

A low-complexity video coder based on the Discrete Walsh Hadamard Transform

R. Costantini,^{a,b} J. Bracamonte,^b G. Ramponi,^a J-L. Nagel,^b M. Ansorge,^b and F. Pellandini^b

^a Image Processing Laboratory, DEEI, University of Trieste, Via A. Valerio 10, I-34127 Trieste, Italy
 Email: costa@ipl.univ.trieste.it

^b Institute of Microtechnology, University of Neuchâtel, Rue A.-L. Breguet 2, CH-2000 Neuchâtel, Switzerland

ABSTRACT

This paper reports the results obtained of applying the Discrete Walsh-Hadamard Transform (DWHT) in a video coding scheme. It will be shown that for low bitrate applications the DWHT can reach performances analogous to those obtained with the Discrete Cosine Transform (DCT) in terms of compression efficiency, PSNR, and visual quality. This suggests the use of the DWHT in video coding systems where the computational cost is a fundamental issue, given the far less computational complexity of the DWHT with respect to the DCT.

1 INTRODUCTION

Due to its good performance in terms of compression efficiency, the DCT has been widely adopted in many standard coding schemes (for example, JPEG, MPEG1, MPEG2, and H.26x), despite its still relatively high computational cost.

The DWHT is an orthogonal transform whose basis functions consist of a set of rectangular discontinuous waveforms that can take the values -1, 0, and 1. It has been largely used for many image processing applications [1] like image enhancement, edge detection, spectral analysis, digital filtering and to some extent image compression. The main advantage of the DWHT is its low computational complexity, since only a small number of additions is required to compute it.

The results in this paper show that in the case of low bitrate video coding the performance of the DWHT is comparable to that obtained with the DCT in terms of compression, PSNR, and visual quality. This indicates that the DWHT can be favorably used in low bitrate applications, since its low computational cost constitutes an advantage with respect to the systems that are based on the DCT.

2 ENERGY COMPACTION PROPERTIES

2.1 Definition

The N -point 1-D DWHT of a discrete-time signal $x(n)$ is defined as :

$$y(k) = \frac{1}{N} \sum_{n=0}^{N-1} x(n)w_k(n) \text{ with } k = 0, 1, \dots, N-1$$

where $w_k(n)$ is the Walsh function [1] of order k , which is defined recursively as:

$$\begin{aligned} w_k[n] &= 0 && \text{for } n < 0 \text{ and } n > N-1 \\ w_0[n] &= 1 && \text{for } n = 0, 1, \dots, N-1 \\ w_{2k}[n] &= w_k[2n] + (-1)^k w_k[2n-1] \\ w_{2k+1}[n] &= w_k[2n] - (-1)^k w_k[2n-1] \end{aligned} \quad (1)$$

for $k = 0, 1, \dots, N-1$. The nature of these functions suggests that the DWHT can be suitable for coding images with high frequency features like edges or detailed texture [2].

2.2 Energy Compaction

Recalling that for linear transforms the energy in the spatial domain is equal to the energy in the frequency domain (Parseval's Relation), we have considered two functions in order to estimate the energy compaction ability of the DWHT.

For this purpose we have processed an input image, first, by dividing it into $N \times N$ -pixel blocks and then, by transforming each block with the DWHT. Thus, these two functions will be defined on a block-of-pixel basis.

The first considered function is the Energy Progressive Function (EPF), which is expressed as:

$$\text{EPF}(n) = \frac{1}{E_0} \sum_{i=0}^n y_i^2 \text{ with } n = 0, \dots, N^2 - 1.$$

It represents the percentage of energy contained in the first n coefficients of the transformed block, when ordered in a 1-D vector of N^2 elements according to a zigzag scanning. E_0 is the energy of the block and y_i the i -th transformed coefficient.

The second considered function is the Transformed Coefficients Variance (TCV) defined as:

$$\text{TCV}(n) = E[(y_n - \bar{y}_n)^2] \text{ with } n = 0, \dots, N^2 - 1$$

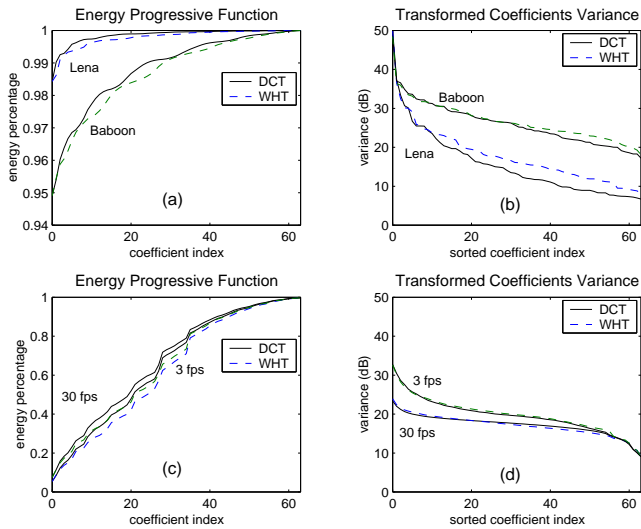


Figure 1: Energy compaction efficiency for the DWHT and the DCT.

