Lapped Multiple Bases Realizations for the Transform Coding of Still Images

Victor DeBrunner, Lixiang Chen and Hongjian Li
vdebrunn@mailhost.ecn.uoknor.edu, phone (405)325–4330 or fax (405)325–7066
School of Electrical Engineering, The University of Oklahoma
Norman, OK 73019

Abstract

We describe a system for still image compression which uses several lapped orthogonal transform (LOT) sets in a multiple bases realization algorithm, the recursive residual projection (RRP) algorithm. Newly developed RRP algorithms are shown to reduce the number of encoded transform coefficients 20% beyond the DCT-only compression standard, JPEG. These algorithms still suffer from the problem of block-discontinuities at the boundaries of the segmented image. We extend these algorithms to use several newly developed LOT bases. Fast hardware implementations of these algorithms are presented.

1: Introduction

We are currently developing block-transform coding image compression systems. The algorithms being discussed here are based on the seminal works in image compression [1,2]. We use multiple bases realizations (MBR) of images in our system. These techniques have been used recently to compress speech [3,4] and image data [5,6]. One motivation for using the mixed representations is their ability to efficiently encode various spectral features of the signal which is to be compressed [7].

The basic motivation behind the development of the Lapped Orthogonal Transform (LOT) [2] comes from one of the major disadvantages of traditional block coding methods – the block effects, which are discontinuities in the reconstructed image. In our algorithms, which are based on modifications of the Recursive Residual Projection (RRP) algorithm [1,3] detailed in [8], we start by dividing the input image into 8x8 sub-images. Each sub-image is then independently compressed using an RRP algorithm. The selected transform coefficients consist of two sets of data, one set containing the basis functions and the other holding the corresponding values of the projections onto the basis functions. The set of values is quantized using a developed γ–quantizer. Finally, the coefficients and their positions are encoded using a lossless coding technique.

Because of the independent processing of each block, coding loss can produce discontinuities in the image since the reconstructed pixels of one block will, most likely, not match with the pixels of the next block. This phenomenon is especially apparent when very low numbers of coefficients are retained in the coding, i.e. in low bit-rate systems. This is called the "blocking" effect. The blocking effect produces a reconstructed image that seems to be built of small tiles, and can be visually annoying. In some applications for which we are pursuing our research, the annoyance becomes more critical, in that vital information may be lost. This paper discusses our approach to these problems: efficient modeling via the MBR which incorporates ideas from the LOT. We see that such a system efficiently encodes the signal while alleviating the blocking problems.

2: MBR Algorithms

MBR algorithms have recently been developed for low bit-rate coding of voice [4,7]. These algorithms are currently being examined and improved [9,10]. These algorithms provide an elegant solution to the problem of encoding signals with varying features. Even more recently, a version of these algorithms has been presented for still image coding [6]. These algorithms use several dominant component (DC) picking strategies, in combination with either a steepest descent or a gradient search algorithm to determine optimal basis function projections. The algorithms rely on a cascaded structure of transform blocks, and we have seen that these algorithms do not always converge, and they may take many iterations before they converge [9,10]. Finally, even if they do converge, the resulting performance is inferior to the RRP algorithm given in [1] and the modifications to the RRP algorithm developed in [8].

A two-transform RRP algorithm flow-graph is given in Figure 1. A modified version, which we call the Look-Back RRP (LBRRP) was derived in [8]. The LBRRP algorithm works similarly, except that we re-examine the values of the previously selected basis functions. If they are not zero, we add the current value to the existing value, and eliminate it once more. Note that this is necessary and justified, because the MBR algorithms do not remove orthogonal basis functions.

3: Fast LOT Algorithms

The LOT is based on decomposing a signal segment of length *N* into *K* blocks of length *blk_sz*. The LOT then is a controlled $M \geq blk\_sz$ to *K* map. For our purposes in this paper, we assume that *M=2*blk_sz*. In this paper, we consider fast LOT algorithms based on the discrete cosine, discrete Walsh and discrete slant transforms. Other transform types have been developed [11].

