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

Document #555
November, 1989

Source : Japan
Title : Buffer Specification
(cf. Section 3.2 of Annex 4 to Doc. #540R)

1 Introduction

At the Stuttgart meeting, it was decided that a tentative buffer specification could be written based on a hypothetical reference decoder (HRD) (Doc. #540R,§8.1). This document discusses a clarification of the HRD, shows some examples of the implementation and proposes an additional specification.

2 Clarification of the HRD

The Stuttgart's description of the HRD has some unclear or unfixed points. Aiming to clarify the HRD operation, the HRD operation defines as shown in Annex.

2.1 Buffer size B

To decide a size of decoder buffer B, following points are discussed.

- The reason for including buffer B in Stuttgart's formula is that the rate of generation of coded bits is not constant but is permitted to vary.
- The size of the buffer is related to the delay time caused by data buffering. If the video channel rate is R (bit/s), the maximum delay time by buffering is B / R (sec).
- The decoder which have a large buffer will permits a large excursion and a large delay by buffering.

Taking into consideration, we propose that the size of buffer be an average number of coded bits per picture, that is

$$B = R \times k / 29.97 \text{ [bit]}$$

where $k / 29.97$ denotes the minimum decodable picture interval [sec]
R denotes the video channel rate [bit/sec]

In this case, the maximum delay time by buffering is $k / 29.97$ (sec).

2.2 Video channel rate

For video channel rate, the following point should be considered.

- During a communication, the video channel rate is not constant but is varied according to a switching of the Transparent User Data (TUD) or the Application Channel (AC).

In practice, the video channel rate is usually designed to be estimated higher, buffer size B also comes to setting larger. Under this condition, the decoder will operate without overflows.

3 How to control the encoder

Applicable schemes of the encoder are introduced in Appendix1 and 2. Those schemes don't need to evaluate the Stuttgart's formula several times to find the applicable value. Hence in practice, the decision of coded frame left to each encoder design.

4 How to control the decoder

An applicable scheme of the decoder is introduced in Appendix3.

5 Delay time

We also discussed to specify a delay time caused by data buffering or a time relation between the time coded frame stores in buffer and the time the reproduced frame display.

However, it is essential to specify the overall system delay, so the decision of delay time assignment within a system should be left to each decoder design.

6 Conclusion

This document proposes that the following additional specification should be written in section 5.2 of the final recommendation.

"Hypothetical reference decoder (HRD) is defined in Annex. Any encoder must not generate the coded data to let this HRD overflow. Any decoder could decode those coded data without the receiving buffer overflow. "

7 Annex

The definition of the hypothetical reference decoder

8 Appendices

Appendix 1 : An example of coding control method based on a buffer occupancy

Appendix 2 : An example of coding control method based on a jitter

Appendix 3 : An example of decoder design

END.

Annex to Document # 555

Hypothetical reference decoder (HRD) is defined below:

- The clock frequency of HRD and that of source coder are the same CIF rate (29.97Hz) and operates synchronously.
- The HRD receiving buffer consists of a fixed size of B and a variable size of X. The number of B bits is defined as;

  $$B = R \times k / 29.97 \quad \text{[bit]}$$

  where R denotes the video channel rate [bit/sec], k / 29.97 denotes the minimum decodable picture interval [sec] (k = 1,2,3,4).

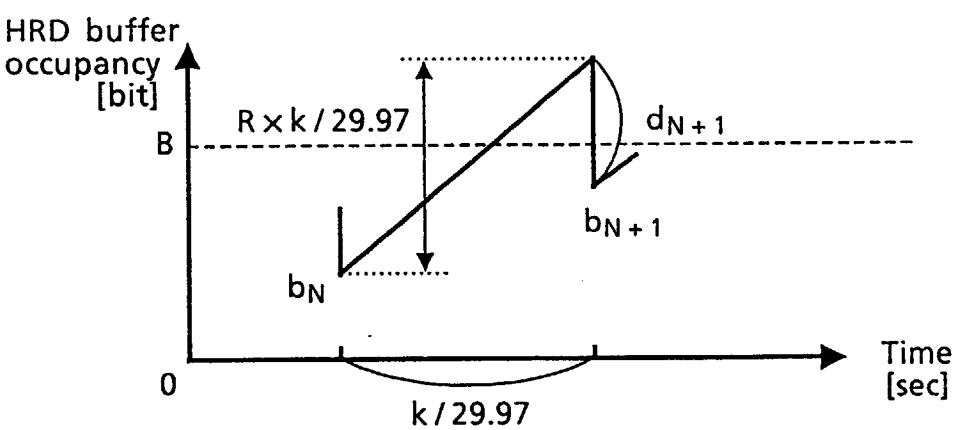
  The value of X is the number of bits of one coded frame.
- After a complete coded frames are stored in the receiving buffer, the HRD removed that coded frame instantaneously and decode it at the CIF rate.
- The coded frame is removed for more than the minimum decodable picture interval. In case that the coded data is very large, the HRD halts removing the next CIF timing until that frame is stored completely.
- The default value of HRD buffer occupancy is set to zero.
- If HRD buffer occupancy after removed the N th coded frame is described b_N, b_N is given by

