Cintra and Bayer *EURASIP Journal on Advances in Signal Processing* (2017) 2017:50
DOI 10.1186/s13634-017-0486-8

EURASIP Journal on Advances in Signal Processing

RESEARCH Open Access

Comments on 'Area and power efficient DCT architecture for image compression' by Dhandapani and Ramachandran

Renato J. Cintra[1*] and Fábio M. Bayer[2]

Abstract

In [Dhandapani and Ramachandran, "Area and power efficient DCT architecture for image compression", EURASIP Journal on Advances in Signal Processing 2014, 2014:180] the authors claim to have introduced an approximation for the discrete cosine transform capable of outperforming several well-known approximations in literature in terms of additive complexity. We could not verify the above results and we offer corrections for their work.

Keywords: DCT approximations, Low-complexity transforms

1 Introduction

In a recent paper [1], a low-complexity transformation was introduced, which is claimed to be a good approximation to the discrete cosine transform (DCT). We wish to evaluate this claim.

The introduced transformation is given by the following matrix:

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

We aim at analyzing the above matrix and showing that it does not consist of a meaningful approximation for the 8-point DCT. In the following, we adopted the same methodology described in [2–11] which the authors also claim to employ.

2 Criticisms

2.1 Inverse transformation

The authors of [1] claim that inverse transformation $\mathbf{T}^{-1}$ is given by

$$\frac{1}{2} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & -1 & 1 \\ 0 & 0 & -1 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ -1 & 1 & 0 & 0 & 1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \end{bmatrix}.$$

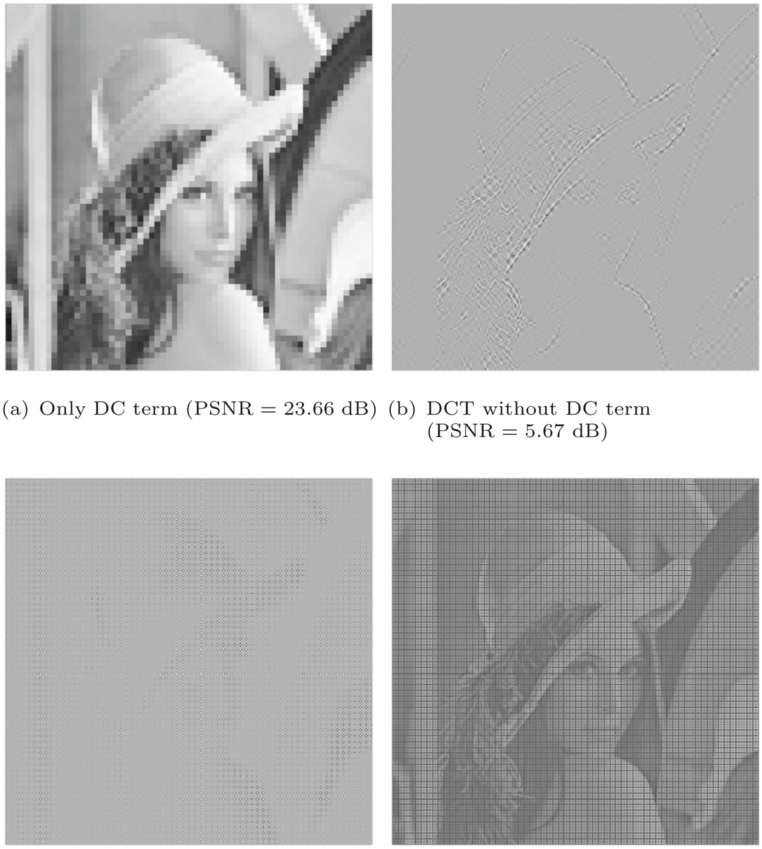
However, simple computation reveal that this is not accurate, being the actual inverse given by:

$$\mathbf{T}^{-1} = \frac{1}{2} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}.$$

