International Telecommunication Union

ITU-R

Radiocommunication Sector of ITU

Recommendation ITU-R BT.1203-1
(01/2007)

User requirements for generic video bit-rate reduction coding of digital TV signals for an end-to-end television system

BT Series

Broadcasting service (television)

ITU International Telecommunication Union

Foreword

The role of the Radiocommunication Sector is to ensure the rational, equitable, efficient and economical use of the radio-frequency spectrum by all radiocommunication services, including satellite services, and carry out studies without limit of frequency range on the basis of which Recommendations are adopted.

The regulatory and policy functions of the Radiocommunication Sector are performed by World and Regional Radiocommunication Conferences and Radiocommunication Assemblies supported by Study Groups.

Policy on Intellectual Property Right (IPR)

ITU-R policy on IPR is described in the Common Patent Policy for ITU-T/ITU-R/ISO/IEC referenced in Annex 1 of Resolution ITU-R 1. Forms to be used for the submission of patent statements and licensing declarations by patent holders are available from http://www.itu.int/ITU-R/go/patents/en where the Guidelines for Implementation of the Common Patent Policy for ITU-T/ITU-R/ISO/IEC and the ITU-R patent information database can also be found.

Series of ITU-R Recommendations

(Also available online at http://www.itu.int/publ/R-REC/en)

Series	Title
BO	Satellite delivery
BR	Recording for production, archival and play-out; film for television
BS	Broadcasting service (sound)
BT	**Broadcasting service (television)**
F	Fixed service
M	Mobile, radiodetermination, amateur and related satellite services
P	Radiowave propagation
RA	Radio astronomy
RS	Remote sensing systems
S	Fixed-satellite service
SA	Space applications and meteorology
SF	Frequency sharing and coordination between fixed-satellite and fixed service systems
SM	Spectrum management
SNG	Satellite news gathering
TF	Time signals and frequency standards emissions
V	Vocabulary and related subjects

Note: *This ITU-R Recommendation was approved in English under the procedure detailed in Resolution ITU-R 1.*

Electronic Publication
Geneva, 2011

RECOMMENDATION ITU-R BT.1203-1

User requirements for generic video bit-rate reduction coding of digital TV signals for an end-to-end television system*

(Question ITU-R 12/6)

(1995-2007)

Scope

This Recommendation is intended to provide user requirements for video bit-rate reduction coding of digital TV signals for an end-to-end television system, with respect to picture formats, coding schemes, picture quality, etc., emphasizing the implementation of H.262 (MPEG-2 Video) and H.264 (MPEG-4 AVC) codecs.

The ITU Radiocommunication Assembly,

considering

a) that rapid progress is being made in bit-rate reduction coding techniques;

b) that video bit-rate reduction coding of digital TV signals has found wide applications for SNG[1]/ENG[2], for studio production, for contribution, for both primary and secondary distribution and for emission by terrestrial and satellite delivery;

c) that in the total chain of broadcasting, a number of codecs will be used in cascade, and this may lead to a loss of picture quality;

d) that ITU-T Recommendations H.262 (MPEG-2 Video) and H.264 (MPEG-4 AVC) are already used or proposed for codecs, for the applications above;

e) that there could be advantages in having a generic (i.e. related) bit-rate reduction coding in the various applications, so that maximum commonality among various applications can be utilized;

f) that interoperability between different video formats and distribution media should be achieved;

g) that digital and analogue broadcasting systems will both exist during a transition period,

recommends

1 that the same picture format or closely related signal formats should be used as far as possible, through a total broadcasting chain;

---

* Radiocommunication Study Group 6 made editorial amendments to this Recommendation in October 2010 in accordance with Resolution ITU-R 1.

1 The definition of SNG (Satellite News Gathering) can be found in Annex 1 § 1.1 of Recommendation ITU-R SNG.770.

2 The definition of ENG (Electronic News Gathering) can be found in Annex 3 § 2 of Recommendation ITU-R SA.1154 and § 3 of Report ITU-R BT.2069.

2 that the picture format of both input and output signals for coding and decoding should be the same;

3 that the same or closely related coding schemes should be used as much as possible for terrestrial and satellite emission and secondary distribution in order to minimize the receiver cost and the quality degradation;

4 that the encoder should as far as possible allow non-changing parameters which may be used in subsequent coding processes, e.g. motion information to be down-loaded into the subsequent codecs;

5 that the interface between the codecs should be simple,

further recommends

1 that the values listed in Table 1 should be used for the input of the encoder and for the output of the decoder;

2 that the functional and operational requirements described in Table 6 should be satisfied;

3 that the benefits of generic coding applied to the total or to parts of the total broadcasting chain should be studied in terms of ease of operation, equipment cost and picture quality;

4 that the choice between ITU-T Recommendation H.262 (MPEG 2 Video) or ITU-T Recommendation H.264 (MPEG-4 AVC)[3] for the source coding method used by individual administrations should depend on a number of considerations including, for instance, interoperability with legacy equipment, efficient use of the bit rate available in the delivery channel, harmonization with source coding methods adopted by neighbouring administrations for digital terrestrial and satellite broadcast channels.

