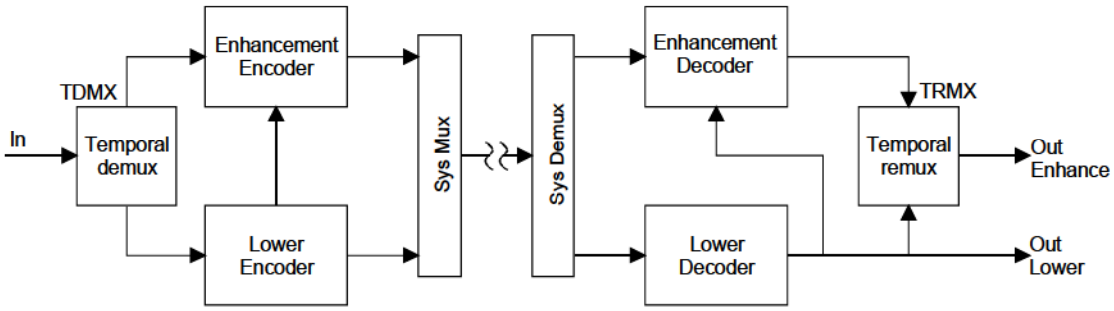



**Figure D.1 – A two-layer codec structure for temporal scalability**

The following forms of temporal scalability are supported and are expressed as higher layer: base layer-to-enhancement layer picture formats.

1) Progressive: progressive-to-progressive Temporal Scalability.
2) Progressive: interlace-to-interlace Temporal Scalability.
3) Interlace: interlace-to-interlace Temporal Scalability.

#### D.7.4.1 Progressive – Progressive-to-progressive temporal scalability

Assuming progressive video input, if it is necessary to code progressive- format video in base and enhancement layers, the operation of *temporal demux* may be relatively simple and involve temporal demultiplexing of input frames into two progressive sequences. The operation of *temporal remux* is inverse, i.e. it performs re-multiplexing of two progressive sequences to generate full temporal rate progressive output (see Figure D.2).

#### D.7.4.2 Progressive – Interlace-to-interlace temporal scalability

Again, assuming full temporal rate progressive video input, if it is necessary to code interlaced format video in base layer, the operation of *temporal demux* may involve progressive to two interlace conversion; this process involves extraction of a normal interlaced- and a complementary interlaced sequence from progressive input video. The operation of *temporal remux* is inverse, i.e. it performs two interlace to progressive conversion to generate full temporal rate progressive output. Figures D.3 and D.4 show operations required in progressive to two interlace and two interlace to progressive conversion.

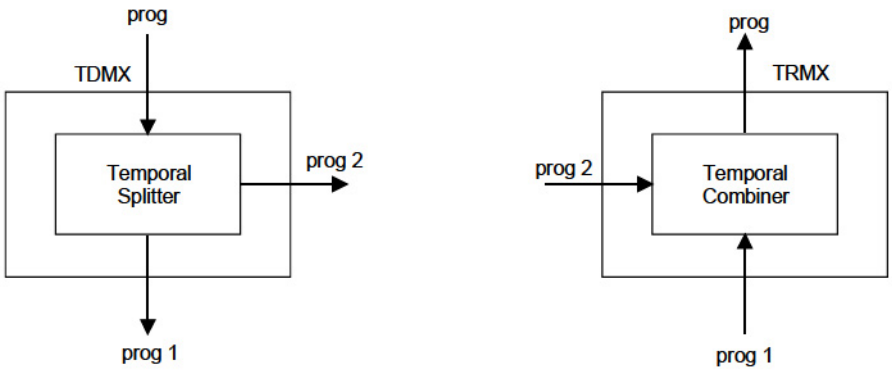



**Figure D.2 – Temporal demultiplexer and remultiplexer for progressive – Progressive-to-progressive temporal scalability**

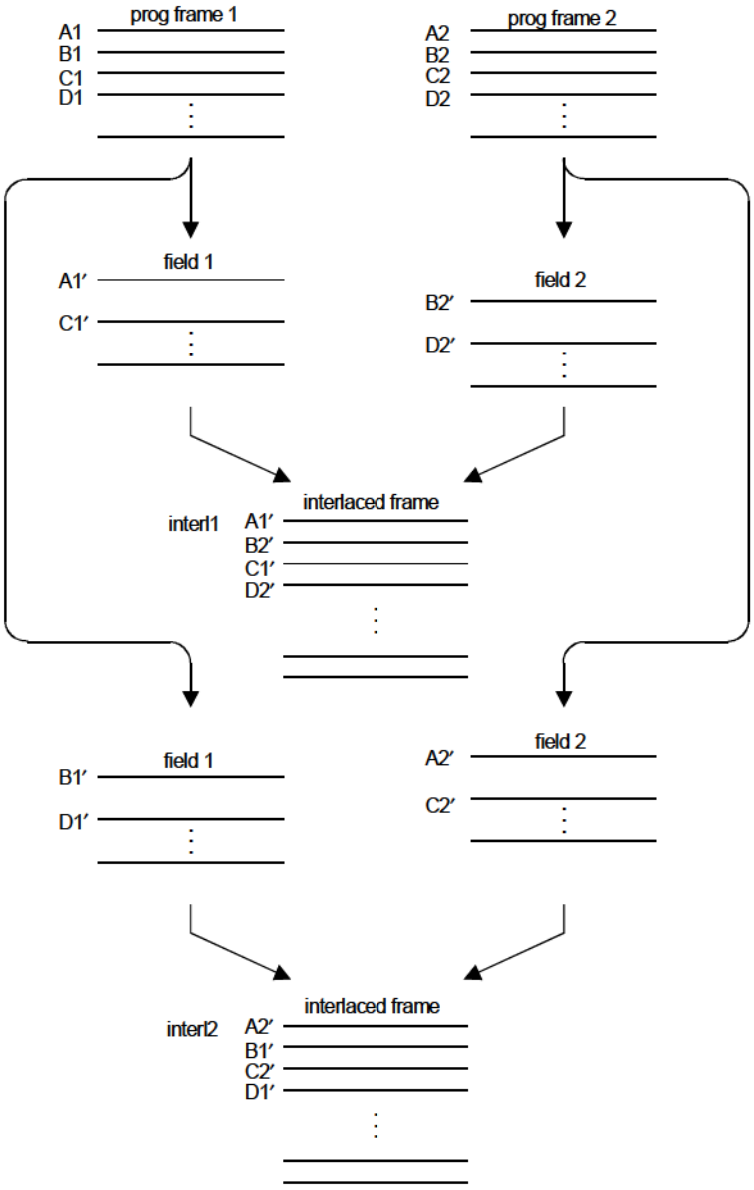



**Figure D.3 – Progressive to two-interlace conversion**

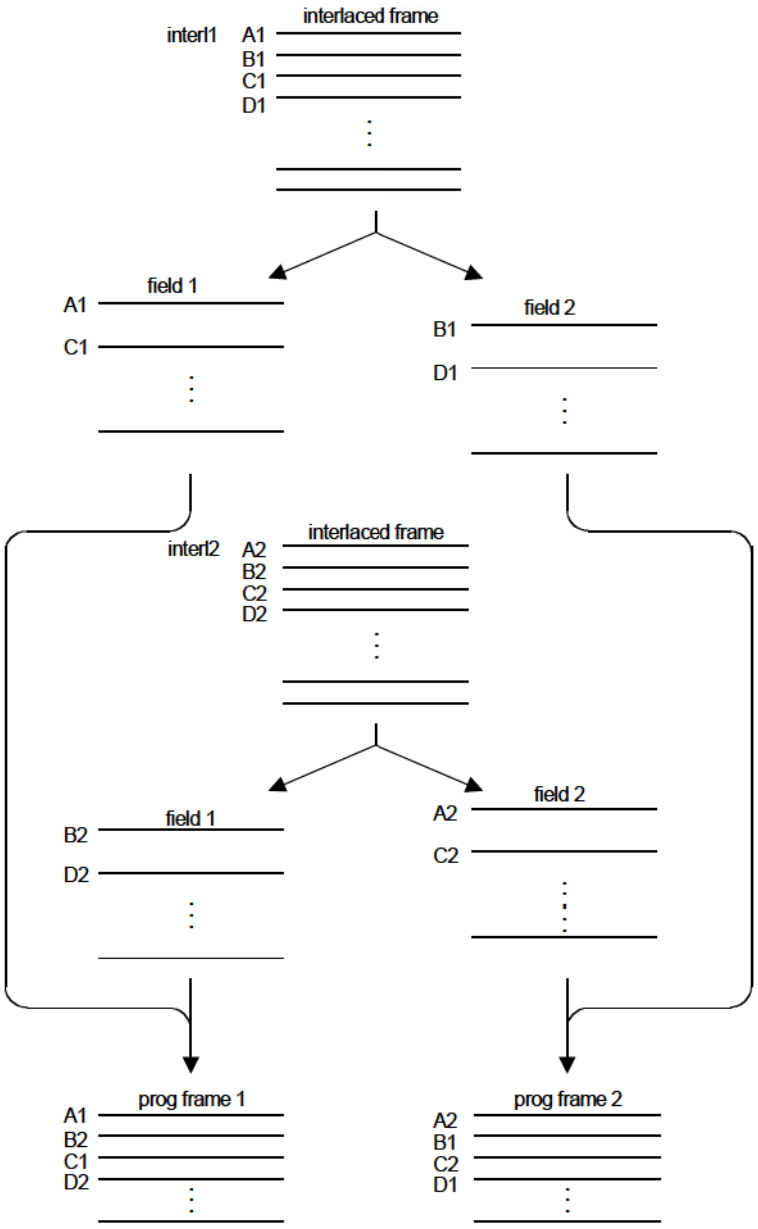



**Figure D.4 – Two-interlace to progressive conversion**

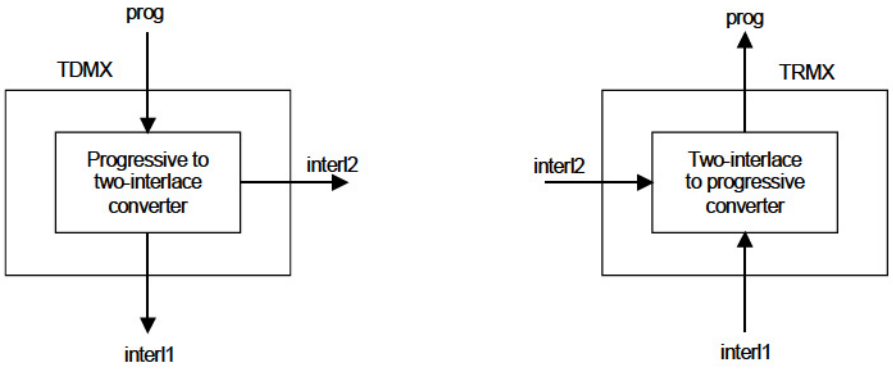



**Figure D.5 – Temporal demultiplexer and remultiplexer for progressive – Interface-to-interface temporal scalability**

### D.7.4.3 Interlace – Interlace-to-interlace temporal scalability

Assuming interlaced video input, if it is necessary to code interlaced- format video in base and enhancement layers, the operation of *temporal demux* may be relatively simple and involve temporal demultiplexing of input frames into two interlaced sequences. The operation of *temporal remux* is inverse, i.e. it performs re-multiplexing of two interlaced sequences to generate full temporal rate interlaced output. The demultiplexing and re-multiplexing is similar to that in Figure D.2.

### D.7.5 Hybrids of the spatial, the SNR and the temporal scalable extensions

This Recommendation | International Standard also allows combinations of scalability tools to produce more than 2 video layers as may be useful and practical to support more demanding applications. Taken two at a time, 3 explicit combinations result. Moreover, within each combination, the order in which each scalability is applied, when interchanged, results in distinct applications. In the hybrid scalabilities involving three layers, the layers are referred to as base layer, enhancement layer 1 and enhancement layer 2.

### D.7.5.1 Spatial and SNR hybrid scalability applications

a) *HDTV with standard TV at two qualities:*

Base layer provides standard TV resolution at basic quality, enhancement layer 1 helps generate standard TV resolution but at higher quality by SNR scalability and the enhancement layer 2 employs HDTV resolution and format which is coded with spatial scalability with respect to high quality standard TV resolution generated by using enhancement layer 1.

b) *Standard TV at two qualities and low definition TV/videophone:*

Base layer provides videophone/low definition quality, using spatial scalability enhancement layer 1 provides standard TV resolution at a basic quality and enhancement layer 2 uses SNR scalability to help generate high quality standard TV.

c) *HDTV at two qualities and standard TV:*

Base layer provides standard TV resolution, using spatial scalability enhancement layer 1 provides basic quality HDTV and enhancement layer 2 uses SNR scalability to help generate high quality HDTV.

### D.7.5.2 Spatial and temporal hybrid scalability applications

a) *High temporal resolution progressive HDTV with basic interlaced HDTV and standard TV:*

Base layer provides standard TV resolution, using spatial scalability enhancement layer 1 provides basic HDTV of interlaced format and enhancement layer 2 uses temporal scalability to help generate full temporal resolution progressive HDTV.

b) *High resolution progressive HDTV with enhanced progressive HDTV and basic progressive HDTV:*

Base layer provides basic progressive HDTV format at temporal resolution, using temporal scalability enhancement layer 1 helps generate progressive HDTV at full temporal resolution and enhancement layer 2 uses spatial scalability to provide high spatial resolution progressive HDTV (at full temporal resolution).

c) *High resolution progressive HDTV with enhanced progressive HDTV and basic interlaced HDTV:*

Base layer provides basic interlaced HDTV format, using temporal scalability enhancement layer 1 helps generate progressive HDTV at full temporal resolution and enhancement layer 2 uses spatial scalability to provide high spatial resolution progressive HDTV (at full temporal resolution).

#### D.7.5.3 Temporal and SNR hybrid scalability applications

a) *Enhanced progressive HDTV with basic progressive HDTV at two qualities:*

Base layer provides basic progressive HDTV at lower temporal rate, using temporal scalability enhancement layer 1 helps generate progressive HDTV at full temporal rate but with basic quality and enhancement layer 2 uses SNR scalability to help generate progressive HDTV with high quality (at full temporal resolution).

b) *Enhanced progressive HDTV with basic interlaced HDTV at two qualities:*

Base layer provides interlaced HDTV of basic quality, using SNR scalability enhancement layer 1 helps generate interlaced HDTV at high quality and enhancement layer 2 uses temporal scalability to help generate progressive HDTV at full temporal resolution (at high quality).

## D.8 Compatibility

This Recommendation | International Standard supports compatibility between different resolution formats as well as compatibility with ISO/IEC 11172-2 (and Recommendation H.261).

### D.8.1 Compatibility with higher and lower resolution formats

This Specification supports compatibility between different resolution video formats. Compatibility is provided for spatial and temporal resolutions with the Spatial Scalability and Temporal Scalability tools. The video is encoded into two resolution layers. A decoder only capable or willing to display a lower resolution video accepts and decodes the lower layer bitstream. The full resolution video can be reconstructed by accepting and decoding both resolution layers provided.

### D.8.2 Compatibility with ISO/IEC 11172-2 (and Recommendation H.261)

The syntax of this Specification supports both backward and forward compatibility with ISO/IEC 11172-2. Forward compatibility with ISO/IEC 11172-2 is provided since the syntax of this Specification is a superset of the ISO/IEC 11172-2 syntax. The Spatial Scalability tool provided by this Specification allows using ISO/IEC 11172-2 coding in the lower resolution, i.e. base layer, thus achieving backward compatibility.

The video syntax contains tools that are needed to implement H.261 compatibility that may be needed for possible future use, however, this is currently not supported by any profile.

Simulcast serves as a simple alternative method to provide backward compatibility with both Recommendation H.261 and ISO/IEC 11172-2.

## D.9 Differences between this Specification and ISO/IEC 11172-2

This subclause lists the differences between MPEG-1 Video and MPEG-2 Video.

All MPEG-2 Video decoders that comply with currently defined profiles and levels are required to decode MPEG-1 constrained bitstreams.

In most instances, MPEG-2 represents a super-set of MPEG-1. For example, the MPEG-1 coefficient zigzag scanning order is one of the two coefficient scanning modes of MPEG-2. However, in some cases, there are syntax elements (or semantics) of MPEG-1 that does not have a direct equivalent in MPEG-2. This Specification lists all those elements.

This Specification may help implementers identify those elements of the MPEG-1 video syntax (or semantics) that do not have their direct equivalent in MPEG-2, and therefore require a special care in order to have guarantee MPEG-1 compatibility.

In this subclause, MPEG-1 refers to ISO/IEC 11172-2 whilst MPEG-2 refers to this Specification.

### D.9.1 IDCT mismatch

MPEG-1 – The IDCT mismatch control consists in adding (or removing) one to each non-zero coefficient that would have been even after inverse quantisation. This is described as part of the inverse quantisation process, in 2.4.4.1, 2.4.4.2 and 2.4.4.3 of ISO/IEC 11172-2 (MPEG-1).

