SMOOTH CONSTRAINED BLOCK MATCHING CRITERION FOR MOTION ESTIMATION

Xuan Jing, Ce Zhu and Lap-Pui Chau

School of Electrical and Electronic Engineering
Nanyang Technological University, Singapore

ABSTRACT

In this paper, a novel and efficient criterion for block matching motion estimation is presented. The proposed criterion is to enhance the conventional Mean Absolute Difference (MAD) scheme with a new smoothness constraint on the residue block. The objective is to reduce the bit rate for encoding the residue image without any degradation of the reconstructed image quality. Simulation results show that by applying the new criterion in motion estimation, both the improvement in Peak Signal-to-Noise Ratio (PSNR) and the reduction in bit rate up to 4.3% can be achieved compared to MAD.

1. INTRODUCTION

Video coding techniques take advantage of data redundancies to reduce the storage or bandwidth needed to represent the visual information. Temporal correlation among consecutive frames is significantly high due to the limited amount of motion. In order to reduce this temporal redundancy, motion estimation has been widely used in various video coding systems. The overall performance of a real-time video coding system depends heavily on the accuracy and efficiency of the motion estimation.

One of the most popular motion estimation approaches is block matching algorithm (BMA) [1]. In the BMA, a frame is firstly divided into blocks of size $N \times N$. Then the basic operation of block matching is to pick up a candidate block in a predefined search range by calculating the matching function between the candidate and the current block. This procedure is repeated until all the candidates have been checked and finally the best match one is selected. A very important part of a MBA is the matching criterion which is the measure of similarity between the current block and the candidate block. The selection of the matching function has a direct impact on the computational complexity and the coding efficiency. So far several popular matching criteria [2-5] that appear frequently in the literature are the Normalized Cross-Correlation Function (NCF), Mean Squared Error (MSE), Mean Absolute Difference (MAD), Pel Difference Classification (PDC) and the Minimized Maximum Error (MiniMax). For the above matching criteria, the NCF and MSE are the most complex because of the intensive computation of multiplications involved. In spite of their accuracy, their hardware realization seems to be far from being feasible. On the other hand, although PDC requires less hardware complexity than MAD, its performance is quite sensitive to the threshold chosen. The less requirement of computation also affects the performance of the MiniMax. Therefore, MAD seems to be the most popular choice in designing practical video coding systems because of its accuracy and simple operations required. However, the MAD criterion is just the average pel distortion without considering the overall smoothness of the residue block. As a result, when conducting motion estimation using MAD, the block which can minimize the distortion is selected without considering how many bits are required to code the residue. In contrast, the following transform coding tries to code the residue using minimum number of bits. Taking this point into consideration, we propose a novel block matching criterion which aims at eliminating the mismatch between motion estimation and the following transform coding. Simulation results show that our proposed criterion performs better than the traditional criteria such as MAD and MiniMax in terms of PSNR and bit rate with comparable computational complexity.

The rest of the paper is organized as follows. In Section 2 we present the proposed matching criterion and its implementation. Section 3 demonstrates the simulation results for performance comparison. Finally, conclusions are given in Section 4.

2. SMOOTH CONSTRAINED MAD (SC-MAD) CRITERION

In motion estimation and compensation, the predicted frame is constructed after predictions for all the Macro-Blocks (MB) in the current frame are obtained. Then the residue frame is calculated and coded using the transform coding technique. According to the characteristics of

Discrete Cosine Transform (DCT), coding a flat error block which has the same value for all the elements within the block only needs a very few bits for the DC coefficient. The remaining AC coefficients should be all zeros thus there is no need for coding them. For those non-flat residue blocks it is clear that the smoother the block is the fewer bits may be required to code the residue. In view of this fact, a good match between the current block and the candidate block requires not only minimizing the total distortions between the two blocks, the smoothness of the residue block is also highly desirable. Usually we can calculate the variance of the residue block to determine its smoothness. However, to calculate variance requires so many computations that it is not suitable for real-time implementation. Instead, we propose another simple factor as the smoothness measurement, i.e. the difference between the maximum residue and the minimum residue within the residue block, denoted as MMD. Obviously, we can reasonably assume that the smoother residue block normally has smaller MMD because all its components will be centralized in a smaller range. In the following we propose a Smooth Constrained MAD (SC-MAD) criterion based on our assumption for choosing the candidate block to improve the coding efficiency of transform coding following the motion estimation.

