A DCT Approximation for Image Compression

Renato J. Cintra, *Senior Member, IEEE*, and Fábio M. Bayer

Abstract—An orthogonal approximation for the 8-point discrete cosine transform (DCT) is introduced. The proposed transformation matrix contains only zeros and ones; multiplications and bit-shift operations are absent. Close spectral behavior relative to the DCT was adopted as design criterion. The proposed algorithm is superior to the signed discrete cosine transform. It could also outperform state-of-the-art algorithms in low and high image compression scenarios, exhibiting at the same time a comparable computational complexity.

Index Terms—DCT approximation, image compression, low-complexity transforms.

I. INTRODUCTION

THE 8-point discrete cosine transform (DCT) is a key step in many image and video processing applications. This particular blocklength is widely adopted in several image and video coding standards, such as JPEG, MPEG-1, MPEG-2, H.261, and H.263 [1]. This is mainly due to its good energy compaction properties, which are closely related to the Karhunen–Loève transform [2], [3].

During decades, much has been done to devise fast algorithms for the DCT. This is illustrated in several prominent works including [4]–[7]. In particular, the DCT design proposed by Arai *et al.* [6] became popular and has been implemented in several different hardware architectures [8], [9]. Nevertheless, all these algorithms require several multiplication operations. Past years have seen very few advances in the proposition of new low-complexity algorithms for the *exact* DCT calculation. A possible exception is the arithmetic cosine transform, whose mathematical background was recently proposed, but much is yet to be developed in terms of practical implementation [10].

In this scenario, signal processing community turned its focus to approximate algorithms for the computation of the 8-point DCT. While not computing the DCT exactly, approximate methods can provide meaningful estimations at low-complexity requirements. Prominent techniques include the signed discrete cosine transform (SDCT) [11], the Bouguezel–Ahmad–Swamy (BAS) series of algorithms [12]–[16], and the level 1 approximation by Lengwehasatit–Ortega [17]. All above mentioned techniques possess extremely low arithmetic complexities.

R. J. Cintra is with the Departamento de Estatística, Universidade Federal de Pernambuco, Recife, PE, Brazil and also with the Department of Electrical and Computer Engineering, University of Akron, Akron OH 44325 USA (e-mail: rjdsc@stat.ufpe.org).

F. M. Bayer is with the Universidade Federal da Fronteira Sul, Erechim, Brazil (e-mail: bayerfm@uffs.edu.br).

In this context, a new theoretical framework for DCT approximate transforms was proposed by Cintra [18]. The implied transformations are orthogonal and are based on polar decomposition methods [18], [19]. The aim of this correspondence is to introduce a new low-complexity DCT approximation for image compression in conjunction with a quantization step. After quantization, the resulting approximate coefficients are expected to be close to the ones furnished by the exact DCT. We restrain our attention to matrices that are good DCT approximations.

II. DCT ROUND-OFF APPROXIMATIONS

The proposed approximation method modifies the standard DCT matrix $\mathbf{C}$ by means of the rounding-off operation. Initially, matrix $\mathbf{C}$ is scaled by two and then submitted to an element-wise round-off operation. Let $[\cdot]$ denote the round-off operation as implemented in Matlab programming environment [20]. Thus, the resulting matrix, $\mathbf{C}_0 = [2 \cdot \mathbf{C}]$, is shaped as follows:

$$\mathbf{C}_0 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & -1 & -1 & -1 \\ 1 & 0 & 0 & -1 & -1 & 0 & 0 & 1 \\ 1 & 0 & -1 & -1 & 1 & 1 & 0 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 \\ 0 & -1 & 1 & 0 & 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & -1 & 1 & -1 & 1 & 0 \end{bmatrix}.$$

Matrix $\mathbf{C}_0$ has some attractive computational properties: i) its constituent elements are 0, 1, or -1, which is an indication of null multiplicative complexity; ii) as a transformation, it requires only additions, being bit-shift operations absent; and iii) its scaled transpose can perform an approximate inversion, making it a quasi-symmetrical tool (cf. [11]). In fact, a coarse approximation for the DCT matrix is achieved by $\hat{\mathbf{C}} = (1/2)\mathbf{C}_0$. Notice that the presence of the scaling factor $1/2$ represents only bit-shifts. In [18], the scaling factor that minimizes the Frobenius norm to the exact DCT matrix was found to be 0.3922.

