CCITT SGXV
Working Party XV/1
Experts Group for ATM Video Coding

Document AVC-46
May 23-31, 1991

SOURCE : JAPAN
TITLE : Considerations on Picture Format in ATM Networks
PURPOSE : Proposal

.....

1 INTRODUCTION

The first meeting of the Experts Group in November 1990 recognized it as a study item whether the idea of a single picture format is applicable to the new high quality video coding standard for B-ISDN (§7.4/AVC-22R). We have investigated possibilities of Super-CIF (subsequently called SCIF) which has a simple relationship with the CIF defined in Recommendation H.261.

We propose that the following items should be further investigated and confirmed toward defining an SCIF;

- Picture quality degradation due to format conversion between local television standard and SCIF is practically negligible.
- Coding efficiency is affected to a practically negligible extent by use of SCIF as the source coding input.
- Processing delay due to format conversion is reasonably small compared to the overall coding/decoding delay.

It should be noted that we used 720 pels per line, 288 lines per field, 59.94 fields per second as a platform to evaluate the SCIF concept, but that these parameter values do not form any part of this proposal.

2 ADVANTAGES OF SCIF

1) SCIF will permit all countries to communicate with each other without bothering to mind regional constraints of other countries, and there will be no need for exceptional signal processing for attaining mutual connectibility between regions or countries.

2) PRE processing and POST processing apart, universal commonality of hardware or software can be attained. This will favor volume production of hardware.

3) The multipoint system can be simplified. There will be no need to distinguish the terminals connected to MCU according to region.

4) The coding specification can be simplified.

5) Good conformity to CIF can be obtained.

3 COMMENTS OF THE SINGLE PICTURE FORMAT

If a single format such as SCIF is to be adopted for both intra-region and inter-region communications, the following problems must be studied for solutions.

① Picture Quality Degradation due to Format Conversion
It has been demonstrated by filter simulation that the deterioration of picture quality due to conversion from 525 lines to 625 lines and back to 525 lines is virtually negligible. Detailed relevant information is given in Annex 1, and the results of simulation will be demonstrated. However, experimental results have so far been unsatisfactory regarding the effects of the frame rate change on spatial resolution, reproduction of movements, and flicker. Detailed relevant information is given in Annex 2, and the results of

simulation will be demonstrated. Further studies are therefore required.

② Effects on Coding Efficiency

Deterioration of picture resolution due to the change of scanning lines and deterioration of temporal resolution due to the change of the frame rate may result in reduction in the coding efficiency, particularly Motion Compensation may be affected. It is therefore necessary to determine the effects of conversion to SCIF in coding operations. For instance, the effects of coding near the signaling rate of 4-9 Mbps should be experimentally determined and analyzed in two cases where the REC601 picture signal is coded as it is and the video signal is coded after conversion to SCIF format.

Detaited preliminary experimental results are given in Annex 3.

③ Processing Delay due to Format Conversion

Conversion to SCIF may involve multiple fields or frames. In this case, signal processing relative to time is inevitable, thus resulting in additional delay. It is necessary to evaluate the magnitude of delay and its disadvantages.

It will be a task of primary importance in the future action program to determine whether such adverse effects of conversion and coding can be held down to a negligible extent.

4 COMMENTS OF MULTIPLE PICTURE FORMATS

If the disadvantages in Par. ① through ③ of Section 3 are not acceptable, multiple picture formats will eventually have to be adopted.

In this case the use of a local format in the same region will eliminate the need of conversion so that picture quality will not be

compromised. On the other hand, however, the five advantages of SCIF listed in Section 2 can be lost or may at least be reduced.
The following two alternatives can be considered if multiple formats are to be adopted.

① SCIF for intra and inter-region communication, and LOCAL for intra-region communication
② LOCAL for intra-region communication and conversion to a suitable format on the receiver's end for inter-region communication

In the case of alternative ①, communication will be ensured if SCIF is selected. However, if high picture quality in intra-region communication is desired, high quality communication ensured by two modes, SCIF and LOCAL, would be a costly trade-off for settling for a single mode.

