Edge Detector Based Fast Level Decision Algorithm for Intra Prediction of HEVC

Wen Shi, Xiantao Jiang, Tian Song and Takashi Shimamoto

Department of Electrical and Electronic Engineering, Graduate School of Engineering, Tokushima University
Minami-Josanjima 2-1, Tokushima City 770-8506, Japan
E-mail: shiwen@ee.tokushima-u.ac.jp, tiansong@ee.tokushima-u.ac.jp, simamoto@ee.tokushima-u.ac.jp

Abstract  High efficiency video coding (HEVC) achieves significant higher coding efficiency than previous video coding standards. In the intra prediction of HEVC, prediction unit sizes from $4 \times 4$ to $64 \times 64$ are employed, respectively defined as levels 1 to 5, to achieve higher coding efficiency. Nevertheless, this is at the expense of high computational complexity. This paper proposes a fast algorithm for the intra prediction of HEVC, which can efficiently decide the level on the basis of the Roberts-cross edge detector. The proposed algorithm utilizes the high correlation between regional texture and prediction unit partitioning. It is mainly composed of a bottom-up level decision method and an efficient decision flow based on an authentic image feature. Furthermore, chrominance information is also employed to decide the prediction unit partitioning. The experimental results for the proposed algorithm demonstrated that it achieved a task with greatly reduced computational complexity compared with the original HEVC. The average time-saving is approximately 37%, while the increase in bit rate and decrease in PSNR are negligible.

Keywords: HEVC, video coding, intra prediction, complexity reduction

1. Introduction

HEVC [1] is a new international standard developed by the Joint Collaborative Team on Video Coding (JCT-VC) that aims to achieve 50% bit rate reduction relative to H.264. In particular, it can support 4K and 8K ultra high-definition (UHD) video, with a resolution of 8,192×4,320. Intra prediction, which is of great significance for video compression in H.264, continues to play an important role in HEVC. Many new features have been proposed to improve the intra coding efficiency of HEVC. The two main features are the size of the prediction unit, which can be defined from $4 \times 4$ up to $64 \times 64$, compared with from $4 \times 4$ to $16 \times 16$ for H.264, and the 35 intra prediction modes. This results in significantly higher complexity for intra prediction.

There have already been some related works aiming to reduce the complexity of HEVC. A fast intra prediction algorithm based on pixel gradient statistics was proposed by Chen et al. [2]. In their paper, pixel gradient statistics are extracted using the Sobel operator to exclude unreasonable modes and unit sizes. The method reduces the encoding time by about 28%. Meanwhile, da Silva et al. developed a gradient based fast intra prediction algorithm using five edge strengths filters. In their paper, the edge strength results decided the corresponding intra prediction mode set. The number of available intra prediction modes was reduced to 9, compared with 35 in HEVC, and a reduction in the processing time of almost 32% was achieved. However, da Silva et al. only utilized edge direction information to reduce the number of modes considered unreasonable [3], and in some cases the list of remaining modes did not contain the optimal mode. In this paper, edge information is considered as a highly sensitive parameter of region complexity conditions and is used for reducing unreasonable unit sizes. This enables our proposed method to save much more time with almost the same bit rate and PSNR as conventional HEVC.

The remainder of this paper is organized as follows. Section 2 introduces and analyzes the intra coding of HEVC. The proposed algorithm for fast level decision is presented in Section 3. Section 4 gives details of the experimental results and performance for the proposed algorithm. Finally, this paper is concluded in Section 5.
2. Analysis of Intra Coding in HEVC

In contrast to the previous video coding standards, HEVC employs a flexible quadtree coding block-partitioning structure, which consists of three kinds of basic units: coding units (CUs), prediction units (PUs) and transform units (TUs). The CU is the fundamental partible unit, whose size can range from 8×8 to the largest coding unit (LCU). A picture is divided into slices and each slice is composed of a sequence of LCUs whose maximum size can be 64×64. Each LCU is recursively split into CUs that constitute the quadtree structure. Figure 1 shows the recursively splitting of CUs in HEVC. The PU is the basic unit for intra and inter prediction and only PU sizes of 2N×2N and N×N are supported in the intra prediction of HEVC. TUs are used for transform and quantization and the TU size can be different from the PU size. The maximum size of PUs and TUs is 64×64, while their minimum sizes are 4×4 [4].