The good features of $\mathbf{C}_0$ could enable such simple approximation matrix $\hat{\mathbf{C}}$ to outperform the SDCT in a wide range of practical compression ratios [21]. However, the suggested approximation has some drawbacks: i) it lacks orthogonality, since $\mathbf{C}_0^{-1} \neq \mathbf{C}_0^{\top}$, where superscript $^{\top}$ denotes matrix transposition, and ii) its resulting approximation is poor when compared with some existing methods (e.g., the BAS algorithms).

This framework encourages a more comprehensive analysis of the discussed approximation. Considering matrix polar decomposition theory [19], an adjustment matrix $\mathbf{S}$ that orthogonalizes $\mathbf{C}_0$ is sought. Indeed, the referred orthogonalization matrix is given by $\mathbf{S} = \sqrt{(\mathbf{C}_0 \cdot \mathbf{C}_0^{\top})^{-1}}$,
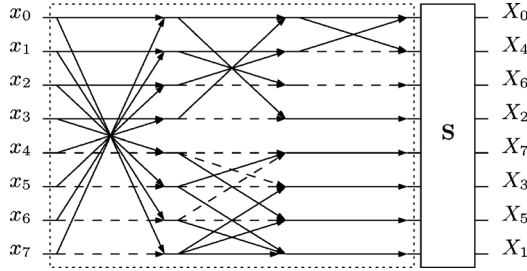
Fig. 1. Flow diagram of the proposed fast algorithm, relating input data x_n, $n = 0, 1, \ldots, 7$ to the corresponding approximate DCT coefficients X_k, $k = 0, 1, \ldots, 7$. Dotted box computes $\mathbf{C}_0$. Dashed arrows represent multiplication by -1.

TABLE I
ARITHMETIC COMPLEXITY ANALYSIS

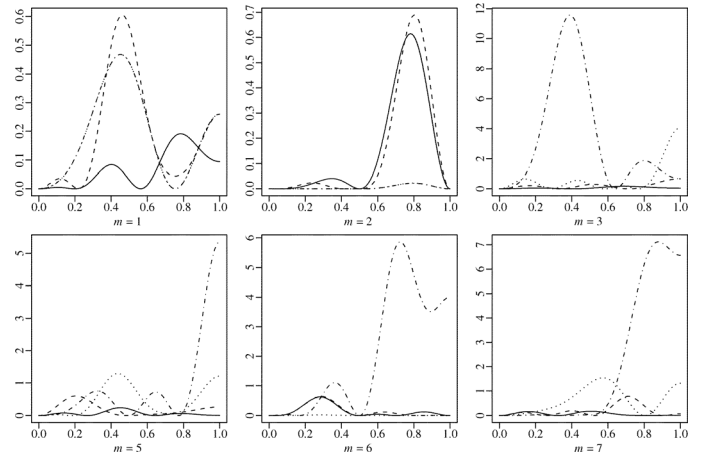| Method	Add.	Mult.	Shifts	Total
Proposed Approximation	22	0	0	22
BAS-2008 Algorithm [13]	18	0	2	20
BAS-2011 Algorithm [16]	18	0	2	20
SDCT [11]	24	0	0	24
BAS [12]	21	0	0	21
BAS [14]	18	0	0	18
BAS [15]	28	0	4	32
Level 1 Approximation [17]	24	0	2	26

Fig. 2. Normalized plots of $D_m(\omega; \mathbf{T})$ for the proposed algorithm (solid line), the SDCT (dashed line), the BAS-2008 algorithm (dotted line), and the BAS-2011 algorithm (dot-dashed line).

where the matrix square root is taken in the principal sense [22, p. 20]. This computation furnishes the diagonal matrix $\mathbf{S} = \mathrm{diag}(1/2\sqrt{2}, 1/\sqrt{6}, 1/2, 1/\sqrt{6}, 1/2\sqrt{2}, 1/\sqrt{6}, 1/2, 1/\sqrt{6})$. Therefore, the DCT matrix can be more adequately approximated by the following proposed matrix: $\hat{\mathbf{C}}_{\mathrm{orth}} = \mathbf{S} \cdot \mathbf{C}_0$. Matrix $\hat{\mathbf{C}}_{\mathrm{orth}}$ possesses useful properties: i) it is orthogonal; ii) it inherits the low computational complexity of $\mathbf{C}_0$; and iii) the orthogonalization matrix $\mathbf{S}$ is diagonal.

