The DLMT. An alternative to the DCT.

Félix Moreno, *Member, IEEE* and David Aledo
Centro de Electrónica Industrial (CEI)-ETSI-Universidad Politécnica de Madrid (Spain)
Email: felix.moreno@upm.es

***Abstract*— In the last recent years, with the popularity of image compression techniques, many architectures have been proposed. Those have been generally based on the Forward and Inverse Discrete Cosine Transform (FDCT, IDCT). Alternatively, compression schemes based on discrete "wavelets" transform (DWT), used, both, in JPEG2000 coding standard and in the next H264-SVC (Scalable Video Coding), do not need to divide the image into non-overlapping blocks or macroblocks.**

This paper discusses the DLMT (*Discrete Lopez-Moreno Transform*). It proposes a new scheme intermediate between the DCT and the DWT (*Discrete Wavelet Transform*). The DLMT is computationally very similar to the DCT and uses quasi-sinusoidal functions, so the emergence of artifact blocks and their effects have a relative low importance. The use of quasi-sinusoidal functions has allowed achieving a multiresolution control quite close to that obtained by a DWT, but without increasing the computational complexity of the transformation. The DLMT can also be applied over a whole image, but this does not involve increasing computational complexity. Simulation results in MATLAB show that the proposed DLMT has significant performance benefits and improvements comparing with the DCT.

I. INTRODUCTION

The reasons for the DCT [1], [2] success are based on the following factors: a) its computational simplicity, b) it is easy to be implemented in hardware and c) the relative quality of its results. However, its main drawback is the yielding of so-called "blocking artifacts" [3], [4], [5]. This effect is due to the implementation of the DCT over data blocks instead of applying it to the whole image (in the case of image compression standards 8x8pels blocks or 16x16pels macroblocks); where the correlation among neighboring blocks or macroblocks is not eliminated. On the other hand, the application of quantization operation (forward and inverse) is a data loss coding technique. This effect is particularly significant when high compression ratios and low bit-rate are needed.

The DWT [3],[6] transforms a time-domain discrete signal into time-frequency domain frequency domains) and "blocking artifacts" do not arise, even though at high compression ratios.

Similarly the Fourier Transform decomposes a signal in a summation of sinusoids at different frequencies; the DWT represents a signal as a superposition of several "wavelets". These are small “waves” that have their power concentrated in certain time windows. Although there is large literature explaining the physical meaning of the DWT (continuous or discrete) based on "wavelets", it would be explained through the concept of scale-resolution which is very common in cartography. That is, the greater the scale of a map (time domain), the lower the resolution (frequency domain); ie, it can cover a large geographic area but with very low detail. However, when decreasing the scale on a map then it is possible to appreciate some details. In the case of time domain signal through a transformation based on "wavelets", decreasing the analysis time (scale) of the signal, then the resolution in the frequency domain signal increases. Conversely, the increasing of the analysis time (scale in the time domain) reduces the resolution in the frequency domain.

The major drawback of the DWT is its high computational cost, compared with the DCT [7], [8], [9], [10], [11]. On the other hand, systems based on DWT compression present their most effective performance when adapting the calculation algorithms to the type of signal (in the case of image compression, the type of the images and their spatial-temporal characteristics). Furthermore, taking into account the calculation algorithms which have been proposed [11], based on the use of multiple stages of signal filtering, some blurring effects and ringing at the edges of the image would arise.

The DLMT proposes a new scheme intermediate between the DCT and the DWT. Firstly, as discussed below, it is computationally very similar to the DCT and the quality of results obtained does not depend on the characteristics of the image being processed. Secondly, the DLMT uses quasi-sinusoidal functions, so the emergence of artifact blocks

978-1-61284-971-3/11/$26.00 ©2011 IEEE

Finally, it is necessary to mention that the use of quasi-sinusoidal functions has allowed to achieve a multiresolution control quite close to that obtained by a DWT, but without increasing the computational complexity of the transformation. In addition, as shown later, the DLMT can also be applied over a whole image, but this does not involve increasing computational complexity. In this case, however, the necessary calculations are increased; but, because of its simplicity, it would be easily implemented in hardware.

II. THE DLMT

The two dimensional DLMT-2D consists of a pair of mathematical transforms, forward and reverse. It is, also, possible to define a pair of one-dimension transforms (DLMT-1D). Given its simplicity, the latter is used to demonstrate the DLMT-2D key properties. Then, those properties are applied to the DLMT-2D which will be the aim of this study because of its application in image compression.

