

Research Article

Warped Discrete Cosine Transform-Based Low Bit-Rate Block Coding Using Image Downsampling

Sarp Ertürk

Kocaeli University Laboratory of Image and Signal Processing (KULIS), Electronics and Telecommunication Engineering Department, University of Kocaeli, 41040 Kocaeli, Turkey

Received 18 May 2006; Revised 30 January 2007; Accepted 6 February 2007

Recommended by Mauro Barni

This paper presents warped discrete cosine transform (WDCT)-based low bit-rate block coding using image downsampling. While WDCT aims to improve the performance of conventional DCT by frequency warping, the WDCT has only been applicable to high bit-rate coding applications because of the overhead required to define the parameters of the warping filter. Recently, low bit-rate block coding based on image downsampling prior to block coding followed by upsampling after the decoding process is proposed to improve the compression performance for low bit-rate block coders. This paper demonstrates that a superior performance can be achieved if WDCT is used in conjunction with image downsampling-based block coding for low bit-rate applications.

Copyright © 2007 Sarp Ertürk. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

Block-based discrete cosine transform (DCT) encoders are incorporated into many image and video coding standards as a result of their high decorrelation performance and the availability of fast DCT algorithms enabling real-time implementation [1]. The DCT is used in the JPEG image coding standard, the MPEG-1 and MPEG-2 video coding standards, as well as the ITU-T H.261 and H.263 recommendations for real-time visual communications.

The warped discrete cosine transform (WDCT) is a cascade connection of conventional DCT and all-pass filters whose parameters are adjusted to provide frequency warping and thereby improve the coding performance [2]. WDCT-based compression is shown to outperform conventional DCT-based compression for high bit-rate applications. However, for low bit-rate applications the overhead required to encode the all-pass filter parameters of each block becomes significant in WDCT and the compression performance falls below conventional DCT.

Recently, it has been shown that downsampling before DCT coding and upsampling after decoding can improve the objective and subjective performance at low bit-rates [3]. If an image frame is downsampled before compression, the available amount of data per pixel that can be encoded in the transform coding increases for a fixed bit-rate so that the reconstructed image quality can be improved. A standard

anti-aliasing filter has been used as decimation filter, and a linear interpolation kernel has been used after decoding in [3]. It has been demonstrated in [4] that the performance can even be enhanced if optimal or near-optimal decimation and interpolation filters are used within this scheme. In [4], the optimal decimation and interpolation filters are determined based on least squares (LS) and it is shown that the performance is improved both visually and quantitatively compared to [3].

This paper demonstrates that WDCT can be used in conjunction with image downsampling-based block coding for low bit-rate applications to achieve a superior compression performance. The applicability of WDCT is enlarged into the low bit-rate range and the performance of downsampling-based block coding is enhanced by using WDCT in conjunction with image downsampling.

2. WDCT-BASED LOW BIT-RATE BLOCK CODING USING IMAGE DOWNSAMPLING

Downsampling-based block coding has been proposed in [3, 4] to improve the performance of block coders for low bit-rate applications. In [3], a standard anti-aliasing filter is used for downsampling, a linear interpolation kernel is used for upsampling, and the downsampling factor (k) is chosen according to analytic predictions. In [4], the downsampling

factor is set to $k = 2$ for simplicity while the decimation and interpolation filters are optimally determined according to least squares. This approach is also utilized in this paper, with WDCT being used instead of conventional DCT to form the compression system given in Figure 1. Note that this is the downsampling-based block encoding approach proposed in [4], with only the conventional DCT changed to WDCT. Here, \mathbf{f} shows the decimation kernel, \mathbf{g} represents the interpolation kernel, and k determines the downsampling/upsampling factor.

