The CCITT Communication Protocol for Videophone Teleconferencing Equipment

Ralf Hinz
Daimler-Benz AG
Institut für Informationstechnik
Tel. 0731 / 505 - 21 32
Fax. 0731 / 505 - 41 04
Wilhelm-Runge-Str. 11
7900 Ulm

Abstract

During the last years, the CCITT has defined the communication protocol for videophone teleconferencing equipment. This paper gives a summary and a survey of the concepts and properties of this protocol.

1 Summary

In the last years, some effort has been made to define a protocol stack which can be used by different terminal equipment for audio, video, and data transmission. This paper gives a summary and a survey of the recommendations of the CCITT communication protocol for videophone teleconferencing equipment.

The described protocol has been defined during the last years by CCITT Study Group XV with cooperation of national working groups. In one of these working groups the author represented the Daimler Benz Research Institute. To get experience in the field of videophone teleconferencing, this institute has developed and produced functional models of ISDN videophones [MAY89].

2 Goal

The quality of transmitted pictures in a videophone session is strongly dependent on the bitrate which is available for the transmission of video information. Since the pictures become better with higher bitrates, the protocol has to be able to cope with

bitrates up to the range of some Mbit/s. On the other hand, it must be assumed that in the near future only 1 or 2 * 64 kbit/s connections (by means of narrowband ISDN) will be available for most of the endusers. For that reason, a protocol designed for both using 64 kbit/s connections, and for working on higher bitrates, too, is desirable.

The goal of the working groups was to define a protocol for audiovisual services (e.g. videophone), which is universal usable, i.e. not specialized for specific networks. Existing recommendations, such as G.704, X.30/I.461, etc. should be taken into account, existing hard- and software should be usable, and the realization of the protocol should be possible on simple microprocessors.

3 Result

The result of the activities is a series of recommendations. To sum up, some of the more important properties are listed below.

In contrast to e.g. the ISDN protocol (D channel), this protocol uses inband signalling, i.e. user data and protocol data are transmitted via one common channel. The protocol makes it possible to multiplex data (merge data from different sources and transmit them over one channel), as well as to split data (send data from one source over more than one channel). It is suitable for application both in networks synchronized centrally and decentrally, and it is designed to work in point-to-point and in point-to-multipoint connections.

The "basic set" of recommendations for this protocol is represented in the following list.

- H.200 - "Framework For Recommendations For Audiovisual Services". This is a survey of 45 recommendations, draft recommendations and titles for planned recommendations, which are defined or have to be defined for audivisual services [CCITT90f].
- H.320 - "Narrowband Visual Telephone Systems And Terminal Equipment". This recommendation describes the technical requirements of terminal equipment to be used for the narrowband videophone service. Narrowband means: The transmission rate is up to 1920 kbit/s [CCITT90a].
- H.221 - "Frame Structure For A 64 To 1920 kbit/s Channel In Audiovisual Teleservices". This recommendation defines the frame structure (syntax and semantics) used by the protocol [CCITT90b].
- H.242 - "System For Establishing Communication Between Audiovisual Terminals Using Digital Channels Up To 2 Mbit/s". This recommendation defines the procedural aspects of the protocol [CCITT90c].
- H.230 - "Frame-Synchronous Control And Indication Signals For Audiovisual Systems". This recommendation describes the exchange of control and indication

information between communication partners using the frame structure defined in H.221 [CCITT90d].

- H.261 - "Video Codec For Audiovisual Services At p * 64 kbit/s". This recommendation describes the coding and decoding algorithm for moving video. The factor p is in the scope of 1 to 30 [CCITT90e].
- G.725 - "System Aspects for the Use of the 7 kHz Audio Codec within 64 kbit/s" [CCITT88]. Some definitions in G.725 form a subset of H.221 and H.242.

For the time being, some of these recommendations are being taken over by ETSI (European Telecommunication Standards Institute) and are adapted to European conditions.

4 Videophone Terminal Equipment

The recommendation H.320 "Narrowband Visual Telephone Systems And Terminal Equipment" describes the technical requirements of terminal equipment for the videophone service with a data transmission rate of up to 1920 kbit/s. This recommendation defines the adequate terminal equipment and divides it into several logical blocks. These blocks are defined in their functional behaviour, and relations to appropiate CCITT recommendations (e.g. H.221, H.242, H.230, H.261 and the I.400 series) are given.

