Early SKIP Mode Decision Based on Bayesian Model for HEVC

Qiang Hu, Zhiru Shi, Xiaoyun Zhang, Zhiyong Gao

Institute of Image Communication and Network Engineering, Shanghai Jiao Tong University, Shanghai, China

{hq2902108007, zhiru.shi, xiaoyun.zhang, zhiyong.gao}@sjtu.edu.cn

Abstract— In High Efficiency Video Coding (HEVC), SKIP mode is an efficient inter prediction tool with high coding performance but low complexity. In this paper, an early SKIP mode decision algorithm is proposed to accelerate the encoding process. The rate-distortion cost (RD-cost) of SKIP mode is used as the decision criterion. A nonparametric density estimation scheme is employed to partition the SKIP RD-cost distribution space into high distinction region (HDR) and low distinction region (LDR). For a given Coding Unit (CU), if the SKIP RD-cost falls into the HDR, the SKIP mode is directly selected as the optimal mode. If the RD-cost maps into LDR, a Bayes risk minimization rule is adopted to ensure RD performance. The statistical parameters are updated according to different QPs and CU depths. Experimental results show that the proposed algorithm can reduce 47% of encoding time with only 0.34% of BD-bitrate increase on average.

Index Terms—inter prediction, SKIP mode decision, Bayes risk minimization, nonparametric density estimation, HEVC.

I. INTRODUCTION

HEVC provides higher compression performance compared to the previous standard H.264/AVC by 50% of bitrate saving for similar perceptual image quality [1], at the price of extreme computational complexity burden. For example, the rate-distortion optimization (RDO) process will be performed for all sizes of Coding Unit (CU), Prediction Unit (PU) and Transform Unit (TU), which results in complexity increasing greatly. Therefore, it’s desired to reduce the complexity of mode decision while maintaining the efficiency.

Recently, a number of approaches have been proposed to solve this problem. An early SKIP detection (ESD) algorithm is proposed in [2]. In ESD, the different motion vectors and coded block flag (CBF) of inter 2Nx2N mode are utilized. This method is adopted into the HEVC reference software model HM14.0 [3] as a fast option. Shen et al. [4] proposed an algorithm to detect early SKIP mode by using prediction information from spatially/temporally adjacent CUs. In [5], a Bayesian decision rule based early CU split decision algorithm is introduced by using relevant and computational-friendly features to assist the decision. In [6], three approaches, i.e., SKIP mode decision (SMD), early CU termination (ECUT), and CU skip estimation (CUSE) are introduced. Thresholds are calculated based on Bayes’ rule with a complexity factor. In [7], a fast HEVC CU size decision algorithm is proposed by using the neighbor and co-located CUs information to decide the depth range.

In our statistical analysis for high-resolution videos, the SKIP mode shows a very high occurrence probability. If the SKIP mode can be correctly determined in advance, a large proportion of RDO process and motion estimation will be skipped. Motivated by these observations, an efficient early SKIP mode decision scheme is proposed in this paper. Unlike the ESD algorithm, which needs to check 2Nx2N mode for CBF comparison, the proposed algorithm only checks SKIP mode and utilizes the SKIP mode RD-cost to make an early decision. Furthermore, considering the RD performance loss caused by incorrect decision, a Bayes risk minimization method has been employed to balance the RD performance loss and the complexity reduction. Based on our observation that the SKIP RD-cost distributions are different for different QPs and CU depths, an online training scheme is adopted to update the probability density distribution for each sequence.

The rest of the paper is organized as follows. Section II gives statistical analysis of SKIP mode. Section III introduces the proposed fast algorithm in detail. Performance evaluations and analyses are presented in Section IV. At last, the conclusion is drawn in Section V.

II. ANALYSIS OF THE SKIP MODE IN HEVC PREDICTION

In HEVC, SKIP mode is a very efficient inter prediction tool, and it can represent a PU without motion estimation and residual coding process if detected early. In this section, statistical analyses are conducted to reveal the property of the SKIP mode in HEVC.

