INTERFRAME VIDEO CODING USING OVERLAPPED MOTION COMPENSATION AND PERFECT RECONSTRUCTION FILTER BANKS

Hirohisa Jozawa †, Hiroshi Watanabe † and Sharad Singhal ‡

† NTT Human Interface Labs., 1-2356 Take, Yokosuka, 238 JAPAN
‡ BELLCORE, 445 South Street, Morristown, NJ 07960, USA

ABSTRACT

Interframe video coders using the Discrete Cosine Transform (DCT) and Motion Compensation (MC) produce block artifacts in the decoded video at low bit-rates. Recent results on Lapped Orthogonal Transforms (LOT) suggest that it can reduce these artifacts. However LOTs are difficult to use efficiently with motion compensation because of block overlap. In this paper, we propose a new video coding algorithm that forms a natural connection between LOTs and motion compensation using two novel concepts: a) Overlapped Motion Compensation (OMC) and b) Overlapped Macroblocks using frequency domain coefficients.

1. INTRODUCTION

Several widely known video coding algorithms, such as the CCITT Recommendation H.261 and the proposed ISO/MPEG 11172 video coding standard, use the Discrete Cosine Transform (DCT) and Motion Compensation (MC). While the DCT/MC coders are efficient and achieve high compression ratios, they also result in blocking artifacts in the picture, especially at low coding rates.

Lapped Orthogonal Transform (LOT) has been proposed [1],[2] as another transform useful for video coding. Unlike the DCT, the LOT basis in a given block overlaps adjacent blocks and thus it produces fewer blocking artifacts than the DCT [3]. However, if motion compensation is used with the LOT, some of its advantages are lost because blocking effects due to motion compensation become visible unless the LOT coefficients are quantized finely.

Windowed Motion Compensation (WMC) [4] has been found useful in reducing the blocking artifacts when used with the DCT. In this paper, we extend the concept of WMC to Overlapped Motion Compensation (OMC) by generalizing the concept of windowing to linear filtering on overlapped motion blocks. We connect WMC and LOT by the concept of overlapped macroblocks and show that the required computation can be reduced by doing the filtering in the coefficient domain.

We describe the basic structure of a predictive coding scheme using LOT/WMC and show how intra blocks can be adaptively coded in the coefficient domain. In addition, we show how the algorithm can be adapted to include loop filtering by truncating LOT coefficients. Other properties, such as layering for robust transmission over ATM and obtaining multiple picture resolution is also described briefly.

2. LAPPED ORTHOGONAL TRANSFORM

The Lapped Orthogonal Transform (LOT) has been developed to reduce the blocking effect due to independent processing of each block as is done in DCT based coding schemes [1]. Although direct comparisons with DCT coding are not available, the coding efficiency of LOT is also reported to be close to Karhunen-Loeve Transform (KLT) [2]. The one-dimensional N-point LOT transforms $2N$ samples x into N coefficients f:

$$f = L_0^T x, \tag{1}$$

and the corresponding ILOT transforms the N coefficients back into $2N$ samples:

$$x' = L_0 f, \tag{2}$$

Note that the samples x' are different from the input samples x, which are reconstructed by superposing and adding spatially adjacent reconstructed vectors $x'(2N \times 1)$ with an overlap of $N/2$ samples. The transform basis $L_0(2N \times N)$ is [1],[2]:

$$L_0 = LZP, \tag{3}$$

$$L = \frac{1}{2} \begin{bmatrix} D_e - D_o & D_e - D_o \\ J(D_e - D_o) & -J(D_e - D_o) \end{bmatrix}. \tag{4}$$

where J is the $(N \times N)$ counter identity matrix, D_e are the $(N \times N/2)$ even vectors from the DCT basis, D_o are the $(N \times N/2)$ odd vectors from the DCT basis, Z is an $(N \times N)$ orthogonal matrix and P is a permutation matrix.

3. OVERLAPPED MOTION COMPENSATION

Windowed Motion Compensation (WMC) [4] has been used to reduce discontinuities at the block boundaries due to motion compensation (MC) in predicted images. WMC reduces the prediction error in the image at MC block boundaries, thus increasing the coding efficiency. As opposed to MC, where the predicted image is assumed to be formed by (possibly shifted) blocks from the previous image, WMC assumes that the predicted image is formed by a superposition of windowed blocks from the previous image.

