A MODIFIED MPEG-1 SYSTEM BASED ON GENLOT

S. H. Oğuz, T. Q. Nguyen and Y. H. Hu

ECE Department, University of Wisconsin-Madison
1415 Johnson Drive, Madison, WI 53706 U.S.A.
Tel: 1 608 2655739; Fax: 1 608 2654623
e-mail: oguz@cae.wisc.edu, nguyen@ece.wisc.edu, hu@engr.wisc.edu

ABSTRACT

In this study, a modification to ISO MPEG-1 and MPEG-2 digital video coding standards is proposed and preliminary results on its performance are reported. The proposed modification aims to improve the visual quality of MPEG-1 and MPEG-2 coding at medium-to-low bit-rate regimes by eliminating the blocking effect caused by the Discrete Cosine Transform. This goal is achieved without introducing a significant change in the MPEG hierarchy and algorithm. The theory of Lapped Orthogonal Transforms which constitutes a rather recently introduced tool for block transform coding suggests that they can reduce the blocking effect to very low levels. Hence, in the modified MPEG-like system, instead of the original two dimensional Discrete Cosine Transform a Lapped Orthogonal Transformation is used as the basic spatial correlation reduction operation and also customized quantization and variable length codeword tables are provided to ensure efficiency. The modified coding algorithm is implemented in software. Simulations are made to compare its performance to the original MPEG-1 algorithm. As performance criteria, PSNR versus compression ratio (equivalently bit-rate) plots and also subjective ratings of visual quality are used.

1. INTRODUCTION

The Discrete Cosine Transform (DCT) has been the most popular orthogonal block transform in signal compression for a long time. The reasons for this are:

• DCT is a signal independent transform whose energy compaction is almost as good as the energy compaction of the statistically optimal (in MSE sense) but signal dependent Karhunen-Loeve Transform, for a large class of signals with low-pass nature.

• DCT can be computed by means of fast algorithms.

However, since the lengths of the basis functions of the DCT are exactly equal to the block size and since these basis functions are not decaying towards zero at both ends, coding with DCT at low bit-rates results in blocking artifacts which are perceived as visible discontinuities in features of the encoded signal across the block boundaries. This constitutes a problem common to all DCT based signal compression codecs (e.g. JPEG standard for still image compression and MPEG-1 for digital video compression).

MPEG-1 and MPEG-2 [1] are well established and popular digital video coding standards which are successful in most of their features. Common to these standards is a very nice hierarchical structure which is aimed at fulfilling certain application and operation environment dictated requirements. MPEG-1 and MPEG-2 syntax introduces three different types of frame coding for normal display purposes. I type coded frames are intra coded using 8x8 sized two dimensional (2-D) DCT applied to non-overlapping sub-images on the frame to be encoded. These frames are used as entry points to independently decodable collections of consecutive frames known as "Groups of Pictures". P and B type coded frames employ variations of motion compensated prediction for temporal redundancy reduction and are therefore essential in achieving a significant compression. The prediction error in P and B type frames is again coded by 8x8 sized 2-D DCT. Consequently, at low bit-rates MPEG-1 and MPEG-2 both suffer from blocking artifacts owing to their use of DCT as the basic spatial correlation reduction technique.

A recently developed class of orthogonal transforms for block based signal coding which is referred to as the Lapped Orthogonal Transforms (LOTs) [2],[3],[4] and their generalized forms known as Generalized Lapped Orthogonal Transforms (GenLOTs) [5] have the potential to reduce the blocking effect to very low levels besides their many other nice features. (We will refer to only GenLOTs from now on since they cover LOTs as a special case.) GenLOTs can reduce the blocking artifacts significantly since their basis functions, which are longer than the block size, overlap adjacent blocks and these basis functions decay toward zero at their both

ends. Besides, GenLOTs have better coding gain than DCT and also have fast algorithms for their implementation. All of these nice features of GenLOTs make them a very good choice to replace DCT and compensate for its artifacts in the MPEG-1 and MPEG-2 video coding standards.

2. THE PROPOSED ALGORITHM

The degradation in the visual quality of MPEG-1 and MPEG-2 coding at low bit-rates is basically due to the blocking effect. Therefore, the proposed modification is essentially the incorporation of a GenLOT into MPEG-1 and MPEG-2 framework to replace the 2-D DCT of the original standard and hence reduce the blocking artifacts.

