Approximating the DCT with the Lifting Scheme: Systematic Design and Applications

Jie Liang and Trac D. Tran
Department of Electrical and Computer Engineering, The Johns Hopkins University
Baltimore, MD 21218. Email: jieliang@jhu.edu, trac@jhu.edu.

Abstract

A systematic approach to design two families of multiplierless approximations of the DCT with the lifting scheme is presented, based on Chen's and Loeffler's factorizations of the DCT matrix, respectively. The analytical values of all the lifting steps are derived, which can be approximated by dyadic values to enable fast implementations with only shifts and additions. Different trade-offs between the complexity and the performance can be easily obtained. A scaled lifting structure is proposed to further reduce its complexity. The performance of the lifting-based DCT implementation is demonstrated in the frameworks of JPEG and H.263+. Besides, lossless compression capability is also presented.

1 Introduction

The Discrete Cosine Transform (DCT) is the underlying technique for many image and video compression standards such as JPEG, H.263 and MPEG, due to its high energy compaction capability and the existence of numerous fast DCT algorithms.

Most fast DCT algorithms can be classified into two categories: (i) fast algorithms taking advantage of the relationships between the DCT and various existing fast transforms such as the FFT, Walsh-Hadamard transform (WHT), and discrete Hartley transform (DHT); (ii) fast algorithms based on the sparse factorizations of the DCT matrix [2, 7].

The theoretical lower bound on the number of multiplications required in the 1D 8-point DCT has been proven to be 11 [4], which is achieved by the method proposed by Loeffler *et. al.* [7]. However, the complexity can be further reduced in lossy image and video coding where quantization is involved in the processing, and hence it is possible to incorporate some multiplications in the quantization steps. This is the so-called *scaled DCT* [8, 1].

However, these fast DCT algorithms still need floating-point or fixed-point multiplications, which are usually costly in both hardware and software implementations. In this paper, we will present two families of multiplierless approximations of the DCT with the lifting scheme [3], a powerful tool for constructing wavelets and wavelet transforms.

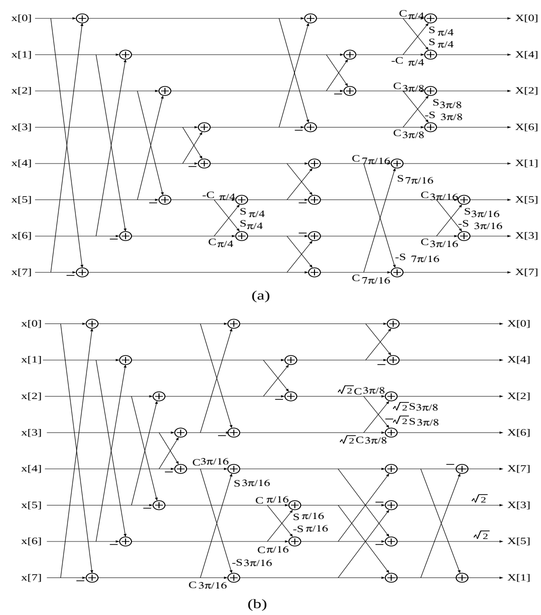


Figure 1. Givens rotation-based factorizations of the 8-point DCT: (a) Chen's factorization; (b) Loeffler's factorization.

Besides, motivated by the idea of the scaled DCT, a scaled lifting structure is used to further reduce the complexity of the approximations. The property of the lifting scheme also enables lossless compression with the new method, which is impossible in other DCT-based algorithms. Experimental results will be presented to demonstrate the performance of these fast DCT algorithms in JPEG and H.263+.