In terms of complexity assessment, matrix $\mathbf{S}$ may not introduce an additional computational overhead. For image compression, the DCT operation is a pre-processing step for a subsequent coefficient quantization procedure. Therefore, the scaling factors in the diagonal matrix $\mathbf{S}$ can be merged into the quantization step. This procedure is suggested and adopted in several works [13], [17]. As a consequence, the computational complexity of $\hat{\mathbf{C}}_{\mathrm{orth}}$ is ultimately confined to $\mathbf{C}_0$.

A fast algorithm for the transformation matrix $\mathbf{C}_0$ was devised and is depicted in Fig. 1. Arithmetic complexity assessment and comparisons are shown in Table I in terms of addition, multiplication, and bit-shift counts. The proposed algorithm is less complex than the SDCT and is only requires two additional operations when compared to the best BAS algorithms.

For an initial comparison screening, we separate i) the SDCT due to its very well documented literature [2], [11]; ii) the BAS algorithms introduced in [13] and [16] (BAS-2008 and BAS-2011, respectively), which are regarded the best BAS methods [14]; and the exact DCT. The BAS-2011 algorithm was considered with its parameter set to 0.5. The approximation proposed in [17] and [15] were not considered because of their comparatively higher computational complexity (Table I).

The proposed approximate matrix is more closely related to the DCT than the other approximations. In fact, following the methodology suggested in [11], the spectral structure, and energy compaction characteristics could be assessed. For such, we

understand each row of the transformation matrix as coefficients of a FIR filter. Thus, the transfer function related to each row of a given transformation matrix $\mathbf{T}$ could be calculated according to

$$H_m(\omega; \mathbf{T}) = \sum_{n=0}^{7} t_{m,n} \exp(-jn\omega), \quad m = 0, 1, \ldots, 7$$

where $j = \sqrt{-1}$, $\omega \in [0, \pi]$, and $t_{m,n}$ is the $(m+1, n+1)$th entry of $\mathbf{T}$.

A useful figure-of-merit is the squared magnitude of the difference between the transfer function of the DCT ($H_m(\omega; \mathbf{C})$) and of each considered approximation ($H_m(\omega; \mathbf{T})$). This measure is clearly energy-related and has the following mathematical expression:

$$D_m(\omega; \mathbf{T}) \triangleq |H_m(\omega; \mathbf{C}) - H_m(\omega; \mathbf{T})|^2, \quad m = 0, 1, \ldots, 7$$

where $\mathbf{T}$ is one of the selected approximate transforms. For $m = 0, 4$, the resulting transfer functions were the same for DCT, SDCT, BAS-2008, BAS-2011, and proposed approximation. Thus, we restrict our comparisons to $m = 1, 2, 3, 5, 6, 7$.

Fig. 2 shows strong similarities between the spectral characteristics of the DCT and the proposed approximation. The best results are presented when $m = 1, 3, 5, 7$. These spectrum similarity demonstrate the good energy related properties of the proposed algorithm. Table II summarized the total error energy departing from the actual DCT for each matrix row. This quantity is given by

$$\epsilon_m(\mathbf{T}) = \int_0^\pi D_m(\omega; \mathbf{T}) \mathrm{d}\omega$$

and was numerically evaluated. On account of these results, we remove the BAS-2011 method from our subsequent analysis because of its higher energy error.

III. APPLICATION TO IMAGE COMPRESSION AND DISCUSSION

The proposed approximation was assessed according to the methodology described in [11] and supported by [13]. A set of 45 512×512 8-bit greyscale images obtained from a standard

TABLE II
ERROR ENERGY $\epsilon_m(\mathbf{T})$ FOR SEVERAL APPROXIMATE TRANSFORMS

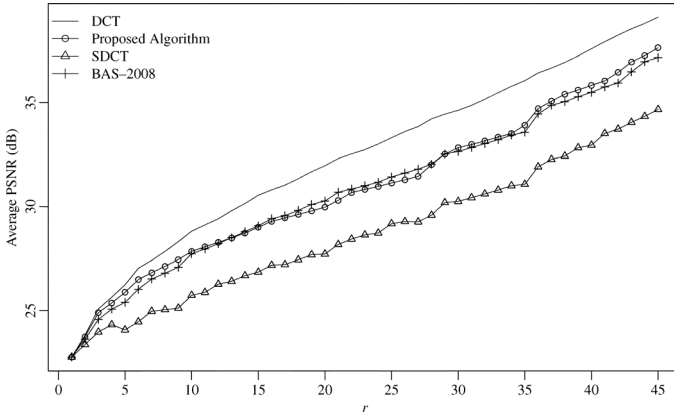| m	BAS-2011	BAS-2008	SDCT	Proposed
1	0.59	0.59	0.53	0.19
2	0.02	0.02	0.44	0.47
3	10.10	1.38	0.58	0.21
5	2.35	1.30	0.57	0.16
6	5.98	0.01	0.40	0.36
7	6.23	1.90	0.38	0.18
Total	25.27	5.20	2.90	1.57