One main consideration in this work was to keep original MPEG syntax intact as much as possible. For this reason we simply interchanged the 2-D 8x8 sized DCT applied at the block level with a 2-D separable size:16 overlap-factor:2 GenLOT applied at the macroblock level. Since macroblock is the basic unit for motion compensation and it is also the highest level in MPEG hierarchy for the entire of which only one mode of coding is chosen, the use of a GenLOT instead of DCT as mentioned above necessitated only a few changes and apart from these indispensable modifications at and below macroblock level, we kept MPEG syntax as it is.

For example, the "coded block pattern" field of original MPEG syntax at the macroblock level is not used in the GenLOT based modified algorithm since the basic data set to be coded is the entire (transformed and quantized) macroblock but not a portion of it.

Definitely, custom quantization tables for intra- and inter-coded macroblocks and a custom variable length codeword (VLC) table for coding the runlength-level pairs were generated.

As an example, for the design of the default intra quantization table, the original MPEG-1 I type coding algorithm was run on some test frames with quantizer scale index set equal to 1. This value of the quantizer scale index parameter is the lowest value allowed and it is also the value which will result in the best reconstruction quality along with the worst compression ratio. The resulting average compression ratio which turned out to be about 2.4:1, was used to determine the bit-rate budget for the default intra quantization table design. The bit-allocation algorithm we used is based on a new non-iterative algorithm [6] which yields the (non-integer) non-negative optimal bit-rate assignments in a single pass. After the determination of initial bit-rate, GenLOT coefficients were generated for intra coded macroblocks over test frames from six different image sequences, namely Calendar & Train, Claire, Miss America, Salesman, Garden and Football. As an initial choice, we used a quasi-optimal improved frequency response GenLOT in our algorithm. The GenLOT coefficient statistics were also fed to the bit allocation algorithm. Only the DC coefficient was handled separately as in the case of MPEG. Due to its increased dynamic range, it was quantized into nine bits and then coded with a lossless DPCM technique. For the remaining AC coefficients the bit allocation algorithm's result, the standard deviations and the dynamic ranges of the coefficients were used to design the associated uniform scalar quantizers. For the inter-coding quantization table a similar procedure was carried out. Then, the transform domain coefficients generated on test frames from the six different image sequences were quantized by the designed quantizers. The resulting indices were scanned in an approximately increasing frequency scan order ("Square Scan") illustrated in Figure 1 for a 9x9 portion of a macroblock. The reason for using this

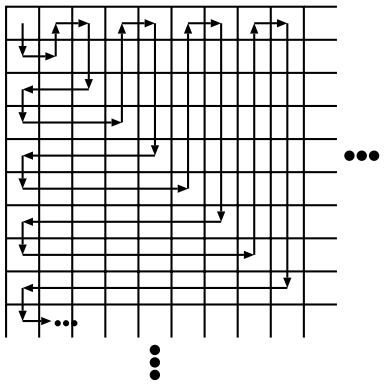

Figure 1: The square scan pattern used to form runlength-level pairs for encoding the quantization indices in the modified algorithm.

type of scanning is that due to structure of the designed default intra quantization table certain coefficients after quantization are known to be zero and these are placed in succession (along vertical strips) by this scanning method to increase the efficiency of the encoding. Runlength-level pair statistics formed in this way from intra- and inter-coded macroblocks were fed to the Huffman coding algorithm to generate a common VLC table for coding the quantization indices.

3. RESULTS AND CONCLUSION

The performance of MPEG coding is very strongly dependent on the coding quality of I type pictures due to the dependence of P and B type pictures on I pictures as their references. If there is no excessive motion or texture in the frames and the coding quality of I pictures is good, then the coding quality of P and B type pictures turn out to be good and they also result in the desired high coding efficiency.

Blocking effect is a serious problem for especially I type pictures which will immediately (propagate to and) effect P and B type pictures. It can typically be observed on I type pictures beyond a compression ratio of 15:1. So our basic effort is to avoid blocking effect in I type coding by the use of GenLOT.