2 DCT Factorizations and the Lifting Scheme

Fig. 1 shows the signal flow graphs of Chen's [2] and Loeffler's [7] factorizations of the 1D 8-point DCT matrix. The overall complexities of these factorizations are 13 multiplications, 29 additions, and 11 multiplications and 29 additions. Note that a scaling factor of $1/2$ and $1/\sqrt{8}$ should be applied to the output of each factorization to obtain true DCT coeffi-

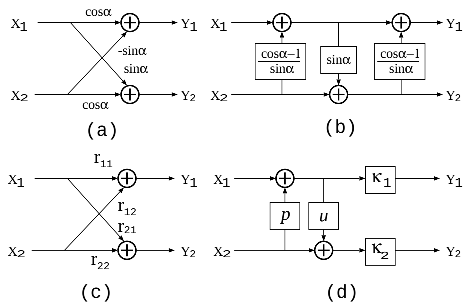


Figure 2. (a) A Givens rotation; (b) Representing by three lifting steps; (c) General form; (d) Representing by the scaled lifting structure.

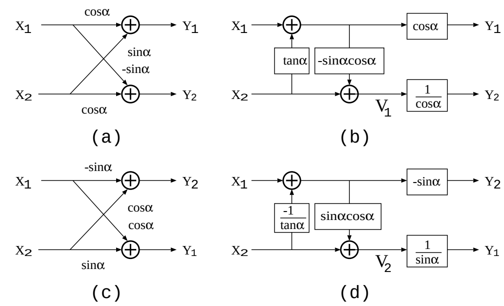


Figure 3. (a) A Givens rotation; (b) Scaled lifting solution; (c) Permuted Givens rotation; (d) Scaled lifting solution for the permuted rotation.

cients.

The common feature of these factorization is that most of the multiplications are in the form of Givers rotations. This enables the application of the lifting scheme. Fig. 2(a) and (b) show the representation of a Givens rotation by three lifting steps (assuming $\alpha \neq 0$) [3]. This can be written in matrix form as:

$$\begin{bmatrix} C\alpha & -S\alpha \\ S\alpha & C\alpha \end{bmatrix} = \begin{bmatrix} 1 & \frac{C\alpha-1}{S\alpha} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ S\alpha & 1 \end{bmatrix} \begin{bmatrix} 1 & \frac{C\alpha-1}{S\alpha} \\ 0 & 1 \end{bmatrix}, \tag{1}$$

where $C\alpha$ and $S\alpha$ stand for $cos\alpha$ and $sin\alpha$, respectively.

Each lifting step is a biorthogonal transform, and its inverse also has a simple lifting structure. Moreover, to inverse a lifting step, we simply need to subtract out what was added at the forward transform. This implies that the original signal can still be perfectly reconstructed when the lifting parameters are approximated by finite-length values, as long as the same value is used in both the forward and the inverse transforms. Hence a multiplierless approximation of the DCT can be obtained if we replace each Givens rotation in the DCT factorizations by lifting steps, and then approximate each lifting parameter by computer-friendly dyadic values, which can be easily implemented by shift and addition operations. Due to this reason, we denote the resulted implementation of the DCT as the *binDCT*.

3 The scaled lifting structure

Note that some of the rotation angles in Fig. 1 are at the end of the transform. This makes it possible to further reduce the complexity of the binDCT by adjusting the aforementioned lifting structure and combining some of the operations in the quantization stage.

In Fig. 2(c) and (d), we represent a Givens rotation by 2 lifting steps and 2 scaling factors. Since the two scaling factors can be absorbed in the quantization, only 2 lifting steps are left in the transform, leading to a more efficient representation than the conventional one as shown in Fig. 2(b).

By comparing Fig. 2(c) and (d), the analytical values of the parameters in the scaled lifting structure can be easily derived as follows:

$$\begin{aligned} p &= \frac{r_{12}}{r_{11}}, \\ u &= \frac{r_{11}\, r_{21}}{r_{11}\, r_{22} - r_{21}\, r_{12}}, \\ \kappa_1 &= r_{11}, \\ \kappa_1 &= \frac{r_{11}\, r_{22} - r_{21}\, r_{12}}{r_{11}}. \end{aligned} \tag{2}$$

With this formula, the Givens rotation as given in Fig. 3(a) can be represented by the scaled lifting structure in Fig. 3(b), which can then be approximated by finite-length values for fast implementations.