![Fig. 1 Recursively splitting of CUs in HEVC](image1)

In HEVC, 35 intra prediction modes are employed to improve the coding efficiency [5]. As illustrated in Fig. 2, the 35 intra prediction modes consist of 33 angular modes, DC mode and planar mode. The specific number of prediction modes varies according to the PU size, as shown in Table 1. To find the optimal CU partitioning and prediction mode, HEVC must examine all the combinations of CU, PU and TU by performing a series of computations to carry out the rate-distortion optimization (RDO) process. In HEVC there are two major steps in achieving the above targets for intra prediction. The first step is to calculate the Hadamard transform absolute difference (HAD) costs of all the supported prediction modes to create a list of candidates with the minimum values of HAD. The number of candidates for different PU sizes is shown in Table 1. In the second step, the optimum PU size is derived by computing the rate-distortion (RD) costs, and the final prediction mode and CU partitioning are determined. The above process has a tremendous computational workload and is time-consuming. In fact, for an LCU with the size of 64×64, 7,552 times HAD costs and 2,623 - 4,923 times RD costs have to be calculated. Thus, if the CU partition can be determined in advance, a great deal of computation time will be saved.

![Fig. 2 Intra prediction modes in HEVC](image2)

<table>
<thead>
<tr>
<th>Size of PU</th>
<th>Number of intra prediction modes</th>
<th>Number of mode candidates</th>
</tr>
</thead>
<tbody>
<tr>
<td>4×4</td>
<td>18</td>
<td>8</td>
</tr>
<tr>
<td>8×8</td>
<td>35</td>
<td>8</td>
</tr>
<tr>
<td>16×16</td>
<td>35</td>
<td>3</td>
</tr>
<tr>
<td>32×32</td>
<td>35</td>
<td>3</td>
</tr>
<tr>
<td>64×64</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

The computation process has a lot of redundancy which can be reduced. As the depths of CUs in HEVC are 0, 1, 2 and 3, recursively considers all possible combinations at each depth in order from depth 0 to depth 3. Finally it reserves only one optimal combination. This means that there is a lot of redundancy in the process. From [6], the proportions of depths for different class sequences are shown in Fig. 3. The average percentage of depth 0 is approximately 60%. Therefore, searching for an effective way to reduce the calculated CU unit size and early termination of the RDO process are very meaningful for saving time.

3. Proposed Algorithm

The CU partitioning of a frame tends to have a high correlation with the regional texture according to experimental results. A homogeneous region is likely
to utilize a larger CU size or low-level CUs, and a complex region will be split into smaller CUs with a high probability [7,8]. An example of this is illustrated in Fig. 4, CUs are partitioned after full RDO process. HEVC utilizes low-level CUs in the top left corner and high-level CUs on the man’s face. Therefore, there is a close relationship between texture complexity and CU partitioning.

The magnitude of the gradient can be defined as follows:

$$\nabla I_{x,y} = G_{x,y} = \sqrt{G_{x}^{2} + G_{y}^{2}}$$ (3)

The angle of orientation can also be defined as follows:

$$\Theta_{x,y} = \frac{G_{x}^{2}}{G_{y}^{2}}$$ (4)

Edge detection can obtain gradient information, which mainly contains two components, the texture complexity and the spatial direction [9]. In this paper, the texture complexity is utilized to determine the PU size for intra prediction. First, the concept of the Roberts-cross edge detector is introduced to obtain the magnitude and orientation which is used for the analysis of the texture.

The calculation utilizes the two convolution masks of 2×2 units shown in Fig. 5. The gradient vector of the unit $G_{x,y}$ is obtained as follows:

$$G_{x_{i,j}} = p_{i,j} - p_{i+1,j+1}$$ (1)

$$G_{y_{i,j}} = p_{i+1,j} - p_{i,j+1}$$ (2)

The texture complexity of the original LCU can be explored using the above Roberts-cross edge detector.

