MERL – A MITSUBISHI ELECTRIC RESEARCH LABORATORY
http://www.merl.com

Adaptive Field/Frame Selection for High Compression Coding

Xi Min Zhang   Anthony Vetro   Huifang Sun
Yun Q. Shi

TR-2003-29   January 2003

Abstract

In practice, interlaced video sequences are typically coded with either a frame-only or field-only structure, irrespective of the content. However, coding in this way will not provide the best coding efficiency. This paper proposes an adaptive picture-level field/frame coding scheme with corresponding rate control. First, a two-pass field/frame decision scheme is proposed. In this scheme, we formulate the field/frame decision as a constrained optimization problem. The actual rate and distortion data are collected and the optimal picture-level coding decision is determined based on this data. An effective rate control for the proposed two-pass algorithm is also presented. However, since the complexity of the two-pass scheme is relatively large since motion estimation must be performed for both the frame-based picture and the field-based picture, we also propose a one-pass field/frame decision scheme. This one-pass scheme calculates the variance of each macroblock in a field and estimates the correlation between two fields. Based on the correlation, a decision to code the picture as a frame or as fields is made. A rate control method for the proposed one-pass scheme is also presented. Simulation results demonstrate that our scheme outperforms frame-only and field-only coding for several sequences coded at a wide range of bit-rates, and the proposed one-pass scheme obtains similar performance as the proposed two-pass scheme.

Published in SPIE Conf. on Image and Video Communications and Processing, Jan 2003.

Adaptive Field/Frame Selection for High Compression Coding

Xi Min Zhang[†], Anthony Vetro[‡], Huifang Sun[‡], Yun Q. Shi[†]

[†]Department of ECE, New Jersey Institute of Technology, Newark, NJ
[‡]MERL - Mitsubishi Electric Research Laboratories, Murray Hill, NJ

ABSTRACT

In practice, interlaced video sequences are typically coded with either a frame-only or field-only structure, irrespective of the content. However, coding in this way will not provide the best coding efficiency. This paper proposes an adaptive picture-level field/frame coding scheme with corresponding rate control. First, a two-pass field/frame decision scheme is proposed. In this scheme, we formulate the field/frame decision as a constrained optimization problem. The actual rate and distortion data are collected and the optimal picture-level coding decision is determined based on this data. An effective rate control for the proposed two-pass algorithm is also presented. However, since the complexity of the two-pass scheme is relatively large since motion estimation must be performed for both the frame-based picture and the field-based picture, we also propose a one-pass field/frame decision scheme. This one-pass scheme calculates the variance of each macroblock in a field and estimates the correlation between two fields. Based on the correlation, a decision to code the picture as a frame or as fields is made. A rate control method for the proposed one-pass scheme is also presented. Simulation results demonstrate that our scheme outperforms frame-only and field-only coding for several sequences coded at a wide range of bit-rates, and the proposed one-pass scheme obtains similar performance as the proposed two-pass scheme.

Keywords: Interlaced video, field/frame decision, two-pass, one-pass, rate-control

1. INTRODUCTION

The advances of video and audio coding technology have successfully changed the people's entertainment style, MP3, DVD have been part of our life. However, the improvement of compression efficiency is still the objective of video coding researchers. The research in this area has been drawn more attention since the DVD Forum decided to adopt low-bit-rate compression for high definition DVD.[1] In the current MPEG video coding standards, only the general coding methodology and format for representing data input to the decoder are specified. Thus, the researchers have many flexibilities to develop their own specified MPEG encoder. Video compression improvement can be achieved by using suitable encoder design method.

In order to obatin an optimal encoder, the research efforts mainly focus on the following areas: image preprocessing; motion estimation; coding mode decison and rate control. Usually, several of them are considered together. Among the previous research work in the above areas, Ramchandran[2] and Lin[3] investigated how to implement dependent quantization among different type frames, such that optimal bit allocation can be obtained. Lee and Dickinson[4] proposed an adaptive frame type selection method for MPEG encoding. Trellis search are Lagrangian multiplier technique are combined to find the optimal frame type arrangement. The MPEG encoding optimization method proposed by Sun and $et\ al.$[5] is perhaps the most comprehensive and practical one. In their method, the macroblock mode decision is jointly optimized with rate control, an effective R/D model is proposed and uniform subjective quality through the picture is assumed. Recently, Naito and *et al.*[6] proposed an optimal MPEG-2 encoder design method for low bit-rate HDTV digital broadcasting. In their scheme, the prior art techniques[4,5] are integrated with a new rate control method. The picture level adaptive field/frame mode decision is not touched in the above works.

Interlaced video is the commonly used scan format for television systems, in which one frame is divided into a top-field and bottom-field. The two interlaced fields represent odd- and even-numbered rows of picture elements in the frame. The fields are sampled at the different times to enhance the temporal smoothness of the video playback.