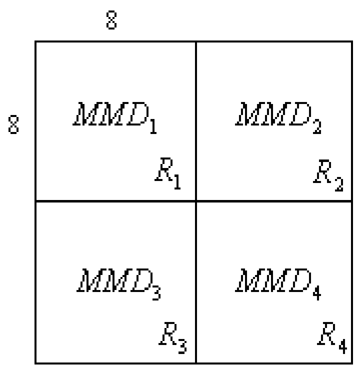


Fig.1 Four 8×8 residue blocks used in the SC-MAD

In the proposed algorithm, we first subtract the current MB, X (16×16) by the candidate MB, Y (16×16) and form the residue MB, R (16 × 16) where $r_{i,j} = x_{i,j} - y_{i,j}$ $i, j = 1...16$. Then we compute the conventional MAD. After that, we subdivide the residue MB, R (16×16) into four Small-Blocks (R_1 to R_4) with size 8×8 and determine the MMDs of the four residue R_m separately. This procedure is depicted in Figure 1. As defined above, for R_m,

$$MMD_m = r_{\max}^m - r_{\min}^m \tag{1}$$

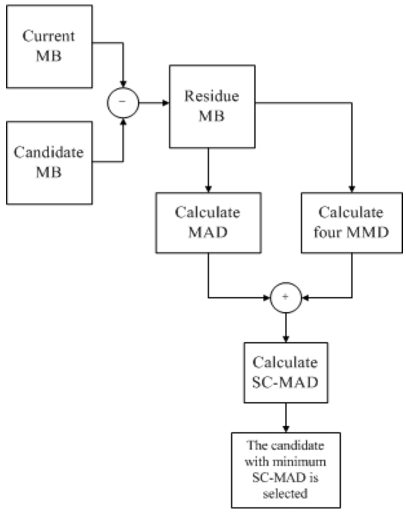


Fig 2. Flow chart of the proposed SC-MAD

where $r_{\max}^m$ and $r_{\min}^m$ are the maximum residue and the minimum residue within a residue block R_m. The reason behind dividing the 16×16 residue MB into four 8×8 blocks for MMD calculation is as follows. In video compression standards, the motion vector is searched on a MB (16×16) basis while the prediction error or residue is coded using a block (8×8) based DCT transform. Thus, it is better to consider the smoothness of the subdivided residue blocks separately and take the sum of those four MMDs as the overall smoothness measurement for the resulting residue MB. Furthermore, the size of the subdivided residue blocks can be adapted to the type of DCT transform used. For instance, if we choose 4×4 based DCT transform, we may divide the residue MB into 4×4 sub-blocks accordingly.

Finally, the proposed distortion measure namely, Smooth Constrained MAD (SC-MAD), is defined as follows:

$$SC\text{-}MAD = MAD + \alpha \times \sum_{m=1}^{4} MMD_m \tag{2}$$

where α is a weighting factor and

$$MAD = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} | x_{i,j} - y_{i,j} | \tag{3}$$

The candidate with the smallest SC-MAD is to be chosen as the best match block. The steps for finding the best match block using this new block matching criterion (SC-MAD) are summarized below:

1. Subtract a candidate MB from the current MB and form the residue MB.
2. Calculate the MAD for the candidate.

3. Subdivide the residue MB into four small residue blocks (R_1 to R_4) and determine the MMD_m based on the residue obtained in step one.

4. Calculate the final distortion measure SC-MAD of the candidate using Eq.(2).

5. Select the candidate block which has the minimum distortion SC-MAD.

The flow chart of the algorithm is depicted in Figure 2. In addition, the computation requirement of the proposed SC-MAD is comparable to MAD only with a slightly increase of computation such as comparisons involved in finding the MMD_m .

