Fast Frame/Field Coding for H.264/AVC

Lejun Yu
Graduate University of Chinese Academy of Sciences
jlyu@ict.ac.cn

Jintao Li
Institute of Computing Technology, Chinese Academy of Sciences
jtli@ict.ac.cn

Yanfei Shen
Institute of Computing Technology, Chinese Academy of Sciences
syf@ict.ac.cn

Abstract

Picture-adaptive frame/field (PAFF) coding and macroblock-adaptive frame/field (MBAFF) coding in the new H.264/AVC can improve compression efficiency for interlaced video coding. It is assumed that the optimal solution can be found by employing a multi-pass strategy, which encodes a macroblock or picture with all possible modes and chooses the one with smallest rate-distortion cost, such as the three-pass coding for PAFF and MBAFF coding in reference of H.264/AVC. However, this coding strategy needs high computational complexity. In this paper, a novel single-pass algorithm based on motion activity detection is proposed to predict the mode of PAFF and MBAFF coding in an analysis stage. The simulation results show that the proposed scheme can reduce computational complexity by approximate 60%-75% compared to the three-pass coding method, while maintaining grate coding efficiency.

1. Introduction

Since the invention of TV over 75 years ago, the interlaced video has been exclusively adopted as the result of tradeoff between transformation bandwidth requirement and picture rate. Interlaced scan separates each frame into two fields: a top filed and a bottom filed. The top field represents all even-numbered scanlines in the frame, and the bottom field represents all odd-numbered scanlines. By means of interlaced scan, picture rate is doubled as compared to that of progressive scan under the same bandwidth requirement, which reduced the perceived flicker.

To compress the interlaced video more efficiently, the H.264/AVC standard [1, 2] adopts two levels of adaptive frame/field coding: picture-adaptive frame/field (PAFF) coding and macroblock adaptive frame/field (MBAFF) coding. With PAFF coding, a frame can be coded as a frame picture or as two field pictures. In field based coding, the top field and bottom field in the interlaced frame are coded as a separated picture, respectively. While in frame based coding, the frame, which is created by interleaving both even-numbered scanlines and odd-numbered scanlines, is coded as a whole picture.

When MBAFF coding is not used, the coded frame, which is called normal frame in this paper, is partitioned into 16x16 macroblocks, and each of them is coded as frame macroblock. When MBAFF coding is supported, the frame picture, which is called MBAFF frame in this paper, is partitioned into 32x16 macroblock pairs, and both macroblocks in each macroblock-pair are coded in frame mode or field mode. MBAFF coding provides great coding efficiency when a scene consists of both stationary and significant motion regions. An example of MBAFF frame coding is given in Figure 1, where the macroblocks in grids are coded in field mode, and the remaining macroblocks are coded in frame mode.

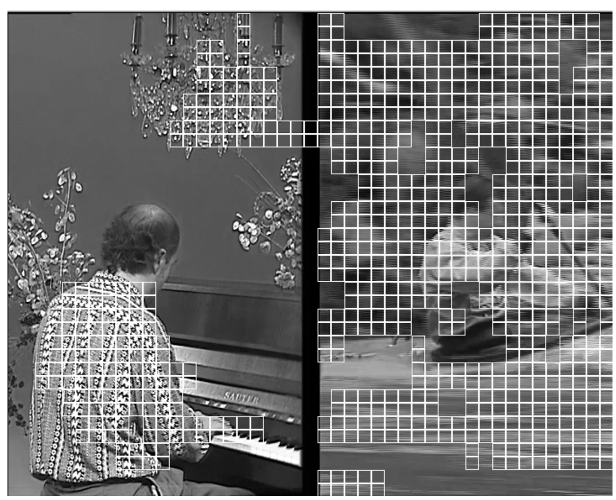

Figure 1. Example of MBAFF

0-7695-2650-0/06/$20.00 (c) IEEE

The H.264/AVC standard does not allow the combination of normal frame and MBAFF frame coding in a bitstream with the same sequence header, but field and frame (normal frame or MBAFF frame) coding can be used together to improve the coding efficiency. This can be realized by performing an adaptive determination of the picture coding mode. In this paper, we focus on the combination of PAFF and MBAFF coding, because this combination can get the best coding efficiency. So the frame coding means MBAFF frame coding in the following text.

