5i’

SECTION 3

INFRASTRUCTURE FOR AUDIOVISUAL SERVICES

Recommendation H.200

FRAMEWORK FOR RECOMMENDATIONS FOR AUDIOVISUAL SERVICES

(Melbourne, 1988)

1 Audiovisual services

A number of services are, or will be, defined in CCITT having as their common characteristic the transmission of speech together with other information reaching the eventual user in visual form. This Recommendation concerns a set of such services which should be treated in a harmonised way; it is convenient to refer to the members of this set as ‘‘audiovisual services’’ (abbreviated to AV services).

2 Harmonisation of audiovisual services

While the various audiovisual services may easily be distinguished in terms of their user-application, common methods are used for the transport of signals representing speech, moving or still pictures, and associated controls/indications, and also telematic auxiliary facilities. The standardisation process seeks the greatest possible harmonisation of these common features, confining the distinction to the application layers wherever possible, in order to:

a) Maximise the possibilities for intercommunication between terminals intended for different applications;

b) Maximise the commonality of hardware and software in the interests of economies of scale. The scope for commonality includes: audio and video input/output parameters, audio and video codecs, the control/indication set, frame structures and multiplexing, call control procedures (including multipoint).

The embodiment of this harmonisation policy will be a consistent set of Recommendations, consistent in the sense that all members of the set take into account all other members.

3 Purpose of this Recommendation

The purpose of this Recommendation H.200 is to define the set that shall be consistent. In fulfilling this function it is important to distinguish, at a given time, between Recommendations and draft Recommendations.

Recommendations | re members of the set by virtue of their consistency with other adopted members of the set: these are listed in Annex A to this Recommendation. It is of course necessary to ensure continued consistency when amendments are introduced.

Draft Recommendations | ange from mere titles or outline contents through varying stages of maturity to a stable final draft. As many different intended members of the Recommendation H.200 set are developed in parallel to ensure consistency they should be treated as ‘‘provisional’’ members of the set. The list of set members including provisional items does not form part of Recommendation H.200, but this Recommendation H.200 should be updated in the future to include new members of the set formally adopted.

4 Framework

Recommendations in the H.200 set are arranged in three main sections:

Service definitions | These specify the service as seen by the user, including basic service, optional enhancements, quality, and intercommunication requirements, together with operational aspects; technical implementation methods are taken into account but not defined herein.

Infrastructure | This section includes all the Recommendations which are applicable to two or more distinct services: these encompass network configuration, frame structures, control/indications, communication/intercommunication, and audio/video coding. The ‘‘infrastructure’’ includes this generality of signals which flow on unrestricted digital bearers on established network connections — it does not include the methods of call establishment and control, orchestrated by signals outside these bearers.

Systems and terminal equipment | This section deals with the technical implementation of specific services: it therefore includes service-specific equipment for the application layer, and draws upon the infrastructure Recommendations to identify the detailed processes required for the particular service.

A *network aspects* | ection is also proposed, to cover those matters which are particular to AV services but, involving out-of-band signals, do not come within the scope of the infrastructure section above.

5 List of audiovisual services covered

The following audiovisual services shall be included in the harmonized set:

- narrowband videophone (1 and 2×64 kbit/s under study);
- broadband videophone (a teleservice for broadband ISDN);
- narrowband videoconferencing ($n \times 384$ kbit/s and $m \times 64$ kbit/s under study);
- broadband videoconferencing (a teleservice for broadband ISDN);
- audiographic teleconferencing;
- telephony (a degenerate case of an AV service, included for intercommunication purposes);
- telesurveillance.

The following audiovisual services are in the process of being defined, and consideration should be given to their inclusion in the set for either of the reasons given in § 2:

- video mail;
- videotex (including pictures and sound);

— video retrieval;

— high resolution image retrieval;

— distribution services.

ANNEX A
(to Recommendation H.200)

Framework for Recommendations for audiovisual services

CCITT Rec. No.

A.1 *Service definition*

AV100 General Recommendation for AV services F.700

AV110 Teleconference services F.710 AV111

AV112

AV120 (Videophone services) AV121 Basic narrow-band videophone service in the ISDN F.721

A.2 *Infrastructure*

AV200 (General Recommendation for AV service infrastructure)

AV210 (Reference network configuration)

AV220 (General Recommendation for frame structures)

AV221 Frame structure for a 64 kbit/s channel

in audiovisual teleservices H.221 AV222 Frame structure for 384-1920 kbit/s channels audiovisual teleservices H.222

AV230 (AV system control and indications)

AV240 (Principles for communiction between AV terminals)

AV241 System aspects for the use of the 7 kHz audio

codec within 64 kbit/s G.725 AV242

AV250 (Audio coding)

AV251 Narrow-band audio coding at 64 kbit/s G.711 AV252 Wideband audio coding in 64 kbit/s G.722 AV253

AV254

AV260 (Video coding)

AV261 $n \times 384$ kbit/s video codec H.261 AV262

A.3 *Systems and terminal equipment*

AV300 (General Recommendations for AV systems and terminals)

AV310 (Requirements for teleconferencing)

AV311

AV312

AV313 (Teleconference protocol)

AV320 (Requirements for videophone services)

A.4 *Network aspects*

AV400

AV410 (Reservation systems)

AV420 (HLC for use in audiovisual calls)

AV430 (Call control command & indication)

AV440 (Multipoint call set-up)

Note 1 — It is intended to merge the substance of existing Recommendations H.100 and H.110 into this framework in the next study period.

Note 2 — Entries in parentheses are indicative of the purpose of the various positions in the framework.

Note 3 — Further Recommendations will be added to the list as they are formally adopted.

Recommendation H.221

FRAME STRUCTURE FOR A 64 kbit/s CHANNEL IN AUDIOVISUAL | fR TELESERVICES

(Melbourne, 1988)

Introduction

The purpose of this Recommendation is to define a frame structure for audiovisual teleservices in a single 64 kbit/s channel which makes the best use of the characteristics and properties of the audio/video encoding algorithms, of the transmission framing structure and of the existing CCITT Recommendations. It offers several advantages:

— It takes into account Recommendations such as G.704, X.30/I.461, etc. It may allow the use of existing hardware or software.

— It is simple, economic and flexible. It may be implemented on a simple microprocessor, using well known hardware principles.