where the mathematical expectation is taken over all the image blocks. It has been reported [2] that the lesser the variance, the better a transform approaches the optimal Karhunen-Loeve Transform (KLT) from the energy compaction point of view; hence, both considered functions give information about energy compaction.

To analyze the energy compaction properties of the DWHT, a comparison with the DCT has been made, by computing the two above described functions for both transforms. Images having different spectral characteristics like Lena, Baboon and a series of prediction error frames obtained from the sequence Hall-Monitor, have been used as test images.

Fig.1(a) and Fig.1(b) show the mean values of EPF(n) and TCV(n) obtained from the $N \times N$ blocks of the images Lena and Baboon. Fig.1(c) and 1(d) report the mean values obtained from the blocks of all the prediction error frames of the sequence Hall-Monitor, measuring the results of using two different frame rates.

These results illustrate that the DCT's and DWHT's performance, in terms of energy compaction, is very similar for the image Baboon and for the prediction error frames; while in the case of the image Lena, the DWHT is less effective than the DCT.

This is consistent with the observation made at the end of Section 2.1 about the characteristics of the DWHT basis functions, since the image Lena has a lesser high-frequency content than the image Baboon.

3 THE DWHT IN STILL IMAGE CODING

3.1 Coding scheme

To analyse the performance of the DWHT in still image coding, a compression scheme based on the JPEG standard has been used. This scheme is briefly described in

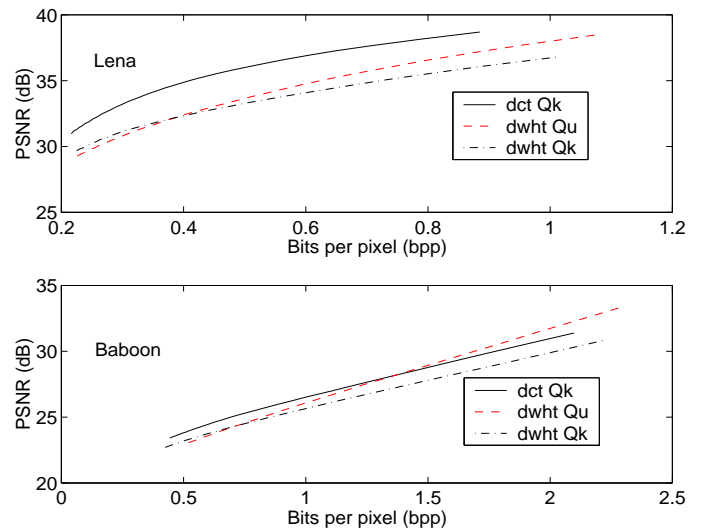


Figure 2: Rate-distortion curves obtained from encoding the images Lena and Baboon.

the following paragraph.

The input image is divided into 8×8 -pixel blocks; each of them is transformed and then quantized. The quantized block is then passed to an entropy coder which first executes a zigzag scanning over the quantized coefficient in order to obtain a 64-point 1-D vector and then performs a run-length code to compact long runs of zeros. The last operation is the Huffman encoding of the resulting symbols according to the established JPEG specifications.

3.2 Results

For comparison purposes we have coded the images Lena and Baboon using the quantization matrix that is proposed in the Annex K Specification of the JPEG Standard [3]. In this paper this matrix will be referred to as Q_k .

Fig.2 shows the rate-distortion curves of the images Lena and Baboon, obtained by varying the scale factor of this quantization matrix within the range $[0.5, 5]$. Since matrix Q_k is not optimized for the DWHT, Fig.2 also presents the resulting rate-distortion curves for the DWHT quantized with a uniform normalization matrix Q_u , i.e., a matrix whose coefficients have all the same value. These values have been allowed to vary from 14 to 90.

It can be noted that the DCT is more efficient than the DWHT for the image Lena in terms of PSNR and compression ratio. Nevertheless the visual quality of the reconstructed images is still comparable.