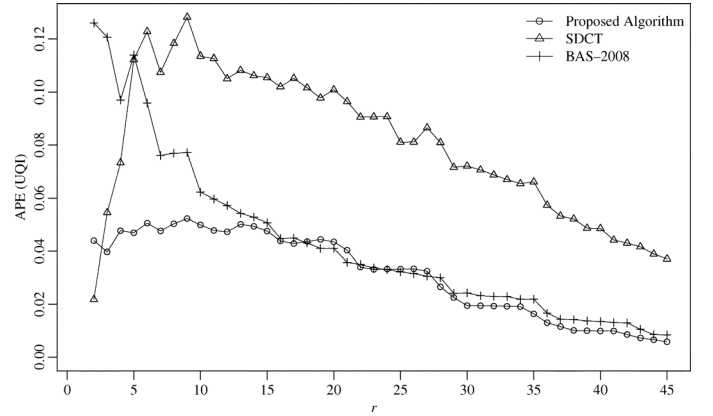Fig. 3. Average PSNR for several compression ratios.

Fig. 4. Average UQI absolute percentage error relative to the DCT for several compression ratios.

Fig. 5. Average MSE absolute percentage error relative to the DCT for several compression ratios.

public image bank [23] was considered. In this set, the images employed in [13] were included.

Each image was divided into 8×8 subblocks, which were submitted to the two-dimensional (2-D) approximate transforms implied by the proposed matrix $\hat{\mathbf{C}}_{\mathrm{orth}}$. An 8×8 image block $\mathbf{A}$ has its 2-D transform mathematically expressed by [24]: $\mathbf{T} \cdot \mathbf{A} \cdot \mathbf{T}^{\top}$, where $\mathbf{T}$ is a considered transformation. This computation furnished 64 approximate transform domain coefficients for each sub-block. According to the standard zigzag sequence [25], only the r initial coefficients were retained, with the remaining ones set to zero. We adopted $1 \leq r \leq 45$. The inverse procedure was then applied to reconstruct the processed data and image degradation is assessed.

As suggested in [13], the peak signal-to-noise ratio (PSNR) was utilized as figure-of-merit. However, in contrast, we are considering the average PSNR from all images instead of the PSNR results obtained from particular images. Average calculations may furnish more robust results, since the considered metric variance is expected to decrease as more images are analyzed [26]. Fig. 3 shows that the proposed approximation $\hat{\mathbf{C}}_{\mathrm{orth}}$ could indeed outperform the SDCT at any compression ratio. Moreover, it could also outperform the BAS-2008 algorithm [13] for high- and low-compression ratios. In the mid-range compression ratios, the performance was comparable. This result could be achieved at the expense of only two additional arithmetic operations.

Additionally, we also considered the universal quality index (UQI) and mean square error (MSE) as assessments tools. The UQI is understood as a better method for image quality assessment [27] and the MSE is an error metrics commonly employed when comparing image compression techniques.

Figs. 4 and 5 depict the absolute percentage error (APE) relative to the exact DCT performance for the average UQI and
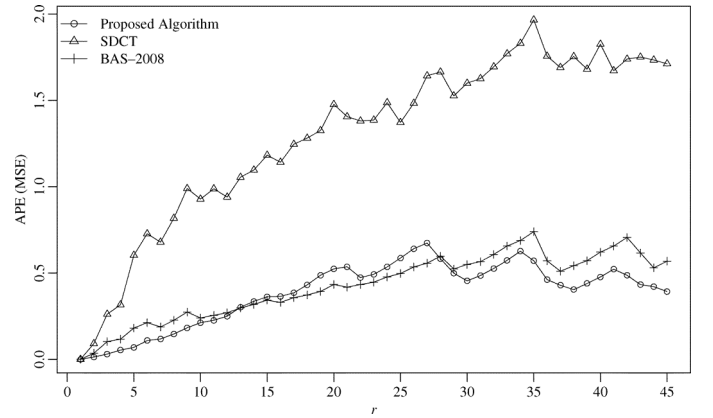
average MSE, respectively. According to these metrics, the proposed approximation lead to a better performance at almost all compression ratios. In particular, for high- and low-compression ratio applications the proposed approximation is clearly superior.