MPEG-2 – The IDCT mismatch control consists in adding (or removing) one to coefficient [7] [7] if the sum of all coefficients is even after inverse quantisation. This is described in 7.4.4 of MPEG-2.

### D.9.2 Macroblock stuffing

MPEG-1 – The VLC code '0000 0001 111' (macroblock_stuffing) can be inserted any number of times before each macroblock_address_increment. This code must be discarded by the decoder. This is described in 2.4.2.7 of MPEG-1.

MPEG-2 – This VLC code is reserved and not used in MPEG-2. In MPEG-2, stuffing can be generated only by inserting zero bytes before a start-code. This is described in 5.2.3 of MPEG-2.

### D.9.3 Run-level escape syntax

MPEG-1 – Run-level values that cannot be coded with a VLC are coded by the escape code '0000 01' followed by either a 14-bit FLC (–127 <= level <= 127), or a 22-bit FLC (–255 <= level <= 255). This is described in Annex B, 2-B5 of MPEG-1.

MPEG-2 – Run-level values that cannot be coded with a VLC are coded by the escape code '0000 01' followed by a 18-bit FLC (–2047 <= level <= 2047). This is described in 7.2.2.3 of MPEG-2.

### D.9.4 Chrominance samples horizontal position

MPEG-1 – The horizontal position of chrominance samples is half the way between luminance samples. This is described in 2.4.1 of MPEG-1.

MPEG-2 – The horizontal position of chrominance samples is co-located with luminance samples. This is described in 6.1.1.8 of MPEG-2.

### D.9.5 Slices

MPEG-1 – Slices do not have to start and end on the same horizontal row of macroblocks. Consequently it is possible to have all the macroblocks of a picture in a single slice. This is described in 2.4.1 of MPEG-1.

MPEG-2 – Slices always start and end on the same horizontal row of macroblocks. This is described in 6.1.2 of MPEG-2.

### D.9.6 D-Pictures

MPEG-1 – A special syntax is defined for D-pictures (picture_coding_type = 4). D-pictures are like I-pictures with only Intra-DC coefficients, no End of Block, and a special end_of_macroblock code '1'.

MPEG-2 – D-pictures (picture_coding_type = 4) are not permitted. This is described in 6.3.9 of MPEG-2.

### D.9.7 Full-pel motion vectors

MPEG-1 – The syntax elements full_pel_forward_vector and full_pel_backward_vector can be set to '1'. When this is the case, the motion vectors that are coded are in integer-sample units instead of half-sample units. Motion vector coordinates must be multiplied by two before being used for the prediction. This is described in 2.4.4.2 and 2.4.4.3 of MPEG-1.

MPEG-2 – The syntax elements full_pel_forward_vector and full_pel_backward_vector must be equal to '0'. Motion vectors are always coded in half-sample units.

### D.9.8 Aspect ratio information

MPEG-1 – The 4-bit pel_aspect_ratio value coded in the sequence header specifies the sample aspect ratio. This is described in 2.4.3.2 of MPEG-1.

MPEG-2 – The 4-bit aspect_ratio_information value coded in the sequence header specifies the display aspect ratio. The sample aspect ratio is derived from this and from the frame size and display size. This is described in 6.3.3 of MPEG-2.

### D.9.9 forward_f_code and backward_f_code

MPEG-1 – The f_code values used for decoding the motion vectors are forward_f_code and backward_f_code, located in the picture_header().

MPEG-2 – The f_code values used for decoding the motion vectors are f_code[s][t], located in the picture_coding_extension(). The values of forward_f_code and backward_f_code must be '111' and are ignored. This is described in 6.3.9 of MPEG-2.

### D.9.10 constrained_parameter_flag and maximum horizontal_size

MPEG-1 – When the constrained_parameter_flag is set to '1', this indicates that a certain number of constraints are verified. One of those constraints is that horizontal_size <= 768. It should be noted that a constrained MPEG-1 video bitstream can have pictures with an horizontal size of up to 768 luminance samples. This is described in 2.4.3.2 of MPEG-1.

MPEG-2 – The constrained_parameter_flag mechanism has been replaced by the profile and level mechanism. However, it should be noted that MP @ ML bitstreams cannot have horizontal size larger than 720 luminance samples. This is described in 8.2.3.1 of MPEG-2.

### D.9.11 Bit_rate and vbv_delay

MPEG-1 – Bit_rate and vbv_delay are set to 3FFFF and FFFF (hex) respectively to indicate variable bitrate. Other values are for constant bitrate.

MPEG-2 – The semantics for bit_rate are changed. In variable bitrate operation, vbv_delay may be set to FFFF (hex), but a different value does not necessarily mean that the bitrate is constant. Constant bitrate operation is simply a special case of variable bitrate operation. There is no way to tell that a bitstream is constant bitrate without examining all of the vbv_delay values and making complicated computations.

Even if the bitrate is constant the value of bit_rate may not be the actual bitrate since bit_rate need only be an upper bound to the actual bitrate.

### D.9.12 VBV

MPEG-1 – VBV is only defined for constant bitrate operation. The STD supersedes the VBV model for variable bitrate operation.

MPEG-2 – VBV is only defined for variable bitrate operation. Constant bitrate operation is viewed as a special case of variable bitrate operation.

### D.9.13 temporal_reference

MPEG-1 – Temporal_reference is incremented by one modulo 1024 for each coded picture, and reset to zero at each group of pictures header.

MPEG-2 – If there are no big pictures, temporal_reference is incremented by one modulo 1024 for each coded picture, and reset to zero at each group of pictures header (as in MPEG-1). If there are big pictures (in low delay bitstreams), then temporal_reference follows different rules.

### D.9.14 MPEG-2 syntax versus MPEG-1 syntax

It is possible to make MPEG-2 bitstreams that have a syntax very close to MPEG-1, by using particular values for the various MPEG-2 syntax elements that do not exist in the MPEG-1 syntax.

In other words, the MPEG-1 decoding process is the same (except for the particular points mentioned earlier) as the MPEG-2 decoding process when :

progressive_sequence = '1' (progressive sequence)

chroma_format = '01' (4:2:0)

frame_rate_extension_n = 0 and frame_rate_extension_d = 0 (MPEG-1 frame-rate)

intra_dc_precision = '00' (8-bit Intra-DC precision)

picture_structure = '11' (frame-picture, because progressive_sequence = '1')

frame_pred_frame_dct = 1 (only frame-based prediction and frame DCT)

concealment_motion_vectors = '0' (no concealment motion vectors)

q_scale_type = '0' (linear quantiser_scale)

intra_vlc_format = '0' (MPEG-1 VLC table for Intra MBs)

alternate_scan = '0' (MPEG-1 zigzag scanning order)

repeat_first_field = '0' (because progressive_sequence = '1')

chroma_420_type = '1' (chrominance is "frame-based", because

progressive_sequence = '1')

progressive_frame = '1' (because progressive_sequence = '1')

## D.10 Complexity

The MPEG-2 Standard supports combinations of high performance/high complexity and low performance/low complexity decoders. This is accommodated by MPEG-2 with the Profiles and Levels definitions which introduce new sets of tool and functionality with every new profile. It is thus possible to trade-off performance of the MPEG-2 coding schemes by decreasing implementation complexity.

Moreover, certain restrictions could allow reducing decoder implementation cost.

## D.11 Editing encoded bitstreams

Many operations on the encoded bitstream are supported to avoid the expense and quality costs of re-coding. Editing, and concatenation of encoded bitstreams with no re-coding and no disruption of the decoded image sequence is possible.

There is a conflict between the requirement for high compression and easy editing. The coding structure and syntax have not been designed with the primary aim of simplifying editing at any picture. Nevertheless, a number of features have been included that enable editing of coded data.

Editing of encoded MPEG-2 bitstreams is supported due to the syntactic hierarchy of the encoded video bitstream. Unique start codes are encoded with different level in the hierarchy (i.e. video sequence, group of pictures, etc.). Video can be encoded with Intra-picture/intra-slices access points in the bitstream. This enables the identification, access and editing of parts of the bitstream without the necessity to decode the entire video.

## D.12 Trick modes

Certain Digital Storage Media (DSM) provide the capability of trick modes, such as Fast Forward/Fast Reverse (FF/FR). The MPEG-2 syntax supports all special access, search and scan modes of ISO/IEC 11172-2. This functionality is supported with the syntactic hierarchy of the video bitstream which enables the identification of relevant parts within a video sequence. It can be assisted by MPEG-2 tools which provide bitstream scalability to limit the access bitrate (i.e. Data Partitioning and the general slice structure). This subclause provides some guideline for decoding a bitstream provided by a DSM.

The decoder is informed by means of a 1-bit flag (DSM_trick_mode_flag) in the PES packet header. This flag indicates that the bitstream is reconstructed by DSM in trick mode, and the bitstream is valid from syntax point of view, but invalid from semantics point of view. When this bit is set, an 8-bit field (DSM_trick_modes) follows. The semantics of DSM_trick_modes are in the ITU-T Rec. H.220.0 | ISO/IEC 13818-1.

### D.12.1 Decoder

While the decoder is decoding PES Packet whose DSM_trick_mode_flag is set to 1, the decoder is recommended to:

- Decode bitstream and display according to DSM_trick_modes.

**Pre-processing**

When the decoder encounters PES Packet whose DSM_trick_mode_flag is set to 1, the decoder is recommended to:

- Clear non trick mode bitstream from buffer.

**Post-processing**

When the decoder encounters PES Packet whose DSM_trick_mode_flag is set to 0, the decoder is recommended to:

- Clear trick mode bitstream from buffer.

**Video Part**

While the decoder is decoding PES Packet whose DSM_trick_mode_flag is set to 1, the decoder is recommended to:

- Neglect vbv_delay and temporal_reference value.
- Decode one picture and display it until next picture is decoded.

The bitstream in trick mode may have a gap between slices. When the decoder encounters a gap between slices, the decoder is recommended to:

- Decode the slice and display it according to the slice vertical position in slice header.
- Fill up the gap with co-sited part of the last displayed picture.

### D.12.2 Encoder

The encoder is recommended to:

- Encode with short size of slice with intra macroblocks.
- Encode with short periodic refreshment by intra picture or intra slice.

**DSM**

DSM is recommended to provide the bitstream in trick mode with perfect syntax.

**Pre-processing**

DSM is recommended to:

- Complete "normal" bitstream at picture_header() and higher syntactic structures.

**System Part**

DSM is recommended to:

- Set DSM_trick_mode_flag to 1 in a PES Packet header.
- Set DSM_trick_modes(8-bit) according to the trick mode.

**Video Part**

DSM is recommended to:

- Insert a sequence_header() with the same parameters as a normal bitstream.
- Insert a sequence_extension() with the same parameters as a normal bitstream.
- Insert a picture_header() with the same parameters as a normal bitstream except that it may be preferable to indicate variable bit rate operation. One way to achieve this is to set vbv_delay to FFFF (hex).

  NOTE – In most cases temporal_reference and vbv_delay are ignored in a decoder, therefore the DSM may not need to set temporal_reference and vbv_delay to correct values.
- Concatenate slices which consists of intra coded macroblocks. The concatenated slices should have slice vertical positions in increasing order.

## D.13 Error resilience

Most digital storage media and communication channels are not error-free. Appropriate channel coding schemes should be used and are beyond the scope of this Specification. Nevertheless the MPEG-2 syntax supports error resilient modes relevant to cell loss in ATM networks and bit errors (isolated and in bursts) in transmissions. The slice structure of the compression scheme defined in this Specification allows a decoder to recover after a residual data error and to re-synchronise its decoding. Therefore, bit errors in the coded data will cause errors in the decoded pictures to be limited in area. Decoders may be able to use concealment strategies to disguise these errors. Error resilience includes graceful degradation in proportion to bit error rate (BER) and graceful recovery in the face of missing video bits or data packets. It has to be noted that all items may require additional support at the system level.

Being an example of a packet-based system, B-ISDN with its Asynchronous Transfer Mode (ATM) is addressed in some detail in the following. Similar statements can be made for other systems where certain packets of data are protected individually by means of forward error-correcting coding.

ATM uses short, fixed length packets, called cells, consisting of a 5 byte header containing routing information, and a user payload of 48 bytes. The nature of errors on ATM is such that some cells may be lost, and the user payload of some cells may contain bit errors. Depending on AAL (ATM Adaptation layer) functionality, indications of lost cells and cells containing bit errors may be available.

As an indication of the impact of cell loss in an ATM environment Table D.2 summarises the average interval between cell losses for a range of CLR and service bitrates based on simple statistical modelling. (A cell payload must be assumed for this. Allowing 1 byte/cell for AAL functions leaves 376 bits = 47 bytes). Note, however, that this summary ignores cell loss bursts and other shorter term temporal statistics.

Bit Error Ratios (BER) corresponding to the above mean times between errors can be calculated easily for the case of isolated bit errors. The BER that would cause the same incidence rate of errors is found by dividing by the cell payload size, i.e. BER = CLR/376.

**Table D.2 – Average interval between cell losses for a range of CLR and service bitrates**

| | Average interval time of error | | | |
|---|---|---|---|---|
| | 5 Mbit/s | 10 Mbit/s | 50 Mbit/s | 100 Mbit/s |
| $10^{-2}$ | 7,52 ms | 3,76 ms | 0,752 ms | 0,376 ms |
| $10^{-3}$ | 75,2 ms | 37,6 ms | 7,52 ms | 3,76 ms |
| $10^{-4}$ | 752 ms | 376 ms | 75,2 ms | 37,6 ms |
| $10^{-5}$ | 7,52 s | 3,76 s | 752 ms | 376 ms |
| $10^{-6}$ | 1,25 m | 37,6 s | 7,52 s | 3,76 s |
| $10^{-7}$ | 12,5 m | 6,27 m | 1,25 m | 37,6 s |
| $10^{-8}$ | 2,09 h | 1,04 h | 12,5 m | 6,27 m |

The following techniques of minimising the impact of lost cells and other error/loss effects are provided for reference, and indicate example methods of using the various tools available in this Specification to provide good performance in the presence of those errors. Note that the techniques described may be applicable in the cases of packets of other sizes (e.g. LANs or certain storage media) or video data with uncorrected errors of different characteristics, in addition to cell loss. It may be appropriate to treat a known erasure [uncorrected bit error(s) known to exist somewhere in a data block] as a lost data block, since the impact of bit errors cannot be predicted. However, this should be a decoder option. The discussion that follows refers generally to "transport packets" where appropriate, to emphasise the applicability to a variety of transport and storage systems. However, specific examples will refer to Cell Loss Ratios (CLR) because cell transport is the most completely defined at the time of preparing this Specification.

The error resilience techniques are summarised in three categories, covering methods of concealing the error once it has occurred, and the restriction of the influence of a loss or error in both space (within a picture) and time (from picture to picture).

### D.13.1 Concealment possibilities

Concealment techniques hide the effect of losses/errors once they have occurred. Some concealment methods can be implemented using any encoded bitstream, while others are reliant on the encoder to structure the data or provide additional information to enable enhanced performance.

#### D.13.1.1 Temporal predictive concealment

A decoder can provide concealment of the errors by estimating the lost data from spatio-temporally adjacent data. The decoder uses information which has been successfully received to make an informed estimate of what should be displayed in place of the lost/errored data, under the assumption that the picture characteristics are fairly similar across adjacent blocks (in both the spatial and temporal dimensions). In the temporal case, this means estimation of errored or lost data from nearby fields or frames.

##### D.13.1.1.1 Substitution from previous frame

The simplest possible approach is to replace a lost macroblock with the macroblock in the same location in the previous picture. This approach is suitable for relatively static picture areas but block displacement is noticeable for moving areas.

The "previous picture" must be interpreted with care due to the use of bi-directional prediction and a difference between picture coded order and picture display order. When a macroblock is lost in a P- or I-picture, it can be concealed by copying the data corresponding to the same macroblock in the previous P-picture or I-picture. This ensures that the picture is complete before it is used for further prediction. Lost macroblocks in B-pictures can be substituted from the last displayed picture, of any type, or from a future I- or P-picture held in memory but not yet displayed.

##### D.13.1.1.2 Motion compensated concealment

The concealment from neighbouring pictures can be improved by estimating the motion vectors for the lost macroblock, based on the motion vectors of neighbouring macroblocks in the affected picture (provided these are not also lost). This improves the concealment in moving picture areas, but there is an obvious problem with errors in macroblocks whose neighbouring macroblocks are coded intra, because there are ordinarily no motion vectors. Encoder assistance to get around this problem is discussed in D.13.1.1.3.

Sophisticated motion vector estimation might involve storage of adjacent macroblock motion vectors from above and below the lost macroblock, for predictions both forward and backward (for B-pictures) in time. The motion vectors from above and below (if available) could then be averaged.

Less complex decoders could use, for example, only forward prediction and/or only the motion vector from the macroblock above the lost macroblock. This would save on storage and interpolation.

#### D.13.1.1.3 Use of Intra MVs

The motion compensated concealment technique outlined in D.13.1.1.2 could not ordinarily be applied when the macroblocks above and below the lost/errored macroblock are Intra-coded, since there is no motion vector associated with Intra-coded macroblocks. In particular, in I-pictures, this type of concealment would not be possible with the normal calculation and use of motion vectors.

