Adaptive Weighted Prediction in Video Coding

Yanfei Shen, Dongming Zhang, Chao Huang, Jintao Li
Institute of Computing Technology of
Chinese Academy of Sciences
{syf,dmzhang,chuang,jtli}@ict.ac.cn

Abstract

In this paper, an adaptive weighted prediction algorithm is proposed to improve the coding efficiency for video scenes that contain global brightness variations caused by fade in/out or local brightness variations caused by camera flashes. In this algorithm, two brightness-variation parameters model is used, which represent multiplier and offset components of the brightness variation. Brightness variations in frame and macroblock level are detected respectively, if there are brightness variations, we use weighted prediction to compensate these brightness variations; otherwise conventional prediction method is used. Simulation results show that the proposed algorithm can improve the coding efficiency sufficiently when the coded scenes contain brightness variations. This technology is adopted by AVS.

1. Introduction

For video coding, interframe prediction is an indispensable technique for the efficient coding of video data, and is used in many video coding standards which include ITU-T Recommendation H.261, H.263[1], MPEG-1, MPEG-2 and MPEG-4[2] etc. block matching technique generally uses criteria to minimize the mean squared or absolute error between two corresponding blocks. However these techniques assume that the brightness in video scenes remains unchanged, so the simple interframe prediction technique can not efficiently code video scenes containing temporal brightness variation caused by fade in/out or camera flashes etc.

As a result of these problems, many methods have been proposed to eliminate the effect of brightness variations, for example JVT presents global weighted prediction to improve coding efficiency for scenes containing fades in H.264/AVC[3], but it is inefficient for scenes containing local brightness variations. In [4], I.J. Cox proposed to use histogram matching to deal with the presence of brightness variations, which is based on the hypothesis that brightness variations are global to the image or at least that they affect pixels of equal grey scale value in the same manner. In practice, however, brightness variations maybe concentrated in different regions of the image and vary spatially. In [5], a set of parameters side information is needed to present brightness variations for every macroblock including multiplier and offset components. In fact it is not necessary for scenes containing global brightness variations and it is too complexity. In this paper, first we detect whether there is brightness variations between reference frame and current frame, if there is brightness variations, a set of parameters is calculated for the whole frame, some or all macroblocks need brightness compensation before motion compensation, otherwise conventional prediction method is used.

This paper is organized as follows. In Section 2 we present brightness variation detection method. We describe brightness variations model in Section 3. An adaptive weighted prediction architecture is proposed in section 4. In section 5, computer simulation results are shown for test sequences containing brightness variations. Finally, in section 6, conclusions are given.

2. Brightness variation detection

In an adaptive weighted prediction algorithm for video coding, brightness variation detection method plays an important role, which depends on the suitable choice of a similarity metric between two frames. Usually histograms show less sensitivity to the frame changes within a similar brightness condition, and histogram-based method is computationally efficient. In this paper, we use histogram difference as the criterion of brightness variations. The histogram difference is defined by

$$D_{HIS} = \sum_j \left| H_{t+1}(j) - H_t(j) \right| \quad (1)$$

0-7803-8603-5/04/$20.00 ©2004 IEEE

where $H_{t+1}(j)$ and $H_t(j)$ signify the histograms in the jth bin for reference frame and current frame separately. By thresholding the histogram difference D_{HIS}, the frames having brightness variations can be detected. The threshold value is selected experimentally.

3. Brightness variation model

In conventional video coding, motion estimation is part of the process to remove temporal redundancies. Block based schemes assume that each block of the current frame is obtained from the translation of some corresponding region in a reference frame. Motion estimation tries to identify this best matching region in the reference frame for every block in the current frame and this is shown in Fig 1.

