CCITT SGXV
Working Party XV/1
Specialists Group on Coding for Visual Telephony

Document #251
Sept.24, 1987

SOURCE: JAPAN
TITLE : FRAME STRUCTURES FOR m x 64 KBIT/S VIDEOPHONE

---

1. INTRODUCTION

The aim of this document is to sort out the requirements for transmitting information in the information transfer phase (Phase C) of a 1B videophone system. It also aims to clarify the problems associated with the present Y.221, and proposes methods of solving them. Specific frame structures which satisfy the system requirements are shown in Annex 2 and Annex 3.

2. REQUIREMENTS

As we have suggested in the companion document (Doc. #250), videophone communications phases can be divided into 5 stages. Phase C is the phase which performs the actual communication, therefore the transmission system in this phase must satisfy several conditions for efficient transmission in the case of 1B in particular. The essential conditions which cannot be satisfied by the present Y.221 are, we believe, as follows.

(1) No Wastage in Bit Rate Allocation

Videophone is multimedia communication, hence various types of data apart from voice and motion picture must be multiplexed and transmitted as shown in Table 1. From the viewpoint of the compression ratio of motion picture signals, 64 kbit/s is near the limit, so it is necessary to allot transmission capacity as much as possible to motion picture signals. Therefore various signals having different bit rates must be transmitted efficiently.

Y.221 basically multiplexes n x 8 kbit/s signals, and it is therefore difficult to efficiently transmit signals at other bit rates or low speed data signals. We shall assume here that data signals are first multiplexed by the 'controller', then they are transmitted through one data port as shown in Fig. 1.

(2) Rapid Changeover of Bit Rate Allocation

As we have mentioned, transmission capacities must be allotted to motion picture signals with consistently high efficiency. In order to be able to efficiently transmit burst data (e.g., from a personal computer or telewriter, etc.), it is necessary to change the bit rate allocation rapidly. As is shown in the Annex 1, the average bit rate allotted to motion picture signals in videophone can be increased by using a voice activation, in which case the changeover has to be made in about 10 - 20 milliseconds.

With Y.221 on the other hand, changeover of bit rate takes from 80 to 160 milliseconds, and a voice activation therefore cannot be used. Transmission efficiency for burst data is also poor.

3. METHODS OF SOLUTION

The following two methods can be envisaged for solving these problems.

<u>Method 1</u>
The basic structure of Y.221 is maintained, but modifications are added so as to satisfy the essential conditions (see Annex 2).

<u>Method 2</u>
Use of Packet Transmission (see Annex 3).

The problem with Method 1 is whether all the conditions can be satisfied. With Method 2, the problem is the cost of implementation of both Y.221 framing and packet transmission by a videophone (as to frame structures and communications procedures for AV terminals, see Table 6 in the companion document, Doc. #250).

4. PROBLEMS REQUIRING FURTHER CONSIDERATION

The following problems remain with regard to the modification of Y.221.

- Is the changeover time sufficient?
- Is the number of the types of channel arrangement sufficient?
- Cost of implementing both the ordinary Y.221 and a modified Y.221 in one terminal.

The following problems remain with regard to the packet system.

- Cost of implementing both the Y.221 transmission system used in Phase B (where no actual signal multiplexing is carried out), and the packet system used in Phase C.
- Cost of having two types of transmission system when communication with other AV terminals, e.g., G.722, is carried out with Y.221 frames.

As a conclusion, one of the two proposals must be adopted as the frame structure in Phase C for 1B videophone after giving due consideration to the above matters.

Further, it should be noted that there would be no problem if we apply the modification of Y.221 as in Method 1 not only to 1B videophone transmission, but also to 2B videophone and all AV terminal transmission systems. If this is done, some of the problems with Method 1 can be solved, and so it is obviously worth considering this idea. Since the BCH codes adopted in Method 1 have 8 bits for information, the BAS codes under consideration in G.72y etc., can be used without modification.

5. CORRESPONDING RECOMMENDATIONS

The frame structure in Phase C should be specified in Recommendation Y.221.