— It is a synchronous procedure. The exact time of a configuration change is the same in the transmitter and the receiver. Configurations can be changed at 20 ms intervals.

— It needs no return link, since a configuration is signalled by a repeatedly transmitted codeword.

— It is very secure in case of transmission errors, since the BAS is protected by a double error correcting code.

— It allows the control of a higher multiplex configuration, into which the basis 64 kbit/s channel is inserted (in the case of $n \times 64$ kbit/s multimedia services such as videoconference).

— It can be used to derive octet synchronization in networks where this is not provided by other means.

— It can be used in multipoint configurations, where no dialogue is needed to negotiate the use of a data channel.

— It provides a variety of data bit-rates (from 6.25 bit/s up to 64 kbit/s) to the user.

1 Basic principle

The 64 kbit/s channel is structured into octets transmitted at 8 kHz. The eighth bit of each octet conveys a subchannel of 8 kbit/s. This subchannel, called service channel (SC), provides end-to-end signalling and consists of three parts (see Figure 1/H.221):

— *Frame alignment signal* (FAS): This signal structures the 64 kbit/s channel into frames of 80 octets each and multiframes (MF) of 16 frames each. Each multiframe is divided into eight 2-frame submultiframes (SMF): In addition to framing and multiframing information, control and alarm information may be inserted, as well as error check information to control end-to-end error performance and to check frame alignment validity. The FAS can be used to derive octet timing when it is not provided by the network.

— *Bit-rate alloction signal* (BAS): This signal allows the transmission of codewords to describe *the capability* of a terminal to structure the residual 62.4 kbit/s capacity in various ways, and to *command* a receiver to demultiplex and make use of the constituent signals in such structures; if other 64 kbit/s channels are associated, as in the case of $n \times 64$ kbit/s services (e.g. videoconference, videophone), this association may also be defined.

Note — For some countries having 56 kbit/s channels, the net available bit rates will be 8 kbit/s less.

— *Application channel* (AC): This channel allows transmission of binary information of the insertion of message type data channel(s) (e.g. for telematic information) at up to 6400 bit/s. A minimum required command and indication channel should be provided and defined as part of the application channel (for further study). The remaining bit rate for the application channel may be added to the sound data or video channel. In this context, compatibility problems among audiovisual services should be considered.

Figure 1/H.221 [T1.221] (a traiter comme tableau MEP), p.

The remaining 56 kbit/s capacity (with fully reserved application channel), carried in bits 1 to 7 of each octet, may convey a variety of signals within the framework of a multimedia service, under the control of the BAS and possibly the AC. Some examples follow:

— Voice, encoded at 56 kbit/s using a truncated form of the PCM of Recommendation G.711 (A-law or μ-law).

— Voice, encoded at 32 kbit/s and data at 24 kbit/s or less.

— Voice, encoded at 56 kbit/s with a bandwidth 50 to 7000 Hz (sub-band ADPCM according to Recommendation G.722). The coding algorithm is also able to work at 48 kbit/s. Data can then be dynamically inserted at up to 14.4 kbit/s.

— Still pictures coded at 56 kbit/s.

— Data at 56 kbit/s inside an audiovisual session (e.g. file transfer for communicating between personal computers).

— Sound and video sharing the 56 kbit/s capacity.

2 Frame alignment

2.1 *General*

An 80-octet frame length produces an 80-bit word in the service channel. These 80 bits are numbered 1 to 80. Bits 2-8 of the service channel in every even frame contain the frame alignment word (FAW) 0011011. These bits are completed by bit 2 in the succeeding odd frame to form the complete frame alignment signal (FAS).

So a pattern similar to the one in Recommendation G.704 is used (see Figure 2/H.221).

Figure 2/H.221 [T2.221] (a traiter comme tableau MEP), p.

2.2 *Multiframe structure*

Each multiframe contains 16 consecutive frames numbered 0 to 15 divided into eight submultiframes of 2 frames each (Figure 3/H.221). The multiframe alignment signal is located in bit 1 of frames 1-3-5-7-9-11 and has the form 001011. Bits 1 of frames 8-10-12-13-14-15 are reserved for future use. Their value is provisionally fixed at 0.

Bits 1 of frames 0-2-4-6 may be used for a modulo 16 counter to number multiframes in descending order. The least significant bit is transmitted in frame 0, and the most significant bit in frame 6. The receiver may use the mutiframe numbering to determine the differential delay of separate 64 kbit/s connections, and to synchronize the received signals. The use of an additional reserved bit in frame 8 to turn on and off the counting procedure is for further study.

2.3 *Loss and recovery of frame alignment*

Frame alignment is defined to have been lost when three consecutive frame alignment signals have been received with an error.

Frame alignment is defined to have been recovered when the following sequence is detected:

— for the first time, the presence of the correct frame alignment word;

— the absence of the frame alignment signal in the following frame detected by verifying that bit 2 is a 1;

— for the second time, the presence of the correct frame alignment word in the next frame.

When the frame alignment is lost, bit 3 (A) of the next odd frame is set to 1 in the transmit direction.

If frame alignment is achieved, but multiframe alignment cannot be achieved, then frame alignment should be sought at another position.

Figure 3/H.221 [T3.221] (à traiter comme tableau MEP), p.3

2.4 *Loss and recovery of multiframe alignment*

Multiframe alignment is needed to validate the bit-rate allocation signal (see § 3). The criteria for loss and recovery of multiframe alignment described below are provisional.

Multiframe alignment is defined to have been lost when three consecutive multiframe alignment signals have been received with an error. It is defined to have recovered when the multiframe alignment signal has been received with no error in the next multiframe. When multiframe alignment is lost, even when an unframed mode is received, bit 3 (A) of the next odd frame is set to 1 in the transmit direction. It is reset to 0 when multiframe alignment is regained again.

2.5 *Procedure to recover octet timing from frame alignment*

When the network does not provide octet timing, the terminal may recover octet timing in the receive direction from bit timing and from the frame alignment. The octet timing in the transmit direction may be derived from the network bit timing and an internal octet timing.

2.5.1 *General rule*

The receive octet timing is normally determined from the FAS position. But at the start of the call and before the frame alignment is gained, the receive octet timing may be taken to be the same as the internal transmit octet timing. As soon as a first frame alignment is gained, the receive octet timing is initialized as the new bit position, but it is not yet validated. It will be validated only when frame alignment is not lost during the next 16 frames.

