THE NEW LAPPED ORTHONORMAL TRANSFORM (NLOT) WITHOUT BLOCKING ARTIFACTS

Le Quang Tuan - Vu Đinh Thanh - Nguyen Kim Sach

(Received on November 8th, 2001)

ABSTRACT:

The block transform-DCT- is commonly used for many video, image, and audio coding standards, such as MPEG and JPEG. However, DCT causes "blocking artifacts" at high compression ratio. Recently, some classes of new transfroms for the reduction of blocking effects have been introduce such as the lapped orthogonal transforms (LOTs or LOT), GenLOTs (Generalized LOTs), VLLOTs (Variable Length LOTs)... Those LOTs, which are based on DCT, reduce blocking effect, but do not remove it perfecty.

The paper will introduce a new class of LOTs, which is called NLOTs. The NLOTs remove perfectly "blocking artifacts" and have the same computational complexity like the LOTs.

I. INTRODUCTION:

Nowadays, Transform Coding (TC) is an efficient and commonly used tool for signal compression. TC is used as a basis for many video, image, and audio coding standards, such as MPEG and JPEG. In the encoding process, the signal is divided into blocks of N samples (or N x N for images). For each block, a transform operator is applied. The resulting transform coefficients are quantized (usually, via scalar quantizers) and entropy encoded. At the decoder, the inverse operations are performed: entropy decoding, dequantization, and inverse transformation. The decoded blocks are combined to form the reconstructed signal.

DCT-based image compression was state-of-the-art, but researchers were uncomfortable with the so-called "blocking artifacts" which are common and annoying artifacts like noise with the checker-board pattern, found in images which were compressed at low bit rates using block transforms. The blocking effect is a natural consequence of the independent processing of each block. Blocking artifacts arise because the concatenation of the reconstructed blocks generates signal discontinuities across block boundaries. It is perceived in images as visible discontinuities in features at the cross block boundaries [1], [2], [3], [4].

Some methods for the reduction of blocking effects have been previously suggested. Three methods were proposed: overlapping, post-filtering and using a lapped orthogonal transform or LOT.

In the overlapping method, the blocks overlap slightly, so that redundant information is transmitted for the samples in the block boundaries. The disadvantage of this approach is the increase in the total number of samples to be processed, and thus an increase in the bit rate. So, it was not applied in image/video compression.

In the filtering method, the coding process at the transmitter is unchanged, and at the receiver, a low-pass filter is applied only to the boundary pixels. Although this method does not increase the bit rate, it blurs the signal across block boundaries.

The lapped orthogonal transforms have the same benefits of the overlapping method, but without an increase in the bit rate.

Malvar [1], [2] gave the LOT an elegant design strategy and a fast algorithm, thus making the LOT practical and a serious contender to replace the DCT for image compression.

Recently, new classes of lapped Transforms-LTs with symmetric bases were developed yielding the class of generalized LOTs (Gen LOT) [3], VLLOTs. The VLLOTs [4] were made to have basis functions of arbitrary length (not a multiple of the block size). GenLOTs, VLLOTs are DCT to develop the new basis functions as Malvar did. Those LOTs reduce blocking effect, but does not remove it perfectly.

The rest of paper is organized as follows: Section 2 introduces The Malvar's LOT and the method to design our new LOT. Some experimental results are presented in section 3. Finally, section 4 concludes the paper.

II. THE OPTIMAL LOT:

2.1 The Malvar's Optimal LOT [1]:

In this section, we review the properties of the Malvar's LOT [1], [2] necessary for its analytical derivation in the explanation of our NLOT.

We assume that the signals to be processed are one-dimensional: extension to two or more dimensions is easily achieved by defining separable transforms based on the one-dimensional profile.

Let us assume that the incoming discrete-time signal is a large segment of MN samples, where N is the block size. In traditional transform coding, M-blocks of length with N samples would be independently transformed and coded. In matrix notation, if we call x_0 the original input vector of length MN, the vector y_0 containing the transform coefficients of all blocks is given by

$$y_0 = T' x_0 , \qquad (1)$$

where T' is the transpose of an MN x MN block-diagonal matrix T in the form

$$T = \begin{bmatrix} D.......0 \\ ...D.....0 \\ \vdots....\ddots... \\ 0.........D \end{bmatrix} \qquad (2)$$

where D is a matrix of order N, whose columns are the basis functions that define the transform of each block.