These results indicate that the proposed approximation is adequate for image compression, specifically for high-compression ratio applications. This scenario is found in low bit rate transmissions [13]. Additionally, some applications operate with large amount of data—which demand fast, low-complexity algorithms — at high compression ratios. For instance, models for face recognition and detection prescribe $r = 15$ or less [28], [29]. This compression ratio is in one of the ranges where our proposed technique excels. Additionally, popular JPEG compression ratio are higher than 75%, which corresponds to $r \leq 16$; again a suitable requirement for the proposed approximation.

Considering the above described compression methods, we also provide a qualitative assessment of the new approximation. Using the 8×8 block size only 5 out of the 64 coefficients in each 8×8 block were retained. Thus, after compression, we derived the reconstructed images according to the DCT, the SDCT, the BAS-2008 algorithm, and the proposed algorithm for three different standard images (Lena, Airplane (F-16), boat.512) obtained from [23]. Fig. 6 shows the resulting images. The resulting reconstructed images using the proposed method are
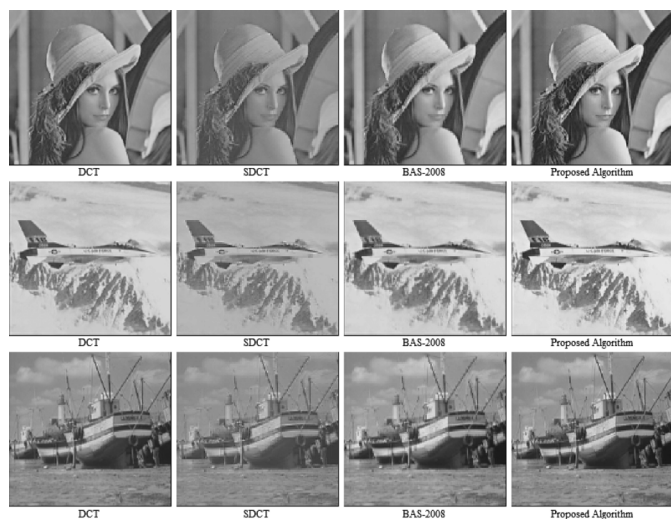
Fig. 6.   Image compression using the DCT, the SDCT, the BAS-2008 algorithm, and the proposed approximation for Lena, Airplane (F-16), and boat.512 images.

close to those obtained via the DCT. The superiority of the proposed algorithm over SDCT is clear. As expected, the reconstructed images have better quality and less blocking artifacts when compared to the BAS-2008 algorithm.

IV. CONCLUSION

This correspondence introduced an approximation algorithm for the DCT computation based on matrix polar decomposition. The proposed method could outperform the BAS-2008 method [13] in high- and low-compression ratios scenarios, according to PSNR, UQI, and MSE measurements. Moreover, the proposed method possesses constructive formulation based on the round-off function. Therefore, generalizations are more readily possible. For example, usual floor and ceiling functions can be considered instead of the round-off function. This would furnish entirely new approximations.

Additionally, the new approximate transform matrix has rows constructed from a different mathematical structure when compared to the BAS series of approximations, for instance. These rows can be considered in the design of hybrid algorithms which may take advantage of the best matrix rows from the existing algorithms aiming at novel improved approximate transforms. Finally, the new approximation offers the users another option for mathematical analysis and circuit implementation.

REFERENCES

[1] N. Roma and L. Sousa, "Efficient hybrid DCT-domain algorithm for video spatial downscaling," *EURASIP J. Adv. Signal Process.*, vol. 2007, no. 2, pp. 30–30, 2007.

[2] V. Britanak, P. Yip, and K. R. Rao, *Discrete Cosine and Sine Transforms*.   New York: Academic, 2007.

[3] M. Effros, H. Feng, and K. Zeger, "Suboptimality of the Karhunen-Loève transform for transform coding," *IEEE Trans. Inf. Theory*, vol. 50, pp. 1605–1619, Aug. 2004.

[4] N. Suehiro and M. Hateri, "Fast algorithms for the DFT and other sinusoidal transforms," *IEEE Trans. Acoust. Speech, Signal Process.*, vol. ASSP-34, no. 6, pp. 642–644, Jun. 1986.

[5] H. S. Hou, "A fast recursive algorithm for computing the discrete cosine transform," *IEEE Trans. Acoust. Speech, Signal Process.*, vol. ASSP-35, no. 10, pp. 1455–1461, Oct. 1987.