Compared to the progressive video scan format, interlaced video has different characteristics and provides more encoding options. For instance, one 16x16 frame-based macroblock can be divided into two 16x8 field-based blocks as shown in Figure 1. In this way, the DCT may be applied to either frames or fields of the video. Also, there is significant flexibility in the way that blocks in the current frame or field are predicted from previous frames or fields. Since these different encoding options provide different compression efficiency, an adaptive method to choose between frame picture coding and field picture coding is desirable. However, in practice, interlaced video sequences are typically coded with either a frame-only or field-only structure, irrespective of the content. Obviously, coding in this way will not provide the best coding efficiency. Although the MPEG-2

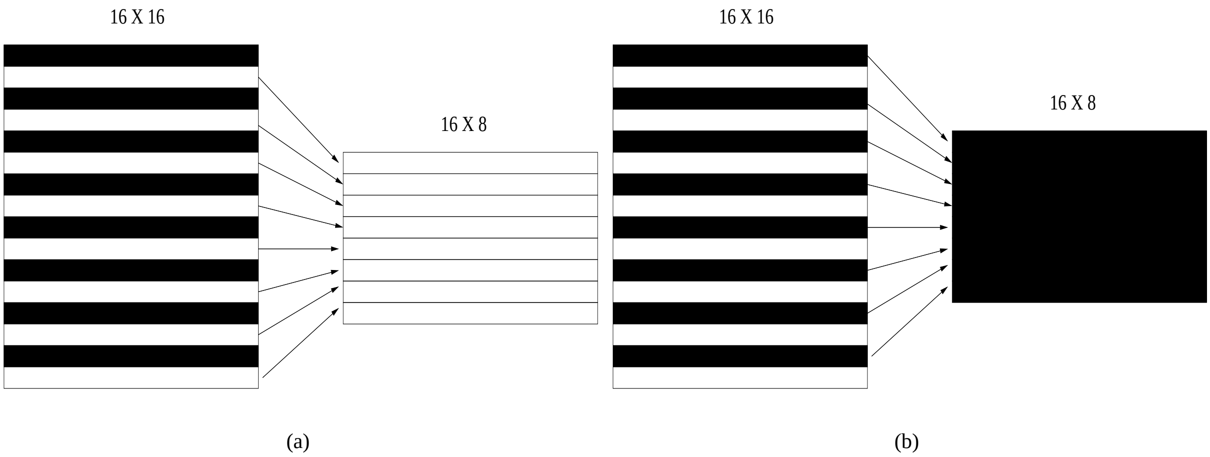

Figure 1. Comparison of quality variation with multiple source transmission and dynamic channel bandwidth. (a) uniform bit allocation, (b) sliding window approach.

standard provides some adaptive mode decision methods in macroblock level, we will show that they are not optimal in Section 2.

To solve the adaptive field/frame mode decision problem, a two pass scheme has been adopted by the Joint Video Team (JVT) reference code.[7] In this scheme, the frame is first encoded by frame mode. The corresponding bit rate and distortion are recorded. This frame is then encoded by field mode. The bit rate and distortion corresponding to field mode are also recorded. After that, a cost function is set up and the costs of two coding modes are compared. The mode with smaller cost is selected. Fixed quantization is used in this method.

In this paper, we focus on the research of optimal picture level field/frame mode decision. We first describe a two-pass adaptive field/frame decision algorithm under a rate budget constraint. In this two-pass scheme, we encode the interlaced video sequence by using field-only mode and frame-only mode, respectively. Rate control is applied to each pass, then a cost function is built based on the corresponding R-D value and the decision is finally made. We also presented an efficient one-pass adaptive field/frame algorithm. In this one-pass scheme, the characteristics of two fields are extracted and considered jointly before the encoding. The normal encoding process starts after the mode decision is made. In this way, only one pass is needed. The rest of this paper is organized as follows. In Section 2, we reviewe and analyze the mode options in MPEG-2 standard. We then propose a two-pass adaptive field/frame decision algorithm in Section 3. A more efficient one-pass algorithm is presented in Section 4. Experimental results show that both of our one-pass and two-pass adaptive algorithms guarantee performance better than frame-only and field-only coding in Section 5. Finally, we summarize our work in Section 6

2. MODE ANALYSIS OF MPEG-2 STANDARD

In the MPEG-2 standard, motion estimation for each picture can be encoded by either frame-coding or field-coding modes. With a given picture level mode, there are various macroblock modes that are associated with it. The relationship between picture coding modes and macroblock coding modes is shown in Figure 2.

MPEG-2 video encoders may use either frame-only coding, where all the frames of a sequence are encoded as frame picture, or field-only coding, where a frame is encoded as two field pictures and two fields of a frame are coded sequentially. In addition to the picture level selection, a selection procedure at the macroblock level is also used to choose the best macroblock-coding mode. One important point to make is that the macroblock modes are not optimized unless the picture level decision is decided optimally.

In MPEG-2, a macroblock can be predicted using a field prediction mode within frame pictures. This is also referred to as adaptive field/frame coding,[8] but it is at the macroblock-level. However, this macroblock-based adaptation is not optimal. This is illustrated in Figure 3. For instance, in the macroblock-based selection, the second I-field can only be encoded with Intra mode, and the P-field and B-field can only be predicted from the previous frame. On the other hand, if the picture level

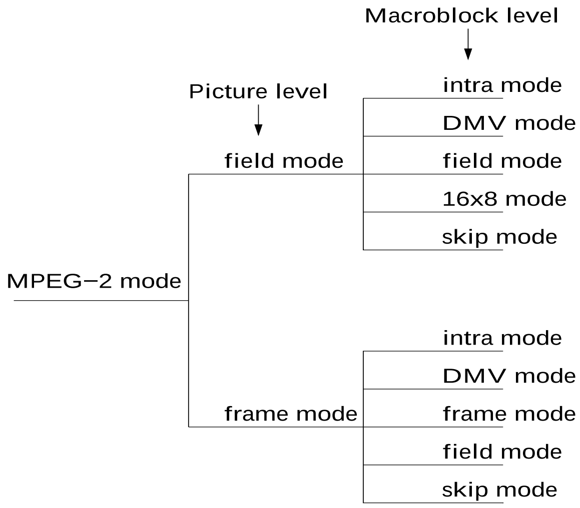

Figure 2. The block diagram of MPEG-2 coding mode options.

mode is field-only, the second I-field can be encoded with Inter mode and predicted from the first I-field; the second P-field can predicted from the first P field even if they are located in the same frame.