Other recommendations which describe terminal equipment suitable for communication with the protocol described in this article, are e.g. G.725 (System Aspects for the Use of the 7 kHz Audio Codec within 64 kbit/s) and H.261 (Video Codec For Audiovisual Services At p * 64 kbit/s).

5 Syntax of Protocol Data Units

Recommendation H.221 "Frame Structure For A 64 To 1920 kbit/s Channel In Audiovisual Teleservices" describes the syntax of protocol data units. The structure defined in this recommendation is not the same as it is used in common protocols like HDLC. Instead, the octet structure of the transmission channels is used for forming eight subchannels (see fig. 1). Each bit of an octet is part of a different subchannel: All bits #1 together build the first subchannel, ... all bits #7 together build the seventh subchannel. The eighth subchannel is called "service channel". It has a more extensive functionality than the other seven subchannels, and it has a special internal structure. Using an ISDN B channel with a transmission rate of 64 kbit/s, each subchannel

provides a transmission rate of 8 kbit/s.

Besides this "vertical" division, there is a "horizontal" division: 80 octets (i.e. 80 bits per subchannel) form a "frame", 16 succeeding frames form a "multiframe". A multiframe consists of eight "submultiframes" (SMF), each consisting of two consecutive frames.

The beginning and the end of this frame structure within the octet stream (the "frame alignment" and the "multiframe alignment") are introduced by the structure of the service channel. The alignment is necessary for the correct interpretation of received data. The information concerning the frame limits are coded within the service channel. For that reason, the bits containing this information are called "Frame Alignment Signal" (FAS, fig. 2). The FAS is constituted by the first eight bits of a frame within the service channel.

For it is also necessary to know, whether or not the communication partner has yet found the frame- and multiframe alignment, each partner uses one bit (the 'A'-bit) to signal his state of synchronization on the frame limits.

Another functionality provided by the FAS is the synchronization of multiple channels. Since a connection can be formed by up to six (ISDN B) channels, some more synchronization steps are necessary: The delay between the channels has to be handled, and the channels have to be treated in the correct order. For the first, a multiframe counter is used. By means of this counter, a relative delay of up to +/- 1.28 seconds between two 64 kbit/s channels can be egalized. For the latter, a channel numbering has been introduced.

Following the FAS, the next bits of the service channel form the "Bit-rate Allocation Signal" (BAS). The BAS is used for the transmission of information concerning the capabilities of a terminal, and it is used for signalling the mode in which the local transmitter is working.

Furthermore, the service channel contains the "Encryption Control Signal" (ECS). These bits may, in future, transmit information regarding encryption. If no information about encryption has to be transmitted, these bits contain user data. User data are also contained in bits 25 - 80.

6 Procedures of the Protocol

Recommendation H.242 "System For Establishing Communication Between Audiovisual Terminals Using Digital Channels Up To 2 Mbit/s" describes an inband signalling protocol. This protocol is used to establish, maintain, and disconnect a connection and to react on errors during these phases.

Since a connection can consist of up to six channels, there has been made a distinction between the channel which has been established first ("Initial Channel") and all other channels ("Additional Channel(s)").

Each channel has to send at least synchronization information by means of the FAS, and user data. The initial channel has some more work to do: Additional tasks before and during a communication are the exchange of capability information, sending of command codes and management of additional channels.

While a connection is established, data for control and indication as described in H.320 ("Frame-Synchronous Control And Indication Signals For Audiovisual Systems") can be exchanged.

6.1 Procedures

There are three basic sequences defined for the protocol which are used as building blocks for more complex procedures. All of them make use of the BAS. These basic procedures are:

- The capability exchange sequence for forcing the communication partners into a defined state, and for informing the communication partner about the own capabilities (fig. 3).
- The mode switching sequence for switching the receiver of the communication partner to a mode conforming to the own transmitter's mode (fig. 4).
- The frame reinstatement sequence for changing the transmission mode from an "unframed" format (i.e. without the FAS, BAS and ECS) back to a framed format (fig. 5).