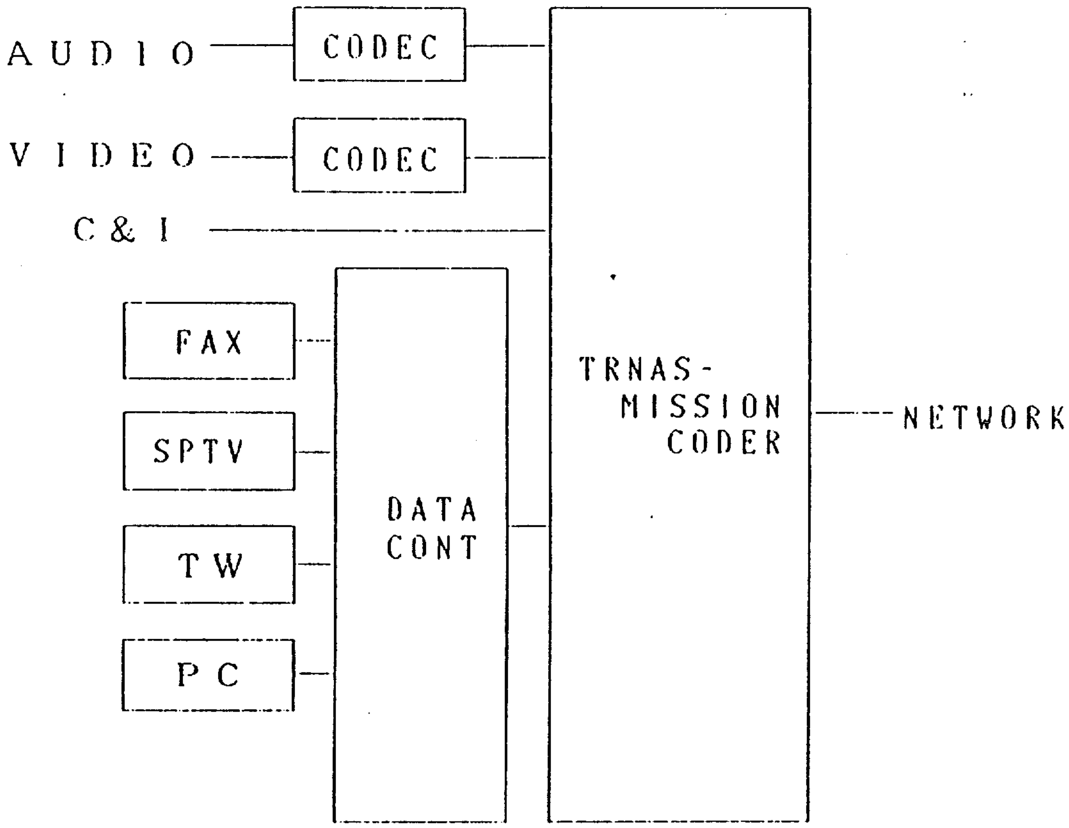
6. CONCLUSION

In this document, we have described the essential conditions required of the information transmission phase (Phase C) in a 1B videophone system, clarified the present problems associated with Y.221, and proposed 2 specific methods of solving them. It was concluded that either one of these 2 methods should be adopted after due consideration of the problems.

Table 1 Signals other than audio or video in videophone

Device	Maximum bit rate	Duration of 1 burst	Rreal time? (Permitted delay)	Relation with AV
FAX	4800 ~9600 b/s	1min	no	simultaneous
	~64 kb/s	4 ~ 5 sec	no	switched
Still picture	~64 kb/s	10 several sec	no	switched
Telewriter (Electronic black-board)	300 ~1200 b/s	several 10ms	yes (0.5 sec)	simultaneous
Pointer, cursor	300 ~1200 b/s	several 10ms	yes (0.5 sec)	simultaneous
Remote camera control	~100 b/s	several sec	yes (0.2~ 0.3 sec)	simultaneous
General purpose data (PC etc)	300 ~9600 b/s	several ms ~ several sec	yes (1 sec ?)	simultaneous
C&I FUR FPR	several b/s ~ several 100 b/s	several ms ~ whole the AV session	yes (several)	simultaneous

FUR : Fast Update Request
FPR : Frelze Pictare Request



C&I : Control and Indication
FAX : Facsimile
SPTV : Still Picture Television
TW : Telewriting
PC : Personal Computer
CONT : Controller

Figure 1 Data signal transmission in videophone

ANNEX 1

Effectiveness of Voice Activation in m x 64 kbit/s Videophone

In normal telephone conversations, both parties do not speak at the same time. Moreover, the voice is not generated during the whole time that party is speaking (voice ratio is less than 1); on average, the voice time occupancy ratio is less than 1/2. Therefore by assigning the voice signal bit rate to motion picture signals when there is no sound (this is known as voice activation), the motion picture bit rate can be increased. This method is particularly effective for 1 x 64 kb/s, as the compression ratio of motion picture signals is extremely high.

