International Telegraph and Telephone
Consultative Committee
(CCITT)

Document N° 71
English only
Date : January 1986

Question 4/XV
Specialist group on coding
for visual telephony

TITLE : **Examples of simulation with vector quantization scheme (VQ)**

SOURCE : **France**

The french simulation results, presented during the meeting in Ipswich, have been achieved according to the coding scheme depicted in document n° 70 "VQ BASED CODER", and the following parameters have been used.

Picture format

COST sequence

Luminance : 360 X 288 Chrominance : 90 x 72
10 pictures/second (intermediate format subsampled 3 to 1).
15 pictures/second in the "TREVOR WHITE" part.

Miss America

Luminance : 360 x 288 Chrominance : 90 x 72
15 pictures/second (intermediate format subsampled 2 to 1).

Checked Jacket

Luminance : 360 x 288 Chrominance : 90 x 72
15 pictures/second (intermediate format subsampled 2 to 1).

Bit-rate

304 kbit/s = 256 kbit/s (Y) + 24 kbit/s (U) + 24 kbit/s (V).

Block-size

6x6 for luminance.
3x3 for each chrominance component.

Buffer size

The size of the buffer is 20 Kbits.

Standard conversion

- Framerepeat from to 10 Hz to 30 Hz for the COST sequence.

- Frame repeat from 15 Hz to 30 Hz for all the sequences.

- Linear interpolation from 30 Hz to 50 Hz according to BT-NTT document n°55.

Preprocessing

Temporal noise reduction.

Postprocessing

Spatial noise reduction.

FRANCE

VQ CODING SIMULATION (Presentation)

1- Pre-processing

The original sequence is temporaly sub-sampled ~~by a ratio of 1/4~~. It is then temporaly filtered. The definition is reduced to 360x288 for each luminance field, and 90x72 for each chrominance component.

2- Codebook generation

2.1- Blocking

Each luminance field of the training set is divided into 6x6 blocks. This size was chosen according to the aimed bit rate of 320 Kbit/s. It leaves us some 11 bits to code each block.

Each chrominance field is divided into 3x3 blocks. Each chrominance block corresponds to 4 luminance blocks.

2.2- Luminance vectorization

The training set is divided into 4 classes before codebook generation. First we ignore the uniform blocks according to the value of their variance. Then we classify the remaining blocks into 3 shape-classes, using an algorithm similar to the method proposed by Gersho [1]. We obtain a set of vertical oriented blocks, a set of horizontal oriented blocks, and a set containing the other blocks. We compute one codebook for each set, excluding the uniform class. Before generation, we normalize each block by dividing it by its mean value.

2.3- Chrominance vectorization

The same distinction is made on the chrominance blocks between uniform and shape blocks. There is no normalization but the values are magnified before computation. Different codebooks are generated for U and V component.

2.4- Codebook generation

We use a tree-structured generation algorithms, in order to decrease the amount of computations. At each level of the codebook, we separate each class of the training set by cutting the space into two parts with an hyperplan. This method exploits the vectorial structure induced by the Euclidian distance, and is independant on the training set.

3- Codec structure

The general codec structure is shown in fig.1. Vector quantization is performed on the video input, the decoded result is kept in a memory, and conditionnal replenishment is used. A buffer of 20 Kbit is used for transmission. Furthermore when the buffer goes empty we send the real values of the blocks which have been decided fix for the most long time.

More detailed information on the vector quantization procedure is provided in fig.2. The main characteristics are :

- The class structure defined during the codebook generation is not taken into account during codebook scanning. The three shape-codebooks are scanned for each block, in order to achieve a better coding accuracy.

- The regulation parameters are the variance threshold, the DC-component threshold, and eventually the codebook scanning depth. The variance threshold controls the ratio of uniform blocks in the picture: for these blocks we need only to transmit the DC-component. The DC-component threshold controls the ratio of fixed blocks in the picture: a block is decided to have changed if its DC-component has changed, or, if it has not, if its codebook address has.

