Document #508 R
June 13,1989

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

SOURCE: JAPAN
TITLE: Improvement of picture in the facial region

1. Introduction

We previously introduced the idea of improving subjective picture quality by detecting the facial region using frame difference and by allocating more bits than other region in Paris (cf.Doc.#361). This time, a more promising result will be reported on this method.

2. Method

2.1 Summary

At the previous simulation shown in Paris, the quantization step size was determined not only from the result of facial region detection but also from buffer contents at each macro block. So, it was not necessarily assured that the SNR of the facial region will be improved for every coded frame when this method was adopted.

This time, we used a frame based coding control method and it can be assured that the SNR of the facial region will be improved for every coded frame.

2.2 Facial region detection

The facial region is detected automatically using frame difference and surrounded by a rectangle as before. The side information of the result of facial region detection is assumed to be sent by TYPE3 and QUANT2.

2.3 Coding control

There are many frame based coding control methods available, but this time we used the following imaginary method for simplicity.

1. First, RM8 simulation is tried out and the mean value of the step size for each frame is caluculated.(We call this step size a standard step size.)
2. Then, we encode the facial region with a smaller stepsize than the standard stepsize and the non-facial region with a larger stepsize than the standard stepsize.

3. Simulation

A comparison between the coding method with the standard step size in a frame and the proposed method is shown in Fig. 1 and Table 1 for the test sequence " Claire ".

Table 1 shows that the mean value of the total bits for the former was about 250 bits more than that for the latter, but this value did not contain the extra bits of TYPE3 and QUANT2. These extra bits are (18 MB line * 2

edges) * (4:extra bits for TYPE3 + 5:extra bits for QUANT2) = 324 bits at the maximum and are thought to be about half of this on the average. Taking this into account, the mean values of the total bits for the two coding method are thought to be almost the same.

Fig. 1 shows the SNR of the facial region and the SNR of the other region for the two coding methods. The solid line shows the SNR of the proposed method and the dotted line shows the SNR of the coding method with the standard step size in a frame.

This result shows that about 1 dB improvement in the facial region SNR is achieved at the expense of about 2 dB degradation in the other region for all coded frames when we use the proposed method.

It will also be demonstrated by VTR that such features as eyes, nose, and mouth are more clearly displayed by the proposed method. Even if the picture quality of the other region is somewhat worse by the proposed method, it may be said that the total subjective picture quality is better by the proposed method when we think of conversation by looking mainly at the facial region.

4. Conclusion

We have improved the coding method reported in Paris and demonstrated a new possibility of improving the subjective picture quality by allocating more bits to the facial region.

