International Telegraph and Telephone
Consultative Committee
(CCITT)
Period 1985-1988

Doc. # 253
Sept. 1987
Original: English

Question: XV/4 Specialists Group

STUDY GROUP XV - CONTRIBUTION

Source: PictureTel Corporation

Title: Frame structure for m x 56/64 kbit/s Video Telephony

Abstract

This document complements our earlier frame structure proposal (Doc. No. 241) with examples on the application to 56/64 kbit/s and 2 x 56/64 kbit/s. A comparison is also made with frame structure Y.221.

The current proposal has important advantages for multimedia applications, allowing flexible bit rate allocation for audio and data channels, and efficient channel utilization. It is applicable to any transmission rate, has a simple procedure to synchronize multiple channels, and is simple to implement. The proposal is compatible with other audiovisual services, and allows multipoint operation.

1 FRAME STRUCTURE PROPOSAL

The frame structure is based on a two-layer packet format presented in Doc. No. 241. The lower layer performs error correction and synchronization using a fixed length packet format. The higher layer multiplexes video, audio, data, codec C and I (Control and Indication) etc using variable length packets.

1.1 Bit Rates

The following types of videophone services have been defined within the CCITT Specialists Group:

Type 1: A single 56/64 kbit/s channel for video, audio, and data.

Type 2: Two 56/64 kbit/s channels; one is used for video and the other for audio.

Type 3: Two 56/64 kbit/s channels, shared between video, audio, and data.

Type 1 (or 3) can be extended to m x 56/64 kbit/s; values of m larger than 2 are mainly of interest in private networks.

1.2 Application Of The Frame Structure To Different Bit Rates

Type 1 and Type 3 utilize the proposed packet format. In pre-ISDN networks and in most ISDN implementations, Type 3 will require synchronization of two 56/64 kbit/s channels. The synchronization is performed by the lower layer; the procedure is described in the previous contribution. In all other aspects, Type 1 and 3 are identical. The audio rate can be fixed or variable, and data channels can also be defined to be synchronous or asynchronous. Actually, the format can be used at any data rate; it is not limited to 64 or 2x64 kbit/s.

Type 2 uses packet format only on the video channel. This makes it possible to use high speed asynchronous data channels for transfer of documents and graphics, e.g., 32 kbit/s. The audio channel uses one of the existing audio standards, e.g., wideband audio at 56 or 64 kbit/s.

2 COMPARISON BETWEEN Y.221 AND PACKET APPROACH

A comparison between Y.221 and two proposals based on a packet format is made in Doc. No. 247. Some further comments are given here based on the requirements stated in Doc. 247.

2.1 Compatibility With Other Audiovisual Services

Type 2 provides interworking with non-video terminals. The audio channel can use mu-law or A-law coded speech at 56/64 kbit/s. It can also use high quality audio according to CCITT G.722. The audio can occupy the whole channel, or audio and data can be multiplexed according to G.72Y.

2.2 Other Network Aspects

The proposed format can be used at any data rate; it is not limited to 64 or 2x64 kbit/s. This makes it possible to use multiples of 56 kbit/s as well as other bit rates. In many private networks, "smart" primary rate multiplexers can provide virtually any bit rate and also change the bit rate in the middle of a session. This is also possible with the proposed frame structure.

In comparison, Y.221 is not defined for bit rates other than 64 kbit/s. Its application to other bit rates, such as multiples of 56 kbit/s is not clear.

2.3 Efficient Channel Utilization

As indicated in Doc. 247, the lower layer packet format uses 2.7 % of the bit rate for framing and error correction. The higher layer typically uses between 1 and 2 % for overhead, depending on the average packet length.

Y.221 uses 2.5 % overhead, excluding error correction. It is not clear, however, whether any portion of the Y.221 service channel is available for video and audio information. It is undesirable to reserve 8 kbit/s (12.5 %) for framing and message channel.

The packet proposal makes it easy to define message channel and C and I as asynchronous channels; they do not need any bits allocated during most of the video session.

Y.221 can only change the bit allocation in multiples of 8 kbit/s, and the configuration can only change every submultiframe (80 ms). Hence, the granularity is 640 bits. This will give loss of efficiency for transmitting asynchronous data channels and audio information at anything else than multiples of 8 kbit/s. Hence, fewer bits are available for video coding.

2.4 Dynamic Bit Allocation For Video, Audio, And Data

The packet format supports dynamic bit allocation between audio, video, and data. It should be noted that most data transfers are of an asynchronous nature, such as graphics, documents, cursor information, PC screen dumps, codec C and I etc. Allowing a variable audio rate makes more bits available for the video, which will improve the quality of the service.

Y.221 can only change the bit allocation in multiples of 8 kbit/s, and the configuration can only change every submultiframe (80 ms). Since Y.221 uses a Bit Allocation Signal (BAS) that will be valid for the next submultiframe, it will take up to 160 ms (2 submultiframes) for a rate change to become effective. This would preclude efficient transmission of audio and asynchronous data without unacceptable delay.