In the case of alternative ② format conversion is required only on the receiver's end, but the multipoint system will be so complicated that receivers will have to identify the regions of other parties.

It is another task of importance in the future action program to study pros and cons of alternatives ① and ②.

5 RELATIONSHIP BETWEEN PICTURE FORMATS AND CODING STANDARDIZATION

Coding standardization has the following as the main elements.

① Picture format
② Source coding and video multiplex coding
③ Transmission coding

It is desirable that coding and multiplex coding be common to all applications where possible. Picture formats, on the other hand, are selectable. For instance, SCIF may be chosen for communication, and

REC601 for other applications. Transmission coding can also have variations between ATM and STM environments.

6 CONCLUSION

A single standardized SCIF picture format would be ideal, but further studies are required to determine whether the picture format should be single or multiple, in consideration of various conceivable applications and what conditions the SCIF format should fulfil.

Annex 1. An Examination of the Line Number Convention.

The source picture sequence for the following process is 'Flower Garden'. Y, Cr, and Cb data were treated equally.

1 480 to 576

Each frame of the sequence underwent the following process and is displayed:

1. The frame is separated into odd and even fields (240 lines each), and both the top and bottom edges of each field are extended by one line by copying the respective edge line. The following process is carried out on each field equally:
2. A three tap transversal filter (filter 'a' below) with a nearly ideal low pass cutoff of 120 cycles per picture height is used to synthesize 288 lines in a 5:6 line phase ratio to the original 240. Only the the center 480 lines (of the total 576) are displayed.

2 576 to 480

Each field generated by the above process underwent the following steps, and the resulting sequences were displayed:

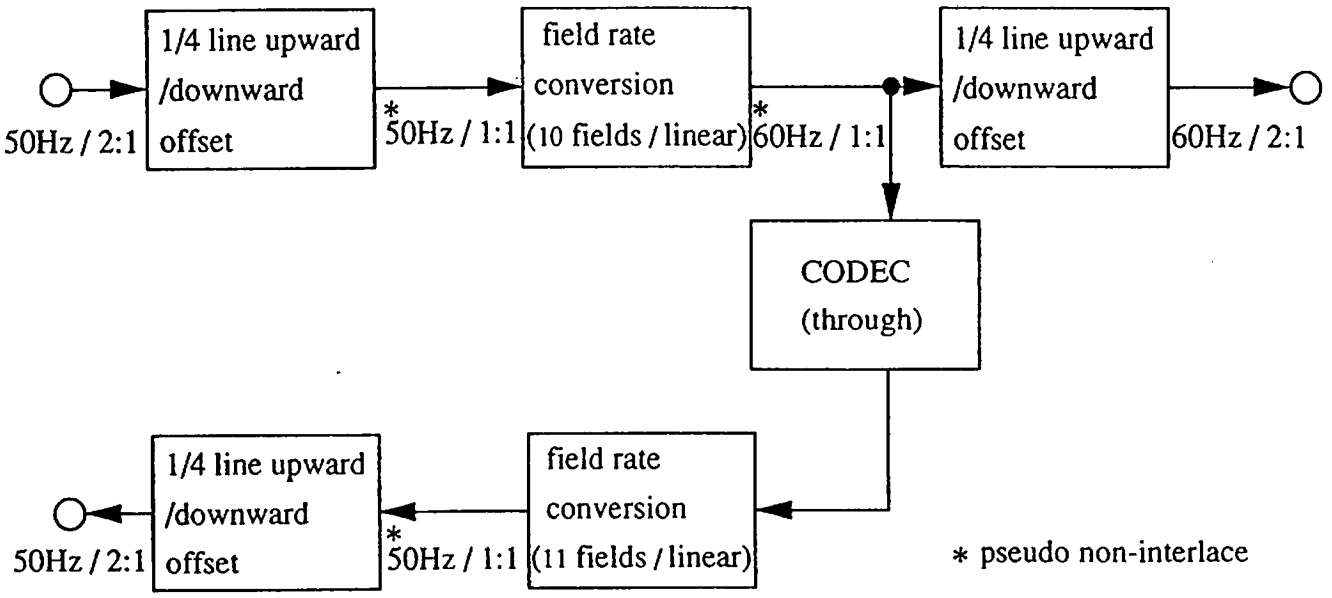
1. First, the three tap transversal filter 'b' with a cutoff of 144 cycles per picture height is used to reconstruct a 240 line field (6:5 line phase ratio).
2. Next, filter 'c' (cutoff at 120 cycles per picture height) is used similarly to filter 'b' in step 1.