The DLMT-1D is defined through its general equations:

$$X(k) = \sum_{n=0}^{N-1} x(n).\alpha(n,k).\emptyset_k^*(n) \quad (1)$$

$$x(n) = \frac{1}{N}\sum_{k=0}^{N-1} X(k).\beta(n,k).\emptyset_k(n) \quad (2)$$

Where X(k) is the forward transform of a sequence x(n) and the functions $\emptyset_k(n)$ and $\emptyset_k^*(n)$ are called base sequences, orthogonal to each other. On the other hand, α(n,k) and β(n,k) are weighting functions of the base sequences.

Traditionally, complex exponentials have been used for the Fourier series development. In the same way, it is also possible to develop functions in sine or cosine series (base sequences, $\emptyset_k(n)$). This is the case of the DCT (Discrete Cosine Transform) and the DST (Discrete Sine Transform). According to [2] and [12] for a given function there are 8 ways to develop in the form of cosine series and 8 ways to develop in sine series. They constitute a family of 16 orthogonal transformations of real sequences, x(n). In this study a sine series development has been chosen. The main reason for that will be showed later when considering the study of the weighting functions α(n, k) and β(n, k).

Accepting an implicit or virtual periodicity of x(n) [12] and, also, an even symmetry (which is always possible by mean of an extension type 2 of the function x(n), according to [2]), the

in a 2N virtual period.

Thus, the pair of transforms (Forward and Inverse) DLMT-2D can be written as:

a) Forward transform, FDLMT-2D:

$$X(k,l) = \frac{1}{(2N)^2}.\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\alpha(i,k).\beta(j.l).x(i,j).\sin\frac{(2i+1)\pi k}{2N}.\sin\frac{(2j+1)\pi l}{2N} \quad (3)$$

b) Inverse transform, IDLMT-2D:

$$x(i,j) = \frac{1}{(2N)^2}.\sum_{k=0}^{N-1}\sum_{l=0}^{N-1}\alpha(i,k).\beta(j.l).X(k,l).\sin\frac{(2i+1)\pi k}{2N}.\sin\frac{(2j+1)\pi l}{2N} \quad (4)$$

Being the weighting functions α(i,k) and β(j,l):

$$\alpha(i,k) = \frac{2N}{(2i+1)\pi k} \qquad \beta(j,l) = \frac{2N}{(2j+1)\pi l} \quad (5)$$

The base sequences used, $\emptyset_k(n)$, are sine functions and therefore orthogonal:

$$\frac{1}{2N}\sum_{n=0}^{N-1}\emptyset_k(n).\emptyset_m^*(n) = \begin{cases}1, & m=k\\ 0, & m\neq k\end{cases}$$

On the other hand, the DLMT-1D is an unitary transformation (Parseval Theorem), so:

$$\sum_{n=0}^{N-1}(x(n))^2 = \frac{1}{2N}\sum_{k=0}^{N-1}(X(k))^2$$

And in the case of the DLMT-2D, it is also an unitary transformation (Parseval Theorem) because:

$$\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}(x(i,j))^2 = \frac{1}{2N}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1}(X(k,l))^2$$

The FDLMT-2D is quite similar to the FDCT-2D in the sense of the property of the low frequency energy compaction [12]. However, the FDLMT-2D is much more efficient, as will be discussed below. This is due to the weighting functions α(i,k) and β(j,l) definition. These weighting functions are multiplied by their corresponding sine functions (see general expressions of the FDLMT-2D and the IDLMT-2D); thus sinc functions are obtained. They have very interesting properties both in the time-domain and in the frequency-domain.

$$\alpha(i,k).\sin\frac{(2i+1)\pi k}{2N} = \begin{cases}1 & ,k=0\\ sinc\frac{(2i+1)k}{2N} & ,k\neq 0\end{cases}$$
$$\beta(j,l).\sin\frac{(2j+1)\pi l}{2N} = \begin{cases}1 & ,l=0\\ sinc\frac{(2j+1)l}{2N} & ,l\neq 0\end{cases} \quad (6)$$

Moreover, these functions will be zero (for k≠0) when:

$$\frac{(2i+1)\pi k}{2N} = m\pi$$

And,

$$\frac{(2j+1)\pi l}{2N} = m\pi$$

for the second dimension.