Obviously, PAFF coding requires that an encoder has the ability to adaptively select the picture coding mode, and MBAFF requires that the encoder has the ability to adaptive select macroblock mode for each macroblock pair in a frame picture. In the H.264/AVC reference software [5], a multi-pass approach is used to perform the PAFF and MBAFF decisions. In this scheme, the frame is first encoded in frame picture mode. When MBAFF coding is used, every macroblock pair is coded as two frame macroblocks or two field macroblocks respectively, and the mode with lower rate-distortion cost is selected, and the corresponding bitrate-distortion cost of all the macroblock pairs in the frame are recorded and summed to get the cost of whole frame picture. It is obvious that this MBAFF frame coding strategy needs two-pass coding With PAFF coding, the same frame is also encoded in field mode, which means the frame is coded as two separate field pictures, and the rate-distortion cost of each macroblock is again recorded and summed to get the cost of two fields in the current frame. So three-pass coding is needed for PAFF and MBAFF coding in reference software [5]. Mode decision of PAFF coding is then made based on summed rate-distortion cost values obtained above.

With the right rate-distortion optimization method, the mode selection scheme in reference software provides with an optimal decision for both MBAFF and PAFF coding, but the computational complexity triples compared to a single-pass approach, so less complicated MBAFF and PAFF coding are desirable. Based on the fact that field picture mode coding provides higher compression efficiency than frame picture mode coding for fast-moving scenes in interlaced video, and vice versa, some fast picture level frame/filed coding algorithms are proposed. A single-pass scheme [3] was previously proposed for performing a picture level frame/field decision for MPEG-2 video coding. Although the scheme is very effective for MPEG2, the method has never been tested for H.264/AVC. And a one pass PAFF coding scheme for H.264/AVC was proposed by Yin [4], which decides frame/field mode of picture by the motion activity of the whole frame. With the 2-field motion detection used by Yin, motion detection is not so accurate because of the different phase between the top field and bottom field in a frame. And MBAFF as a new tool for interlaced video coding in H.264, fast algorithm is not considered in all these papers available.

As we known, very few papers have discussed the fast PAFF and MBAFF coding together. In this paper, a novel single-pass algorithm is proposed, which can simultaneously perform both PAFF and MBAFF coding mode decisions for interlaced video based on the motion detection. A pre-analysis stage is performed to decide the frame/field mode of each macroblock-pair, and the PAFF coding mode is decided by all the macroblock-pair modes in the frame. The detailed algorithm is described in section 2, and simulation results and conclusions are given in section 3 and section 4 respectively.

2. Fast PAFF and MBAFF coding

Based on the posteriori result of using the RD based three-pass scheme, it can be confirmed that the PAFF and MBAFF coding modes are directly related to the motion activity within each scene or parts of scene. The frames or parts of a frame with high motion activity tend to prefer field coding, whereas frames characterized by low motion activity benefit from frame coding. To achieve fast frame/field coding, a simple and accurate motion detection algorithm is need for frame/field mode pre-decision at both picture and macroblock level. In this paper, we extend the 4-field motion detection algorithm used in de-interlacing technology [7].

Let the pixel at (x, y) in current frame to be $c(x, y)$, whose corresponding pixel with the same position in the forward frame is $f(x, y)$, so the motion status of pixel at (x, y) is defined as

$$p(x,y)=\begin{cases}1 & if\ |c(x,y)-f(x,y)|>T_1\\ 0 & otherwise\end{cases} \quad (1)$$

where T_1 is a predefined threshold, and operator $|\cdot|$ denotes to calculate the absolute value. If the difference between $c(x, y)$ and $f(x, y)$ is larger than T_1, then the motion status $p(x, y)$ equals to 1, which means the current pixel is in a moving region. Otherwise, $c(x, y)$ is in a static region.

Set the *i-th* macroblock pair to be MBP_i, and the motion activity of MBP_i is defined as the number of moving pixels in it, that is to say,

$$m_i = \sum_{(x,y)\in MBP_i} p(x,y) \quad (2)$$

The motion detection of macroblock pair MBP_i is achieved by compare m_i with a predefined threshold T_2

$$D_i = \begin{cases} 1 & if(m_i(x,y) > T_2) \\ 0 & otherwise \end{cases} \quad (3)$$

If D_i equals to 1, then the macroblock pair MBP_i is in a moving region, which should be coded in field mode when MBAFF is used. Otherwise, MBP_i is in a static region and frame macroblock coding is favorable for MBAFF coding.