3. SIMULATION RESULTS

To test the performance of the proposed criterion, we compare it with those conventional criteria such as MAD and MiniMax. Four video sequences (*Foreman* CIF 352×288, *Tennis* CIF 352×288, *Mother and Daughter* CIF 352×288, *Miss America* QCIF 176×144) were used in the simulation. Among those test videos *Foreman* and *Tennis* have relatively fast motion while the other two have moderate movements. We conducted the simulation based on a H.261 [6] codec framework. In motion estimation the block size was fixed to 16×16 and the search strategy was Full Search (FS) with a maximum displacement of $\pm$ 15 in both directions. A fixed quantization parameter of 15 was adopted in order not to affect the results with rate control operations. The frame rate was 30 fps and the frame coding structure was I frame for the first frame followed by consecutive P frames without frame skipping. In P frames, all MBs were forced to be inter-coded in order to ensure a fair comparison by using same number of inter-coded MBs for different criteria.

Performances of different criteria are depicted in Table 1. in terms of average PSNR of luminance component and the number of bits needed for coding the inter blocks (Error or Residue Bits). From the table it is clearly shown that there has been a significant improvement in all sequences by taking SC-MAD as a new matching criterion. We obtained savings in bit rate of around 2~4% over using the MAD. Especially for *Foreman*, the error bits were reduced by 4.3%. In addition, by using the new criterion we achieved the improvement in image quality for all four testing videos from 0.04 dB to 0.13 dB compared to MAD. Figure 3 and Figure 4 demonstrate the frame by frame comparisons of different criteria for *Foreman* and *Mother &Daughter* sequences.

4. CONCLUSIONS

In this paper, we have presented a new motion estimation block matching criterion called Smooth Constrained MAD (SC-MAD) which enhances the conventional MAD function by adding a smoothness control for the residue block. The main advantage of our new scheme over MAD is that it can further distinguish candidate blocks when they have similar MADs by taking the smoothness constraint MMD_m into consideration. The simulation results have indicated that SC-MAD performs better than MAD in terms of reconstructed image quality and the number of bits needed for coding the residue and much better than the MiniMax criterion.

5. REFERENCES

[1] H. M. Musmann, P. Pirsch, and H. J. Gravoert, “Advances in picture coding,” *Proc. IEEE,* vol. 73, pp. 523-548, Apr. 1985.

[2] S. Kappagantula and K. R. Rao, “Motion compensated interframe image prediction,” *IEEE Trans. on Communication.,*vol. COM-33, no. 9, pp. 1011-1015, Sept. 1985.

[3] T. Koga, K. Iinuma, A. Hirano, Y. Iiyima and T. Ishi-guro, “Motion compensated interframe coding for video conferencing,” *Proc. NTC81*, pp. G5.3.5, 1981.

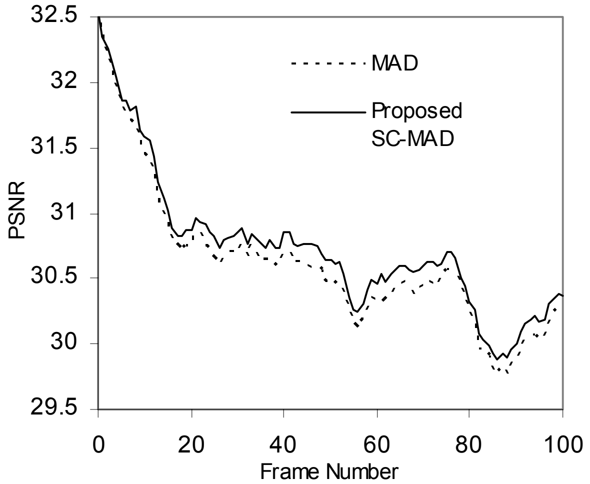
[4] H. Gharavi and M. Mills, “Block-matching motion estimation algorithm-New results,” *IEEE Trans. on Circuits and Systems*, vol. 37, no. 5, pp. 649-651, May 1990.

