Title	Multiresolution motion compensation coding for video compression
Author(s)	Choi, KT; Chan, SC; Ng, TS
Citation	International Conference On Signal Processing Proceedings, Icsp, 1996, v. 2, p. 1059-1061
Issued Date	1996
URL	http://hdl.handle.net/10722/46002
Rights	Creative Commons: Attribution 3.0 Hong Kong License

Multiresolution Motion Compensation Coding for Video Compression

K.T.Choi, S.C.Chan, and T.S.Ng,

The Department of Electrical and Electronics Engineering,
University of Hong Kong
Pokfulam Road, Hong Kong
Email: scchan@hkueee.hku.hk

Abstract

Motion estimation using block matching algorithm (BMA) is frequently used to reduced the temporal redundancy found in video coding. The performance of the mean squared error (MSE) is usually better than the mean absolute difference (MAD). However, the computational complexity is higher due to the squaring operations. In this paper, we propose a multiple candidates approach to shorten the gap between the two methods with slight increase in computational complexity. Application of the proposed algorithm to subband motion-estimation is studied and simulation results demonstrate that great improvement in performance can be obtained with similar computational complexity as the MAD measure.

1. Introduction

Motion estimation/compensation is an effective method for reducing temporal redundancy found in video sequence compression. Due to its simplicity, block matching algorithm (BMA) is frequently used to estimate the motion of a given block to be encoded. In BMA, all pixels in a block are assumed to undergo a translational motion. An estimate of the current block can therefore be obtained by searching similar blocks in the previous encoded (or original image) frame around the current block. In most video coding standards like H.261 and MPEG, the block size for motion estimation is (16x16). A straight forward but computational very expensive mean is to use the full search algorithm (FS) which estimates motion vectors by searching all the points in the search area. A candidate block in the search frame is shifted along all the positions within the search area to find the best match to the reference block of the current frame.

In general, the full search using mean squared error (MSE) as error criterion performs better than using mean absolute difference (MAD). However, the computational complexity using MSE is much larger due to the squaring operations. In this paper, we propose a multiple candidates approach to shorten the gap between the MAD and the MSE methods with a slight increase in computational complexity. We shall also investigate the use of the algorithm to subband motion estimation using different approaches.

The paper is organized as follows: In Section 2, the proposed method for motion estimation is presented. The application of the proposed algorithm to subband motion estimation will be discussed in Section 3. The simulation results of different algorithms and the analysis of their performance are given in Section 4. Finally, the conclusion is presented in Section 5.

2. Proposed algorithm

In a hybrid video coder based on the traditional motion compensation scheme, motion estimation is performed by block matching between blocks in the original frame and in the previously reconstructed frame. There are a number of matching criterion for BMA but the mean squared error (MSE) and mean absolute difference (MAD) are commonly used. The MSE criterion is defined as follows:-

$$MSE(x,y) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\left(f(x,y) - f(x+i,y+j)\right)^2 \quad (1)$$

0-7803-2912-0

While the MAD criterion is defined as follows:-

$$MAD(x,y) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} |f(x,y) - f'(x+i, y+j)| \qquad (2)$$

where MxN is the block size, f(x,y) is the reference block of the current frame and f'(x+i, y+i) is the candidate block within the search area. Motion vector is the value (i,j) which results in the minimum value of distortion measure within the search area. The PSNR performance of the MSE measure is better than the MAD measure, but its computational complexity is much higher due to the squaring operations. Therefore, in order to reduce the computational complexity of the encoding system, the MAD criterion is usually used in motion estimation.

Our algorithm is based on the following observation: Though the best MAD candidate cannot achieve the best MSE result, the probability of getting a much better MSE result by examining the best k MAD candidates is very high. Therefore, our algorithm is divided into 2 stages. In the first stage, we compute the MAD for all candidate blocks as usual. At the same time, we keep a small number of candidates with the best MAD score. In the second stage, we calculate the MSE for these candidates and select the best one as the motion vector.

In summary, our algorithm involves the following steps:-

(1) Performs MAD calculation on all motion vector candidates within the search window.
(2) Select the best n candidates.
(3) Find the best candidate with the minimum MSE.

The computation complexity of the proposed method is close to that of traditional MAD method. The additional complexity depends on the number of candidates that are being kept in (2). Computer simulation shows that only small number of candidates is required to achieve a performance close to that of the MSE method.