In low bit rate motion picture coding, picture elements are conditionally replenished based on interframe prediction. If the object moves rapidly, picture quality deteriorates and even after the object has stopped moving, some time is required for the quality to recover. By using voice activation, however, this time can be reduced, and as favorable conditions are maintained after recovery up to the next violent movement, average picture quality is also improved. Although dynamic variations of picture quality are unavoidable with low bit rate coding, they can thus be minimized by voice activation.

Speech chopping is a problem with voice activation, but in the case of videophone, they can be dealt with in the following way. The coding of motion picture signals normally takes about 200 msec, whereas it takes only 50 - 100 msec to process voice. If we are concerned with lip synchronization, therefore, we must add a delay of 100 -150 msec to voice. As shown in Fig. A-1, by inserting a delay circuit (approx. 50 msec) in the voice encoder, and by tapping the voice level detection signal earlier than the voice signal, therefore, speech chopping can be avoided. This is moreover easy to realize in practice.

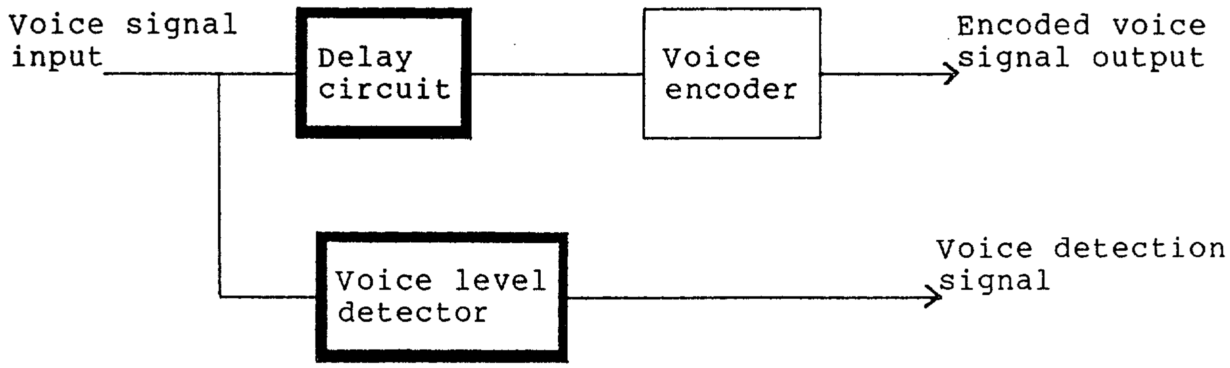


Figure A-1 Additional Circuit For Avoiding Voice Chopping

ANNEX 2

Method 1 (Modification of Y.221)

In the current Y.221, protection to transmission errors are carried out by changing BAS codes in a unit of submultiframe (8 frames), and making majority decisions (5/8 or more is taken to be correct). The changeover therefore requires no less than 80 msec and no more than 160 msec.

Instead of taking majority decisions in the direction of the time axis, however, the changeover can be made in a minimum of 20 msec and maximum of 40 msec by using 16 bits contained in 2 frames. Here, after proceeding to the videophone information transfer phase, the BAS coding is made by another method, i.e., by 2 frame units using a shortened BCH (16, 8) code* (double error correcting). Table A-1 shows the modes for 1B. In the videophone information transfer phase, either one of these BAS codes is always set.

Several types of data channel are prepared corresponding to the bit rates required by the data signals to be transmitted. Of these, data 4 is used for such media that transmits a large amount of information in a short time, as G4 facsimile and still pictures. In this case, video signal transmission is temporarily suspended.

If double error correction coding is used, and if the circuit random error rate is about 10^{-4}, the average interval between BAS code errors will be of the order of 10^4 hours, which in practice presents no problems. Further, since the BAS codes are interleaved every 8 bits, there is a some degree of tolerance to burst errors. The code which returns the system to the normal BAS mode can, for example, be inserted during whole the multiframe (8 frames) for further error protection.