With the LOT, each block has L samples, with L>N, so that neighboring blocks overlap by L – N samples. The LOT maps the L samples of each block into N transform coefficients. With the number of transform coefficients being equal to the block size there is no increase in the data rate. The LOT can be defined as in (1), with T given by

$$T = \begin{bmatrix} P_1 \ldots\ldots\ldots 0 \\ \ldots P_0 \ldots\ldots 0 \\ \ldots\ldots \ddots \ldots \\ \ldots\ldots\ldots P_0 \\ 0 \ldots\ldots\ldots P_2 \end{bmatrix} \quad (3)$$

where P_0 is an L x N matrix that contains the LOT basis functions for each block. We have assumed $L \leq 2N$, i.e., the length of each basic function is at most twice the block size. This choice will be justified later. The matrices P_1 and P_2 are introduced, because the first and last blocks of a segment have only one neighboring block. And thus the LOT for the first and last blocks must be defined in a slightly different way, to guarantee that none of the basis functions extends beyond the segment boundaries. We shall concentrate on P_0 now.

We note that the LOT of a single block is not invertible, since P_0 is not square. Nevertheless, in terms of reconstructing the whole segment x_0, all we need is the invertibility of T. Orthogonality of T is also a desirable property, as with all transforms in traditional transform coding, since it guarantees good numerical stability. In order for T to be orthogonal, the columns of P_0 must be orthogonal.

$$P'_0 P_0 = I, \quad (4)$$

and the overlapping functions of neighboring blocks must also be orthogonal.

$$P'_0 W P_0 = P'_0 W' P_0 = 0, \quad (5)$$

where I is the identity matrix, and the shift operator W is defined as:

$$W = \begin{bmatrix} 0I \\ 00 \end{bmatrix} \quad (6)$$

The above identity matrix is of order L – N, and we have assumed L = 2N. We will say that a LOT matrix P_0 is feasible if it satisfies (4) and (5).

Malvar suggested a direct approach [1], for the derivation of an optimal LOT when L = 2N, i.e., the basis functions of neighboring blocks overlap by N samples. The key point is to start with a feasible LOT matrix P that is not necessarily optimal. Then, the matrix

$$P_0 = PZ, \quad (7)$$

is also a feasible LOT for any orthogonal Z, since

$$P'_0 P_0 = Z'P'PZ = Z'Z = I, \quad (8)$$

$$P'_0 W P_0 = Z'P'WPZ = 0, \quad (9)$$

He introduced a feasible LOT from the DCT, by

$$P = (1/2) \begin{bmatrix} D_e - D_o & D_e - D_o \\ J(D_e - D_o) & -J(D_e - D_o) \end{bmatrix} \quad (10)$$

where D_e and D_o are the N x N/2 matrices containing the even and odd DCT-II functions, respectively.

With P as in (10), what he needed to obtain an optimal LOT is to find an optimal Z in (7) with maximal transform coding gain, G_{TC} [1].

$$G_{TC} = \frac{\frac{1}{N}\sum_{i}\sigma_i^2}{(\prod_{i=1}^{N}\sigma_o^2)^{1/N}} \quad (11)$$

where σ_i^2 , σ_o^2 is the its diagonal entry of the matrix R_{xx} , R_{yy}. R_{xx} is the autocorrelation matrix of original image. R_{yy} is the autocorrelation matrix of reconstructed image.

We can choose Z to diagonalize R_{yy}

$$R_{yy} = Z'P'_0R_{xx}P_0Z, \quad (12)$$

With P and R_{xx} fixed, it is clear the G_{TC} is maximized when R_{yy} is diagonal, i.e., when the columns of Z are the eigenvectors of $P'R_{xx}P$. With such a Z, the LOT matrix P_0 is optimal.

It is important to point out that optimization approach leads to an optimal LOT that is tied to the choice of the initial matrix P. Since each column of P has L elements, with L>N, they span an N-dimensional subspace of $\Re^L$. For any Z, the matrix PZ will always belong to that subspace and so will the optimal LOT. However, there may exist a feasible LOT P that does not belong to the subspace spanned by the columns of P, i.e., it cannot be generated by (7).

For the first-order Gauss-Markov model with the inter-sample correlation coefficients $\rho = 0.95$, the columns of the optimal P_0, Malvar's LOT, are shown in Fig.1a.

Comment on Malvar's LOT:

- Following Malvar' papers [1][2], the key to reduce blocking effects is that the low-order basis functions decay toward zero at their ends.

-The blocking effects disappear (or with the lowest effects) if the low-order basis functions are zeros at their ends.

-In Malvar's LOT, the first basis function, for example, has a boundary value that is 5.83 times lower than its value at the center. Therefore, the discontinuity from zero to the boundary value is much lower than that of the standard DCT functions, and this is one of the main reasons why blocking effects are reduced.