- Chrominance coding of a block uses the variance information of the 4 corresponding luminance blocks to decide if the block is uniform. It uses the replenishment information of these 4 blocks to decide if it has remained unchanged.

4- Other characteristics

Temporal sub-sampling is performed when the ratio of uncoded blocks in the picture becomes too high.

The change of scene is automatically detected and the codebook changed, yet the codebook transmission time has not been simulated.

Fig.3 gives us some indications about the codebook scanning activity. The branches of the tree are figured in dashed lines if they have been scanned, in full line in the other case.

5- Reference

[1] Image Vector Quantization with a perceptually-based classifier
Ramamurtha & Gersho
IEEE 1984

[2] H. Van Helden PCS 84

Fig.1 General Codec structure

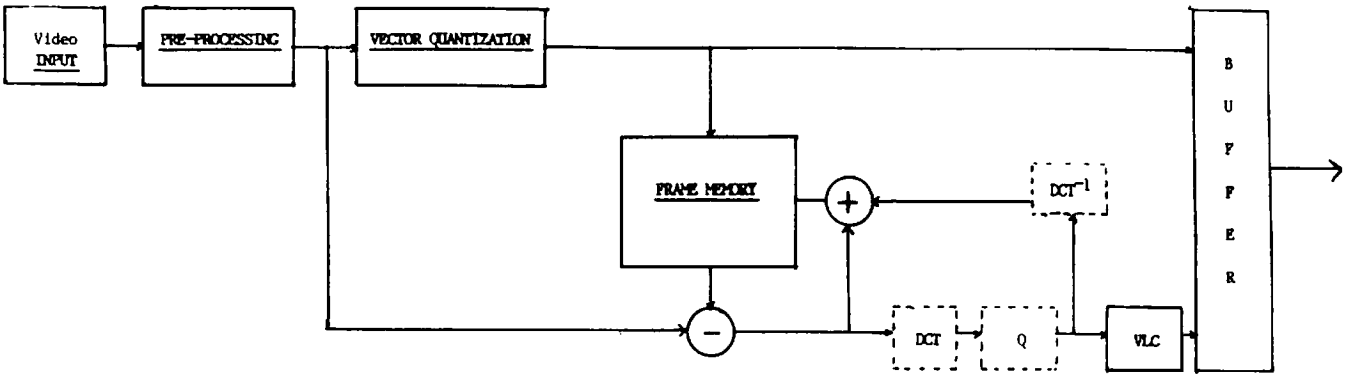
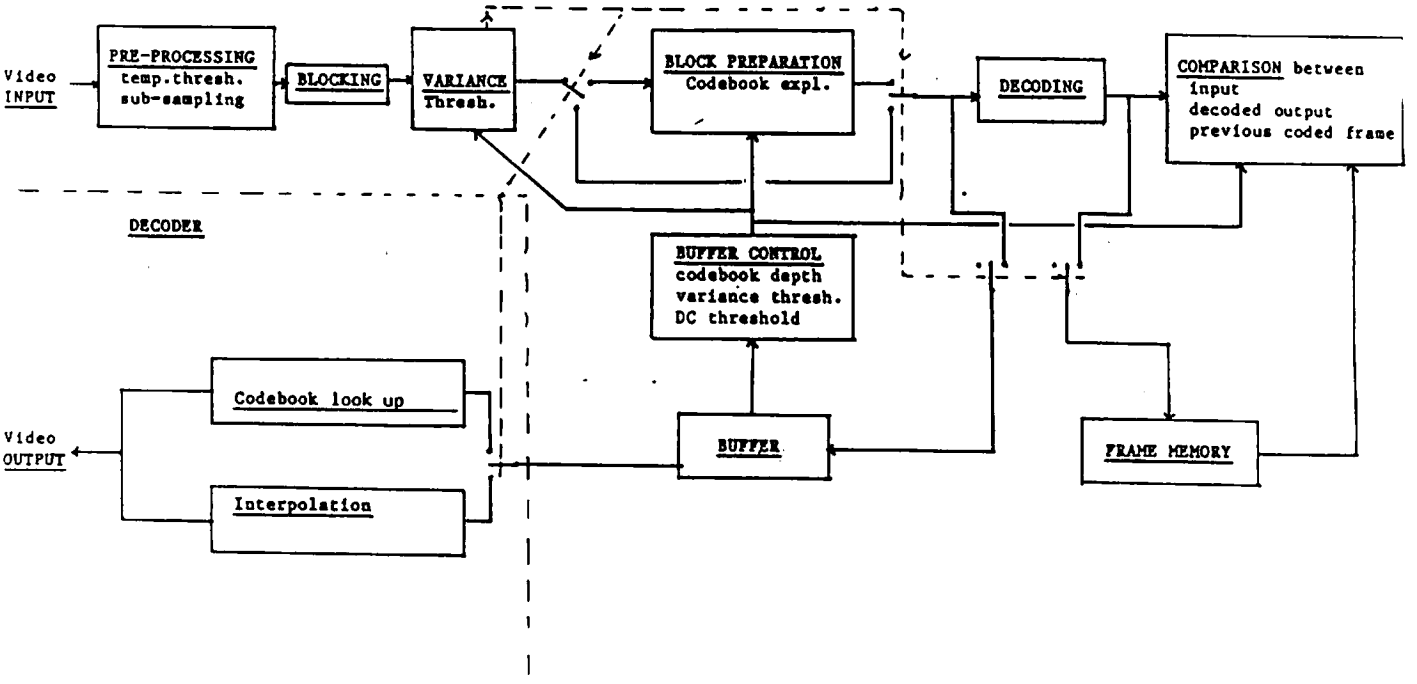


