CCITT SG XV
Working Party XV/1
Specialist Group on Coding for Visual Telephony

Document # 281
Jan. 15, 1988

Source: Compression Labs, Inc.
Title: Inverse Transform Compatibility

---

1. INTRODUCTION

The method of implementation of the inverse discrete cosine transform (IDCT) must be specified to ensure compatibility between codecs. Document number 255 demonstrated that the distortion of the received image grows as a function of time when a mismatch exists between the IDCTs of the encoder and the decoder. Errors that produce insignificant distortion at the encoder will accumulate at the decoder. It is proposed to define a maximum amount of mismatch that is allowable, as well as a minimum amount of refresh required to limit the error accumulation.

2. PROPOSAL

2.1 DISCUSSION

During the Red Bank meeting a small group meeting was held to discuss the transform specification. The results of this meeting were summarized in TD# 7, which included a list of four possible approaches to specifying the IDCT. These four approaches are shown below, along with the advantages and disadvantages of each approach:

1. Specify the direct matrix multiplication approach with a defined amount of precision.

   ADVANTAGES - Allows for a small number of bits for implementation. Similar to the flexible hardware approach.

   DISADVANTAGES - Not algorithm independent. Direct matrix multiplication is not an efficient algorithm.

2. Allow some amount of mismatch. Define a minimum refresh rate and a maximum amount of mismatch.

   ADVANTAGES - Algorithm independent. Allows for many different methods of implementation.

   DISADVANTAGES - Refresh required to limit accumulated error. Can not use parity to detect error in memory. Some loss in coding efficiency due to forced refresh.

3. Specify a fast algorithm.

67

ADVANTAGES - Efficient implementation.

DISADVANTAGES - Not algorithm independent.

4. Choose a chip design as the standard.

ADVANTAGES - Efficient implementation. IC design already available.

DISADVANTAGES - Not algorithm independent. Possible licensing problems.

The proposal presented in this document combines approaches 1 and 2. Approach 1 is used as a baseline for approach 2. This allows manufacturers to use various methods for implementing the IDCT while still maintaining compatibility.

2.2 PROPOSED SPECIFICATION

The proposed specification consists of two major portions. One portion specifies the transform, and the other specifies the refresh requirement.

2.2.1 TRANSFORM SPECIFICATION

The transform is specified in terms of a baseline transform and the degree that the actual transform may deviate from the baseline.

