2000年度 第13回 信號處理合同學術大會 論文集 第13卷 1號

효율적인 비디오 부호화를 위한 영역 선택적 가변크기 움직임 예측

최 윤 정, 손 광 훈
연세대학교 전기컴퓨터공학과

New variable size motion estimation technique by selecting the active motion region

Yunjung Choi and Kwanghoon Sohn
Department of Electrical and Computer Engineering, Yonsei University

Abstract

In this letter, The proposed algorithm has processed the motion estimation and compensation only in the region which has the motion. By considering the temporal redundancies between the successive frames, the bit cost for representing the motion are reduced efficiently. The performance of them proposed technique show the superior result to the existing methods' in terms of both PSNR and coding cost.

Index terms: temporal redundancies, motion estimation and compensation, video coding

I. Introduction

The typical methods for reducing temporal redundancies of video signals are generally based on the Block Matching Algorithm (BMA). However, it is not possible to properly carry out in the case of blocks that include boundaries, since only a single motion vector is assigned to each block. As a result, several Variable size Block Matching Algorithms(VBMA) have been proposed to improve the performance of the motion compensated transform coding algorithms [1-2]. However, since they require additional bits to encode information on block sizes, the computational complexities are increased.

Recently, an object-based BMA has been 1proposed to reduce temporal redundancy. However, this technique requires segmentation and shape coding [3].

In this letter, a rate-distortion constrained region selective VBMA(RD-SVBMA) for efficient video coding is proposed for the reduction of temporal redundancies. If the current frame has only the local motions that are restricted to the region which contains the moving object, it becomes possible to efficiently reduce the computational complexity and coding cost in the quadtree splitting process by selecting only the active motion regions. In addition, since the several motions in the homogeneous region have negligible effects on quality, the homogeneous regions can be excluded from the active motion regions. The current frame that includes global motions (i.e. zooming, panning) has non-zero motions in the entire region. Thus, the active motion regions in global motion frames are selected less tightly. When a frame, either a global or a local motion frame is classified by analyzing the characteristics of the difference of successive frames, each subregion that satisfies a certain criteria is determined as an active motion region. Thus, it is possible to increase the coding efficiency and to decrease computational complexity by selectively encoding the active motion regions.

II. R-D constrained region selective VBMA(RD-SVBMA)

iii. The frame differences in video sequences are negligibly small except in boundaries and edges. Most regions with motions appear in the boundary between the objects or between the object and the background. Thus, the temporal redundancies can be efficiently reduced by region selecting methodology for motion estimation, and the coding rate and computational complexity as well. The preprocessing for selecting the active motion region consists of two stages. At the first stage, the characteristics of the frame are determined as to whether it is a global motion frame or a local motion frame by eq(1):

$$I_t = \begin{cases} I_{t,global} & m_t \geq T_M, v_t \geq T_V \\ I_{t,local} & , \text{ otherwise} \end{cases} \quad (1)$$

$$m_t = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} | I_t(i,j) - I_{t-1}(i,j) | \quad (1\text{-}1)$$

$$v_t = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (m_t - | I_t(i,j) - I_{t-1}(i,j) |)^2 \quad (1\text{-}2)$$

, where T_M and T_V are the empirically obtained threshold values, at the next stage, active motion regions are determined as follows:

$$R_{active}(k,l) = \{R(k,l) \mid m_{block}(k,l) \geq T_{active} \quad (2)$$

, where $R_{active}(k,l)$ represents a squared active motion region starting at (k,l).

$$m_{block}(k,l) = \frac{1}{N_r^2} \sum_{i=0}^{N_r-1} \sum_{j=0}^{N_r-1} | I_t(i+k, j+l) - I_{t-1}(i+k, j+l) | \quad (2\text{-}1)$$

, where $m_{t,block}(k,l)$ represents absolute difference mean of $N_r \times N_r$ square block starting at (k,l).

$$T_{active} = \alpha \, m_t \quad (2\text{-}2)$$

, where α is a weight factor for selecting the active motion regions with respect to the frame type. The smaller α is used for the $I_{t,global}$, rather than for the $I_{t,local}$ to select active motion regions less tightly (empirically α=0.35 in the $I_{t,global}$ and α=1 in the $I_{t,local}$). Once the active motion regions are classified, they are split into guadtree by $Nr/4 \times Nr/4$ according to the rate-distortion(R-D) constraint. In other words, a R-D constrained regularization technique is used as a splitting criteria [2][4]. We assume that inactive motion regions have the same pixel values as the previous frame. Thus, it is not necessary to encode them.

