Document #361
September 19,1988

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

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

1. Introduction

By incorporating into the coding process knowledge about a person's location in the picture, we can improve picture quality in subjectively important regions using only a few extra overhead bits.

2. Method

2.1 Improving picture quality in the facial region

First, we assume that the subject in the picture is a person. We then use the frame difference signal to estimate the location of the persons face. Finally we allocate a greater number of bits to the facial region yielding improved picture quality in this area.

2.2 An example of the facial region detection method

There is a need for more study on the facial region detection method, but Fig. 1 gives one example.
This method consists of the following steps:

(1) Divide frame difference picture into fixed sized blocks
(2) Count the number of pixels which have a frame difference greater than the 1st threshold, th1
(3) From the top, left and right sides of the picture respectively, search for the first block whose pixel count from step (2) is greater than the 2nd threshold, th2
(4) Determine these 3 blocks as the top, left and right sides of the facial region
(5) Determine the height of the facial region by
height = width x fixed ratio

The facial region is a rectangle, and so expressed by 3 coordinates (x1,x2,y1 in Fig. 1).
When using RM5, we detect the facial region with macro-block accuracy.

2.3 Coding control

Bit allocation control is done by multiplying the quantizer step size by a coefficient larger than 1 when coding blocks outside the facial region.

3. Simulation

A comparison between RM5 and the proposed method is shown in Fig. 2 and Fig. 3. Fig. 2 shows the SNR of the facial region and Fig. 3 shows the SNR of the non-facial region.
These results show that about 1 dB improvement in the facial region SNR is achieved at the expense of an SNR degradation in the non-facial region. Total SNR is a little bit lower in this method but the

subjective impression is better.

4. Conclusion

We have demonstrated the possibility of detecting the facial region and using this information to improve subjective picture quality over that of RM5.

In the present method we have treated the facial region as a rectangle, so the bits required to express it are few. We propose that some means to send this extra information with each frame should be taken into account.