If the 8-point DCT $\{C_0, C_1, \dots, C_7\}$ of the input vector $[x_0, x_1, \dots, x_7]$ is defined as [5]

$$C_k = U(k) \frac{1}{\sqrt{8}} \sum_{n=0}^7 x_n \cos \frac{(2n+1)k\pi}{16}, \quad k = 0, \dots, 7, \quad (1)$$

where

$$U(k) = \begin{cases} \frac{1}{\sqrt{2}}, & k = 0 \\ 1, & \text{otherwise} \end{cases} \quad (2)$$

then it is possible to carry out the DCT computation using a filter bank, where each filter is given by

$$F_k(z^{-1}) = U(k) \frac{1}{\sqrt{8}} \sum_{n=0}^7 \cos \frac{(2n+1)k\pi}{16} z^{-n}, \quad k = 0, \dots, 7 \quad (3)$$

so that the i th coefficient of $F_k(z^{-1})$ is the (k, i) th element of the DCT matrix. Note that in this case the signal block should be time reversed before filtering. While the conventional DCT performs well for inputs with low-frequency components, the coding efficiency deteriorates in cases of high-frequency content. It has been proposed in [2] to warp the input frequency to adjust the frequency distribution of the input to be more suitable for DCT. A first-order all-pass filter with transfer function

$$A(z) = \frac{-\alpha + z^{-1}}{1 - \alpha z^{-1}} \quad (4)$$

is used to perform the warping by replacing z^{-1} in (1) with $A(z)$. The frequency warping is controlled using the α parameter, and therefore it is required to send this parameter as side information. The WDCT can be expressed using a filter bank in the form of

$$F_k(A(z)) = U(k) \sum_{n=0}^7 \cos \frac{(2n+1)k\pi}{16} (A(z))^n, \quad k = 0, \dots, 7. \quad (5)$$

Two methods have been suggested in [2], for the implementation of the WDCT. The first approach expands the filters $(A(z))^k$ for every k , and then obtains the WDCT matrix by a matrix-vector multiplication using the conventional DCT matrix. The second approach consists of constructing an 8-tap FIR filter that approximates $F_k(A(z))$ using eight equally spaced samples of $F_k(A(e^{j\Omega}))$ computed using the

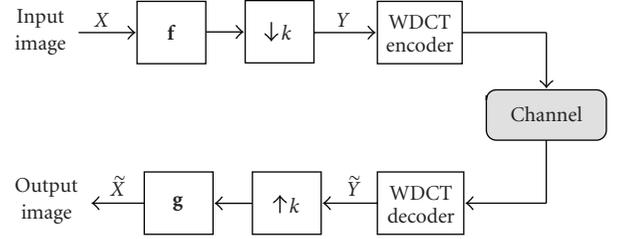


FIGURE 1: WDCT-based low bit-rate block coding using image downsampling.

inverse discrete Fourier transform (IDFT). The second approach is noted to provide a slightly better performance and is therefore utilized in this paper.

In a similar approach to [2], $2N$ approximated WDCT matrices are prepared using a set of warping parameters with values $\alpha = n/10N$, $n = -N, \dots, N - 1$. The WDCT matrix for $n = 0$ will be equal to the conventional DCT matrix, and therefore the conventional case will be included in the WDCT. For each image block, every WDCT matrix is tried and the one that gives the lowest reconstruction error is selected.

The index of the WDCT matrix (i.e., the index corresponding to the best value of the control parameter α) is sent to the decoder as side information and therefore results in data overhead. If the number of WDCT matrices is increased (a larger N is used) the warping process is enhanced resulting in superior transform coding, however, the required side information as well as the computational load will be increased as well. It is shown in [2] that for a constant bit-rate there is actually no gain in increasing the number of WDCT matrices beyond 16, and therefore $N = 8$ (corresponding to 16 WDCT matrices) is also utilized in this paper.

