Multiplierless approximate 4-point DCT VLSI architectures for transform block coding

F.M. Bayer, R.J. Cintra, A. Madanayake and U.S. Potluri

Two multiplierless algorithms are proposed for 4×4 approximate-discrete cosine transform (DCT) for transform coding in digital video. Computational architectures for one-dimensional (1D)/2D realisations are implemented using Xilinx field programmable gate array devices. CMOS synthesis at the 45 nm node indicates real-time operation at 1 GHz yielding 4×4 block rates of 125 MHz at < 120 mW of dynamic power consumption.

Introduction: Video and multimedia processing based on signal and image compression such as the high-efficiency video coding and H.265 reconfigurable video codecs require two-dimensional (2D) transform block coding for block sizes $N \times N$ where $N \in \{4, 8, 16, 32, 64\}$ [1]. The transform coding stage requires algorithms for the N-point discrete cosine transform (DCT) of types II and IV. The associate transformation matrices are defined, respectively, according to [2]

$$[C_{\mathrm{II}}]_{(m,n)} = \sqrt{\frac{2}{N}}\, \alpha_m \cos\left[\left(m - \frac{1}{2}\right) \frac{\pi(n-1)}{N}\right]$$

$$[C_{\mathrm{IV}}]_{(m,n)} = \sqrt{\frac{2}{N}} \cos\left[\left(m - \frac{1}{2}\right)\left(n - \frac{1}{2}\right)\frac{\pi}{N}\right]$$

where $m, n = 1, 2, \ldots, N$, $\alpha_1 = 1/\sqrt{2}$ and $\alpha_m = 1$, for $m > 1$.

In this Letter, our goal is to propose multiplication-free approximations for the 4-point DCT-II and -IV as well as its fast algorithms. We also aim at very large scale integration (VLSI) realisations of both the 1D and 2D versions of the derived approximate transforms, while maintaining high numerical accuracy and low computational complexity.

Optimisation and orthogonalisation: Let $\mathcal{M}_P(4)$ be the set of all 4×4 matrices whose entries are defined over $P = \{-1, 0, 1\}$. In this set, all the matrices represent multiplierless transforms. Our goal is to find matrices in $\mathcal{M}_P(4)$ that satisfactorily approximate C_{II} and C_{IV}.

Therefore, we propose the following multivariate nonlinear optimisation problem over $\mathcal{M}_P(4)$:

$$C_k^* = \underset{A \in \mathcal{M}_P(4)}{\arg\min}\, error(A, C_k), \quad k \in \{\mathrm{II}, \mathrm{IV}\} \qquad (1)$$

where C_k^* are the optimal matrices and error$(\cdot, \cdot)$ is an error measure between a given candidate matrix and the exact matrices C_{II} and C_{IV}.

Let $h_i[n]$ be the discrete signal formed by the ith row of a given matrix T and the discrete-time Fourier transform (DTFT) of $h_i[n]$ be denoted by $H_i(\omega; T)$. As discussed in [3, 4], we adopted the total error energy as the error measure. This particular measure is defined as follows:

$$\epsilon(A, C_k) = \sum_{m=1}^{4} \int_0^\pi \left| H_m(\omega;\, A) - H_m(\omega;\, C_k) \right|^2 \, d\omega$$

for $k \in \{\mathrm{II}, \mathrm{IV}\}$. In other words, $\epsilon(A, C_k)$ quantifies the sum of the energy error in the DTFT domain – between A and C_k – when the entries of a given matrix row are interpreted as the filter coefficients [3, 4]. This quantity can be computed numerically by standard quadrature methods [5].

As an additional constraint to (1), we impose that the matrix $A \cdot A^\top$ must be a diagonal matrix to ensure that orthogonality can be achieved in the obtained approximations [6]. The resulting constrained optimisation problem is algebraically intractable and we resorted to an exhaustive computational search.