3.1 Bottom-up level-decision method

It is well known that humans are less sensitive to changes in hue (chrominance) than to changes in brightness (luminance), and coding technology has made use of this feature. In this paper, pixel luminance samples of the original source are utilized to calculate the gradient magnitude. On the basis of bottom-up theory, the calculation starts from the smallest unit size (4×4). Because the gradient magnitude is tiny, a homogeneous region and four units can be merged to a larger unit (8×8) which means to be decided as the next level. On the other hand, if the derived gradient magnitude is not neglected, which means the region is complex, the four units will not be merged and decided as high-level.

In the meantime, to assist in the merging decision, chrominance sample information is introduced. By way of illustration, we will explain the chroma-assisting decision. As shown in Fig. 6, a search for the maximum and minimum chrominance of the samples from all 64 pixels in PU1, PU2, PU3 and PU4 (size 4×4, level 4) is carried out to decide whether or not the image chrominance changes sharply. If some of the four differences between the maximum and minimum are large, the four PU will be assigned to size 4×4, otherwise, the four PUs have the possibility of being assigned to a larger PU (8×8) depending to the above gradient magnitude.

As described in Algorithm 1, the calculation starts from a unit size of 2×2. Then, the chrominance of a level 4 sample is calculated and the sum of sharply varying chroma units (SCU) is evaluated. If the value
of SCU is not 0, the decision process will terminate and the level will be set as 4. Otherwise, the luminance decision process will start and the sum of homogeneous units (SHU) will be evaluated. If SHU is 4, the four PUs will be merged into a larger PU of size 8×8 and the level is set as 3. It is noteworthy that the four PUs involved in the calculation are those continuously merged from the smallest unit size (4×4); not all the PUs have a size of 8×8. Therefore, redundant calculations are avoided and the encoding time is reduced. The same process is performed at unit sizes of 16×16 and 32×32. In this way, a fast bottom-up level decision is achieved.

Algorithm 1 Bottom-up Level Decision

1: function LevelDecision
2: for every 4 continuous 2×2 units in LCU do
3: edge detection calculation
4: end for
5: for every 4 continuous 4×4 units in LCU do
6: chrominance decision if SCU ≠ 0 then
7: Level ← 4
8: else if SCU = 0 do
9: luminance decision if SHU ≠ 4 then
10: Level ← 4
11: else if SHU = 0 then
12: Level ← 3
13: end if
14: end if
15: end for
16: end function

![Fig. 6 Chroma-assisting decision](image)

3.2 Authentic-feature-based level decision method

The second part of the proposed algorithm is based on an authentic image feature, as we observed from statistical results that the PU size is mainly 16×16 and 8×8. Therefore, the same calculation as above is started from level 2, which has the higher probability with a PU size of 16×16. As shown in Fig. 7, the program is executed in two directions: to level 0 with a PU size of 64×64, and to level 4 with a PU of size 4×4. Here, we preferentially perform the calculation whose orientation is level 0, and then achieve the level decision with a top-down design starting from level 2.

![Fig. 7 Authentic-image-feature-based level decision procedure](image)

3.3 Integrated fast level decision algorithm

In general there are five levels from 4×4 to 64×64 in HEVC. The proposed algorithm can rapidly decide the level and terminate the CU compression process early. Firstly, a rough gradient calculation for a frame with unit size 64×64 is carried out to obtain the complexity of the frame. A threshold (TH) is used to choose the bottom-up level decision method or authentic-image-feature-based level decision method for the frame. In this study, a series of experiments are conducted with the sequence RaceHorses (416×240) and the results are shown in Table 2. The quantization parameter is set to 32. The increasing in bitrate (ΔBR), decrease in PSNR (ΔPSNR) and time saving in encoding (TS) are considered in selecting the threshold value. Finally, the threshold value is set to...
Table 2 Experimental results for threshold value