[3] *Note from the Secretariat*: ITU-T Recommendations H.262 and ITU-T H.264 are available in electronic form at the following address: http://www.itu.int/md/R03-WP6A-C-0110/en.

Annex 1

TABLE 1

Signal format for codecs

Items	SNG/ENG		Contribution	Studio production	Primary distribution	Primary distribution	Terrestrial emission	Satellite emission	Secondary distribution
	Mode 1(1)	Mode 2(1)			Case 1(2)	Case 2(3)			
No. of samples/line and No. of lines/frame	Examples are listed in Table 2								
Colour format	4:2:2 or 4:4:4 should be used for the digital interface				4:2:2 should be used for the digital interface 4:2:0 may be used for internal coding				

(1) Mode 1: good transmission conditions.

Mode 2: poor transmission conditions.

(2) Case 1: digital primary distribution followed by analogue secondary distribution or emission.

(3) Case 2: digital primary distribution followed by digital secondary distribution or emission.

TABLE 2

Examples of picture formats

Recommendation ITU-R	24 Hz environment		50 Hz environment		60 Hz environment	
BT.709/HDTV (16:9)	1 920 × 1 080(1)	P	1 920 × 1 080(1) 1 920 × 1 080(1)	P I	1 920 × 1 080(1) 1 920 × 1 080(1)	P I
BT.1543 (16:9)					1 280 × 720	P
BT.601 and BT.1358/SDTV (4:3 or 16:9)			720 × 576 720 × 576	P I	720 × 483(2) 720 × 483(2)	P I

I: interlaced structure

P: progressive structure

(1) For internal coding, the number of samples/line may be reduced to 1 440.

(2) The number of coded lines may be 480 in the case of emission and secondary distribution applications, although aspect ratio and the picture centre are defined using 483 active lines.

NOTE 1 – Tables 3 and 4 show information about profiles and levels defined in ITU-T Recommendation H.262.

NOTE 2 – Table 5 shows information about levels defined in ITU-T Recommendation H.264.

TABLE 3

Upper bounds for sampling density and luminance sample rate currently in use among those specified in ITU-T Recommendation H.262

Level		Profile				
		Simple	Main	High		4:2:2
				(4:2:0)	(4:2:2)	
High	Samples/line lines/frame frames/s samples/s		1 920 1 088 60 62 668 800	1 920 1 088 60 83 558 400	1 920 1 088 60 62 668 800	1 920 1 088 60 62 668 800
High-1 440	Samples/line lines/frame frames/s samples/s		1 440 1 088 60 47 001 600	1 440 1 088 60 62 668 800	1 440 1 088 60 47 001 600	
Main	Samples/line lines/frame frames/s samples/s	720 576 30 10 368 000	720 576 30 10 368 000	720 576 30 14 745 600	720 576 30 11 095 200	720 608(1) 30 11 095 200
Low	Samples/line lines/frame frames/s samples/s		352 288 30 3 041 280			

(1) 512 lines/frame for 525/60, 608 lines/frame for 625/50.

TABLE 4

Upper bounds for bit rates (Mbit/s) currently in use among those specified in ITU-T Recommendation H.262

Level	Profile			
	Simple	Main	High	4:2:2
High		80	100	300
High-1 440		60	80	
Main	15	15	20	50
Low		4		

TABLE 5

Levels to be used among those specified in ITU-T Recommendation H.264

Level number	Picture size (samples/line × lines/frame)	Frame rate (frames/s)	Maximum video bit rate(1)	Maximum number of reference frames for picture size
3	720 × 480 720 × 576	30 25	10 Mbit/s	5
3.1	720 × 480 720 × 576 1 280 × 720	60 50 30	14 Mbit/s	5
3.2	1 280 × 720	60	20 Mbit/s	4
4	1 920 × 1 080	30, 25, 24	20 Mbit/s	4
4.1	1 920 × 1 080	30, 25, 24	50 Mbit/s	4
4.2	1 920 × 1 080	60, 50	50 Mbit/s	4

(1) For High, High 10 and High 4:2:2 profiles, the following bit-rate multipliers apply:
High: × 1.25, High 10: × 3, High 4:2:2: × 4