To predict the PAFF coding mode, global motion detection of the whole frame should be performed, which can be realized based on the results of motion detection of macroblock pairs. With the motion detection results { D_0 , D_1 ,…, D_{N-1} } of all the macroblock pairs in the current frame, where N is the total number of macroblock pairs in the current frame, the motion activity of the current frame is defined as

$$m_{frm} = \sum_{i=0}^{N-1} D_i \quad (4)$$

The frame/field coding mode of current frame is decided by comparing m_{frm} with the pre-defined threshold T_3,

$$D_{frm} = \begin{cases} 1 & if(m_{frm} < T_3) \\ 0 & otherwise \end{cases} \quad (5)$$

If D_{frm} equals to 1, there is global motion in current frame, and field coding mode should be chosen for PAFF coding, that is to say, the top and bottom fields are coded separately. If D_{frm} equals to 0, most regions of current frame are static, and frame mode should be chosen for PAFF. When MBAFF is used, the macroblock level frame/field coding mode of each macroblock-pair MBP_i is decided by the motion decision result D_i, which has been discussed above.

To sum up, with the motion detection algorithm introduced above, the PAFF and MBAFF coding can be achieved with only single-pass, the block diagram of which is illustrated in Figure 2. The motion detection results of MBAFF are transferred to frame picture coding module to control the coding mode of macroblock pairs. And the motion detection results of PAFF control the switch between field picture coding module and frame picture coding module.

3. Simulation results

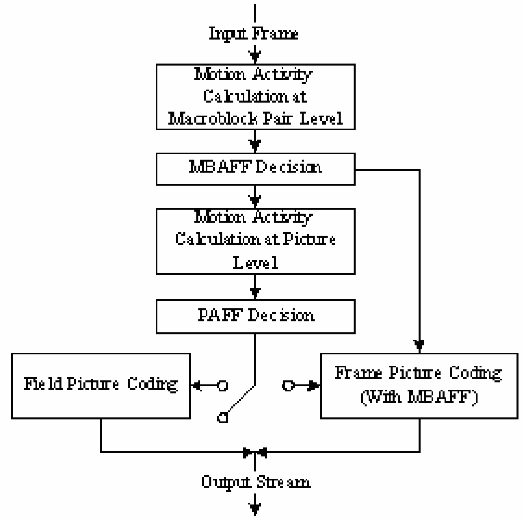

Figure 2. Single-pass encoder with PAFF and MBAFF coding

The proposed fast frame/field coding was integrated within the H.264 reference software JM8.0 [5]. The sequences in Table 1, Football, Boat, Musician are 720x576 interlaced videos at 50 fields per second. The sequences *Football* and *Boat* are fast motion sequences, while *Musician* with just lower motion of zoom. To get one sequence with both fast moving regions and static regions, the sequence *Musiboat* is combined of the left part of sequence of *Boat* and the right part of *Musician*. Only the first frame was coded as intra frame composed of two field pictures, and all the remaining pictures are coded as B or P frames in the pattern like BBPBBP…. With the main profile, quantization values 24, 28, 36 and 40 were test. Performance of the proposed scheme was compared with the rate-distortion based three-pass decision in reference software JM8.0.

The simulation results are shown in Table 1. The performance is compared with the PAFF+MBAFF coding, PAFF coding and MBAFF coding, and the corresponding results are shown in rows labeled as (1), (2) and (3) respectively. The peak signal noise ratio gain (ΔPSNR) and percentage of bitrate saving (Δbitrate) are obtained by [6]. The percent of the time saving (Speedup), calculated with $\left({t_{proposed}}/{t_{(i)}} - 1 \right) \times 100\%$ ($t_{proposed}$ is the running time of proposed fast algorithm, $t_{(i)}$, i=1,2,3, is the running time of PAFF+MBAFF, PAFF, MBAFF respectively.), is used to evaluate the computational complexity.

Because the fast motion in them, sequences *Football* and *Boat* can get better coding efficiency by PAFF coding than by MBAFF coding. And the sequences of *Musician* and *Musiboat* benefit from MBAFF coding because of large static regions in them. These means only PAFF coding or only MBAFF can not get the best compression efficiency. The simulation result show that, for all the tested sequence, PAFF+MBAFF coding is the best one of all the coding schemes Compared with all the PAFF+MBAFF, the proposed single-pass coding algorithm reduce about 60%-75% computational complexity, with the cost of 0.06-0.21dB PSNR decrease or 1.34%-3.96% bitrate increase.

