Further Directional Search on Integer-pixel Variable Block Size Motion Estimation for H.264/AVC

Pei-Kang Song and Wan-Chi Siu

Department of Electronic and Information Engineering,
Hong Kong Polytechnic University

Abstract— Many fast motion estimation algorithms have been proposed in order to facilitate the real-time application of H.264 standard. Due to the need of developing the High Performance Video Codec (HVC), this is the time for us to review the best available fast group of algorithms and add novelty into it to make the deployment for future standards, such as the H.265. In this paper, we propose a scheme-based motion estimation algorithm which is able to describe the characteristic of block motion activities through the information of motion vector prediction from the previous frame. With the aid of this information, different search strategies can be carried out. Also, this algorithm is able to skip and save unnecessary motion search of small block types. Experimental results shows that the proposed algorithm is able to speed up the motion estimation time for almost 1.8 times compared to EPZS, but with a similar RD performance.

I. INTRODUCTION

The H.264/AVC is the newest video coding standard, which is developed by the Joint Video Team (JVT) from ISO/IEC MPEG-4 Advanced Video Coding (AVC) and ITU-T H.264. Some new features are included in the H.264. For example, the variable block size motion estimation (VBS-ME) supports up to 4x4 block partitions; and the sub-pixel motion estimation can achieve quarter-pixel accuracy. All of these advanced coding methods allow H.264 to achieve 50% bit-rate reduction. However, the improved RD performance of H.264 is at the cost of highly complex design of the coding algorithm. It is estimated that more than 80% of the encoding time is used for motion estimation.

Since the primitive Full Search Algorithm (FSA) has a huge computational burden, a large amount of research has been carried out in order to find a motion estimation algorithm with fast computational time but an acceptable decrease in RD performance. The Fast Full Search Algorithm (FFS) using techniques such as the Partial Distortion Search [1] and Successive Elimination Algorithm [2][3] is able to reduce the ME time and achieves the same RD performance of FS; however, the speed up time is not remarkable. Fast search using geometric pattern such as Three Step Search (TSS) [4] and Diamond Search (DS) [5] makes use of the characteristic of center-bias to decrease the number of candidates. With a similar idea, Pixel Decimation Algorithm [6] utilizes only a limited number of pixels for block evaluation. Both fast search by pattern and by pixel decimation can speed up the ME time substantially, whereas the algorithms might be easily trapped into the local minima, especially for sequences with fast motion activities.

Recent years, schema-based motion estimation algorithms have attracted more and more attention. One example is the Enhanced Predictive Zonal Search (EPZS) [7]. EPZS divides the search scheme into several stages, in which the focus of the search has been transferred to find the search center through various MV predictors. The pattern search stage is only a refinement of the best MV predictor. The EPZS is an efficient fast search algorithm; however, it is failed to reduce the computational burden brought by VBS-ME.

In this paper, we propose a motion estimation algorithm called Further Directional Search (FDS), which is built upon the structure of Directional Search in [8]. FDS makes use of the information from the search of previous frame, to feature the block motion activity of the current frame. Different strategies will then be used based on the judgment of motion activity of the blocks. Also, this scheme-based estimation tries to predict the best mode of variable block size and save the computational time for unnecessary search of VBS-ME.

II. ALGORITHM DEVELOPMENT

A. Block Classification in VBS-ME

In [8], Block Classification (BC) technique is used to describe the characteristic of block motions. If the final MV of one block is different from the best MV predictor, we will consider this block as a special block. If the final MV is the same as its best predictor, it will be considered as an ordinary block. Before the pattern search stage, we predict the block type by a process called prejudgment. Once the collocated MB is special, this MB, together with its 8 neighbors will be prejudged as special blocks. For prejudged special blocks, since they are difficult to predict, we rely more on the pattern search stage to find a better MV. For prejudged ordinary blocks, the center is trustable and less work can be done in the pattern search stage.

In this paper, we modified BC technique in order to better distinguish special blocks and ordinary blocks. From the experiment, we found that the BC result of small block size was related to the BC result of its upper-level block size. For example, in order to prejudice 16x8 and 8x16 blocks, we can make use of the BC result of its parent 16x16 block. If previously the 16x16 is ordinary, we also consider its 16x8 and

Acknowledgment: This work is supported by the Centre for Signal Processing of the Hong Kong PolyU under 1-BB9A, and the Research Grant Council of the Hong Kong SAR Government, Q07G.
8x16 blocks as ordinary, and vice versa. On the other hand, for the prejudgment of 16x16 block, we have to rely on the information from the previous frame. From the experiment, we found that the final MV of the collocated block could influence the prejudgment result of its neighboring blocks. Hence, we propose to make use of the direction of collocated MV as a tool to do the prejudgment. For example, in Figure 1-(a), if the collocated block of a special block contains horizontal movement, then we set both its left and right blocks as prejudged special. Figure 1-(b) is the same as (a), except that the collocated MV is vertical. In (c), since the collocated MV contains both horizontal and vertical components, we have to set all the neighboring blocks as special.