However, it can be shown that when the rotation angle is close to $\pi/2$, the performance of the above approximation is very sensitive to the roundoff error of the lifting parameters [6]. Besides, the lifting parameter $\tan\alpha$ is quite large in this case, which increases the dynamic range of the transform. In this case, we can apply a permutation to the output signals, as shown in Fig. 3(c). The corresponding scaled lifting structure is given in Fig. 3(d), which can provide robust approximations as well as smaller dynamic range.

Since perfect reconstruction is guaranteed by the lifting structure itself, the remaining problem is to determine the lifting parameters such that the binDCT could have similar coding performance to that of the true DCT. In this paper, the performance is measured by the coding gain, which represents the energy compaction capability of a transform. Under the assumption of the $AR(1)$ input with zero-mean, unit variance and inter-sample autocorrelation coefficient $\rho = 0.95$, the coding gain of the 8-point DCT is $8.8259\,dB$.

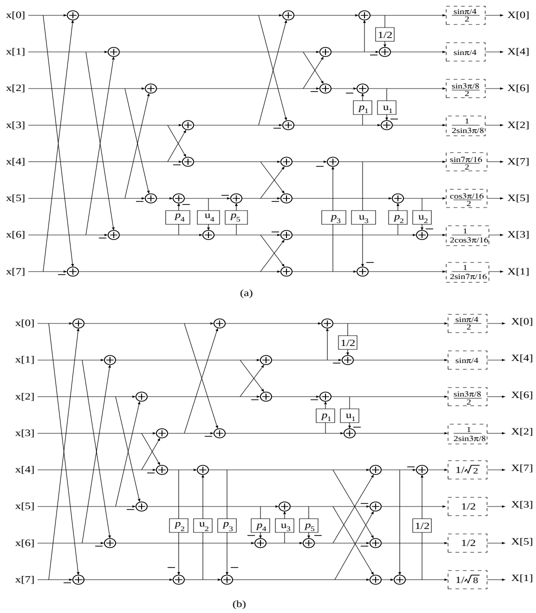


Figure 4. General structure of the binDCT from (a) Chen's factorization; (b) Loeffler's factorization.

4 General Structures and Design Examples

From the above analysis, we can obtain the general structures of the binDCT from Chen's and Loeffler's factorizations, as shown in Fig. 4, where the intermediate rotation angles are implemented by the normal 3-lifting structure, and those rotations at the end of the signal flow are replaced by more efficient scaled lifting structures. Note that the permuted version of the scaled lifting structure is used for angles close to $\pi/2$.

The scaling factors in the dash boxes will be absorbed by the quantization step in the implementation. They will be bypassed if lossless compression is desired. Note that some sign manipulations are involved here to make all the scaling factors positive.

The analytical values of all the lifting parameters can be obtained from the above formulas. These results can be approximated by dyadic values, which enables simple implementation with only shift and addition operations. This has tremendous advantages in software and hardware implementations, especially for low-power hand-held devices such as PDA and cellular phones.

In Table 1 and 2 , we present several possible configurations of the two types of binDCT. C_g stands for the coding gain of the binDCT. In general, with increasing complexity, the multiplierless binDCT can approximate the floating DCT more and more accurately. Note that even the 9-shift version has a pretty good coding gain of $8.7686\,dB$ already. It also shows that given the same complexity, the binDCT from Loeffler's factorization has a slightly better performance than that from the Chen's factorization, due to the reduced number of Givens rotations in the factorization.

5 Experimental results

The proposed binDCT families have been implemented in the frameworks of JPEG and H.263+, based on popular public domain software [5, 10]. We replace the DCT by the proposed binDCT, and all quantization-related parts are modified to incorporate the binDCT scaling factors.

5.1 Performance of the binDCT in JPEG

Fig. 5 plots the PSNRs of the reconstructed Lena image with different DCT implementations. The binDCT-C1 and binDCT-C4 in Table 1 are compared with the floating DCT and the fast integer DCT in the IJG's code [5]. The latter is a scaled DCT and requires 5 fixed-point multiplications and 29 additions [8, 1]. In obtaining these results, the same kind of transform is used in both the forward and the inverse DCT.