The average peak signal to noise ratio (PSNR) in decibels (dB) versus average compression ratio (CR) curves for MPEG and GenLOT based MPEG (LOTBMPEG) I type coding on Salesman sequence are shown in Figure 2. Figure 3 shows the difference

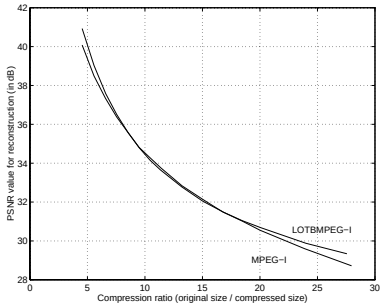

Figure 2: Average PSNR (dB) versus average CR curves for MPEG and LOTBMPEG I type coding. (Results are based on the Salesman sequence).

PSNR $_{LOTBMPEG}$ (dB) - PSNR $_{MPEG}$ (dB) versus average compression ratio (for the data of Figure 2) in more detail. As can be seen from Figure 3, GenLOT

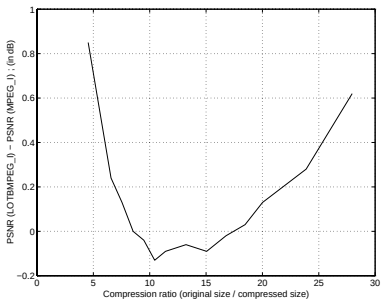

Figure 3: The plot of the difference PSNR $_{LOTBMPEG}$ (dB) - PSNR $_{MPEG}$ (dB) versus average CR for the data of Figure 2.

based MPEG-like algorithm performs significantly better on the average for high and low bit rate regimes in terms of PSNR as compared to standard MPEG. Based on careful subjective evaluations, we also concluded that GenLOT based MPEG-like algorithm successfully removed the blocking artifacts. However, due to the longer basis functions of GenLOT, ringing effects were observed at low bit rates. Perceptual weighting of the transform coefficient variances before the quantizer design for intra coded macroblocks could improve this drawback of using GenLOTs.

Figure 4 shows the average PSNR (dB) versus average compression ratio curves for MPEG and LOTBMPEG P type coding on Salesman sequence. As can

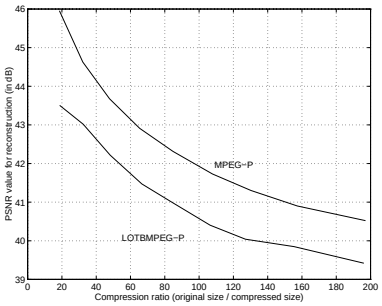

Figure 4: Average PSNR (dB) versus average CR curves for MPEG and LOTBMPEG P type coding. (The results are based on the Salesman sequence.

be seen from this plot, LOTBMPEG follows closely MPEG in P type coding but has an almost constant 1.5 dB inferior performance in terms of PSNR. The reason for this is lying in the initial design of the runlength-level vlc codebook for coding quantization indices. Basically, in this initial design equal weights were given to statistics from intra- and inter-coded macroblocks. However, for a typical application the number of inter-coded macroblocks are significantly larger than the number of intra-coded macroblocks both at the P or B picture level and also at the group of pictures (GOP) level. So, we are confident that this 1.5 dB inferior performance in PSNR levels can be compensated for by a careful new design of the relevant codebook. Figure 5 shows the 5th frame from the Salesman sequence. Figure 6 shows the result of applying MPEG P type

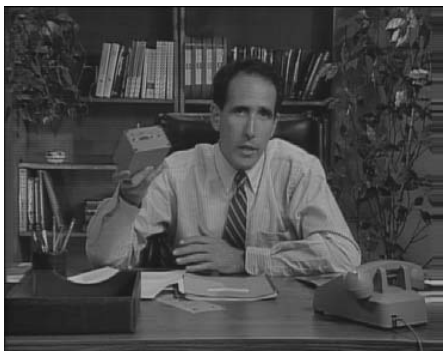

Figure 5: Original 5th frame from Salesman sequence.

coding to the frame shown in Figure 5 at compression ratio 197:1 with a corresponding PSNR value of 40.52 dB. Similarly, Figure 7 shows the result of applying LOTBMPEG P type coding to the same original frame at the same compression ratio as above and with corresponding PSNR 39.42 dB. Notice the blocking effects in Figure 6 and also the ringing effects in Figure 7.