2.2 Lack of DC component

The first point to be noticed is that the matrix $\mathbf{T}$ lacks a row of constant entries. Therefore, it is not capable of

*Correspondence: rjdsc@stat.ufpe.org
[1]Departmento de Estatística, Universidade Federal de Pernambuco, Recife, Brazil
Full list of author information is available at the end of the article



(a) Only DC term (PSNR = 23.66 dB) (b) DCT without DC term (PSNR = 5.67 dB)

(c) Only first row of **T** [1] (PSNR = 4.70 dB) (d) **T** without the first row (PSNR = 4.74 dB)

Fig. 1 Reconstructed Lena image based **a** only on the DC component of the DCT, **b** on all DCT transform coefficient except the DC component, **c** only on the first row of the **T** [1], and **d** on all **T** coefficients, except the first row

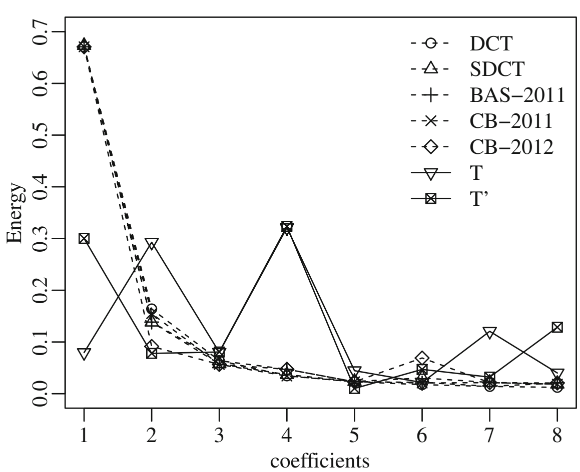


Fig. 2 Energy distribution of transform coefficients

Table 1 Transform coding assessment

Method	Transform efficiency	Coding gain (dB)
DCT [12]	93.99	8.83
SDCT [2]	82.62	6.02
BAS-2008 [4]	84.95	6.01
BAS-2009 [6]	85.38	7.91
BAS-2010 [8]	88.22	8.33
BAS-2011 [9]	85.38	7.91
CB-2011 [13]	87.43	8.18
CB-2012 [3]	80.90	7.33
Transformation in [14]	**34.93**	**−1.65**
Transformation in [15]	**33.67**	**−4.08**
T [1]	**28.95**	**−1.86**
$\mathbf{T}_{\mathrm{SFG}}$ [1]	**28.95**	**−1.86**

Table 2 PSNR of reconstructed images ($r = 6w0$)

Transform	Lena	Boat	Goldhill	Barbara	Lighthouse	Mandrill	Grass
DCT [12]	51.400	46.531	49.497	47.097	49.719	41.147	44.264
SDCT [2]	45.708	41.593	44.308	40.532	43.044	35.956	36.517
BAS-2008 [4]	43.996	39.498	42.449	38.304	41.139	33.886	34.364
BAS-2009 [6]	48.096	44.828	46.470	40.143	44.035	37.982	36.869
BAS-2010 [8]	50.976	46.483	48.912	46.657	48.193	40.617	42.486
BAS-2011 [9]	48.010	44.874	46.328	40.073	44.690	38.085	37.191
CB-2011 [13]	49.537	45.353	47.892	43.163	46.455	39.668	39.815
CB-2012 [3]	46.621	44.217	45.027	39.763	41.939	36.486	35.223
Transformation in [14]	30.193	29.635	32.107	29.411	29.777	26.575	20.612
Transformation in [15]	27.895	27.463	29.797	27.260	27.547	24.530	18.445
T [1]	30.560	30.034	32.565	29.851	30.090	26.862	20.982
$\mathbf{T}_{\mathrm{SFG}}$ [1]	30.889	29.867	32.920	29.117	30.189	26.779	20.900

