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

Document #366
Sep. 1988

Source : NTT, KDD, NEC and FUJITSU

Title : Precision for Inverse DCT Calculation

---

1. Introduction

An issue of IDCT mismatch error and its countermeasure have been discussed in Doc.#281, #284, #321. According to these documents, reduction of visible degradation is possible by using intra-frame mode, and evaluation of IDCT calculation error is required to reproduce picture with sufficient quality. This document describes estimation of IDCT calculation error and the countermeasure for the mismatch error problem. Moreover, procedures are exemplified for IDCT specification.

The following points are discussed. (1) Whether random data is appropriated or not for estimation of IDCT calculation error. (2) Two mismatch suppression methods, filter in the coding loop and bit length expansion for IDCT output. (3) Procedures for decision of IDCT specification.

2. Evaluation of IDCT calculation error

Two evaluation methods for IDCT calculation error are compared in this section. One uses a random data sequence, and the other uses an actual picture sequence. The random data is generated by "C rand()". The picture sequence is a prediction error sequence which is obtained from 'Claire(part)' with RM4.

Result is shown in Table.1, where evaluated values are represented by mean square error(M.S.E.). The evaluation was carried out at the point 'comparison-1' in Fig.1, where IDCT output data is not rounded off. Two groups of evaluation data, one for random data and the other for picture data almost coincide with each other. Therefore, The evaluation with random data may be considered appropriate.

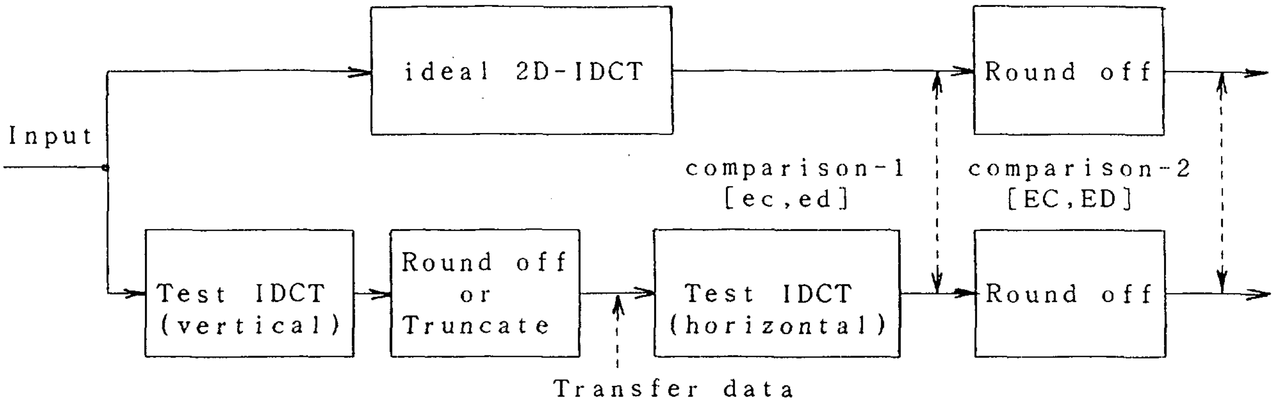


Fig.1 Evaluation of the IDCT Calculation Error

Transform Matrix Data Length (Below binary Point)	Transfer Data (Below binary Point)	Evaluation with Random Data (M.S.E.)	Evaluation with Picture Data (M.S.E.)
15bit	1bit(round off)	0.021	0.0170
15bit	2bit(round off)	0.005	0.0044
15bit	3bit(round off)	0.001	0.0011

Table.1 IDCT calculation Error

3. Suppression of mismatch error by filter in the coding loop

A filter in the coding loop can suppress accumulation of the mismatch error. If the filter in the coding loop is applied to every significant block at a constant rate, accumulation value will be saturated.

However, there can be some significant blocks to which filter is not applied. For example, if the filter is controlled by motion vectors, there may be no filtered blocks in still picture area even though noisy. Mismatch error is not suppressed in such a case. VTR demonstration of encoded color-bar in the annex to Doc.#321(in the Hague) should be this example. Generally, mismatch error accumulation value in back ground parts is greater than moving part with this reason.