Fig.2 Codec scheme



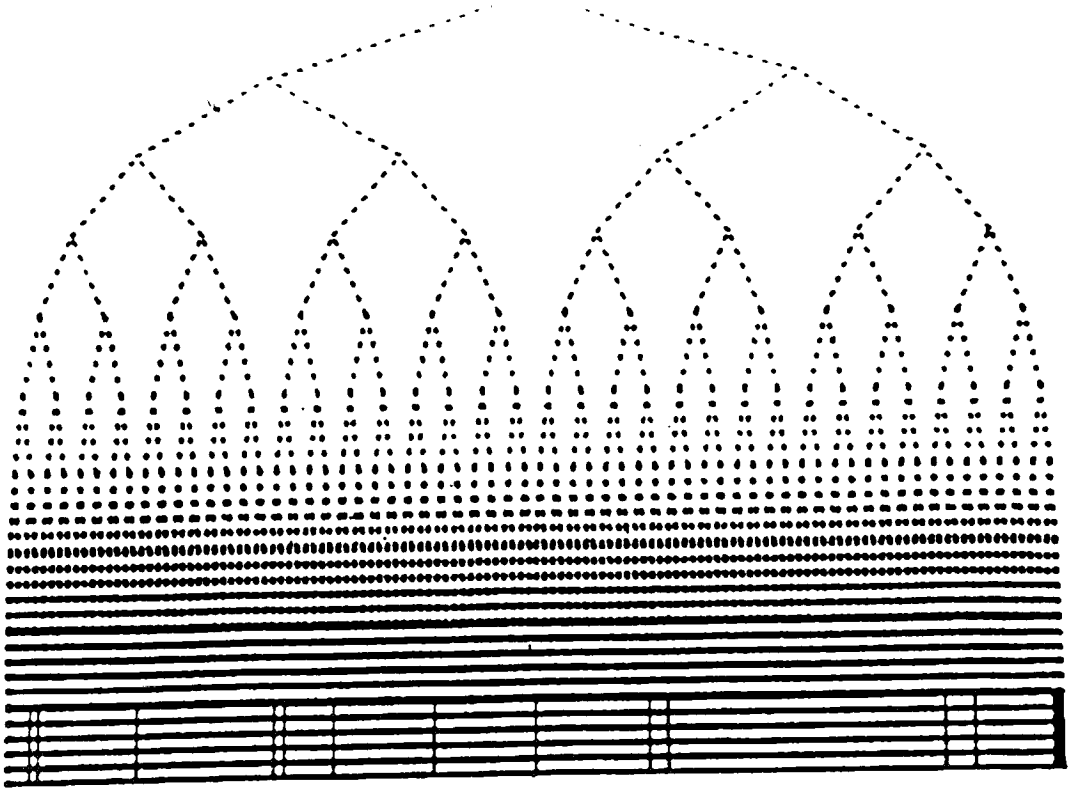


Fig. 3: Codebook