We now show how the concepts used in the LOT allow us to extend the ideas in windowed motion compensation to the more general case of linear filtering for overlapped data. The LOT and ILOT described in the previous section can be treated as linear filtering on the image data because the span of the LOT basis is twice that of the original data length. From Eq.(1) and (2), the inverse transformed data x' can be represented as follows:

$$x' = L_0 L_0^T x, \tag{5}$$

It can be shown [2] that

$$LL^T = \begin{bmatrix} G_u & 0 \\ 0 & G_l \end{bmatrix}. \tag{6}$$

where

$$G_u = \frac{1}{2}\,(I - D_e D_o^T - D_o D_e^T), \tag{7}$$

and

$$G_l = \frac{1}{2}\,(I + D_e D_o^T + D_o D_e^T). \tag{8}$$

The transfer matrix $L_0 L_0^T$ (Eq.(5)) does not give a windowing operation, but can be regarded as a filtering operation on the input data. If the input blocks are shifted blocks from the previous frame, with the shift being computed using block matching techniques, we can regard the above operations as Overlapped Motion Compensation (OMC). As in WMC, the sub-matrices G_u and G_l of the transfer matrix need to satisfy the following equations to reconstruct the overlapped area properly in neighboring blocks that have the same motion vectors:

$$G_u + G_l = I, \tag{9}$$

$$G_u = J G_l J. \tag{10}$$

Equation (9) is directly derived from the perfect reconstruction condition for FIR filter banks [5], and Eq.(10) states that each basis vector of the LOT is even or odd symmetric around its center. These equations can be regarded as generalized conditions for the OMC in the image domain.

The transfer matrix $L_0 L_0^T$ emulates a symmetric window operation to certain inputs such as flat data having a constant value in the block. In general cases, the superimposed transfer ratio with the neighboring block is not one. However, the advantage of overlapped motion compensation by LOT (OMC-LOT) can be seen around the block boundaries since a smooth mixture is obtained by superimposing outputs of OMC-LOT from two blocks in different positions. This effect is the same as WMC because of the overlapped blocking.

3.1 Adaptive Loop Filtering by Coefficient Truncation

Video sequences typically contain camera noise which changes over time because of changes in zoom, contrast, lighting etc. and appears as random texture changes in the picture. This results in noise in the high frequency transform coefficients and causes the inter-frame correlation between these coefficients to be low. A low-pass filter in the prediction loop is often used to decrease this noise and avoid using bits to encode it unnecessarily. OMC-LOT can implement an adaptive loop filter in the coefficient domain by truncating transformed coefficients. This adaptive filter can avoid unnecessary interframe prediction in the high frequency area. The loop filter can be implemented by inserting a truncation matrix M_m in the transfer matrix to truncate the $N - m$ high frequency coefficients, i.e.,

$$\hat{y} = L_0 M_m L_0^T x \tag{11}$$

where

$$M_m = \begin{bmatrix} I_m & 0 \\ 0 & 0 \end{bmatrix} \tag{12}$$

and I_m is an identity matrix of order m. The truncation point m can be chosen to minimize the camera noise without removing too much detail from the image. The coefficient truncation has to be performed at both the encoder and the decoder. Thus, if m is changed during coding, it has to be transmitted in order to be used at the decoder.

The OMC described above is based on the basic LOT. However, any other analysis-synthesis filter banks which satisfy the perfect reconstruction conditions for FIR filters can be used for OMC. For instance, the fast LOT [2] is another such class of analysis-synthesis filter banks. For the fast LOT, we can simply replace Z in Eq. (3) by

$$Z = \begin{bmatrix} I & 0 \\ 0 & C^{II}_{N/2}\ /\ S^{IV}_{N/2} \end{bmatrix} \tag{13}$$

where $C^{II}_{N/2}$ and $S^{IV}_{N/2}$ are the DCT-II and DST-IV matrices [2]. Eq.(13) is an orthogonal matrix and can be used for OMC since the relation in Eq. (5) is maintained. In general, OMC can be realized by perfect reconstruction filter banks [5] which are a generalized case of orthogonal transforms.