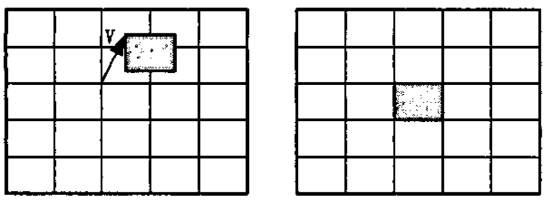


Fig.1. Illustration of motion estimation

In Fig.1, the gray block on the right corresponds to the current block being coded and the gray area on the left represents the best match found for the current block in the previous frame. In these schemes, no brightness variation is assumed, in practice, this assumption can not be guaranteed. If brightness variation happened between successive frames, the RD algorithm will select intra coding mode, which result in low coding efficiency. If we use a suitable model to simulate the brightness variation, it will reduce the impaction of coding efficiency brought by brightness variation.

In this paper brightness variation is modeled by

$$I(x,y,t) = a \cdot I(x+\Delta x, y+\Delta y, t-1) + b \quad (2)$$

which is based phase-correlation-based techniques in [6], where $I(x,y,t)$ is the brightness intensity of pixel point (x,y) at time t, a and b are the parameters corresponding to the multiplier and offset components of brightness variation model, $(\Delta x, \Delta y)$ is the motion vector of pixel point (x,y) in current frame.

We extend this model to block-matching motion estimation video coding, that is, we estimate brightness variation parameters a and b macroblock by macroblock. We assume the motion vector V of the current macroblock is represented by $(\Delta x, \Delta y)$, R is the search region used to estimate motion vector. Parameters a and b are calculated by minimizing the following evaluation function

$$E = \sum_{x,y \in R} (I(x,y,t) - (a \cdot I(x+\Delta x, y+\Delta y, t-1) + b))^2 \quad (3)$$

that is, we need to solve linear equation

$$a = \frac{N \cdot S - W \cdot Q}{N \cdot T - W^2} \quad (4)$$

$$b = \frac{T \cdot Q - W \cdot S}{N \cdot T - W^2} \quad (5)$$

where

$$W = \sum_{i=0}^{m} \sum_{j=0}^{n} I(i+x+\Delta x, j+y+\Delta y, t-1) \quad (6)$$

$$Q = \sum_{i=0}^{m} \sum_{j=0}^{n} I(i+x, j+y, t) \quad (7)$$

$$S = \sum_{i=0}^{m} \sum_{j=0}^{n} [I(i+x+\Delta x, j+y+\Delta y, t-1) \times I(i+x+\Delta x, j+y+\Delta y, t)] \quad (8)$$

$$T = \sum_{i=0}^{m} \sum_{j=0}^{n} I^2(i+x+\Delta x, j+y+\Delta y, t-1) \quad (9)$$

and N is the number of pixels in the macroblock, m and n is the width and height of the macroblock. If there are global brightness variations between current frame and reference frame, the set of parameters a and b are common to all macroblocks. In practice, the values diverge because of various local brightness variations or the infection of occasional noise. In order to determine the unique weighted prediction parameters for the whole frame, we use the frequency of the local brightness-variation parameters obtained for all macroblock blocks of the frame in the same way that[7] uses the frequency of the motion vectors. That is, the most common set of brightness-variations parameters a and b is chosen as the unique weighted prediction parameters. If the macroblock uses weighted prediction, the final prediction value is generated as

$$\hat{I}(x,y,t) = a \cdot I(x+\Delta x, y+\Delta y, t-1) + b \quad (10)$$

We assume the weighted prediction multiplier parameter a takes values from 1/8 to 8, whose quantization step is 1/32, and offset parameter b takes values from -127 to +127, whose quantization step is 1, so we can use 8 bits to represent parameters a and b.