Proposed 4-point DCT approximations: By solving (1), we obtained the following new DCT approximations:

$$C_{\mathrm{II}}^* = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & -1 \\ 1 & -1 & -1 & 1 \\ 0 & -1 & 1 & 0 \end{bmatrix} \text{ and } C_{\mathrm{IV}}^* = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & -1 & -1 \\ 1 & -1 & 0 & 1 \\ 0 & -1 & 1 & -1 \end{bmatrix}$$

Although possessing very low complexity, these matrices are not orthogonal. In several contexts, such as image processing for coding, orthogonality is often a desirable property [2]. By adopting the

orthogonalisation methods detailed in [6], new orthogonal matrices $\widehat{C}_{\mathrm{II}}$ and $\widehat{C}_{\mathrm{IV}}$ can be derived based on C_{II}^* and C_{IV}^*, respectively. These orthogonal matrices are given by

$$\widehat{C}_{\mathrm{II}} = D_{\mathrm{II}} \cdot C_{\mathrm{II}}^* \text{ and } \widehat{C}_{\mathrm{IV}} = D_{\mathrm{IV}} \cdot C_{\mathrm{IV}}^*$$

where $D_{\mathrm{II}} = \sqrt{\left[C_{\mathrm{II}}^* \cdot \left(C_{\mathrm{II}}^*\right)^\top\right]^{-1}}$ and $D_{\mathrm{IV}} = \sqrt{\left[C_{\mathrm{IV}}^* \cdot \left(C_{\mathrm{IV}}^*\right)^\top\right]^{-1}}$. Explicitly, we obtain that

$$D_{\mathrm{II}} = \mathrm{diag}\left(\frac{1}{2},\ \frac{1}{\sqrt{2}},\ \frac{1}{2},\ \frac{1}{\sqrt{2}}\right)$$

and

$$D_{\mathrm{IV}} = \frac{1}{\sqrt{3}} \cdot I_4$$

where I_4 is the identity matrix of size 4. In the image compression context, the scaling matrices D_1 and D_2 may not introduce any computational overhead, because they can be merged into the quantisation step, as described in [4, 7–9].

The signal flow graph for C_{II}^* and C_{IV}^* is shown in Fig. 1. We note that the C_{II}^* and C_{IV}^* transforms require only 6 and 8 additions, respectively. Multiplications or bit-shifting operations are totally absent. The resulting approximations $\widehat{C}_{\mathrm{II}}$ and $\widehat{C}_{\mathrm{IV}}$ are very close to the respective ideal DCT and offer extremely low complexities. In Table 1, we show the error measure and arithmetic complexity for the proposed transforms, the exact DCT computation [2] and the well-known signed DCT [3].
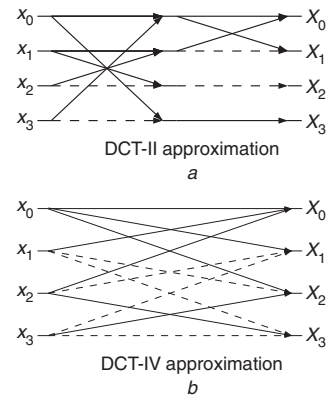
Fig. 1 *Signal flow graph for proposed transforms*
a DCT-II approximation
b DCT-IV approximation

Table 1: Total error energy and arithmetic complexity analysis

Methods	Error energies	Complexities		
		Add.	Mult.	Total
Exact 4-point DCT-II [2]	0.000	8	4	12
4-point signed DCT-II [3]	0.957	8	0	8
Proposed $\widehat{C}_{\mathrm{II}}$	0.957	6	0	6
Exact 4-point DCT-IV [2]	0.000	12	8	20
4-point signed DCT-IV [3]	2.359	10	0	10
Proposed $\widehat{C}_{\mathrm{IV}}$	0.838	8	0	8

Table 2: Resource consumption on Xilinx XC6VSX475T-2FF1156

Proposed approx.	CLB	FF	LUT	Slices	$F_{\max}$ (MHz)	D_{p} (W)
1D DCT-II	56	76	92	35	743.5	0.535
1D DCT-IV	76	132	128	52	735.3	0.574
2D DCT-II	166	408	330	108	704.2	0.884
2D DCT-IV	210	528	472	148	689.2	0.921

Field programmable gate array (FPGA) prototypes: The approximate DCTs were realised as an architecture for the 4-point 1D transforms and extended to the 4×4 2D transform. The inputs were assumed at an 8 bit resolution. Rapid prototypes were realised on a Xilinx Virtex-6 FPGA device and tested to ensure correct on-chip functionality. The results concerning the consumption of configurable logic blocks

(CLBs), flip–flops (FFs), look-up tables (LUTs) and slices are shown in Table 2. The maximum operating frequency (F_{max}) and dynamic power consumption (D_p) are also displayed.