  $$b_N = \sum_{n=1}^{N} (R \times k / 29.97 - d_n)$$

  where dn is the number of bits of the n th coding frame.
  To avoid HRD buffer overflow, the number of bits of the (N + 1) th coding frame d_{N+1} must satisfy (See below figure):

  $$d_{N+1} \geqq b_N + R \times k / 29.97 - B = b_N$$

  This is a requirement on the source coder output bit stream.



Appendix 1 to Document # 555

As an example of coding control scheme, a method based on the buffer occupancy is described below.

Definition of symbols

b_N[bit]	HRD buffer occupancy after decoding the N th coded frame
b_{N+1}[bit] ...	HRD buffer occupancy after decoding the (N+1)th coded frame
R[bit/sec] ...	video channel rate
k / 29.97[sec]	minimum decodable picture interval (k = 1, 2, 3, 4)
B[bit]	HRD buffer size ($B = R \times k / 29.97$)
d_{N+1}[bit] ...	number of bits of the (N+1)th coding frame

Buffer occupancy b_N is given by

$$b_N = \sum_{n=1}^{N} (R \times k \ / \ 29.97 - d_n)$$

where the default value of HRD decoder buffer occupancy b_0 is set to 0.

Coding control method

According to d_{N+1}, b_{N+1} is defined as follows :

① Ordinary case .... If d_{N+1} is satisfied next equation

$$b_N \leqq d_{N+1} \leqq b_N + R \cdot k / 29.97$$

then b_{N+1} is defined as

$$b_{N+1} = b_N + R \cdot k / 29.97 - d_{N+1}$$

② Case of HRD overflow If d_{N+1} is satisfied next equation

$$d_{N+1} \leqq b_N$$

then filler bits are inserted until $b_{N+1} = B (= R \times k / 29.97)$.

③ Case of HRD underflowIf d_{N+1} is satisfied next equation

$$d_{N+1} \geqq b_N + R \cdot k / 29.97$$

HRD cannot decode the (N+1)th coded frame for the past m /29.97 [sec] after decoded the N th coded frame. Then b_{N+1} is defined as

$$b_{N+1} = b_N + (R / P - d_{N+1}) + R \cdot m / 29.97$$

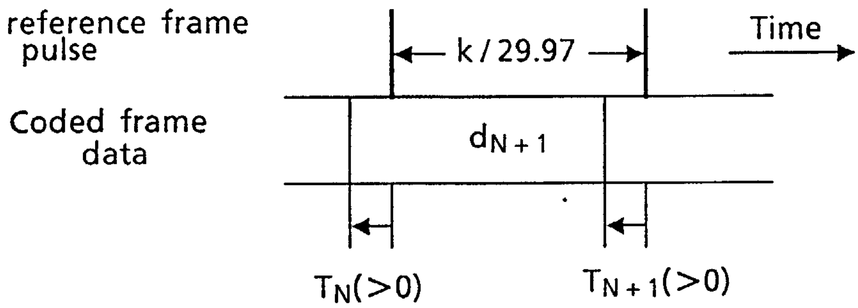
Appendix 2 to Document # 555

As an example of coding control scheme, the method based on the jitter between the reference frame pulse and the transmission timing of picture start code (PSC) is described below.

The reference frame pulses which show the timing to transmit the coded data are generated with minimum decodable picture interval $k / 29.97$ [sec], ($k = 1, 2, 3, 4$). The jitter is given by the next equation

$$T_{N+1} = T_N + \left(\frac{k}{29.97} - \frac{d_{N+1}}{R} \right) \qquad (1)$$

where T_i is the jitter for ith coded frame and $T_1 = 0$. If the transmission timing of PSC precedes a reference frame pulse, the jitter $T_i > 0$, which is illustrated as follows :



Coding control method

① In case of $T_N < 0$, the reference frame pulse is repeatedly shifted by CIF interval (1/29.97 [sec]) to satisfy the next equation

$$T_{N+1} = T_N + \left(\frac{k}{29.97} - \frac{d_{N+1}}{R} \right) + \frac{m}{29.97} > 0 \qquad (2)$$

where m is minimum value being $T_{N+1} > 0$.