As noted in [4], the optimal decimation filter is typically obtained to be a lowpass filter with extremely high cutoff frequency, which is essentially an identity filter. Therefore it is basically possible to ignore the optimization process for the decimation filter \mathbf{f} and simply avoid filtering prior to downsampling, in order to preserve the texture that dominates the image. In this case it is very simple to obtain the interpolation filter \mathbf{g} by least squares by

$$\min_{\mathbf{g}} \|X - \tilde{X}\|_2^2 = \min_{\mathbf{g}} \|X - \mathbf{g} * \tilde{Y}_u\|_2^2, \quad (6)$$

where \tilde{Y}_u shows the upsampled WDCT decoded image and $*$ represents convolution with the filter kernel.

3. EFFECT OF QUANTIZATION

It is noted in [2] that frequency weighting is already accomplished by the warping process, and therefore it is more appropriate to utilize a uniform quantizer instead of the standard JPEG quantization matrix given in Table 1. The JPEG quantization matrix is designed by taking the visual response to luminance variations into account, as a small variation in

TABLE 1: JPEG quantization table for the luminance channel.

16	11	10	16	24	40	51	61
12	12	14	19	26	59	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

intensity is more visible in slowly varying regions (i.e., low spatial frequency) than in busier ones (i.e., high spatial frequency) [6]. As the WDCT accomplishes frequency warping, it is noted in [2] that a uniform quantizer is more appropriate for WDCT.

In order to evaluate the influence of the quantization matrix on the proposed approach, compression results using the downsampling-based WDCT approach proposed in this paper are evaluated for both quantization approaches. Figure 2 shows the peak signal-to-noise ratio (PSNR) against compression bit-rate results for the *Barbara* image with different types of quantizers. It is seen that uniform quantization indeed provides a better performance compared to the standard quantization approach generally, particularly for higher bit-rates. However, for extremely low bit-rates the standard quantizer shown in Table 1 can outperform uniform quantization. As [2] uses WDCT for medium- to high-bit-rate applications it is therefore natural that uniform quantization is preferred in [2]. Because it is aimed in this paper to utilize WDCT for low-bit-rate applications, however, uniform quantization does not seem to be the best solution.

Uniform quantization performance clearly falls below standard quantization performance for extremely low bit-rates. It is therefore proposed in this paper to utilize an “in-between” quantization approach. While the quantization matrix should still quantize low frequencies with higher resolution compared to high frequencies (because of the visual response to luminance variations), the balance should not be as excessive as in the standard case because frequency warping is already accomplished. Hence the quantization matrix given in Table 2 is utilized.

The compression performance of the proposed downsampling-based WDCT approach with the quantizer given in Table 2 is also shown in Figure 2 for the *Barbara* image. It is seen that the proposed quantizer performs as well as uniform quantization for higher bit-rates and outperforms both uniform as well as standard quantization for low bit-rates, while the performance is similar to standard quantization at extremely low bit-rates.

4. EXPERIMENTAL RESULTS

In order to evaluate the performance of the proposed approach, compression results using standard JPEG, downsampling-based DCT as proposed in [4] (denoted as DS-DCT), and downsampling-based WDCT as proposed in

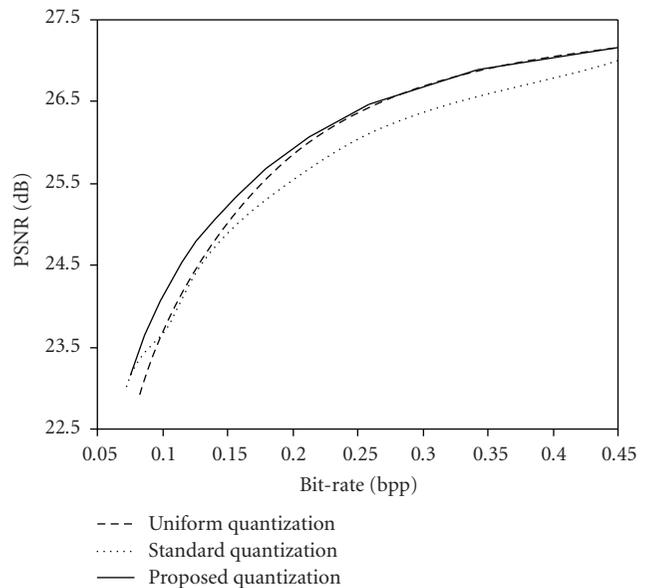
FIGURE 2: PSNR versus bit-rate for the *Barbara* image of size 512×512 for various quantization approaches.