Taking one of these two expressions (because both are identical), it follows that:

$$k = \frac{2Nm}{(2i+1)}$$

For m=1, the first zero of the function will be produced, then it is easily deductible that zeros of the sinc function, and therefore zeros of the pair of transformed FDLMT-2D and IDLMT-2D, depend on the value of N. That is, the size of one of the two dimensions of the signal x(i,j), (8x8pels block, 16x16pels macroblock, etc). 2N is also the inherent or virtual periodicity of the signal to be processed, but also 4N, and 8N etc. This allows to control perfectly what it has been called scale (time domain) of the signal (just at the introduction of this work), and its resolution in the frequency-domain. Since taking periods multiples of the inherent or virtual periodicity 2N or taking different values for N, it is possible to control the first zero of the sinc function. In other words, when increasing the value of N, the scale (time-domain) of the signal is also increasing; while the resolution (frequency-domain) is decreasing. Conversely, when decreasing the value of N the scale in the time-domain is decreasing while the resolution in the frequency-domain is increasing. This behavior is quite similar and the results are also comparable to those obtained by the DWT as discussed below. Also, in terms of pixel decorrelation and energy compaction when using sinc functions, in the frequency domain an explicit linear filterbank is obtained, with a variable and controllable cut-off frequency. This cut-off frequency depends on the inherent periodicity 2N and on the DLMT indexes "i" and "j". In summary, the proposed DLMT is biortogonal and multiresolution.

III. FDLMT-2D AND IDLMT-2D PROCESSING ALGORITHMS

In what follows an algorithm for calculating the FDLMT-2D and the IDLMT-2D that facilitates its implementation in hardware and software is proposed. Its study and its multi-

The solution is simple and is presented in the case of the DLMT-2D by means of the property of separation of the two dimensions, because it is a biortogonal transformation. Therefore, the DLMT-2D is separable, being able to express directly the FDLMT-2D as:

$$X(k,l) = \sum_{i=0}^{N-1} \frac{1}{2N}.\alpha(i,k).sin\frac{(2i+1)\pi k}{2N}. \sum_{j=0}^{N-1} x(i,j).\frac{1}{2N}.\beta(j,l).sin\frac{(2j+1)\pi l}{2N} \quad (7)$$

From the above expression yields the coefficient matrix [C]:

$$[C] = \sum_{i=0}^{N-1} \frac{1}{2N}.\alpha(i,k).sin\frac{(2i+1)\pi k}{2N} \quad (8)$$

And the transposed matrix $[C^T]$:

$$[C^T] = \sum_{j=0}^{N-1} \frac{1}{2N}.\beta(j,l).sin\frac{(2j+1)\pi l}{2N} \quad (9)$$

Being X(k,l), the product of both matrices and x(i,j):

$$X(k,l) = [C][x(i,j)][C^T] \quad (10)$$

On the other hand, from the expressions for the FDLMT-2D and the IDLMT-2D the equation for calculating the inverse transformation can be derived:

$$x(i,j) = [C^T]^{-1}[X(k,l)][C]^{-1} \quad (11)$$

Those expressions are computationally very similar to those used traditionally (especially in hardware implementations) for the FDCT-2D and the IDCT-2D and much simpler than those used in the calculation of the DWT-2D.

IV. ILLUSTRATIVE EXAMPLES

In the following, results obtained using a training set of 256x256pels grey-scale still image are presented (lena, cameraman, goldhill and baboon). Traditionally, the performance is evaluated by the PSNR (Peak Signal to Noise Ratio) [7], [9], [10] and the CR (Compression Ratio):

$$\text{PSNR}_{\text{dB}} = 10\log_{10}\frac{\text{Max. Val.}^2}{\text{MSE}}$$

$$CR = \frac{(Nboi - Nbci)x100}{Nboi}$$

Where Max.Val. is the maximum value of the input data (in the case of a luminance matrix it would be 235), Nboi is the Number of bits of the original image, Nbci is the number of bits of

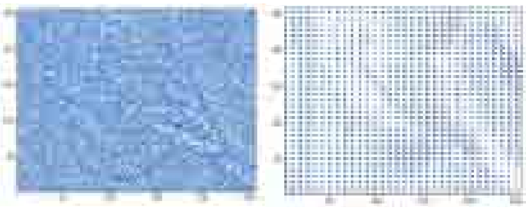
$$\mathrm{MSE} = \frac{1}{\mathrm{NxN}} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \left(x(i,j) - x'(i,j)\right)^2$$

Where x(i,j) is the input data matrix (8x8pels of luminance and 8bit per pel) and x'(i,j) is the processed data matrix. In the following examples Q means Quantized and NonQ means Non-Quantized.