A. PU Modes Distribution

In inter prediction, to decide the best PU mode in a given CU depth, all the inter PU modes are checked and compared by computing the rate-distortion cost. The RD-cost (J) function used in HM is evaluated as follow:

\[ J = SSE + \lambda \cdot Bit \]  

(1)

where Bit is the bitrate cost dependent on each PU case, SSE is the sum of squared errors between the current block and the matching block, \( \lambda \) is the Lagrange multiplier.

In order to reveal the distributions of the optimal PU mode for each CU depth level, HM14.0 is used to encode high definition sequence BQTerrance at random access configuration. The PU modes distribution results are shown in Fig. 1. It shows that SKIP is the most selected mode and takes up the major proportion ranging from 40% to 90%.
We can also see that the percentage of SKIP mode is various for different QP values and CU depths. For the low QP values (QP>22), residual coefficients are fewer quantized to zeros, so that less SKIP modes are encoded. While for the high QP situation, residual coefficients are likely to be quantized into zeros, and thus SKIP mode takes up much more. When considering CU depths, the smaller the blocks are, the more correlation can be obtained from neighboring blocks. In this case, more SKIP modes are selected for smaller CU block size. Motivated by these, if the SKIP mode can be early detected, the vast of computational complexity can be reduced, while RD performance will not be effected.

B. Analysis of RD-cost Distribution of Different PUs

In general, different PU modes exhibit different RD-cost distributions. The distributions of average RD-cost of different PU modes are illustrated in Fig. 2. In particular, the average RD-cost value of SKIP mode, Merge mode, 2Nx2N, 2NxN and Nx2N are studied. Since the random access configuration enable QP offset for each frame, the PU RD-cost is various for different frames.

According to Fig. 2, the average RD-cost difference between SKIP mode and other PU modes is large for each CU depth. As the RD-cost of other inter prediction modes are mixed together, only the SKIP mode can be distinguished by RD-cost. Based on this feature, a fast early SKIP mode determination is proposed using SKIP mode RD-cost distribution.

III. PROPOSED ALGORITHM

Based on the previous analysis, the RD-costs of SKIP and non-SKIP modes are quite different. Since the SKIP mode RD-cost can be firstly and easily obtained with lowest complexity, it is chosen to help make the early decision. In our proposed early SKIP mode decision algorithm, the SKIP mode RD-cost ($J^d_S$) in the given CU depth $d$ is calculated and used as the criterion for early decision making.

A. Skip Mode Decision in High Distinction Region

The early SKIP mode decision can be treated as a classification problem with two categories $W = \{\omega^d_S, \omega^d_{nS}\}$. The category $\omega^d_S$ denotes the SKIP mode is determined as the best in the given CU depth $d$. While category $\omega^d_{nS}$ denotes the optimal PU is non-SKIP mode in a certain CU depth $d$. The histogram distributions of SKIP RD-cost $J^d_S$ in terms of categories $\omega^d_S$ and $\omega^d_{nS}$ are investigated and illustrated in Fig. 3 using the sequence BQTerrace with QP=27 at CU depth 0.

According to the Fig. 3, the histogram shapes are various, but the tendencies are similar. $J^d_S$ of SKIP mode class $\omega^d_S$ shows a concentrated distribution centered in a narrow RD-cost range. Meanwhile, $J^d_{nS}$ of non-SKIP class $\omega^d_{nS}$ distributes in a relatively wide and flat RD-cost range. Since the possibility of non-SKIP mode is corresponding low at low $J^d_S$ range, the PUs with small $J^d_S$ are more likely to be SKIP mode. According to this, we define this RD-cost range of $J^d_S$ as the high distinction region (HDR), in which the PU can be directly determined as SKIP mode. The dotted line in Fig. 3 is the boundary of HDR, which is denoted as $th^d_{HDR}$. The $J^d_S$ mapping into HDR refers that the current CU is homogenous and easy to be predicted. In practice, the 3 percentiles of non-SKIP mode is used as $th^d_{HDR}$ in the proposed scheme by performing extensive experiments. Experimental results have confirmed this threshold can achieve good complexity reduction and guarantee RD performance.