The encoding process can be extended to include motion vectors for intra macroblocks. Of course, the motion vector and coded information for a particular macroblock must be transmitted separately (e.g. in different packets) so that the motion vector is still available in the event that the image data is lost.

When "concealment_motion_vectors" = 1, motion vectors are transmitted with Intra macroblocks, allowing improved concealment performance of the decoders. The concealment motion vector associated with an Intra-coded macroblock is intended to be used only for concealment (if necessary) of the macroblock located immediately below the Intra-coded macroblock.

For simplicity, concealment motion vectors associated with Intra-coded macroblocks are always forward, and are considered as frame motion-vectors in Frame pictures and field motion-vectors in field pictures.

Therefore, encoders that choose to generate concealment motion vectors should transmit, for a given Intra-coded macroblock, the frame- or field-motion vector that should be used to conceal (i.e. to predict, with forward frame- or field-based prediction respectively) the macroblock located immediately below the Intra-coded macroblock.

Concealment motion vectors are intended primarily for I- and P-pictures, but the syntax allows their use in B-pictures. Concealment in B-pictures is not critical, since B-pictures are not used as predictors and so errors do not propagate to other pictures. Therefore, it may be wasteful to transmit concealment motion vectors in B-pictures.

Concealment motion vectors transmitted with Intra macroblocks located in the bottom row of a picture cannot be used for concealment. However, if "concealment_motion_vectors" = 1, those concealment motion vectors must be transmitted. Encoders can use the (0, 0) motion vector to minimise the coding overhead.

When concealment motion vectors are used, it is a good idea to have one slice contain one row of macroblocks (or smaller), so that concealment can be limited to less than one row of macroblocks when a slice, or part of a slice, is lost. This means that the loss of macroblocks in two successive rows is much less likely, and therefore the chances of achieving effective concealment using concealment motion vectors is improved.

NOTE – When "concealment_motion_vectors" = 1, Predictors for Motion Vectors (PMV) are NOT reset when an Intra macroblock is transmitted. Ordinarily, an Intra macroblock would reset the PMVs.

### D.13.1.2 Spatial predictive concealment

The generation of predicted, concealment macroblocks is also possible by interpolation from neighbouring macroblocks within the one picture (see Annex F, Bibliography, reference [17]). This is best suited to areas of high motion, where temporal prediction is not successful, or as an alternative means of concealment for Intra macroblocks when concealment motion vectors (see D.13.1.1.3) are not available. It also could be particularly useful for cell loss after scene changes.

There are several possible approaches to spatial interpolation, and it could be carried out in the spatial or DCT domain, but normally it is only feasible and useful to predict the broad features of a lost macroblock, such as the DC coefficient and perhaps the lowest AC coefficients. Spatial prediction of fine detail (high frequencies) is likely to be unsuccessful and is of little value in fast-moving pictures anyway.

Spatial predictive macroblock concealment may also be useful in combination with layered coding methods (i.e. Data Partitioning or SNR scalability, see D.13.1.3). If in the event of cell loss some DCT coefficients in a macroblock are recovered from the lower layer, it is possible to use all information available (DCT coefficients recovered in the same macroblock from the lower layer and all DCT coefficients received in the adjacent macroblocks) for error concealment. This is especially useful if the lower layer only contains DC coefficients due to bandwidth constraints.

### D.13.1.3 Layered coding to facilitate concealment

It is possible to assist the concealment process further by arranging the coded video information such that the most important information is most likely to be received. The loss of the less important information can then be more effectively concealed. This approach can gain from use of a transmission medium or storage device with different priority levels (such as priority-controlled cell-based transmission in the B-ISDN, or where different error protection or correction is provided on different channels). The components produced by the coding process can be placed in a hierarchy of importance according to the effect of loss on the reconstructed image. By indicating the priority of bitstream components and treating the individual components with due importance, superior error concealment performance may be possible.

Strategies available for producing hierarchically ordered bitstreams, or layers, include:

- **Data partitioning** – The coded macroblock data is partitioned into multiple layers such that partition zero contains address and control information and lower order DCT coefficients, while partition one contains high frequency DCT coefficients.
- **SNR scalability** – Two sets of coefficients are dequantised and then added together at the receiver before decoding. One set of coefficients could be a refinement of the quantisation error of the other, but other combinations (including an emulation of data partitioning) are possible.
- **Spatial scalability** – The lower layer may be coded without regard for the enhancement layer, and could use other standard coding methods (ISO/IEC 11172-2, etc.). The enhancement layer contains the coded prediction error from a prediction based on the lower layer.
- **Temporal scalability** – The enhancement layer defines additional pictures which, when re-multiplexed with the base layer, provides a combined picture sequence of greater picture rate.

These strategies produce layers which, when added progressively, produce increasing quality of the reconstructed sequence. While some of these source coding techniques may result in a bitrate increase compared to the system without layering, the performance of the layered systems, when subjected to channel errors, may be greater.

Considering error resilience alone, the hierarchically ordered layers should be handled with due quality, such that some function (such as picture quality for a given total bitrate) is optimised. The bitstream components may be treated differently at one or more of the following locations:

- *Encoder* – Different channel coding might be used.
- *Channel* – The channel may be able to provide different cell/packet loss probabilities or error characteristics to the different bitstream components.
- *Decoder* – Error concealment could be performed differently within each bitstream.

#### D.13.1.3.1 Use of data partitioning

Data partitioning allows a straightforward division of macroblock data into two layers. The Priority Break Point (PBP) pointer determines the contents of each layer. Ordinarily, data partition 0 contains the address and control information and the low frequency DCT coefficients, while data partition 1 contains the high frequency DCT coefficients.

At the encoder the value of the PBP pointer may be different for each slice such that the distribution of bits between the two layers may be controlled (e.g. maintained constant). The distribution may be different for I-, P-, and B-frames. The management of rate between the layers could mean that, for some macroblocks, data partition 0 contains no DCT coefficients or motion vectors.

Good tolerance to errors can be achieved if channel errors are distributed so that data partition 1 receives most errors.

It is assumed that errors can be detected at the decoder, so that actions can be taken to prevent errored data from being displayed. For data partition 1, errored data is simply not displayed (i.e. only data partition 0 is used). Losses or errors in data partition 0 should be minimised through use of high reliability transport. Decoder concealment actions may also be necessary.

#### D.13.1.3.2 Use of SNR scalable coding

SNR scalable coding provides two layers with the same spatial resolution but different image quality, depending on whether one or both layers are decoded. This technique is mainly intended to provide a lower-quality layer that is usable even when the enhancement layer is absent. However, it also provides good error resilience if the errors can be mainly confined to the enhancement layer.

In case of errors in the enhancement layer, the lower layer can be used alone for the affected image area. Especially in the case of frequent errors, temporary loss or permanent unavailability of the enhancement layer this concealment is very effective, since the displayed signal can be made relatively free of non-linear distortions like blocking or motion jerkiness.

If the enhancement layer is permanently unavailable and so only the lower layer is decoded, a small drift may occur in the case where only one MC prediction loop is implemented in the encoder. However, this drift is likely to be invisible in most configurations (e.g. M = 3, N = 12 would normally provide correction often enough).

The lower-layer of an SNR Scalable system is well suited to concealment in the case of a very high error rate, temporary or permanent loss of the enhancement-layer signal. However, the enhancement-layer quality in the error-free case does not achieve that of a sub-band like layered scheme (e.g. data partitioning).

#### D.13.1.3.3 Use of spatial scalable coding

Spatial scalable coding allows the lower layer to be coded without regard for the enhancement layer, and other standard coding methods (ISO/IEC 11172-2, etc.) could be used. The enhancement layer contains the coded prediction error from a prediction based on the lower layer. In case of errors in the enhancement layer, the upconverted lower layer can be used directly as concealment information for the affected image area. Especially in case of frequent errors or temporary loss of the enhancement layer, this concealment data is relatively free of non-linear distortions like blocking (which could arise if high frequency DCT coefficients are completely absent from the lower layer) or motion jerkiness (if the motion information is omitted from the high priority layer).

In the error-free case the upconverted lower layer is used as an additional source of predictions in a macroblock-adaptive way to improve the enhancement-layer coding performance. The enhancement layer bitstream therefore consists of the quantised temporal or lower layer prediction errors.

Spatial scalable coding provides a lower layer that is very suitable for concealment in case of a high error rate or temporary loss of the enhancement layer. However, the quality of the enhanced picture when both layers are available will not, in general, be as good as other layered coding approaches.

#### D.13.1.3.4 Use of temporal scalable coding

Temporal scalability is a coding technique that allows layering of video frames. The spatial resolution of frames in each layer is the same but the temporal rates of each layer are lower than that of the source; however the combined temporal rate of the two layers results in full temporal rate of the source. In case of errors in the enhancement layer, the base layer of full spatial resolution can be easily used for concealment. Especially in case of frequent errors or temporary loss of the enhancement layer, the base layer offers good concealment properties.

In some telecommunications applications a high degree of error resilience might be achieved with temporal scalability by encoding the base layer using the same spatial resolution but only half the temporal resolution of the source; the remaining frames corresponding to the other half of the temporal resolution are coded in the enhancement layer. Typically, the enhancement layer data may be assigned lower priority and when lost, the base layer decoded frames can be used for concealment by frame repetition. This type of concealment leads to only a temporary loss of full temporal resolution while maintaining full spatial quality and full spatial resolution.

In HDTV applications such as those using high temporal resolution progressive video format as source, high degree of error resilience can be achieved with temporal scalability. Such an application is envisaged to require 2 layers, a base layer and an enhancement layer, each of which process same picture formats (either both progressive or both interlaced) but at half the temporal rates. Temporal re-multiplexing of the base and enhancement layers irrespective of their chosen formats always results in full progressive temporal resolution of the source. In HDTV transmission, if the lower priority enhancement layer is corrupted, the base layer can be used for concealment, either directly, as in case of progressive format base layer or after reversal of parity of fields for interlaced format base layer.

Typically, the enhancement layer data may be assigned lower priority and when lost, the base layer decoded frames can be used for concealment by either frame repetition or frame averaging. This type of concealment leads to only a temporary indistinguishable loss in temporal resolution while maintaining full spatial quality and full spatial resolution.

### D.13.2 Spatial localisation

Spatial localisation encompasses those methods aimed at minimising the extent to which errors propagate within a picture, by providing early resynchronisation of the elements in the bitstream that are coded differentially between macroblocks.

Isolated bit errors may be detected through invalid codewords and so a decoder designer may choose to allow an errored sequence to be decoded. However, the effect on the picture is difficult to predict (legal, but incorrect, codewords could be generated) and it may be preferable to control the error through concealment of the entire affected slice(s) even when only one bit is known to be in error somewhere in a block of data.

When long consecutive errors occur (e.g. packet or cell loss), virtually the only option is to discard data until the next resynchronisation point is located (a start code at the next slice or picture header). By providing more resynchronisation points, the area of the screen affected by a loss or error can be reduced, in turn reducing the demands on the concealment techniques and making the errors less visible at the expense of coding efficiency. Spatial localisation of errors is therefore dependent on controlling the slice size since this is the smallest coded unit with resynchronisation points (start codes).

#### D.13.2.1 Small slices

The most basic method for achieving spatial localisation of errors is to reduce the (fixed) number of macroblocks in a slice. The increased frequency of resynchronisation points will reduce the affected picture area in the event of a loss. It is effective in any transport or storage media, and in any profile since the slice structure is always present in MPEG coded video.

The method results in a small loss of coding efficiency due to the increase of overhead information. The loss is about 3% for 11 Macroblocks per slice and 12% for 4 Macroblocks per slice based on Recommendation ITU-R BT.601 picture format at 4 Mbit/s, (percentages calculated relative to a system using 44 Macroblocks, or one picture width, per slice). The efficiency loss results in degradation of picture quality up to about 1 dB with 4 Macroblocks per slice and 0,2 dB with 11 Macroblocks per slice without errors at 4 Mbit/s. However, the method performs approximately 1 to 5 dB better at CLR = $10^{-2}$, depending on the concealment method used (simple macroblock replacement or motion compensated concealment).

From the view point of perceived picture quality, the performance of this method is generally dependent on the relative size of slice size and picture. Therefore, the slice size should be decided by considering the picture size (in macroblocks) and the trade-off between coding efficiency and visual degradation due to errors.

#### D.13.2.2 Adaptive slice size

There is a significant variation in the number of bits required to code a picture slice, depending on the coding mode, picture activity, etc. If slices contain only a few macroblocks, it will be possible that one transport packet, even a short packet or cell, could contain several slices. Offering multiple resynchronisation points in the same transport packet serves no purpose. Another problem with the simplistic short slice approach is that, because no account is taken of the transport packet structure, the first valid transport packet after a loss could contain most of the information for a slice, but it is unusable because the start code was lost.

An improvement over the small slice method may be to use adaptive slice sizes. As the encoder is producing the bitstream, it keeps track of the data contents within transport packets. The start of a slice is placed at the first opportunity in every transport packet (or in every second, third, …). This approach can achieve about the same spatial localisation of errors as small, fixed size slices, but with a greater efficiency.

However, this method ONLY gives an advantage for cell or packet based transmission, or where error detection occurs over a large block of data. The frequent resynchronisation points of small slice localisation are only wasteful if more than one is lost in the event of an error. If isolated bit errors affect just one slice anyway, then there is no advantage in adapting the slice size.

Furthermore, the adaptive slice size technique requires an intimate connection between encoder and packetiser, to allow a new slice for a new packet or cell. As such, it may not be appropriate for some applications (e.g. stored video intended to be distributed by multiple means) because only one transport packet structure would be assumed during encoding.

### D.13.3 Temporal localisation

Temporal localisation encompasses those methods aimed at minimising the extent to which errors propagate from picture to picture in the temporal sequence, by providing early resynchronisation of pictures that are coded differentially. An obvious way to do this is to make use of intra mode coding.

#### D.13.3.1 Intra pictures

By use of intra pictures a single error will not stay in the decoded picture longer than (N + M –1) pictures if every Nth picture is coded intra and (M-1) B-pictures are displayed before each I-picture.

While the intra pictures, normally used as "anchors" for synchronising the video decoding part way through a sequence, are useful for temporal localisation, care should be taken in adding extra intra pictures (i.e. reducing N) for error resilience. Intra pictures require a large number of bits to code, take up a relatively large proportion of the encoded bitstream and, as a result, are more likely to be affected by losses or errors themselves.

#### D.13.3.2 Intra slices

To avoid the additional delay caused by intra pictures, some applications requiring low delay may want to update the picture by coding only parts of the picture intra. This may provide the same kind of error resilience as intra pictures. As an example assume that a constant number of slices per picture from top to bottom are intra coded so that the whole picture is updated every P-pictures. Three aspects of this kind of updating should be kept in mind:

- While an errored portion of the scene will ordinarily be erased within P-pictures (with an average duration of about P/2), it is possible that motion compensation will allow the disturbance to bypass the intra refresh and it may persist as long as 2P-pictures.

- To ensure that errors are not propagating into the updated region of the picture, restrictions could be put on motion vectors, limiting the vertical vector components to ensure that predictions are not made from the "oldest" parts of the picture.

- The visual effect of clearing errors can be similar to a windscreen wiper clearing water. This *windscreen wiper* effect can become noticeable in some cases in the error free sequence, unless the rate control mechanism ensures that the quality of the intra slice is close to that of the surrounding non-intra macroblocks.

### D.13.4 Summary

Table D.3 summarises the above error resilience techniques, with a guide to their applicability.

**Table D.3 – Summary of error concealment techniques**

| Category | Technique | Profile/Applicability |
|---|---|---|
| Concealment | Temporal predictive– Substitution from previous picture | Any profile. Most suited to static pictures. |
| | Temporal predictive – Motion compensated | Any profile. Choice of sophistication in motion vector estimation. |
| | Temporal predictive – Using concealment MVs | Any profile, but calculation of Intra MVs is an encoder option. |
| | Spatial predictive | Any profile. Not suitable for static, complex pictures. |
| | Data Partitioning | Not currently used in a profile, but may be added as post/pre-processing. Minimal overhead and complexity. Depending on bitrate allocation, lower layer may not provide usable pictures by itself. |
| | SNR Scalability | SNR SCALABLE, SPATIALLY SCALABLE, HIGH profiles. Suitable for very high error rates or temporary unavailability of the enhancement layer. Relatively simple to implement. |
| | Spatial Scalability | SPATIALLY SCALABLE and HIGH profiles. Suitable for very high error rates or temporary unavailability of the enhancement layer. |
| | Temporal Scalability | Not currently used in a profile. Suitable for very high error rates or temporary unavailability of the enhancement layer. |
| Spatial Localisation | Small Slices | Any profile |
| | Adaptive slice sizes | Any profile, but requires knowledge of transmission characteristics when packet size is decided. |
| Temporal Localisation | Intra pictures | Any profile, but has delay implications. |
| | Intra slices | Any profile, but errors may persist longer than for Intra picture method. |

It is not possible to provide a concise indication of error resilience performance, because assessments must necessarily be subjective and application dependent, and so should be taken as nothing more than a guide. It is also true that several different approaches to error resilience are likely to be used in combination. However, the following descriptions are provided as some guidance to performance. They are the results of cell loss experiments, looking only at cell-based transmission of video information.