Hence, either frame-only or field-only coding may lead to coding inefficiency. By adaptive the frame/field coding decision at the picture level, an input frame can be coded as one frame picture or two field pictures considering content characteristics and any external constraints such as bit-rate. The rules for picture level adaptive coding include: (1) a picture header indicates whether the current picture is coded as one frame or two fields, and (2) for field-only coding, two fields of a frame are coded sequentially. If the picture type is Intra (I-type), it is divided into one I-field and one P-field. If the picture type is Inter (P-type or B-type), it is divided into two P-fields or two B-fields.

Field prediction in frame mode (top: first field, bot: second field)

Org / Ref	I−top	I−bot	P−top	P−bot	B−top	B−bot
Pre−top	No	No	Yes	Yes	Yes	Yes
Pre−bot	No	No	Yes	Yes	Yes	Yes
Cur−top	No	No	No	No	No	No
Cur−bot	No	No	No	No	No	No

Field prediction in field mode (top: first field, bot: second field)

Org / Ref	I−top	I−bot	P−top	P−bot	B−top	B−bot
Pre−top	No	No	Yes	No	Yes	No
Pre−bot	No	No	Yes	Yes	Yes	Yes
Cur−top	No	Yes	No	Yes	No	Yes
Cur−bot	No	No	No	No	No	No

Figure 3. The illustration of the difference between macroblock field prediction and field coding and frame coding modes.

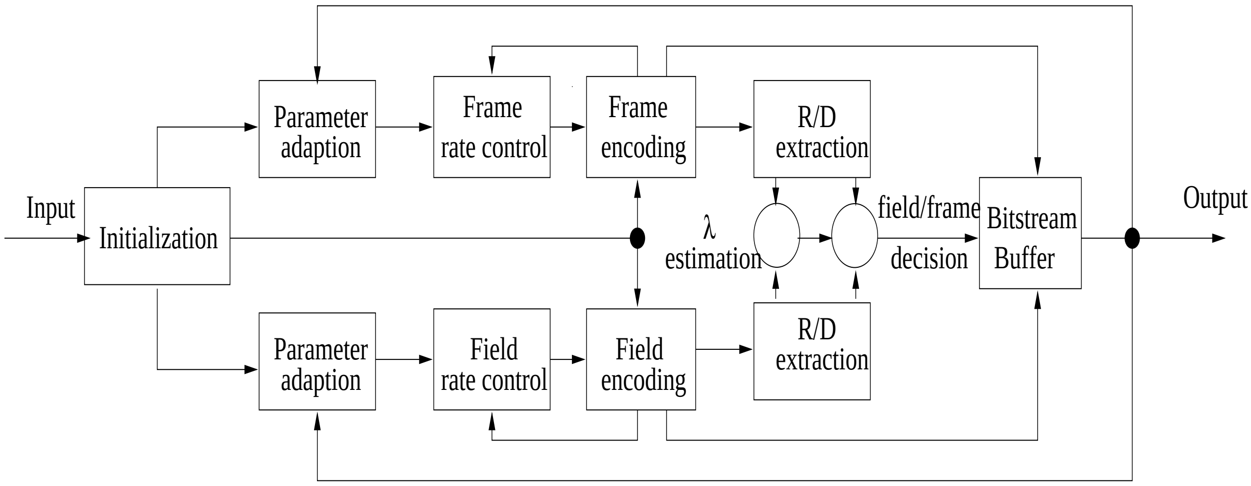

Figure 4. The principle of our proposed two-pass video encoder with adaptive field/frame picture decisions.

3. TWO-PASS ADAPTIVE FIELD/FRAME CODING ALGORITHM

Our proposed adaptive field/frame coding scheme is shown in Figure 4. In our scheme, the first block is used to initialize the parameters for the encoding of current frame, where the reference picture for the future motion estimation, the number of bits left in the buffer, and the number of bits used are determined. The current frame is then encoded using two paths. In both upper and lower paths, parameters such as the size of the picture, the number of P/B pictures left in the GOP, are updated. After all of the parameters are fixed, the current frame is encoded by using frame-only coding in the upper path and field-only coding in the lower path. In both paths, rate control is applied according to the rate budget for the current frame. The generated bitstreams are stored in two buffers separately. The number of bits used for the current frame is recorded respectively for two paths. We then calculate the two distortions based on the reconstructed frames. The two distortion values and the corresponding bits used are used to build a cost function. The values of the cost function according to different coding modes are then used to decide whether we select frame or field coding. After the decision is made, either the frame coded bitstream or field coded bitstream is selected and send out as the output. The corresponding information is fed back to the parameter initialization block for the encoding of next frame. In our scheme, the criterion for selecting either frame or field coding per frame is entirely based on rate-distortion (R-D) characteristics.

3.1. R-D based decision making

Previous methods on optimal rate allocation have provided ways to minimize the overall distortion (rate) subject to rate (distortion) constraints. Using a Lagrange multiplier to find the optimal solution is the most common approach. In general, to minimize the overall distortion, we can equivalently minimize the following cost function,

$$J(\lambda) = \sum_{i=0}^{N-1} D_i(R_i) + \lambda \sum_{i=0}^{N-1} R_i, \quad \text{subject to} \quad \frac{F_s}{N} \sum_{i=0}^{N-1} R_i \leq R_{budget} \tag{1}$$

where N is the total frames in the sequence.