III. Experimental results

The computer simulation is performed for 'Clair' and 'Table Tennis' sequences. 'Clair' is a typical head-and-shoulder sequence with a CIF (352*288) format, which has the local motion and static background. 'Table Tennis' is a sequence that includes a zooming operation and a fast motion with a SIF (352*240) format. Fig.1 shows the active motion regions represented as white. Fig. 2 shows the corresponding quadtree splitting maps in the 5rh frame of the "clair" sequences which has been decided as a local motion frame. As shown in the figures, the active motion regions reduce the temporal redundancies of the successive frames. The performance of the proposed method is shown in Fig.3 and Fig.4. For the 'Clair' sequences, as shown in the Fig 4., the proposed algorithm is improved by 1.4dB and 0.9dB compared to the Fixed size Block Matching Algorithm(FBMA) and the RD_VBMA on the average at the 0.01bpp. For 'Table Tennis' sequences, it is improved by 2.4dB and 1.7dB compared to the FBMA and the RD_ VBMA on the average at 0.018bpp.

IV. Conclusion

In this letter, we report on an improvement in the existing RD_VBMA by efficiently reducing temporal redundancy. First, by analyzing the characteristics of the difference of successive frames, it is possible to classify the local motion frames and global motion frames and select the active motion regions in each frame. RD_VBMA is then selectively performed only in the active motion regions. The findings here confirm that it outperformed other existing algorithms in terms of both PSNR and coding cost.

Reference

1. Jong Won Kim and Sang Uk Lee, " Video Coding With R-D Constrained Hierarchical Variable Block Size Motion Estimation ", Journal of Visual Communication & Image Representation, V.9 N.3, September 1998
2. Johanna V. Gisladottir, Kannan Ramchandran, Michael Orchard, "Motion-based Representation of Video Sequences Using Variable Block Sizes", Visual Communications and Image Processing '96-SPIE Vol.2727 Part 1/3, March 1996
3. Michael K Steliaros, Graham R Martin, Roger A Packwood, "Composite Shape and Motion Coding for Arbitrary-Shaped Video Objects", Proceedings of the Picture Coding Symposium, pp. 161-164, April 1999
4. Hao Bi and Wai-Yip Chan, " Rate-Distortion Opimization of Hierachical Displacement Field", IEEE Trans. on Circuit and Systems for Video Technology, vol. 8, no. 1, pp. 18-24, Februry 1998

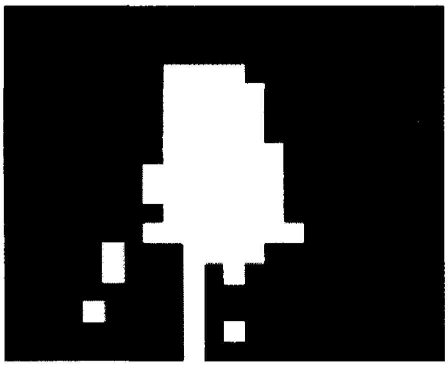

Fig. 1. Selected active motion region ("Clair" #26)

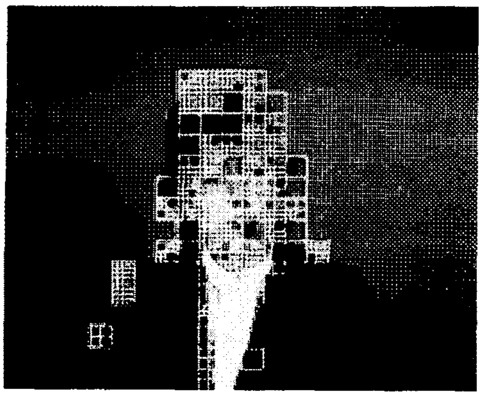

Fig. 2. Quadtree splitting map in the local motion frame ("Clair" #26)

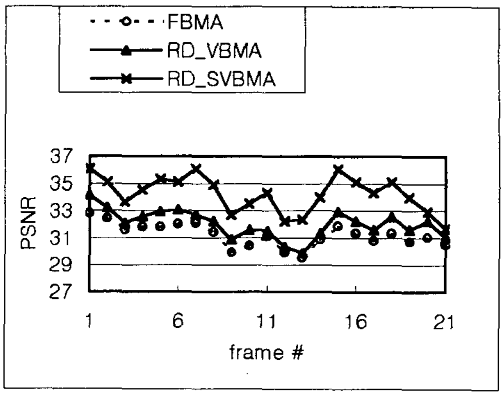

Fig.4 PSNR performance comparison between RD_SVBMA, RD_VBMA and FBMA ("Table tennis")

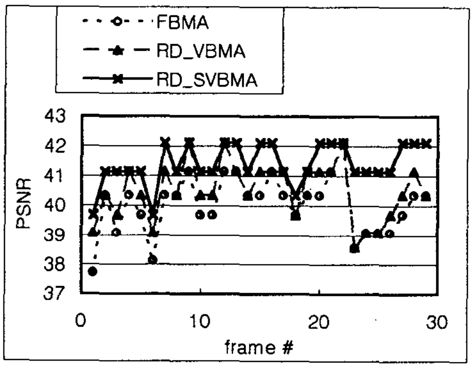

Fig.3 PSNR performance comparison between RD_SVBMA, RD_VBMA and FBMA ("Clair)