computing the mean value or the DC component of a signal under analysis. In terms of image compression, the DC value is the single most important coefficient concentrating most of the image energy. To illustrate this fact, Fig. 1 shows the reconstructed standard Lena image by means of (i) the DC component of the standard DCT, (ii) all DCT coefficients, except the DC component, (iii) the first row of matrix **T** [1], and (iv) all **T** coefficients, except the first row, respectively. In [12], Britanak meticulously cataloged dozens of DCT approximations; all of them computed the DC component. The lack of the DC component computation suggests that compressed images resulting from the application of **T** are expected to be severely degraded in terms of perceived image quality. The associated PSNR values in Fig. 1 also show the poor quality of the reconstructed images using **T**. Considering $M \times N$ pixel images, the PSNR measure is calculated by:

$$\mathrm{PSNR} = 10 \cdot \log_{10}\left(\frac{\mathrm{MAX}^2}{\mathrm{MSE}}\right),$$

where $\mathrm{MSE} = \frac{1}{M \cdot N}\sum_{i=1}^{M}\sum_{j=1}^{N}(a_{i,j} - b_{i,j})^2$, $a_{i,j}$ and $b_{i,j}$ are the (i,j)-th element of the original and reconstructed images, respectively; and MAX is the maximum pixel valye. For 8-bit greyscale images, $\mathrm{MAX} = 255$.

2.3 Fast algorithm

In the 'Fig. 1' of [1], the authors display a signal flow graph (SFG) which does not correspond to the computation implied by their proposed matrix. Their proposed SFG consists of two addition butterfly sections and one final permutation, which correspond to the following matrices, respectively:

$$\mathbf{A}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

Table 3 PSNR of reconstructed images ($r = 10$)

Transform	Lena	Boat	Goldhill	Barbara	Lighthouse	Mandrill	Grass
DCT [12]	32.088	28.971	30.656	24.752	25.549	22.832	19.893
SDCT [2]	27.443	25.570	27.543	23.488	23.348	21.095	17.019
BAS-2008 [4]	29.509	27.150	28.994	24.285	24.444	22.279	18.849
BAS-2009 [6]	29.916	27.354	29.288	24.520	24.381	22.223	18.661
BAS-2010 [8]	31.143	28.292	30.072	24.666	25.063	22.581	19.376
BAS-2011 [9]	29.916	27.354	29.288	24.520	24.381	22.223	18.661
CB-2011 [13]	30.446	27.861	29.612	24.460	24.756	22.516	19.157
CB-2012 [3]	27.015	25.190	27.141	23.595	23.087	21.596	17.170
Transformation in [14]	**2.159**	**1.856**	**2.877**	**2.936**	**2.582**	**1.992**	**1.981**
Transformation in [15]	**−6.927**	**−7.213**	**−6.205**	**−6.120**	**−6.442**	**−7.053**	**−6.914**
T [1]	**2.163**	**1.867**	**2.880**	**2.951**	**2.596**	**2.001**	**1.981**
$\mathbf{T}_{\mathrm{SFG}}$ [1]	**4.686**	**4.380**	**5.399**	**5.454**	**5.086**	**4.481**	**4.355**