TABLE 2: Proposed quantization table for the luminance channel.

40	40	40	40	60	60	80	80
40	40	40	40	60	60	80	80
40	40	40	40	60	60	80	80
40	40	40	40	60	60	80	80
60	60	60	60	60	60	80	80
60	60	60	60	60	60	80	80
80	80	80	80	80	80	80	80
80	80	80	80	80	80	80	80

this paper with uniform quantization (denoted as DS-WDCT-Qu) as well as the proposed quantization (denoted as DS-WDCT-Qp) for various images and various bit-rates are evaluated.

Figure 3 shows peak signal-to-noise ratio (PSNR) results for the *Barbara* image. It is seen that the proposed downsampling-based WDCT approach outperforms downsampling based conventional DCT and also outperforms standard JPEG at low bit-rates. The proposed quantization approach improves the performance of the proposed DS-WDCT approach for very low bit-rates compared to uniform quantization.

Figure 4 shows peak signal-to-noise ratio (PSNR) results for the *Lena* image. The proposed downsampling-based WDCT technique is again seen to provide a superior compression performance compared to downsampling-based conventional DCT. For very low bit-rates, downsampling-based DS-WDCT performs significantly better than standard JPEG. It is seen that the proposed quantization approach is again more suitable than uniform quantization.

Figure 5 shows the peak signal-to-noise ratio (PSNR) versus bit-rate results for the *Cameraman* image. The proposed

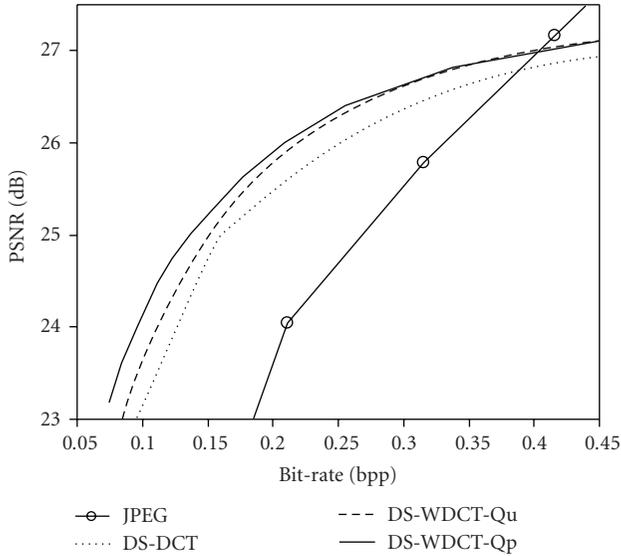


FIGURE 3: PSNR versus bit-rate for the Barbara image of size 512×512 .

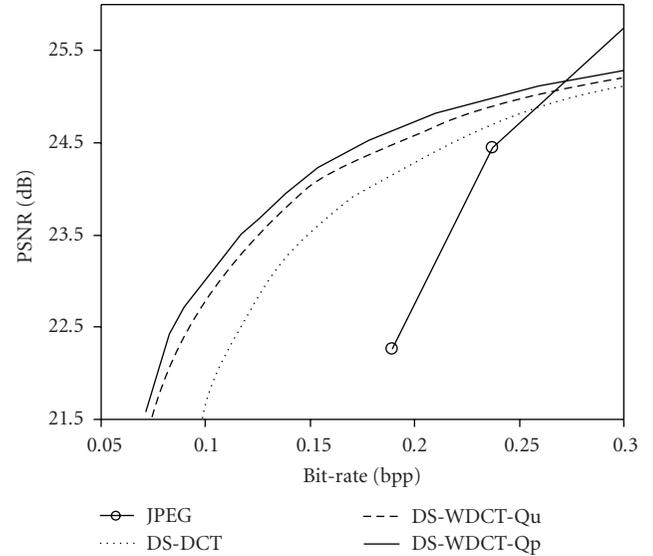


FIGURE 5: PSNR versus bit-rate for the Cameraman image of size 256×256 .