* BCH (17, 9) is shortened to (16, 8).

Table A-1 Bit rate allocation for videophone modes
(for 1B type, number of bits per 10ms frame)

BAS code	FAS	BAS	AUDIO	VIDEO	DATA	C&I	Data rate (Note)
01	8	8	160	456	0	8	
02	8	8	160	464	0	0	
03	8	8	0	616	0	8	
04	8	8	0	624	0	0	
05	8	8	160	453	3	8	Data 1= 300 bit/s
06	8	8	160	461	3	0	Data 1= 300 bit/s
07	8	8	0	613	3	8	Data 1= 300 bit/s
08	8	8	0	621	3	0	Data 1= 300 bit/s
09	8	8	160	444	12	8	Data 2= 1200 bit/s
0A	8	8	160	452	12	0	Data 2= 1200 bit/s
0B	8	8	0	604	12	8	Data 2= 1200 bit/s
0C	8	8	0	612	12	0	Data 2= 1200 bit/s
0D	8	8	160	408	48	8	Data 3= 4800 bit/s
0E	8	8	160	416	48	0	Data 3= 4800 bit/s
0F	8	8	0	568	48	8	Data 3= 4800 bit/s
10	8	8	0	576	48	0	Data 3= 4800 bit/s
11	8	8	160	0	452	8	Data 4= 45.2 kbit/s
12	8	8	160	0	464	0	Data 4= 46.4 kbit/s
13	8	8	0	0	616	8	Data 4= 61.6 kbit/s
14	8	8	0	0	624	0	Data 4= 62.4 kbit/s
00	Transfer to normal BAS						

Note: Data bit rates show an example. Actual values and number of classes need further study.

ANNEX 3

Method 2 (Packet Transmission)

1. Frame Structure

In order to transmit different types of signals such as voice, motion picture, and data efficiently, each signal is multiplexed in units of a fixed length frame with a transmission header. No dummy bits are inserted between frames. When the circuit speed is Vc, the frame data length is Li, and the identification header length is Lh, the frame cycle Tf is therefore given by the equation:

$$Tf = (Li + Lh)/Vc$$

where Lh is 8 bits and Li approx. 640 bits.

2. Frame Type

There are 5 types of frames, viz. motion picture frames, voice frames, data frames, T1 frames and T2 frames. T1 frames are used when establishing initial synchronization; T2 frams are used when there is no other information to be sent, or for executing the mode changeover procedure. Under normal conditions, therefore, T2 frames must not be transmitted.

3. Frame Priority

On the transmission side, frames are sent out with the priority (1) voice, (2) data, (3) motion picture, and (4) T2 (Fig. A-2). In this case, the decision as to which frame should be sent next is made as late as the processing time permits.

4. Transmission Headers

Transmission headers consist of 8 bits. In order to decrease the probability of a header error, BCH (7,3) codes (single error correction, double error detection) are applied to the transmission headers except the T1 frame. One bit of the header can be used as a hook information bit if so required.

5. T2 Frames

T2 frames are used as (1) normal idle frames, or (2) mode changeover frames, depending on the identifier which immediately follows the transmission header. Fig. A-3 shows the structure of T2 frames.

6. Establishment of Frame Synchronization

Before beginning the communication, frame synchronization is set up between terminals (Fig. A-4).

Terminals first continuously send T1 frames, and simultaneously perform receive frame synchronization procedure. If the frame (T1 or T2) is received correctly m times in succession, receive frame synchronizasion is established, and a T2 frame is then sent continuously for a minimum of m frames. In this state, if a normal frame except T1 is received from the distant terminal k times in succession, frame synchronization in both directions is established (terminals therefore search for a T1 or T2 frame in an initial receive frame synchronization procedure). During frame synchronization procedure, only headers without error are permitted (i.e., BCH correction is not performed).

If the system starts from the Y.221 framing mode, this procedure is unnecessary, because frame synchronization already exists between terminals.