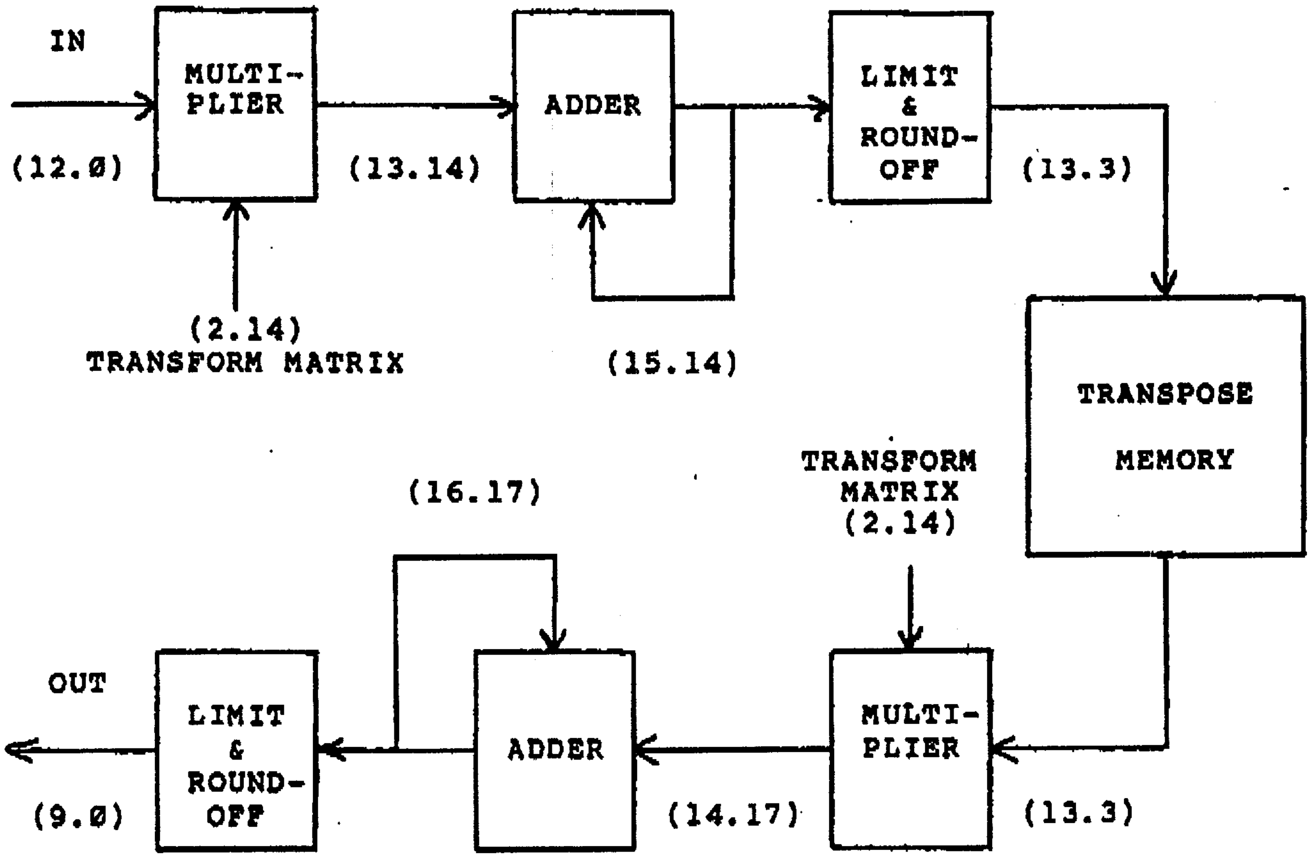
2.2.1.1 BASELINE TRANSFORM

The baseline implementation is similar to that defined for the flexible hardware. The baseline transform consists of a direct matrix multiplication as shown in Figure 1. The coefficients have a precision of 16 bits and are given in Table 1. The transform matrix has been defined to have perfect even and odd symmetry so that a baseline transform implementation may take advantage of this feature. The coefficients are scaled by a factor of 2.828 times those given in the flexible hardware specification to provide for slightly greater precision. The number of bits used at each point is shown in parentheses.

The results produced by the multipliers are retained at full precision, which is 27 bits for the vertical transform and 31 bits for the horizontal transform. The accumulators can handle the maximum possible range which requires 29 and 33 bits for vertical and horizontal transforms, respectively. After the vertical transform, the values are limited to 16 bits by rounding and clipping to the most positive and most negative 16-bit values. After the horizontal transform the values are limited to 9 bits by dividing by 8, rounding, and clipping to a 9-bit range.

Figure 1. Block diagram of baseline IDCT

VERTICAL INVERSE TRANSFORM



HORIZONTAL INVERSE TRANSFORM

Table 1. Inverse transform coefficients (scaled by 16384)

16384	16384	16384	16384	16384	16384	16384	16384
22725	19266	12873	4520	-4520	-12873	-19266	-22725
21407	8867	-8867	-21407	-21407	-8867	8867	21407
19266	-4520	-22725	-12873	12873	22725	4520	-19266
16384	-16384	-16384	16384	16384	-16384	-16384	16384
12873	-22725	4520	19266	-19266	-4520	22725	-12873
8867	-21407	21407	-8867	-8867	21407	-21407	8867
4520	-12873	19266	-22725	22725	-19266	12873	-4520

2.2.1.2 TRANSFORM MISMATCH

An implementation of the IDCT may vary from the baseline by a defined amount. This amount of variation is called the mismatch error (MME). The MME is measured by computing the MSE between the output of the baseline IDCT against the actual IDCT. For the purpose of measuring the MME it is assumed that the input to a perfect forward DCT consist of uniformly distributed, uncorrelated, 9 bit data. This data is transformed and rounded to 12 bits before passing through the IDCT. It is proposed that the MME should be no greater than 0.02.