The prejudged BC result can be used to decide what pattern should be used. For prejudged ordinary block, we use small diamond pattern which contains 5 points; for prejudged special block, the square pattern which contains 9 points is used.

It can be easy to find out that compared to the original BC method described at the beginning of this part, the new way results in less special blocks for 16x16 prejudgment. In order to evaluate the consequences, we compared the two schemes by recording the percentage that the final MVs of those prejudged ordinary blocks in the new scheme but prejudged as special in original one are the same. Ten sequences were tested and the statistical results in Table 1 show that for those blocks having different prejudgment result in two schemes, the MVs are very similar, which means the new scheme do not decrease the accuracy much.

Table 1. Percentage that MVs are the same for blocks having different prejudged result in two schemes

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Percentage that MVs are the same</th>
</tr>
</thead>
<tbody>
<tr>
<td>Akiyo</td>
<td>98.62%</td>
</tr>
<tr>
<td>Mobile</td>
<td>99.50%</td>
</tr>
<tr>
<td>Bus</td>
<td>95.41%</td>
</tr>
<tr>
<td>Average</td>
<td>97.80%</td>
</tr>
</tbody>
</table>

To see the effectiveness of BC technique, we performed ME for two times - one using small diamond pattern and the other using square pattern. The experimental result indicates that more than 97% of the prejudged ordinary blocks obtain the same MV by using small diamond pattern and square pattern, whereas the same situation only happens on 85% of the prejudged special blocks. This would validate the feasibility of using small diamond pattern for prejudged ordinary blocks and the necessity of using square pattern for prejudged special blocks.

![Figure 1. Prejudgment of block classification](image)

B. MV Prediction

In this algorithm, we use a large MV Predictor SET (PSET) for the search. These predictors are
1) MV(0,0),
2) MV from the collocated block,
3) MVs from left, up, up-right blocks, and the median MV of neighboring blocks,
4) MVs from bottom and right blocks in reference frame.

Naturally increasing the number of predictors obtains more accurate initial search point. However, using more predictors means a higher computational cost.

![Figure 2. Probability of best predictors for VBS](image)

C. Mode Termination

In [8], the algorithm of mode termination is carried out by comparing the motion cost of P16x16 with P16x8 and P8x16. If the P16x16 mode is chosen as the best mode temporarily with a relatively low motion cost, then P8x8 mode will be skipped.

From our experiment, around 93% of the MBs will not be encoded as P8x8 mode if their temporary best mode is P16x16. However, to decrease the potential harm of this technique leading to a poor coding performance, we propose to do an MV refinement for P8x8 mode. Here we set a flag named flag8 Skip = 1 if the above mode termination criterion is fulfilled; otherwise flag8_skip = 0. We recorded the percentage that blocks in P8x8 mode have the final MVs within...
blocks. If the temporary best mode is P16x16 and the prejudged special blocks, respectively. After P16x16 stage, the algorithm will continue to do the search for P16x8 and P8x16 square pattern are assigned to prejudged ordinary blocks and search pattern have to be checked. The search pattern is judged about 96% of the blocks that are required to take MV refinement still get the MV pointing to the center position. Therefore, we propose to make use of the BC result of 16x8 and 8x16 blocks to judge whether MV refinement should be taken. If any one of 16x8 and 8x16 blocks is classified as special block after pattern stage and mode8_skip = 1, we have to do the MV refinement for P8x8 blocks; otherwise, we skip the motion search for P8x8 blocks.

III. OVERALL STRUCTURE OF THE ALGORITHM

Here is the whole flow of the algorithm described so far. We firstly check all the predictors in PSET. The one with the minimum SAD score will be chosen to as the initial position. The early termination threshold, which is the minimum SAD value of neighboring blocks, is used to judge whether pattern search stage should be carried out. Besides, the following conditions should adhere to the early termination threshold criteria:

(i) The block is prejudged ordinary block.
(ii) The SAD of the best MV predictor is smaller than the minimum SAD value of neighboring blocks.
(iii) The SAD of the best MV predictor is smaller than the collocated SAD score in reference frame.