A. Image: lena, 256x256pels. Block size: 8x8pels. Periodicity: 2N=16pels.

	DLMT Q	DLMT NonQ	DCT Q	DCT NonQ
CR	75%	70%	74%	72%
MSE	$1x10^3$	$1x10^3$	$1x10^3$	$1.9x10^3$
PSNR	17.9dB	17.8dB	17.8dB	15.3dB


Original Image DLMT processed



a) b)

Fig. 1 Energy compaction, Matlab contour view: a) FDCT 8x8pels regularity and b) FDLMT 8x8pels regularity

B. Image: lena, 256x256pels. DCT with different block sizes in pels.

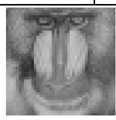
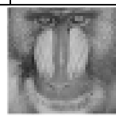
	16x16	32x32	64x64	128x128	256x256
CR	71%	63%	49%	20%	4%
MSE	$2.6x10^3$	$4.6x10^3$	$8.9x10^3$	$1.76x10^4$	$3.51x10^4$
PSNR	14dB	11.5dB	8.65dB	5.68dB	2.68dB

C. Image: lena, 256x256pels. DLMT with different block sizes in pels. Periodicity: 2N=16pels. (Q: Quantized. NonQ: Non-Quantized)

	16x16	32x32	64x64	128x128	256x256
CR(NonQ)	74%	75%	75%	75%	75%
CR(Q)	75%	75%	75%	75%	75%
MSE.Q/ NonQ	545/ 540	273/ 272.9	137/ 137	68.5/ 68.4	34.2/34.2
PSNR.Q/ NonQ(dB)	20.8/ 21	23.8/ 23.7	26.8/ 26.7	29.8/ 29.8	32.8/ 32.8

D. Image: baboon, 256x256pels. Block size: 8x8pels. Periodicity: 2N=16pels.

	DLMT Q	DLMT NonQ	DCT Q	DCT NonQ
CR	75%	71%	72%	34%
MSE	1.2x103	1.1x103	1.14x103	2.1x103
PSNR	17.5dB	17.7dB	17.56dB	15.9dB

Original Image DLMT processed

E. Image: cameraman, 256x256pels. Block size: 8x8pels. Periodicity: 2N=16pels.

	DLMT Q	DLMT NonQ	DCT Q	DCT NonQ
CR	74%	71%	73%	57%
MSE	1.1x103	1.1x103	1.1x103	2x103
PSNR	17.6dB	17.8dB	17.7dB	15.1dB


Original Image DLMT processed

F. Image: goldhill, 256x256pels. Block size: 8x8pels. Periodicity: 2N=16pels

	DLMT Q	DLMT NonQ	DCT Q	DCT NonQ
CR	75%	68%	74%	73%
MSE	930	895	919	1.63x103
PSNR	18.4dB	18.6dB	18.5Db	16dB


Original Image DLMT processed

V. PRELIMINARY FPGA IMPLEMENTATION RESULTS

The initial prototype DLMT-2D system implementation accomplished has been just a proof of concept, so no noticeable effort was made in optimizing hardware system implementation. The very first synthesis results obtained have shown that those FPGA with integrated hardware multiplier would be necessary in order to achieve real-time processing. The implemented algorithm (see section III) is doing 32 multiplications by means of a typical pipelining architecture, which has a measured initial latency of 64 clock cycles.

On the other hand, a 640x480pels (30fps) Toshiba TCM38230MD(A) video camera has been also used (not only still images, see section IV). So, processing one pel per clock cycle, each frame takes 24.576µsec. As a result of that, the maximum processing time is 29.696µsec which is less than the video frame rate 1/30fps=33msec.

The DLMT-2D algorithm implemented needs a little bit larger memory to store coefficients matrix and coefficients inverse matrix; however no quantization process is necessary at all, as in the case of the DCT-2D.

VI. CONCLUSION

This paper describes a new image coding scheme based on the DLMT. Just from the analysis of the previous results it can be derived that the FDLMT-2D has a better low frequency energy compaction than that obtained by the FDCT-2D (See Fig. 1 and the results of the section IV). On the other hand, these results have shown that, except in the first line and first column of the coefficient matrix we calculate for the FDLMT-2D, where a light effect of the sinc function can be observed, the rest of the coefficient matrix is much more uniform than the FDCT-2D. That makes the operation of quantization unnecessary (one of the sources of losses in image coding), which is typical in those coding systems based on the FDCT-2D. The effect of this can be observed in the first row and first column of the coefficient matrix we have mentioned before, which is still much smoother than the one in the DCT-2D. Nevertheless, in the case of the FDLMT-2D, the coefficient matrix can be divided by a constant (DLMT quantization), which is not strictly a quantization to improve the compression factor (although there is not a noticeable improvement-see experiments-) when applying RLE (Run-Length Encoding) and Huffman encoding. Besides, it would be possible to remove (or, at least, to simplify) the typical algorithm for reading the information (Zig-zag Scan), which would further improve the processing speed due to the simplification of the process.