The decision making problem for the selection of field only mode or frame only mode is similar as the above optimal rate allocation problem. For instance, by using field only mode in one frame, it may cost less bits compared with the frame only mode. However, the distortion of this frame may be worse than the result by using frame only mode. The optimal decision should base on both of the distortion and rate. In this paper, we adopted the similar approach as the above used by optimal rate allocation. The cost is defined as

$$cost = Distoriton + \lambda Bitrate. \tag{2}$$

If $cost(frame) < cost(field)$, we select the frame coding and vice versa. Now the problem left is how to find the suitable value of λ. In order to determine λ, we need to model the R/D relationship. Here we adopt the exponential model

$$D(R) = a\sigma^2 2^{-2R}. \tag{3}$$

This is consistent with the classic theory on this subject.[9]

Applying this model to the above cost function, we can obtain that

$$\lambda = 2a\sigma^2 2^{-2R_i} ln2 = 2D(R_i) ln2. \tag{4}$$

where R_i denotes the optimal rate allocated to frame i. So it is reasonable to use the distortion of the current encoded frame to estimate the value of λ. In our approach, the following equation is used to estimate λ in the first frame.

$$\lambda = (D_{frame}(R_{frame}) + D_{field}(R_{field})) ln2. \tag{5}$$

Then we update it by using the following equation for the following frames.

$$\lambda = W_1 \times \lambda_{current} + W_2 \times \lambda_{previous}. \tag{6}$$

In Equation (2), $\lambda_{current}$ is calculated by using Equation (1), $\lambda_{previous}$ is the estimate λ in the previous frame, W_1 and W_2 are weights, where $W_1 + W_2 = 1$. It is noted that the calculation of the I frame only use Equation 5.

The key difference between the prior art method and our method is as follows. In the prior art method, uniform quantization is used, while in our scheme, adaptive quantization is used. Also, in the prior art method, the parameter in the cost function depends on the knowledge of the quantization, while in our scheme, the parameter in the cost function is independent of the quantization. Since it is impossible to accurately estimate the motion and texture information before the encoding, real-time rate control cannot be obtained by using uniform quantization. On the other hand, the parameters in our scheme are obtained from the coding result, where the quantizer scale can be adapted according to a rate control strategy described further below. Therefore, effective rate control is realized in our scheme. In the following, we develop our rate-control procedure for the proposed two-pass adaptive field/frame decision scheme.

3.2. Rate Control for Two-Pass Algorithm

There are many rate control methods have been proposed for MPEG coding techniques, including prior art two-pass rate control methods that use the first pass to collect information and the second pass to apply rate control. The scenario is totally different in our two-pass algorithm, where the rate control is applied simultaneously to both paths, and is based on the same set of parameters transferred from the previous picture. The parameters of the existing rate control methods have not considered coding mode transitions during the encoding process. For instance, in the well-known TM5 rate control,[10] no specification is provided on how to adapt its parameters when transitioning from frame-to-field or field-to-frame so that a better budget allocation per field or frame can be achieved.

According to the description above, we do not need information pertaining to the quantizer used for encoding in our two-pass scheme. Consequently, an effective rate control scheme is developed within the context of our algorithm. In the following, we will present an effective constant bit-rate (CBR) rate control scheme for our two-pass scheme, which is partly based on the well-known TM5 rate control.

In this rate control scheme, we first initialize the rate budget R, I-picture activity X_i, P-picture activity X_p, B-picture activity X_b, I-picture buffer fullness $d0_i$, P-picture buffer fullness $d0_p$ and B-picture buffer fullness $d0_b$ by using the procedure for frame coding mode in TM5 rate control. Then all of the above rate control parameters are stored in a rate controller.

If the current frame is the first in the GOP, calculate the number of P-frames in the current GOP. Np, the number of B frames in the current GOP, N_b, then do the following:

- For the upper path: Encode the current frame by using frame coding mode, TM5 rate control and the parameters in the rate controller. Store the updated rate control parameters in a buffer BUF_{frame}.
- For the lower path: Let $N_p = 2 \times N_p + 1$, $N_b = 2 \times N_b$, then encode the current frame by using field coding mode, TM5 rate control and the parameters in the rate controller. Store the updated rate control parameters in a buffer BUF_{field}.
- If frame coding mode is selected, update the parameters in the rate controller by using the data in BUF_{frame}; if field coding mode is selected, update the parameters in the rate controller by using the data in BUF_{field}.

If the current frame is not the first in the GOP, do the following:

- For the upper path: Let $N_p = N_p/2$, $N_b = N_b/2$. Encode the current frame by using frame coding mode, TM5 rate control and the parameters in the rate controller. Remove the contents in BUF_{frame} and store the new updated rate control parameters in it.
- For the lower path: Let $N_p = (N_p + 1) \times 2$, $N_b = (N_b + 1) \times 2$. Encode the current frame by using field coding mode, TM5 rate control and the parameters in the rate controller. Remove the contents in BUF_{field} and store the new updated rate control parameters in it.
- If frame coding mode is selected, update the parameters in the rate controller by using the data in BUF_{frame}; if field coding mode is selected, update the parameters in the rate controller by using the data in BUF_{field}.