4. INTERFRAME VIDEO CODING BASED ON OVERLAPPED MACROBLOCKS

4.1 Overlapped Macroblocks in The Image Domain

Interframe coders typically code the picture in small blocks called macroblocks. Motion is computed on a macroblock basis and the difference signal between the current and predicted macroblock is coded using one or more DCT blocks. In case it is more efficient to code the input signal directly, the macroblock is intra coded and the receiver is also signaled to use intra values for decoding. This is possible because the data in each macroblock is independent of the neighboring blocks. If WMC is used with DCT or if the LOT is used with normal motion compensation, it becomes difficult to do this because of data overlap in the macroblocks. However, OMC-LOT followed by LOT can implement overlapped macroblocks that avoid this problem.

We now consider overlapped macroblocks that are formed in the image domain. If y is the input image, $\hat{y}$ is the predicted image and e is the error image, the LOT coefficients of the error image are given by

$$f_e = L_0^T e, \tag{14}$$

where

$$e = y - \hat{y}, \tag{15}$$

and $\hat{y}$ is obtained from the frame memory by OMC-LOT:

$$\hat{y} = L_0 M_m L_0^T x(s). \tag{16}$$

where $x(s)$ is data from the frame memory shifted by the motion vector s. Two LOTs and one ILOT are required in the prediction loop. The same number of operations are needed at the decoder. At the decoder, the decoded image $Q(y)$ is computed by adding the decoded (quantized) error signal $Q(e)$ to the predicted image $\hat{y}$:

$$Q(y) = \hat{y} + Q(e), \tag{17}$$

where

$$Q(e) = L_0(Q(f_e)). \tag{18}$$

The intra/inter decision is performed by comparing the prediction error e (Eq.(15)) and the current image y. For the intra coding at the encoder, intra LOT coefficients f_y are obtained by

$$f_y = L_0^T y, \tag{19}$$

and quantized. However, at the decoder, extra processing is required to take into account the overlapped area of the intra block and the decoder has to compute

$$Q(y) = L_0(Q(f_y)) + x(0) - L_0 L_0^T x(0) \tag{20}$$

in order to reconstruct the image.

4.2 Overlapped Macroblocks in The Coefficient Domain

The computation in the prediction loops as well as the extra processing for intra blocks at the decoder can be reduced if we work in the coefficient domain. Consider the case when the data of the overlapped macroblocks is represented by its LOT coefficients. Substituting Eqs.(15),(16) into (14) and using (19), we get

$$f_e = f_y - f_{x(s)}, \tag{21}$$

where

$$f_{x(s)} = M_m L_0^T x(s). \tag{22}$$

This reduces the computation in the prediction loop to only two LOTs. The decoder loop also works in the coefficient domain with the quantized coefficients $Q(f_e)$. Substituting (16) and (18) into (17), we obtain

$$Q(y) = L_0 \left(Q(f_e) + M_m L_0^T x(s) \right). \tag{23}$$

The computation complexity in this loop is also reduced by one ILOT.

Note that the block overlap is only in the image domain; there is no overlap in the coefficients representing each block. Thus, the intra/inter decision can be made by comparing the current image coefficients f_y and the prediction error coefficients f_e and quantizing and sending the coefficients which result in higher efficiency. At the decoder, the received image can simply be computed in the intra mode as

$$Q(y) = L_0(Q(f_y)), \tag{24}$$

and superimposed with the data of the inter-coded blocks. Intra-block processing in the coefficient domain results in a much simpler procedure. The complete coder/decoder block diagram using the coefficient domain macroblocks is shown in Fig. 1.

5. FEATURES AND EXTENSIONS

5.1 Two-Layer Coding

The ATM network is expected to allow several transmission classes in terms of cost, priority and expected packet loss performance. Layered coding can take advantage of this structure by using the low priority class (with higher packet loss) to transmit part of the data. The simplest layered coding model is a two-layer model which uses one high priority and one low priority channel for transmission. At the encoder, a truncation matrix similar to that used in loop filtering in OMC (section 3) is used to separate the LOT coefficients to be quantized and transmitted. The lower order coefficients are used in the prediction loop and are transmitted over the high priority channel. The remaining coefficients (those truncated by the truncation matrix) are coded in intra-mode and transmitted over the low priority channel. Loss of coefficients in the low priority channel does not degrade the reconstructed picture seriously.