B. SKIP Mode Decision in Low Distinction Region

In the low distinction region (LDR), the SKIP mode curve and non-SKIP mode curve are overlapped. The incorrect SKIP early decision may cause RD performance loss. It is important to notice that the RD performance loss is only caused by the incorrect decision of SKIP mode $\omega^d_S$. In the remaining wrong decision (i.e. non-SKIP mode is wrongly chosen as the optimal mode), the RDO process for the optimal PU mode is not early terminated. In this case, the RD performance is not affected without computation reduction.

Based on this analysis, we adopted the Bayes risk minimization rule to solve this binary classification problem. For SKIP/non-SKIP class $\{\omega^d_S, \omega^d_{nS}\}$, the posterior probability $p(\omega^d|J^d_S)$ is introduced. Assuming $p(\omega^d|J^d_S)$ is the probability of the observing current CU being determined into class $\omega^d_i$, $i \in \{S, nS\}$. In order to minimize the loss incurred in the early SKIP termination, a loss matrix is given as follow:

$$
\begin{pmatrix}
L^d_{i,S} & L^d_{i,nS} \\
L^d_{nS,i} & L^d_{nS,nS}
\end{pmatrix}
$$

where $L^d_{i,j}, i, j \in \{S, nS\}$ is the loss of making decision $\omega^d_i$.
when the truth is $\omega_i^d$. There is no RD loss when making correct decision, therefore $L_{NS,i}^d = 0$. For the incorrect decision as non-SKIP, there is no RD performance loss but only complexity reduction loss. In this case, we denote $L_{NS,i}^d = 1$. Meanwhile, the loss of making decision as SKIP mode while the correct decision is non-SKIP is defined as:

$$L_{NS,i}^d = (f_i^d - f_i^d) / L_{NS,i}^d$$

where $f_i^d$ is the RD-cost of non-SKIP mode for current depth. Thus the expected RD performance loss function is given as:

$$R(\omega_i^d | f_i^d) = \sum_j L_{NS,i}^d p(\omega_i^d | f_i^d)$$

(4)

Then, we can rewrite the Bayes risk in equation (4) as:

$$R(\omega_i^d | f_i^d) = L_{NS,i}^d p(\omega_i^d | f_i^d)$$

(5)

The optimal solution is the one which minimizes the loss function. The risk minimization rule can be defined as follow:

$$\begin{cases} R(\omega_i^d | f_i^d) < R(\omega_{NS,i}^d | f_i^d), & \text{choosing } \omega_i^d \\ R(\omega_i^d | f_i^d) \geq R(\omega_{NS,i}^d | f_i^d), & \text{choosing } \omega_{NS,i}^d \end{cases}$$

(6)

The posterior probability $p(\omega_i^d | f_i^d)$ can be calculated by the Bayes’ rule:

$$p(\omega_i^d | f_i^d) = \frac{p(f_i^d | \omega_i^d) p(\omega_i^d)}{p(f_i^d)} , i \in \{S, nS\}$$

(7)

where $p(f_i^d | \omega_i^d)$ denotes the class-conditional probability density function (PDF) of $f_i^d$ given $\omega_i^d$ and $p(\omega_i^d)$ is the prior probability of class $\omega_i^d$. Substituting (7) into (5), the decision rule can be rewritten as:

$$\begin{cases} \frac{p(f_i^d | \omega_i^d)}{p(f_i^d | \omega_{NS,i}^d)} < \frac{L_{NS,i}^d p(\omega_i^d)}{L_{NS,i}^d p(\omega_{NS,i}^d)}, & \text{choosing } \omega_i^d \\ \text{else }, & \text{choosing } \omega_{NS,i}^d \end{cases}$$

(8)

where $p(\omega_i^d)$ can be estimated as follow:

$$p(\omega_i^d) = \frac{N_i^d}{N_i} , i \in \{S, nS\}$$

(9)

$N_S^d$ and $N_{NS}^d$ in Eq.(9) represent the number of SKIP and non-SKIP CUs for the current depth $d$, respectively. The conditional PDF $p(f_i^d | \omega_i^d) , i \in \{S, nS\}$ is estimated using nonparametric estimation method [8].

$$p(f_i^d | \omega_i^d) = \frac{H_i(k)}{N_i^d} , i \in \{S, nS\}$$

(10)

where $H_i(k)$ is the numbers of PUs that are determined as class $\omega_i^d$ for $i$th interval of normalized histogram. It should be noticed that the $f_i^d$ distributions are different according to the sequences, CU depths and the QPs. The $f_i^d$ distributions estimation should consider the impact of these factors.

