

The invention relates to a method for the correction in accordance with the first part of claim 1.

An approximation of a matrix is described at pages 114 - 120 of the 1990 dissertation thesis at the TH Aachen "Adaptive Transformationskodierung von digitalisierten Bildsignalen" by Dipl. -Ing Wolfgang Mauersberger. Because of the real value matrix elements, an arithmetical processor using floating point arithmetic is proposed. The matrix elements $A(I, J)$ are approximated (page 115, second paragraph).

A method for the transformation with the aid of matrix multiplications is described in EP-A-0 077 089 in which a rounding, resulting from the limited number of places in the calculation, is compensated by a reversible unambiguous mapping of digit locations into other non-occupied digit locations. This method is only suitable however for such special transformations in which the coefficients can be formed by the building up of sums and differences of binary digits wherein two even or two odd results must always occur.

It is the object of the invention to provide a method for the compensation of rounding errors in matrix multiplications in general transformations.

In a transformation coding, analogue signals are digitalised and the digitalised signals transformed from a time domain into the frequency domain. For this, the digitalised signals are arranged in matrices. For a multiplication of matrices, matrix elements are multiplied together column-wise and row-wise and the products are summed. For the binary multiplication, the number of places of the product results from the sum of the number of places of the factors. After a plurality of successive multiplications have been carried out, a processing using all the places can no longer be realised in an arithmetical processor utilising only simple hardware. In accordance with the desired degree of accuracy, rounding errors in the multiplications can be tolerated if it can be retrospectively ensured that these rounding errors are compensated.

For the better understanding of the invention, an embodiment will be explained hereinafter with the help of the Figures. These show in:

Fig. 1 a communications system,

Fig. 2 a source encoder,

Fig. 3 a source decoder

Fig. 4 a screen,

Fig. 5 a matrix having transformed values,

Fig. 6 a matrix multiplication,

Fig. 7 a block circuit diagram for an inverse discrete cosine transformation,

Fig. 8 a correcting circuit following a discrete cosine transformation.

Fig. 1 shows a communications source 1, an ADU 2 (analogue to digital converter), a source encoder 3, a channel coder 4, a channel 5, a channel decoder 6, a source decoder 7, a DAU 8 (digital to analogue converter) and a communications sink 9. The channel 5 may, for example, stand for a HF transmission path or a digital video recorder. A digital video recorder may however also represent for example, the ADU 2, the source encoder 3, the channel coder 4, the channel 5, the channel decoder 6, the source decoder 7 and the DAU 8. The source is then formed by a HF transmission path, the channel 5 by a magnetic tape and a magnetic head of the video recorder, the sink being a television apparatus.

Fig. 2 shows a source encoder 3 having a discrete cosine transformer 10, a quantiser 11 and a control means 12. Discrete digital signals arranged in blocks are transformed from the time domain into the frequency domain in the discrete cosine transformation 10. This corresponds to a mathematical multiplication of matrices. The values transformed into the frequency domain are quantised in the quantiser 11. Quantisation, see also page 29 et seq in the thesis of Mauersberger. The control 12 effects the insertion of control and address codes. The output of the quantiser 11 is applied to the channel encoder 4.

Fig. 3 shows a source decoder having a decoder 13, including an inverse discrete cosine transformation 14 and a controller 15. The controller 15 sorts out the control and address codes from the incoming transmitted values and ensures that the decoder only passes on data words (as opposed to control words and addresses) to the inverse transformer 14. The inverse transformer 14 transforms the quantised frequency values transmitted over the channel 5 into the time domain.

Fig. 4 shows a television apparatus 16 having a screen 17. Lines 18 can be recognised on the screen 17. Three television lines 18 pass through each block 19. The blocks 19 encompass a defined number of picture elements and are arranged in a square. In this example, the blocks 19 encompass a block having 3×3 picture elements. In practice, a block of 8×8 picture elements is utilised. The grey or colour values of the picture elements in the blocks form the elements of the matrices of the display region.