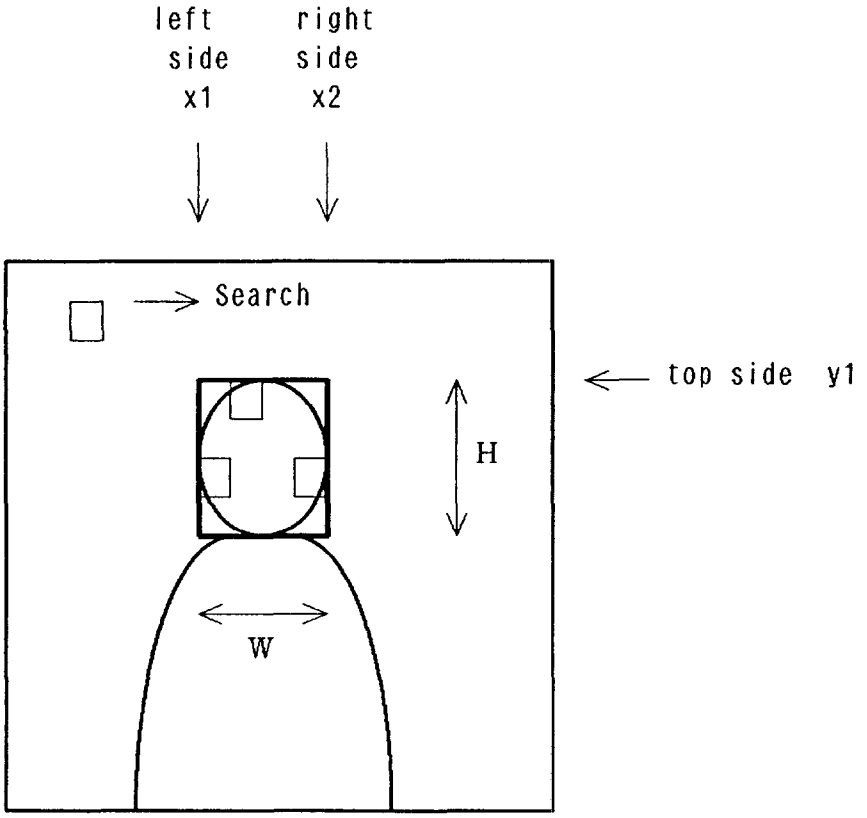
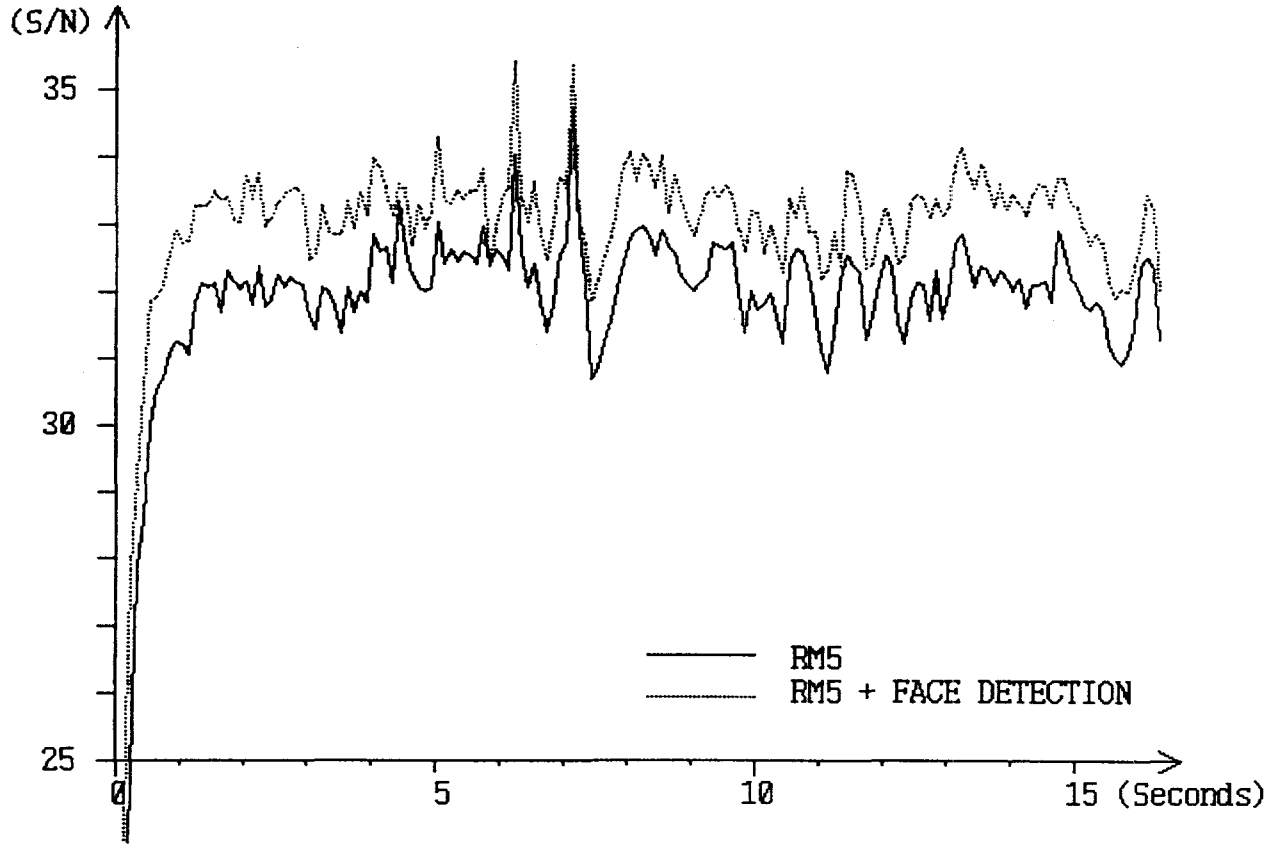
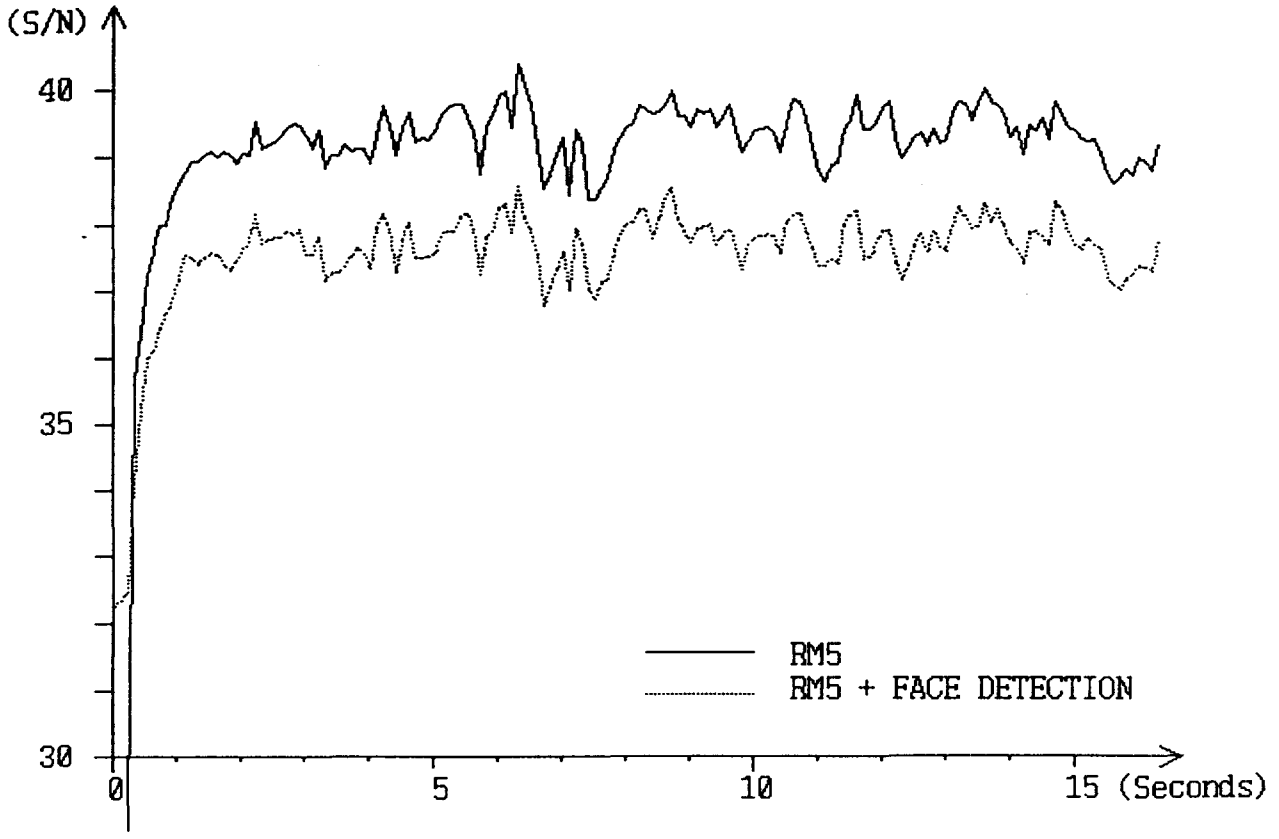


Fig.1 Face region detection


(S/N)
35
30
25
0
5
10
15 (Seconds)
RM5
RM5 + FACE DETECTION


(S/N)
40
35
30
0
5
10
15 (Seconds)
RM5
RM5 + FACE DETECTION