2.2.2 REFRESH SPECIFICATION

The refresh specification is as follows - After a block has been interframe coded more than 30 times, it must be intraframe coded within the next 30 coded frame times.

This specification ensures that a block will not be interframe coded more than 60 times before being intraframe coded. Also a block need not be refreshed until after it has been interframe coded more than 30 times since the last intraframe coding. Blocks that are continually in the fixed mode never have to be refreshed, although in practice some minimum amount of refresh should be used. This specification allows one to use a simple cyclic refresh technique, or to use a somewhat more complicated technique that refreshes only the blocks that require it.

3. EXPERIMENTAL VERIFICATION

A study was performed to verify that the proposed specification produces reasonable results. When a mismatch exists the error at the decoder will grow uncontrollably if there is no refresh or leak factor. The error can be limited by using some amount of refresh, at the expense of additional bits required to code the refreshed blocks. Figure 2 shows four curves produced by coding the "Salesman" sequence at 768 Kbps with RM4. The coded frame rate is 15 frames per second.

The top curve gives the SNR when there is no mismatch between the encoder and the decoder, and no refresh is used. The bottom curve gives the results when the encoder uses the baseline IDCT and the IDCT in the decoder differs by rounding to 12 bits, instead of 16, between the vertical and horizontal inverse transforms. The relative error between the encoder and the decoder grows to about 1.75 dB in 15 seconds.

The second curve from the top shows the performance when no mismatch exists, but a refresh period of 61 frames, or slightly more than 4 seconds, is used. This causes a loss of about 0.2 dB. The refresh was performed by forcing every 61st block into the intraframe mode for each frame.