② In the case of $T_{N+1} > Th$ (where Th defines k/29.97[sec]), dummy bits are transmitted until $T_{N+1} = Th$.

An example of this control method is shown in Figure 1.

For example, Th = 2/29.97[sec] and B = 2R/29.97[bit]

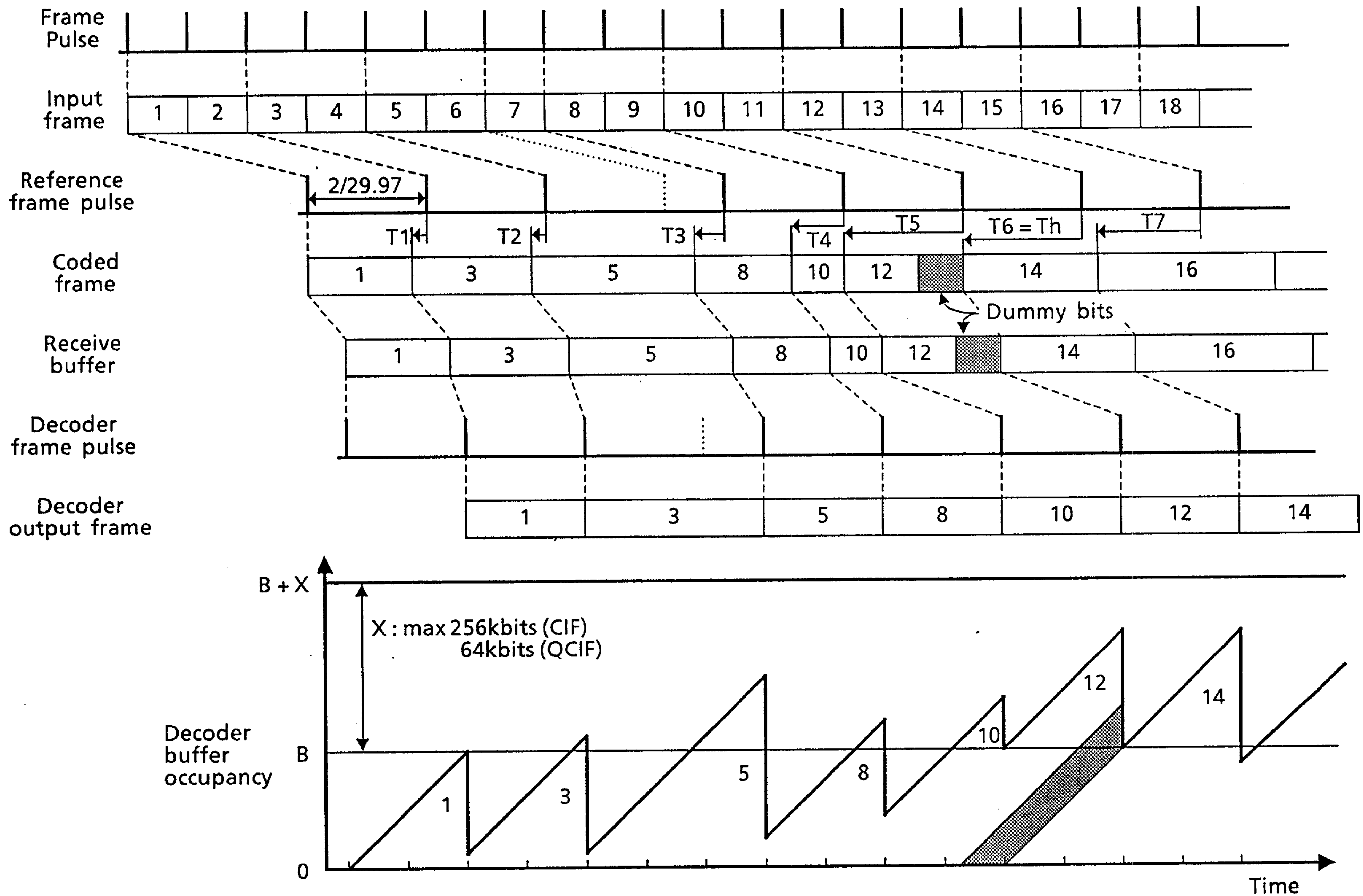


Figure 1. An example of coding contol scheme based on the jitter

Appendix 3 to Document # 555

An example of decoder scheme is described below.

- The receiving buffer is sufficiently large, e.g. 256 kbits + mR × k /29.97 bits.
- Source decoder clock frequency is higher than the CIF rate (29.97Hz) by a small amount, e.g. 1000 ppm.
- Source decoder works only after data for more than a complete coded picture is stored in the receiving buffer.
- Reproduced picture is stored in display buffer, and will be displayed referring to the Temporal Reference or immediately.