$$a = \begin{pmatrix} -53 & 180 & 129 \\ -45 & 226 & 75 \\ -19 & 252 & 23 \\ 23 & 252 & -19 \\ 75 & 226 & -45 \\ 129 & 180 & -53 \end{pmatrix} \quad b = \begin{pmatrix} 154 & 154 & -52 \\ 92 & 214 & -50 \\ 28 & 251 & -23 \\ -23 & 251 & 28 \\ -50 & 214 & 92 \end{pmatrix} \quad c = \begin{pmatrix} 146 & 146 & -36 \\ 100 & 170 & -14 \\ 56 & 183 & 17 \\ 17 & 183 & 56 \\ -14 & 170 & 100 \end{pmatrix}$$

Coefficients of filters

Demonstration of field rate conversion

1) Configuration of field rate conversion



α_n is designed so that the ringing distortion becomes less than 3 per cent where the encoding/decoding total impulse response of the 1/4 line offset filter and field rate conversion filter is denoted as (... α_n ... α_2 α_1 1 α_1 α_2 ... α_n ...).

2) Demonstration list

(1) horizontal and vertical lines	original	25 sec
(2) horizontal and vertical lines	50 Hz to 60 Hz	25 sec
(3) horizontal and vertical lines	60 Hz to 50 Hz	25 sec
(4) "Flower Garden"	original	25 sec
(5) "Flower Garden"	50 Hz to 60 Hz	25 sec
(6) "Flower Garden"	60 Hz to 50 Hz	25 sec
(7) "Table Tennis"	original	25 sec
(8) "Table Tennis"	50 Hz to 60 Hz	25 sec
(9) "Table Tennis"	60 Hz to 50 Hz	25 sec

Annex 3 to Document AVC-46

Comparison of Coding Efficiency between 525 line and converted 625 line pictures based on RM8 coding

1. Introduction

Possibility has been pointed out that coding efficiency of converted 625 line pictures deteriorates compared with that of original 525 line pictures because of the increased number of pixels. A study of this problem is done using RM8 based coding.

2. Picture Format

Two kinds of picture formats are involved in this study. Coded areas of luninance are as follows.

-525 lines : 720 pixels × 240 lines × 59.94 fields, interlace

-625 lines : 720 pixels × 288 lines × 59.94 fields, interlace

(For chrominance, both pixels and lines are half of them.)

Intra-field linear 5 tap filters are used to convert 525 lines to 625 lines. For the reconversion from 625 to 525, intra-field linear 6 tap filters are used. Table.1 shows SNRs of 525 original signals and 525 converted-reconvertedsignals.

Here, "SNR of A and B" is defined by;

$$\text{SNR of A and B} = 10 \log \frac{255**2}{\text{MSE}}$$

$$\text{where, MSE} = \frac{\Sigma(B-A)**2}{\text{Number of pixels}}$$

Values presented are averages of 180 fields.

3. Simulation results

Following four cases are simulated using RM8 based coding.

Case 1 : simple coding of 525 lines. SNRs of 525 original and 525 coded-decoded are calculated.