It can be seen that the performances of both binDCT configurations are very close to that of the floating DCT. In particular, when the quality factor is below 90, the maximal difference between binDCT-C4 and the floating DCT is less than $0.1dB$, which is negligible. The maximal performance degradation of binDCT-C1 is also below $0.5dB$. Besides, when the quality factor is above 90, the performance of the proposed binDCT is much better than the fast integer DCT in the IJG's code.

As previously mentioned, lossless compression can be easily achieved with the lifting-based binDCT by bypassing the scaling factors. To improve the compression ratio, we replace all the butterflies in the binDCT by the lifting steps as shown in Fig. 6, which minimizes the dynamic range of the transform. For instance, the DC coefficient in the modified structure is the average of all the inputs, while in the original structure, it is the summation of the inputs.

The lossless binDCT has been implemented together with two coding methods, Huffman and SPIHT [9]. To use Huffman coding, a new Huffman table is designed, since the default Huffman table defined in the JPEG standard is not optimal for the lossless binDCT, as the statistic distribution of the binDCT coefficients are different from that of the original DCT coefficients. When SPIHT is used, we rearrange the binDCT coefficients according to the pattern of the wavelet transform coefficients before applying zerotree processing.

The binDCT-C4-based lossless transforms are compared with some advanced context model-based prediction methods, including the HP LOCO-I and CALIC [11, 12], and the results are summarized in Table 3, showing that the overall compression ratio of the binDCT-based method is not as good

Table 1. Some configurations of the binDCT from Chen's factorization.

Config.	p_1	u_1	p_2	u_2	p_3	u_3	p_4	u_4	p_5	Shifts	Adds	C_g (dB)
binDCT-C1	1/2	1/2	1	1/2	1/4	1/4	1/2	3/4	1/2	9	28	8.7686
binDCT-C2	1/2	3/8	7/8	1/2	3/16	1/4	7/16	3/4	3/8	14	33	8.8033
binDCT-C3	3/8	3/8	7/8	1/2	3/16	3/16	7/16	11/16	3/8	17	36	8.8159
binDCT-C4	7/16	3/8	5/8	7/16	3/16	3/16	7/16	11/16	3/8	19	37	8.8220
binDCT-C5	13/32	11/32	11/16	15/32	3/16	3/16	7/16	11/16	3/8	21	40	8.8233
binDCT-C6	7/16	3/8	5/8	7/16	3/16	3/16	13/32	11/16	13/32	21	39	8.8240
binDCT-C7	13/32	11/32	11/16	15/32	3/16	3/16	13/32	11/16	13/32	23	42	8.8251

Table 2. Some configurations of the binDCT from Loeffler's factorization.

Config.	p_1	u_1	p_2	u_2	p_3	p_4	u_3	p_5	Shifts	Adds	C_g (dB)
binDCT-L1	1/2	1/2	1/4	1/2	1/4	1/8	1/4	1/8	10	28	8.7716
binDCT-L2	3/8	1/2	1/4	1/2	1/4	1/8	3/16	3/32	13	31	8.8027
binDCT-L3	7/16	3/8	1/4	9/16	5/16	1/8	3/16	3/32	16	34	8.8225
binDCT-L4	13/32	11/32	5/16	9/16	5/16	3/32	3/16	3/32	20	38	8.8242
binDCT-L5	13/32	11/32	19/64	9/16	19/64	3/32	3/16	3/32	22	40	8.8257