4. ONE-PASS ADAPTIVE FIELD/FRAME CODING

By using the proposed two-pass adaptive field/frame decision scheme, improved coding efficiency is obtained. However, in the two-pass algorithm, the encoding time is almost twice of the traditional MPEG-2 encoder. For some applications with very limited resources and very sensitive to the delays, a low complexity adaptive field/frame decision scheme is desirable. According to the results of using the proposed two-pass algorithm, it is observed that the decision to code a field or frame picture is directly related to the motion activities of each sequence, more specifically, each frame. The sequences or parts of sequence with high motion activities favour field-only coding, the sequences or parts of sequence with low motion activities favour frame-only coding. Now the problem left open is how to estimate the motion activity of each frame. Traditionaly, block matching based search methods are used to estimate the motion among frames and fields in the video compression. Although these kinds of methods can provide accurate motion estimation, their computational costs are high. On the other hand, we do not need the accurate value of the motion in the coding mode decision. As long as the motion level is known, the decision can be made. A simple motion estimation emthod is needed. During our investigation, we found the amount of motion can be approximated by the difference between the pixel characteristics, specifically the correlation among the top and bottom fields. Given an macroblock and partition it into two fields as shown in Figure 1, if no motion exists, most likely, the variance of the top field is similar as it of the bottom field. If high motion is detected, most likely, the variance of the top field is different with it of the bottom field. Motivated by these observations, we propose an efficient one-pass adaptive field/frame decision algorithm.

For I-frames the procedure is as follows. In the MPEG-2 standard, I-framea consists of two fields. We denote them as I-top and I-bottom, where I-top includes all of the odd lines and I-bottom includes all of the even lines. If the current frame is set to field mode, then either the top-field or the bottom-field is set as the first field and a header is added to indicate if the current field is first or second. By using field mode, the second field can be encoded as inter and predicted from the first field. We have found that it is always more efficient to predict the second I-field from the first I-field, rather than coding the entire I-frame as intra. Based on this obsearvation, the picture coding mode for I-frames is always set to field picture in our algorithm. It is noted that this does not mean that all of the macroblocks in the second field will be encoded using inter mode. Through the macroblock-based mode decision, blocks that would be coded more efficiently with intra, may be coded in that way.

The one-pass coding scheme for non-Intra coded frames described by this paper is shown in Figure 5. The input picture is sent to a field separator that produces the top-field and bottom-field. The activity is estimated for each field, where activity is defined in more detail below. The activities from each field are operated on to trigger the decision to perform either field-based motion estimation or frame-based motion estimation based on the input picture and relevant reference pictures. Depending on the picture coding decision, coding of the field-based residue or frame-based residue is coded via subsequent DCT, Quantization and VLC processes. According to the usual encoding process, P-pictures are reconstructed from the coded data and used as a reference for future pictures.

For P-pictures and B-pictures, the procedure described in more detail as follows. We consider each 16x16 macroblock in the current picture. For each macroblock, it is divided into its top-field and bottom-field. The top-field is a 16x8 block that consists of eight odd lines, and the bottom-field is a 16x8 block that consists of eight even lines. Then, our algorithm can be implemented as the following steps:

1. We first set two counters MB_{field} and MB_{frame}, and initialize both of them as zero.
2. For each 16x16 macroblock, the variance of the top-field 16x8 block and the variance of the bottom-field 16x8 block are calculated by
$$Var = \sum_i (P_i - E(P_i))^2 \ ,$$
where P_i denotes the pixel value and $E(P_i)$ denotes the mean value of the corresponding 16x8 block.

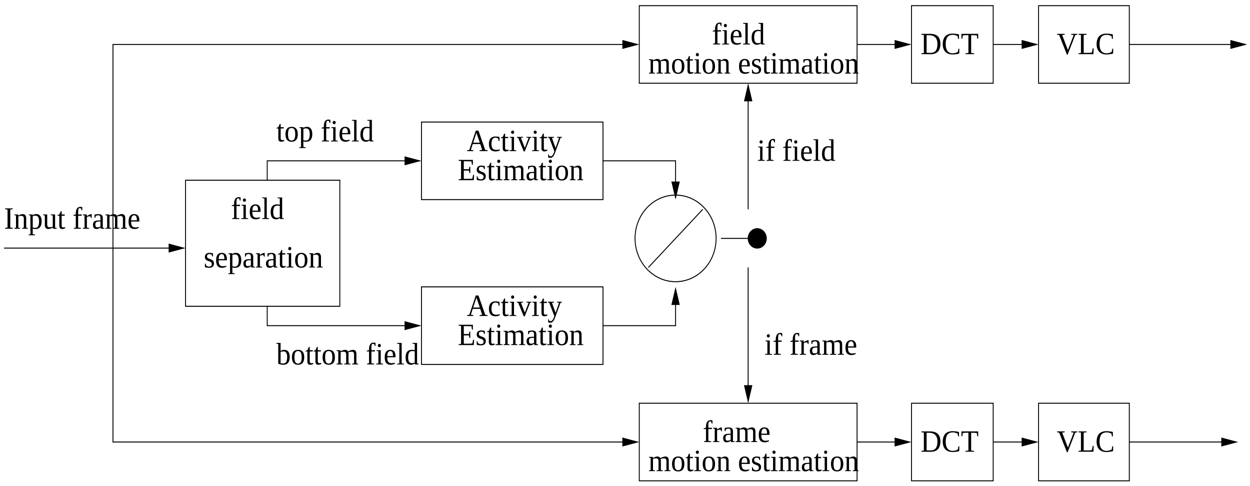

Figure 5. The principle of our proposed one-pass video encoder with adaptive field/frame picture decisions.