Annex 2

TABLE 6

Functional and operational requirements for ITU-T Recommendation H.262 codecs*

Item	SNG/ENG Mode 1	SNG/ENG Mode 2	Contribution	Studio production	Primary distribution Case 1	Primary distribution Case 2	Terrestrial emission	Satellite emission	Secondary distribution
No. of audio channels BT.709/HDTV BT.1543 BT.601 and BT.1358/SDTV	Minimum 2 Minimum 2 Minimum 2		Maximum 8 Maximum 8 Maximum 6		Maximum 6 Maximum 6 Maximum 6				
Range of bit rates BT.709/HDTV BT.1543 BT.601 and BT.1358/SDTV	Up to 140 Mbit/s Up to 140 Mbit/s Up to 34 or 45 Mbit/s				Corresponds to SNG and contribution bit rates	Corresponds to secondary distribution	Up to 80 Mbit/s Up to 80 Mbit/s Up to 15 Mbit/s		
Prediction mode[1]	I, P				(I, B, P) and (I, P) are used in non-live and live broadcasting, respectively				
Picture quality[2]	12%[3]	36%[3]	12%[4]		12%[4]	12%[4]	12%[5]		
Compatibility	Not required				Desirable				
Hierarchical coding	Not required				Required only for the graceful degradation system				
Scalability	Not required, however if needed then lower quality can be obtained with a spatial interpolator				Desirable, needed for hierarchical coding				
Interoperability	Not required				A decoder should decode bit streams with as many picture formats as possible, but not necessarily all				

TABLE 6 (*end*)

<table>
<tr><th>Item</th><th>SNG/ENG Mode 1</th><th>SNG/ENG Mode 2</th><th>Contribution</th><th>Studio production</th><th>Primary distribution Case 1</th><th>Primary distribution Case 2</th><th>Terrestrial emission</th><th>Satellite emission</th><th>Secondary distribution</th></tr>
<tr><td>Editability</td><td></td><td></td><td></td><td>Required in the bit stream domain</td><td></td><td colspan="4"></td></tr>
<tr><td>Bit-rate flexibility</td><td colspan="9">A decoder should decode bit streams at any bit rate described in the item of “range of bit rates”</td></tr>
<tr><td>Codec delay</td><td colspan="9">An overall delay of less than 300 ms would be desirable for interactive talk-back applications</td></tr>
<tr><td>Recovery time
(after a break of 50 ms)</td><td colspan="2">≤ 1 s</td><td>≤ 500 ms</td><td></td><td colspan="5">≤ 500 ms</td></tr>
<tr><td>Acquisition time</td><td colspan="9">The major contributions to acquisition time are the decoding delay and the interval between I pictures. A desirable figure for this value is less than 500 ms</td></tr>
<tr><td>Error concealment</td><td colspan="4">Required, a decoder should support this functionality and should also provide a signalling function of error conditions for studio applications</td><td colspan="5">Desirable</td></tr>
<tr><td>Graceful degradation</td><td colspan="5">Not required</td><td colspan="4">Desirable, essential for mobile and portable reception</td></tr>
<tr><td>Channel hopping latency</td><td colspan="5">Not required</td><td colspan="4">Less than 550 ms</td></tr>
<tr><td>Relative delay between sound and vision</td><td colspan="6">± 2 ms per codec</td><td colspan="3">Under study</td></tr>
</table>

* Functional and operational requirements for ITU-T Recommendation H.264 codecs are under study.

(1) I: I-picture, P: P-picture, B: B-picture.

(2) DSCQS (Double Stimulus Continuous Quality Scale) method is used.

(3) See Recommendation ITU-R BT.1868.

(4) See Recommendation ITU-R BT.1868.

(5) See Recommendation ITU-R BT.1122.

Appendix[4]

Definition and explanation of items listed in Tables 1, 3 and 6

Generic coding: digital coding of pictures based on family of related coding methods.

No. of samples/line: number of luminance samples per active line.

No. of lines/frame: number of vertical lines per active frame.

Colour format: ratio between the number of the luminance pixels and the number of the co-sited chroma difference pixels or the ratio between the colour pixels *R*, *G* and *B*.

No. of audio channels: total number of sound channels per programme, together with a description how these channels can be combined for different applications.

Range of bit rates: minimum and maximum encoder output bit rates for several input formats.

Prediction mode: type of prediction used inside the encoder. This influences very strongly the maximum achievable picture quality of following codecs.

Picture quality: results of the subjective evaluation of the encoding and decoding performance in an error-free channel.

Compatibility: description whether the bit stream syntax allows the separate signal processing of parts of the total bit stream in subsequent codecs.

Hierarchical coding: method to achieve different resolution layers on the decoder side.

Scalability: access to several picture qualities in a single bit stream.

Interoperability: description of the grade of commonality between different bit streams inside the broadcasting chain.

Editability: ability to edit a programme taking into account the structure of the encoder output data.

Bit-rate flexibility: the coding algorithm may allow the use of either CBR (constant bit rate) – or VBR (variable bit rate) – coding.

Codec delay: the delay introduced by the coding/decoding algorithm.

Recovery time: the time period between a physical interruption inside the broadcasting chain and the achievement of full functionality.

Acquisition time: the maximum acceptable waiting time from start of the decoding process until the display of the picture. This might influence the choice of the generic coding scheme.

Error concealment: possibility of the decoder to react in a specified way to alarm signals coming from the FEC part of the decoder.

Graceful degradation: to avoid an abrupt degradation of the picture quality on the decoder side, the output of scalable encoders can be protected by different FEC schemes or by non-uniform modulation schemes. A combination of both methods is also possible.

Channel hopping latency: waiting time introduced by the switching between different TV channels.

---

[4] The definitions of this Appendix only relate to this Recommendation.