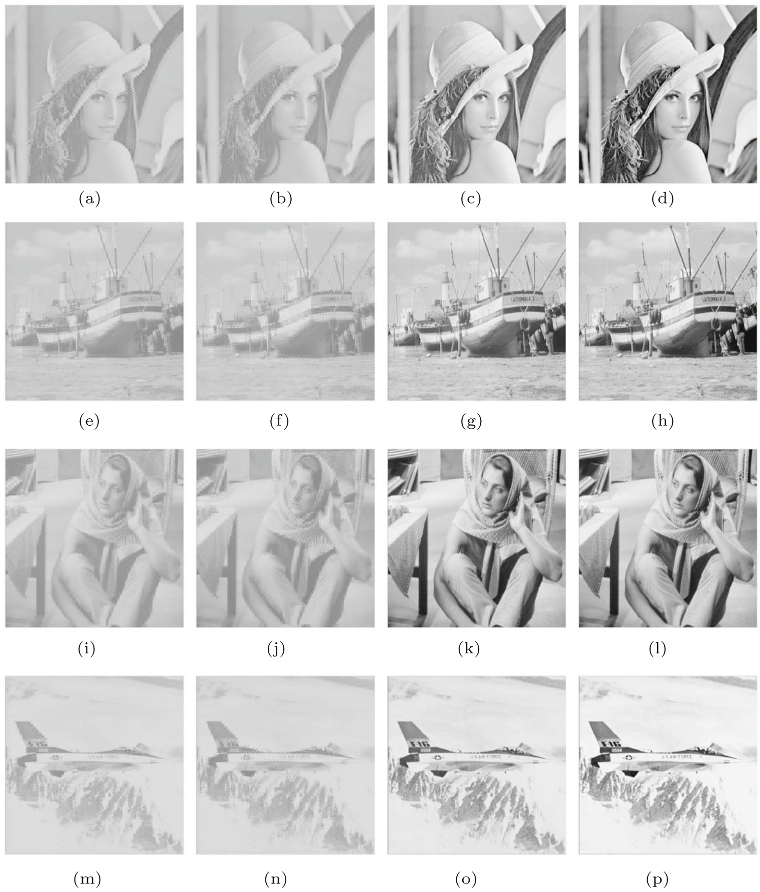


Fig. 3 Reconstructed images for $r = 45$. **a** $\mathbf{T}$ (PSNR=23.743), **b** $\mathbf{T}_{\mathrm{SFG}}$ (PSNR=24.244), **c** CB-2012 (PSNR=36.977), **d** BAS-2011 (PSNR=39.835), **e** $\mathbf{T}$ (PSNR=23.350), **f** $\mathbf{T}_{\mathrm{SFG}}$ (PSNR=22.926), **g** CB-2012 (PSNR=34.782), **h** BAS-2011 (PSNR=37.152), **i** $\mathbf{T}$ (PSNR=23.512), **j** $\mathbf{T}_{\mathrm{SFG}}$ (PSNR=23.653), **k** CB-2012 (PSNR=29.913), **l** BAS-2011 (PSNR=31.298), **m** $\mathbf{T}$ (PSNR=22.282), **n** $\mathbf{T}_{\mathrm{SFG}}$ (PSNR=22.184), **o** CB-2012 (PSNR=36.785) and **p** BAS-2011 (PSNR=40.433)

$$\mathbf{A}_2 = \begin{bmatrix} 1&1&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0 \\ 0&0&1&1&0&0&0&0 \\ 0&0&0&1&0&0&0&0 \\ 0&0&0&0&1&1&0&0 \\ 0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&1&1 \end{bmatrix}, \mathbf{P} = \begin{bmatrix} 0&0&0&0&0&0&0&1 \\ 0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&1&0&0 \\ 0&0&0&0&1&0&0&0 \\ 0&0&0&1&0&0&0&0 \\ 0&0&1&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0 \\ 1&0&0&0&0&0&0&0 \end{bmatrix}.$$

However, this fast algorithm induces to the following matrix:

$$\mathbf{P} \cdot \mathbf{A}_2 \cdot \mathbf{A}_1 = \begin{bmatrix} 1&1&0&0&0&0&1&1 \\ 0&1&0&0&0&0&1&0 \\ 0&0&1&0&0&1&0&0 \\ 0&0&1&1&1&1&0&0 \\ 0&0&0&1&-1&0&0&0 \\ 0&0&1&1&-1&-1&0&0 \\ 0&1&0&0&0&0&-1&0 \\ 1&1&0&0&0&0&-1&-1 \end{bmatrix} = \mathbf{T}_{\mathrm{SFG}},$$

which is different from $\mathbf{T}$. Therefore, the SFG is incorrect and does not correspond to the proposed matrix. We assume that the intended method is $\mathbf{T}_{\mathrm{SFG}}$, which is the matrix implied by the fast algorithm. Indeed, this transformation is shown again in the schematics of the hardware realization of their work. Nevertheless, hereafter, we analyze both matrices: $\mathbf{T}$ and $\mathbf{T}_{\mathrm{SFG}}$. Similar to $\mathbf{T}$, the matrix $\mathbf{T}_{\mathrm{SFG}}$ does not evaluate the DC value, being subject to the criticism detailed in the previous subsection.

