POSSIBILITIES OF USING SEVERAL DISCRETE TRANSFORMS FOR THE VIDEO COMPRESSION

Ivo Chromý
Institute of Radio Electronics
Faculty of Electrical Engineering and Computer Science
Brno University of Technology
Purkyňova 118, 612 00 Brno,
Czech Republic
E-mail: chromy@urel.fee.vutbr.cz

Abstract

This paper treats of compression of digital video signals and the main aim is to probe the possibility of using several discrete transforms besides well-known discrete cosine transform (DCT). Mathematical descriptions of these transforms are presented in the article. DCT and other orthogonal transforms are described here. Matrix notation of computation of two-dimensional transform is also presented. The comparison of the results obtained by application of the transforms on pictures is introduced. A Normalised Root Mean Square Error is defined for mathematical comparison of results. The aim of the work is to probe applicability of the transforms introduced above for video compression in good quality and the best compression ratio.

Keywords

discrete transform, compression ratio, Normalised Root Mean Square Error, visual quality, ...

1. Introduction

The transmission of video signals in digital form is not possible without using compression techniques because the resultant bit flow of non-compressed signal would be too high.

They are used many methods of the compression but the most used ways are transformation methods using the discrete cosine transform (DCT). The DCT and some other transforms will be described in the article.

Pictures are two-dimensional objects therefore the use of the two-dimensional transforms is necessary for their compression. The two-dimensional cases of the transform will be presented below.

2. Mathematical description of the transforms

This chapter is devoted to the description of the discrete cosine transform, two types of Walsh-Hadamard transforms and Haar transform.

Discrete cosine transform (DCT) of a data sequence X(m), $m = 0, 1, ..., N-1$ is defined as [1]

$$\mathbf{L_x}(n) = \mathbf{H_{DCT}}(n)\mathbf{X}(n)^T , \tag{2.1}$$

where $L_x(k)$ is the k-th DCT coefficient, $\mathbf{H_{DCT}}$ is the kernel of size NxN of DCT, $\mathbf{X}(n)$ is an N-vector of data X(m) and $n = \log_2 N$.

Walsh-Hadamard transform (WHT) [1] is transform which uses the set of Walsh functions as the set of basis functions. Two types of WHT are described below. The first is Hadamard-ordered Walsh-Hadamard transform $(\mathrm{WHT})_h$ and the second is Walsh-ordered Walsh-Hadamard transform $(\mathrm{WHT})_w$.

Haar transform (HT) is the last transform which will be presented in the paper.

All $(\mathrm{WHT})_h$, $(\mathrm{WHT})_w$ and HT of a data sequence X(m), m = 0, 1, ..., N-1, in matrix notation are defined as

$$\mathbf{B_x}(n) = \frac{1}{N}\mathbf{H}_i(n)\ \mathbf{X}(n)^T ,\ n = \log_2 N , \tag{2.2}$$

where $\mathrm{B_x}(k)$ is the k-th WHT or HT coefficient, $\mathbf{H}_i(n)$ is the kernel of $(\mathrm{WHT})_h$ - $\mathbf{H_h}(n)$, $(\mathrm{WHT})_w$ - $\mathbf{H_w}(n)$ or HT - $\mathbf{H_{HT}}(n)$, and $\mathbf{X}(n)$ is an N-vector of data X(m).

Fast algorithms to compute all the transforms described above are developed and presented in [1].

Two-dimensional transform computing of DCT is given by [1]

$$\mathbf{L_{xx}}(u_1,u_2) = \mathbf{H_{DCT}}(n_1)\mathbf{X}(m_1,m_2)\ \mathbf{H_{DCT}}(n_2)^T \tag{2.3}$$

where $\mathbf{X}(m_1, m_2)$ denotes the $(N_1 \times N_2)$ data matrix and $\mathbf{L_{xx}}(u_1, u_2)$ denotes the discrete cosine transform matrix.

The two-dimensional case of other transforms described above is given by [1]

$$\mathbf{S_{xx}}(u_1,u_2) = \frac{1}{N_1 N_2}\mathbf{H}_i(n_1)\mathbf{X}(m_1,m_2)\mathbf{H}_i(n_2)^T \tag{2.4}$$

where $\mathbf{S_{xx}}(u_1, u_2)$ is the $(\mathrm{WHT})_h$, $(\mathrm{WHT})_w$ or HT transform matrix, $\mathbf{H_i}(n_j)$ is the kernel of one of these transforms and $\mathbf{X}(m_1, m_2)$ denotes the data matrix.