[6] Y. Arai, T. Agui, and M. Nakajima, "A fast DCT-SQ scheme for images," *Trans. IEICE*, vol. E-71, pp. 1095–1097, Nov. 1988.

[7] C. Loeffler, A. Ligtenberg, and G. Moschytz, "Practical fast 1D DCT algorithms with 11 multiplications," in *Proc. Int. Conf. Acoustics, Speech, and Signal Processing*, 1991, pp. 988–991.

[8] V. S. Dimitrov, K. Wahid, and G. A. Jullien, "Multiplication-free 8×8 DCT architecture using algebraic integer encoding," *Electron. Lett.*, vol. 40, no. 20, pp. 1310–1311, 2004.

[9] H. L. P. Arjuna Madanayake, R. J. Cintra, D. Onen, V. S. Dimitrov, and L. T. Bruton, "Algebraic integer based 8×8 2-d dct architecture for digital video processing," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, May 2011, pp. 1247–1250.

[10] R. J. Cintra and V. S. Dimitrov, "The arithmetic cosine transform: Exact and approximate algorithms," *IEEE Trans. Signal Process.*, vol. 58, pp. 3076–3085, Jun. 2010.

[11] T. I. Haweel, "A new square wave transform based on the DCT," *Signal Process.*, vol. 82, pp. 2309–2319, 2001.

[12] S. Bouguezel, M. O. Ahmad, and M. N. S. Swamy, "A multiplication-free transform for image compression," in *2nd Int. Conf. Signals, Circuits and Systems*, Nov. 2008, pp. 1–4.

[13] S. Bouguezel, M. O. Ahmad, and M. N. S. Swamy, "Low-complexity 8×8 transform for image compression," *Electron. Lett.*, vol. 44, pp. 1249–1250, Sep. 2008.

[14] S. Bouguezel, M. O. Ahmad, and M. N. S. Swamy, "A fast 8×8 transform for image compression," in *2009 Int. Conf. Microelectronics (ICM)*, Dec. 2009, pp. 74–77.

[15] S. Bouguezel, M. O. Ahmad, and M. N. S. Swamy, "A novel transform for image compression," in *53rd IEEE Int. Midwest Symp. Circuits and Systems (MWSCAS)*, Aug. 2010, pp. 509–512.

[16] S. Bouguezel, M. O. Ahmad, and M. N. S. Swamy, "A low-complexity parametric transform for image compression," in *Proc. 2011 IEEE Int. Symp. Circuits and Systems*, 2011.

[17] K. Lengwehasatit and A. Ortega, "Scalable variable complexity approximate forward DCT," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, pp. 1236–1248, Nov. 2004.

[18] R. J. Cintra, "An integer approximation method for discrete sinusoidal transforms," *Circuits, Systems, Signal Process.*, 2011, to be published.

[19] N. J. Higham, "Computing the polar decomposition—with applications," *SIAM J. Sci. Statist. Comput.*, vol. 7, pp. 1160–1174, Oct. 1986.

[20] MathWorks, MATLAB.   Natick, MA, The MathWorks, Inc., 2011.

[21] F. M. Bayer and R. J. Cintra, "Image compression via a fast DCT approximation," *IEEE Latin America Trans.*, vol. 8, pp. 708–713, Dec. 2010.

[22] N. J. Higham, *Functions of Matrices: Theory and Computation*.   Philadelphia, PA: SIAM, 2008.

[23] The USC-SIPI Image Database Univ. Southern California, Signal and Image Processing Inst., 2011 [Online]. Available: http://sipi.usc.edu/database/

[24] T. Suzuki and M. Ikehara, "Integer DCT based on direct-lifting of DCT-IDCT for lossless-to-lossy image coding," *IEEE Trans. Image Process.*, vol. 19, pp. 2958–2965, Nov. 2010.

[25] I.-M. Pao and M.-T. Sun, "Approximation of calculations for forward discrete cosine transform," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, pp. 264–268, Jun. 1998.

[26] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*.   Upper Saddle River, NJ: Prentice-Hall, 1993, vol. 1, Prentice-Hall Signal Processing Series.

[27] Z. Wang and A. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, 2002.

[28] C. Sanderson and B. C. Lovell, "Multi-region probabilistic histograms for robust and scalable identity inference," *Lecture Notes in Computer Science*, vol. 5558, pp. 199–208, 2009.

[29] S. Eickeler, S. Müller, and G. Rigoll, "Recognition of JPEG compressed face images based on statistical methods," *Image Vis. Comput.*, vol. 18, pp. 279–287, 2000.