CCITT Study Group XV

Working Party on Visual Telephony

Specialists Group on Coding for Visual Telephony

Document no. # 39

English only

Date : September 1985

Source: The Netherlands, British Telecom

Anthropomorfic Transform Codec

OUTLINE OF THIS PAPER

In the introduction a brief description of the coder is given. The basic block structure is shown and a description of the functional aspects of the blocks is given. Further details are available for future participants, interested in collaborative research for optimization of parts of the coder.
References, mainly to the Anthropomorfic Transform are given.

INTRODUCTION

This paper is a first specification of a coder in the range of 64 kbit/s - 2 Mbit/s capable of transmitting and receiving component signals.
The coder is based on the Anthropomorfic Transform and a motion-speed-based Replenishment.

BASIC STRUCTURE OF THE ANTHROPOMORFIC CODER

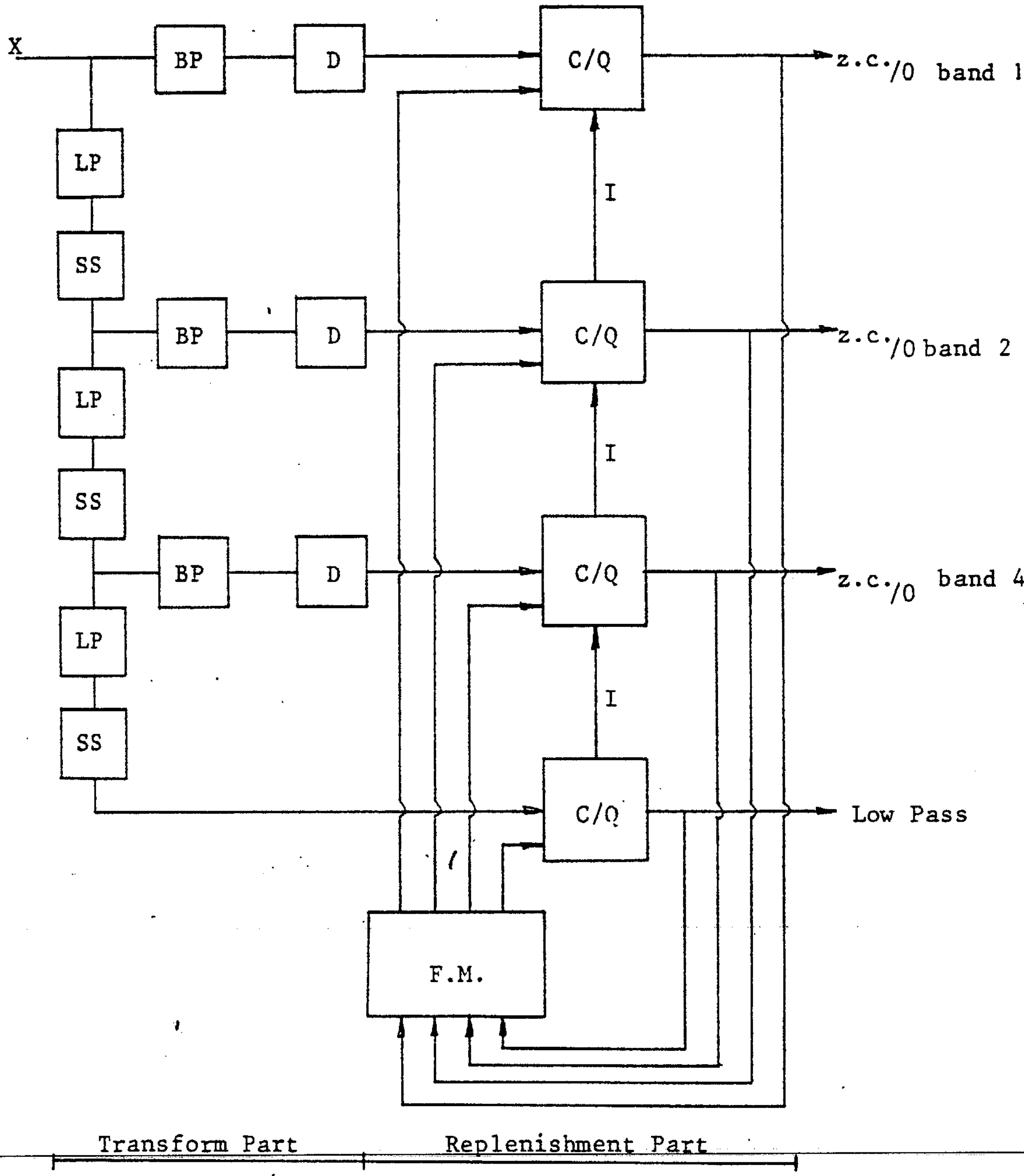

Figure 1. Blockscheme coder

BP : Bandpass filter

LP : Lowpass filter

D : Zerocrossing Detector

C/Q: Comparator/Quantizer

SS : Subsampling, factor 4:1

FM : Frame Memory: contains updated zero-crossing information

I : Inhibition

z.c./0: zerocrossing or no zerocrossing

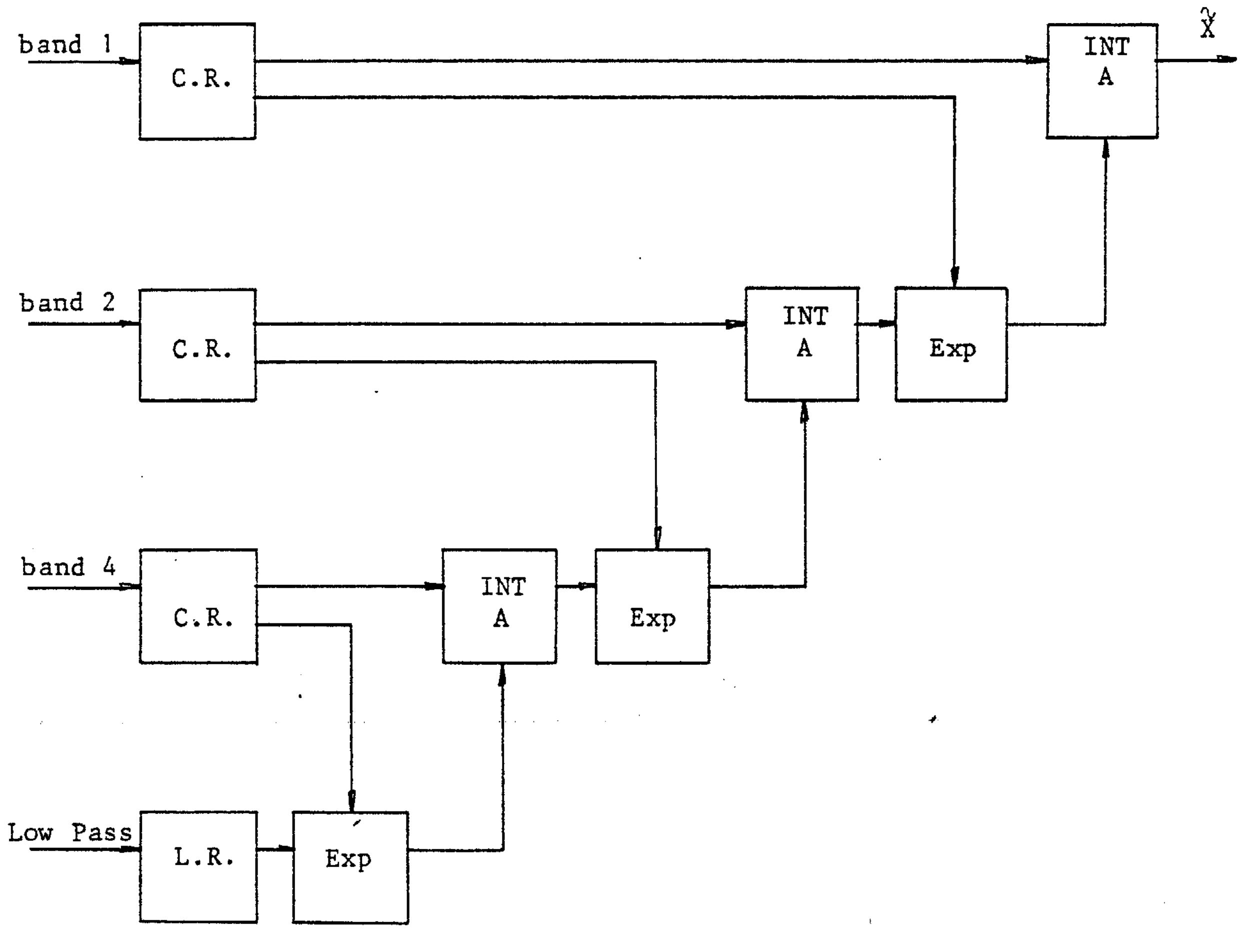

Figure 2. Blockscheme decoder: Reconstruction Part

CR : Contour Reconstruction

Exp: Expansion, 1:4

A : Adder

Int: Interpolation

THE TRANSFORM PART

This part of the coder is a close approximation to a specific, complete, non-orthogonal, self similar Hierarchical Transform [1] with a true inverse $T^{-1}=T$. This approximation is called the Anthropomorfic Transform [2] and is based on a model of the early pathway of the human visual system. The Anthropomorfic Transform uses only additions, no multiplications. A hardware design has already been made.