-However, Malvar's low-order basis functions are not exactly zeros at their ends. For example, the first order basis function of Malvar's LOT is shown in Fig.1a.

2.2. Our LOT with the zero-end basis functions:

We suggest an algorithm to design a new LOT, NLOT, with the zero-end basis functions. Our design is also based on D_e and D_o as Malvar did, but we use an appropriate window function to force the ends of NLOT basis functions to zeroes. The function of the window function is weighing to P. In order to preserve the symmetry properties of NLOT, the weighing function must be evenly symmetrical and have zeros at their ends.

Since a modification of a basis function affects the orthonornality of the other basis functions, we apply the Gram-Schmidt orthogonalization procedure to rebuild the system of the orthonormal basis functions ...

That algorithm initiates from the Malvar's P matrix, as followed:

1. Calculating the window functions such as Blackman, Bartlett, Hamming or Hanning.

2. For every basis of P, we do:

- Weighting that basis with the window function. This step will force the ends of basis function to zeroes.
- Applying the Gram-Schmidt orthogonalization procedure to P. After this step, the ends of the above basis function keep zeroes as well.

3. Creating the optimal NLOT P_0, by finding the matrix Z which is the eigenvectors of $P' R_{xx} P$, as Malvar did. That is $P_0 = P*Z$.

The matrix multiplication of PZ still preserves the zero-ends of basis functions of P_0.

Our one-dimensional basis functions with the Blackman window (the columns of P_0) can be seen in the Fig.1b. For the comparison between NLOT and Malvar's LOT, we also introduce the first order basis function of NLOT in Fig.1b. According to the type of window functions, our basis functions have the boundary values those are more than 100 times (or towards infinity when the ends are zeroes) lower than their values at the center.

III. SIMULATION RESULTS:

NLOT and Malvar's LOT coder and decoder are created in MATLAB, in order to compare the blocking effects with DCT coder and decoder that were also implemented. In those coder and decoder, the mean of the test image is first removed. The processed image is then transformed using 8 by 8 blocks (with 8-pixel overlap for the LOT). The coefficients are rounded to the nearest integer values. The entropy is calculated to yield the storage bit-rate, given in bits per pixel. In practice the better quantizer and entropy en/decoder could be used to further reduce the bit rate, but is not done here for simplicity.

In our simulations, the testing images Lena (Fig.3a) and Barbara (Fig.4a) are used and give the following results:

- Fig.2 give the comparisons of Transform Coding Gain (G_{TC}) (11) in dB between LOTs and DCT. From these tables, G_{TC} of our NLOT is better than that of DCT, but lower than that of Malvar's LOT. However, the important problem is the visual quality of reconstructed images.

- Fig.3b and Fig.4b are the reconstruction images with low bit rate, using the DCT. Fig.3c, d and Fig.4c, d, are the low bit rate reconstruction images using the Malvar's and our NLOT.

With the nearly similar PSNR gain, the difference of visual quality between DCT and LOTs is evident. We cannot accept the reconstructed image with DCT because of the blocking effect. Therefore, the parameter PSNR could not show the perceptual quality of visuality in low bit-rate.

- The difference of visual quality between Malvar's and our NLOTs is subtle, but can be noticed upon careful examination of the smooth areas of the images. It seems that the images with Malvar's LOT are suffered from a type of "spotted noise" which spreads over the surfaces of images.

- We think that the spotted noise is caused by the non-zero boundary values of basis functions in Malvar's LOT. Our NLOT, with the non-zero boundary values of basis functions, improves the perceptual quality in low bit-rate. Our NLOT can replace DCT in low bit-rate compression for video/image.

IV. CONCLUSION:

We have derived a new optimal set of overlapping basis functions, NLOT. Unlike the Malvar's derivation, where the boundary values of basis functions are non-zero, we have obtained the zero-end LOT as the solution to solve fully the blocking effect in low bit-rate reconstruction.

We believe that the NLOT introduced in this paper allows the implementation of block coding systems at low bit rates (below 1.0 bits per sample) with much less noticeable blocking effects than DCT or Malvar's LOT-based transform coding.

However, our NLOT may not be easily factorable such as DCT or Malvar's LOT. In the future, we try to find a fast algorithm for our LOT.

ACKNOWLEDGMENT

The authors would like to thank Dr. H.S.Malvar for much valuable documentation.