3.1: A LOT Based on the Discrete Cosine Transform (DCT)

The fast algorithm used to compute the LOT and inverse LOT based on the discrete cosine transform (LOTDCT/ILOTDCT) is based on the algorithm given in [2]. We see that the LOTDCT is markedly superior to the DCT for representing high frequencies. A flow-graph for one 8-point LOTDCT is shown in Figure 2. Notice two important items: 1) the LOTDCT shown operates on an interior data segment, the endpoints must be treated differently. For details, we refer the reader to [11]. 2) The algorithm shown is "fast" because it relies on fast transforms and 2x2 butterflies. A point we will discuss in more detail in Section 5 is that the performance of the compaction is sensitive to the angles of rotation used in the 2x2 butterflies.

3.2: A LOT Based on the Discrete Walsh Transform (DWT)

The fast algorithm used to compute the LOT and inverse LOT based on the discrete Walsh transform (LOTDWT/ILOTDWT) is developed using the ideas given in [2]. Here, the DWT basis functions are re-ordered so that even and odd functions alternate in increasing sequency to form the LOT basis. 2x2 butterflies are given to complete the transform and its inverse. We do not see much difference in energy compaction between the DWT and the LOTDWT. Note that the LOTDWT computation consists of ± 1 "multiplications" and adds, with only the added *blk_sz–1* plane rotations in the 2x2 butterflies being true multiplications.

3.3: A LOT Based on the Discrete Slant Transform (DST)

The fast algorithm used to compute the LOT and inverse LOT based on the discrete Slant transform (LOTDST/ILOTDST) is based on the algorithm given in [2]. Computation of the LOTDST basis functions takes the form of the LOTDWT computations. The LOTDST is superior to the DST in its energy compaction performance.

4: Image Compression Using Multiple LOT Bases Functions

Our RRP-based algorithms employ combinations of two different basis functions, usually the DCT and either the DWT or the DST. In our RRP algorithm, the image is first broken into a set of sub–images. The DC value (the average) of each sub-images is then removed. The created image is transformed into various LOT sets of basis functions. For instance, the LOTDCT transform of the input image contains a set of LOTDCT matrices of dimension *blk_sz* and the LOTDWT transform of the input image contains a set of LOTDWT matrices of dimension *blk_sz*. At each iteration, the RRP algorithm searches for the largest projection on both transform sets. The largest projection along with the corresponding basis function is stored in the output data structure, and a new residual image is created. This process is repeated until some stopping criterion is satisfied. The difference between the RRP using the several LOT bases and the original RRP using conventional transforms is that the operation on each LOT transform of the sub-image results in changes to neighboring residual sub-images. The performance of each combination of transforms is determined by the properties of the image which is to be compressed. The developed "look-back" algorithm, which examines projections onto previously removed basis functions, may be nearly optimal. Previously removed projections move away from zero because of the non-orthogonal decomposition. We add these non-zero projections to their previous value at each iteration of the search. The added cost seems worthwhile, because the energy packing is significantly improved [8].

5: Simulation Results

The method is coded in the "C" language on an RS/6000, and used to decompose various still images with different statistical properties. We have used the techniques on "natural" images, radiological images, aerial photographs, maps and computer generated graphics. In all cases except the computer generated graphics, the RRP algorithms required between 15% and 25% less coefficients to encode the image to a specified quality as measured by mean square error when compared to DCT only coding [8]. For the LOT coding, the performance difference is less significant, though still significant when low numbers of coefficients per block are selected. For instance, we have examined the "natural" image Lena (or U.S.C. Girl, see Figure 3), when several different numbers of coefficients are selected. The Mean Square Error results in percentage are tabulated in Table 1, and the resulting figures for the first row of the table are given in Figure 4.

Table 1: LOTRRP, LOTDCT and DCT comparisons, %MSE

# coefs/ block	DCT	LOT-DCT	LOT-DCT/ LOT-DWT	LOT-DCT/ LOT-DST
1	69.53%	66.44%	65.97%	65.90%
2	56.11%	54.17%	53.82%	53.54%
3	47.83%	47.01%	46.97%	46.36%
4	41.97%	41.93%	42.59%	41.29%
5	37.59%	38.08%	39.54%	38.01%

The results indicate that significant energy packing without blocking effects can be achieved over the more traditional LOTDCT-only coding when low bit rates are desired. At higher bit rates, the DCT-only coding is superior to all LOT-based coding methods.

