Intra Prediction Algorithm Using Adaptive Modes

Kyungmin Lim*, Jaeho Lee*, Seongwan Kim*, Dachyun Pak*, Sangyoun Lee°

ABSTRACT

H.264/AVC has shown high coding efficiency by using various coding tools, including intra and inter prediction. However, there are still many more redundancy components in intra prediction than in inter prediction. In this paper, a novel intra prediction method is proposed with adaptive mode selection. The combined intra prediction modes and simplified gradient modes are added in order to refine the directional feature and gradation region. Suitable modes are selected according to the neighboring blocks that provide a high compression rate and lower computational complexity. The improvement of the proposed method is 1.96% in terms of the bitrate, 0.25 dB in PSNR, and 1.72 times in terms of the computational complexity.

I. Introduction

H.264/AVC is the latest video standard that has been developed by ITU-T and ISO/IEC[1]. It has shown greater coding efficiency than the existing video coding standards because it uses various coding tools including multiple reference frames, variable block sizes, rate-distortion optimization (RDO), and so on[2,3]. There are two types of prediction, intra and inter prediction, which reduce the spatial and temporal redundancy, respectively[4]. Inter prediction uses multiple frames to de-correlate redundant components by using a motion compensation technique. Intra prediction reduces the
spatial redundancy by using the pixels from the
already coded neighboring blocks [5]. Although intra
prediction in H.264/AVC shows high coding
efficiency, the compression ratio of intra prediction
is lower than that of inter prediction [6].

Many studies have been done to improve the
coding efficiency of intra prediction [6-13]. Tsukuba et
al. [7] selected one of five predictor sets consisting of
nine prediction modes on each macroblock (MB).
Each prediction set is designed to do a better job of
predicting local specific edges. Shiodera et al. [8]
suggested a change of the sub-block coding order in
the MB and Bi-directional Intra Prediction (BIP),
which is generated using a weighted combination of
the prediction from the upper/left pixels and the
prediction from the bottom/right pixels. Ye et al. [9]
proposed enhanced intra prediction methods using
nine BIP modes that are frequently used, directional
transforms, and improved residual coefficient coding
in CAVLC. Intra prediction with a spatial gradient
and multiple reference lines is proposed [10]. Tian et
al. [10] used temporal-spatial prediction for intra
coding. To perform the spatial prediction, it selects
a block from the previous frame to obtain the
reference data. Thiow Keng Tan et al. [11] proposed a
template matching method as a sample predictor
using a region of reconstructed pixels. A recursive
prediction and an enhanced block-matching
algorithm are proposed [12]. Kwang Su Jung et al. [13]
proposed a new 4x4 intra coding method by
unidirectional prediction for improvement.

Although the previous studies have been dedicated
to the improvement of intra coding efficiency, they
cause very high computational complexity because
of the exhaustive search range or because of the
many intra mode decisions. In this paper, a novel
intra prediction method is proposed which improves
the coding efficiency with only a small increase in
the computational complexity. The proposed
algorithm provides prediction modes that are suitable
for refining the direction or the region of gradation.
Moreover, the highly probable modes are decided
according to the prediction modes of neighboring
blocks using the adaptive mode selection method.

This paper is organized as follows. In Section II,
the proposed method is described in detail. The
simulation results are shown in Section III. Finally,
Section IV concludes this paper.

II. Proposed Algorithm

2.1. Overall Algorithm

The coding loss in gradation or refine region is
mainly from the lack of intra prediction modes
which is only 8 directional modes and DC mode in
H.264/AVC. In this paper, various intra prediction
modes which are suitable for gradation or refine
region are proposed and adaptively selected
according to neighboring blocks. Fig. 1 represents
the flowchart of the encoding of one macroblock based on the proposed algorithm. The eight
combined modes and the sixteen simplified gradient
modes which will be described in Section II-2 are
added to the 9 conventional prediction modes in
Intra4x4 and Intra8x8 as shown in Fig. 2. The best
modes of the neighboring blocks are classified
according to 5 directional groups which will be
described in Section II-3. According to each group,
the 17 candidate modes are adaptively selected. All
of the candidate modes are tested to find the best
intra mode using the Rate Distortion optimization
process. To make bitstream decodable, the encoding
method of the intra mode also changed. It will be
described in Section II-4.