Compared with PAFF coding or MBAFF coding, the proposed single-pass frame/field coding is more efficient when all the test sequences are considered, and computational complexity is reduced about 30% to 60% respectively. To show the compression efficiency more obviously, the rate-distortion performance for sequence *Musiboat* is illustrated in Figure 3, where normal frame coding ("Frame") and field ("Field") are also provided. We can tell the proposed algorithm can maintain great coding efficiency.

Table 1. Performance of the proposed scheme

Sequences		Δbitrate (%)	ΔPSNR (dB)	Speedup (%)
Football	(1)	+3.96	-0.21	-63.25
	(2)	+1.59	-0.09	-34.48
	(3)	-6.26	+0.34	-51.04
Boat	(1)	+4.41	-0.27	-62.09
	(2)	+2.83	-0.18	-33.88
	(3)	-10.60	+0.69	-44.95
Musician	(1)	+1.34	-0.06	-74.18
	(2)	-3.36	+0.16	-55.35
	(3)	+2.54	-0.11	-62.13
Musiboat	(1)	+3.45	-0.19	-68.65
	(2)	-1.94	+0.11	-45.48
	(3)	+4.41	-0.24	-53.47

Note: data in the rows labeled as (1), (2), (3) are results compared with PAFF+MBAFF coding, PAFF coding and MBAFF coding respectively

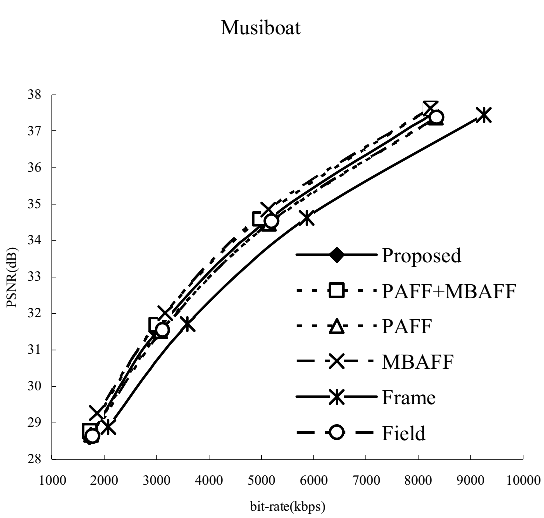

Figure 3. Rate-distortion performance

4. Conclusions

A novel fast frame/field algorithm is proposed in this paper, in which both PAFF and MBAFF modes are decided with motion detection in the pre-analysis stage. The proposed single-pass frame/filed coding scheme can reduce computational complexity by about 60%-75% as compared with the three-pass method in reference software of H.264/AVC, while great coding efficiency is maintained.

5. Acknowledgements

This work was supported in part by the National Nature Science Foundation of China (60473002), the Key Project of Beijing Natural Science Foundation (4051004) and the Key Project of International Science and Technology Cooperation (2005DFA11060).

6. References

[1] T. Wiegand, G.J. Sullivan, G. Bjontegaard and A. Luthra, "Overview of the H.264/AVC Video Coding Strandard," *IEEE Transactions on Circuits and Systems for Video Tecnology*, vol.13, no.7, July 2003, pp. 560-576.

[2] "Advanced video coding for generic audiovisual services," ITU-T Recommendation H.264, May. 2005.

[3] X.M. Zhang, A. Vetro and H. Sun, "Adaptive field/frame selection for high compression coding," *Proceedings of the 2003 SPIE Conf. on Image and Video Communications and Processing*, vol.5022, Jan 2003, pp.583-593.

[4] P. Yin, A.M. Tourapis, J. Boyce, "Fast decision on picture adaptive frame/field coding for H.264," *Visual Communications and Image Processing 2005,* SPIE, Bellingham, 2005, pp. 2092-2099.

[5] JVT Reference Software version JM8.0, http://bs.hhi.de/~suehring/tml/download

[6] G.Bjontegaard, "Calculation of average PSNR differences between RD-Curves", VCEG-M33, Mar. 2001.

[7] C. Lee, S. Chang and C. Jen, "Motion adaptive pro-scan conversion," *IEEE Internal Symposium on Circuits and Systems*, 1991, Pages :666-669.