2.4 Lack of energy concentration

Contrary to the expected behavior for a data compression transformation, the matrix $\mathbf{T}$ does not exhibit good decorrelation and energy concentration properties. Energy concentration can be quantified by submitting data to a considered transformation and then computing the energy distribution along transform-domain coefficients. Thus, we considered a Monte Carlo simulation with 10,000 8-point input vectors modeled after the first-order Markov process with correlation coefficient of 0.95 [12]. For comparison, we considered the following transformations: the DCT [12], the SDCT [2] the BAS-2013 [9], and the BC-2012 [3]. Obtained mean values are displayed in Fig. 2. In clear contrast with the other methods, transformations $\mathbf{T}$ and $\mathbf{T}_{\text{SFG}}$ perform very poorly.

Moreover, the lack of energy concentration in the first transform coefficients indicates that the standard zigzag pattern employed in the quantization step is not adequate for this transformation. Nevertheless, the authors claim to employ the zigzag pattern with success. We could not verify this claim.

To further assess the claim of good coding capabilities, we considered the unified coding gain and the transform efficiency as measures to quantify the coding performance [12] of $\mathbf{T}$ in comparison with bona fide transforms, such as: DCT, SDCT [2], BAS-2008 [4], BAS-2009 [6], BAS-2010 [8], BAS-2011 [9], CB-2011 [13], and CB-2012 [3]. In addition, we also considered the transformation in [14] and transformation in [15]. Results are shown in Table 1. We emphasize in bold the unfavorable measurements associated to the transformations proposed by the authors. Such transformations are not expected to be suitable for image compression, since both coding measures resulted in very low values.

2.5 Irreproducibility of results

The results shown by Dhandapani and Ramachandran could not be repeated. The authors state that they employ simultaneously a quantization step, which corresponds to variable bitrate encoding, and a fixed number of retained transform-domain coefficients, which suggests constant bitrate. This seems contradictory. However, to examine the transformation suggested by the authors, we adopted