2.2. Additional Intra Prediction Modes

In intra prediction, Intra4x4 and Intra8x8 provide
8 directional modes and one DC mode as shown in
Fig. 3. In addition to the conventional intra modes,
additional prediction modes that are well-suited for
the refined direction and the part of the gradation
region are proposed. First, the combined modes are
proposed in order to consider a more detailed
directional feature between two adjacent modes. As
shown in Fig. 4, eight modes are introduced using
the following equation:

\[ S_{\text{combined mode}}(i,j) = (S_q(i,j) + S_q(i,j) + 1) \geq 1 \] 

where \( S_{\text{combined mode}}(i,j) \) represents the predicted
pixel values of the combined mode, and $S_v(i, j)$ and $S_h(i, j)$ are the predicted pixel values of the two adjacent directional modes in H.264/AVC. The variables $i$ and $j$ range from 0 to 3 in Intra4x4, and from 0 to 7 in Intra8x8. The eight modes described following represent the combined modes according to each of the two adjacent modes.

Mode 9: vertical (0) + vertical right (5)
Mode 10: vertical (0) + vertical left (7)
Mode 11: horizontal (1) + horizontal up (8)
Mode 12: horizontal (1) + horizontal down (6)
Mode 13: diagonal down left (3) + vertical left (7)
Mode 14: diagonal down left (3) + horizontal up (8)
Mode 15: diagonal down right (4) + horizontal
The proposed intra prediction modes.

Fig. 2.

Mode 16: diagonal down right (4) + vertical right (5)

Second, the gradient modes are added for the gradation region. In general, the gradient mode has increasing or decreasing prediction values according to the distance from the reference pixel. In the proposed scheme, simplified gradient modes are introduced, which use only sixteen gradient modes. Two increasing and decreasing gradient predictions are provided according to each direction of the horizontal, vertical, diagonal down right, and diagonal down left modes. The simplified gradient mode prevents high encoding complexity and reduces the number of overhead bits for intra mode coding. Fig. 5 shows the sixteen gradient modes with conventional H.264/AVC intra modes.

The gradient modes can be obtained as follows:

\[ y = \alpha \times d + \beta \]

where \( \alpha \) is the gradient coefficient that can be selected from \{ +2, +1, -1, -2 \}, \( d \) is the distance of the predicted pixel from the reference pixel, and \( \beta \) represents the reference pixel values. Each of the gradient modes is defined as follows.

Mode 17: +1 gradient to vertical (0)
Mode 18: -1 gradient to vertical (0)
Mode 19: +1 gradient to horizontal (1)
Mode 20: -1 gradient to horizontal (1)
Mode 21: +1 gradient to diagonal down left (3)
Mode 22: -1 gradient to diagonal down left (3)
Mode 23: +1 gradient to diagonal down right (4)
Mode 24: -1 gradient to diagonal down right (4)
Mode 25: +2 gradient to vertical (0)
Mode 26: -2 gradient to vertical (0)
Mode 27: +2 gradient to horizontal (1)
Mode 28: -2 gradient to horizontal (1)
Mode 29: +2 gradient to diagonal down left (3)
Mode 30: -2 gradient to diagonal down left (3)
Mode 31: +2 gradient to diagonal down right (4)

Mode 32: -2 gradient to diagonal down right (4)

From the above additional prediction methods, a total of 24 additional modes are added to the 9 conventional prediction modes. This is applied to both Intra4x4 and Intra8x8 prediction.

2.3. Adaptive Mode Selection of Prediction Modes

Intra prediction uses the spatial correlation of neighboring blocks because it is highly correlated to the best mode of the current block. For example, the MPM (Most Probable Mode), which is frequently selected as the best mode, is decided from the neighboring blocks. The reference pixels that are
Fig. 7. The four groups of prediction modes according to the main direction.