<table>
<thead>
<tr>
<th>Threshold value</th>
<th>∆BR(%)</th>
<th>∆PSNR(dB)</th>
<th>TS(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>122</td>
<td>1.72</td>
<td>-0.0503</td>
<td>22.40</td>
</tr>
<tr>
<td>124</td>
<td>1.69</td>
<td>-0.0523</td>
<td>25.50</td>
</tr>
<tr>
<td>126</td>
<td>1.47</td>
<td>-0.0575</td>
<td>25.61</td>
</tr>
<tr>
<td>128</td>
<td>1.31</td>
<td>-0.0508</td>
<td>27.04</td>
</tr>
<tr>
<td>130</td>
<td>1.12</td>
<td>-0.0479</td>
<td>25.35</td>
</tr>
<tr>
<td>132</td>
<td>1.12</td>
<td>-0.0569</td>
<td>25.24</td>
</tr>
<tr>
<td>134</td>
<td>1.04</td>
<td>-0.0491</td>
<td>23.69</td>
</tr>
</tbody>
</table>

128 in this paper. In the rough gradient calculation, if the value for the merged 4×4 unit is above the threshold value, the authentic-image-feature-based level decision method will be utilized. Otherwise, the bottom-up level decision method will be utilized. Then the bottom-up level decision or authentic-image-feature-based level decision is carried out to decide the level and terminate the CU compression. A flowchart of the proposed algorithm is shown in Fig. 8.

4. Experimental Results

The proposed edge-detector-based fast level-decision algorithm is integrated with the reference software of the HEVC test model (HM) 12.1 [10]. The simulation platform is Intel(R) Core(TM) 2 Quad CPU Q8400 @ 2.66GHz with 4 cores and 2.00 GB RAM. Class A, B, C, D and E sequences are employed for performance comparison. As the algorithm is mainly applied to intra prediction, we set the period of I frames to 1 to ensure that all the frames are intra encoded. The simulation conditions are defined in [11] and the quantization parameters are set to 22, 27, 32 and 37. The performances of da Silva’s algorithm compared with HM and our proposed algorithm compared to the HM are shown in Table 3. In addition, the reduction in complexity of the proposed algorithm is derived from the time saving of encoding, which is defined as follows:

\[ TS = \frac{T_H - T_{PA}}{T_H} \times 100\% \]  

where \( T_H \) denotes the coding time of HM, and \( T_{PA} \) denotes the time used by the proposed algorithm. \( \Delta PSNR \) is the difference in PSNR between the proposed algorithm and the original HM. \( \Delta BR \) denotes the percentage increased in the bit rate of the proposed algorithm compared with the original HM.

As shown in Table 3, the comparison performance between the proposed algorithm and da Silva’s algorithm gives clear results for the bit rate (\( \Delta BR \)), \( \Delta PSNR \) and time saving (TS). On average, the proposed algorithm achieves a time saving of 37.16% for intra encoding, while the average increase in the bit rate is 1.46% and the decrease in PSNR is only 0.0635 dB, which is negligible. The RD performance is shown

<table>
<thead>
<tr>
<th>Sequences</th>
<th>da Silva’s Algorithm</th>
<th>Proposed Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>∆BR(%)</td>
<td>∆PSNR(dB)</td>
</tr>
<tr>
<td>Class A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2560×1600</td>
<td>SteamLocomotiveTrain</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>PeopleOnStreet</td>
<td>1.98</td>
</tr>
<tr>
<td>Class B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1920×1080</td>
<td>Kimono</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>BQTerrace</td>
<td>2.26</td>
</tr>
<tr>
<td></td>
<td>ParkScene</td>
<td>0.03</td>
</tr>
<tr>
<td>Class C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>832×480</td>
<td>BQMall</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>BasketballDrill</td>
<td>NA</td>
</tr>
<tr>
<td>Class D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>416×240</td>
<td>RaceHorses</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>BlowingBubbles</td>
<td>NA</td>
</tr>
<tr>
<td>Class E</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1280×720</td>
<td>Vidiyo1</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Vidiyo4</td>
<td>NA</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.11</td>
<td>-0.0254</td>
</tr>
</tbody>
</table>

Table 3 Performance of the proposed algorithm and da Silva’s algorithm
in Fig. 9-11. From the RD curves for the proposed algorithm and HEVC, it is clear that our proposed algorithm achieves almost the same PSNR value for different bit rates as HM.