70

Figure 3. Distortion versus elapsed time

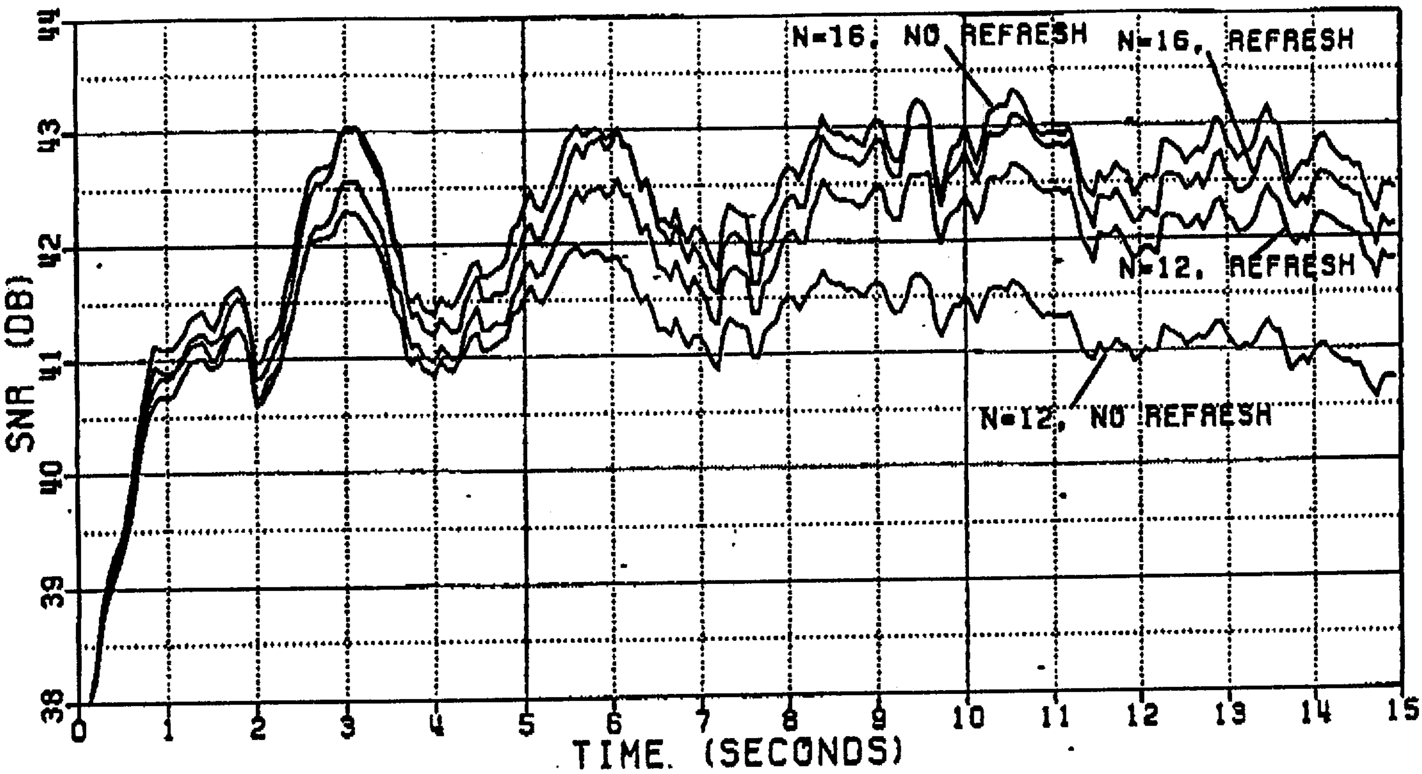
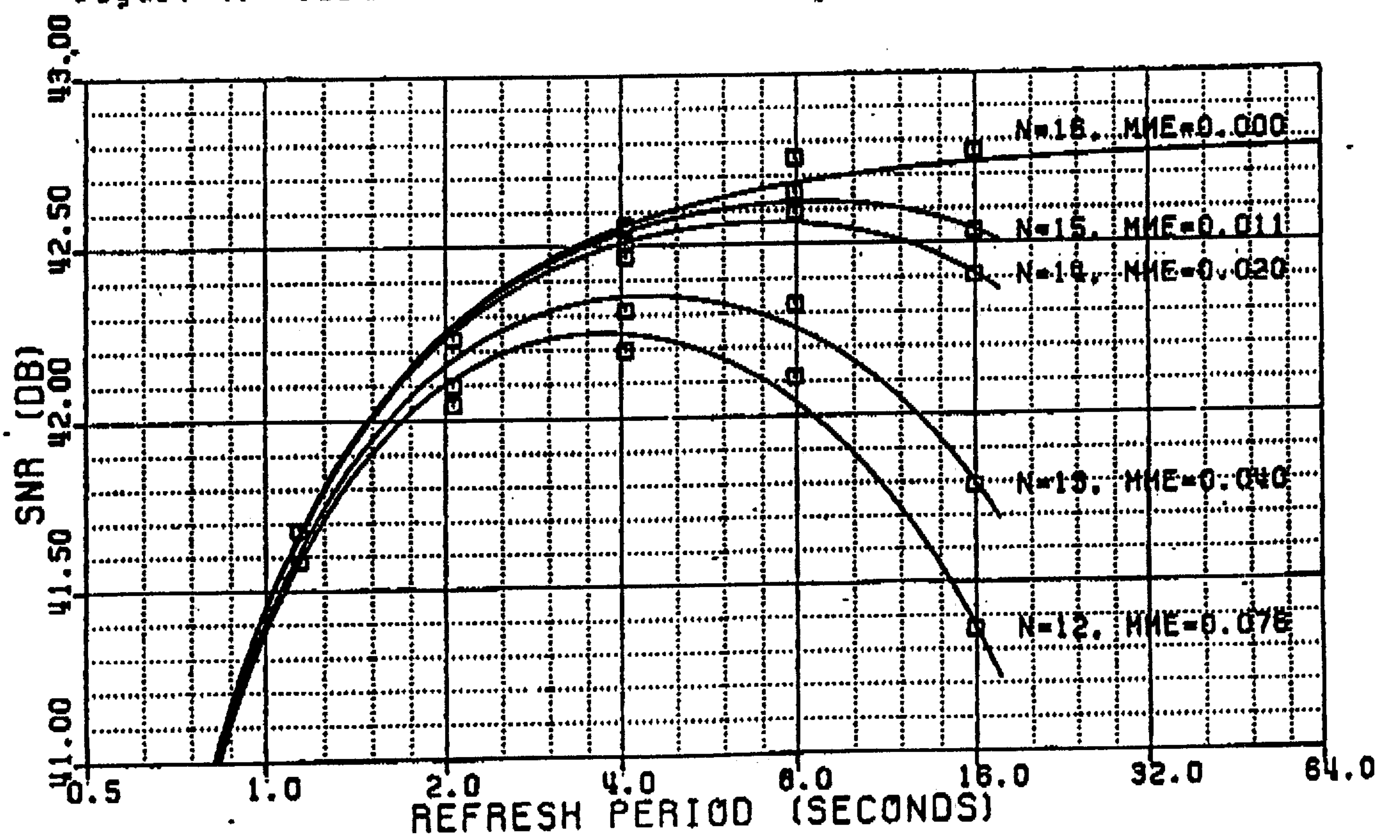