When a communication between two or more partners starts, the partners have to find a common mode for operation. This common mode should make use of as much of the terminal's capabilities as possible, since only in this case all of their facilities can be used. For this, the mode initialization procedure has been defined.

If, during a connection, one of the communication partners wants to change his transmission mode (e.g. from voice to picture), he can do so by using the procedure for dynamic mode switching. It enables him to make use of different capabilities of the connected terminals.

Prior to disconnecting or to using special network services (e.g. call transfer), it is necessary for the connected terminals to work in a mode where 3.1 kHz PCM[1] audio signals can be decoded. For this case, the mode 0 forcing procedure has been defined.

In case of errors due to mode mismatch, the mode mismatch recovering procedure forces a mode reinitialization.

Also, a set of possbile error conditions and the reactions on them has been included into recommendation H.242.

---

[1] pulse code modulation

6.2 Relation to Network Signalling

The establishment and disconnection of channels cannot be achieved by means of the described inband signalling protocol. Instead, network signalling protocols have to be used. In case of an ISDN, this is the D channel protocol.

Prior to disconnecting a channel, the terminals are forced to work in the simplest mode ("mode 0") for being able to e.g. decode audio signals coming from a PABX[2].

7 Literature

[CCITT88] CCITT SGXVIII; Recommendation G.725: System Aspects for the Use of the 7 kHz Audio Codec within 64 kbit/s; Melbourne: July 1988

[CCITT90a] CCITT SGXV; Recommendation H.320: Narrowband Visual Telephone Systems And Terminal Equipment; COM XV-R 37-E, August 1990

[CCITT90b] CCITT SGXV; Recommendation H.221: Frame Structure For A 64 To 1920 kbit/s Channel In Audiovisual Teleservices; COM XV-R 37-E, August 1990

[CCITT90c] CCITT SGXV; Recommendation H.242: System For Establishing Communication Between Audiovisual Terminals Using Digital Channels Up To 2 Mbit/s; COM XV-R 37-E, August 1990

[CCITT90d] CCITT SGXV; Recommendation H.230: Frame Synchronous Control And Indication Signals For Audiovisual Systems; COM XV-R 37-E, August 1990

[CCITT90e] CCITT SGXV; Recommendation H.261: Video Codec For Audiovisual Services At p * 64 kbit/s; COM XV-R 37-E, August 1990

[CCITT90f] CCITT SGXV; Framework For Recommendations For Audiovisual Services; Geneva, 16 – 27 July 1990

[MAY89] May, Franz: Algorithmen und Realisierung eines ISDN-Bildtelefon-Codecs; ntz Bd. 42 (1989) Heft3, S. 130-133

[2] private automatic branch exchange

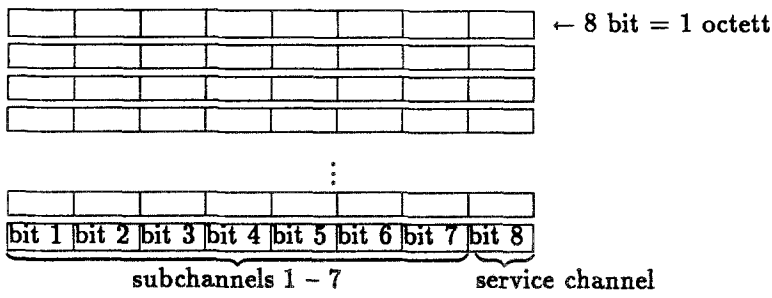

Fig. 1: H.221 subchannels

subchannel #1	subchannel #2	subchannel #3	subchannel #4	subchannel #5	subchannel #6	subchannel #7	subchannel #8 = service channel	octet number
							FAS	1 . 8
							BAS	9 . 16
							ECS	17 . 24
								25 . . 80