Therefore, if the filter in the coding loop is used to suppress the mismatch error, the filter must be controlled taking this into consideration. It should be noted that excessive use of the filter reproduces blurred image.

It is suggested for safety that the mismatch error specification should be decided without the filter in the coding loop.

4. Bit length expansion for IDCT output

A relation between IDCT calculation error and the mismatch error is shown as follows.

$$\overline{(EC - ED)^2} = \sqrt{2/3}\,\sqrt{\overline{EC^2} + \overline{ED^2} - (1/6)L^2}\cdot L$$
$$= \sqrt{2/3}\,\sqrt{\overline{ec^2} + \overline{ed^2}}\cdot L \qquad (4.1)$$

This is a modified equation based on the equation in Doc.#321. (In Doc.#321, L=1.) In this equation, L represents round-off precision for IDCT output, EC and ec for the IDCT calculation error in coder, ED and ed for the error in decoder. EC and ED include the last round-off error, while ec ,ed do not[See Fig.1]. $\overline{EC^2} = \overline{ec^2} + (1/12)L^2$ and $\overline{ED^2} = \overline{ed^2} + (1/12)L^2$.

Although Eq.(4.1) is not adequate if $\overline{ec^2} > (1/12)L^2$ or $\overline{ed^2} > (1/12)L^2$, the following condition holds in most cases, $\overline{ec^2} < (1/12)L^2$ or $\overline{ed^2} < (1/12)L^2$. Eq.(4.1) is shown by a solid line in Fig.2. Measured values from computer simulation are also plotted in this figure. These show a good match between the Eq.(4.1) and the experimental results.

The Eq.(4.1) shows that mismatch error can be suppressed with a smaller value of L. L represents accuracy of IDCT output. If IDCT output data accuracy is expanded by n bits below binary point, mean square mismatch error is reduced to 2^{-n} times as small. Computer simulation was carried out to prove it. A simulation result shows that expansion by 1 bit and 2 bits makes mean square mismatch error reduced to 1/2 and 1/4, respectively, as shown in Fig.3. Intra-frame mode is not used in this simulation. In Fig.4, relation between the mismatch error and round-off width L is shown. This shows that mismatch error can be well suppressed when n is taken to be up to three. For longer n values the suppression effect seems to saturate.