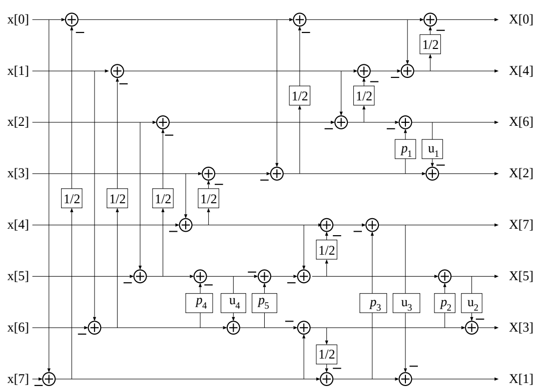
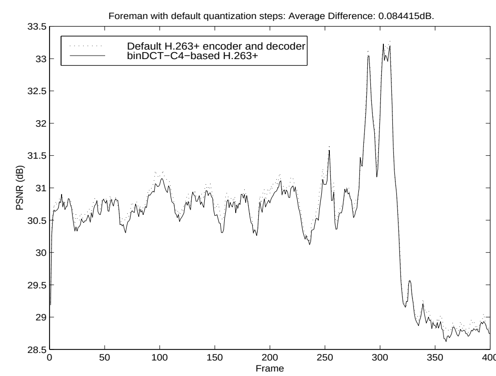


Figure 6. Lossless binDCT from Chen's factorization with minimized dynamic range.

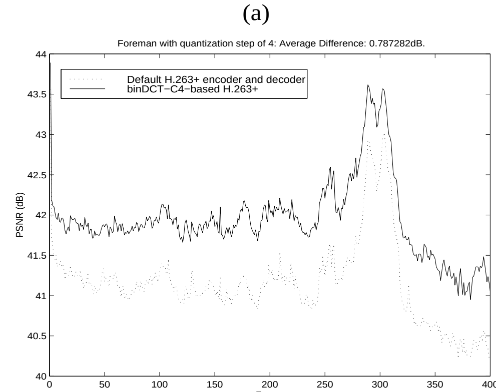
as these methods. However, the proposed binDCT is much simpler.

5.2 Performance of the binDCT in H.263+

The forward DCT in the selected H.263+ software is based on Chen's factorization with floating-point multiplications, and the inverse DCT is a scaled version of the same method with fixed-point multiplications. In this experiment, the binDCT-C4 is employed to compare with the original H.263+, and some results are shown in Fig. 7. It can be seen that with the default quantization steps (40 for I frames and 26 for P frames), the average PSNR of the binDCT-C4 based H.263+ is $0.08dB$ below that of the default H.263+. However, the compression ratio is improved to $102.67 : 1$ from $101.03 : 1$. The opposite scenario happens when small quantization step is used. For example, with a quantization



(a)



(b)

Figure 7. Comparison of the H.263+ and the binDCT-C4-based H.263+: (a) PSNRs of the reconstructed sequence with the default quantization; (b) PSNRs with a quantization step of 4 for all frames.

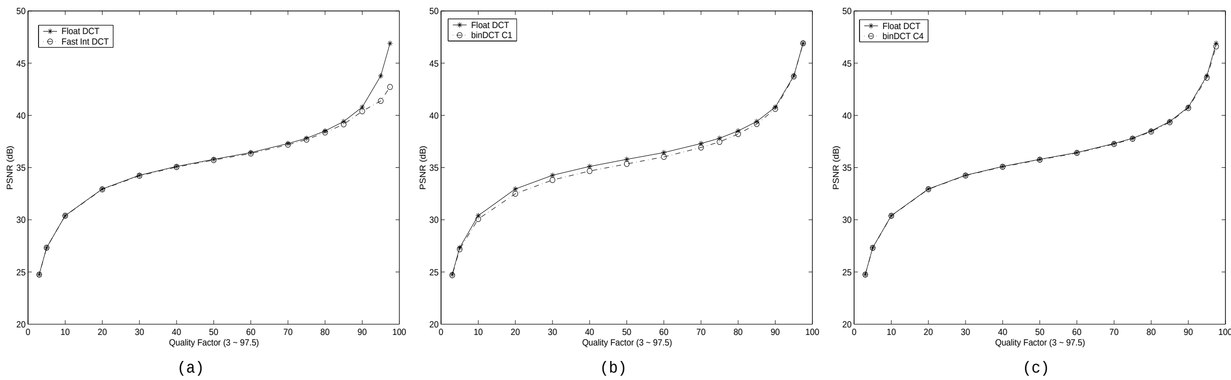


Figure 5. Comparison of the PSNRs resulted from (a) the floating DCT and the fast integer DCT in [5]; (b) the floating DCT and the binDCT-C1; (c) the floating DCT and the binDCT-C4.