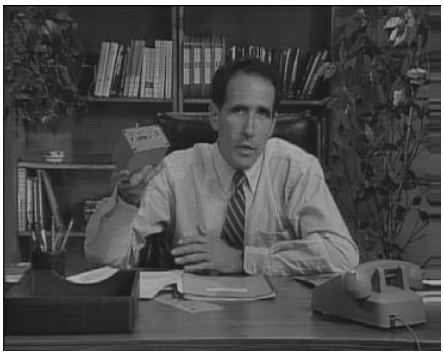

Figure 6: Reconstruction from the frame in Figure 5 coded by MPEG P type coding. CR= 197:1, PSNR= 40.52 dB.

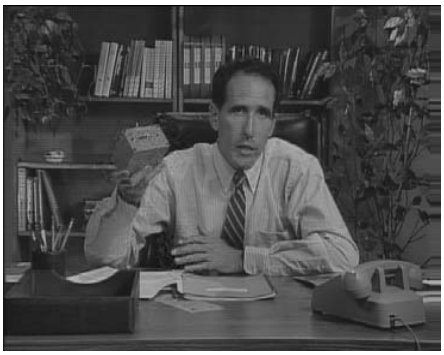

Figure 7: Reconstruction from the frame in Figure 5 coded by LOTBMPEG P type coding. CR= 197:1. PSNR= 39.42 dB.

Finally, Figure 8 shows the average PSNR (dB) versus average compression ratio curves for MPEG and LOTBMPEG B type coding applied to Salesman sequence. It is interesting to note that this time LOTBMPEG B type coding is performing slightly better than MPEG B type coding in spite of its poor codebook. This could be explained by noting that the coding type decision differences between B and P type picture coding algorithms might have shifted the data to be coded towards a more efficiently codable region.

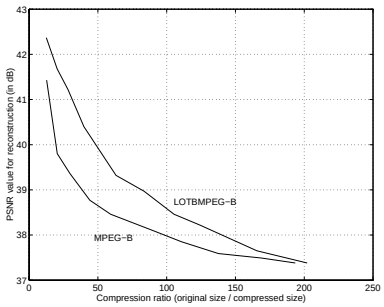

Figure 8: Average PSNR (dB) versus average CR curves for MPEG and LOTBMPEG B type coding. (The results are based on the Salesman sequence).

Based on these preliminary results, it is clear that GenLOT as a substitute for DCT in MPEG coding algorithm was able to remove blocking effects significantly. However, in return it introduced ringing effects which is another severe artifact at low bit rates. A careful further study to control ringing effects associated with GenLOTs may open the way to fully justify the use of GenLOTs instead of DCT in popular transform coding based codecs.

4. FUTURE WORK

Perceptual weighting of GenLOT coefficient variances for default intra quantizer design and the design of two different VLC tables for coding runlength-level pairs from intra- and inter-coded macroblocks are two indispensable future work items and these two points are very likely to significantly enhance the performance of the proposed algorithm. Also, further study on the ringing effects associated with GenLOTs and extensive simulations and comparisons between the two algorithms are necessary for a firm conclusion on the effectiveness of using GenLOTs instead of DCT. Finally, besides subjective evaluations, both systems' performances should be evaluated objectively in terms of blocking effects through some recently reported analytic measures of blocking effect visibility [7].

REFERENCES

[1] Recommendation H.262, ISO/IEC 13818, Generic Coding of Moving Pictures and Associated Audio, Draft International Standard of MPEG-2.

[2] H. S. Malvar, *Signal Processing with Lapped Transforms.* Boston: Artech House, 1992.

[3] 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, April 1989.

[4] H. S. Malvar, "Lapped Transforms for Efficient Transform/Subband Coding," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. 38, pp. 969-978, June 1990.

[5] 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, March 1996.

[6] K. Gosse, F. Moreau de St. Martin, and P. Duhamel, "Filter Bank Design for Minimum Distortion in Presence of Subband Quantization" *Proc. IEEE ICASSP-96, The 1996 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Atlanta, USA, May 1996.

[7] S. A. Karunasekera and N. G. Kingsbury, "A Distortion Measure for Blocking Artifacts in Images Based on Human Visual Sensitivity," *IEEE Trans. on Image Processing*, Vol. 4, pp. 713-724, June 1995.