2.5.2 *Particular cases*

a) When, at the initiation of a call, the terminal is in a forced reception mode, or when the frame alignment has not yet been gained, the terminal may temporarily use the transmit octet timing.

b) When frame alignment is lost after being gained, the receive octet timing should not change until frame alignment is recovered.

c) As soon as frame and multiframe alignment have been gained once, the octet timing is considered as valid for the rest of the call, unless frame alignment is lost and a new frame alignment is gained at another bit position.

d) When the terminal switches from a framed mode to an unframed mode (by means of the BAS), the octet timing, previously gained, must be kept.

e) When a new frame alignment is gained on a new position, different from that previously validated, the receive octet timing is reinitialized to the new position but not yet validated and the previous bit position is stored. If no loss of frame alignment occurs in the next 16 frames, the new position is validated; otherwise the stored old bit position is reutilized.

2.5.3 *Search for frame alignment signal (FAS)*

Two methods may be used: sequential or parallel. In the sequential method, each of the eight possible bit positions for the FAS is tried. When FAS is lost after being validated, the search must resume starting from the previously validated bit position. In the parallel method, a sliding window, shifting one bit for each bit period, may be used. In that case, when frame alignment is lost, the search must resume starting from the bit position next to the previously validated one.

2.6 *Description of the CRC4 procedure*

In order to provide an end-to-end quality monitoring of the 64 kbit/s connection, a CRC4 procedure may be used and the four bits C1, C2, C3 and C4 computed at the source location are inserted in bit positions 5 to 8 of the odd frames. In addition, bit 4 of the odd frames, noted E, is used to transmit an indication about the received signal in the opposite direction whether the most recent CRC block has been received with errors or not.

When the CRC4 procedure is not used, bit E shall be set to 0, and bits C1, C2, C3 and C4 shall be set to 1 by the transmitter. Provisionally, the receiver may disable reporting of CRC errors after receiving eight consecutive CRCs set to all 1s, and it may enable reporting of CRC errors after receiving two consecutive CRCs each containing a 0 bit. (This method of eanbling and disabling CRC error reporting must be verified and is for further study.)

2.6.1 *Computation of the CRC4 bits*

The CRC4 bits C1, C2, C3 and C4 are computed from the whole 64 kbit/s channel, for a block made of two frames: one even frame (containing the FAW) followed by one odd frame (not containing the FAW). The CRC4 block size is then 160 octets, i.e. 1280 bits, and the computation is performed 50 times per second.

2.6.1.1 *Multiplication division process*

A given C1-C4 word located in block N is the remainder after multiplication by x^4 and then division (modulo 2) by the generator polynominal x^{D}lFBOCAD15 4 + x + 1 of the polynomial representation of block (N—1).

When representing contents of a block as a polynominal the first bit in the block should be taken as being the most significant bit. Similarly C1 is defined to be the most significant bit of the remainder and C4 the least significant bit of the remainder.

This process can be realized with a four-stage register and two exclusive-ors.

2.6.1.2 *Encoding procedure*

i) The CRC bit positions in the odd frame are initially set at zero, i.e. C1 = C2 = C3 = C4 = 0.

ii) The block is then acted upon by the multiplication-division process referred to above in § 2.6.1.1.

iii) The remainder resulting from the multiplication-division process is stored ready for insertion into the respective CRC locations of the next odd frame.

Note — These CRC bits do not affect the computation of the CRC bits of the next block, since the corresponding locations are set at zero before the computation.

2.6.1.3 *Decoding procedure*

i) A received block is acted upon by the multiplication division process, referred above in § 2.6.1.1, after having its CRC bits extracted and replaced by zeros.

ii) The remainder resulting from this multiplication-division process is then stored and subsequently compared on a bit-by-bit basis with the CRC bits received in the next block.

iii) If the decoded calculated remainder exactly corresponds to the CRC bits sent from the encoder, it is assumed that the checked block is error-free.

2.6.2 *Consequent actions*

2.6.2.1 *Action on bit E*

Bit E of block N is set to 1 in the transmitting direction of bits C1-C4 detected in the most recent block in the opposite direction have been found in error (at least one bit in error). In the opposite case, it is at zero.

2.6.2.2 *Monitoring for incorrect frame alignment*

In case of a long simulation of the FAW, the CRC4 information can be used to re-invite a search for frame alignment. For such a purpose, it is possible to count the number of blocks CRC in error within 2 s (100 blocks) and to compare this number with 89. If the number of CRC blocks in error is greater than or equal to 89, a search for frame alignment should be re-initiated.

These values of 100 and 89 have been chosen in order that:

— for a random transmission error rate of 10^{D}lF261^{3}, the probability of incorrectly re-initiating a search for frame alignment because of 89 or more blocks in error, be less than 10^{D}lF261^{4};

— in case of simulation of frame alignment, the probability of not re-initiating a search of frame alignment after a 2 s period be less than 2.5%.

2.6.2.3 *Monitoring for error performance*

The quality of the 64 kbit/s connection can be monitored by counting the number of CRC blocks in error within a period of one second (50 blocks). For instance, a good evaluation of the proportion of seconds without errors as defined in Recommendation G.821 can be provided.

For information purposes, the following proportions of CRC block in error can be computed for randomly distributed errors of error rate Pe, as shown in Table 1/H.221.

H.T. [T4.221]

TABLE 1/H.221

Pe	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
{ Proportion of CRC blocks in error }	70%	12%	1.2%	0.12%	0.012%

Tableau 1/H.221 [T4.221], p.

By counting the received E bits, it is possible to monitor the quality of the connection in the opposite direction.

3 Bit-rate allocation signal (BAS) and switching between configurations

The bit-rate allocation signal (BAS) occupies bits 9-16 of the service channel in every frame. An eight bit BAS code (b_0, b_1, b_2, b_3, b_4, b_5, b_6, b_7) is complemented by eight error correction bits (p_0, p_1, p_2, p_3, p_4, p_5, p_6, p_7) to implement a (16,8) double error correcting code. This error correcting code is obtained by shortening the (17,9) cyclic code with generator polynomial:

$$g(x) = x^8 + x^7 + x^6 + x^4 + x^2 + x + 1$$

The error correction bits are calculated as coefficients of the remainder polynomial in the following equation:

$$p_0 x^7 + p_1 x^6 + p_2 x^5 + p_3 x^4 + p_4 x^3 + p_5 x^2 + p_6 x + p_7 = RES_{g(x)} [b_0 x^{15} + b_1 x^{14} + b_2 x^{13} + b_3 x^{12} + b_4 x^{11} + b_5 x^{10} + b_6 x^9 + b_7 x^8]$$

where $RES_{g(x)}[f(x)]$ represents the residue obtained by dividing $f(x)$ by $g(x)$.