Fig. 5 shows a 4×4 matrix having values, the matrix coefficients, transformed into the frequency domain. The matrix value 20, which is arranged in the first column and the first row i.e. the element $K(1, 1)$, is a measure of the dc coefficient, hereinafter referred to as the average value, of a block.

Fig. 6 shows a mathematical equation including matrices wherein the matrices 21 and 22 are multiplied together and produce the matrix 23.

Fig. 7 shows an inverse discrete cosine transformer 14, in which frequency values are transformed back into the time domain. The frequency domain values are buffered in the buffer store 24 and supplied to the arithmetical processor 25. The frequency domain values arranged block-wise in the form of a matrix are transformed back into the display region in the arithmetical processor 25. To this end, the arithmetical processor 25 accesses the PROM 26 and calls up the matrices stored therein. The formation of such matrices for a discrete transformation or a discrete cosine transformation (DCT) are described for example (at pages 20 - 25) of the thesis "Ein Beitrag zur Informationsreduktion bei Fernsehbildsignalen mit Transformationscodierung und adaptiver Quantisierung" by Robert Sell, 1979 at the Gesamthochschule, Wuppertal. The method of operation of arithmetical processors is explained in the TTL Cookbook from Texas Instruments, e.g. at pages 226 - 228. The buffer store 27 receives the grey or colour values transformed into the display region and passes them on to a monitor or a television apparatus.

Fig. 8 shows the discrete cosine transformer 10, in which the display region values are transformed into the frequency domain. Display region values are buffered in the buffer store 28 and supplied from there to the arithmetical processor 29. The arithmetical processor 29 accesses the further PROM store 30 and is supplied with matrix elements by the PROM 30. The

arithmetical processor 29 multiplies the matrices together and thus produces frequency domain values, hereinafter called coefficients, from the display region values. The block-wise arranged coefficients are passed on via a summing circuit 32 to the buffer store 33. For a transformation, e.g. DCT, the following calculation is carried out:

$$(K) = (DCT) * (B) * (DCT)^{-1}$$

For a reverse transformation or inverse transformation one calculates as follows:

$$(B) = (DCT)^{-1} * (K) * (DCT)$$

where

- (B) = picture element matrix
- (K) = coefficients matrix
- (DCT) = transformation matrix
- (DCT)⁻¹ = inverse transformation matrix

In order to simplify the arithmetical processor 29 as much as possible, calculations are made there with integral numbers. Microprocessors or arithmetical processors generally operate with $2 \exp n$ place digits, as a rule with eight or sixteen bits. Display region values and frequency domain values are preferably tailored from real numerical values into integer numbers which are natural numbers having positive and negative prefixes and the numerical value zero i.e. they are multiplied by an arbitrary but defined factor in order to exclude the decimal points.

Because with the four matrix multiplications the place digits of the products increase i.e. after a multiplication, one or more extra bits are needed each time by the arithmetical processor, and because the arithmetical processor is required to function with the smallest possible word-length e.g. of 8 or 16 bits, there is a rounding after some or all of the multiplications, or the number of places is deliberately limited. In general, due to leaving out the last place, there is a rounding down because this can be effected without additional expenditure. However, the amplitude of the picture elements, especially that of the average value (element K(1, 1) in the coefficients matrix) becomes smaller and smaller in the course of the calculation in comparison with a calculation without rounding.

The correcting circuit 31, which to a large extent corrects the average value, prevents a too small an amplitude. It has transpired in practice that it is sufficient to add a correction

value only to the average value of the coefficients matrix. The correction value MK is calculated from the following numbers:

- the element $K(1, 1)$ of the coefficients matrix itself,
- the number N of the bits which are available in the arithmetical processor for the maximum word length,
- the number T of the bits which is stored in the PROM 30 as the maximum word length for the transformation matrix elements (DCT),
- the number D of the bits which results from the row (x) and column length (y) of the picture element matrix (x and y).