Fig. 2: H.221 frame structure

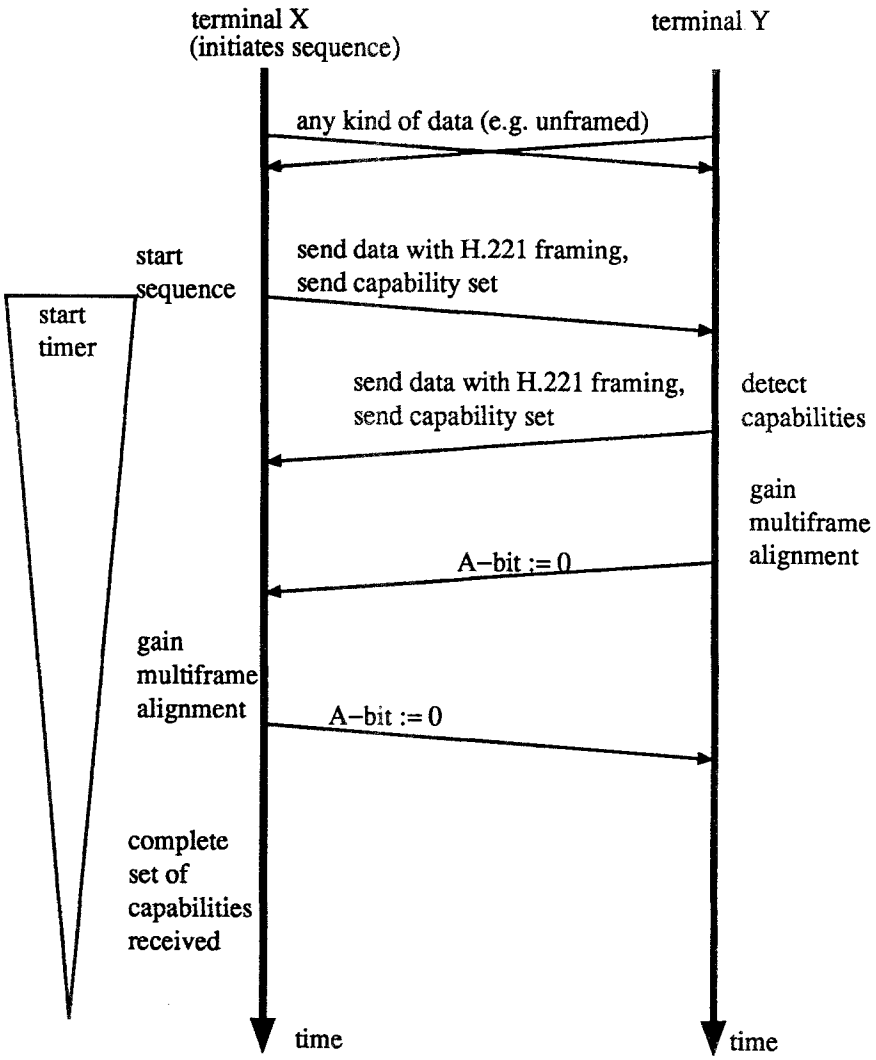

Fig. 3: Successful capability exchange sequence