The output of the bandpass systems is the input of a zerocrossing detector, that establishes the positions where the value of the bandpass signal crosses the zero-level. These points are not isolated points but usually connected in zero crossing-contours. The output of the detector gives the position, local gradient and local orientation of a zerocrossing contour. These data can be looked at as Transform coefficients.

THE CONDITIONAL REPLENISHMENT PART

In this part of the coder the zerocrossings of the actual frame are compared with the updated zerocrossings of the past. [These are the same zerocrossings that the decoder has used to construct the previous frame]. Such a comparison gives an output: changed or not-changed. As an effect of the self similarity and the hierarchy in the system the changed or not-changed information in one level can be used in a next higher level to build a motion speed detector, simply by an inhibition rule. As a consequence the sharpness of an edge is reduced more whenever it moves faster. This part also uses additions only. A hardware design has not been made yet.

QUANTIZATION

The quantization rules are based on the ability of the human visual system to discriminate between edges with different steepness. Investigations have shown that the quantization of the gradient must be different for each band. Using 2 à 3 bit for band 1 up to 4 à 5 bit for band 4 no visual degradation can be seen.

ADRESSING THE ZEROCROSSING POSITION

The zerocrossing density is a typical parameter for the information content of the image. In video conferencing scenes the pictures are usually not fine-detailed and only partly moving, so the average zerocrossing density is very low. Speaking in Transform coding language: the number of coefficients that are definitly zero is relatively high.

In order to reach a low bitrate, an efficient adressing scheme for the sparse positions of the zerocrossings is needed. Because of the connectivity of zerocrossings a tree-like adressing scheme is proposed.