4. Adaptive weighted prediction

We divide brightness variation into two cases, one brightness variation case is caused by post processing of video content, for example fade in/out of the whole frame. In this case, all macroblocks of the current frame need weighted prediction. The other brightness variation case is caused by camera flashes, camera-iris adjustment or focus lamp, and so on. In this case only partial macroblocks of the current frame need weighted prediction, we can select whether the macroblock uses weighted prediction based on the squared error sum produced. The flowchart of the proposed weighted prediction is show in fig.2

The refined steps are detailed using the following procedure:

1) Brightness variation detection using histogram difference between current frame and reference frame.
2) If there is brightness variation, variable BrightVar is set 1, then go to next step, otherwise variable BrightVar is set 0 and end the procedure.
3) Calculate brightness variation parameters set a and b for all macroblocks of the current frame.
4) If brightness variation parameters a and b of all macroblocks are common, variable AllMBWeight is set 1, that is, all macroblocks need weighted prediction. Otherwise variable AllMBWeight is set 0, that is, only partial macroblocks of the current frame need weighted prediction based on the squared error sum produced by them.
5) Calculate weighted prediction parameters based on the frequency of brightness-variation parameters a and b of all macroblocks.

Consequently Macroblocks are compensated based on the variables BrightVar, AllMBWeight, parameter a and b.

5. Experiments

We performed computer simulation to evaluate the performance of the proposed weighted prediction algorithm. Two CIF video sequences consisting of three frames are used, brightness of the second frame is different from the others. In scene 1, as shown in Fig.3 (a), an artificial digital fade-out effect was performed on standard test sequence forman, whose brightness variation can be sufficiently represented by (3). In scene 2, as shown in Fig.3 (b), it is cropped from 1280×720 test sequence crew containing local brightness variation caused by camera's flash.

Histograms for two scenes containing three frames are compared in Fig.4. Fig.4 (a) represents the histograms for scene 1, where frame 0 and frame 2 have similar histogram distribution, which is evidently different from frame 1. Although scene 2 only contain partial brightness, the histogram of frame 1 is different from the others. So histogram difference can detect brightness variation between neighbor frames. Moreover, calculation of histogram is very simple. In our simulation, the threshold is set quarter of image size, for example, the threshold is set quarter of 352×288 for CIF frame.

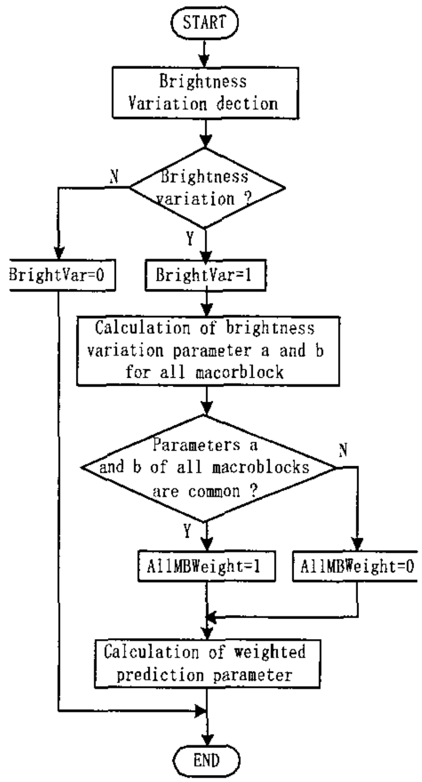
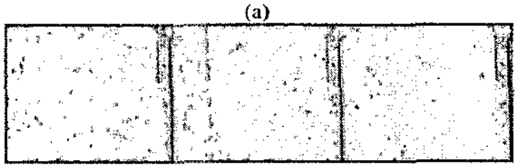


Fig.2. Flowchart of the proposed weighted prediction algorithm

We use AVS (Audio Video Coding Standard Working Group of China) reference software RM3.0[8] to validate our proposed algorithm, where QP is set 25 and the first frame is intra coding. In order to reduce the complexity of computing brightness variation parameters a and b for all the macroblocks of the current frame, instead of motion estimating, we directly compute the brightness variation parameters a and b of co-location macroblocks between reference frame and current frame, that is, Δx and Δy in (3) are set zero. Chroma components are processed in the same way as for luminance detailed above.