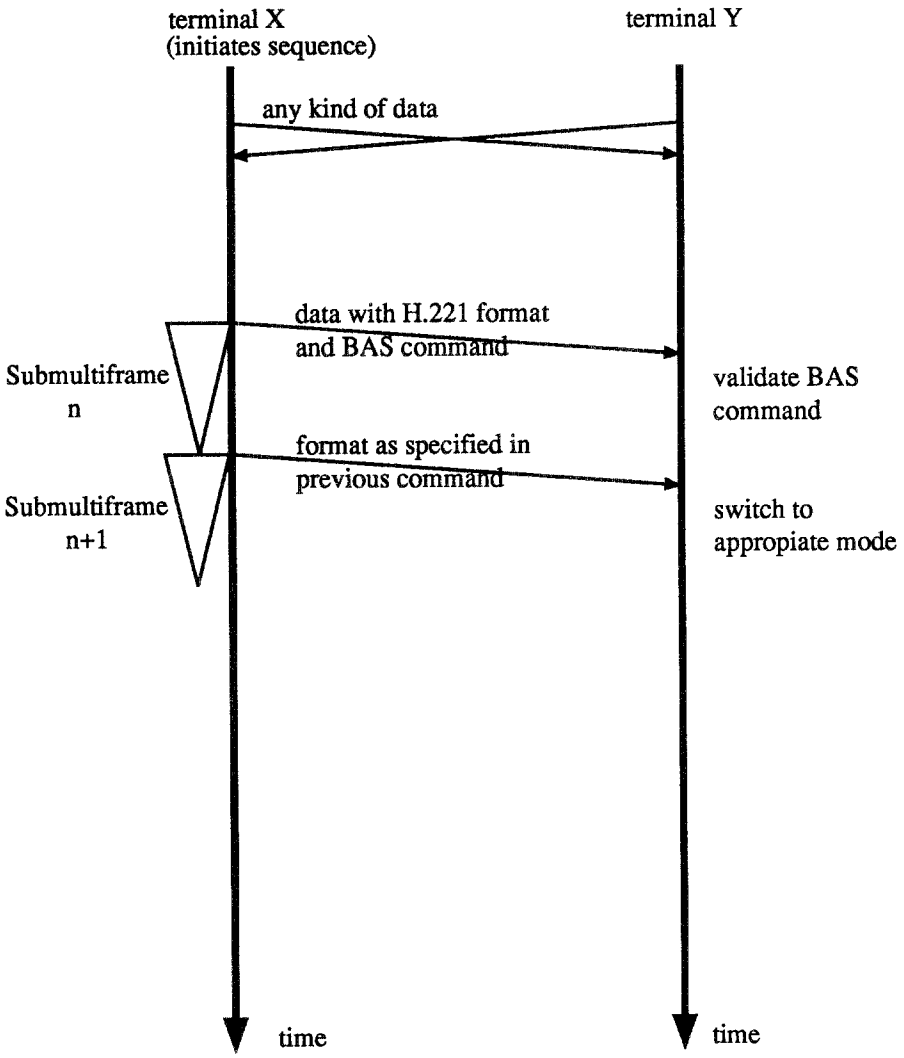

Fig. 4: Mode switching sequence

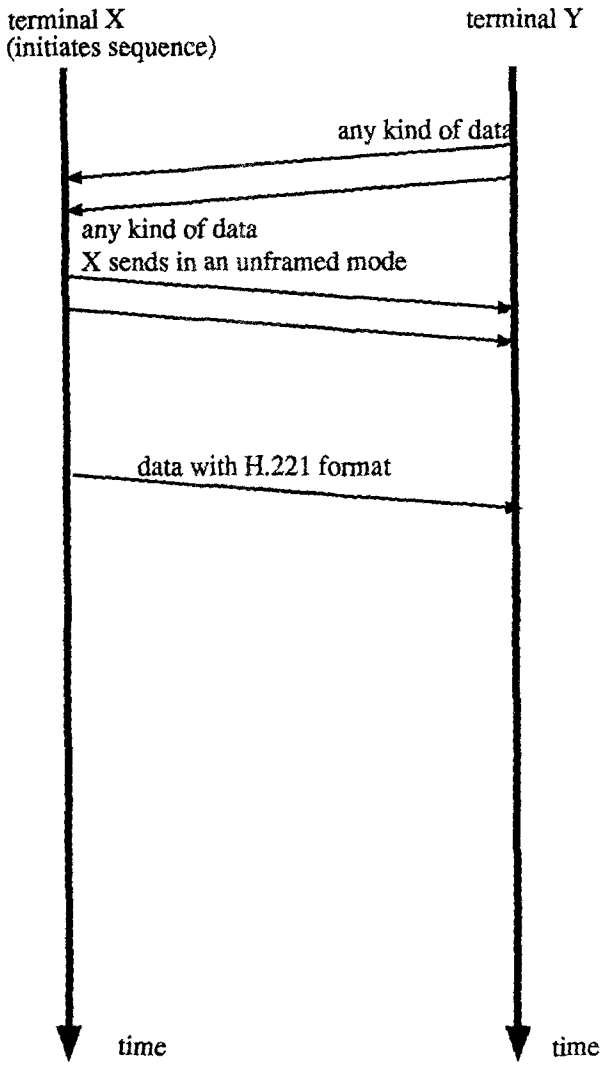

Fig. 5: Frame reinstatement sequence