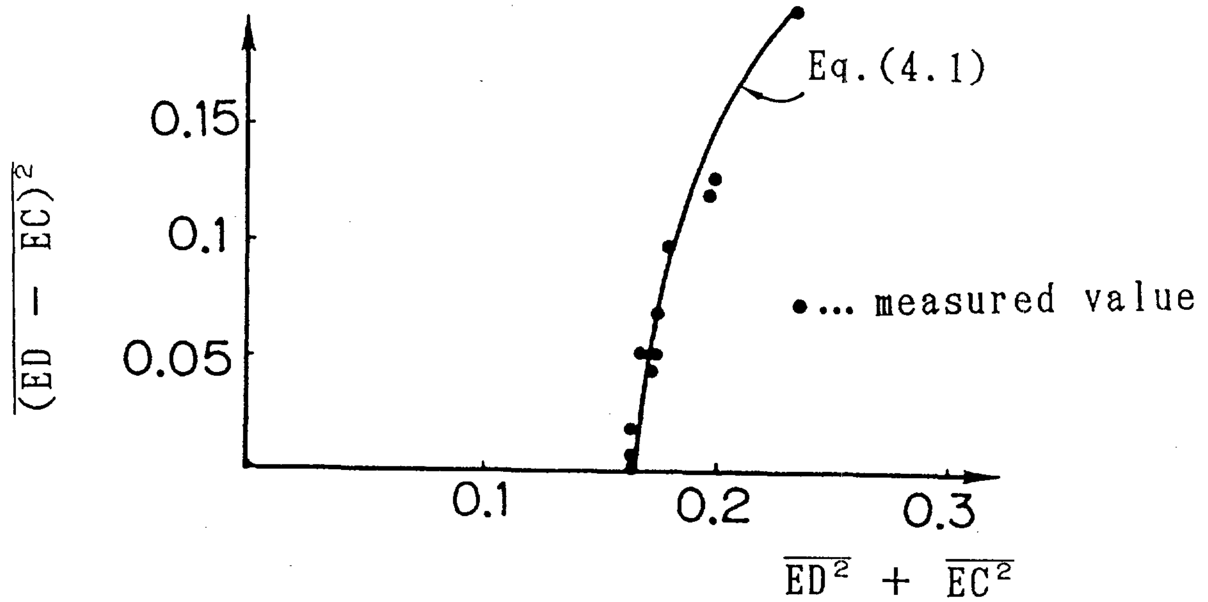


Fig.2 Relation between mean square mismatch error and IDCT calculation error

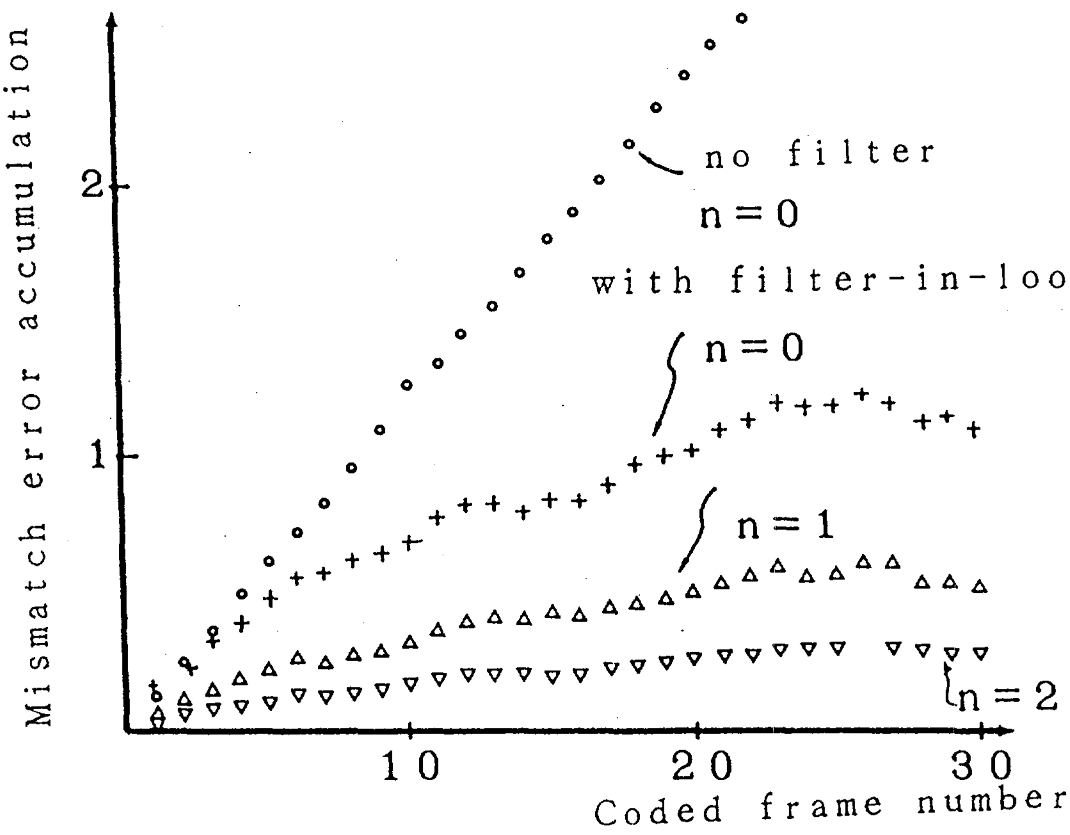
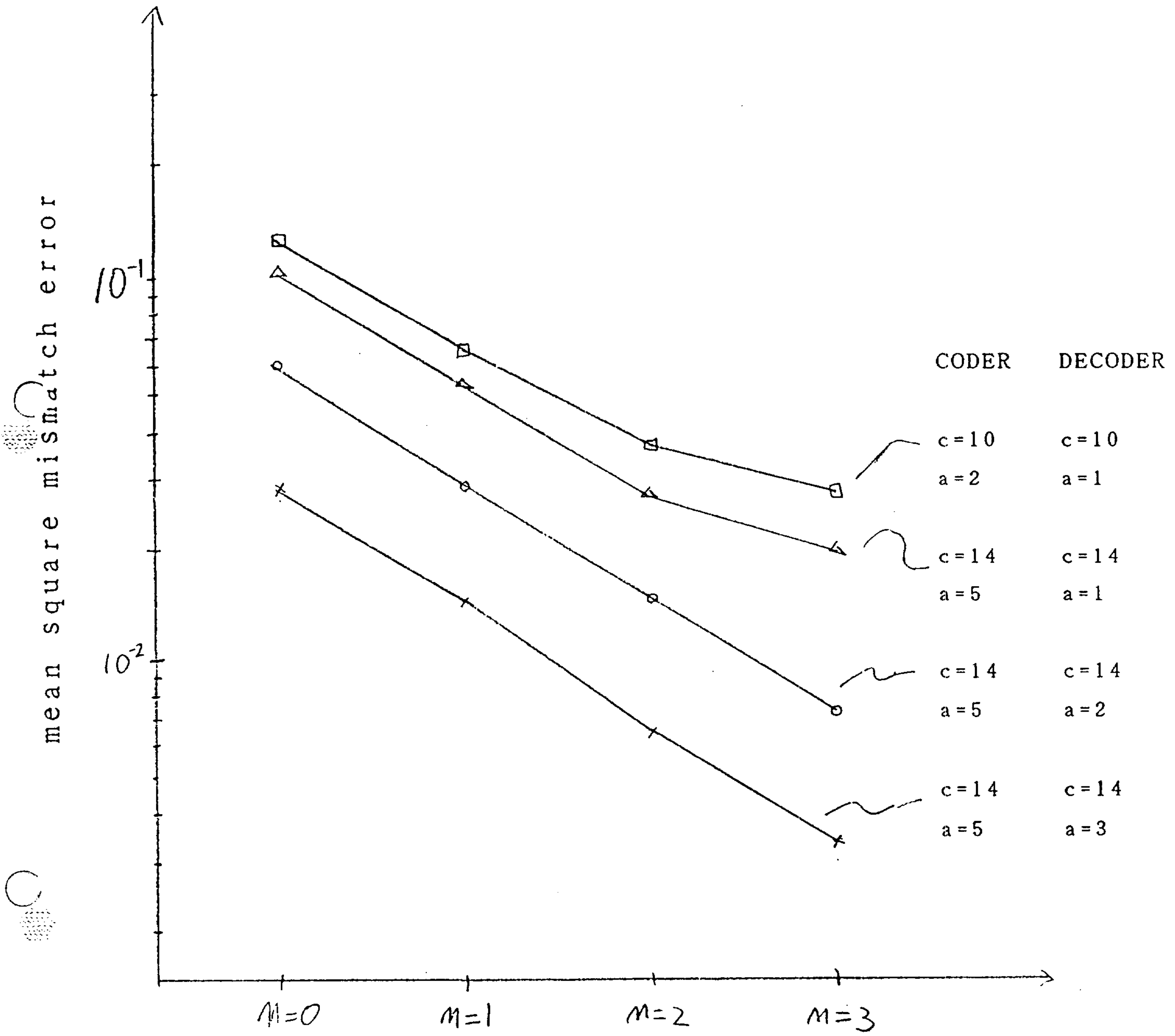


Fig.3 Simulation result(1)



IDCT output data precision
(n-bit accuracy bellow binary point)

a: transfer data length (below binary point)
c: transform matrix element length
(below binary point)

Fig.4 Simulation result(2)

This suppression technique is effective, though it requires expansion of data length in frame memories.