The equation (2.5) describe the two-dimensional case of inverse DCT and equation (2.6) describe the two-dimensional case of other inverse transforms which are needed for decompression of pictures. The two-dimensional inverse DCT can be written as [1]

$$\mathbf{X}(m_1, m_2) = \mathbf{H}_{\mathbf{DCT}}^T(n_1)\mathbf{L}_{\mathbf{xx}}(u_1, u_2)\mathbf{H}_{\mathbf{DCT}}(n_2) \qquad (2.5)$$

and the two-dimensional case of other inverse transforms can be defined by [1]

$$\mathbf{X}(m_1, m_2) = \frac{1}{N_1 N_2}\mathbf{H}_i^T(n_1)\mathbf{S}_{\mathbf{xx}}(u_1, u_2)\mathbf{H}_i(n_2). \qquad (2.6)$$

3. Mathematical evaluation of results

It is necessary to evaluate two factors for comparison of results obtained by compression and decompression of original pictures. The first is a compress ratio and the second is a quality of decompressed pictures.

In whole this paper, the compression ratio (CR) is calculated as a ratio of number of all pixels in the picture f_{all} to number of all non-zero frequency coefficients f_{nz} obtained by transformation the picture from spatial to frequency domain. It can be written as

$$CR = \frac{f_{all}}{f_{nz}}. \qquad (3.1)$$

Two possibilities exist for the quality evaluation. The results can be compare visually and mathematically. The Normalised Root Mean Square Error (NRMSE) is defined for the mathematical comparison and it is given by [2]

$$NRMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(X_i - \hat{X}_i\right)^2}{\sum_{i=1}^{n}\left(X_i^2\right)}} \qquad (3.2)$$

where X_i are original pixel values, $\hat{X}_i$ are pixel values after decompression, and n is the total number of pixels in picture.

4. Experimental results of using the DCT, WHT and HT on few black and white pictures

The transforms were tried on selected pictures, which are shown in Fig. 4.1 and 4.2. The pictures were divided into blocks and block size varied from 8x8 to 64x64 pixels. The achieved compression ratio was changed too and it is necessary to recall that only one type of quantization table was used for all transforms. The changes of the compression ratio were performed by changing of threshold value after quantization.

The dependence of NRMSE on the block size and CR for both „The rowanberries“ and „The vertical stripes“ using the DCT and $(WHT)_H$ respectively is diagrammatized in Figs. 4.3 till 4.6. The $(WHT)_W$ gives very similar results to $(WHT)_H$ and HT gives worse results in every cases. Therefore, this results are not presented in this paper in the concrete.

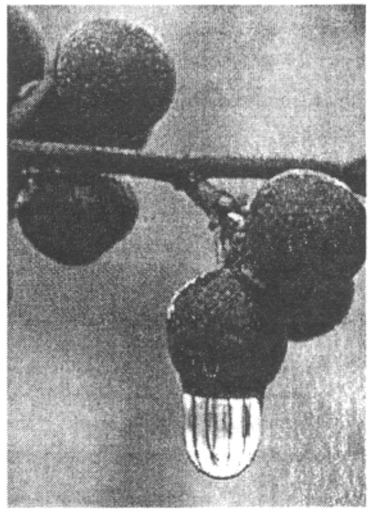

Fig. 4.1: The rowanberries

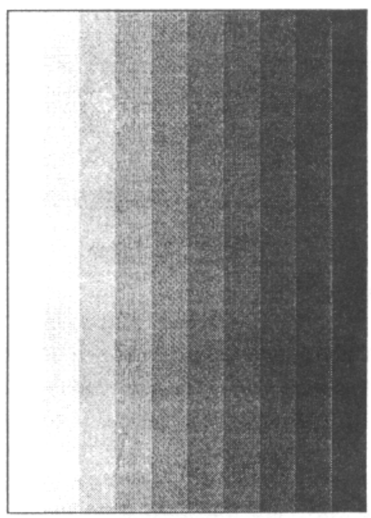

Fig. 4.2: The vertical stripes

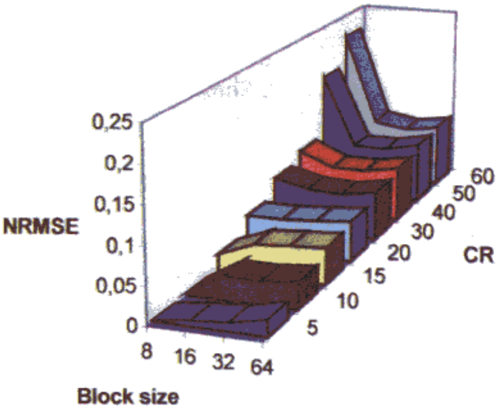

Fig. 4.3: The relation between NRMSE, Block size and CR for "The rowanberries" using the DCT

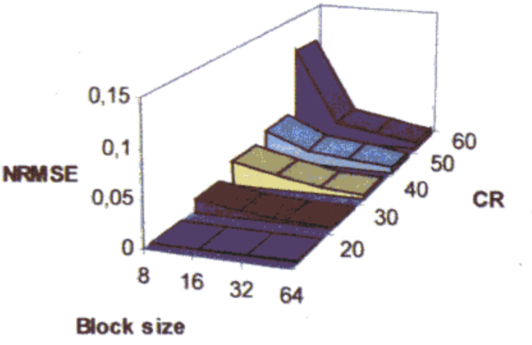

Fig. 4.4: The relation between NRMSE, Block size and CR for "The vertical stripes" using the DCT