If any of the above criteria is not fulfilled, candidates in search pattern have to be checked. The search pattern is judged based on the prejudgment of BC. Small diamond pattern and square pattern are assigned to prejudged ordinary blocks and prejudged special blocks, respectively. After P16x16 stage, the algorithm will continue to do the search for P16x8 and P8x16 blocks. If the temporary best mode is P16x16 and the BC result of 16x16 block is ordinary, then mode8_skip will be set to one. This flag, together with the BC result of P16x8 and P8x16 blocks, decides whether MV refinement job should be taken or not. After all inter and intra predictions, the MB will be classified based on the BC result of the best mode. This information will be used for BC prejudgment of the next frame.

IV. SIMULATION RESULT

The H.264/AVC JM12.2 was used for our experimental work, with RDOptimization option set to 1 and with the search range set to 32. To test the performance of our proposed FDS algorithm, totally 12 sequences were tested using integer-pixel motion estimation. The comparison is between the proposed algorithm and the EPZS, with reference to FSA. Results of two versions of our Further Directional Search are enclosed in this paper, which have some difference in mode termination strategy as mentioned in section II-C. FDS_1 carries out MV refinement for all P8x8 blocks if mode8_skip = 1, whereas FDS_2 performs MV refinement only when mode8_skip = 1 and the BC result of P16x8 or P8x16 block is special.

In Table 3, it is seen that the bit/pixel of FDS is almost the same as that of EPZS. One version of the proposed algorithm, FDS_1, even achieves a higher PSNR performance. On the other hand, the speed-up time of proposed algorithm, especially FDS_2 is significantly better as compared to that of the EPZS.

Let us also compare the rate-distortion performance of our algorithm with others. Figure 3 is the average rate-distortion curves from 12 sequences whose QP varied from 26–34. From this figure, we can see that the RD performance of FDS_1 is the same as EPZS, and these is a slight drop of RD curve for FDS_2, but as mentioned, the computational time of FDS_2 is reduced considerably.

V. CONCLUSION

In this paper, we have proposed a scheme-based motion estimation algorithm for VBS-ME. The searching strategy for every block is related to the information obtained from both the previous frame and the current frame. Also, the information from the upper-level block mode is used to predict the result of the proceeding mode. Extensive experiment has shown that our algorithm is much faster than EPZS in JM12.2 of the H.264 and the RD performance is almost the same as that of the EPZS. The approach is much faster than other fast algorithms available in the literature, and it is also one of our deployments for the coming H.265 standard[11-12].

REFERENCES

Table 3. Comparison between proposed algorithm and EPZS with reference of FSA

<table>
<thead>
<tr>
<th>QP = 28</th>
<th>Full Search</th>
<th>EPZS</th>
<th>FDS_1</th>
<th>FDS_2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PSNR</td>
<td>Bit/pixel</td>
<td>Speed up</td>
<td>PSNR</td>
</tr>
<tr>
<td>Akiyo</td>
<td>39.66</td>
<td>0.049</td>
<td>1</td>
<td>39.63</td>
</tr>
<tr>
<td>Bus</td>
<td>35.30</td>
<td>0.738</td>
<td>1</td>
<td>35.31</td>
</tr>
<tr>
<td>CoastGuard</td>
<td>35.13</td>
<td>0.576</td>
<td>1</td>
<td>35.12</td>
</tr>
<tr>
<td>Container</td>
<td>36.20</td>
<td>0.139</td>
<td>1</td>
<td>36.17</td>
</tr>
<tr>
<td>Football</td>
<td>37.39</td>
<td>0.499</td>
<td>1</td>
<td>37.42</td>
</tr>
<tr>
<td>Foreman</td>
<td>37.78</td>
<td>0.281</td>
<td>1</td>
<td>37.75</td>
</tr>
<tr>
<td>Hall</td>
<td>37.86</td>
<td>0.093</td>
<td>1</td>
<td>37.81</td>
</tr>
<tr>
<td>Mobile</td>
<td>35.22</td>
<td>1.287</td>
<td>1</td>
<td>35.18</td>
</tr>
<tr>
<td>MotherDaughter</td>
<td>38.60</td>
<td>0.066</td>
<td>1</td>
<td>38.55</td>
</tr>
<tr>
<td>Paris</td>
<td>35.87</td>
<td>0.263</td>
<td>1</td>
<td>35.86</td>
</tr>
<tr>
<td>Stefan</td>
<td>36.14</td>
<td>0.727</td>
<td>1</td>
<td>36.13</td>
</tr>
<tr>
<td>WaterFall</td>
<td>34.55</td>
<td>0.328</td>
<td>1</td>
<td>34.54</td>
</tr>
<tr>
<td>Average</td>
<td>36.56</td>
<td>0.420</td>
<td>1.00</td>
<td>36.54</td>
</tr>
</tbody>
</table>

Figure 3. Average rate-distortion performance of 12 sequences.