2.5 Procedure To Synchronize Two 56/64 Kbit/s Channels

The packet format supports synchronization of two channels. It is important that this procedure is defined within the current study period and not left for further study.

2.6 Real Time Transmission For Low Bit Rate Data

Y.221 can implement synchronous channels without delay; however, most data channels used with videophone service are asynchronous. As pointed out above, Y.221 will give a delay of 160 ms for asynchronous data. This is significantly longer than the packet approach, where the packet length can be chosen to match the delay requirements.

2.7 Transmission Delay

2.7.1 Higher Layer - The higher level delay depends on the packet length. If a short packet length is used, video and audio are "interleaved", and the delay properties are similar to Y.221, see Fig. 1 a).

An interesting possibility with the packet format is that only one audio packet is sent per video frame, see Fig. 1 b). The audio is delayed one frame period in this scheme; however, it needs to be delayed anyway to achieve lip synchronization. Since the video data is available earlier, the decoded frame can be displayed earlier. If the video frame period is T, and the audio uses the fraction a of the total bit rate, the processing delay can be reduced by aT.

```
a)
           V   A   V   A   V   A   V   A   V   A   V   A   V   A
          |---|-|---|-|---|-|---|-|---|-|---|-|---|-|

b)
                      Video                          Audio
          |------------------------------|----------|

           <--------    (1-a)T    -------->  <-- aT -->
```

Fig. 1 Comparison of short and long packet lengths.

2.7.2 Lower Layer - When operating over a single channel, the packet format does not give any delay by itself. The error correction needs a delay of one block; the block length has been chosen equal to the packet length (255 octets). A trade-off must be done between block length and efficiency of the error correcting code. The proposal gives a delay of 32 ms at 64 kbit/s (packet period Tp = 255x8/64 ms). If the frame structure would be used on a channel with extremely low error rate, e.g., less than 10**-9, error correction is probably unnecessary, and no delay is incurred.

Y.221 assumes that error correction is performed at a higher layer. Since the information is spread out when multiplexed into the frame structure, the same error correction block length will give longer delay.

When two channels are used, some delay is incurred by the synchronization procedure. If the channels operate at the same rate, (i.e., no slip) the packets can be sent with a constant relative position, as indicated in Fig. 2. It can be shown that the synchronization needs a delay of Tp/2. The error correction gives an additional delay of Tp/2. Hence, the delay for two channels is the same as for a single channel: 32 ms for 2x64 kbit/s.

If the channels are not running off the same clock, the worst case phase relationship must be considered. This is when the packets on the two channels are in phase. In that case, the synchronization needs a delay of Tp. The error correction still needs Tp/2, resulting in a total delay of 1.5 Tp.

```
Channel 1:     |-------|-------|-------|-------|---

Channel 2:     ---|-------|-------|-------|-------|
```

Fig. 2. Phase relationship between packets

2.8 Error Correction

An error correction procedure has been defined. Syndrome and parity symbol calculation benefit from dedicated hardware for efficient implementation. Because of the hardware impact, the item should be agreed upon at the same time as the other elements of the frame structure.

Y.221 allows error correction at a higher layer, although with some reduced efficiency as pointed out in the previous section.

2.9 Multipoint Operation Using MCU

A multipoint conference unit (MCU) provides multipoint capability in a point-to-point network. The MCU performs audio bridging and selects the broadcaster depending on audio levels or explicit control information.

In multipoint operation, a fixed bit rate is allocated to audio. A low-rate synchronous data channel is also defined for multipoint control (message channel). It contains information exchanged between MCU and video terminals.

Multipoint operation can be performed for both Types 1, 2, and 3. In Type 2, it is possible to use audio bridging facilities developed for audio teleconferencing using mu-law, A-law, or G.722 audio.

2.10 Implementation

The proposed frame structure is byte-oriented to facilitate a microprocessor implementation.

The merging of video, audio, and data bitstreams bit by bit in Y.221 would require 10-30 times the number of instructions needed for the current proposal. Synchronization is also more complex, since the frame alignment signal is spread out over 8 octets.

3 PROBLEMS IDENTIFIED WITH Y.221

The following is a summary of the problems we have identified with the application of Y.221 to video telephony.

1. Not defined for bit rates other than 64 kbit/s, e.g., m x 56 kbit/s.
2. Not possible to change bit rate during a session (autobaud).
3. Inefficient transmission of asynchronous data.
4. Not possible to employ efficient audio coding with variable bit rate.

5. Multiplexing of multiple data side channels has not been defined.

6. Not clear whether any portion of the Y.221 service channel (8 kbit/s) is available for video and audio information.

7. Not clear whether synchronization of two channels will be supported.

8. No error correction procedure has been defined.

9. The format is bit- rather than byte-oriented (unsuitable for microprocessor implementation).