The BAS code is sent in the even-numbered frame, while the associated error correction bits are sent in the subsequent odd-numbered frame. Each bit of BAS code or the error correction is transmitted in the order shown in Table 2/H.221, to avoid emulation of the frame alignment signal.

H.T. [T5.221]

TABLE 2/H.221

Bit position	Even frame	Odd frame
9	b 0	P 2
10	b 3	P 1
11	b 2	P 0
12	b 1	P 4
13	b 5	P 3
14	b 4	P 5
15	b 6	P 6
16	b 7	P 7

Tableau 2/H.221 [T5.221], p.

The decoded BAS value is valid if:

— the receiver is in frame and multiframe alignment, and

— the FAS in the same submultiframe was received with 2 or fewer bits in error.

Otherwise, the decoded BAS value is ignored. When the receiver actually loses frame alignment, it should undo any changes caused by the three previously decoded BAS values and revert to the state determined by the fourth previously decoded BAS value.

The encoding of BAS is made in accordance with the attribute method.

The first three bits (b_0, b_1, b_2) represent the attribute number, which describes the general command or capability, and the next five bits (b_3, b_4, b_5, b_6, b_7) identify the specific command or capability. The following attributes are defined:

000 Audio coding command: values defined in Annex A

001 Transfer rate command: values defined in Annex B

010 Video and other command: values defined in Annex D

011 Data command: values defined in Annex E

100 Terminal capability: values defined in Annex C

Annex A defined a number of modes, according to the audio coding type and bit rate. Since a validated value of BAS command code applies to the next submultiframe, a change in configuration can occur every 20 ms. This applies equally to the use of video and data command BAS, controlling sub-modes of various configurations of the remaining capacity.

When the incoming bit A (see § 2.3) is set to 1, the distant rceiver is not in multiframe alignment and will not immediately validate a new BAS value.

Capability BAS require a response from the distant terminal and should not be sent unnecessarily when the incoming signal is unframed.

See Recommendations G.725 for further information on signalling procedures.

4 Application channel (AC)

It occupies bits 17-80 of the service channel in each frame, providing a user available bit rate of 6.4 kbit/s. According to the application, different kinds of information may be inserted herein. In particular, information concerning forward error correction or end-to-end encryption which both depend on the application, could take place in the application channel.

The AC may be used to convey a message channel conforming to the OSI protocols where appropriate. With this message channel, a transport and a session protocol may be used to control the use of audio and data channels. For example, once the command/response procedure has agreed to open a connection, if necessary, the BAS is used to adjust the capability available for data.

Examples for the use of AC are given in Appendix I.

5 Access to non-audio information within bits 1-7

Use of attribute (000) according to Annex A provides for the static or dynamic allocation of ‘‘data channels’’ of up to 56 kbit/s capacity; in some applications, it may be desirable to combine the application channel with the data channel in order to have a single user-data path, of capacity up to 62.4 kbit/s.

Unless BAS codes (010), (011) are used to direct otherwise, the ‘‘data channel’’ is treated as a single stream of non-video information; in this case access may be realised according to standardised procedures (e.g. Recommendations I.461, I.462, I.463). Data is transmitted in the order received from the data terminal equipment or data terminal adaptator.

In the presence of a non-zero video command BAS (010) the data channel is assigned to moving picture information, except that some part may be subtracted for other data purposes by application of a non-zero data command BAS (011).

ANNEX A
(to Recommendation H.221)

Attribute 000 used for BAS encoding

H.T. [T6.221]

{
Attribute
Bits b
0 | (hy |
2
}
{
Attribute value
Bits b
3 | (hy |
7
}
Meaning

000 Audio coding
00000 {
Neutralised channel (the 62.4 kbit/s user date are unused)
PCM [G.711] (truncated to 7 bits)
(Note 1)
PCM [G.711] (truncated to 7 bits)
(Note 2)
}

S0010 S0011 {
A-law; data at 0 or 6.4 kbit/s
Mode OF
µ-law; data at 0 or 6.4 kbit/s
Mode OF
}

S0001 {
32 kbit/s ADPCM data at 24 or 30.4 kbit/s
(Note 3)
}

{
64 kbit/s unframed mode
(Note 4)
}

00100 PCM A-law Mode 0
00101 PCM µ-law Mode 0
00110 {
SB-ADPCM [G.722]
Mode 1
SB-ADPCM [G.722]
(Note 5)
}

00111 {
0 kbit/s; data at 64 kbit/s
Mode 10
}

{
Variable bit-rate audio coding
}

S1000 {
G.722 56 kbit/s; data at 0 or 6.4 kbit/s
Mode 2
}

S1001 {
G.722 48 kbit/s; data at 8 or 14.4 kbit/s
Mode 3
}

S1010 {
?04
Reserved for audio coding
}

. | | ?05 at bit rates less than
S1110 ⌡ 48 kbit/s (Note 6)
S1111 {

0 kbit/s; data at 56 or 62.4 kbit/s Mode 9 0 kbit/s; data at 56 or 62.4 kbit/s (Note 7) }		
	10000 101xx	Free Free

Note 1 — The 8th bit is fixed to 0 in the audio PCM decoder.

Note 2 — The S bit set to 1 indicates that the application channel is merged with the data channel to form a single user-data path. The method for merging the two channels is shown in Figure A-1/H.221 for the 14.4 kbit/s case.

Note 3 — The coding law and respective place of data and audio in each byte of the 64 kbit/s channel is under study.

Note 4 — Attribute values 001xx imply the switching to an unframed mode. In the receive direction, reverting to a framed mode can only be achieved by recovering frame and multiframe alignment, which might take up to 2 multiframes (i.e. 320 ms).