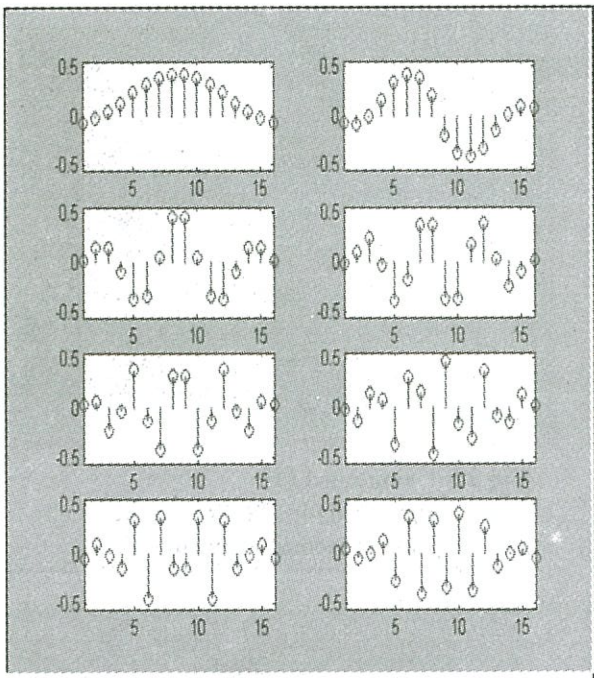

Fig.1a: Malvar's LOT basis functions with N=8, L=2N=16.
The first Malvar's basis-vector:
[**-0.0686, -0.0320**, 0.0372, 0.1285, 0.2276, 0.3177, 0.3849, **0.4189, 0.4189**, 0.3849, 0.3177, 0.2276, 0.1285, 0.0372, **-0.0320, - 0.0686**]

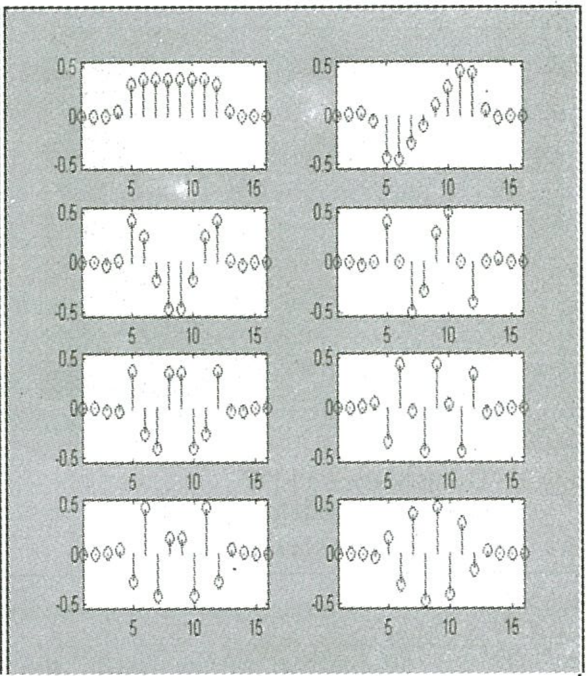

Fig.1b: Our LOT basis functions with N=8, L=2N=16 and "Blackman window".
The first our basis-vector:
[**0.0, -0.0019**, -0.0109, 0.0436, 0.3124, 0.3723, 0.3615, 0.3620, **0.3620, 0.3615**, 0.3723, 0.3124, 0.0436, -0.0109, **-0.0019, 0.0**]

Type of Transform	G_{TC} (dB)
Malvar's LOT	9.24
DCT II	8.83

Our LOTs	G_{TC} (dB)
Type of window	
Hamming	9.07
Hanning	9.06
Blackman	8.95
Bartlett	9.09

Fig 2: Comparisons of Transform Coding Gain (G_{TC})

Fig.3a: Original image, "Barbara", (image zoomed in two times)

Fig.3b: Using DCT II; bpp = 0.701 PSNR = 25.297 (image zoomed in two times)

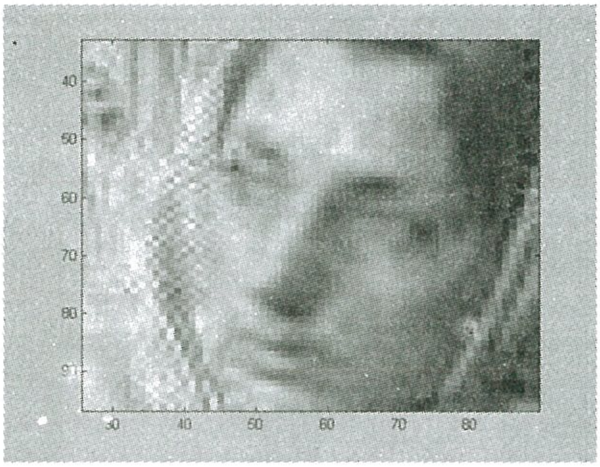

Fig.3c: Using Malvar's LOT
bpp =0.616, PSNR = 25.638
(image zoomed in two times)