Table 3. Lossless coding results (bits/pixel)

Image	binDCT-C4 + Huffman	binDCT-C4 + SPIHT	HP LOCO-I	CALIC
Balloon	3.78	3.58	2.90	2.78
Zelda	4.44	4.33	3.89	3.69
Hotel	5.20	5.07	4.38	4.18
Barbara	5.22	5.11	4.69	4.31
Board	4.34	4.24	3.68	3.51
Girl	4.60	4.50	3.93	3.72
Gold	5.20	5.04	4.48	4.35
Boats	4.67	4.56	3.93	3.78
Average	4.68	4.55	3.99	3.79

step of 4 for all frames, the PSNR given by the binDCT-C4 is $0.79dB$ higher than that of the default H.263+, but the file size increases by 2.5%. Hence the overall performance of the binDCT-C4 is very similar to the default DCT in the selected H.263+ implementation.

The compatibility of the binDCT-based H.263+ with the default H.263+ is also satisfactory. When the test sequence Foreman is encoded with the default DCT and default quantization steps, and decoded by the binDCT-C4, the average PSNR would be $30.42dB$, which is only $0.12dB$ below the $30.54dB$ provided by the default H.263+ decoder.

6 Conclusion

A systematic approach to design two families of fast multiplierless approximations of the DCT with the lifting scheme is presented, based on factorizations of the DCT matrix. The analytical values of all lifting parameters are derived, enabling appropriate finite-length approximations, which can be implemented with shift and addition operations. A scaled lifting structure is also proposed to further reduce the complexity of the transform. The performance of the proposed fast DCT is demonstrated in the framework of JPEG and H.263+. Experimental results show that the binDCT is a promising fast DCT algorithm that is especially suitable for software and hardware implementations on low-power handheld devices.

References

[1] Y. Arai, T. Agui, and M. Nakajima. A fast dct-sq scheme for images. *Trans. IEICE*, E-71(11):1095, 1988.

[2] W. Chen, C. Harrison, and S. Fralick. A fast computational algorithm for the discrete cosine transform. *IEEE Trans. Communications*, Vol. COM-25, No. 9:1004–1011, 1977.

[3] I. Daubechies and W. Sweldens. Factoring wavelet transforms into lifting step. *Journal Fourier Anal. Appl.*, Vol. 4, No. 3:247–269, 1998.

[4] E. Feig and S. Winograd. On the multiplicative complexity of discrete cosine transform. *IEEE Trans. Information Theory*, Vol. 38:1387–1391, Jul. 1992.

[5] IJG. *ftp://ftp.uu.net/graphics/jpeg*.

[6] J. Liang and T. D. Tran. Fast multiplierless approximation of the dct with the lifting scheme. *Proceeding of the 45th SPIE Annual Meeting*, Aug. 2000.

[7] C. Loeffler, A. Lightenberg, and G. Moschytz. Practical fast 1d dct algorithms with 11 multiplications. *Proc. IEEE ICASSP*, Vol. 2:988–991, Feb. 1989.

[8] W. Pennebaker and J. Mitchell. *JPEG Still Image Data Compression Standard*. Van Nostrand Reinhold, New York, 1993.

[9] A. Said and W. Pearlman. A new fast and efficient image codec based on set partitioning in hierarchical trees. *IEEE Trans, Circuits and Systems for Video Tech.*, Vol 6, No. 3, June 1996.

[10] UBC. *http://spmg.ece.ubc.ca/h263plus/h263plus.html.*

[11] M. J. Weinberger, G. Seroussi, and G. Sapiro. The loco-i lossless image compression algorithm: principles and standardization into jpeg-ls. *IEEE Trans. Image Processing*, Vol. 9 No. 8, Aug. 2000.

[12] X. Wu and N. D. Memon. Context-based, adaptive, lossless image coding. *IEEE Trans. Commun.*, Apr. 1997.