On the other hand, for the image Baboon the DWHT's performance is very close to the DCT's, even in the rate-distortion sense. This result confirms again the previous observation regarding the energy compaction properties of the DWHT, i.e., its performance is com-

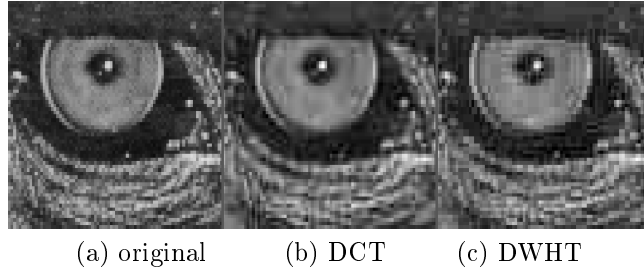


Figure 3: Region of the image Baboon.

parable to the DCT's when applied to high frequency content images, as shown in Fig.3.

Fig.3(b) and Fig.3(c) show the reconstruction quality of a region of the image Baboon. The complete image was compressed/decompressed with the DCT and the DWHT with a coding bitrate of 0.866 bpp in both cases. It can be noted that the low-frequency zones (e.g., the iris) are better reconstructed by the DCT, while the DWHT performs better on the high-frequency regions.

4 THE DWHT IN VIDEO SEQUENCE CODING

To reinforce the latter observation, a large number of prediction error frames has been coded following the reported scheme and using both transforms. Fig.4 reports a representative example of the results obtained. It can be noted that the DWHT's rate-distortion performance is very close to that obtained by using the DCT. This fact suggests to analyze the application of the DWHT to encode prediction error frames.

4.1 Coding Scheme

The motivation of using DWHT is to take advantage of its low computational cost. In order to keep the low complexity target in video coding, a very simple scheme has been used to study the performance of the DWHT. This scheme is based on Motion JPEG (MJPEG).

The first frame of the video sequence is intraframe coded, i.e., it is coded as an ordinary still image using the JPEG standard. The subsequent frames are interframe coded by applying the JPEG algorithm to the prediction error frames using a uniform quantization matrix.

To analyse the DWHT performance, we have applied both transforms in this video coding scheme and compared the mean PSNR of all the frames of the reconstructed sequences obtained at different bitrates. The QCIF Hall-Monitor sequence has been used as a test sequence.

When the first frame of the sequence is encoded using the same quantization matrix Q_k for both the DCT and the DWHT, the corresponding PSNR of the reconstructed images are different. Since in predictive coding the quality of the first reconstructed frame influences the mean PSNR over all the remaining frames, in order

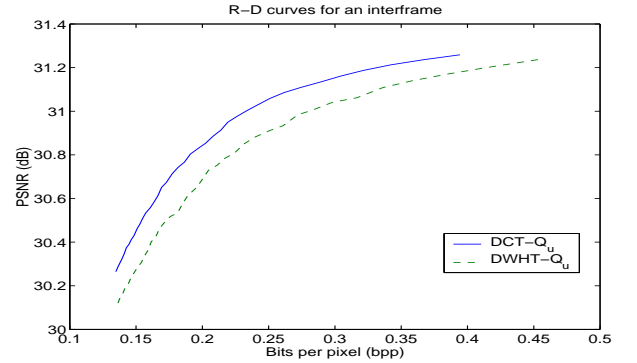


Figure 4: Rate-distortion curves obtained by coding a single error prediction frame.

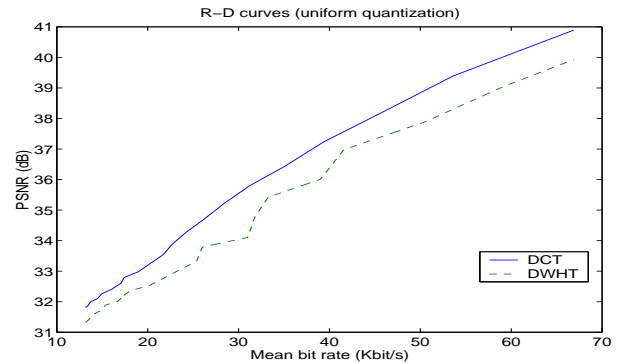
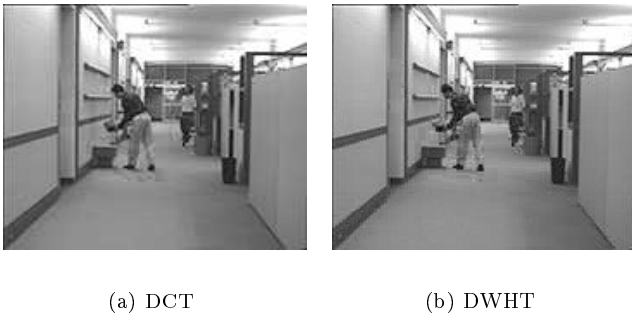


Figure 5: Coding performance of the DCT and the DWHT for the Hall-Monitor sequence.

to make a fair comparison, it has been chosen to start with a first reconstructed frame having the same PSNR for both cases. To accomplish this task two different quantization matrices have been used in the coding process of the first frame: Q_k for the DCT and Q_k scaled by a factor 0.7 for the DWHT. In this way the PSNR of the first reconstructed image is the same for both transforms.