The register transfer language code corresponding to the FPGA-verified designs were targeted to a 45 nm CMOS standard cell process using Cadence encounter. The CMOS designs were realised up to the synthesis and place-and-route levels leading to the estimated results in Table 3. Area–time complexities AT and AT^2 were adopted and measured in μm^2 ns and μm^2 ns^2, respectively.

Table 3: Resource consumption for 45 nm CMOS

Proposed approx.	ASIC gates	Area (μm^2)	F_{max} (GHz)	D_p (mW)	AT	AT^2
1D DCT-II	849	3386.9	1.10	6.31	3160	2948
1D DCT-IV	1207	4870.4	1.00	8.62	4846	4822
2D DCT-II	7400	31 217.8	0.95	59.33	7770	8159
2D DCT-IV	13 770	59 052.5	0.94	115.66	14 596	15 472

Conclusion: Numerical optimisation methods have led to the 4-point approximations for DCT-II and DCT-IV. Such matrices are tailored for minimal computational complexity and are adequate for computing realisations linked to coding operations with applications in digital video and multimedia. Fast algorithms were derived and the associate physical realisations do not require VLSI area- and power-intensive multiplier circuits. Both 1D and 2D realisations were proposed with the FPGA prototypes for architecture validation and CMOS synthesis results at the 45 nm node. The results indicate a real-time blockrate of 125 MHz for processing 4×4 blocks at a 1 GHz clock frequency.

F.M. Bayer (*Departamento de Estatística and Laboratório de Ciências Espaciais de Santa Maria (LACESM), Universidade Federal de Santa Maria, Santa Maria, RS, Brazil*)

R.J. Cintra (*Signal Processing Group, Departamento de Estatística, Universidade Federal de Pernambuco, PE, Brazil*)

E-mail: rjdsc@stat.ufpe.org

A. Madanayake and U.S. Potluri (*ECE, The University of Akron, Akron, OH, USA*)

References

1 Sullivan, G.J., Ohm, J.-R., Han, W.-J., and Wiegand, T.: 'Overview of the high efficiency video coding (HEVC) standard', *IEEE Trans. Circuits Syst. Video Technol.*, 2012, **22**, pp. 1649–1668
2 Britanak, V., Yip, P., and Rao, K.R.: 'Discrete cosine and sine transforms' (Academic Press, Elsevier, Oxford, UK, 2007)
3 Haweel, T.I.: 'A new square wave transform based on the DCT', *Signal Process.*, 2001, **82**, pp. 2309–2319
4 Cintra, R.J., and Bayer, F.M.: 'A DCT approximation for image compression', *IEEE Signal Process. Lett.*, 2011, **18**, pp. 579–582
5 Piessens, R., deDoncker-Kapenga, E., Uberhuber, C., and Kahaner, D.: 'Quadpack: a subroutine package for automatic integration' (Springer-Verlag, 1983)
6 Cintra, R.J.: 'An integer approximation method for discrete sinusoidal transforms', *J. Circuits Syst. Signal Process.*, 2011, **30**, (6), pp. 1481–1501
7 Bouguezel, S., Ahmad, M.O., and Swamy, M.N.S.: 'Low-complexity 8×8 transform for image compression', *Electron. Lett.*, 2008, **44**, pp. 1249–1250
8 Bayer, F.M., and Cintra, R.J.: 'DCT-like transform for image compression requires 14 additions only', *Electron. Lett.*, 2012, **48**, pp. 919–921
9 Bouguezel, S., Ahmad, M.O., and Swamy, M.N.S.: 'A low-complexity parametric transform for image compression'. Proc. 2011 IEEE Int. Symp. Circuits and Systems, Rio de Janeiro, Brazil, May 2011