5.2 Motion Estimation for Overlapped Macroblock

Conventional motion estimation uses a block matching method. The block matching method is a comparison of data in rectangular block of two images. On the other hand, OMC-LOT needs data outside of the conventional motion estimation block. Thus the estimation criterion for OMC blocks should be different from the one for rectangular blocks. The optimum criterion can be chosen to match the one used for the adaptive prediction, although exhaustive calculation is needed for locating the suitable motion vector for an overlapped block.

5.3 Varying Resolution using OMC-LOT

As with other subband or transform coding techniques, resolution of the decoded image can easily be changed by a factor of 2 by decimating the higher order coefficients and using a half length transform to reconstruct the picture. This method avoids additional filtering to change the picture size. Thus, M/N decimated images can be derived by performing M ($M = N/2^i$, i:integer) point inverse LOT over the low order M coefficients. Discarding high order LOT coefficient is equivalent to low pass filtering. The frequency characteristics of the decimated LOT have a sharp transition around the cut off frequency compared with N-tap FIR filters and its characteristic is much better than DCT [3]. Thus, decimation in LOT coefficient domain causes much less aliasing than coefficient truncation in the DCT.

6. CONCLUSION

A new interframe coding scheme that is expected to allow better coding than DCT based coders, especially at low rates is proposed in this paper. The algorithm uses two new concepts: a) overlapped macroblocks in the coefficient domain, and b) overlapped motion compensation. Encoding and decoding schemes exploiting OMC-LOT and LOT in which predicted image is created in the coefficient domain are presented. The algorithm can be adapted to include loop filtering by truncating LOT coefficients. In addition, the coefficients can be layered for robust transmission in presence of packet loss over ATM. Picture resolution can be varied easily at the decoder by inverse transforming using half length bases, since the inverse transforming truncated coefficients is equivalent to decimation using a sharp subsampling filter. Future work will examine the performance of this algorithm at different bit rates and compare it to DCT based coders.

REFERENCES

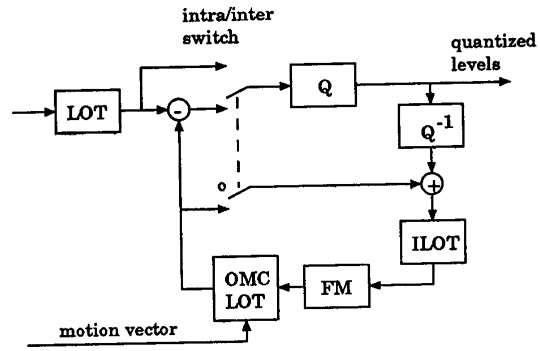
[1] H.S.Malvar and D.H.Staelin: "The LOT: Transform coding without blocking effects", IEEE Trans., vol.ASSP-37, No.4, pp.553-559 (Apr.1989).

[2] H.S.Malvar: "Lapped transforms for efficient transform/subband coding", IEEE Trans., vol.ASSP-38, No.6, pp.969-978 (June 1990).

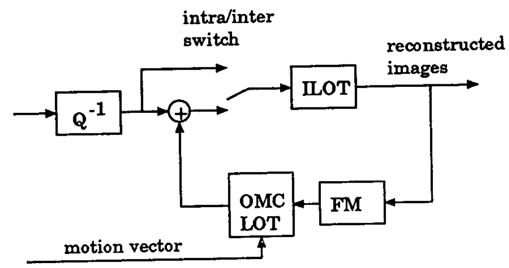
[3] H.Jozawa and H.Watanabe: "Intrafield/Interframe adaptive lapped transform for compatible HDTV coding", 1991 HDTV Workshop 2B-4 (Sep.1991).

[4] H.Watanabe and S.Singhal: "Windowed Motion Compensation", SPIE Vol.1605 Visual Communications and Image Processing '91, pp.582-589 (Nov.1991).

[5] M.Vetterli and D.LeGall: "Perfect reconstruction FIR filter banks: Some properties and factorizations", IEEE Trans., vol.ASSP-37, No.7, pp.1057-1071 (July 1989).



(a) Encoder block diagram



(b) Decoder block diagram

Figure 1. Encoder and decoder scheme by OMC+LOT in the coefficient domain.