The DLMT used here attempts to exploit the effect of the processed block size (NxNpels) and/or the inherent periodicity of the FDLMT (See section II). When increasing the value of N, the scale (time-domain) of the signal is also increasing while its resolution (frequency-domain) is decreasing (the cut-off frequency of [illegible]

images) the DLMT allows using smaller values of N (smaller scale, time-domain) and increasing resolution (frequency-domain) while maintaining good visual quality. At the same time it is even possible to increase the virtual periodicity (4N, 8N or 16N, etc) so; as a consequence of that, the CR approaches 75% and the PSNR exceeds 30dB, improving the energy compaction. Finally, the experiments A, B and C demonstrate that the DLMT is producing better results than the DCT when increasing block size and using a constant periodicity of 2N.

This method is well adapted to very low bit rate compression schemes. Furthermore, using a simple quantization operation provides good results, comparable (or even better) to the best results currently published concerning the DCT or the DCT-like schemes [6],[9],[10].

Further research should include some new comparative studies such as the DLMT against wavelet-based image coding. This encouraging coding scheme is very suitable for implementation in hardware without a noticeable increase of the computational load and further results will be reported.

AKNOWLEDGMENT

The author wishes to express his deep gratitude for advice he has received from Prof. J. Lopez of the Universidad de Las Palmas de Gran Canaria (Spain). He proposed the initial ideas and my appreciation for that can be envisaged in the DLMT acronyms. Special thanks go to Prof. Ramirez-Angulo of the New Mexico State University (NMSU-USA) for his proofreading efforts and for his support and advice during my sabbatical period at the NMSU in 2010.

REFERENCES

[1] K. R. Rao and P. Yip, Discrete Cosine Transform. Algorithms, Advantages, Applications, Academic Press Inc., 1990, ISBN 012-580-203X.

[2] Stephen A. Martucci, Symmetric Convolution and the Discrete Sine and Cosine Transform, IEEE Transactions on Signal Processing, Vol. 42, No. 5, pp. 1038-1051, May 1994.

[3] Marc Antonini, Michel Barlaud, Pierre Mathieu and Ingrid Daubechies, Image Coding Using Wavelet Transform, IEEE Transactions on Image Processing, Vol. 1, No. 2, pp. 205-220, April 1992.

[4] Jianmin Jiang and Guocan Feng, The

Transactions on Signal Processing, Vol. 50, No. 5, pp. 1160-1169, May 2002.
[5] A. N. Skodras, Direct Transform to Computation, IEEE Signal processing Letters, Vol. 6, pp. 202-204, Aug. 1999.
[6] J. R. Hernandez, M. Amado and F. P. Gonzalez, DCT-domain watermarking techniques for still images: detector performance analysis and a new structure, IEEE Transactions on Image Processing, Vol. 9, pp. 55-68, Jan. 2000.
[7] V. Srinivasa Rao, Dr. P. R. Kumar, G. V. H. Prasad, M. Prema Kumar and S. Ravichand, Discrete Cosine Transform Vs Discrete Wavelet Transform: An Objective Comparison of Image Compression techniques for JPEG Encoder, International Journal of Advanced Engineering and Applications, pp. 87-90, Jan. 2010.
[8] Zixiang Xiong, Kannan Ramchandran, Michael T. Orchard, and Ya-Qin Zhang, A Comparative Study of DCT- and Wavelet-Based Image Coding, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 9, No. 5, pp. 692-695, Aug. 1999.
[9] Ali Al-Haj, Combined DWT-DCT Digital Image Watermarking, Journal of Computer Science 3 (9), pp. 740-746, 2007. ISSN 1549-3636.
[10] Wen-Chien Yan and Yen-Yu Chen, DCT-Based Image Compression Using Wavelet-Based Algorithm with Efficient Deblocking Filter, 14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, Sept. 4-8, 2006.
[11] Hznnu Olkkonen and Juuso T. Olkkonen, Discrete Wavelet Transform Structures for VLSI Architecture Design, InTech Publ. 2010, ISBN: 978-953-049-0.
[12] Alan V. Oppenheim, Roland W. Schafer and John R. Buck, Discrete-Time Signal Processing, 2nd Edition, Ed. Prentice Hall, 2008, ISBN 978-84-205-2987-5.