Note 5 — The allocation of bits in each byte of the 64 kbit/s channel is as follows:

Audio bit-rate	1	2	3	4	5	6	7	8
64 kbit/s	H	H	L	L	L	L	L	L
56 kbit/s	H	H	L	L	L	L	L	S
48 kbit/s	H	H	L	L	L	L	D	S

S Service channel

D Data channel

H High band audio

B Low band audio

Tableau [T6.221], p.

Figure A-1/H.221 [T7.221] (a traiter comme tableau MEP), p.

ANNEX B
(to Recommendation H.221)

Attribute 001 used for BAS encoding

H.T. [T8.221]

{ Attribute Bits b 0 | (hy | 2 }	{ Attribute value Bits b 3 | (hy | 7 }	Meaning
001 Transfert rate	00000	64 kbit/s
	00001	{ 64 kbit/s (audio) + 64 kbit/s (data/video) }
	00010	{ 64 kbit/s (audio) + 64 kbit/s (data/video) treated as a single 128 kbit/s channel }
	01010	{ 384 kbit/s: 64 (audio) + 320 (video) }
	01011	{ **384 kbit/s:** 64 (audio) + 256 (video) **384 kbit/s:** + 64 (data) }
	01100	{ 768 kbit/s: 64 (audio) + 704 (video) }
	01101	{ **768 kbit/s:** 64 (audio) + 640 (video) **768 kbit/s:** + 64 (data) }
	01110	{ 1152 kbit/s: 64 (audio) + 1088 (video) }
	01111	{ **1152 kbit/s:** 64 (audio) + 1024 (video) **768 kbit/s:** + 64 (data) }
	10000	{ 1536 kbit/s: 64 (audio) + 1472 (video) }
	10001	{ **1536 kbit/s:** 64 (audio) + 1408 (video) **1536 kbit/s:** + 64 (data) }
	10010	{ 1920 kbit/s: 64 (audio) + 1856 (video) }
	10011	{ **1920 kbit/s:** 64 (audio) + 1792 (video) **1920 kbit/s:**

+ 64 (data) }			

Tableau [T8.221], p.

ANNEX C
(to Recommendation H.221)

Attribute 100 used for BAS encoding

H.T. [T9.221]

Attribute Bits b 0 | (hy | 2	Attribute value Bits b 3 | (hy | 7	Meaning
100	00000	| eutral (Note 1)
Terminal	00001	| .725 Type 0 — A-law (Note 2)
capability	00010	| .725 Type 0 — μ-law
	00011	| .725 Type 1 — G.722
	00100	| .725 Type 2 — G.722 + data
	00101	?04
	| | |	?05 Reserved for audio capabilities
	00110	
	00111	| eserved for national use
	01000	| on-standard video capability (Note 3)
	01001	?04
	. | |	?05 Reserved for video capabilities
	01110	
	01111	| eserved for national use
	10000	| on-standard system capability (Note 3)
	10001	| B transfer rate capability (Note 4)
	10010	| B transfer rate capability (Note 4)
	10011	| B transfer rate capability (Note 4)
	10100	| B transfer rate capability (Note 4)
	10101	| B transfer rate capability (Note 4)
	10110	| eserved for transfer rate capability
	10111	| eserved for national use
	11000	| 00 bit/s data capability (Note 5)

| 200 bit/s data capability (Note 5) }	11001	{
| 400 bit/s data capability (Note 5) }	11010	{
| 800 bit/s data capability (Note 5) }	11011	{
| 400 bit/s data capability (Note 5) }	11100	{
| 000 bit/s data capability (Note 5) }	11101	{
| 600 bit/s data capability (Note 5) }	11110	{
| 4 | 00 bit/s data capability (Note 5) }	11111	{

Note 1 — The neutral value indicates no change in the current capabilities of the terminal.

Note 2 — Types 0, 1 and 2 are defined according to Recommendation G.725 § 2.

— Type 0 terminal can work in mode 0 (PCM) only.

— Type 1 terminal preferably works in mode 1 (G.722) but is able to work in mode 0.

— Type 2 terminal preferably works in mode 2 (G.722 + H.221) but is able to work in modes 1 and 0.

Note 3 — If sent (additional), an improved video algorithm decoding or whole system capability is indicated; it is specified elsewhere.

Note 4 — A capability to use several B channels implies the capability to use fewer channels.

Note 5 — A data capability specifies only one rate; if multiple rates are possible the capabilities are sent individually.

Tableau [T9.221], p.

ANNEX D
(to Recommendation H.221)

Attribute 010 used for BAS encoding

H.T. [T10.221]

{ Attribute Bits b 0 | (hy | 2 } Attribute value Bits b 3 | (hy | 7 }	{ Meaning	
010 Video and other Standard video for m × 64 kbit/s }	00000 00001	No video; video switched OFF {
command Video ON, using improved algorithm }	00010	{
Standard video to Recommendation H.261 }	00011	{
Transfer to non-standard system mode }	. | | 11111	{

Tableau [T10.221], p.

ANNEX E
(to Recommendation H.221)

Attribute 011 used for BAS encoding

H.T. [T11.221]

Attribute Bits b 0-b 2 Attribute value Bits b 3-b 7 }	{ Meaning	
011 | o data; data switched OFF } Data command	00000	{
| 00 bit/s in AC assigned to data | bit 8 of last three octets in each frame) }	00001	{
| 200 bit/s in AC assigned to data | bit 8 of last 12 octets in each frame) }	00010	{
| 800 bit/s in AC assigned to data | bit 8 of last 48 octets in each frame) }	00011	{
| 400 bit/s in AC assigned to data (whole of AC) }	00100	{
| 000 bit/s assigned to data (bit 7) }	00101	{
| 600 bit/s assigned to data (bit 7 + bit 8 of last 16 octets | n each frame) }	00110	{
| 4.4 kbit/s assigned to data (bit 7 + AC) }	00111	{
?04 Reserved for communicating the status }	. | | 10000	{
?05 of the data terminal equipment }	to	{
	10111	⌡ interfaces
Variable rate data; data switched ON (Note) }	. | | 11111	{

Note — When video is switched on, the entire variable data capacity is used for video.

Tableau [T11.221], p.

APPENDIX I
(to Recommendation H.221)

Examples for the use of the application channel

I.1 *Binary information*

Each bit of the application channel may be used to convey the information of a 100 kbit/s channel, repeated 100 times per second. If odd and even frames are identified, each bit may carry the 150 Hz bit/s channels. If multiframing is used, each bit may carry the information of 16 channels, each at 6.25 bit/s.