3. The ratio between the variance of the top-field and variance of the bottom field is calculated. Then

$$if \quad \frac{Var(top - field)}{Var(bottom - field)} > Threshold_1 \ , \quad MB_{field} + = 1 \ ;$$

$$else\,if \quad \frac{Var(top - field)}{Var(bottom - field)} < Threshold_2 \ , \quad MB_{field} + = 1 \ ;$$

$$else \qquad MB_{frame} + = 1 \ .$$

4. After looping through all macroblocks, the following picture coding decision is made: If $MB_{field} > MB_{frame}$, then field mode is selected; otherwise, if $MB_{field} <= MB_{frame}$, frame mode is selected.

In the above procedures, the values of two thresholds are obtained by large amounts of experiments applying on some typical video sequences. In summary, an effective block-based correlation method is proposed to estimate the motion activities of the current picture in our one-pass scheme. This measure is computed using a ratio of the block-based variances for each field. In doing so, the computational expensive motion estimation is avoided. The decision to code a picture as a frame or as two fields depends on the motion activity of the majority of the macroblocks in the current picture.

4.1. Rate Control for One-Pass Algorithm

As we mentioned before, the current rate control methods have not considered the coding mode transition during the encoding process in the two-pass coding algorithm. However, mode transitioning from frame-to-field or field-to-frame often happens in our one-pass algorithm. Under these circumstances, some rate-control parameters must be adapted.

The rate-control scheme for our one-pass algorithm is based on the TM5 approach and can be realized using the following procedure.

- Using TM5 to control the encoding process of the I frame (first frame in GOP), which is always field coding mode.
- If the current frame adopt frame coding mode, then
  - If the previous frame adopt frame coding mode, use the normal procedure of TM5.
  - If the previous frame adopt field coding mode, let $N_p = N_p/2$, $N_b = N_b/2$ and use TM5.
- If the current frame adopt field coding mode, then
  - If the previous frame adopt frame coding mode, let $N_p = 2 \times N_p$, $N_b = 2 \times N_b$ and use TM5.
  - If the previous frame adopt field coding mode, use the normal procedure of TM5.

Table 1. Common Sequences for Interlace Testing

	Format (4:2:0)	Length	F-code	Frames/second
Bus	720 × 480	120	H:3, V:2	30
Football	720 × 480	120	H:3, V:2	30
Stefan	720 × 480	120	H:3, V:2	30
Stefan-Football	720 × 480	120	H:3, V:2	30

5. EXPERIMENTAL RESULTS

To validate the effectiveness of the methods that we have described, we embed them in the TM5 reference MPEG-2 encoder. Simulations are carried out on a set of interlace video sequences. As shown in Table 5, Football, Stefan and Bus are the common sequence for interlace testing, and Stefan-Football is a GOP-by-GOP concatenated sequence of Stefan and Football, i.e., one GOP of Stefan, one GOP of Football, one GOP of Stefan, and so on. Football is a high motion sequence while Stefan is slow motion, but panning sequence. Frame, field and adaptive coding were performed for each of those sequences separately. A set of five rates were tested per coding method and per sequence (2Mbps, 3Mbps, 4Mbps, 5Mbps, 6Mbps).

Figures 6 (a) to (d) compare the performance of our two-pass adaptive field/frame decision algorithm to frame-only and field-only mode. In this experiment, the value of parameter W_1 and W_2 in Equation 6 is 0.9 and 0.1 respectively. The PSNR is the average of 120 frames with respect to different bit rate for the four test sequences for the picture structure of I and P only. The curves with square marks are for frame coding, the curves with star marks for field coding, and the curves with triangular marks for picture level adaptive coding. It is seen that Stefan sequence demonstrates better results with frame-only coding, Football and Bus sequences demonstrate better resutls with field-only coding. This suggests that either frame-only or field-only coding is not a good solution for interlace video material because frame-only coding performs better for some sequences while field-only coding better for others. Picture level adaptive coding is seen to be slightly better than frame-only coding for Stefan, and slightly better than field-only coding for Football and Bus. One the other hand, the parts in the concatenated sequence with Football favor field coding and the parts with Stefan favor frame coding. Adaptive coding gives a much better performance than both frame and field coding. Also, effective rate control is obtained. In the experiment, we set the rate budget as 2Mbps, 3Mbps, 4Mbps, 5Mbps, 6Mbps separately. After the encoding, the target bit rate is obtained with negligible difference.

To test the performance of our one-pass algorithm, we encode the above mentioned four sequences by using our one-pass and two-pass algorithm separately. Figures 7 (a) to (d) compare the performance of our two-pass and one-pass adaptive field/frame decision algorithm. The simulation is conducted on our optimized MPEG-2 encoder with the same conditions as above. The value of $Threshold_1$ and $Threshold_2$ is 1.35 and 0.75 respectively. It is seen our one-pass algorithm obtains similar performance as our two-pass algorithm. Effective rate control is also obtained for the one-pass algorithm.