A simple macroblock substitution from a previous frame combined with the small-slice method (4 macroblocks per slice) will provide adequate picture quality for most sequences in the presence of rather low error rates of around CLR = $10^{-5}$ (in a reference 4 Mbit/s, Main Profile, Main Level system).

Including sophisticated motion compensated concealment (with full spatial and temporal interpolation of motion vectors for lost macroblocks, and concealing losses in P-pictures that use intra slice updating, i.e. N = infinity, M = 1) provides adequate picture quality at CLR = $10^{-3}$ (again, in a reference 4 Mbit/s, Main Profile, Main Level system).

Operation in environments with greater loss may require use of one of the layered coding methods. With adequate protection of the high priority information, these schemes can provide adequate performance in the face of CLRs as high as $10^{-2}$ or even $10^{-1}$. Data partitioning, implemented as a post-processing function to a 4 Mbit/s Main Profile, Main Level system, with 50% of the rate allocated to each partition and no loss in the base layer, has been shown in one example to give approximately 0,5 dB loss in SNR at a CLR of $10^{-3}$, about 1,5 dB loss at $10^{-2}$, and with almost no visible degradation in either case.

Given the range of different layered coding approaches that are possible, some general comments may be useful. In general, it is not expected that inclusion of the most complex layered coding methods could be justified purely on the basis of error resilience. Instead, they could be utilised for error resilience if they were required to satisfy other system requirements. Data partitioning is very simple to implement and is likely to provide error resilience very nearly the same as any of the other methods except in the case of extremely high error rates (> 10% loss) or where the enhancement layer could be lost completely. SNR scalability is slightly more complex, and has slightly lower efficiency than data partitioning, but it is easier to produce lower layers of a usable quality when the enhancement layer is absent. Spatial scalability is more complex again, but provides a good lower layer picture quality at the expense of overall (two layer) efficiency.

## D.14 Concatenated sequences

Sequence concatenation occurs when an elementary stream contains a sequence ending with a sequence_end_code that is followed by another sequence starting with a sequence_start_code. Any parameter including but not limited to profile, level, VBV buffer size, frame rate, horizontal size, vertical size, or bitrate, which is not allowed to change within a single sequence may change from sequence to sequence.

The behaviour of the decoding process and display process for concatenated sequences is not within the scope of this Recommendation | International Standard. An application that needs to use concatenated sequences must ensure by private arrangement that the decoder will be able to decode and play concatenated sequences.

Applications should ensure that decoders will have an acceptable behaviour when parameters change. For example changes to the

- frame size;
- frame rate;
- field parity of the first displayed field of the new sequence versus the field parity of the final displayed field of the previous sequence;
- buffer status,

amongst others may cause problems.

# Annex E

## Profile and level restrictions

(This annex does not form an integral part of this Recommendation | International Standard)

## E.1 Syntax element restrictions in profiles

This annex tabulates all of the syntactic elements defined in this Specification. Each is classified to indicate whether it is required to be supported by a decoder compliant to a particular profile and level. Note that normative specifications for compliance are given in ISO/IEC 13818-4.

NOTE – This annex is informative and is simply intended as a summary of the normative restrictions set out in clause 8. If, because of an error in the preparation of this text, a discrepancy exists between clause 8 and Annex E, the normative text in clause 8 shall always take precedence.

In the Tables E.2 to E.19 a number of abbreviations are used as shown in Table E.1.

**Table E.1 – Abbreviations used in the Tables of Annex E**

| Abbreviation | Used in | Meaning |
|---|---|---|
| x | Status | Must be supported by the decoder. |
| o | Status | Need not be supported by the decoder. |
| D | Type | Item with Level-dependent parameters. |
| I | Type | Item independent of the Level in the Profile. |
| P | Type | Item for post-processing after decoding; the decoder must be capable of decoding bitstreams which contain these items, but their use is beyond the scope of this Recommendation \| International Standard. |
| NOTE – "Status" is kept blank if an entry is not a syntactic element. | | |

**Table E.2 – Sequence header**

| # | Syntactic elements | Status | | | | | | | Type | Comments |
|---|---|---|---|---|---|---|---|---|---|---|
| | | SIMPLE | MAIN | SNR | SPATIAL | HIGH | 4:2:2 | Multi-view | | |
| 01 | **horizontal_size_value** | x | x | x | x | x | x | x | D | Table 8-11 |
| 02 | **vertical_size_value** | x | x | x | x | x | x | x | D | Table 8-11 |
| 03 | **aspect_ratio_information** | x | x | x | x | x | x | x | P | |
| 04 | **frame_rate_code** | x | x | x | x | x | x | x | D | Table 8-11 |
| 05 | (sample rate)<br>NOTE – This is not a syntactic element | | | | | | | | D | Table 8-12. Sample rate is a product of samples/line, lines/frame and frames/sec. |
| 06 | **bit_rate_value** | x | x | x | x | x | x | x | D | Table 8-13 |
| 07 | **vbv_buffer_size_value** | x | x | x | x | x | x | x | D | Table 8-14 |
| 08 | **constrained_parameters_flag** | x | x | x | x | x | x | x | I | Set to "1" if ISO/IEC 11172-2 constrained,<br>Set to "0" if ITU-T Rec. H.262 \| ISO/IEC 13818-2 |
| 09 | **load_intra_quantiser_matrix** | x | x | x | x | x | x | x | I | |
| 10 | **intra_quantiser_matrix[64]** | x | x | x | x | x | x | x | I | |
| 11 | **load_non_intra_quantiser_matrix** | x | x | x | x | x | x | x | I | |
| 12 | **non_intra_quantiser_matrix[64]** | x | x | x | x | x | x | x | I | |
| 13 | sequence_extension() | x | x | x | x | x | x | x | I | Always present if ITU-T Rec. H.262 \| ISO/IEC 13818-2 |
| 14 | sequence_display_extension() | x | x | x | x | x | x | x | P | |
| 15 | sequence_scalable_extension() | o | o | x | x | x | o | x | I | Table 8-9 for maximum number of scalable layers |
| 16 | user_data() | x | x | x | x | x | x | x | I | Decoder may skip this data |

**Table E.3 – Sequence extension**

| # | Syntactic elements | Status | | | | | | | Type | Comments |
|---|---|---|---|---|---|---|---|---|---|---|
| | | SIMPLE | MAIN | SNR | SPATIAL | HIGH | 4:2:2 | Multi-view | | |
| 01 | **profile_and_level_indication** | x | x | x | x | x | x | x | D | Profile: one of 8 values<br>Level: one of 16 values<br>Escape bit: one of 2 values |
| 02 | **progressive_sequence** | x | x | x | x | x | x | x | I | |
| 03 | **chroma_format** | x | x | x | x | x | x | x | I | Table 8-5 |
| 04 | **horizontal_size_extension** | x | x | x | x | x | x | x | D | Input picture size related |
| 05 | **vertical_size_extension** | x | x | x | x | x | x | x | D | Input picture size related |
| 06 | **bit_rate_extension** | x | x | x | x | x | x | x | D | Input picture size related |
| 07 | **vbv_buffer_size_extension** | x | x | x | x | x | x | x | D | Input picture size related |
| 08 | **low_delay** | x | x | x | x | x | x | x | I | |
| 09 | **frame_rate_extension_n** | x | x | x | x | x | x | x | I | Set to "0" for all defined profiles |
| 10 | **frame_rate_extension_d** | x | x | x | x | x | x | x | I | Set to "0" for all defined profiles |

**Table E.4 – Sequence display extension elements**

| # | Syntactic elements | Status | | | | | | | Type | Comments |
|---|---|---|---|---|---|---|---|---|---|---|
| | | SIMPLE | MAIN | SNR | SPATIAL | HIGH | 4:2:2 | Multi-view | | |
| 01 | **video_format** | x | x | x | x | x | x | x | P | |
| 02 | **colour_description** | x | x | x | x | x | x | x | P | Input format related |
| 03 | **colour_primaries** | x | x | x | x | x | x | x | P | |
| 04 | **transfer_characteristics** | x | x | x | x | x | x | x | P | |
| 05 | **matrix_coefficients** | x | x | x | x | x | x | x | P | |
| 06 | **display_horizontal_size** | x | x | x | x | x | x | x | P | Input format related |
| 07 | **display_vertical_size** | x | x | x | x | x | x | x | P | Input format related |

**Table E.5 – Sequence scalable extension**

| # | Syntactic elements | Status: SIMPLE | MAIN | SNR | SPATIAL | HIGH | 4:2:2 | Multi-view | Type | Comments |
|---|---|---|---|---|---|---|---|---|---|---|
| 01 | **scalable_mode** | o | o | x | x | x | o | x | I | SNR Profile: SNR Scalability<br>Spatial and High Profile: SNR and/or Spatial Scalability<br>Multi-view Profile: Temporal Scalability |
| 02 | **layer_id** | o | o | x | x | x | o | x | I | |
| | if (spatial scalable) | | | | | | | | | |
| 03 | **lower_layer_prediction_ horizontal_size** | o | o | o | x | x | o | o | D | Table 8-12 for luminance sampling density |
| 04 | **lower_layer_prediction_ vertical_size** | o | o | o | x | x | o | o | D | Table 8-12 for luminance sampling density |
| 05 | **horizontal_subsampling_ factor_m** | o | o | o | x | x | o | o | I | |
| 06 | **horizontal_subsampling_ factor_n** | o | o | o | x | x | o | o | I | |
| 07 | **vertical_subsampling_ factor_m** | o | o | o | x | x | o | o | I | |
| 08 | **vertical_subsampling_ factor_n** | o | o | o | x | x | o | o | I | |
| | if (temporal scalable) | | | | | | | | | |
| 09 | **picture_mux_enable** | o | o | o | o | o | o | x | I | |
| 10 | **mux_to_progressive_sequence** | o | o | o | o | o | o | x | I | |
| 11 | **picture_mux_order** | o | o | o | o | o | o | x | I | |
| 12 | **picture_mux_factor** | o | o | o | o | o | o | x | I | |

**Table E.6 – Group of pictures header**

| # | Syntactic elements | Status | | | | | | | Type | Comments |
|---|---|---|---|---|---|---|---|---|---|---|
| | | SIMPLE | MAIN | SNR | SPATIAL | HIGH | 4:2:2 | Multi-view | | |
| 01 | **time_code** | x | x | x | x | x | x | x | I | Decoder may skip this data |
| 02 | **closed_gop** | x | x | x | x | x | x | x | I | |
| 03 | **broken_link** | x | x | x | x | x | x | x | I | |

**Table E.7 – Picture header**

| # | Syntactic elements | Status | | | | | | | Type | Comments |
|---|---|---|---|---|---|---|---|---|---|---|
| | | SIMPLE | MAIN | SNR | SPATIAL | HIGH | 4:2:2 | Multi-view | | |
| 01 | **temporal_reference** | x | x | x | x | x | x | x | I | |
| 02 | **picture_coding_type** | x | x | x | x | x | x | x | I | Simple Profile: I, P at Main level, I, P, B at Low level<br>Main, SNR, Spatial, High and Multi-view Profile: I, P, B |
| 03 | **vbv_delay** | x | x | x | x | x | x | x | I | |
| 04 | **full_pel_forward_vector** | x | x | x | x | x | x | x | I | Set to "0" for ITU-T Rec. H.262 \| ISO/IEC 13818-2 |
| 05 | **forward_f_code** | x | x | x | x | x | x | x | I | Set to "111" for ITU-T Rec. H.262 \| ISO/IEC 13818-2 |
| 06 | **full_pel_backward_vector** | x | x | x | x | x | x | x | I | Set to "0" for ITU-T Rec. H.262 \| ISO/IEC 13818-2 |
| 07 | **backward_f_code** | x | x | x | x | x | x | x | I | Set to "111" for ITU-T Rec. H.262 \| ISO/IEC 13818-2 |
| 08 | **extra_information_picture** | x | x | x | x | x | x | x | I | |
| 09 | picture_coding_extension() | x | x | x | x | x | x | x | I | |
| 10 | quant_matrix_extension() | x | x | x | x | x | x | x | I | |
| 11 | picture_display_extension() | x | x | x | x | x | x | x | P | |
| 12 | picture_spatial_scalable_extension() | o | o | o | x | x | o | o | I | |
| 13 | picture_temporal_scalable_extension() | o | o | o | o | o | o | x | I | |
| 14 | camera_parameters_extension() | o | o | o | o | o | o | x | P | |

**Table E.8 – Picture coding extension**

| # | Syntactic elements | Status | | | | | | | Type | Comments |
|---|---|---|---|---|---|---|---|---|---|---|
| | | SIMPLE | MAIN | SNR | SPATIAL | HIGH | 4:2:2 | Multi-view | | |
| 01 | **f_code[0][0] (forward horizontal )** | x | x | x | x | x | x | x | D | Low Level [1:7]<br>Main Level [1:8]<br>High-1440 and High Level [1:9] |
| 02 | **f_code[0][1] (forward vertical)** | x | x | x | x | x | x | x | D | Low Level [1:4]<br>Main, High-1440 and High Level [1:5] |
| 03 | **f_code[1][0](backward horizontal)** | x | x | x | x | x | x | x | D | Low Level [1:7]<br>Main Level [1:8]<br>High-1440 and High Level [1:9] |
| 04 | **f_code[1][1] (backward vertical )** | x | x | x | x | x | x | x | D | Low level [1:4]<br>Main, H-14 and High Level [1:5] |
| 05 | **intra_dc_precision** | x | x | x | x | x | x | x | I | Simple, Main, SNR, Spatial and Multi-view Profile: [8:10]<br>High Profile: [8:11]<br>4:2:2 Profile: [8:11] |
| 06 | **picture_structure** | x | x | x | x | x | x | x | I | |
| 07 | **top_field_first** | x | x | x | x | x | x | x | I | |
| 08 | **frame_pred_frame_dct** | x | x | x | x | x | x | x | I | |
| 09 | **concealment_motion_vectors** | x | x | x | x | x | x | x | I | |
| 10 | **q_scale_type** | x | x | x | x | x | x | x | I | |
| 11 | **intra_vlc_format** | x | x | x | x | x | x | x | I | |
| 12 | **alternate_scan** | x | x | x | x | x | x | x | I | |
| 13 | **repeat_first_field** | x | x | x | x | x | x | x | I | |
| 14 | **chroma_420_type** | x | x | x | x | x | x | x | P | |
| 15 | **progressive_frame** | x | x | x | x | x | x | x | P | |
| 16 | **composite_display_flag** | x | x | x | x | x | x | x | P | |
| 17 | **v_axis** | x | x | x | x | x | x | x | P | |
| 18 | **field_sequence** | x | x | x | x | x | x | x | P | |
| 19 | **sub_carrier** | x | x | x | x | x | x | x | P | |
| 20 | **burst_amplitude** | x | x | x | x | x | x | x | P | |
| 21 | **sub_carrier_phase** | x | x | x | x | x | x | x | P | |

**Table E.9 – Quant matrix extension**

| # | Syntactic elements | Status | | | | | | | Type | Comments |
|---|---|---|---|---|---|---|---|---|---|---|
| | | SIMPLE | MAIN | SNR | SPATIAL | HIGH | 4:2:2 | Multi-view | | |
| 01 | **load_intra_quantiser_matrix** | x | x | x | x | x | x | x | I | |
| 02 | **intra_quantiser_matrix[64]** | x | x | x | x | x | x | x | I | |
| 03 | **load_non_intra_quantiser_matrix** | x | x | x | x | x | x | x | I | |
| 04 | **non_intra_quantiser_ matrix[64]** | x | x | x | x | x | x | x | I | |
| 05 | **load_chroma_intra_quantiser_ matrix** | o | o | o | o | x | x | o | I | |
| 06 | **chroma_intra_quantiser_ matrix[64]** | o | o | o | o | x | x | o | I | |
| 07 | **load_chroma_non_intra_ quantiser_matrix** | o | o | o | o | x | x | o | I | |
| 08 | **chroma_non_intra_quantiser_ matrix[64]** | o | o | o | o | x | x | o | I | |

**Table E.10 – Picture display extension**

| # | Syntactic elements | Status | | | | | | | Type | Comments |
|---|---|---|---|---|---|---|---|---|---|---|
| | | SIMPLE | MAIN | SNR | SPATIAL | HIGH | 4:2:2 | Multi-view | | |
| 01 | **frame_centre_horizontal_offset** | x | x | x | x | x | x | x | P | Input format related |
| 02 | **frame_centre_vertical_offset** | x | x | x | x | x | x | x | P | Input format related |