7. Monitoring and Recovery of Frame Synchronization

Terminals monitor synchronization by checking the transmission header in incoming frames. If the header causes a 1 or 2 bit-error j times in succession, the system is deemed to be out of synchronization. In such a case, synchronization is reestablished by searching for the header pattern, but in order to avoid false synchronization, BCH error correction is not carried out and only bit patterns without errors are searched. If a frame is received normally m times in succession, frame synchronization is deemed to have been reestablished. If the system remains out of synchronization for a time Ts or longer, the T1 frame is sent continuously. IF i or more T1 frames are received continuously during the communication, the distant terminal is deemed to have caused loss of synchronization. In this case, a T2 frame is sent out continuously as above instead of normal information frames, and the synchronization procsdure is executed from the initial state.

8. Mode Changeover

When the communication is over and the system returns to the normal Y.221 mode, a T2 frame which has a identifier of mode changeover is used.

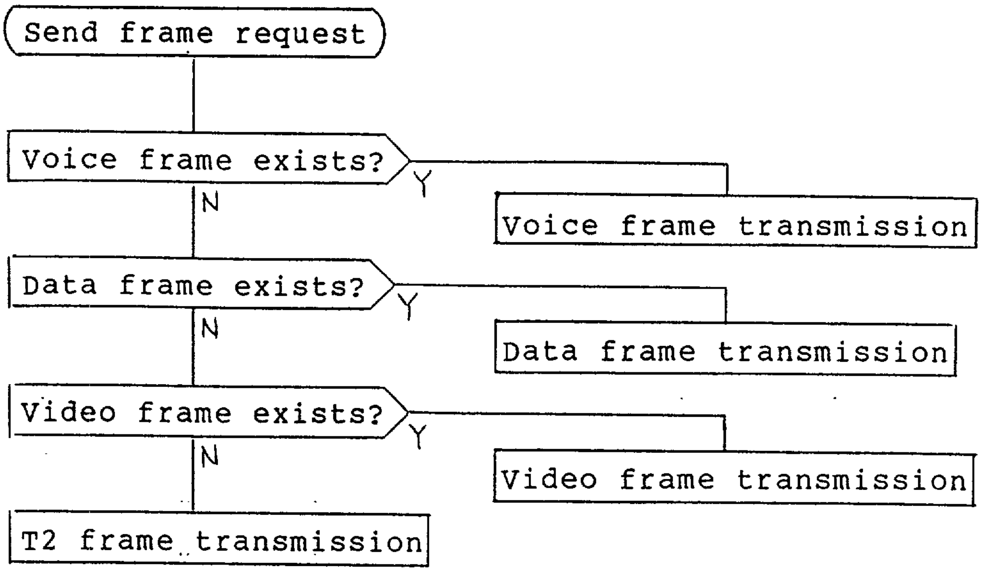


Figure A-2 Priority for sending transmission frames

Transmission Header	Identifier	Dummy bits

Figure A-3 Structure of T2 Frame

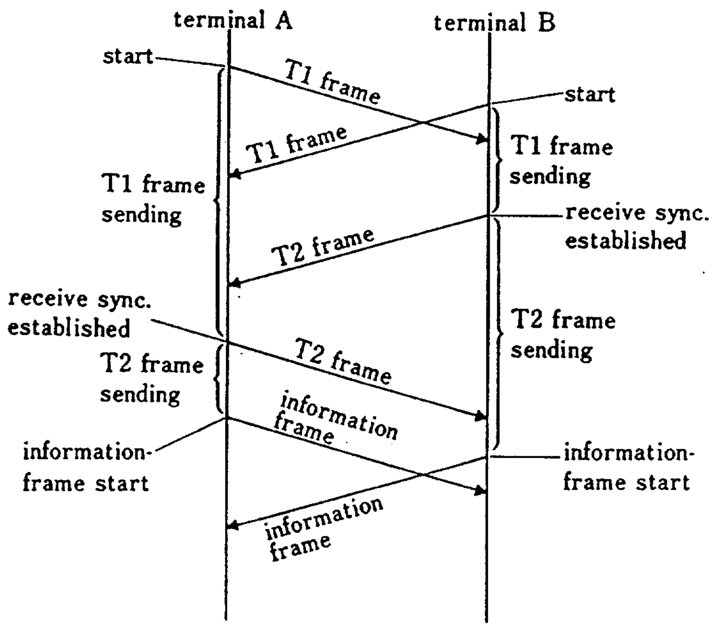


Figure A-4 Procedure for Initial Frame Synchronization