As shown in Fig. 7, 32 directional modes are divided into 4 sets according to the main direction. Each of the four groups is described as follows:

- Horizontal group
  = {modes 8, 11, 1, 12, 6, 19, 20, 27, and 28}
- Diagonal Right group
  = {modes 15, 4, 16, 23, 24, 31, and 32}
- Vertical group
  = {modes 5, 9, 0, 10, 7, 17, 18, 25, and 26}
- Diagonal Left group
  = {modes 13, 3, 14, 21, 22, 29, and 30}

The candidate intra prediction modes are selected according to the best modes of blocks A and B shown in Fig. 8. If blocks A and B are encoded with the prediction modes of the same group, such as the Horizontal group, the candidate modes of the current block are selected as the 9 conventional prediction modes and additional modes with corresponding directions, such as the horizontal direction in this case.

Fig. 9(a) shows the candidate modes when both blocks A and B are encoded with the mode of the Horizontal group. In this manner, if both blocks A and B are predicted to be the same group such as Diagonal Right, Vertical, or Diagonal Left, the current block uses 17 prediction modes with the proper additional modes of the corresponding directions as shown in Fig. 9. On the other hand, if blocks A and B are encoded with the modes of different groups, the candidate modes are decided as depicted in Fig. 4. By using the adaptive mode selection method, the number of candidate modes is always decided to be 17 modes. The candidate mode list (CML) is obtained as follows:
\[ CML = \begin{cases} 
\text{Modes} + \text{Comb}_{\text{Hor}} + \text{Grad}_{\text{Hor}} \\
(M_A \in G_{\text{Hor}}, M_B \in G_{\text{Hor}}) \\
\text{Modes} + \text{Comb}_{\text{Ver}} + \text{Grad}_{\text{Ver}} \\
(M_A \in G_{\text{Ver}}, M_B \in G_{\text{Ver}}) \\
\text{Modes} + \text{Comb}_{\text{DR}} + \text{Grad}_{\text{DR}} \\
(M_A \in G_{\text{DR}}, M_B \in G_{\text{DR}}) \\
\text{Modes} + \text{Comb}_{\text{DL}} + \text{Grad}_{\text{DL}} \\
(M_A \in G_{\text{DL}}, M_B \in G_{\text{DL}}) \\
\text{Modes} + \text{Comb}_{\text{all}} \\
(\text{Group of } M_A \neq \text{Group of } M_B) 
\end{cases} \]

where \text{Modes} is the conventional 9 intra prediction modes, \text{Comb}, \text{Grad} and \text{G} are combined modes, gradient modes and group of corresponding direction, respectively. \text{M}_A and \text{M}_B are intra prediction modes of neighboring block. As the equation (3), the candidate mode sets according to each cases are listed as follows:

Case I. Both blocks A and B are encoded with modes in the Horizontal group.
\[ = \{14, 3, 7, 0, 5, 4, 15, 6, 12, 20, 28, 1, 19, 27, 11, 8, 2\} \]

Case II. Both blocks A and B are encoded with modes in the Diagonal Right group.
\[ = \{3, 7, 0, 9, 5, 16, 24, 32, 4, 23, 31, 15, 6, 12, 1, 8, 2\} \]

Case III. Both blocks A and B are encoded with modes in the Vertical group.
\[ = \{3, 13, 7, 10, 18, 26, 0, 17, 25, 9, 5, 16, 4, 6, 1, 8, 2\} \]

Case IV. Both blocks A and B are encoded with modes in the Diagonal Left group.
\[ = \{14, 22, 30, 3, 21, 29, 13, 7, 10, 0, 5, 4, 6, 1, 11, 8, 2\} \]

Case V. Blocks A and B are encoded with modes in different groups.