An example of this kind of information is, in teleconference, the use of a bit to synchronize the encoder clock on the receive clock, or to indicate the microphone number, or to signal the use of the grahics mode, etc.

I.2 *Synchronous message-type channel*

As each bit of the application channel represents a bit-rate of 100 bit/s, any synchronous channel working at $n \times 100$ bit/s may be inserted in the application channel. An example is, in videoconference, the message channel at 4 kbit/s which is used for multipoint management.

Another possibility is the insertion of data channels at one of the bit rates defined in Recommendation X.1, according to Recommendations X.30/I.461: ‘‘Support of X.21 and X.21*bis* | ased DTEs by an ISDN’’. The present frame structure is consistent with the Recommendations X.30/I.461 frame structure in a double way:

— it has the same length (80 bits by bearer channel at 8 kbit/s);

— it needs 63 bits per frame (17 bits are used for framing information not to be transmitted), which fits into the 64 bits available in this frame structure.

I.3 *Asynchronous message-type channel*

In case of asynchronous terminals, Recommendation X.1 bit-rates are relevant, too. The applicable standard is that specified in [1]. This standard also uses the same 80-bit frame structure as Recommendations X.30/I.461 mentioned above. The application channel will therefore allow adoption of this ECMA standard if needed.

I.4 *Error correction and encryption*

When needed, forward error correction and encryption information may be transmitted in the application channel. The bit-rate and the protocol to be used will depend on the application.

Reference

[1] ECMA-TAxx *Bit-rate adaption for the support of synchronous and asynchronous terminal equipment using the V-Series interfaces on a PSTN* .

Recommendation H.222

FRAME STRUCTURE FOR 384-1920 kbit/s CHANNELS IN | fR AUDIOVISUAL TELESERVICES

(Melbourne, 1988)

1 Scope

This Recommendation provides a mechanism to multiplex multimedia signals such as audio, video, data, control and indication, etc., for audiovisual teleservices using an $n \times 384$ kbit/s ($n = 1$-5) channel.

2 Basic structure

The multiplex structure is based upon multiple octets transmitted at 8 kHz as in Recommendation I.431.

As $n \times 384$ kbit/s channel consists of $6 \times n$ time slots of 64 kbit/s (see Figure 1/H.222). The first 64 kbit/s time slot has a frame structure conforming to Recommendation H.221, containing frame alignment signal (FAS), bit rate allocation signal (BAS) and application channel (AC).

Figure 1/H.222, p.

3 BAS codes

Particular codes for allocating audio, video and data signals in an $n \times 384$ kbit/s channel are given in Annex B to Recommendation H.221 for Attribute ‘‘001’’.

4 Data transmission

A 64 kbit/s data channel can be allocated to the fourth time slot in the $n \times 384$ kbit/s channel if controlled by the corresponding BAS code.

Provision of more than one 64 kbit/s data channels is under study.

5 Bit assignment in application channel

Application channel conveys control and indication signals, message channel, etc., for audiovisual teleservices using $n \times 384$ kbit/s transmission. Bit assignment is under study.

Recommendation H.261

CODEC FOR AUDIOVISUAL SERVICES AT n × 384 kbit/s

(Melbourne, 1988)

The CCITT,

considering

(a) that there is significant customer demand for videoconference service;

(b) that circuits to meet this demand can be provided by digital transmission using the H_0 rate or its multiples up to the primary rate;

(c) that ISDNs are likely to be available in some countries that provide a switched transmission service at the H_0 rate;

(d) that the existence of different digital hierarchies and different television standards in different parts of the world complicates the problems of specifying coding and transmission standards for international connections;

(e) that videophone services are likely to appear using basic ISDN access and that some means of interconnection of videophone and videoconference terminals should be possible;

(f) that Recommendation H.120 for videoconferencing using primary digital group transmission was the first in an evolving series of Recommendations,

appreciating

that advances are being made in research and development of video coding and bit rate reduction techniques which will lead to further Recommendations for videophone and videoconferencing at multiples of 64 kbit/s during subsequent Study Periods, so that this may be considered as the second in the evolving series of Recommendations.

and noting

that it is the basic objective of CCITT to recommend unique solutions for international connections,

recommends

that in addition to those codecs complying with Recommendation H.120, codecs having signal processing and interface characteristics described below should be used for international videoconference connections.

Note 1 — Codecs of this type are also suitable for some television services where full broadcast quality is not required.

Note 2 — Equipment for transcoding from and to codecs according to Recommendation H.120 is under study.

Note 3 — It is recognised that the objective is to provide interworking between $n \times 384$ kbit/s codecs and $m \times 64$ kbit/s codecs as defined in the H-Series Recommendations. Interworking will be on the basis of $m \times 64$ kbit/s, where the values of m are under study.

1 Scope

This Recommendation describes the coding and decoding methods for audiovisual services at the rates of $n \times 384$ kbit/s, where n is 1 to 5. Possible extension of this scope to meet the objective in Note 3 above is under study.

2 Brief specification

An outline block diagram of the codec is given in Figure 1/H.261.

2.1 *Video input and output*

To permit a single Recommendation to cover use in and between 625 and 525-line regions, pictures are coded in one common intermediate format. The standards of the input and output television signals, which may, for example, be composite or component, analogue or digital and the methods of performing any necessary conversion to and from the intermediate coding format are not subject to recommendation.

Figure 1/H.261, p.13

2.2 *Digital output and input*

Digital access at the primary rate of 1544 or 2048 kbit/s is with vacated time slots in accordance with Recommendation I.431.

Interfaces using ISDN basic accesses are under study (see Recommendation I.420).

2.3 *Sampling frequency*

Pictures are sampled at an integer multiple of the video line rate. This sampling clock and the digital network clock are asynchronous.

2.4 *Source coding algorithm*

A hybrid of inter-picture prediction to utilize temporal redundancy and transform coding of the remaining signal to reduce spatial redundancy is adopted. The decoder has motion compensation capability, allowing optional incorporation of this technique in the coder.

2.5 *Audio channel*

Audio is coded according to mode 2 of Recommendation G.722. This is combined with control and indication information and conveyed in one 64 kbit/s time slot which conforms to Recommendation H.221.