4.2 Results

Figure 5 shows the rate-distortion curves obtained by using DCT and DWHT. It can be noted that for low bitrate values (10-20 Kbit/s) the performance is very similar for both transforms.



(a) DCT

(b) DWHT

Figure 6: Reconstructed frame of sequence Hall-Monitor.

8x8 2-D DCT		
method	multiplications	additions
Arai [4], 1-D (x2)	80	464
Linzer [5], 1-D (x2)	160	416
Winograd [6], 2-D	54	462
Dimitrov [7]	-	636
8x8 2-D DWHT		
method	multiplications	additions
Manz [8], 1-D (x2)	-	384

Table 1: Computational cost for various DCT and DWHT algorithms.

The reconstructed sequences were also evaluated from a visual quality point of view. It was found that the resulting DCT's and DWHT's reconstructed frames were nearly indistinguishable. Figure 6 shows a reconstructed frame of the sequence Hall-Monitor obtained by using both transforms. These results indicate that the DWHT could be favorably used in a video coding scheme, especially when high compression ratios are required in high frequency content frames.

5 COMPUTATIONAL COMPLEXITY

This section considers the computational complexity of the DCT and the DWHT, in terms of additions and multiplications required to compute the desired transform.

To compute the N -point 1-D DWHT many fast methods have been designed. Most of them present the same computational cost of $N \log_2 N$ additions [1], and they are all characterized by their regularly structured addition scheme. Their main difference is rather on the output order of the transformed coefficients. The 2-D DWHT can be obtained with the classical row-column approach requiring $2N^2 \log_2 N$ additions.

To compute the N -point 1-D fast DCT many algorithms have been proposed. Table 1 reports a short survey of such methods.

As noticed from the table, the DWHT does not re-

quire any multiplication and it involves a lesser number of additions than those needed by any of the DCT's fast schemes, which as pointed out before, shows the advantage of the DWHT from a computational point of view.

6 CONCLUSIONS

This paper presented the performances of the DWHT when applied to a video coding scheme. It has been shown that in low bitrate coding, the performance of the DWHT is comparable to that obtained with the DCT.

Thus, due to its good coding efficiency in video coding and to its low computational cost, the DWHT could be efficiently used in low bitrate video coding schemes, that demand low complexity implementations. Furthermore, the regularity of its fast algorithms can also be exploited to design reduced-die-area VLSI implementations. Particular applications of interest include tele-surveillance systems and video coding for miniature mobile systems.

Acknowledgements

The first author acknowledges the support received by COST 254 in the frame of a Short Term Scientific Mission, and by the Swiss Federal Office for Education and Science under Grant OFES C97.0050.

References

- [1] K.G. Beauchamp, "Applications of Walsh and Related Functions", Academic Press, London, UK, 1984.
- [2] N.S. Jayant and P. Noll "Digital coding of waveforms", Prentice-Hall, Englewood Cliffs, NJ, USA, 1984.
- [3] ITU-T Recommendation T.81, "Digital Compression and Coding of Continuous-tone Still Images", Sept. 1992.
- [4] Y. Arai, T. Agui and M. Nakajima, "A Fast DCT-SQ Scheme for Images", Trans. of the IEICE, Vol. E 71, No. 11, pp. 1095-1097, Nov. 1988.
- [5] E. Linzer and E. Feig, "New scaled DCT algorithms for fused multiply/add architectures", Proc. ICASSP-91, pp. 2201-2204, Toronto, Ont., Canada, May 1991.
- [6] E. Feig and S. Winograd, "Fast Algorithms for the Discrete Cosine Transform", IEEE Trans. on Signal Processing, Vol. 40, No. 9, pp. 2174-2193, Sept. 1992.
- [7] V.S. Dimitrov, G.A. Jullien and W.C. Miller, "A new DCT algorithm based on encoding algebraic integers", ICASSP-98, 12-15 May 1998, Seattle, WA, USA.
- [8] J.W. Manz, "A Sequence-Ordered Fast Walsh Transform", IEEE Trans. on Audio and Electroacoustics, Vol. AU-20, No. 3, pp. 204-206, Aug. 1972.