scanning example

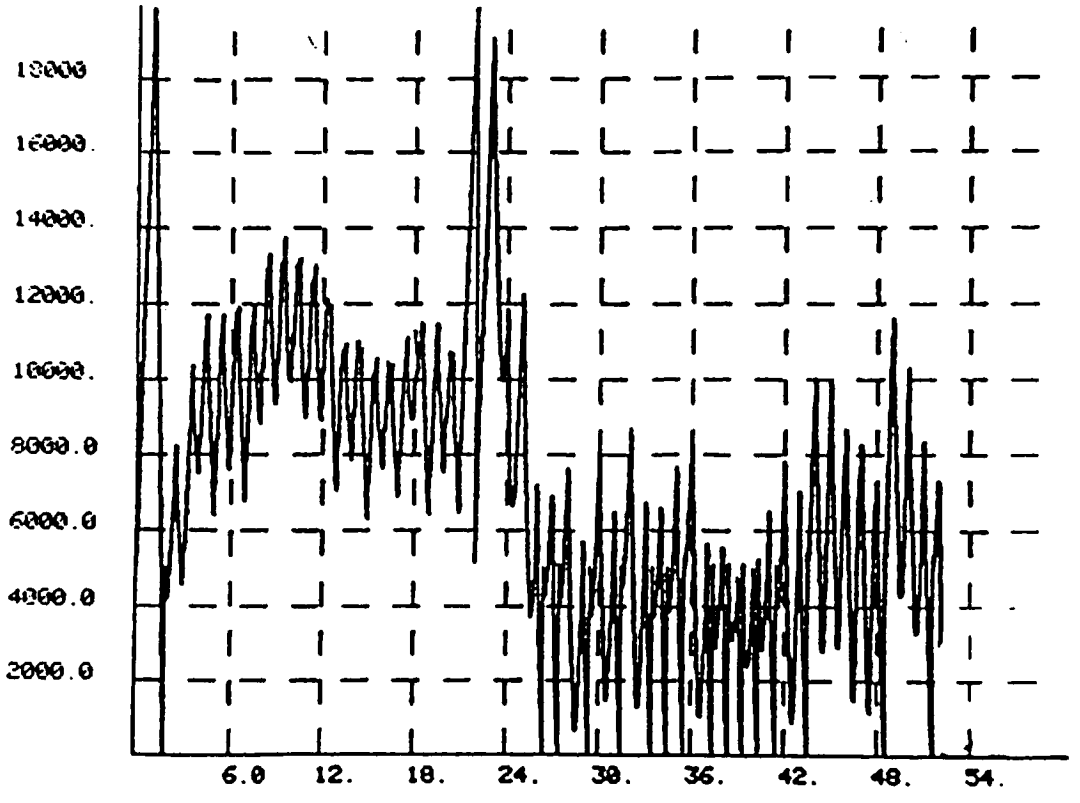


Fig. 4: Buffer filness