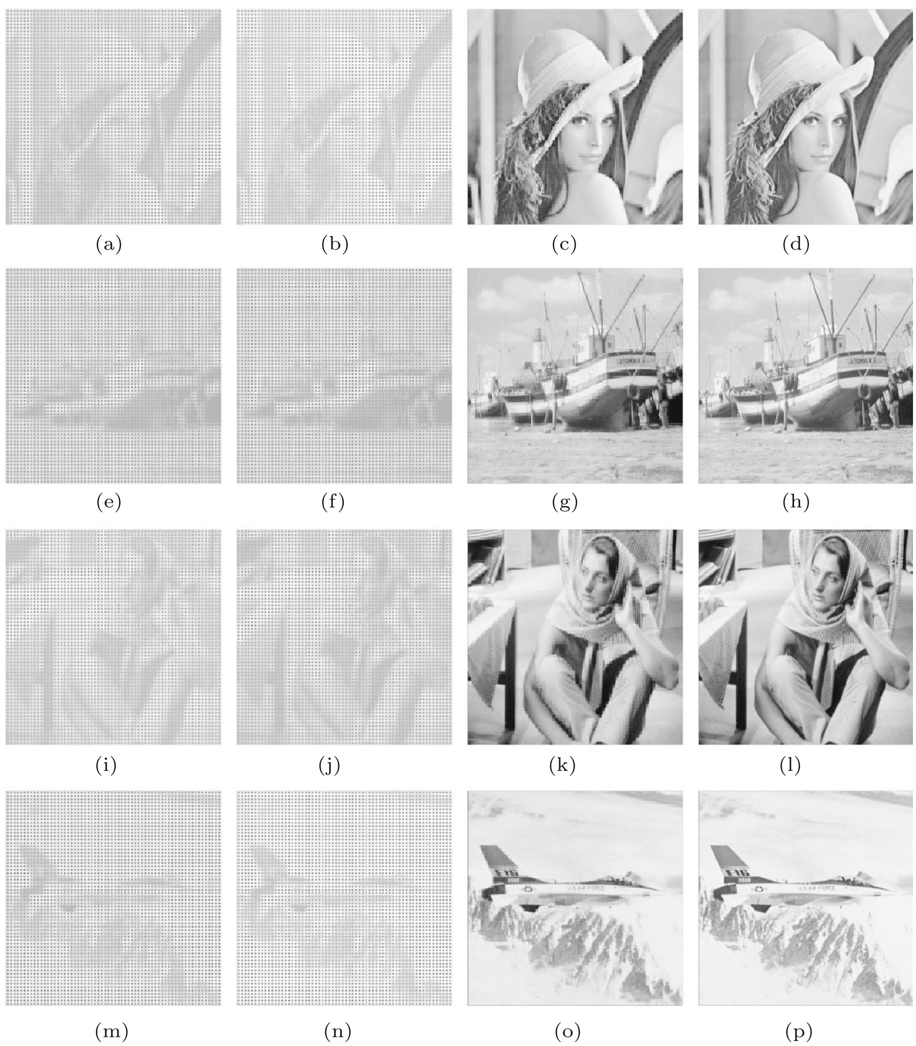


Fig. 4 Reconstructed images for $r = 15$. **a** $\mathbf{T}$ (PSNR=8.572), **b** $\mathbf{T}_{\text{SFG}}$ (PSNR=8.578), **c** CB-2012 (PSNR=27.500), **d** BAS-2011 (PSNR=31.271), **e** $\mathbf{T}$ (PSNR=8.231), **f** $\mathbf{T}_{\text{SFG}}$ (PSNR=8.239), **g** CB-2012 (PSNR=25.777), **h** BAS-2011 (PSNR=28.602), **i** $\mathbf{T}$ (PSNR=9.274), **j** $\mathbf{T}_{\text{SFG}}$ (PSNR=9.281), **k** CB-2012 (PSNR=24.116), **l** BAS-2011 (PSNR=25.275), **m** $\mathbf{T}$ (PSNR=5.711), **n** $\mathbf{T}_{\text{SFG}}$ (PSNR=5.715), **o** CB-2012 (PSNR=26.201) and **p** BAS-2011 (PSNR=29.152)

a constant bitrate encoding based on the retention of r transform-domain coefficients, as suggested originally by Haweel and others [2–11].

Although the authors do not explicitly inform the number of retained coefficients (r) in their computations. Only for high values of r we could obtain similar values. We calculated the PSNR values considering $r = 60$. Notice that for such a high value of r data is practically not compressed. This is because only 4 coefficients are discarded, implying a compression rate of only 6.25%. Table 2 shows the results. Additionally, at low compression, most orthogonal transforms tend to behave similarly. However, even under this scenario, the transformation proposed by the authors performed poorly—roughly 10 dB lower PSNR measurements. Indeed, the pivotal character of a good transform is its behavior in a wide range of compression rates, specially at high compression. For instance, considering the more realistic case of $r = 10$, as suggested in [2], we obtain the PSNR values shown in Table 3. Notice that the transformation proposed by the authors exhibits extremely high errors, which are emphasized in bold. We also report that the results linked to the transformations described in [14] and [15] display also acutely poor results as shown in Table 3.