A flowchart of the proposed algorithm is shown in Fig. 4. The proposed scheme is composed of the training stage and the early SKIP mode decision stage. The first frames coded by different QPs in every $n$ frames are performed as training process. During this process, statistical parameters $p(f_i^d | \omega_i^d), p(\omega_i^d), \text{avg}_i f_i^d$ and $	ext{avg}_i f_{NS,i}^d$ are collected based on various QPs and CU depths. The normalized histogram is established for computing HDR and LDR boundary $th_{HDR}^d$ and PDF $p(f_i^d | \omega_i^d)$. Then, Risk-Minimizing decision is conducted based on these information. In practice, periodical frame number $n=50$ is used in our algorithm.

IV. EXPERIMENTAL RESULTS

The performance of the proposed algorithm is evaluated in terms of the encoding time reduction and Bjontegaard Delta bitrate (BDBR) [9]. It is implemented on the HEVC test model HM14.0, using the “encoder_randomaccess_main” (RA-Main) setting and the common test conditions recommended by the ICT-VC [10]. The experiments are conducted both for HDR only and the overall algorithm (i.e. algorithm HDR+LDR), and compared with the ESD algorithm [2]. The experiment results are given in Table 1.

It shows that the proposed HDR algorithm can reduce 37%
The proposed early SKIP mode determination algorithm has more improvements of time saving for homogeneous sequences (such as Traffic and Fourpeople). Because the SKIP mode takes more percentages in these homogeneous sequences, more CUs can be early terminated by the proposed scheme. While for the sequences with non-homogeneous regions (such as PeopleOnStreet and RaceHorse) the proposed algorithm can always achieve better time saving than ESD, which means more efficient decision can be made.

The frame by frame encoding time comparison is also demonstrated in Fig. 5 for HM, ESD and the proposed overall algorithm. The frames highlighted in the red rectangle are the training stages of proposed algorithm, where the full mode decisions are conducted and encoding time is the same as HM. After these periods, the time consumption of the proposed one is decreased dramatically.

V. CONCLUSION

In this article, an early SKIP mode decision algorithm is proposed to reduce the complexity of HEVC. The main idea is using SKIP mode RD-cost to solve the SKIP/non-SKIP classification problem. In particular, each CU is mapped into HDR or LDR. SKIP mode is directly determined when CU is mapped into HDR. For the LDR, the Bayes risk minimization rule plays an important role to keep RD performance. Experimental results show that the proposed algorithm can effectively reduce the computational complexity from 23% to 70% without noticeable quality degradation.

VI. ACKNOWLEDGEMENT

This work was supported in part by National Natural Science Foundation of China (61221001, 61133009, 61301116), the 111 Project (B07022), the Shanghai Key Laboratory of Digital Media Processing and Transmissions STCSM 12DZ2272600), National Key Technology R&D Program of China (2013BAH53F04).

REFERENCES


Table 1. Comparison of proposed algorithms to the ESD algorithm in terms of BDBR and ΔT.