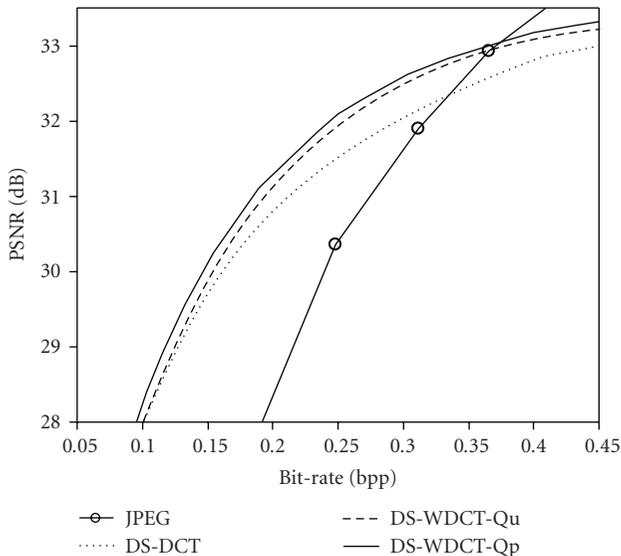


FIGURE 4: PSNR versus bit-rate for the Lena image of size 512×512 .

downsampling-based WDCT encoding approach again outperforms DS-DCT as well as standard DCT. The advantage of the proposed quantization approach is also observed in these results.

Figure 6 shows sample decoded versions of the *Barbara* image in order to provide visual evaluation. It is seen that the proposed quantization matrix not only improves the PSNR of the reconstructed decoded image but also provides superior visual results.

The main reason why WDCT can be used for low bit-rates when combined with image downsampling to provide an improved performance, while the WDCT performance

without downsampling is lower than standard DCT for low-bit-rates and WDCT surpasses standard DCT only for high bit-rates is the overhead amount. In WDCT, typically the quantized control parameter has to be sent as side information for each block. Because a total of 16 WDCT matrices are utilized, the overhead is 4 bits per block. Normally this overhead prohibits the use of WDCT for standard block coding low bit-rate applications (i.e., if downsampling is not utilized), as in this case a 4-bit overhead is required for each 8×8 block so that the WDCT parameter overhead equals to $4/(8 \times 8) = 0,0625$ bits per pixel which is naturally an important overhead if it is desired to encode at very low bit-rates in the range of typically 0.1–0.3 bpp in the first place.

The downsampling process reduces the total number of blocks that need to be transform coded and therefore significantly reduces the WDCT overhead per pixel. For a downsampling factor of $k = 2$, which is used in the results provided in this paper, a control parameter has to be sent as side information for each 8×8 block of the downsampled image and as this corresponds to a block of size 16×16 in the original image size, the overhead of WDCT will only be $4/(16 \times 16) = 0,015625$ bits per pixel. Hence it becomes possible to utilize WDCT to achieve enhanced representation of transform coefficients to improve the compression performance.

5. CONCLUSION

This paper shows that a superior compression performance for low-bit-rate applications can be achieved by using WDCT in conjunction with image downsampling-based block coding. Instead of using conventional DCT, the frequency warping of WDCT enhances the reconstructed image quality and therefore results in improved performance. While an overhead is required to send the control parameter of each block