In ‘Fig. 3’ of their work, the authors show reconstructed compressed images according to the following transformations: $\mathbf{T}_{\text{SFG}}$, CB-2012, and BAS-2011. All images showed high PSNR values with $\mathbf{T}_{\text{SFG}}$ offering PSNR values greater than 41 dB. We could no reproduce these results. The authors does not detail the employed parameters, in particular the value of r. However, for $r = 45$, we could obtain comparable PSNR measurements in terms of the traditional DCT approximations. Considering $\mathbf{T}$ or $\mathbf{T}_{\text{SFG}}$ the image degradation is very high, as shown in Fig. 3. For $r = 15$, a more realistic value, we obtain the images shown in Fig. 4. Images associated to $\mathbf{T}$ or $\mathbf{T}_{\text{SFG}}$ are severely degraded—roughly 25–30 dB lower than the typical values offered by traditional approximations. These results are evidence that the transformation proposed by the authors is not suitable for image compression.

Authors also show in ‘Fig. 4’ of their paper a curve relating PSNR measurements of compressed images to the parameter r. We could not reproduce their results. Figure 5 shows the curves that we obtained considering the same images as the authors. Our results are compatible to the computations independently found in [2, 4–11]. The curves associated to $\mathbf{T}$ and $\mathbf{T}_{\text{SFG}}$ indicate

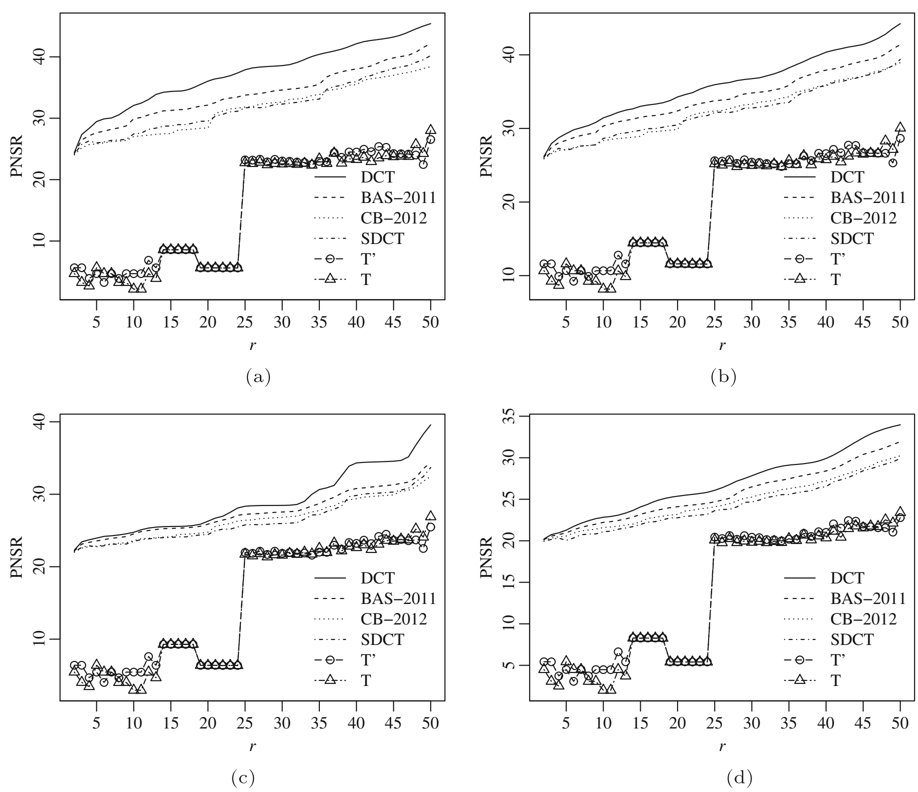


Fig. 5 PSNR measurements in terms of the number of retained coefficients r for selected standard images. **a** Lena, **b** Cameraman, **c** Barbara and **d** Mandrill

a significantly lower performance. For $r < 25$—a more realistic scenario—the PSNR loss compared to the traditional transformations is roughly 20 dB. Such evidence points towards the ineffectiveness of $\mathbf{T}$ and $\mathbf{T}_{\text{SFG}}$ for image compression.