3. Application to Subband Motion Estimation

There are in general three approaches in applying analysis-synthesis systems to video coding. In 3D subband coding [1,2], the video is decomposed into spatio-temporal subbands. The motion and spatial details are encoded by allocating bits to the appropriate subbands. This can be viewed as a generalization of subband coding for still images to three dimension. 3D subband coding has the advantage of low implementation complexity because no motion compensation is required. However, larger storage buffer is needed. The second approach is to replace the DCT in a motion-compensation transform coder [3]. The prediction residual after motion compensation is passed through the analysis-synthesis system instead of the DCT. Due to low correlation and blocky nature of the prediction residual, the performance of such type of coder is limited. The third approach is to perform the motion compensation in the subband domain [4,5]. Each image frame is first decomposed into spatial subbands. Motion estimation is then performed on each subband to produce the prediction residual for coding.

Here, we shall consider two different methods to perform motion estimation in the subband. The first scheme is proposed by Fuldseth and Ramstad [5], in which motion estimation is performed on the undecimated subband signals. Since the current block to be encoded is decimated, each candidate block is also decimated before calculating the error criterion with the down-sampled original block. All the subband samples are involved in determining the motion vectors. Here, the motion vector is found from motion estimation using the low-low subband only. The same motion vector is then applied to other frequency bands.

In the second scheme, the subbands are interpolated to full resolution using bilinear interpolation to perform motion estimation. In both cases, only the low-low band of previous frame is used for finding the motion vector, and all other subbands will use the same motion vectors for reconstruction.

4. Experiments and Results

We combine our proposed algorithm for motion estimation with the two subband coding schemes. We perform one level of subband decomposition using the 9/7 wavelet filter. The block size is (8x8) in the decimated domain and the maximum search range in both the vertical and horizontal directions is ±8 pels. The Carphone sequence at QCIF format (176x144pixels) is used for testing.. A total of 150 frames was tested and motion estimation is performed on the luminance component only. Figure 2 show the MSE comparison of different algorithms. It can be seen that scheme 2 outperforms scheme 1 in terms of MSE (see Fig. 2). However, it is found that motion compensated pictures from scheme 1 is much sharper as compared with that of scheme 2 (see Fig. 1) .

We also compare the proposed motion estimation method with the algorithm using MSE and the one using MAD as error criterion. The performance of our method is very close to that of the MSE method, while the prediction error using MAD criterion is higher than the others. Figure 3 shows the MSE results of our proposed algorithms with different number of candidates. It can be observed that the performance of our method improves as the number of candidates increase. This shows the robustness of our algorithm. It can also be seen that 2 to 4 candidates are sufficient for most applications with only slight increase in computational complexity.

5. Conclusion

In this paper, we have presented a multiple candidates approach to improve the performance of using the MAD measure. Simulation results on two multiresolution motion estimation scheme demonstrated that similar performance as the MSE measure can be achieved with only slight increase in computational complexity. It is also found that 2 to 4 candidates are sufficient for most applications.

References

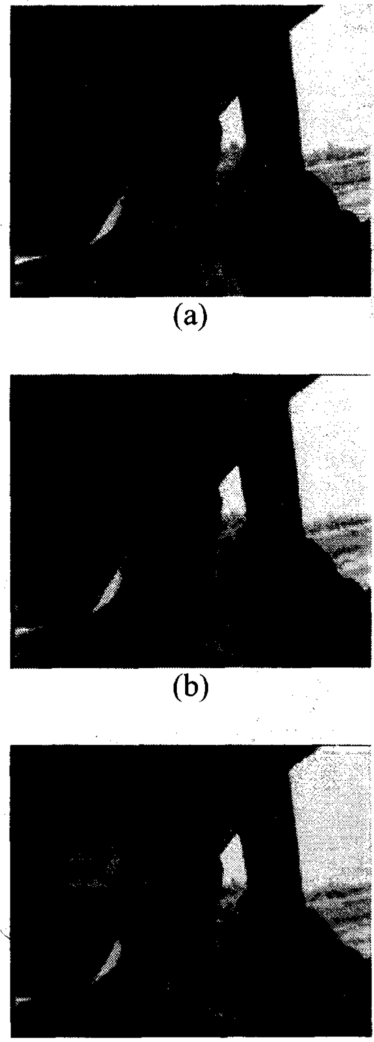
[1] C.T. Podilchuk, N.S. Jayant, and Farvardin, " Three-dimensional subband coding of video," IEEE Trans Image Processing, vol. 4, pp. 125 - 139, Feb. 1995.

[2] G.Karlsson and M.Vetterli, "Three dimensional subband coding of video," in Proc. ICASSP, 1988.

[3] M.Ohta and S.Nogaki, "Hybrid picture coding with wavelet transform and overlapped motion-compensation interframe prediction coding," IEEE Trans. Signal Processing, vol. 41, pp. 3416 - 3424, Dec. 1993.

[4] Y.Q.Zhang, and S.Zafar, "Motion-compensated wavelet transform coding for color video compression," IEEE Trans. Circuit and systems for Video Technology, vol. 2, pp 285 - 296, Sept. 1992.