<table>
<thead>
<tr>
<th>Sequences</th>
<th>BDBR ESD (%)</th>
<th>ΔT (%)</th>
<th>BDBR Proposed HDR (%)</th>
<th>ΔT (%)</th>
<th>BDBR Proposed HDR+LDR (%)</th>
<th>ΔT (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A PeopleOnStreet</td>
<td>0.32</td>
<td>26.13</td>
<td>0.27</td>
<td>27.34</td>
<td>0.30</td>
<td>37.94</td>
</tr>
<tr>
<td>Traffic</td>
<td>0.11</td>
<td>44.96</td>
<td>0.12</td>
<td>46.01</td>
<td>0.19</td>
<td>55.18</td>
</tr>
<tr>
<td>B BasketballDrive</td>
<td>0.34</td>
<td>31.80</td>
<td>0.21</td>
<td>32.23</td>
<td>0.28</td>
<td>41.15</td>
</tr>
<tr>
<td>BQTerrace</td>
<td>0.18</td>
<td>38.98</td>
<td>0.11</td>
<td>36.78</td>
<td>0.14</td>
<td>44.22</td>
</tr>
<tr>
<td>Cactus</td>
<td>0.18</td>
<td>32.34</td>
<td>0.20</td>
<td>33.12</td>
<td>0.30</td>
<td>41.85</td>
</tr>
<tr>
<td>Kimono</td>
<td>0.15</td>
<td>26.53</td>
<td>0.17</td>
<td>26.04</td>
<td>0.27</td>
<td>33.18</td>
</tr>
<tr>
<td>ParkScene</td>
<td>0.20</td>
<td>36.76</td>
<td>0.22</td>
<td>37.02</td>
<td>0.28</td>
<td>43.67</td>
</tr>
<tr>
<td>B BasketballDrill</td>
<td>0.17</td>
<td>31.81</td>
<td>0.24</td>
<td>33.24</td>
<td>0.40</td>
<td>41.23</td>
</tr>
<tr>
<td>BQMall</td>
<td>0.23</td>
<td>31.64</td>
<td>0.22</td>
<td>32.38</td>
<td>0.29</td>
<td>40.88</td>
</tr>
<tr>
<td>PartyScene</td>
<td>0.14</td>
<td>28.04</td>
<td>0.16</td>
<td>26.94</td>
<td>0.66</td>
<td>38.91</td>
</tr>
<tr>
<td>RaceHorses</td>
<td>0.49</td>
<td>23.20</td>
<td>0.24</td>
<td>21.16</td>
<td>0.40</td>
<td>27.88</td>
</tr>
<tr>
<td>D BasketballPass</td>
<td>0.35</td>
<td>41.75</td>
<td>0.28</td>
<td>39.15</td>
<td>0.46</td>
<td>53.57</td>
</tr>
<tr>
<td>BQSquare</td>
<td>0.11</td>
<td>40.96</td>
<td>0.09</td>
<td>38.98</td>
<td>0.25</td>
<td>49.71</td>
</tr>
<tr>
<td>BlowingBubbles</td>
<td>0.23</td>
<td>26.93</td>
<td>0.22</td>
<td>25.08</td>
<td>0.35</td>
<td>34.84</td>
</tr>
<tr>
<td>RaceHorses</td>
<td>0.37</td>
<td>17.85</td>
<td>0.36</td>
<td>17.42</td>
<td>0.49</td>
<td>23.49</td>
</tr>
<tr>
<td>FourPeople</td>
<td>0.01</td>
<td>53.17</td>
<td>0.10</td>
<td>52.57</td>
<td>0.26</td>
<td>66.95</td>
</tr>
<tr>
<td>Johnny</td>
<td>0.07</td>
<td>57.73</td>
<td>0.14</td>
<td>59.21</td>
<td>0.64</td>
<td>71.04</td>
</tr>
<tr>
<td>E KristenAndSara</td>
<td>0.13</td>
<td>53.53</td>
<td>0.16</td>
<td>54.21</td>
<td>0.21</td>
<td>67.18</td>
</tr>
<tr>
<td>Vidyo1</td>
<td>0.25</td>
<td>52.98</td>
<td>0.20</td>
<td>51.56</td>
<td>0.27</td>
<td>67.92</td>
</tr>
<tr>
<td>Vidyo3</td>
<td>0.42</td>
<td>48.01</td>
<td>0.26</td>
<td>49.98</td>
<td>0.36</td>
<td>61.31</td>
</tr>
</tbody>
</table>

Average 0.22 37.26 0.20 37.07 0.34 47.11