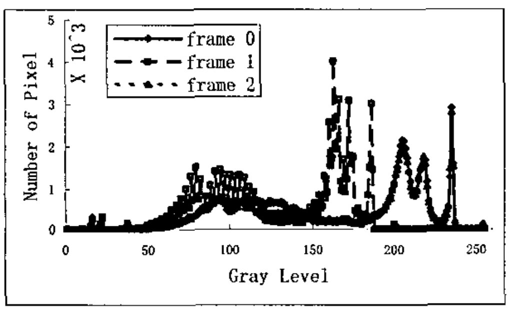
(a)



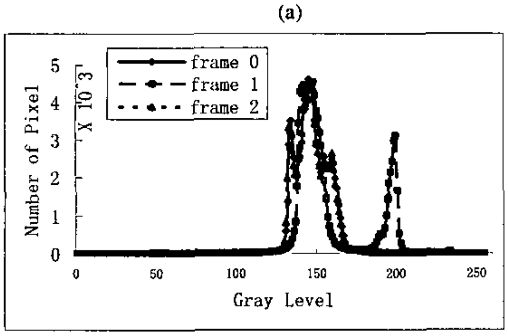
(b)

Fig.3. Test video sequences
(a) Three frames for scene 1
(b) Three frames for scene 2

Table 1 shows the performance of RM 3.0 with and without weighted prediction for two scenes containing brightness variations. In scene 1, our algorithm can estimate brightness parameter exactly as we used to perform fade-out effect, all macroblocks perform weighted prediction. In scene 2, only partial macroblocks perform weighted prediction. Simulation results listed in table 1 show that the proposed weighted prediction can yield higher PSNR value or lower bitrate compared with the method without weighted prediction.



(a)



(b)

Fig.4. Comparison of histograms for two scenes
(a) Histograms of scene 1 for three frames
(b) Histograms of scene 2 for three frames

Table.1. PSNR and Bit rates for two sequences

Method	PSNR(dB)		Bitrates(Bit/Pic)	
	Scene 1	Scene 2	Scene 1	Scene 2
With weighted prediction	39.85	41.99	146512	51144
Without weighted prediction	39.78	41.49	165400	95224

6. Conclusions

In this paper, we propose an adaptive weighted prediction algorithm for video coding that improves coding efficiency for video scenes containing global brightness variations caused by fade in/out effects or local brightness variations caused by camera-focus or flicker etc. Simulation results show that the proposed method can improve coding efficiency for video sequences containing brightness variations.

Further research will focus on the selection of weighted parameters which can exactly simulate the brightness variations.

7. References

[1] Draft ITU-T Recommendation H.263, "Video coding for low bit-rate communication," Dec. 1995.

[2] MPEG "MPEG-4 video verification model version 3.0," ISO/IEC JTC1/SC29/WG11 N1277, July 1996.

[3] "Draft ITU-T recommendation and final draft international standard of joint video specification (ITU-T Rec. H.264/ISO/IEC 14 496-10 AVC," in Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, JVTG050, 2003.

[4] I.J. Cox, S. Roy, S.L. Hingorani, "Dynamic histogram warping of image pairs for constant image brightness", in: Proceeding ICIP, Washington, USA, 1995, pp. 366-369.

[5] Sang Hyun Kim, Rae-Hong Park, "Fast local motion-compensation algorithm for video sequences with brightness variations", IEEE trans. Circuits Syst. Video Technol., Vol. 13, No. 4, April 2003, pp. 289-299.

[6] C. D. Kuglin and D.C. Hines, "the phase correlation image alignment method", in Proc. Int. Conf. Cybern. Soc., Sept. 1975, pp.163-163.

[7] M.hotter, "differential estimation of the global motion estimation parameters zoom and pan", signal processing, vol. 16, mar. 1989, pp 249-265.

[8] AVS reference software RM3.0 http://www.avs.org.cn