6. CONCLUSION

The optimal picture-level field/frame mode decision is investigated in this paper. We first formulate the field/frame decision as a constrained optimization problem. Based on this formulation, a two-pass adaptive field/frame mode decision scheme is proposed under a rate budget constraint. In this scheme, the interlaced video sequence is encoded parallel in two pass by using field-only mode and frame-only mode. The actual rate and distortion data are collected. The optimal picture-level coding decision is determined based on these data, and the bitstream in the corresponding pass is selected. An effective rate control for the proposed two-pass scheme is presented. To further reduce the computational complexity of the two-pass scheme, we then propose an efficient one-pass adaptive field/frame decision scheme. In this one-pass scheme, the variance of each macroblock is calculated in two fields separately. Their difference are used to estimate the motion activity of the frame. Based on the estimation, a decision to code the picture as a frame or as fields is made. A rate control method for the proposed one-pass scheme is also presented. The normal encoding process starts after the mode decision is made. In this way, only one pass is needed. Simulation results demonstrate the good performance of both proposed two-pass and one-pass algorithm. Although this two schemes are proposed and tested on the MPEG-2 plateform, we believe their principles can be applied to other coding techniques.

REFERENCES

1. EE Times, *http://www.eetimes.com*. March 5, 2002.

2. K. Ramchandran, A. Ortega, and M. Vetterli, “Bit allocation for dependent quantization with applications to multiresolution and MPEG video coders,” *IEEE Transaction on Image Processing*, vol. 3, no. 5, pp. 533–545, 1994.
3. L.-J. Lin and A. Ortega, “Bit-Rate Control Using Piecewise Approximated Rate-Distortion Characteristics,” *IEEE Trans. Circuits and Systems for Video Technology*, vol. 8, no. 4, pp. 446–459, 1998.
4. J. Lee and B. W. Dickinson, “Rate Distortion Optimized Frame Type Selection for MPEG Encoding,” *IEEE Trans. Circuits and Systems for Video Technology*, vol. 7, no. 3, pp. 501–510, 1997.
5. H. Sun, W. Kwok, M.Chien, and C. H. J. Ju, “MPEG Coding Performance Improvement by Jointly Optimizing Coding Mode Decisions and Rate Control,” *IEEE Transaction on Circuits and Systems for Video Technology*, vol. 7, no. 3, pp. 449–458, 1997.
6. S. Naito, A. Koike, M. Wada, and S. Matsumoto, “Optimal MPEG-2 encoder design for low bit-rate HDTV digital broadcasting,” *Proc. International Conference on Image Processing ICIP02*, Rochester, NY, USA, Sep. 2002.
7. Joint Video Team, *Adaptive Frame/Field Coding for JVT*. ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, JVT-B071, 2002.
8. A. Puri, R. Aravind, and B. G. Haskell, “Adaptive frame/field motion compensated video coding,” *Signal Processing: Image Communications*, vol. 5, pp. 39–58, Feb. 1993.
9. N. Jayant and P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs, NJ: Prentice Hall, 1984.
10. MPEG-2 Video Test Model Editing Committee, *MPEG-2 Video Test Model 5*. ISO-IEC/JTC1/SC29/WG11, N0400, Apr. 1993.