At the present time, we believe that this phenomena is due to the non-optimal choices of rotation angles in the 2x2 output butterfly stages of the LOT computations. These angles should be chosen based on the statistics of the image being compressed. However, in our algorithms, these statistics change throughout the RRP algorithm, because the transforms operate on residual images to determine successive coefficients. We have just begun to examine this issue.

6: Conclusions and Further Research

Efficient coding schemes without blocking effect for still images are developed and examined. One algorithm, which we call the "look backward" RRP, may be nearly optimal [8]. We will be examining choices of optimal rotation angles on the output butterflies of the individual LOT computational elements.

Support Acknowledgement

This work was supported, in part, by the MICOM Research, Development and Engineering Center (Mr. Michael Pitruzzello) under the auspices of the U. S. Army Research Office Scientific Services Program administered by Battelle (Delivery Order 1099, Contract No. DAAL03-91-C-0034)

7: References

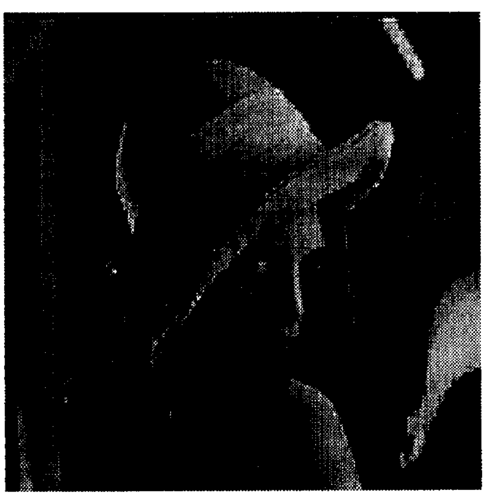
1. R. Khanna, *Image Data Compression Using Multiple Bases Representation,* M.S. Thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA, 24061, September 1990. Advisor: A. A. (Louis) Beex.
2. H. S. Malvar and D. H. Staelin, "The LOT: Transform Coding Without Blocking Effects," *IEEE Trans. on Acoust., Speech, and Sig. Proc.*, vol. 37, no. 4, pp. 553-559, April 1989.
3. F. G. Safar, *Signal Compression and Reconstruction Using Multiple Bases Representation,* M.S. Thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA, 24061, September 1988. Advisor: A. A. (Louis) Beex.
4. A. S. Spanias and P. C. Loizou, "Mixed Fourier/ Walsh transform scheme for speech coding at 4.0 Kbits/s," *IEE Proceedings-I*, vol. 139, no. 5, pp. 473-481, October 1992.
5. V. DeBrunner, "Digital Image Compression," *Final Report, Contract DAAL03-91-C-0034, Delivery Order 818, TCN 93257*, Battelle, Research Triangle Park Office for the U.S. Army Research Office, August 1993.
6. W. B. Mikhael and A. Ramaswamy, "Application of Multitransforms for Lossy Image Representation," *IEEE Trans. on Cir. & Sys—II*, vol. 41, no. 6, pp. 431-434, June 1994.
7. W. B. Mikhael and A. Ramaswamy, "Residual error representation and adaptive minimization for representing nonstationary signals using mixed transforms," *IEEE Trans. on Cir. & Sys.—II*, vol. 39, no. 7, pp. 489-492 , July 1992.
8. H. Li, *An Image Compression System Based on the Recursive Residual Projection Algorithm,* M.S. Thesis, The University of Oklahoma, Norman, OK, 73019, May 1994. Advisor: V. E. DeBrunner.
9. W. Lou and V. E. DeBrunner, "A Constrained Multiple Bases LMS Algorithm for Non-Stationary Signal Representation," *submitted to the technical program, ICASSP'95.*
10. V. E. DeBrunner, W. Lou and J. Thripuraneni, "Dominant Component Picking Strategies for Constrained Multiple Bases LMS Algorithms," *submitted to the technical program, ICASSP'95.*
11. V. E. DeBrunner and L. Chen, "LOT MBR Still Image Compression System," Midterm Report, *Delivery Order 1099, Contract No. DAAL03-91-C-0034*, Battelle, Research Triangle Park Office for the U.S. Army Research Office, May 1994.