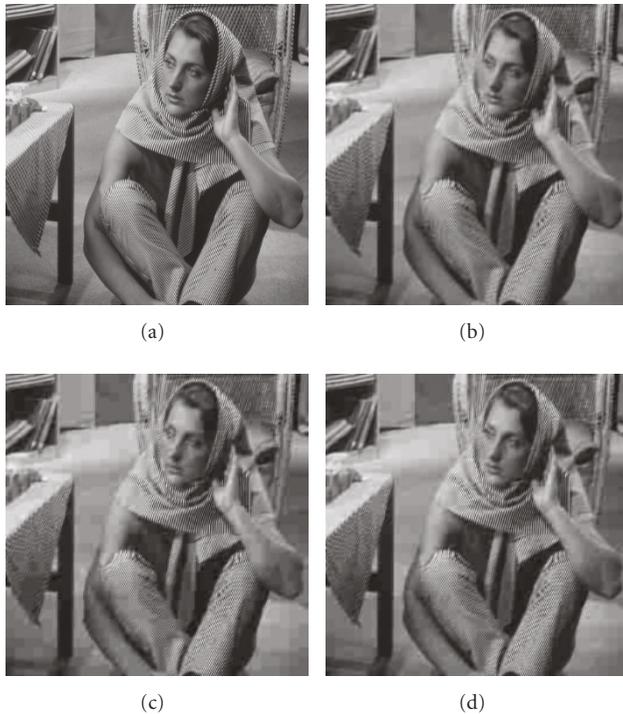


FIGURE 6: Visual results for the Barbara image of size 512×512 . (a) Original, (b) encoded using DS-DCT at 0.175 bpp, PSNR 25.21 dB, (c) encoded using DS-WDCT-Qu at 0.175 bpp, PSNR 25.46 dB, (d) encoded using DS-WDCT-Qp at 0.175 bpp, PSNR 25.56 dB.

in the case of WDCT, the downsampling process reduces the number of blocks that are transform coded and therefore significantly reduces the total overhead, thereby facilitating the use of WDCT in low-bit-rate applications. Because it is possible to implement the WDCT using standard DCT hardware [2], the proposed approach can be utilized in systems that already have conventional DCT hardware installed so as to improve the performance for low bit-rate applications.

REFERENCES

- [1] T. Sikora, "Trends and perspectives in image and video coding," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 6–17, 2005.
- [2] N. I. Cho and S. K. Mitra, "Warped discrete cosine transform and its application in image compression," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 8, pp. 1364–1373, 2000.
- [3] A. M. Bruckstein, M. Elad, and R. Kimmel, "Down-scaling for better transform compression," *IEEE Transactions on Image Processing*, vol. 12, no. 9, pp. 1132–1144, 2003.
- [4] Y. Tsaig, M. Elad, P. Milanfar, and G. H. Golub, "Variable projection for near-optimal filtering in low bit-rate block coders," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 1, pp. 154–160, 2005.
- [5] R. Gonzales and R. Woods, *Digital Image Processing*, Addison-Wesley, Reading, Mass, USA, 1992.
- [6] A. M. Tekalp, *Digital Video Processing*, Prentice-Hall, Upper Saddle River, NJ, USA, 1995.

Sarp Ertürk graduated in 1995 from the Electrical-Electronic Engineering Department of Middle East Technical University (M.E.T.U.). In 1996, he completed his M.S. degree at Essex University, UK, in telecommunication and information systems with a T.E.V. scholarship. He earned his Ph.D. degree again from Essex University in 1999, in the field of electronics system engineering with a Y.Ö.K. scholarship. He started his compulsory military service in 1999, which he completed in April 2001 as a Lecturer at the Army Academy. From April 2001 to November 2002, he worked as Assistant Professor at the Electronics and Telecommunication Engineering Department of the University of Kocaeli. He has been appointed as Associate Professor in the same department since November 2002. In the beginning of 2003, he established KULIS (Kocaeli University Laboratory of Image and Signal processing) research laboratory. He has lectured a postgraduate class and directed research at Chung-Ang University, Korea, between March–September 2006. He is carrying out research in the areas of image and video processing, signal processing, and digital telecommunications, and has directed national and international projects and published numerous journal and conference papers.