According to Table 3, in high-resolution video sequences, da Silva’s algorithm cannot save as much time as the proposed algorithm, while the advantages in terms of PSNR loss and the increase in bit rate are less evident than our proposed method. In the sequence PeopleOnStreet and BQTerrace, both of da Silva’s proposed algorithm and our algorithm have lower performance compared with other sequences. The two sequences have the same feature that there is a lot of shade, which influences the accuracy of edge detection. This leads to PU partition tending to occur at a low level and termination early, resulting in both algorithms missing the optimal.

Fig. 9 RD curves of ParkScene sequence

Fig. 10 RD curves of BQMall sequence

Fig. 11 RD curves of RaceHorses sequence

5. Conclusions

We proposed a fast level decision algorithm for the intra prediction of HEVC using an edge detector. By using the result of gradient detection, the times required for RDO processes are considerably reduced and the coding efficiency is increased. The proposed algorithm achieved a large reduction in computation complexity compared with the original HEVC decision algorithm. Experimental results showed that our proposed algorithm reduced the coding time by about 37.16%, while the corresponding increase in the bit rate was only 1.46% bit rate and the PSNR loss was 0.0635 dB. Compared with da Silva’s algorithm, our proposed algorithm achieves greater time saving and a similar performance in terms of the increase in bit rate and decrease in PSNR. In future, our main work will to develop novel intra prediction modes and optimize the decision of the prediction mode combination with the fast-level-decision algorithm to further increase coding efficiency.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Numbers 24700048 and 24500117.

References


algorithm achieved a large reduction in computation and the coding efficiency is increased. The proposed required for RDO processes are considerably reduced using the result of gradient detection, the times re-intra prediction of HEVC using an edge detector. By at a low level and termination early, resulting in both detection. This leads to PU partition tending to occur a lot of shade, which influences the accuracy of edge lower performance compared with other sequences. Silva’s proposed algorithm and our algorithm have sequence PeopleOnStreet and BQTerrace, both of da in terms of PSNR loss and the increase in bit rate time as the proposed algorithm, while the advantagesquences, da Silva’s algorithm cannot save as much ɹ

According to Table 3, in high-resolution video se-

different bit rates as HM. 

algorithm achieves almost the same PSNR value for algorithm and HEVC, it is clear that our proposed

in Fig. 9-11. From the RD curves for the proposed

5. Conclusions

We proposed a fast level decision algorithm for the

Fig. 9 RD curves of ParkScene sequence

Fig. 10 RD curves of BQMall sequence

References


[10] https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/


(Received August 25, 2014; revised December 19, 2014)

Wen Shi received his B.E. degree in electrical engineering and automation from Harbin Institute of Technology, Harbin, China, in 2012. He is working toward his masters degree in electrical and electronic engineering at the Graduate School of Advanced Technology and Science, Tokushima University, Tokushima, Japan. His current research interests include video coding algorithms and hardware design. vspace2cm

Xiantao Jiang received his M.S. degree from Shanghai Maritime University in 2012. Now he is currently working toward his Ph.D degree in computer science and technology at the University of Tongji, Shanghai China. He is also working toward his joint Ph.D degree in electrical and electronic engineering from Tokushima University, Tokushima, Japan. His research areas of interest are the generation of video coding (HEVC), pattern recognition, and computer vision.

Tian Song received his B.E. degree from Dalian University of Technology, China, in 1995, and his M.E. and Dr.E. degrees from Osaka University in 2001 and 2004, respectively. He joined Tokushima University in 2004 as an Assistant Professor. Presently, he is an Associate Professor of the Department of Electrical and Electronic Engineering, Graduate School of Advanced Technology and Science, Tokushima University. He is a member of IEICE, IEEE and RISP. His current research interests include heuristic algorithms for VLSI CAD. His research interests include video coding algorithms, VLSI architectures, and system design methodology.