Experiments with such schemes show that the total amount of bits needed for the adressing will not exceed the amount of bits for the gradient and the orientation.

TEMPORAL FILTER AND SUBSAMPLING

384 kbit/s coders use a temporal filter and a subsampling (4:1). In this coder such a filter and subsampling procedure is an intrinsic part of the hierarchical coding scheme. Other temporal filters do not affect the coding schemes. It just takes the reduced picture format as the input.

TRANSLATABILITY

This coding scheme can be used without specific rearrangements for a large variety of bitrates.

IMPROVEMENTS

At the moment specific parts of the coder are still under investigation. Some parts are very robust to the human observer. Other parts don't use exhaustively the interdependence of the data. For example in the described coder no use is made of the dependences between the zerocrossing outputs of different bands. It is feasible that a prediction of a band-output can be made based on the lower band output. Transmitting only the difference between a bandoutput and its prediction gives a lower bitrate without effecting the quality. Such schemes are under investigation.

THE RECONSTRUCTION PART

This part consists of local operations, based only on additions.

REFERENCES

[1] Heideman G.H.L.M., Tattje H.E.P., Linden E.A.R. van der, Rijks D.: Self Similar Hierarchical Transforms; a bridge between Block-Transform coding and coding with a model of the human visual system.
2nd. International Technical Symposium on Optical and Electro-Optical Applied Science and Engineering.
Image Processing Symposium, France, Cannes 2-6 december 1985.

[2] Heideman G.H.L.M., Wanschers L.H.: Applications and performance of a human observer oriented Transform for image coding with low bitrates.
International Zurich Seminar on Digital Communications. Zurich, Switserland, March 1984.