2.6 *Data channels*

Recommendation H.221 permits part of the 64 kbit/s time slot carrying the audio to be used for auxiliary data transmission.

Additionally, one of the time slots normally used for video may be reassigned as a 64 kbit/s data channel. The possibility of further such channels is under study.

2.7 *Symmetry of transmission*

The codec may be used for bidirectional or unidirectional audiovisual communication.

2.8 *Error handling*

Under study.

2.9 *Propagation delay*

Under study.

2.10 *Additional facilities*

Under study.

3 Source coder

3.1 *Source format*

The source coder operates on non-interlaced pictures occurring 30000/1001 (approximately 29.97) times per second. The tolerance on picture frequency is ± | 0 ppm.

Pictures are coded as luminance and two colour difference components (Y, C_R et C_B). These components and the codes representing their sampled values are as defined in CCIR Recommendation 601.

Black = 16

White = 235

Zero colour difference = 128

Peak colour difference = 16 and 240.

These values are nominal ones and the coding algorithm functions with input values of 0 through to 255.

For coding, the luminance sampling structure is 288 lines per picture, 352 pels per line in an orthogonal arrangement. Sampling of each of the two colour difference components is at 144 lines, 176 pels per line, orthogonal. Colour difference samples are sited such that their block boundaries coincide with luminance block boundaries as shown in Figure 2/H.261. The picture area covered by these numbers of pels and lines has an aspect ratio of 4 | | and corresponds to the active portion of the local standard video input.

Note — The number of pels per line is compatible with sampling the active portions of the luminance and colour difference signals from 525 to 625-line sources at 6.75 and 3.375 MHz, respectively. These frequencies have a simple relationship to those in CCIR Recommendation 601.

Figure 2/H.261, p.

3.2 *Video source coding algorithm*

The video coding algorithm is shown in generalised form in Figure 3/H.261. The main elements are prediction, block transformation, quantization and classification.

Figure 3/H.261, p.

The prediction error (INTER mode) or the input picture (INTRA mode) is subdivided into 8 pel by 8 line blocks which are segmented as transmitted or non-transmitted. The criteria for choice of mode and transmitting a block are not subject to recommendation and may be varied dynamically as part of the data rate control strategy. Transmitted blocks are transformed and resulting coefficients are quantized and variable length coded.

3.2.1 *Prediction*

The prediction is inter-picture and may be augmented by motion compensation (§ 3.2.2) and a spatial filter (§ 3.2.3).

3.2.2 *Motion compensation*

Motion compensation is optional in the encoder. The decoder will accept one vector for each block of 8 pels by 8 lines. The range of permitted vectors is under study.

A positive value of the horizontal or vertical component of the motion vector signifies that the prediction is formed from pels in the previous picture which are spatially to the right or below the pels being predicted.

Motion vectors are restricted such that all pels referenced by them are within the coded picture area.

3.2.3 *Loop filter*

The prediction process may be modified by a two-dimensional spatial filter which operates on pels within a predicted block.

The filter is separable into one dimensional hironzontal and vertical functions. Both are non-recursive with coefficients of 1/4, 1/2, 1/4. At block edges, where one of the taps would fall outside the block, the peripheral pel is used for two taps. Full arithmetic precision is retained with rounding to 8 bit integer values at the 2-D filter output. Values whose fractional part is one half are rounded up.

The filter may be switched on or off on a block by block basis. The method of signalling is under study.

3.2.4 *Transformer*

Transmitted blocks are coded with a separable 2-dimensional discrete cosine transform of size 8 by 8. The input to the forward transform and output from the inverse transform have 9 bits. The arithmetic procedures for computing the transforms are under study.

Note — The output from the forward and input to the inverse are likely to be 12 bits.

3.2.5 *Quantization*

The number of quantizers, their characteristics and their assignment are under study.

3.2.6 *Clipping*

To prevent quantization distortion of transform coefficient amplitudes causing arithmetic overflow in the encoder and decoder loops, clipping functions are inserted. In addition to those in the inverse transform, a clipping function is applied at both encoder and decoder to the reconstructed picture which is formed by summing the prediction and the prediction error as modified by the coding process. This clipper operates on resulting pel values less than 0 or greater than 255, changing them to 0 and 255 respectively.

3.3 *Data rate control*

Sections where parameters which may be varied to control the rate of generation of coded video data include processing prior to the source coder, the quantizer, block significance criterion and temporal subsampling. The proportions of such measures in the overall control strategy are not subject to recommendation.

When invoked, temporal subsampling is performed by discarding complete pictures. Interpolated pictures are not placed in the picture memory.

3.4 *Forced updating*

This function is achieved by forcing the use of the INTRA mode of the coding algorithm. The update interval and pattern are under study.

4 Video multiplex coder

4.1 *Data structure*

Note 1 — Unless specified otherwise, the most significant bit is transmitted first.

Note 2 — Unless specified otherwise, bit 1 is transmitted first.

Note 3 — Unless specified otherwise, all unused or spare bits are set to ‘1‘.

4.2 *Video multiplex arrangement*

4.2.1 *Picture header*

The structure of the picture header is shown in Figure 4/H.261. Picture headers for dropped pictures are not transmitted.

Figure 4/H.261 [T1.261] (a traiter comme tableau MEP), p.

4.2.1.1 *Picture start code (PSC)*

A unique word of 21 bits which cannot be emulated by error-free data. Its value is under study.

4.2.1.2 *Temporal reference (TR)*

A five bit number derived using modulo-32 counting of pictures at 29.97 Hz.

4.2.1.3 *Type information (TYPE1)*

Information about the complete picture:

Bit 1 Split screen indicator. ‘0‘ off, ‘1‘ on.

Bit 2 Document camera. ‘0‘ off, ‘1‘ on.

Bit 3 Freeze picture release. Under study.

Bit 4 Under study. Possible uses include signalling of the use of motion compensation and the method of switching the loop filter.

Bit 5 Number of classes. ‘0‘ one, ‘1‘ four.

Bits 6 to 12 Under study.

4.2.1.4 *Extra insertion information (PEI)*

Two bits which signal the presence of the following two optional data fields.

4.2.1.5 *Parity information (PARITY)*

For optional use and present only if the first PEI bit is set to ‘1‘. Eight parity bits each representing odd parity of the aggregate of the corresponding bit planes of the locally decoded PCM values of Y, C_R and C_B in the previous picture period.