Figure 4. Distortion versus refresh period and IDCT accuracy



The third curve from the top shows the performance when the same mismatch exists as in the lowest curve, but the same refresh scheme is used as for the second curve. This produces an additional Ø.35 dB decrease in SNR, or a total of Ø.55 dB loss. However the error is now contained.

Figure 3 shows a family of curves giving the SNR as a function of refresh time and computational precision. The precision is controlled by the number of bits used between the vertical and horizontal transforms. This varies from 12 to 16 bits.

When a mismatch exists there is certain amount of refresh that gives the optimum performance. When the refresh rate is too low the mismatch error will dominate. When the refresh rate is too high the additional burden of intraframe coding will dominate. For an MME between Ø.Ø11 and Ø.Ø76, the optimal refresh period is in the range of 3 to 9 seconds. This implies a refresh period between 45 and 135 frames at 15 frames per second.

The effect of IDCT mismatch and refresh will differ from one sequence to another. The accumulation of error between the encoder and decoder is also a function of the percentage of interframe coded blocks, since intraframe and fixed blocks do not contribute to the accumulation. The use of the loop filter decreases the mismatch error for both the interframe and fixed modes.

The mismatch error will also depend on the number of nonzero coefficients. As the quantization factor decreases the number of nonzero coefficients will increase. It is expected that this will also tend to increase the mismatch error.

The "Salesman" sequence was chosen for this study because it contains a fairly highly detailed background and quite a bit of motion. These features should tend to produce a large accumulation of mismatch error and also be quite sensitive to the refresh period.

4. SUMMARY

A proposal for the IDCT has been given in this document that allows for some degree of freedom in the choice of the IDCT algorithm. This allows for various methods of implementations that can take advantage of presently available technologies as well as more advanced implementational techniques in the future.

It was shown that the proposed specification allows for this flexibility with only about Ø.3 dB loss in coding efficiency for the "Salesman" sequence at 768 Kbps. Since the baseline IDCT is defined with 16 bits of precision for the coefficients and multiplications, other IDCT algorithms should be implementable using 16 bit precision as well.