**Table E.11 – Picture temporal scalable extension**

| # | Syntactic elements | Status | | | | | | | Type | Comments |
|---|---|---|---|---|---|---|---|---|---|---|
| | | SIMPLE | MAIN | SNR | SPATIAL | HIGH | 4:2:2 | Multi-view | | |
| 01 | **reference_select_code** | o | o | o | o | o | o | x | I | |
| 02 | **forward_temporal_reference** | o | o | o | o | o | o | x | I | |
| 03 | **backward_temporal_reference** | o | o | o | o | o | o | x | I | |

**Table E.12 – Picture spatial scalable extension**

| # | Syntactic elements | Status | | | | | | | Type | Comments |
|---|---|---|---|---|---|---|---|---|---|---|
| | | SIMPLE | MAIN | SNR | SPATIAL | HIGH | 4:2:2 | Multi-view | | |
| 01 | **lower_layer_temporal_reference** | o | o | o | x | x | o | o | I | |
| 02 | **lower_layer_horizontal_offset** | o | o | o | x | x | o | o | D | Input format related |
| 03 | **lower_layer_vertical_offset** | o | o | o | x | x | o | o | D | Input format related |
| 04 | **spatial_temporal_weight_code_ table_index** | o | o | o | x | x | o | o | I | |
| 05 | **lower_layer_progressive_frame** | o | o | o | x | x | o | o | I | |
| 06 | **lower_layer_deinterlaced_field_ select** | o | o | o | x | x | o | o | I | |

**Table E.12-1 – Camera parameters extension**

| # | Syntactic elements | Status | | | | | | | Type | Comments |
|---|---|---|---|---|---|---|---|---|---|---|
| | | SIMPLE | MAIN | SNR | SPATIAL | HIGH | 4:2:2 | Multi-view | | |
| 01 | reserved | o | o | o | o | o | o | x | P | |
| 02 | camera_id | o | o | o | o | o | o | x | P | |
| 03 | marker_bit | o | o | o | o | o | o | x | P | |
| 04 | height_of_image_device | o | o | o | o | o | o | x | P | |
| 05 | marker_bit | o | o | o | o | o | o | x | P | |
| 06 | focal_length | o | o | o | o | o | o | x | P | |
| 07 | marker_bit | o | o | o | o | o | o | x | P | |
| 08 | f_number | o | o | o | o | o | o | x | P | |
| 09 | marker_bit | o | o | o | o | o | o | x | P | |
| 10 | vertical_angle_of_view | o | o | o | o | o | o | x | P | |
| 11 | marker_bit | o | o | o | o | o | o | x | P | |
| 12 | camera_position_x_upper | o | o | o | o | o | o | x | P | |
| 13 | marker_bit | o | o | o | o | o | o | x | P | |
| 14 | camera_position_x_lower | o | o | o | o | o | o | x | P | |
| 15 | marker_bit | o | o | o | o | o | o | x | P | |
| 16 | camera_position_y_upper | o | o | o | o | o | o | x | P | |
| 17 | marker_bit | o | o | o | o | o | o | x | P | |
| 18 | camera_position_y_lower | o | o | o | o | o | o | x | P | |
| 19 | marker_bit | o | o | o | o | o | o | x | P | |
| 20 | camera_position_z_upper | o | o | o | o | o | o | x | P | |
| 21 | marker_bit | o | o | o | o | o | o | x | P | |
| 22 | camera_position_z_lower | o | o | o | o | o | o | x | P | |
| 23 | marker_bit | o | o | o | o | o | o | x | P | |
| 24 | camera_direction_x | o | o | o | o | o | o | x | P | |
| 25 | marker_bit | o | o | o | o | o | o | x | P | |
| 26 | camera_direction_y | o | o | o | o | o | o | x | P | |
| 27 | marker_bit | o | o | o | o | o | o | x | P | |
| 28 | camera_direction_z | o | o | o | o | o | o | x | P | |
| 29 | marker_bit | o | o | o | o | o | o | x | P | |
| 30 | image_plane_vertical_x | o | o | o | o | o | o | x | P | |
| 31 | marker_bit | o | o | o | o | o | o | x | P | |
| 32 | image_plane_vertical_y | o | o | o | o | o | o | x | P | |
| 33 | marker_bit | o | o | o | o | o | o | x | P | |
| 34 | image_plane_vertical_z | o | o | o | o | o | o | x | P | |
| 35 | reserved | o | o | o | o | o | o | x | P | |

**Table E.13 – Slice layer**

| # | Syntactic elements | Status | | | | | | | Type | Comments |
|---|---|---|---|---|---|---|---|---|---|---|
| | | SIMPLE | MAIN | SNR | SPATIAL | HIGH | 4:2:2 | Multi-view | | |
| 01 | **slice_vertical_position_extension** | x | x | x | x | x | x | x | D | Input format related |
| 02 | **priority_breakpoint** | o | o | o | o | o | o | o | I | Only required for data partitioning |
| 03 | **quantiser_scale_code** | x | x | x | x | x | x | x | I | |
| 04 | **slice_extension_flag** | x | x | x | x | x | x | x | I | |
| 05 | **intra_slice** | x | x | x | x | x | x | x | I | Decoder may skip this data |
| 06 | **slice_picture_id_enable** | x | x | x | x | x | x | x | I | Decoder may skip this data |
| 07 | **slice_picture_id** | x | x | x | x | x | x | x | I | Decoder may skip this data |
| 08 | **extra_bit_slice** | x | x | x | x | x | x | x | I | Decoder may skip this data |
| 09 | macroblock() | x | x | x | x | x | x | x | I | |

**Table E.14 – Macroblock layer**

| # | Syntactic elements | Status | | | | | | | Type | Comments |
|---|---|---|---|---|---|---|---|---|---|---|
| | | SIMPLE | MAIN | SNR | SPATIAL | HIGH | 4:2:2 | Multi-view | | |
| 01 | **macroblock_escape** | x | x | x | x | x | x | x | I | |
| 02 | **macroblock_address_increment** | x | x | x | x | x | x | x | I | |
| 03 | macroblock_modes() | x | x | x | x | x | x | x | I | |
| 04 | **quantiser_scale_code** | x | x | x | x | x | x | x | I | |
| 05 | motion_vectors(0) | x | x | x | x | x | x | x | I | Forward motion vector |
| 06 | motion_vectors(1) | o | x | x | x | x | x | x | I | Backward motion vector |
| 07 | coded_block_pattern() | x | x | x | x | x | x | x | I | |
| 08 | block(i) | x | x | x | x | x | x | x | I | |

**Table E.15 – Macroblock modes**

| # | Syntactic elements | Status: SIMPLE | MAIN | SNR | SPATIAL | HIGH | 4:2:2 | Multi-view | Type | Comments |
|---|---|---|---|---|---|---|---|---|---|---|
| 01 | **macroblock_type** | x | x | x | x | x | x | x | I | |
| 02 | **spatial_temporal_weight_code** | o | o | o | x | x | o | o | I | |
| 03 | **frame_motion_type** | x | x | x | x | x | x | x | I | 01: Field-based prediction<br>10: Frame-based prediction<br>11: Dual-prime |
| 04 | **field_motion_type** | x | x | x | x | x | x | x | I | 01: Field-based prediction<br>10: 16 × 8 MC<br>11: Dual-prime |
| 05 | **dct_type** | x | x | x | x | x | x | x | I | |

**Table E.16 – Motion vectors**

| # | Syntactic elements | Status: SIMPLE | MAIN | SNR | SPATIAL | HIGH | 4:2:2 | Multi-view | Type | Comments |
|---|---|---|---|---|---|---|---|---|---|---|
| 01 | **motion_vertical_field_select** | x | x | x | x | x | x | x | I | |
| 02 | motion_vector() | x | x | x | x | x | x | x | I | |

**Table E.17 – Motion vector**

| # | Syntactic elements | Status: SIMPLE | MAIN | SNR | SPATIAL | HIGH | 4:2:2 | Multi-view | Type | Comments |
|---|---|---|---|---|---|---|---|---|---|---|
| 01 | **motion_horizontal_code** | x | x | x | x | x | x | x | I | |
| 02 | **motion_horizontal_r** | x | x | x | x | x | x | x | I | |
| 03 | **dmv_horizontal** | x | x | x | x | x | x | x | I | |
| 04 | **motion_vertical_code** | x | x | x | x | x | x | x | I | |
| 05 | **motion_vertical_r** | x | x | x | x | x | x | x | I | |
| 06 | **dmv_vertical** | x | x | x | x | x | x | x | I | |

**Table E.18 – Coded block pattern**

| # | Syntactic elements | Status: SIMPLE | MAIN | SNR | SPATIAL | HIGH | 4:2:2 | Multi-view | Type | Comments |
|---|---|---|---|---|---|---|---|---|---|---|
| 01 | **coded_block_pattern_420** | x | x | x | x | x | x | x | I | |
| 02 | **coded_block_pattern_1** | o | o | o | o | x | x | o | I | 4:2:2 |
| 03 | **coded_block_pattern_2** | o | o | o | o | o | o | o | I | 4:4:4 |

**Table E.19 – Block layer**

| # | Syntactic elements | SIMPLE | MAIN | SNR | SPATIAL | HIGH | 4:2:2 | Multi-view | Type | Comments |
|---|---|---|---|---|---|---|---|---|---|---|
| 01 | DCT coefficients | x | x | x | x | x | x | x | I | |
| 02 | End of block | x | x | x | x | x | x | x | I | |

## E.2 Permissible layer combinations

Tables E.20 to E.46 illustrate the parameter limits that may be applied in each layer of a bitstream, and the corresponding appropriate profile_and_level_indication that should be used. Each table describes the limits of a single compliance point in the profile / level matrix.

The following notation has been adopted:

<profile abbreviation>@<level abbreviation>

The abbreviations are defined in Table E.20.

**Table E.20 – Abbreviations for profile and level names**

| Profile | <profile abbreviation> | Level | <level abbreviation> |
|---|---|---|---|
| Simple | SP | Low | LL |
| Main | MP | Main | ML |
| SNR Scalable | SNR | High-1440 | H-14 |
| Spatially Scalable | Spt | High | HL |
| High | HP | | |
| Multi-view | MVP | | |
| ISO/IEC 11172-2 constrained parameters | | | ISO/IEC 11172-2 |

**Table E.21 – Simple profile @ Main level**

| No. of layers | layer id | Scalable mode | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|
| 1 | 0 | Base | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |

**Table E.22 – Main profile @ Low level**

| No. of layers | layer id | Scalable mode | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|
| 1 | 0 | Base | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |

**Table E.23 – Main profile @ Main level**

| No. of layers | layer id | Scalable mode | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|
| 1 | 0 | Base | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |

**Table E.24 – Main profile @ High-1440 level**

| No. of layers | layer id | Scalable mode | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|
| 1 | 0 | Base | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | MP@H-14 |

**Table E.25 – Main profile @ High level**

| No. of layers | layer id | Scalable mode | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|
| 1 | 0 | Base | 1920/1088/60 | 62 668 800 | 80 | 9 781 248 | MP@HL |

**Table E.26 – SNR profile @ Low level**

| No. of layers | layer id | Scalable mode | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|
| 2 | 0 | Base | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | SNR | 352/288/30 | 2 534 400 | 4 | 475 136 | SNR@LL |
| 2 | 0 | Base | 352/288/30 | 3 041 280 | 3 | 360 448 | SP@ML |
| | 1 | SNR | 352/288/30 | 3 041 280 | 4 | 475 136 | SNR@LL |
| 2 | 0 | Base | 352/288/30 | 3 041 280 | 3 | 360 448 | MP@LL |
| | 1 | SNR | 352/288/30 | 3 041 280 | 4 | 475 136 | SNR@LL |

**Table E.27 – SNR profile @ Main level**

| No. of layers | layer id | Scalable mode | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|
| 2 | 0 | Base | 720/576/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | SNR | 720/576/30 | 2 534 400 | 15 | 1 835 008 | SNR@ML |
| 2 | 0 | Base | 720/576/30 | 10 368 000 | 10 | 1 212 416 | SP@ML |
| | 1 | SNR | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SNR@ML |
| 2 | 0 | Base | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | SNR | 352/288/30 | 3 041 280 | 15 | 1 835 008 | SNR@ML |
| 2 | 0 | Base | 720/576/30 | 10 368 000 | 10 | 1 212 416 | MP@ML |
| | 1 | SNR | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SNR@ML |

**Table E.28 – Spatial profile @ High-1440 level (Base Layer + SNR)**

| No. of layers | layer id | Scalable mode | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|
| 2 | 0 | Base | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | SNR | 352/288/30 | 2 534 400 | 60 | 7 340 032 | Spt@H-14 |
| 2 | 0 | Base | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | SNR | 720/576/30 | 10 368 000 | 60 | 7 340 032 | Spt@H-14 |
| 2 | 0 | Base | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | SNR | 352/288/30 | 3 041 280 | 60 | 7 340 032 | Spt@H-14 |
| 2 | 0 | Base | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | SNR | 720/576/30 | 10 368 000 | 60 | 7 340 032 | Spt@H-14 |
| 2 | 0 | Base | 1440/1088/60 | 47 001 600 | 40 | 4 882 432 | MP@H-14 |
| | 1 | SNR | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |

**Table E.29 – Spatial profile @ High-1440 level (Base Layer + Spatial)**

| No. of layers | layer id | Scalable mode | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|
| 2 | 0 | Base | 768/576/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 1440/1088/30 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| 2 | 0 | Base | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | Spatial | 1440/1088/30 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| 2 | 0 | Base | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 1440/1088/30 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| 2 | 0 | Base | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | Spatial | 1440/1088/30 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| 2 | 0 | Base | 1440/1088/60 | 47 001 600 | 40 | 4 882 432 | MP@H-14 |
| | 1 | Spatial | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |

**Table E.30 – Spatial profile @ High-1440 level (Base Layer + SNR + Spatial)**

| No. of layers | layer id | Scalable mode | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|
| 3 | 0 | Base | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | SNR | 352/288/30 | 2 534 400 | 4 | 475 136 | SNR@LL |
| | 2 | Spatial | 1440/1088/30 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| 3 | 0 | Base | 352/288/30 | 3 041 280 | 3 | 360 448 | SP@ML |
| | 1 | SNR | 352/288/30 | 3 041 280 | 4 | 475 136 | SNR@LL |
| | 2 | Spatial | 1440/1088/30 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| 3 | 0 | Base | 352/288/30 | 3 041 280 | 3 | 360 448 | MP@LL |
| | 1 | SNR | 352/288/30 | 3 041 280 | 4 | 475 136 | SNR@LL |
| | 2 | Spatial | 1440/1088/30 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| 3 | 0 | Base | 720/576/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | SNR | 720/576/30 | 2 534 400 | 15 | 1 835 008 | SNR@ML |
| | 2 | Spatial | 1440/1088/30 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| 3 | 0 | Base | 720/576/30 | 10 368 000 | 10 | 1 212 416 | SP@ML |
| | 1 | SNR | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SNR@ML |
| | 2 | Spatial | 1440/1088/30 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| 3 | 0 | Base | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | SNR | 352/288/30 | 3 041 280 | 15 | 1 835 008 | SNR@ML |
| | 2 | Spatial | 1440/1088/30 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| 3 | 0 | Base | 720/576/30 | 10 368 000 | 10 | 1 212 416 | MP@ML |
| | 1 | SNR | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SNR@ML |
| | 2 | Spatial | 1440/1088/30 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| 3 | 0 | Base | 1440/1088/60 | 10 368 000 | 15 | 1 835 008 | MP@H-14 |
| | 1 | SNR | 1440/1088/60 | 10 368 000 | 40 | 4 882 432 | Spt@H-14 |
| | 2 | Spatial | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |

**Table E.31 – Spatial profile @ High-1440 level (Base Layer + Spatial + SNR)**

| No. of layers | layer id | Scalable mode | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|
| 3 | 0 | Base | 768/576/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 1440/1088/30 | 47 001 600 | 40 | 4 882 432 | Spt@H-14 |
| | 2 | SNR | 1440/1088/30 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| 3 | 0 | Base | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | Spatial | 1440/1088/30 | 47 001 600 | 40 | 4 882 432 | Spt@H-14 |
| | 2 | SNR | 1440/1088/30 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| 3 | 0 | Base | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 1440/1088/30 | 47 001 600 | 40 | 4 882 432 | Spt@H-14 |
| | 2 | SNR | 1440/1088/30 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| 3 | 0 | Base | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | Spatial | 1440/1088/30 | 47 001 600 | 40 | 4 882 432 | Spt@H-14 |
| | 2 | SNR | 1440/1088/30 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| 3 | 0 | Base | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@H-14 |
| | 1 | Spatial | 1440/1088/60 | 47 001 600 | 40 | 4 882 432 | Spt@H-14 |
| | 2 | SNR | 1440/1088/30 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |

**Table E.32 – High profile @ Main level (Base Layer)**

| No. of layers | layer id | Scalable mode | Chroma Format | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|---|
| 1 | 0 | Base | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | HP@ML |
| 1 | 0 | Base | 4:2:2 | 720/576/30 | 11 059 200 | 20 | 2 441 216 | HP@ML |