[5] T.A.Ramstad and A.Fuldsethm, "Subband Video Coding with Smooth Motion Compensation" Proc. ICASSP, May 7-9, 1996.



(a)

(b)

(c)

Figure. 1 Frame 124 of Carphone for (a) Original, (b) Undecimated Method, and (c) Bilinear Method

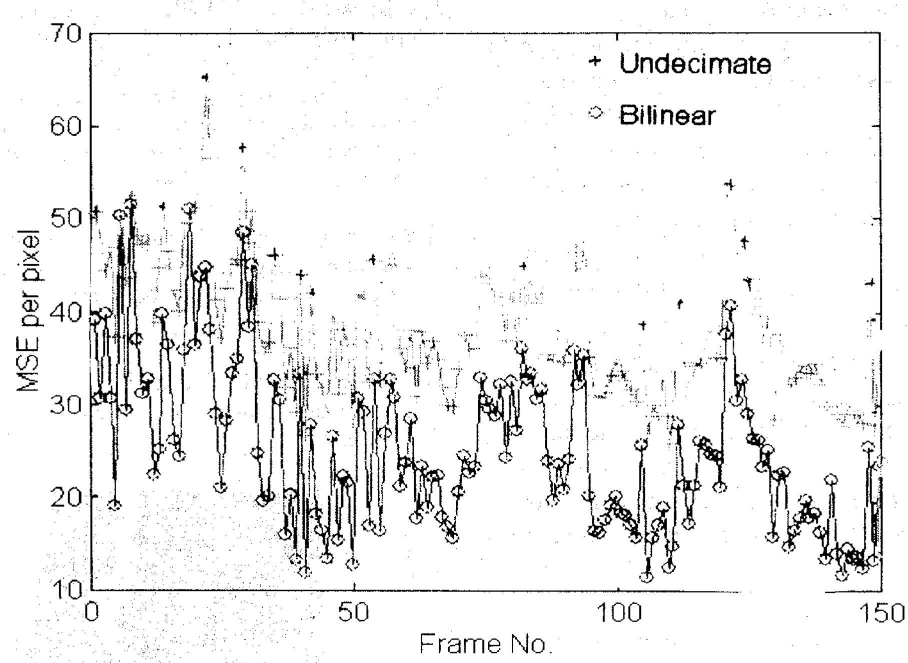


Figure 2. MSE per pixel for Carphone sequence (176x144) using Undecimated Method and Bilinear Method.

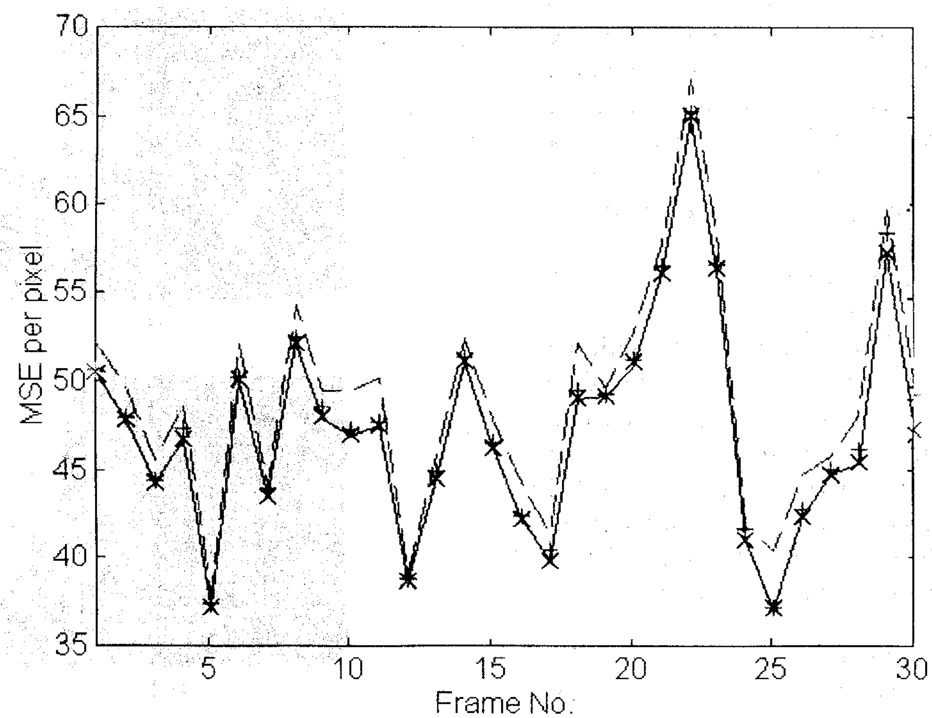


Figure 3. MSE per pixel for Carphone sequence (176x144) using MAD method (dash line), MSE method (solid line), and Proposed method with 2 candidates (symbol +) and 4 candidates (symbol x).