DOC. # 83

C.C.I.T.T.
STUDY GROUP XV WP XV/1
Specialist Group on Picture Coding
Tokyo 25-28 March 1986

English Version

Source: AT&T Bell Laboratories

Title: Comparison of Discrete Cosine Transform with Other Transforms
Some Computer Simulation Results

This contribution consists of videotape demonstrations of full motion television image quality obtained after digital coding and compression down to the range of 300-400 kilobits/second, followed by decompression and conversion back to analog NTSC color TV. These results were obtained by computer simulation using the "Miss America" test image provided by the USA.

The first step for the simulation was to convert the test image into the "Common Intermediate Format (CIF)" agreed upon at the Torino, Italy meeting of the CCITT Specialists Group on Coding for Video Telephony. This format is basically, 288 line, 360 pel, 30 fields/sec noninterlaced. The CIF color difference signals for motion video are not specified. Thus, we made a reasonable choice of 4:1 horizontal and vertical subsampling of both the B-Y and R-Y color signals.

Next we reduced the field rate to 15 Hz incorporating some simple temporal low pass filtering, and each field was divided into 8x8 pel blocks. Alternate fields were then processed using block quantization of frame differences between the alternate fields. In future implementations, we may use motion compensated frame differences obtained by minimizing the average magnitude of the motion compensated differential signal[1].

The intermediate fields not processed above were then coded using conditional field subsampling[2]. With this technique a prediction of the intermediate field is made based on the two adjacent fields. If for a given block of pels, the prediction error is larger than some predetermined threshold, then the block is transmitted. Otherwise, it is not. In the future we may use motion compensated interpolation[3] for these intermediate fields.

Two methods of block quantization were implemented for comparison purposes, the first being the well known Discrete Cosine Transform (DCT). The DCT is one of a family of orthogonal transforms which has received much study over the years for coding single images, and some study for coding moving images. It is reasonably close in theoretical and practical performance to the optimum Karhunen-Loeve Transform for a variety of imagery under the mean square error (MSE) distortion criterion. However, it is widely known that the MSE criterion can be a misleading indicator of actual perceived picture quality, especially if there are large and sudden signal transitions due to sharp edges in the scene. The problem which arises is that even though most of the picture is portrayed well, a few small areas may be represented poorly. These tend to attract the viewer's attention and contribute to a lower overall quality rating than might otherwise result.

Study of the DCT and other block quantization methods for coding motion compensated frame differential signals has only begun. Theoretical results are largely nonexistant at this time, due mainly to the highly nonstationary statistics of the signal and to the relative nonutility of the MSE as a quality measure. Experimental efforts are proceeding mostly on a cut-and-try basis.

Since it is not at all clear that the DCT is anywhere near optimal for coding motion compensated frame differential signals, and also since orthogonal transform implementations generally tend to be more complicated than picture element (pel) domain approaches, we have begun studying some alternatives to orthogonal transforms for block quantization of pictorial data.

One such technique was used to produce the second sequence on the videotape. It codes the pels of the block to be sent, using linear combinations of other pels as it goes along. Typically for most blocks, not all pels need to be sent, and this is where most of the data compression occurs. The transmitted information can be quantized adaptively and values can be run-length-coded just as with the DCT. The main difference lies in the relative implementation simplicity compared with the DCT.

One hope of this study is to demonstrate that simple methods for block quantization might perform about as well as more complicated orthogonal transforms. Of course, final evaluation must await hardware simulation experiments.

[1] "Predictive Video Encoding Using Measured Subject Velocity", B.G. Haskell and J.O. Limb, U.S. Patent 3,632,865, Jan. 4, 1972

[2] "Conditional Vertical Subsampling in a Video Redundancy Reduction System", R.F.W. Pease, U.S. Patent 3,736,373, May 29, 1973

[3] "Video Signal Interpolation Using Motion Estimation", A.N. Netravali and J.D. Robbins, U.S. Patent 4,383,272, May 10, 1983