[5] L. G. Chen, M. J. Chen, and T. D. Chiueh, “A new block-matching criterion for motion estimation and its implementation,” *IEEE Trans. on Circuits and Systems for Video Technology.* vol. 5, no. 3, June 1995.

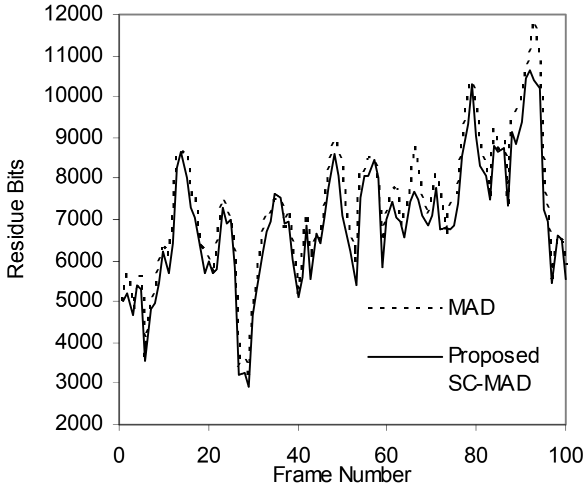
[6] ITU-T, “Video codec for audiovisual services at $p \times 64$ kbits,” ITU-T Recommendation H.261, Mar 1993.

Table 1. Performance comparisons for different criteria

Sequences		MAD	MiniMax	SC-MAD
Foreman	PSNR	30.59	29.77	30.72
	ErrorBits	725098	890644	695039
Tennis	PSNR	27.94	27.43	27.98
	ErrorBits	3173653	3730522	3120409
Mother & Daughter	PSNR	31.44	30.99	31.57
	ErrorBits	435618	538236	421031
Miss America	PSNR	34.69	34.06	34.79
	ErrorBits	51011	62974	50076

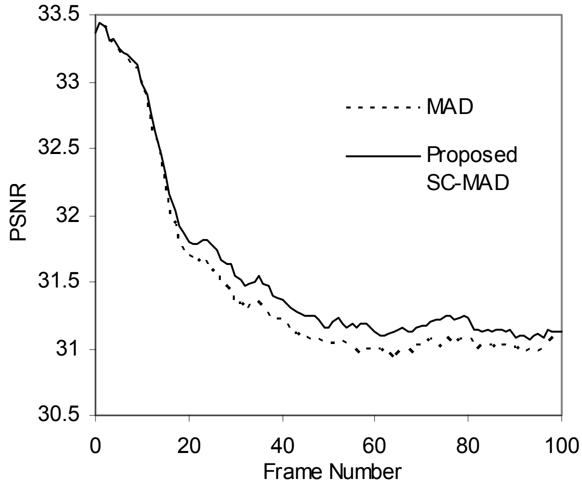


a) PSNR comparison

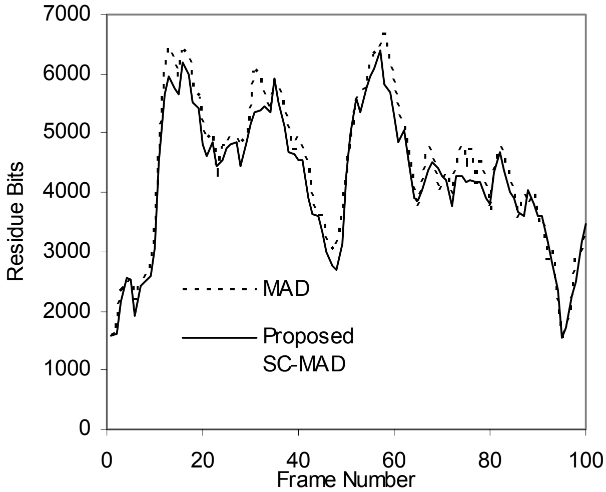


b) Bit rate comparison

Fig.2 Frame by frame comparisons for *Foreman* sequence (CIF 352×288 first 100 frames)



a) PSNR comparison



b) Bit rate comparison

Fig.3 Frame by frame comparisons for *Mother&Daughter* sequence (CIF 352×288 first 100 frames)