Fig. 9. The candidate prediction modes when neighboring blocks are coded with modes in (a) Horizontal group, (b) Diagonal Right group, (c) Vertical group, and (d) Diagonal Left group.
In terms of intra mode coding, either one or four bits are used in H.264/AVC depending on whether or not the predicted mode is MPM. When the best mode that is determined from the rate distortion process is MPM, one bit is signaled. If it is not MPM, one bit is non-signaled, and the additional 3 bits are used to indicate one of the 8 remaining modes. Therefore, it uses a maximum of 4 bits in conventional mode coding. In the proposed algorithm, it needs some changes because of the 17 extended candidate modes. To make the bit stream decodable, it changes one bit for MPM or five bits for non-MPM in order to distinguish the intra prediction modes in the decoder.

2.4. The decision of MPM

In H.264/AVC, the MPM is set to be the minimum mode of blocks A and B. In the proposed scheme, the same MPM decision method as in H.264/AVC cannot guarantee a decodable bit stream, because each of the blocks uses different candidate modes according to the neighboring blocks. Therefore, the MPM for which the number is bigger than that of mode 16 is set into the corresponding main directional mode. The proposed decision method of MPM is described as follows:

Case I. \( \min(A, B) \leq 16 \)
- Set MPM = \( \min(A, B) \).

Case II. \( \min(A, B) = \{22, 30, 21, 29\} \)
- Set MPM as mode 3.

Case III. \( \min(A, B) = \{18, 26, 17, 25\} \)
- Set MPM as mode 0.

Case IV. \( \min(A, B) = \{24, 32, 23, 31\} \)
- Set MPM as mode 4.

Case V. \( \min(A, B) = \{20, 28, 19, 27\} \)
- Set MPM as mode 1.

III. Experimental Results

The simulation is performed using JM11.0 reference software. The various resolutions of video sequences such as the QCIF, CIF, and HEVC test sequences, which are Class A (4K), Class B (1080p), Class C (WVGA), Class D (WQVGA), and Class E (720p), are used. CABAC is used for entropy coding, and four different QPs are tested. The detailed simulation conditions are described in Table 1. A 3.20GHz processor with 6GB RAM is used for the simulation.

The results are presented according to the Bjontegaard Delta Bit Rate (BDBR) and Bjontegaard Delta PSNR (BDPSNR). The (+) sign in BDPSNR indicates the gain, and the (-) sign in BDBR indicates the bitrate reduction. The computational complexity, which is defined as \( \Delta Time \), can be represented by:

\[
\Delta Time = \frac{\text{Time}_{\text{proposed}}}{\text{Time}_{\text{JM11.0}}}. \tag{4}
\]

Table 2 shows the simulation results of the proposed algorithm and the proposed algorithm without adaptive mode selection. The proposed algorithm without adaptive mode selection shows 0.29 dB gain in BDPSNR and 2.28% reduction in BDBR while the encoding complexity is increased about 2.56 times on average. The proposed algorithm shows 0.25 dB in BDPSNR and 1.96% BDBR reduction. It is especially noteworthy that approximately 0.35 dB gain and -2.59 % bitrate are achieved for the tempete sequence. But the average encoding time is 1.72 times which is faster than the proposed algorithm without adaptive mode selection. It is because the number of searching mode in intra prediction is different. The proposed algorithm always provides 17 modes while the proposed algorithm without adaptive mode selection provides 33 modes. Table 3 presents the average simulation results of the proposed algorithm as compared with

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Profile</td>
<td>High Profile</td>
</tr>
<tr>
<td>Entropy coding</td>
<td>CABAC</td>
</tr>
<tr>
<td>QPs</td>
<td>7, 12, 17, 22</td>
</tr>
<tr>
<td>RD optimization</td>
<td>High Complexity Mode</td>
</tr>
<tr>
<td>Sequence Type</td>
<td>1 1 1 1...</td>
</tr>
<tr>
<td>Total Frames</td>
<td>100</td>
</tr>
</tbody>
</table>
Table 2. Performances of the proposed method in various sequences