4.2.1.6 *Spare information (PSPARE)*

Sixteen bits are present when the second PEI bit is set to ‘1‘. The use of these bits is under study.

4.2.2 *Group of blocks header*

A group of blocks consists of 2k lines of 44 luminance blocks each, k lines of 22 C_R blocks and k lines of 22 C_B blocks. The value of k is under study.

The structure of the group of blocks header is shown in Figure 5/H.261. All GOB headers are transmitted except those in dropped pictures.

Figure 5/H.261 [T2.261] (a traiter comme tableau MEP), p.

4.2.2.1 *Group of blocks start code (GBSC)*

A word of 16 bits, 0000 0000 0000 0001.

4.2.2.2 *Group number (GN)*

An *m* bit number indicating the vertical position of the group of blocks. The value of *m* is the smallest integer greater than or equal to $\log_2(18/k)$. GN is 1 at the top of the picture.

Note — GBSC plus the following GN is not emulated by error-free video data.

4.2.2.3 *Type information (TYPE2)*

TYPE2 is *p* bits which give information about all the transmitted blocks in a group of blocks. The value of *p* is under study.

Bit 1 When set to '1' indicates that all the transmitted blocks in the GOB are coded in INTRA mode and without block addressing data.

Bits 2 to *p* Spare, under study.

4.2.2.4 *Quantizer information (QUANT1)*

A *j* bit code word which indicates the blocks in the group of blocks where QUANT2 code words are present. These blocks, their code words and the value of *j* are under study.

Whether QUANT1 is in the GOB header or the picture header is under study.

4.2.2.5 *Extra insertion information (GEI)*

Under study.

4.2.2.6 *Group of blocks global motion vector (GGMV)*

Under study.

4.2.2.7 *Spare information (GSPARE)*

Under study.

4.2.3 *Block data alignment*

The structure of the data for *n* transmitted blocks is shown in Figure 6/H.261. The values of *n* and the order are under study. Elements are omitted when not required.

Figure 6/H.261 [T3.261] (a traiter comme tableau MEP), p.

4.2.3.1 *Block address (BA)*

A variable length code word indicating the position of *n* blocks within a group of blocks. VLC code words using a combination of relative and absolute addressing are under study.

The transmission order and addressing of blocks are under study.

When bit 1 of TYPE2 is ‘1‘, BA is not included and up to 132k blocks beginning with and continuing in the above transmission order are transmitted before the next GOB header.

4.2.3.2 *Block type information (TYPE3)*

Variable length code words indicating the types of blocks and which data elements are present. Block types and VLC code words are under study.

4.2.3.3 *Quantizer (QUANT2)*

A code word of up to q bits signifying the table(s) used to quantize transform coefficients. The value of q and the code words are under study. QUANT2 is present in the first transmitted block after the position indicated by QUANT1.

4.2.3.4 *Classification index (CLASS)*

CLASS is present if bit 5 of TYPE1 is set to ‘1‘ and indicates which of the four available transmission sequence orders is used for luminance block coefficients. If bit 5 of TYPE1 is set to ‘0‘ then luminance block coefficients are transmitted in the default sequence order.

Chrominance block coefficients are transmitted in one sequence order.

The CLASS code words and sequence orders are under study.

4.2.3.5 *Motion vector data (MVD)*

Calculation of the vector data is under study.

When the vector data is zero, this is signalled by TYPE3 and MVD is not present.

When the vector data is non-zero, MVD is present consisting of a variable length code word for the horizontal component followed by a variable length code word for the vertical component.

Variable length coding of the vector components is under study.

4.2.3.6 *Transform coefficients (TCOEFF)*

The quantized transform coefficients are sequentially transmitted according to the sequence defined by CLASS. The DC component is always first. Coefficients after the last non-zero one are not transmitted.

The coding method and tables are under study.

4.2.3.7 *End of block marker (EOB)*

Use of and code word for EOB are under study. An EOB without any transform coefficients for a block is allowed.

4.3 *Multipoint considerations*

4.3.1 *Freeze picture request*

Causes the decoder to freeze its received picture until a picture freeze release signal is received. The transmission method for this control signal is under study.

4.3.2 *Fast update request*

Causes the encoder to empty its transmission buffer and encode its next picture in INTRA mode with coding parameters such as to avoid buffer overflow. The transmission method for this control signal is under study.

4.3.3 *Data continuity*

The prototocl adopted for ensuring continuity of data channels in a switched multipoint connection is handled by the message channel. Under study.

5 Vide data buffering

The size of the transmission buffer at the encoder and its relationship to the transmittion rate are under study.

Transmission buffer overflow and underflow are not permitted. Measures to prevent underflow are under study.

6 Transmission coder

6.1 *Bit rate*

The net bit rate including audio and optional data channels is an integer multiple of 384 kbit/s up to and including 1920 kbit/s.

The source and stability of the encoder output clock are under study.

6.2 *Video clock justification*

Video clock justification is not provided.

6.3 *Frame structure*

6.3.1 *Frame structure for 384-1920 kbit/s channels*

The frame structure is defined in Recommendation H.222.

6.3.2 *Bit assignment in application channel*

Under study.

6.3.3 *Time slot positioning*

According to Recommendation I.431.

6.4 *Audio coding*

Recommendation G.722 56/48 kbit/s audio, 0/8 kbit/s data and 8 kbit/s service channel in the first time slot.

The delay of the encoded audio relative to the encoded video at the channel output is under study.

6.5 *Data transmission*

One or more time slots may be allocated as data channels of 64 kbit/s each. The first channel uses the fourth time slot.

Positioning of the other channels, and possible restrictions on availability at lower overall bit rates are under study. The BAS codes used to signal that these data channels are in use are specified in Recommendation H.221.

6.6 *Error handling*

Under study.

6.7 *Encryption*

Under study.

6.8 *Bit sequence independence restrictions*

Under stydy.

6.9 *Network interface*

Access at the primary rate is with vacated time slots as per Recommendation I.431.

For 1544 kbit/s interfaces the default H_0 channel is time slots 1 to 6.

For 2048 kbit/s interfaces the default H_0 channel is time slots 1-2-3-17-18-19.

Interfaces using ISDN basic accesses are under study (see Recommendation I.420).