END

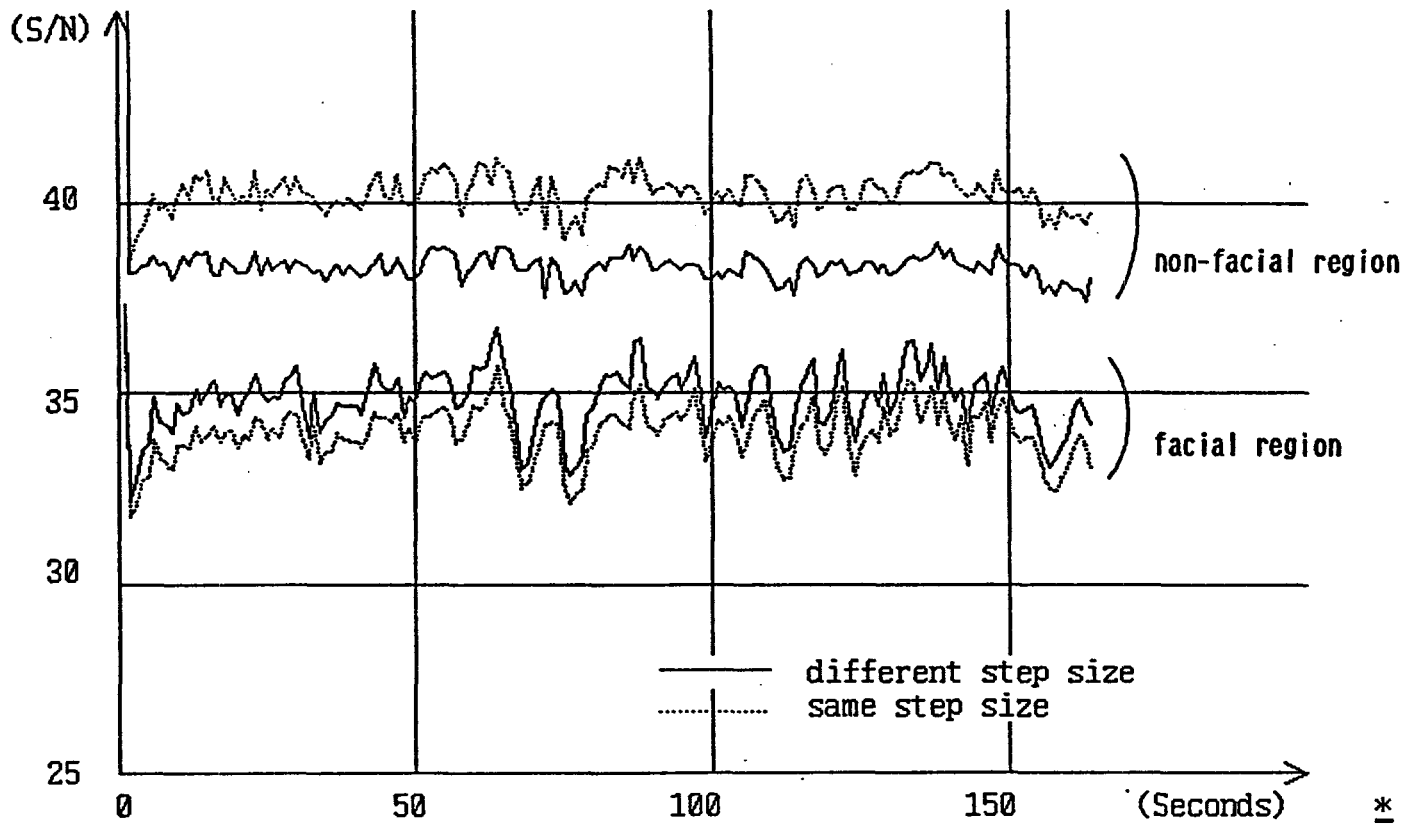
(S/N)
40
35
30
25
0
50
100
150
(Seconds)
non-facial region
facial region
different step size
same step size

Fig. 1

INSTITUTE TOSHIBA

STATISTICS RH8 with frame based coding control
SEQUENCE : claire
MODIFICATION : facial region detection

DATE: June 13, 1989

FRAME RATE : 10Hz

ITEM			same step size	different step size
1. RMS for luminance			2.99	3.36
2. SNR for luminance			38.65dB	37.62dB
for chrominance (U)			39.05dB	38.38dB
for chrominance (V)			41.74dB	41.69dB
3. Mean value of step size			17.8	
4. Mean value of the number of non-zero coefficients			2.9	3.1
5. Mean value of the number of zeroes before the last NZ-coefficient			5.5	5.5
6. Block type of MACRO	FIXED		278.5	301.8
	CODED MC		46.1	44.9
	FIXED MC		6.5	10.5
	CODED		64.7	38.7
	INTRA		0.1	0.1
	MC + Q			
	CODED + Q			
	INTRA + Q			
7. Block type of Y	FIXED		1269.9	1293.3
	CODED MC		115.5	119.6
	FIXED MC		95.3	102.2
	CODED		102.9	68.5
8. Block type of UV	FIXED		759.2	759.3
	CODED		32.7	32.6
9. Number of bits	Macro attributes		1061.9	897.2
	End of block		503.7	442.7
	Motion vectors		334.5	362.1
	Coefficients	Y	3404.6	3333.6
		U	137.6	138.9
		V	54.8	70.9
		Total	3597.0	3543.4
	Total		5503.3	5250.9

Table 1