**Table E.33 – High profile @ Main level (Base Layer + SNR)**

| No. of layers | layer id | Scalable mode | Chroma Format | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|---|
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | SNR | 4:2:0 | 720/576/30 | 10 368 000 | 20 | 2 441 216 | HP@ML |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | SNR | 4:2:2 | 720/576/30 | 10 368 000 | 20 | 2 441 216 | HP@ML |
| 2 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | SNR | 4:2:0 | 352/288/30 | 3 041 280 | 20 | 2 441 216 | HP@ML |
| 2 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | SNR | 4:2:2 | 352/288/30 | 3 041 280 | 20 | 2 441 216 | HP@ML |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | SNR | 4:2:0 | 720/576/30 | 10 368 000 | 20 | 2 441 216 | HP@ML |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | SNR | 4:2:2 | 720/576/30 | 10 368 000 | 20 | 2 441 216 | HP@ML |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 14 745 600 | 15 | 1 835 008 | HP@ML |
| | 1 | SNR | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | HP@ML |
| 2 | 0 | Base | 4:2:2 | 720/576/30 | 11 059 200 | 15 | 1 835 008 | HP@ML |
| | 1 | SNR | 4:2:2 | 720/576/30 | 11 059 200 | 20 | 2 441 216 | HP@ML |

**Table E.34 – High profile @ Main level (Base Layer + Spatial)**

| No. of layers | layer id | Scalable mode | Chroma Format | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|---|
| 2 | 0 | Base | 4:2:0 | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | HP@ML |
| 2 | 0 | Base | 4:2:0 | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 4:2:2 | 720/576/30 | 11 059 200 | 20 | 2 441 216 | HP@ML |
| 2 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | SP@ML |
| | 1 | Spatial | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | HP@ML |
| 2 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | SP@ML |
| | 1 | Spatial | 4:2:2 | 720/576/30 | 11 059 200 | 20 | 2 441 216 | HP@ML |
| 2 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | HP@ML |
| 2 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 4:2:2 | 720/576/30 | 11 059 200 | 20 | 2 441 216 | HP@ML |

**Table E.35 – High profile @ Main level (Base Layer + SNR + Spatial)**

| No. of layers | layer id | Scalable mode | Chroma Format | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|---|
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 3 | 360 448 | SP@ML |
| | 1 | SNR | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | SNR@LL |
| | 2 | Spatial | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | HP@ML |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 3 | 360 448 | SP@ML |
| | 1 | SNR | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | SNR@LL |
| | 2 | Spatial | 4:2:2 | 720/576/30 | 11 059 200 | 20 | 2 441 216 | HP@ML |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 3 | 360 448 | MP@LL |
| | 1 | SNR | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | SNR@LL |
| | 2 | Spatial | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | HP@ML |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 3 | 360 448 | MP@LL |
| | 1 | SNR | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | SNR@LL |
| | 2 | Spatial | 4:2:2 | 720/576/30 | 11 059 200 | 20 | 2 441 216 | HP@ML |

**Table E.36 – High profile @ Main level (Base Layer + Spatial + SNR)**

| No. of layers | layer id | Scalable mode | Chroma Format | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|---|
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 4:2:0 | 720/576/30 | 14 745 600 | 15 | 1 835 008 | HP@ML |
| | 2 | SNR | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | HP@ML |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 4:2:0 | 720/576/30 | 11 059 200 | 15 | 1 835 008 | HP@ML |
| | 2 | SNR | 4:2:2 | 720/576/30 | 11 059 200 | 20 | 2 441 216 | HP@ML |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 4:2:2 | 720/576/30 | 11 059 200 | 15 | 1 835 008 | HP@ML |
| | 2 | SNR | 4:2:2 | 720/576/30 | 11 059 200 | 20 | 2 441 216 | HP@ML |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | SP@ML |
| | 1 | Spatial | 4:2:0 | 720/576/30 | 14 745 600 | 15 | 1 835 008 | HP@ML |
| | 2 | SNR | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | HP@ML |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | SP@ML |
| | 1 | Spatial | 4:2:0 | 720/576/30 | 11 059 200 | 15 | 1 835 008 | HP@ML |
| | 2 | SNR | 4:2:2 | 720/576/30 | 11 059 200 | 20 | 2 441 216 | HP@ML |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | SP@ML |
| | 1 | Spatial | 4:2:2 | 720/576/30 | 11 059 200 | 15 | 1 835 008 | HP@ML |
| | 2 | SNR | 4:2:2 | 720/576/30 | 11 059 200 | 20 | 2 441 216 | HP@ML |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 4:2:0 | 720/576/30 | 14 745 600 | 15 | 1 835 008 | HP@ML |
| | 2 | SNR | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | HP@ML |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 4:2:0 | 720/576/30 | 11 059 200 | 15 | 1 835 008 | HP@ML |
| | 2 | SNR | 4:2:2 | 720/576/30 | 11 059 200 | 20 | 2 441 216 | HP@ML |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 4:2:2 | 720/576/30 | 11 059 200 | 15 | 1 835 008 | HP@ML |
| | 2 | SNR | 4:2:2 | 720/576/30 | 11 059 200 | 20 | 2 441 216 | HP@ML |

**Table E.37 – High profile @ High-1440 level (Base Layer)**

| No. of layers | layer id | Scalable mode | Chroma Format | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|---|
| 1 | 0 | Base | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| 1 | 0 | Base | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |

**Table E.38 – High profile @ High-1440 level (Base Layer + SNR)**

| No. of layers | layer id | Scalable mode | Chroma Format | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|---|
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | SNR | 4:2:0 | 720/576/30 | 10 368 000 | 80 | 9 781 248 | HP@H-14 |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | SNR | 4:2:2 | 720/576/30 | 10 368 000 | 80 | 9 781 248 | HP@H-14 |
| 2 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | SNR | 4:2:0 | 352/288/30 | 3 041 280 | 80 | 9 781 248 | HP@H-14 |
| 2 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | SNR | 4:2:2 | 352/288/30 | 3 041 280 | 80 | 9 781 248 | HP@H-14 |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | SNR | 4:2:0 | 720/576/30 | 10 368 000 | 80 | 9 781 248 | HP@H-14 |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | SNR | 4:2:2 | 720/576/30 | 10 368 000 | 80 | 9 781 248 | HP@H-14 |
| 2 | 0 | Base | 4:2:0 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | MP@H-14 |
| | 1 | SNR | 4:2:0 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 2 | 0 | Base | 4:2:0 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | MP@H-14 |
| | 1 | SNR | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 1 835 008 | HP@ML |
| | 1 | SNR | 4:2:0 | 720/576/30 | 14 745 600 | 80 | 9 781 248 | HP@H-14 |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 1 835 008 | HP@ML |
| | 1 | SNR | 4:2:2 | 720/576/30 | 14 745 600 | 80 | 9 781 248 | HP@H-14 |
| 2 | 0 | Base | 4:2:2 | 720/576/30 | 11 059 200 | 20 | 1 835 008 | HP@ML |
| | 1 | SNR | 4:2:2 | 720/576/30 | 11 059 200 | 80 | 9 781 248 | HP@H-14 |
| 2 | 0 | Base | 4:2:0 | 1440/1088/60 | 62 668 800 | 60 | 7 340 032 | HP@H-14 |
| | 1 | SNR | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| 2 | 0 | Base | 4:2:0 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | HP@H-14 |
| | 1 | SNR | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 2 | 0 | Base | 4:2:2 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | HP@H-14 |
| | 1 | SNR | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |

**Table E.39 – High profile @ High-1440 level (Base Layer + Spatial)**

| No. of layers | layer id | Scalable mode | Chroma Format | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|---|
| 2 | 0 | Base | 4:2:0 | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| 2 | 0 | Base | 4:2:0 | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | Spatial | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 2 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| 2 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | Spatial | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | MP@H-14 |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | MP@H-14 |
| | 1 | Spatial | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | HP@ML |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | HP@ML |
| | 1 | Spatial | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 2 | 0 | Base | 4:2:2 | 720/576/30 | 11 059 200 | 20 | 2 441 216 | HP@ML |
| | 1 | Spatial | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |

**Table E.40 – High profile @ High-1440 level (Base Layer + SNR + Spatial)**

| No. of layers | layer id | Scalable mode | Chroma Format | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|---|
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 3 | 360 448 | SP@ML |
| | 1 | SNR | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | SNR@LL |
| | 2 | Spatial | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 3 | 360 448 | SP@ML |
| | 1 | SNR | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | SNR@LL |
| | 2 | Spatial | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 10 | 1 212 416 | SP@ML |
| | 1 | SNR | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SNR@ML |
| | 2 | Spatial | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 10 | 1 212 416 | SP@ML |
| | 1 | SNR | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SNR@ML |
| | 2 | Spatial | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 3 | 360 448 | MP@LL |
| | 1 | SNR | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | SNR@LL |
| | 2 | Spatial | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 3 | 360 448 | MP@LL |
| | 1 | SNR | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | SNR@LL |
| | 2 | Spatial | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 10 | 1 212 416 | MP@ML |
| | 1 | SNR | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SNR@ML |
| | 2 | Spatial | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 10 | 1 212 416 | MP@ML |
| | 1 | SNR | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SNR@ML |
| | 2 | Spatial | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | SNR | 4:2:2 | 720/576/30 | 10 368 000 | 20 | 2 441 216 | HP@ML |
| | 2 | Spatial | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 14 745 600 | 15 | 1 835 008 | HP@ML |
| | 1 | SNR | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | HP@ML |
| | 2 | Spatial | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 11 059 200 | 15 | 1 835 008 | HP@ML |
| | 1 | SNR | 4:2:2 | 720/576/30 | 11 059 200 | 20 | 2 441 216 | HP@ML |
| | 2 | Spatial | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:2 | 720/576/30 | 11 059 200 | 15 | 1 835 008 | HP@ML |
| | 1 | SNR | 4:2:2 | 720/576/30 | 11 059 200 | 20 | 2 441 216 | HP@ML |
| | 2 | Spatial | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |

**Table E.41 – High profile @ High-1440 level (Base Layer + Spatial + SNR)**

| No. of layers | layer id | Scalable mode | Chroma Format | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|---|
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| | 2 | SNR | 4:2:0 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| | 2 | SNR | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 62 668 800 | 60 | 7 340 032 | HP@H-14 |
| | 2 | SNR | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 4:2:2 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | HP@H-14 |
| | 2 | SNR | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| | 2 | SNR | 4:2:0 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| | 2 | SNR | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 62 668 800 | 60 | 7 340 032 | HP@H-14 |
| | 2 | SNR | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | Spatial | 4:2:2 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | HP@H-14 |
| | 2 | SNR | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| | 2 | SNR | 4:2:0 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| | 2 | SNR | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 62 668 800 | 60 | 7 340 032 | HP@H-14 |
| | 2 | SNR | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 4:2:2 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | HP@H-14 |
| | 2 | SNR | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| | 2 | SNR | 4:2:0 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |

**Table E.41 – High profile @ High-1440 level (Base Layer + Spatial + SNR)** *(concluded)*

| No. of layers | layer id | Scalable mode | Chroma Format | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|---|
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| | 2 | SNR | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 62 668 800 | 60 | 7 340 032 | HP@H-14 |
| | 2 | SNR | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | Spatial | 4:2:2 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | HP@H-14 |
| | 2 | SNR | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | HP@ML |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 62 668 800 | 60 | 7 340 032 | HP@H-14 |
| | 2 | SNR | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | HP@ML |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | HP@H-14 |
| | 2 | SNR | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | HP@ML |
| | 1 | Spatial | 4:2:2 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | HP@H-14 |
| | 2 | SNR | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| 3 | 0 | Base | 4:2:2 | 720/576/30 | 11 059 200 | 20 | 2 441 216 | HP@ML |
| | 1 | Spatial | 4:2:2 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | HP@H-14 |
| | 2 | SNR | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |

**Table E.42 – High profile @ High level (Base Layer)**

| No. of layers | layer id | Scalable mode | Chroma Format | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|---|
| 1 | 0 | Base | 4:2:0 | 1920/1088/60 | 83 558 400 | 100 | 12 222 464 | HP@HL |
| 1 | 0 | Base | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |

**Table E.43 – High profile @ High level (Base Layer + SNR)**

| No. of layers | layer id | Scalable mode | Chroma Format | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|---|
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | SNR | 4:2:0 | 720/576/30 | 10 368 000 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | SNR | 4:2:2 | 720/576/30 | 10 368 000 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | SNR | 4:2:0 | 352/288/30 | 3 041 280 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | SNR | 4:2:2 | 352/288/30 | 3 041 280 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | SNR | 4:2:0 | 720/576/30 | 10 368 000 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | SNR | 4:2:2 | 720/576/30 | 10 368 000 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | MP@H-14 |
| | 1 | SNR | 4:2:0 | 1440/1088/60 | 47 001 600 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | MP@H-14 |
| | 1 | SNR | 4:2:2 | 1440/1088/60 | 47 001 600 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 1920/1088/60 | 62 668 800 | 80 | 9 781 248 | MP@HL |
| | 1 | SNR | 4:2:0 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 1920/1088/60 | 62 668 800 | 80 | 9 781 248 | MP@HL |
| | 1 | SNR | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 1 835 008 | HP@ML |
| | 1 | SNR | 4:2:0 | 720/576/30 | 14 745 600 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 1 835 008 | HP@ML |
| | 1 | SNR | 4:2:2 | 720/576/30 | 14 745 600 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:2 | 720/576/30 | 11 059 200 | 20 | 1 835 008 | HP@ML |
| | 1 | SNR | 4:2:2 | 720/576/30 | 11 059 200 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| | 1 | SNR | 4:2:0 | 1440/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| | 1 | SNR | 4:2:2 | 1440/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| | 1 | SNR | 4:2:2 | 1440/1088/60 | 47 001 600 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 1920/1088/60 | 83 558 400 | 80 | 9 781 248 | HP@HL |
| | 1 | SNR | 4:2:0 | 1920/1088/60 | 83 558 400 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 1920/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@HL |
| | 1 | SNR | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:2 | 1920/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@HL |
| | 1 | SNR | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |

**Table E.44 – High profile @ High level (Base Layer + Spatial)**

| No. of layers | layer id | Scalable mode | Chroma Format | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|---|
| 2 | 0 | Base | 4:2:0 | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 4:2:0 | 1920/1088/60 | 83 558 400 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | Spatial | 4:2:0 | 1920/1088/60 | 83 558 400 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 4:2:0 | 1920/1088/60 | 83 558 400 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | Spatial | 4:2:0 | 1920/1088/60 | 83 558 400 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 960/576/30 | 19 660 800 | 25 | 3 047 424 | MP@H-14 |
| | 1 | Spatial | 4:2:0 | 1920/1088/60 | 83 558 400 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 960/576/30 | 19 660 800 | 25 | 3 047 424 | MP@H-14 |
| | 1 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | HP@ML |
| | 1 | Spatial | 4:2:0 | 1920/1088/60 | 83 558 400 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | HP@ML |
| | 1 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:2 | 720/576/30 | 11 059 200 | 20 | 2 441 216 | HP@ML |
| | 1 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 960/576/30 | 19 660 800 | 25 | 3 047 424 | HP@H-14 |
| | 1 | Spatial | 4:2:0 | 1920/1088/60 | 83 558 400 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:0 | 960/576/30 | 19 660 800 | 25 | 3 047 424 | HP@H-14 |
| | 1 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 2 | 0 | Base | 4:2:2 | 960/576/30 | 14 745 600 | 25 | 3 047 424 | HP@H-14 |
| | 1 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |

**Table E.45 – High profile @ High level (Base Layer + SNR + Spatial)**

| No. of layers | layer id | Scalable mode | Chroma Format | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|---|
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 3 | 360 448 | SP@ML |
| | 1 | SNR | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | SNR@LL |
| | 2 | Spatial | 4:2:0 | 1920/1088/60 | 83 558 400 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 3 | 360 448 | SP@ML |
| | 1 | SNR | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | SNR@LL |
| | 2 | Spatial | 4:2:2 | 1920/1088/60 | 83 558 400 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 10 | 1 212 416 | SP@ML |
| | 1 | SNR | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SNR@ML |
| | 2 | Spatial | 4:2:0 | 1920/1088/60 | 83 558 400 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 10 | 1 212 416 | SP@ML |
| | 1 | SNR | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SNR@ML |
| | 2 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | SNR | 4:2:2 | 720/576/30 | 10 368 000 | 20 | 2 441 216 | HP@ML |
| | 2 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 3 | 360 448 | MP@LL |
| | 1 | SNR | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | SNR@LL |
| | 2 | Spatial | 4:2:0 | 1920/1088/60 | 83 558 400 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 3 | 360 448 | MP@LL |
| | 1 | SNR | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | SNR@LL |
| | 2 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 10 | 1 212 416 | MP@ML |
| | 1 | SNR | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SNR@ML |
| | 2 | Spatial | 4:2:0 | 1920/1088/60 | 83 558 400 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 10 | 1 212 416 | MP@ML |
| | 1 | SNR | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SNR@ML |
| | 2 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | SNR | 4:2:2 | 720/576/30 | 10 368 000 | 20 | 2 441 216 | HP@ML |
| | 2 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 960/576/30 | 19 660 800 | 20 | 2 441 216 | MP@H-14 |
| | 1 | SNR | 4:2:0 | 960/576/30 | 19 660 800 | 25 | 3 047 424 | Spt@H-14 |
| | 2 | Spatial | 4:2:0 | 1920/1088/60 | 83 558 400 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 960/576/30 | 19 660 800 | 20 | 2 441 216 | MP@H-14 |
| | 1 | SNR | 4:2:0 | 960/576/30 | 19 660 800 | 25 | 3 047 424 | Spt@H-14 |
| | 2 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 960/576/30 | 14 745 600 | 20 | 2 441 216 | MP@H-14 |
| | 1 | SNR | 4:2:2 | 960/576/30 | 14 745 600 | 25 | 3 047 424 | HP@H-14 |
| | 2 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |

**Table E.45 – High profile @ High level (Base Layer + SNR + Spatial)** *(concluded)*

| No. of layers | layer id | Scalable mode | Chroma Format | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|---|
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 14 745 600 | 15 | 1 835 008 | HP@ML |
| | 1 | SNR | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | HP@ML |
| | 2 | Spatial | 4:2:0 | 1920/1088/60 | 83 558 400 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 14 745 600 | 15 | 1 835 008 | HP@ML |
| | 1 | SNR | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | HP@ML |
| | 2 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 11 059 200 | 15 | 1 835 008 | HP@ML |
| | 1 | SNR | 4:2:2 | 720/576/30 | 11 059 200 | 20 | 2 441 216 | HP@ML |
| | 2 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:2 | 720/576/30 | 11 059 200 | 15 | 1 835 008 | HP@ML |
| | 1 | SNR | 4:2:2 | 720/576/30 | 11 059 200 | 20 | 2 441 216 | HP@ML |
| | 2 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 960/576/30 | 19 660 800 | 20 | 2 441 216 | HP@H-14 |
| | 1 | SNR | 4:2:0 | 960/576/30 | 19 660 800 | 25 | 3 047 424 | HP@H-14 |
| | 2 | Spatial | 4:2:0 | 1920/1088/60 | 83 558 400 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 960/576/30 | 19 660 800 | 20 | 2 441 216 | HP@H-14 |
| | 1 | SNR | 4:2:0 | 960/576/30 | 19 660 800 | 25 | 3 047 424 | HP@H-14 |
| | 2 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 960/576/30 | 14 745 600 | 20 | 2 441 216 | HP@H-14 |
| | 1 | SNR | 4:2:2 | 960/576/30 | 14 745 600 | 25 | 3 047 424 | HP@H-14 |
| | 2 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:2 | 960/576/30 | 14 745 600 | 20 | 2 441 216 | HP@H-14 |
| | 1 | SNR | 4:2:2 | 960/576/30 | 14 745 600 | 25 | 3 047 424 | HP@H-14 |
| | 2 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |

**Table E.46 – High profile @ High level (Base Layer + Spatial + SNR)**

| No. of layers | layer id | Scalable mode | Chroma Format | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|---|
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| | 2 | SNR | 4:2:0 | 1440/1088/60 | 47 001 600 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| | 2 | SNR | 4:2:2 | 1440/1088/60 | 47 001 600 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | HP@ML |
| | 2 | SNR | 4:2:0 | 720/576/30 | 14 745 600 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | HP@ML |
| | 2 | SNR | 4:2:2 | 720/576/30 | 14 745 600 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 4:2:2 | 720/576/30 | 11 059 200 | 20 | 2 441 216 | HP@ML |
| | 2 | SNR | 4:2:2 | 720/576/30 | 11 059 200 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| | 2 | SNR | 4:2:0 | 1440/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| | 2 | SNR | 4:2:2 | 1440/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| | 2 | SNR | 4:2:2 | 1440/1088/60 | 47 001 600 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 4:2:0 | 1920/1088/60 | 83 558 400 | 80 | 9 781 248 | HP@HL |
| | 2 | SNR | 4:2:0 | 1920/1088/60 | 83 558 400 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 4:2:0 | 1920/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@HL |
| | 2 | SNR | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 2 534 400 | 1.856 | 327 680 | ISO 11172 |
| | 1 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@HL |
| | 2 | SNR | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| | 2 | SNR | 4:2:0 | 1440/1088/60 | 47 001 600 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| | 2 | SNR | 4:2:2 | 1440/1088/60 | 47 001 600 | 100 | 12 222 464 | HP@HL |

**Table E.46 – High profile @ High level (Base Layer + Spatial + SNR)** *(continued)*

| No. of layers | layer id | Scalable mode | Chroma Format | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|---|
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| | 2 | SNR | 4:2:0 | 1440/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| | 2 | SNR | 4:2:2 | 1440/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | Spatial | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| | 2 | SNR | 4:2:2 | 1440/1088/60 | 47 001 600 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | Spatial | 4:2:0 | 1920/1088/60 | 83 558 400 | 80 | 9 781 248 | HP@HL |
| | 2 | SNR | 4:2:0 | 1920/1088/60 | 83 558 400 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | Spatial | 4:2:0 | 1920/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@HL |
| | 2 | SNR | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@HL |
| | 2 | SNR | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| | 2 | SNR | 4:2:0 | 1440/1088/60 | 47 001 600 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| | 2 | SNR | 4:2:2 | 1440/1088/60 | 47 001 600 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | HP@ML |
| | 2 | SNR | 4:2:0 | 720/576/30 | 14 745 600 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 4:2:0 | 720/576/30 | 14 745 600 | 20 | 2 441 216 | HP@ML |
| | 2 | SNR | 4:2:2 | 720/576/30 | 14 745 600 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 4:2:2 | 720/576/30 | 11 059 200 | 20 | 2 441 216 | HP@ML |
| | 2 | SNR | 4:2:2 | 720/576/30 | 11 059 200 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| | 2 | SNR | 4:2:0 | 1440/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| | 2 | SNR | 4:2:2 | 1440/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |

**Table E.46 – High profile @ High level (Base Layer + Spatial + SNR)** *(continued)*

| No. of layers | layer id | Scalable mode | Chroma Format | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|---|
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| | 2 | SNR | 4:2:2 | 1440/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 4:2:0 | 1920/1088/60 | 83 558 400 | 80 | 9 781 248 | HP@HL |
| | 2 | SNR | 4:2:0 | 1920/1088/60 | 83 558 400 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 4:2:0 | 1920/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@HL |
| | 2 | SNR | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@HL |
| | 2 | SNR | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| | 2 | SNR | 4:2:0 | 1440/1088/60 | 47 001 600 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | Spt@H-14 |
| | 2 | SNR | 4:2:2 | 1440/1088/60 | 47 001 600 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| | 2 | SNR | 4:2:0 | 1440/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | Spatial | 4:2:0 | 1440/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@H-14 |
| | 2 | SNR | 4:2:2 | 1440/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | Spatial | 4:2:2 | 1440/1088/60 | 47 001 600 | 80 | 9 781 248 | HP@H-14 |
| | 2 | SNR | 4:2:2 | 1440/1088/60 | 47 001 600 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | Spatial | 4:2:0 | 1920/1088/60 | 83 558 400 | 80 | 9 781 248 | HP@HL |
| | 2 | SNR | 4:2:0 | 1920/1088/60 | 83 558 400 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | Spatial | 4:2:0 | 1920/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@HL |
| | 2 | SNR | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@HL |
| | 2 | SNR | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 960/576/30 | 19 660 800 | 25 | 3 047 424 | HP@H-14 |
| | 1 | Spatial | 4:2:0 | 1920/1088/60 | 83 558 400 | 80 | 9 781 248 | HP@HL |
| | 2 | SNR | 4:2:0 | 1920/1088/60 | 83 558 400 | 100 | 12 222 464 | HP@HL |

**Table E.46 – High profile @ High level (Base Layer + Spatial + SNR)** *(concluded)*

| No. of layers | layer id | Scalable mode | Chroma Format | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|---|
| 3 | 0 | Base | 4:2:0 | 960/576/30 | 19 660 800 | 25 | 3 047 424 | HP@H-14 |
| | 1 | Spatial | 4:2:0 | 1920/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@HL |
| | 2 | SNR | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:0 | 960/576/30 | 19 660 800 | 25 | 3 047 424 | HP@H-14 |
| | 1 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@HL |
| | 2 | SNR | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |
| 3 | 0 | Base | 4:2:2 | 960/576/30 | 14 745 600 | 25 | 3 047 424 | HP@H-14 |
| | 1 | Spatial | 4:2:2 | 1920/1088/60 | 62 668 800 | 80 | 9 781 248 | HP@HL |
| | 2 | SNR | 4:2:2 | 1920/1088/60 | 62 668 800 | 100 | 12 222 464 | HP@HL |

**Table E.47 – Multi-view profile @ Low level**

| No. of layers | layer id | Scalable mode | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|
| 2 | 0 | Base | 352/288/30 | 3 041 280 | 4 | 475 136 | MP@LL |
| | 1 | Temporal | 352/288/30 | 3 041 280 | 8 | 950 272 | MVP@LL |

**Table E.48 – Multi-view profile @ Main level**

| No. of layers | layer id | Scalable mode | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|
| 2 | 0 | Base | 720/576/30 | 10 368 000 | 15 | 1 835 008 | SP@ML |
| | 1 | Temporal | 720/576/30 | 10 368 000 | 25 | 3 047 424 | MVP@ML |
| 2 | 0 | Base | 720/576/30 | 10 368 000 | 15 | 1 835 008 | MP@ML |
| | 1 | Temporal | 720/576/30 | 10 368 000 | 25 | 3 047 424 | MVP@ML |

**Table E.49 – Multi-view profile @ High-1440 level**

| No. of layers | layer id | Scalable mode | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|
| 2 | 0 | Base | 1440/1088/60 | 47 001 600 | 60 | 7 340 032 | MP@H-14 |
| | 1 | Temporal | 1440/1088/60 | 47 001 600 | 100 | 12 222 464 | MVP@H-14 |

**Table E.50 – Multi-view profile @ High level**

| No. of layers | layer id | Scalable mode | Maximum sample density (H/V/F) | Maximum sample rate | Maximum total bit rate /1000000 | Maximum total VBV buffer | Profile and level indication |
|---|---|---|---|---|---|---|---|
| 2 | 0 | Base | 1920/1088/60 | 62 668 800 | 80 | 9 781 248 | MP@HL |
| | 1 | Temporal | 1920/1088/60 | 62 668 800 | 130 | 15 898 480 | MVP@HL |

# Annex F

## Bibliography

(This annex does not form an integral part of this Recommendation | International Standard)


[1] NETRAVALI (A.N.) and HASKELL (B.G.): Digital Pictures, representation and compression, *Plenum Press*, 1988.

[2] LEGALL (D.): MPEG: A Video Compression Standard for Multimedia Applications, *Trans. ACM*, April 1991.

[3] LOEFFLER (C.), LIGTENBERG (A.), MOSCHYTZ (G.S.): Practical fast 1-D DCT algorithms with 11 multiplications, *Proceedings IEEE ICASSP-89*, Vol. 2, pp. 988-991, February 1989.

[4] See the Normative Reference for ITU-R Recommendation BT.601.

[5] See the Normative Reference for IEC Publication 60461.

[6] See the Normative Reference for Recommendation H.261.

[7] See the Normative reference for IEEE Standard Specification P1180-1990.

[8] CCITT Recommendation T.81 (1992) | ISO/IEC 10918-1:1994 (Previously JPEG).

[9] VISCITO (E.) and GONZALES (C.): A Video Compression Algorithm with Adaptive Bit Allocation and Quantization, *Proc. SPIE Visual Communications and Image Proc 91,* Vol. 1605 205, Boston, MA, 10-15 November 1991.

[10] PURI (A.) and ARAVIND (R.): Motion Compensated Video Coding with Adaptive Perceptual Quantization, *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 1, pp. 351, December 1991.

[11] GONZALES (C.) and VISCITO (E.): Flexibly scalable digital video coding, *Image Communications*, Vol. 5, Nos. 1-2, February 1993.

[12] JOHNSON (A.W.), SIKORA (T.) and TAN (T.K.): Filters for Drift Reduction in Frequency Scalable Video Coding Schemes, *Transmitted for publication to Electronic Letters.*

[13] MOKRY (R.) and ANASTASSIOU (D.): Minimal Error Drift in Frequency Scalability for Motion-Compensated DCT Coding, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 4, No. 4, pp. 392-407, August 1994.

[14] NGAN (K.N.), ARNOLD (J.), SIKORA (T.), TAN (T.K.) and JOHNSON (A.W.): Frequency Scalability Experiments for MPEG-2 Standard, *Asia-Pacific Conference on Communications*, Korea, August 1993.

[15] SIKORA (T.), TAN (T.K.) and NGAN (K.N.): A Performance Comparison of Frequency Domain Pyramid Scalable Coding Schemes Within the MPEG Framework, *Proc. PCS, Picture Coding Symposium*, pp. 16.1-16.2, March 1993, Lausanne, Switzerland.

[16] IWAHASHI (M.): Motion Compensation Technique for 2:1 Scaled-down Moving Pictures. 8-14, *Picture Coding Symposium 93*.

[17] SIKORA, (T.) and PANG, (K.): Experiments with Optimal Block-Overlapping Filters for Cell Loss Concealment in Packet Video, *Proc. IEEE Visual Signal Processing and Communications Workshop*, pp. 247-250, 21-22 September 1993, Melbourne.

[18] PURI (A.): Video Coding Using the MPEG-2 Compression Standard, *Proc. SPIE Visual Communications and Image Proc. 93*, Boston, MA, November 1993.

[19] PURI (A.) and WONG (A.): Spatial Domain Resolution Scalable Video Coding, *Proc. SPIE Visual Communications and Image Proc. 93*, Boston, MA, November 1993.

# Annex G

## Registration Procedure

(This annex does not form an integral part of this Recommendation | International Standard)

### G.1 Procedure for the request of a Registered Identifier (RID)

Requesters of an RID shall apply to the Registration Authority. Registration forms shall be available from the Registration Authority. Information which the requester shall provide is given in G.3. Companies and organizations are eligible to apply.

### G.2 Responsibilities of the Registration Authority

The primary responsibilities of the Registration Authority administrating the registration of copyright_identifiers is outlined in this subclause; certain other responsibilities may be found in the JTC 1 Directives. The Registration Authority shall:

a) implement a registration procedure for application for a unique RID in accordance with Annex H of the JTC 1 Directives;

b) receive and process the applications for allocation of the work type code identifier from Copyright Registration Authority;

c) ascertain which applications received are in accordance with this registration procedure, and to inform the requester within 30 days of receipt of the application of their assigned RID;

d) inform application providers whose request is denied in writing within 30 days of receipt of the application, and also inform the requesting party of the appeals process;

e) maintain an accurate register of the allocated RID. Revisions to the contact information and technical specifications shall be accepted and maintained by the Registration Authority;

f) make the contents of this register available upon request to any interested party;

g) maintain a database of RID request forms, granted and denied. Parties seeking technical information on the format of private data which has a copyright_identifier shall have access to such information which is part of the database maintained by the Registration Authority;

h) report its activities to JTC 1, the ITTF, and the JTC 1/SC 29 Secretariat, or their respective assignees, annually on a schedule mutually agreed upon.

#### G.2.1 Contact information of the Registration Authority

Organization Name:

Address:

Telephone:

Fax:

### G.3 Responsibilities of parties requesting an RID

The party requesting an RID for the purpose of copyright identification shall:

a) apply using the form and procedures supplied by the Registration Authority;

b) provide contact information describing how a complete description of the copyright organization can be obtained on a non-discriminatory basis;

c) include technical details of the syntax and semantics of the data format used to describe the audiovisual works or other copyrighted works within the additional_copyright_info field. Once registered, the syntax used for the additional copyright information shall not change;

d) agree to institute the intended use of the granted copyright_identifier within a reasonable time-frame;

e) maintain a permanent record of the application form and the notification received from the Registration Authority of each granted copyright_identifier.

## G.4 Appeal procedure for denied applications

The Registration Management Group is formed to have jurisdiction over appeals relating to a denied request for an RID. The RMG shall have a membership who are nominated by P and L members of the ISO technical body responsible for this Recommendation | International Standard. It shall have a convenor and secretariat nominated from its members. The Registration Authority is entitled to nominate one non-voting observing member.

The responsibilities of the RMG shall be:

a) to review and act on all appeals within a reasonable time-frame;

b) to inform, in writing, organizations which make an appeal for reconsideration of its petition of the RMG's disposition of the matter;

c) to review the annual report of the Registration Authority summary of activities;

d) to supply ISO member bodies with information concerning the scope of operation of the Registration Authority.

# Annex H

## Registration Application Form

(This annex does not form an integral part of this Recommendation | International Standard)

### H.1 Contact information of organization requesting a Registered Identifier (RID)

Organization Name:

Address:

Telephone:

Fax:

E-mail:

### H.2 Statement of an intention to apply the assigned RID

RID application domain: using guidelines to be provided by the Registration Authority.