Figure 3: Original Lena (U.S.C. Girl)

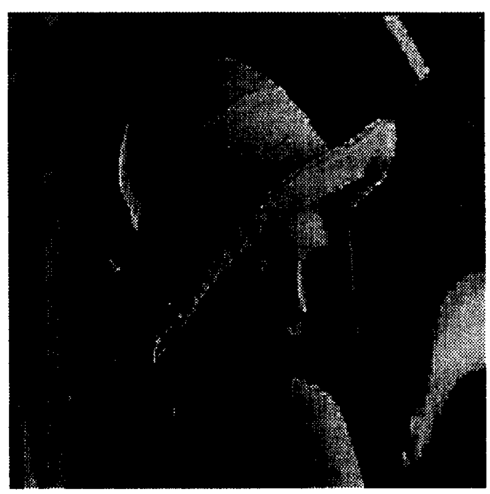
Figure 4: "Lena" coded with 1 coefficient/8x8 block

DCT

LOTDCT



LOTDCT/LOTDWT



LOTDCT/LOTDST

Figure 1: RRP based on two LOT transforms

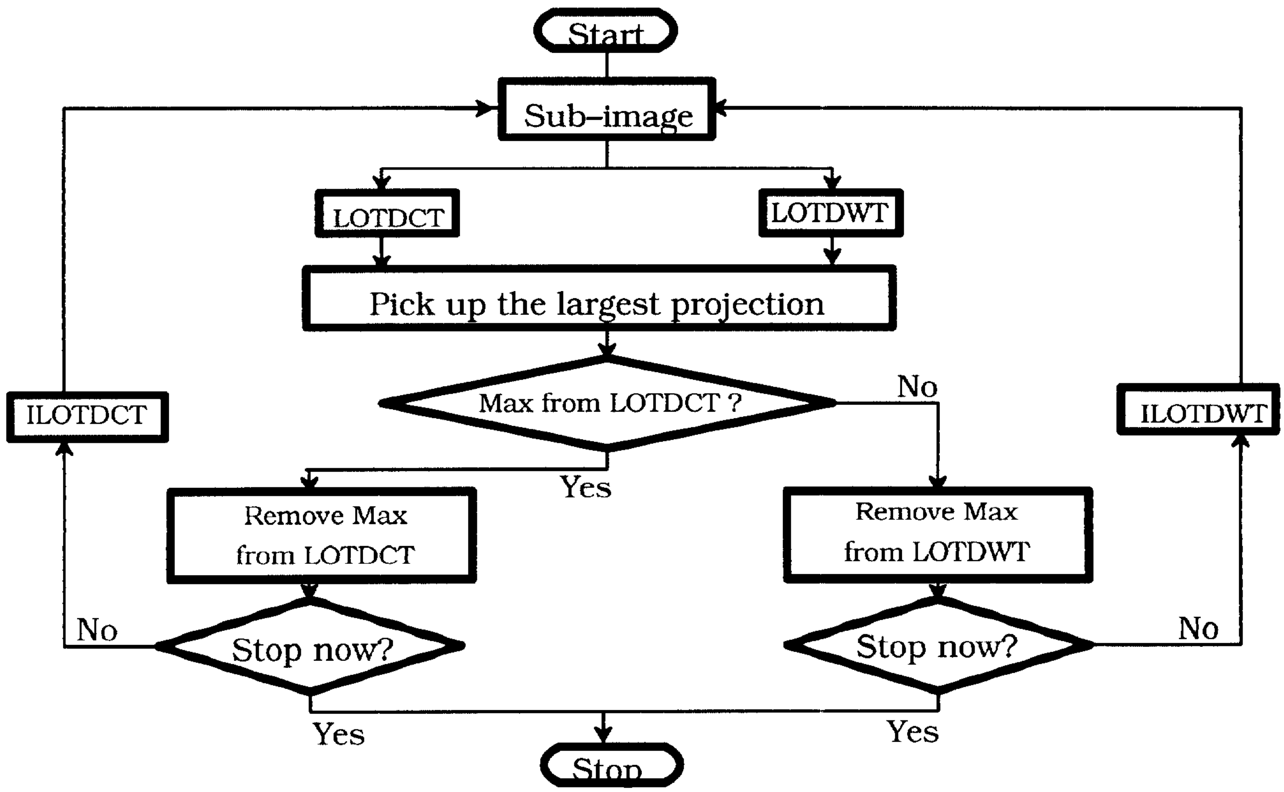


Figure 2: A Single Block LOTDCT Transform with Butterflies