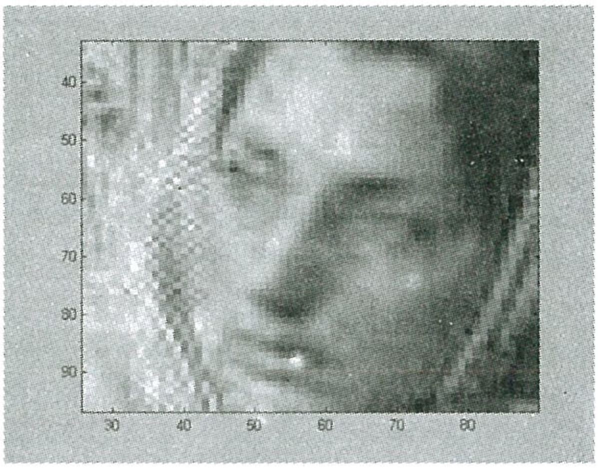

Fig.3d: Using our LOT
bpp = 0.634; PSNR = 24.815
(image zoomed in two times)

Fig.4a: Original image, "Lena"
(image zoomed in two times)

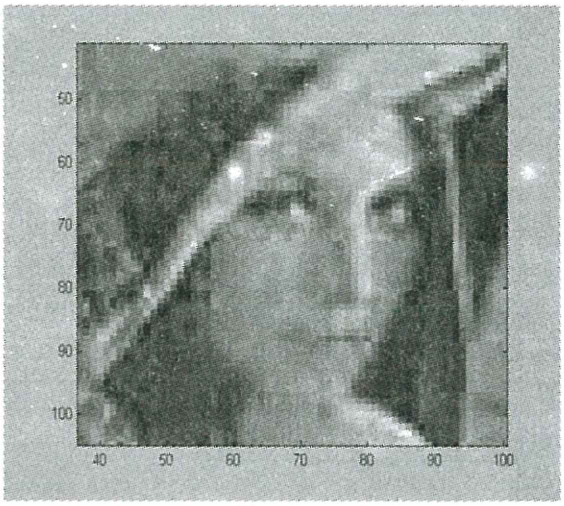

Fig.4b: Using DCT II, (image zoomed in two times) bpp = 0.452 , PSNR = 27.048

Fig.4c: Using Malvar's LOT; (image zoomed in two times) bpp = 0.384 ; PSNR = 27.322

Fig.4d: Using our LOT; (image zoomed in two times) bpp = 0.398; PSNR =26.880

BIẾN ĐỔI TRỰC CHUẨN LOT MỚI KHÔNG GÂY RA MÉO KHỐI

Lê Quang Tuấn, Vũ Đình Thành - Trường Đại Học Bách Khoa, ĐHQG-HCM
Nguyễn Kim Sách - TT Nghiên cứu Ứng Dụng & Phát Triển Truyền Hình

TÓM TẮT:

Biến đổi khối DCT hiện nay đang được sử dụng phổ biến trong các chuẩn nén ảnh JPEG và MPEG. Tuy nhiên, khi nén với tỉ lệ cao, DCT sẽ gây ra méo khối nghiêm trọng. Thời gian qua, nhiều lớp biến đổi mới được xây dựng để giảm bớt loại méo này, như LOT,GenLOT(Generalized LOT),VLLOT (Variable Length LOT)... Những lớp biến đổi này đều được xây dựng trên cơ sở DCT nhưng chỉ làm giảm, không triệt tiêu hoàn toàn méo khối.

Bài viết này trình bày một lớp biến đổi LOT mới, NLOT, không gây ra méo khối và có cùng mức độ phức tạp tính toán như LOT.

REFERENCES

[1] H.S.Malvar and D.H.Staelin, "The LOT: transform coding without blocking effects," IEEE, Trans. on Acoustics, Speech , and Signal Processing, vol.37, pp.553-559, Apr.1989

[2] H.S.Malvar, "Lapped biorthogonal transforms for transform coding with reduced blocking and ringing artifacts," Proc.IEEE Int.Conf. on Acounstics, Speech, and Signal Processing Munich, pp.2421-2424, Apr.1997.

[3] R.L.de Queiroz, T.Q.Nguyen, and K.R.Rao, "The GenLOT: generalized linear-phase Lapped orthogonal transform," IEEE, Trans. on Signal Processing, vol.44, pp.497-507, Mar.1996.

[4] T.D.Tran, "Linear phase perfect reconstruction filter banks: Theory, Structure, Design, and Application in image compression," Ph.D.thesis, University of Wiscosin, Madison, W.I, May.1998.