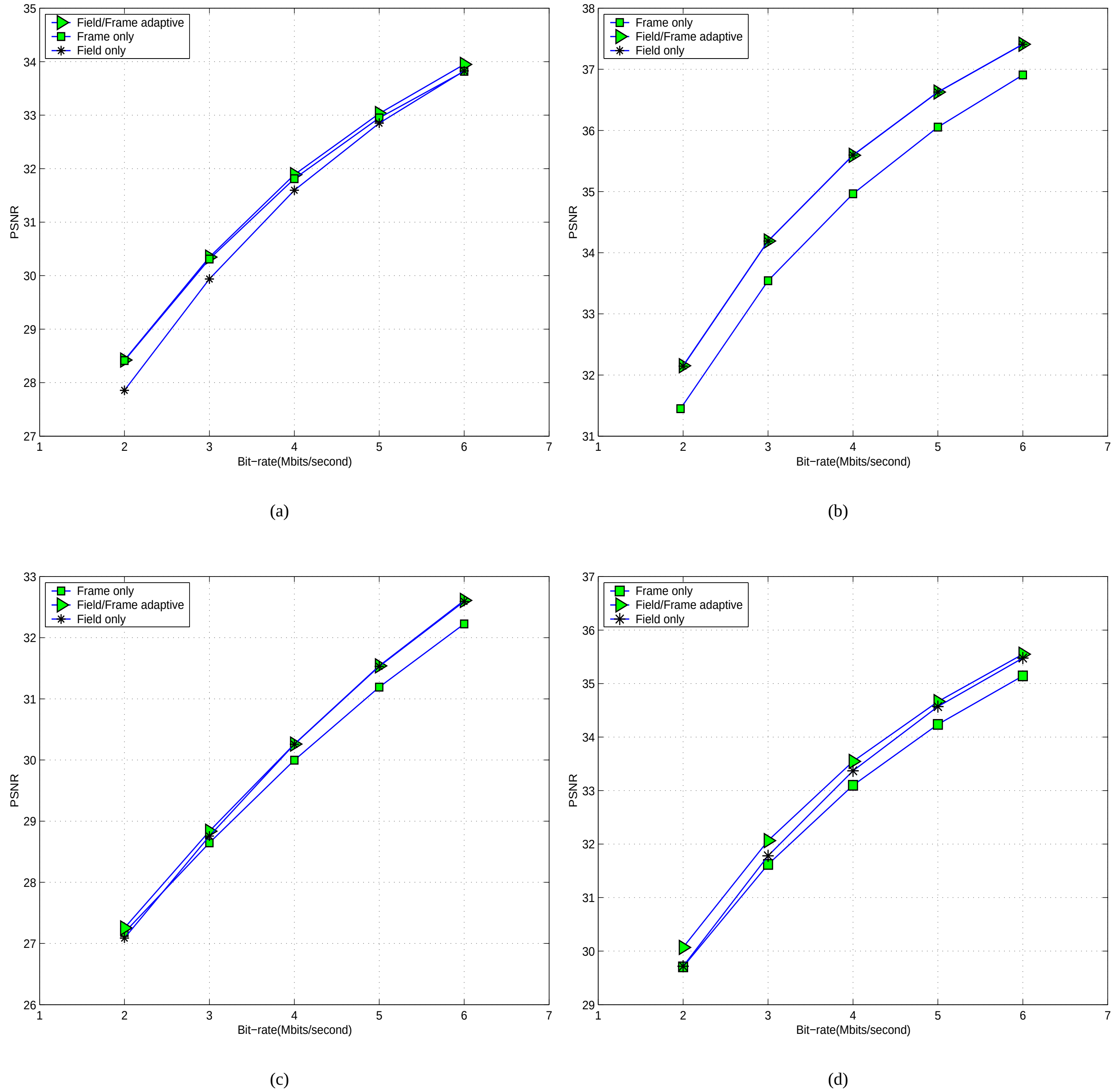

Figure 6. Comparison of coding efficiency with different coding mode. (a) Stefan sequence with rate control, (b) Football sequence with rate control, (c) Bus sequence with rate control, (d) Stefan-Football sequence with rate control.

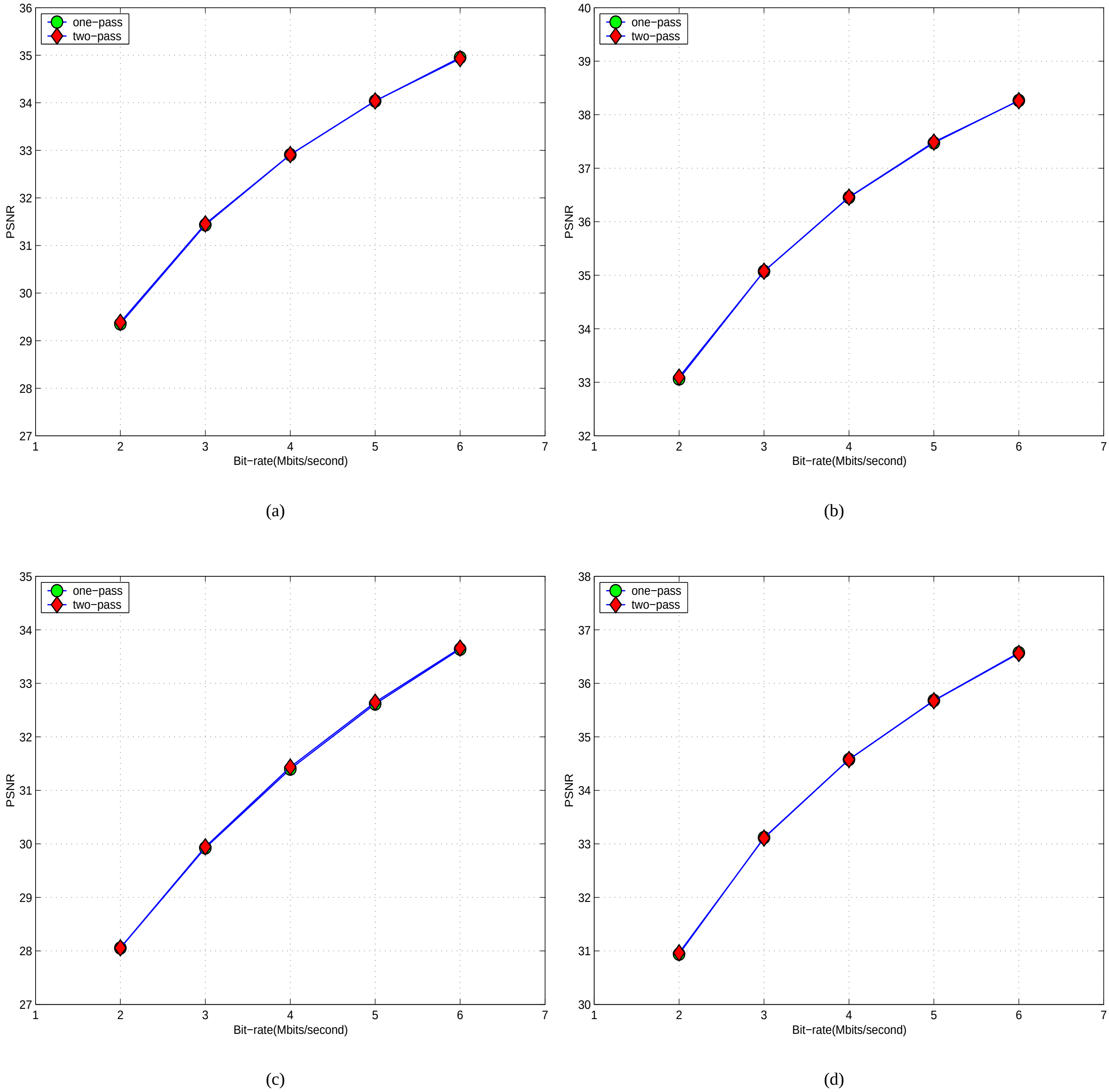

Figure 7. Comparison of performance between one-pass algorithm and two-pass algorithm. (a) Stefan sequence with rate control, (b) Football sequence with rate control, (c) Bus sequence with rate control, (d) Stefan-Football sequence with rate control.