3 Conclusion

The transformation proposed in [1] performs poorly when compared to archived DCT approximations. The results in [1] could not be reproduced and some corrections are supplied for the benefit of community.

Acknowledgements
This work was partially supported by CNPq, FACEPE, and FAPERGS.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]Departmento de Estatística, Universidade Federal de Pernambuco, Recife, Brazil. [2]Departmento de Estatística and LACESM, Universidade Federal de Santa Maria, Santa Maria, Brazil.

Received: 24 November 2015 Accepted: 18 June 2017
Published online: 10 July 2017

References

1. V Dhandapani, S Ramachandran, Area and power efficient DCT architecture for image compression. EURASIP J. Adv. Signal Process. **180**, 1–9 (2014)
2. TI Haweel, A new square wave transform based on the DCT. Signal Process. **82**, 2309–2319 (2001)
3. FM Bayer, RJ Cintra, DCT-like transform for image compression requires 14 additions only. Electron. Lett. **48**(15), 919–921 (2012). doi:10.1049/el.2012.1148
4. S Bouguezel, MO Ahmad, MNS Swamy, Low-complexity 8×8 transform for image compression. Electron. Lett. **44**(21), 1249–1250 (2008). doi:10.1049/el:20082239
5. S Bouguezel, MO Ahmad, MNS Swamy, in *2nd International Conference on Signals, Circuits and Systems (SCS)*. A multiplication-free transform for image compression, (2008), pp. 1–4. doi:10.1109/ICSCS.2008.4746898
6. S Bouguezel, MO Ahmad, MNS Swamy, in *2009 International Conference on Microelectronics (ICM)*. A fast 8×8 transform for image compression, (2009), pp. 74–77. doi:10.1109/ICM.2009.5418584
7. S Bouguezel, MO Ahmad, MNS Swamy, in *Proceedings of 2010 IEEE International Symposium on Circuits and Systems (ISCAS)*. Image encryption using the reciprocal-orthogonal parametric transform, (2010), pp. 2542–2545. doi:10.1109/ISCAS.2010.5537110
8. S Bouguezel, MO Ahmad, MNS Swamy, in *53rd IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*. A novel transform for image compression, (2010), pp. 509–512. doi:10.1109/MWSCAS.2010.5548745
9. S Bouguezel, MO Ahmad, MNS Swamy, in *Proceedings of the 2011 IEEE International Symposium on Circuits and Systems*. A low-complexity parametric transform for image compression, (2011)
10. S Bouguezel, MO Ahmad, MNS Swamy, Binary discrete cosine and Hartley transforms. IEEE Trans. Circuits Syst. I: Regular Papers. **60**(4), 989–1002 (2013)
11. N Brahimi, S Bouguezel, in *7th International Workshop on Systems, Signal Processing and Their Applications (WOSSPA)*. An efficient fast integer DCT transform for images compression with 16 additions only, (2011), pp. 71–74. doi:10.1109/WOSSPA.2011.5931415
12. V Britanak, P Yip, KR Rao, *Discrete cosine and sine transforms*. (Academic Press, Oxford, 2007)
13. RJ Cintra, FM Bayer, A DCT approximation for image compression. IEEE Signal Process. Lett. **18**(10), 579–582 (2011)
14. D Vaithiyanathan, R Seshasayanan, in *Proceeding of the International Conference on Advanced Computing and Communication Systems (ICACCS)*. Low power DCT architecture for image compression, (Coimbatore, Tamil Nadu, India, 2013), pp. 1–6. doi:10.1109/ICACCS.2013.6938745
15. D Vaithiyanathan, R Seshasayanan, S Anith, K Kunaraj, in *Proceeding of the International Conference on Green Computing, Communication and Conservation of Energy (ICGCE 2013)*. A low-complexity DCT approximation for image compression with 14 additions, (Chennai, Tamil Nadu, India, 2013), pp. 303–307. doi:10.1109/ICGCE.2013.6823450