### H.3 Date of intended implementation of the RID

### H.4 Authorized representative

Name:

Title:

Address:

Signature _____

### H.5 For official use only of the Registration Authority

Registration Rejected _____

Reason for rejection of the application:

Registration Granted _____ Registration Value _____

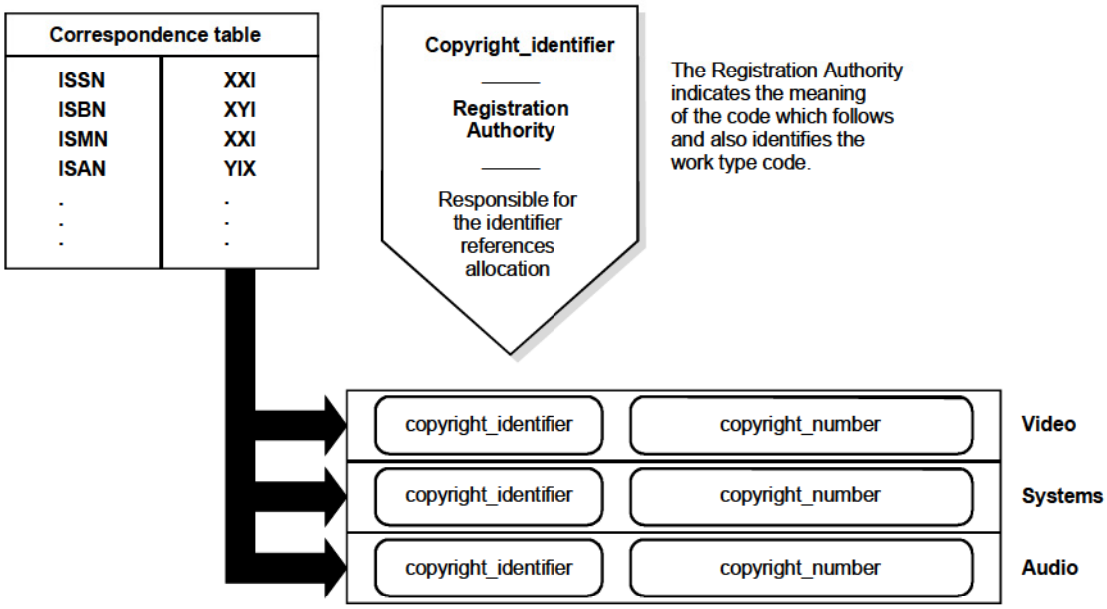
Attachment 1 – Attachment of technical details of the registered data format.

Attachment 2 – Attachment of notification of appeal procedure for rejected applications.

# Annex I

(This annex does not form an integral part of this Recommendation | International Standard)

**Registration Authority**
**Diagram of administration structure**




**Examples**

| copyright_identifier | copyright_number |
|---|---|
| I.S.B.N. (for books) | 2-11- 0725 575 (ISBN Number) |
| I.S.A.N. (for audiovisual works) | 1234567890123456 (ISAN Number) |

All the copyright_identifiers are registered by the Registration Authority, uniquely for copyright_numbers standardized by ISO. Each organization which allocates copyright_numbers, requests a specific copyright_identifier from the Registration Authority, e.g. Staatsbibliothek Preussischer Kulturbesitz, designated by ISO to manage I.S.B.N., asks for a specific copyright_identifier from the R.A. for book numbering.

# Annex J

# 4:2:2 Profile test results

(This annex does not form an integral part of this Recommendation | International Standard)

## J.1 Introduction

This annex provides guidance to users regarding the applicability of the 4:2:2 Profile at Main Level to applications which may require:

- higher quality than Main Profile at Main Level;
- better chroma resolution than Main Profile at Main Level;
- post processing after compression and decompression;
- multiple generations of compression and decompression;
- short Group of Pictures (GOP) for editability;
- capability to pass all active video;
- capability to pass vertical blanking interval information.

It should be noted that application of this Profile is an area of ongoing progress. Results presented here reflect varying degrees of algorithm refinement, so further improvement can be expected.

### J.1.1 Test sequences

The test sequences were generated using computer simulation of the ITU-T Rec. H.262 | ISO/IEC 13818-2 compression and decompression. For 525/60, the test material included:

- Gwen;
- Trailblazers;
- Mobile and Calendar;
- Dissolve.

For 625/50, the test material included:

- Balls of Wool;
- Cactus and Comb;
- Basketball;
- Wall;
- Renata and Butterfly;
- Mobile and Calendar.

"Gwen" is a chroma key test sequence with a woman in the foreground keyed over a forest scene in the background. "Gwen" is a difficult sequence to chroma key but an easy sequence to compress. Both "Cactus and Comb" and "Balls of Wool" are chroma key sequences which were used with a coloured background. "Trailblazers" is a rapid motion basketball sequence shot with an unshuttered CCD camera. "Basketball" is also a rapid motion sports sequence. Both are typical program material and moderately difficult to compress. "Wall" consists of a woman standing in front of a wall made of many small stones. "Renata" consists of a woman in front of a complex background with a dissolve to a complex image of butterflies. "Mobile and Calendar" is a particularly difficult compression test sequence with saturated colours and complex motion. "Dissolve" consists of two segments of "Mobile and Calendar" with a one second fade between the two segments and is also difficult to compress.

Test sequences were supplied by:

- ITU-R;
- Portland Trailblazers;
- SMPTE;
- Tektronix.

### J.1.2 Test procedures

MPEG has conducted experiments to verify the performance of the 4:2:2 Profile. The results of those experiments are presented here. There are separate tests for 525/60 and 625/50. The 525/60 tests explore a broad range of data rates and GOP structures, while the 625/50 tests include more variety of test material but less combinations of data rate, GOP structure, and number of generations. The parameters chosen for the experiments are for example only, and do not cover the entire range of allowed parameter values. The examples are not intended as specific recommendations. Each application should use the combination of parameters that is most appropriate, depending on its requirements for quality, editability, and cost.

The tests include both a single generation and eight generations of cascaded compression and decompression. For the eight generation tests, separate tests were done with no shifts, with two spatial shifts, and with two temporal shifts. Spatial shifting means that the picture was shifted horizontally and vertically by two pixels and two spatial lines between the first and second generations and then back between the fifth and sixth generations. Spatial shifting represents the effects of picture repositioning which might occur in a DVE. Temporal shifting means that the GOP structure was shifted one frame between the first and second generations and again between the fifth and sixth generations. Temporal shifting represents the effect of multiple generations which have different GOP alignment.

Chroma key experiments were done by processing the foreground with blue screen through compression and decompression. After decompression the component digital signal was chroma keyed to add the background. The background image was not compressed.

Mixed environment tests for 525/60 used ITU-T Rec. H.262 | ISO/IEC 13818-2 4:2:2 compression and decompression cascaded with a compressed digital VTR using 2:1 intra-field compression. The tests used a total of eight generations of compression. The four odd number generations were MPEG and the four even number generations were compressed digital VTR. There were no shifts between generations.

Mixed environment tests for 625/50 used only MPEG compression. The tests used a total of three generations of compression. The first and third generations were ITU-T Rec. H.262 | ISO/IEC 13818-2 4:2:2 compression with IBBP-GOP structure at 20 Mbits/s, while the second generation was ITU-T Rec. H.262 | ISO/IEC 13818-2 4:2:2 compression with I-only GOP structure at 50 Mbits/s. A temporal shift of one frame was included between the second and third generations.

Compression and decompression processing were contributed by:

- CCETT;
- FTZ;
- IRT;
- JVC;
- Sony;
- Technical University of Braunschweig/BTS;
- Tektronix.

Editing and duplication of test tapes were contributed by:

- RAI;
- Tektronix.

### J.1.3 Subjective assessment

The subjective assessment used the DSCQS method described in ITU-R Rec. BT.500-6. Both expert and non-expert viewing sessions were conducted at a number of sites around the world. All of the expert viewing results were combined, and all of the non-expert viewing results were combined. Both expert and non-expert results are presented here. Only subjective test results are presented, as signal to noise ratio is not regarded as a reliable measure of picture quality in these cases.

Expert subjective assessment viewing sessions were conducted by:

- NHK;
- SMPTE.

Non-expert subjective assessment viewing sessions were conducted by:

- CCETT;
- JVC/MPT/NHK/NTV;
- RAI;
- Technical University of Braunschweig/BTS.

### J.1.4 Test results

Test results are presented in the following order (see Tables J.1 and J.2):

- 525/60 Homogeneous Environment;
- 525/60 Non-Homogeneous Environment;
- 625/50 Homogeneous Environment;
- 625/50 Non-Homogeneous Environment.

The tables of test results are organized with higher data rates presented first and lower data rates presented last. Within a given bit rate, results are organized by GOP structure, number of generations, and type of shifting. The mean and confidence interval are given for each test sequence.

These tests used the continuous quality scale specified in ITU-R Rec. BT.500-6. The subjective assessments were done on a continuous 0 to 100 scale. The mean differences between original and compressed sequence ratings were calculated, on a 0 to 100 scale, with differences inferior or equal to 0 representing no degradation through compression and 100 being the worst possible rating.

Hereinafter the average of the differences between original and compressed sequence ratings, calculated over the subjects, will be referred to as diff-grade.

The results presented here are based on the following quality definitions:

- **transparency**: diff-grade for all test sequences does not exceed 12% of the scale;
- **near transparency**: the diff-grade of the 25% of the test sequences is between 12% and 18%, while all the other diff-grades do not exceed 12% of the scale;
- **good quality in most of the material**: the mean diff-grades calculated over the test sequences do not exceed 18%, while the diff-grades of 25% of the test sequences exceed 18% of the scale;
- **difficulties in some materials**: all other cases.

**Table J.1 – Subjective test results for the 525/60 system**

| Compression parameters | | Viewer ratings | |
|---|---|---|---|
| **525/60 50 Mbits/s** | | **Expert viewers** | **Non-expert viewers** |
| GOP = I | 1 generation | Transparency | Transparency |
| | 8 generations, No shifts | Good quality in most test materials | Transparency |
| | 8 generations, 2 spatial shifts | Good quality in most test materials | Near-transparency |
| GOP = IB | 1 generation | Transparency | Transparency |
| | 8 generations, No shifts | Transparency | Transparency |
| | 8 generations, 2 spatial shifts | Good quality in most test materials | Transparency |
| | 8 generations, 2 temporal shifts | Transparency | Transparency |
| **525/60 30 Mbits/s** | | **Expert viewers** | **Non-expert viewers** |
| GOP = I | 1 generation | Difficulties in some materials | Good quality in most test materials |
| | 8 generations, No shifts | Difficulties in some materials | Transparency |
| | 8 generations, 2 spatial shifts | Difficulties in some materials | Difficulties in some materials |
| GOP = IB | 1 generation | Good quality in most test materials | Transparency |
| | 8 generations, No shifts | Good quality in most test materials | Transparency |
| | 8 generations, 2 spatial shifts | Difficulties in some materials | Near-transparency |
| | 8 generations, 2 temporal shifts | Difficulties in some materials | Good quality in most test materials |
| **525/60 20 Mbits/s** | | **Expert viewers** | **Non-expert viewers** |
| GOP = IB | 1 generation | Difficulties in some materials | Difficulties in some materials |
| | 8 generations, No shifts | Difficulties in some materials | Transparency |
| | 8 generations, 2 spatial shifts | Difficulties in some materials | Difficulties in some materials |
| | 8 generations, 2 temporal shifts | Difficulties in some materials | Difficulties in some materials |
| GOP = IBBP | 1 generation | Transparency | Transparency |
| | 8 generations, No shifts | Difficulties in some materials | Good quality in most test materials |
| | 8 generations, 2 spatial shifts | Difficulties in some materials | Difficulties in some materials |
| | 8 generations, 2 temporal shifts | Difficulties in some materials | Difficulties in some materials |
| **525/60 mixed environment** | | **Expert viewers** | **Non-expert viewers** |
| 30 Mbits/s GOP = I | 8 generations, No shifts | Difficulties in some materials | Good quality in most test materials |
| 20 Mbits/s GOP = IB | 8 generations, No shifts | Difficulties in some materials | Good quality in most test materials |

**Table J.2 – Subjective test results for the 625/50 system**

| Compression parameters | | Viewers | |
|---|---|---|---|
| **625/50 50 Mbits/s** | | **Expert viewers** | **Non-expert viewers** |
| GOP = I | 1 generation | Transparency | Transparency |
| | 8 generations, 2 spatial shifts | Transparency | Transparency |
| **625/50 30 Mbits/s** | | **Expert viewers** | **Non-expert viewers** |
| GOP = I | 1 generation | Transparency | Transparency |
| | 8 generations, 2 spatial shifts | Good quality in most test materials | Difficulties in some materials |
| GOP = IB | 8 generations, 2 spatial shifts | Good quality in most test materials | Near-transparency |
| | 8 generations, 2 temporal shifts | Good quality in most test materials | Near-transparency |
| **625/50 20 Mbits/s** | | **Expert viewers** | **Non-expert viewers** |
| GOP = IB | 1 generation | Transparency | Transparency |
| | 8 generations, 2 spatial shifts | Difficulties in some materials | Difficulties in some materials |
| | 8 generations, 2 temporal shifts | Difficulties in some materials | Difficulties in some materials |
| GOP = IBBP | 8 generations, 2 spatial shifts | Difficulties in some materials | Difficulties in some materials |
| | 8 generations, 2 temporal shifts | Good quality in most test materials | Good quality in most test materials |
| **625/50 mixed environment Cascaded 20 Mbits/s + 50 Mbits/s + 20 Mbits/s** | | **Expert viewers** | **Non-expert viewers** |
| GOP = IBBP + I + IBBP | 3 generations, 1 temporal shift | Good quality in most test materials | Good quality in most test materials |

# Annex K

## Patents

(This annex does not form an integral part of this Recommendation | International Standard)

(This annex is only applicable to ISO/IEC 13818-2)

**Patent List**

The user's attention is called to the possibility that - for some of the processes specified in this Recommendation | International Standard conformance with this Recommendation | International Standard may require use of an invention covered by patent rights.

By publication of this part of Recommendation | International Standard, no position is taken with respect to the validity of this claim or of any patent rights in connection therewith. However, each company listed in this annex has undertaken to file with the Information Technology Task Force (ITTF) a statement of willingness to grant a license under such rights that they hold on reasonable and non-discriminatory terms and conditions to applicants desiring to obtain such a license.

Information regarding such patents can be obtained from the following organizations.

The table summarizes the formal patent statements received and indicates the parts of the standard to which the statement applies. Three "N"s in a row corresponding to a company means that the statement from the company did not mention any part. The list includes all organizations that have submitted informal patent statements. However, if no "X" is present, no formal patent statement has yet been received from that organization.

**Table -- List of companies supplying patent statements**

| Company | ISO/IEC 13818-1 | ISO/IEC 13818-2 | ISO/IEC 13818-3 |
|---|---|---|---|
| AT&T | X | X | X |
| BBC Research Department | | X | |
| Bellcore | | X | |
| Belgian Science Policy Office | N | N | N |
| BOSCH | X | X | X |
| CCETT | | | |
| Columbia University in the City of New York | N | N | N |
| Compression Labs Incorporated | N | N | N |
| CSELT | | X | |
| David Sarnoff Research Center | X | X | X |
| Deutsche Thomson-Brandt GmbH | N | N | N |
| France Telecom CNET | N | N | N |
| Fraunhofer Gesellschaft | N | N | N |
| FUJITSU Limited | | X | |
| GC Technology Corporation | N | N | N |
| General Instruments | N | N | N |
| Goldstar | N | N | N |
| Hitachi, Ltd. | N | N | N |
| International Business Machines Corporation | X | X | X |
| IRT | | | X |
| KDD Co. Ltd. | | X | |
| Massachusetts Institute of Technology | X | X | X |
| Matsushita Electric Industrial Co., Ltd. | N | N | N |
| Mitsubishi Electric Corporation | N | N | N |
| National Transcommunications Limited | N | N | N |
| NEC Corporation | N | N | N |
| Nippon Hoso Kyokai | | X | |
| Nippon Telegraph and Telephone | N | N | N |
| Nokia Research Center | | X | |

| Company | ISO/IEC 13818-1 | ISO/IEC 13818-2 | ISO/IEC 13818-3 |
|---|---|---|---|
| Norwegian Telecom Research | | X | |
| OKI Electric Industry Co. Ltd | N | N | N |
| Philips Consumer Electronics | N | N | N |
| Qualcomm Incorporated | | X | |
| Royal PTT Nederland N.V., PTT Research (NL) | | | |
| Samsung Electronics | X | X | X |
| Scientific Atlanta | X | X | X |
| Siemens AG | N | N | N |
| Sharp Corporation | N | N | N |
| Sony Corporation | N | N | N |
| Texas Instruments | N | N | N |
| Thomson Consumer Electronics | N | N | N |
| Toshiba Corporation | | X | |
| TV/Com | X | X | X |
| Victor Company of Japan Limited | X | X | X |

ICS 35.040
Price based on 208 pages

© ISO/IEC 2000 – All rights reserved