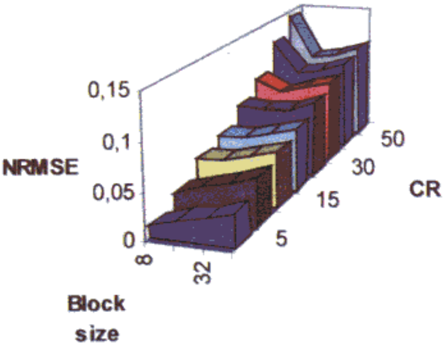

Fig. 4.5: The relation between NRMSE, Block size and CR for "The rowanberries" using the $(WHT)_H$

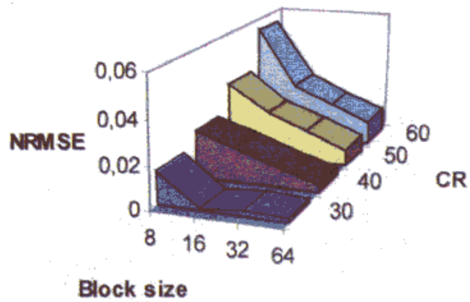

Fig. 4.6: The relation between NRMSE, Block size and CR for "The vertical stripes" using the $(WHT)_H$

5. Experimental results of using the DCT on coloured pictures

Coloured pictures was compressed and decompressed too. Fig. 5.1 to 5.4 show „Coloured rowanberries" and „Seattle" after compression and decompression. The pictures 5.1 and 5.3 was compressed using the same threshold level for luminance and both chrominance components while the threshold level for the chrominance components of the pictures 5.2 and 5.4 was 10× higher than for the luminance component. The CR reached for the pictures 5.2 and 5.4 is about twice higher than for the pictures 5.1 and 5.3. However, the quality of all the pictures is nearly the same.

DCT, CR=10.8, NRMSE=0.0356,
bs=32x32, Y, R-Y, B-Y-the same thr. 1

Fig. 5.1: Picture "Coloured rowanberries" – the same threshold level for the luminance and both chrominance components

DCT, CR=16.2, NRMSE=0.0338, bs = 32x32, R-Y, B-Y - 10x higher thr. level

Fig. 5.2: Picture „Coloured rowanberries" – the threshold level for the chrominance components 10× higher then for the luminance component

DCT, CR=9.2, NRMSE=0.1338, bs=64x64, Y, R-Y, B-Y - the same threshold level

Fig. 5.3: Picture „Seattle" – the same threshold level for the luminance and both chrominance components

DCT, CR=17.4, NRMSE=0.1398, bs=64x64, R-Y, B-Y - 10x higher threshold level

Fig. 5.4: Picture „Seattle" – the threshold level for the chrominance components 10× higher then for the luminance component

6. Conclusion

Many runs of the transforms were performed and many results were obtained. Some of them were presented in this paper. Both (WHT)h and (WHT)w give comparable results with DCT for pictures containing the large areas of the same brightness but generally the DCT yields the best issues of all transforms described above. The pictures containing the large areas of the same brightness can be compressed using larger pixel blocks while the pictures containing smaller areas have to be compressed using smaller pixel block. Therefore, it is suitable to use a variable block size dividing of pictures for achievement of high CR and very good quality [3]. An edge detection of pictures or limitation of maximum NRMSE are needed for this method.

It was shown that the chrominance components of coloured pictures can be compressed using 10x higher threshold level than the luminance component.

References

[1] N. Ahmed, K. R. Rao, "Orthogonal Transforms for Digital Signal Processing," Springer-Verlag, Berlin, 1975

[2] R. Westwater, B. Furth, "Real-Time Video Compression," Kluwer Academic Publishers, Massachusetts, 1997

[3] Y. B. Yu, M. H. Chan, and A. G. Constantinides, "Low bit rate video coding using variable block size model," in ICASSP-90, Int. Conf. Acoust., Speech, Signal Processing, Albuquerque, NM, Apr. 1990, pp. 2229-2232

[4] V. Chen, H. Smith, S. Fralick, "A Fast Computational Algorithm for the Discrete Cosine Transform," IEEE Transactions on Communications, Vol. COM-25, No. 9, September 1977, pp. 1004-1009

[5] V. Vít, "Televizní technika - barevné přenosové soustavy," BEN, Praha, 1997

About author

Ivo Chromý was born in 1974 in Znojmo, Czech Republic. He received an Ing. (MSc) degree in electrical engineering from Brno University of Technology, Czech Republic in 1997. Since October 1997 he has been studying at Institute of Radio Electronics, Faculty of Electrical Engineering, Brno University of Technology toward Ph.D. degree. His research interests include digital image processing and digital video compression.