Case 2 : coding of 625 lines converted from 525 lines. SNRs of original 525 lines and 525 lines of converted, coded, decoded and reconverted are calculated.

Case 3 : The same as case 2, except the rough exclusion of influence caused by line conversion-reconversion process.
SNR is calculated by following in this case.

$$SNR = 10 \log \frac{255**2}{MSE(total) - MSE(conv\text{-}reconv)}$$

where, MSE(total) : MSE cased by entire process, that is conversion, coding, decoding and reconversion.

MSE(conv-reconv): MSE cased by line conversion-reconversion only

Case 4 : coding of 625 lines. SNRs of converted 625 lines and coded-decoded 625 lines are calculated.

Average luminance SNRs vs total average transmission rate are plotted in Fig. 1～3. 180 fields are used for each case.

4. Conclusion

In Fig.1～3, curves of case 1 and case 3 almost adjoin. According to the RM8 based study, coding efficieny deterioration is relatively small in comparison with conversion-reconversion deterioration.

Table 1 Conversion Filter Characteristics

Sequence	SNR of Original 525 Lines and Reconverted 525 Lines
Susie	53.58 dB
Football	44.91 dB
Flower Garden	37.15 dB

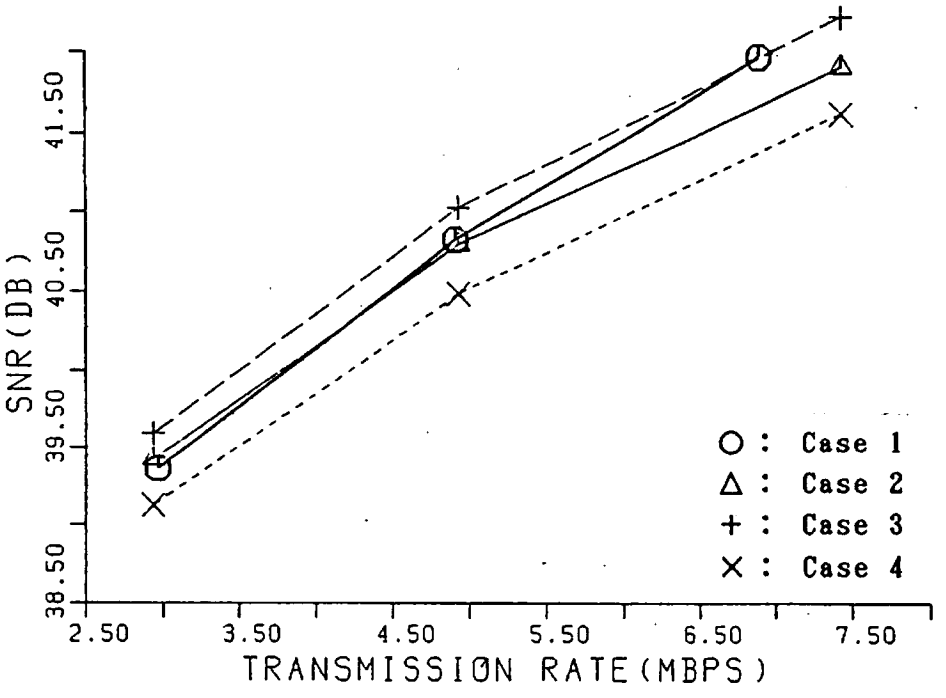


Fig.1 SNR - Rate Characteristics
Sequence: Susie

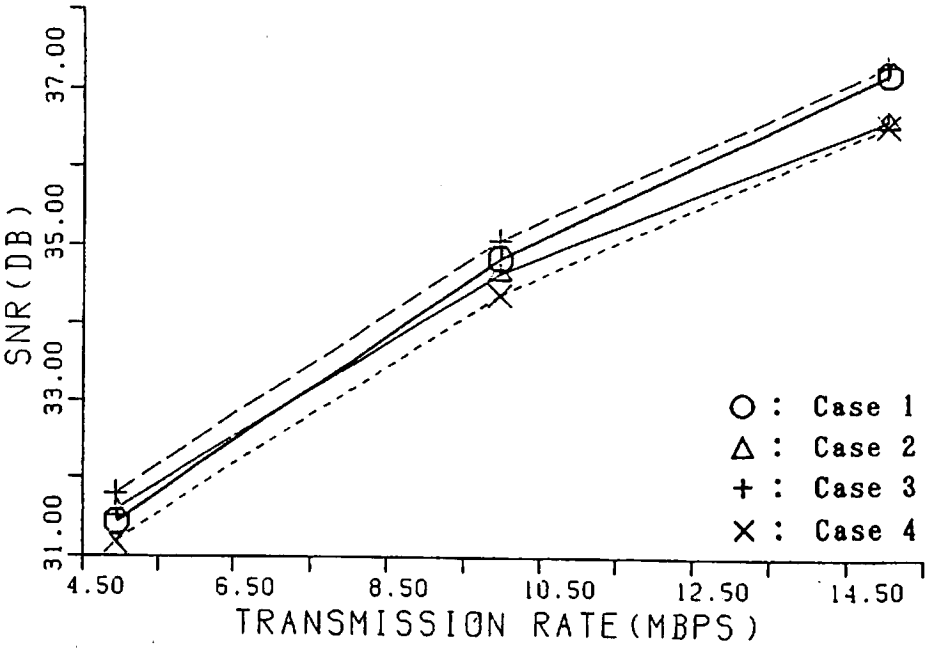


Fig.2 SNR - Rate Characteristics
Sequence: Football

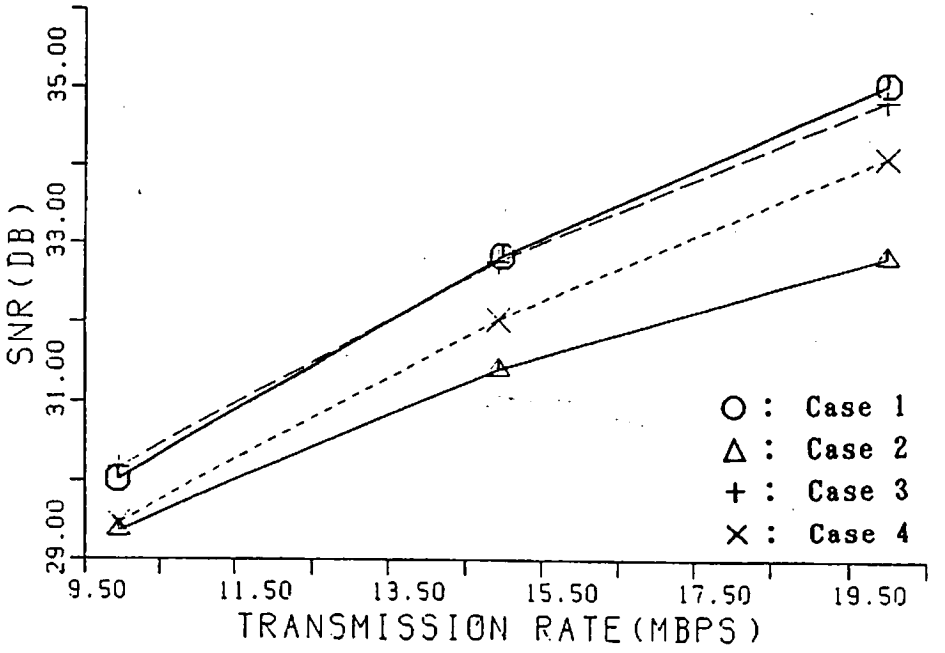


Fig.3 SNR - Rate Characteristics
Sequence: Flower Garden