5. Procedure for IDCT specification

Procedures are exemplified to decide IDCT specification in this section. These procedures have to be carried out in as many organizations as possible, since these procedures require subjective estimation.

Degradation caused by mismatch error can be classified as follows.

a. Off-set error accumulation at a particular position in each block.
   For example, vertical stripes appear with truncation of transfer data.
b. Off-set error accumulation on whole blocks.
   In this case, picture becomes darker/brighter gradually.
c. Random error accumulation.
   Picture quality degradation is observed as granular noise.

These types of degradation can be suppressed by limiting the following items, respectively.

A. maximum mean error
B. over-all mean error
C. mean square error

Stimulus threshold for these values can be decided with the procedures mentioned in what follows using Flexible Hardware.

5-A. Specification for maximum mean error

I. Produce off-set mismatch error which is concentrated in a particular part of a block.
For example, if transfer data (between vertical IDCT and horizontal IDCT) is rounded off in coder and truncated in decoder, off-set error concentrate to left edge of each block.

II. Evaluate Maximum mean error of IDCT calculation in a coder and a decoder with random data. The difference of these two values is defined as 'MEmme'.

III. Find maximum frame number 'Nmme' when picture quality degradation becomes visible under the conditions:
No cyclic refresh,
No intra-frame mode,
No filter in the coding loop.

IV. Calculate the threshold 'THmme' by

$$THmme = (MEmme * Nmme) / RI$$

where RI is an allowable refresh interval.

V. Maximum mean error of IDCT should not exceed THmme/2.

5-B. Specification for over-all mean error

I. Produce off-set mismatch error which occurs equally in each block.
For example, add off-set value to DC coefficients in decoder. In this case, reproduced picture becomes darker /brighter gradually in decoder.

II. Evaluate over-all mean error of IDCT calculation in a coder and in a decoder with random data. The difference of these two values is defined as 'MEome'.

III. Find maximum frame number 'Nome' when picture quality degradation becomes visible under the following conditions:
No cyclic refresh,
No intra-frame mode,
No filter in the coding loop.

IV. Calculate the threshold 'THome' by

THome = (MEome * Nome) / RI

where RI is an allowable refresh interval.

V. Over-all mean error of IDCT should not exceed THome/2.

5-C. Specification for mean square error

I. Produce mismatch error which does not include off-set error.

II. Evaluate mean square errors of IDCT calculation in a coder and in a decoder with random data, called '$\overline{ec^2}$', '$\overline{ed^2}$' or '$\overline{EC^2}$', '$\overline{ED^2}$'.

III. Calculate the mean square mismatch error 'MEmse' with

$$\overline{MEmse^2} = \sqrt{2/3}\,\sqrt{\overline{EC^2} + \overline{ED^2} - (1/6)L^2} \cdot L$$
$$= \sqrt{2/3}\,\sqrt{\overline{ec^2} + \overline{ed^2}} \cdot L,$$

where L is round off width for IDCT output.

IV. Find maximum frame number 'Nmse' where picture quality degradation becomes visible under the following conditions:
No cyclic refresh,
No intra-frame mode,
No filter-in-loop.

V. To calculate the threshold 'THmse' with

$$\overline{THmse^2} = (MEmse * Nmse) / RI$$

where RI is an allowable refresh interval.

VI. Specification for mean square error of IDCT is giving by

$$\overline{ec^2}, \overline{ed^2} < (3/4)\ (THmse / L)^2$$

or

$$\overline{EC^2}, \overline{ED^2} < (3/4)\ (THmse / L)^2 + (1/12)L^2.$$

6. Conclusion

Random data sequence is appropriate to evaluate IDCT calculation error. Filter in the coding loop is effective to suppress mismatch error accumulation. However, it should not be taken into consideration on IDCT specification, since constant use of the filter brings intolerable blurring to reproduced picture. Expanding bit length for IDCT output is an effective method, though it requires a little bit greater hardware complexity. Finally, procedures are exemplified in deciding IDCT specifications.