The picture element matrix may be one dimensional ($x = 1$ or $y = 1$), it may be a 2 dimensional square ($x = y, \geq 2$) or it may be a variable 2 dimensional one ($x \neq y, x \geq 2, y \geq 2$).

With

$$D = \text{int} (\log_2 (2 * (\max (x, y)) - 1))$$

the correction value becomes

$$MK = K(1, 1) + K(1, 1) / 2 \exp (N - T - D + 1))$$

for a one dimensional and a square two dimensional matrix, and becomes

$$MK = K(1, 1) + K(1, 1) / 2 \exp (N - T - D))$$

for the two dimensional matrix having rows and columns of different lengths.

The element $K(1, 1)$ of a frequency domain matrix stands for the dc coefficient and thus represents a measure for the brightness of a block. The correction value MK for this dc coefficient is calculated in the correcting circuit 31 and the dc coefficient and the correction value are added in the summing circuit 32 and the sum of the two is supplied to the buffer store 33. A delay member 34 delays the average value and all of the remaining coefficients by the time t_0 which is necessary in order to calculate the correction value MK for the average value. The correcting circuit is constructed from arithmetical processors which are described in the TTL-Cookbook in accordance with the equations for MK. To raise to the power of, means to shift a value in a register. D and the factor $(2 \exp (N - T - D + 1))$ or the factor $(2 \exp (N - T - D))$ are not calculated in the correcting circuit 31 but rather are specified as defined magnitudes by the matrices being used.

Claims

1. Method for the correction of discrete values formed blockwise by one or more matrix multiplications (Fig. 6) in which the maximum number of places of the formed values normally resulting from these multiplications is limited to a predetermined lower number of places (N) by the circuits (25, 29) used for the matrix multiplication, wherein at least one of the matrices involved contains predetermined coefficients (DCT, 26, 30), characterised in that, a correcting value (MK) is formed from an element (20) formed by the matrix multiplication which represents a dc coefficient of the appertaining elements of the matrix and is added to this element, wherein in addition, for the calculation of this correcting value (MK), the number of places (N) predetermined by the circuits (25, 29), the maximum number of places (T) of the predetermined coefficients and a value (D) resulting from the maximum length of the columns or rows of the input signal matrix is used.

2. Method in accordance with claim 1, characterised in that, the value (D) resulting from the maximum length of the columns or rows of the input signal matrix is calculated in the following manner:

$$D = \text{int} (\log_2 [2 * \max(x, y) - 1]),$$

wherein x is the length of the rows and y the length of the columns.

3. Method in accordance with claim 1 or 2, characterised in that, for a one dimensional matrix or for a quadratic input signal matrix, the correcting value (MK) is calculated in the following manner:

$$MK = K(1,1) / 2^{(N-T-D+1)},$$

wherein K(1,1) is the dc coefficient and N the number of places predetermined by the circuits (25, 29) and T the maximum number of places of the predetermined coefficients.



4. Method in accordance with claim 1 or 2, characterised in that, for a two dimensional non-quadratic input signal matrix, the correcting value (MK) is calculated in the following manner:

$$MK = K(1,1) / 2^{(N-T-D)},$$

wherein $K(1,1)$ is the dc coefficient and N the number of places predetermined by the circuits (25, 29) and T the maximum number of places of the predetermined coefficients.

5. Method in accordance with one or more of claims 2 to 4, characterised in that, the values for D and/or $2^{(N-T-D+1)}$ and/or $2^{(N-T-D)}$ are stored or hard wired-in.

6. Method in accordance with one or more of claims 1 to 5, characterised in that, the input signal matrix contains grey and/or colour values of a video signal.

7. Circuit for a method in accordance with one or more of claims 1 to 6, provided with an arithmetical processor (29) for the matrix multiplications, with read-store means (30) which contain the predetermined coefficients connected to the arithmetical processor (29), with a correcting circuit (31) arranged after the arithmetical processor (29) which each time calculates a correcting value (MK) in accordance with claim 3 or 4 and adds to the corresponding dc coefficient in a summing circuit (32).