<table>
<thead>
<tr>
<th>Format</th>
<th>Test Sequence</th>
<th>Proposed Algorithm</th>
<th>Proposed Algorithm without Adaptive Mode Selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>QCIF</td>
<td>container</td>
<td>0.22</td>
<td>-1.88</td>
</tr>
<tr>
<td></td>
<td>foreman</td>
<td>0.27</td>
<td>-2.22</td>
</tr>
<tr>
<td></td>
<td>silent</td>
<td>0.28</td>
<td>-2.24</td>
</tr>
<tr>
<td></td>
<td>mobile</td>
<td>0.30</td>
<td>-1.55</td>
</tr>
<tr>
<td></td>
<td>paris</td>
<td>0.24</td>
<td>-1.81</td>
</tr>
<tr>
<td></td>
<td>tempeste</td>
<td>0.35</td>
<td>-2.59</td>
</tr>
<tr>
<td>CIF</td>
<td>Traffic</td>
<td>0.22</td>
<td>-2.30</td>
</tr>
<tr>
<td></td>
<td>PeopleOnStreet</td>
<td>0.27</td>
<td>-2.56</td>
</tr>
<tr>
<td></td>
<td>NebutaFestival</td>
<td>0.52</td>
<td>-0.58</td>
</tr>
<tr>
<td></td>
<td>StreamLocomotiveTrain</td>
<td>0.45</td>
<td>-0.82</td>
</tr>
<tr>
<td>Class A (4K)</td>
<td>Traffic</td>
<td>0.22</td>
<td>-2.30</td>
</tr>
<tr>
<td></td>
<td>PeopleOnStreet</td>
<td>0.27</td>
<td>-2.56</td>
</tr>
<tr>
<td></td>
<td>NebutaFestival</td>
<td>0.52</td>
<td>-0.58</td>
</tr>
<tr>
<td></td>
<td>StreamLocomotiveTrain</td>
<td>0.45</td>
<td>-0.82</td>
</tr>
<tr>
<td>Class B (1080p)</td>
<td>Kimono</td>
<td>0.19</td>
<td>-2.87</td>
</tr>
<tr>
<td></td>
<td>ParkScene</td>
<td>0.21</td>
<td>-1.94</td>
</tr>
<tr>
<td></td>
<td>Cactus</td>
<td>0.22</td>
<td>-1.91</td>
</tr>
<tr>
<td></td>
<td>BasketballDrive</td>
<td>0.12</td>
<td>-1.32</td>
</tr>
<tr>
<td></td>
<td>BQTerrace</td>
<td>0.16</td>
<td>-1.39</td>
</tr>
<tr>
<td>Class C (WVGA)</td>
<td>BasketballDrill</td>
<td>0.23</td>
<td>-2.02</td>
</tr>
<tr>
<td></td>
<td>BQMall</td>
<td>0.23</td>
<td>-2.04</td>
</tr>
<tr>
<td></td>
<td>PartyScene</td>
<td>0.26</td>
<td>-1.59</td>
</tr>
<tr>
<td></td>
<td>RaceHorsesC</td>
<td>0.26</td>
<td>-2.27</td>
</tr>
<tr>
<td>Class D (WQVGA)</td>
<td>BasketballPass</td>
<td>0.21</td>
<td>-1.99</td>
</tr>
<tr>
<td></td>
<td>BQSquare</td>
<td>0.19</td>
<td>-1.34</td>
</tr>
<tr>
<td></td>
<td>BlowingBubbles</td>
<td>0.30</td>
<td>-2.25</td>
</tr>
<tr>
<td></td>
<td>RaceHorses</td>
<td>0.28</td>
<td>-2.28</td>
</tr>
<tr>
<td>Class E (720p)</td>
<td>Vido1</td>
<td>0.19</td>
<td>-2.69</td>
</tr>
<tr>
<td></td>
<td>Vido3</td>
<td>0.15</td>
<td>-2.09</td>
</tr>
<tr>
<td></td>
<td>Vido4</td>
<td>0.18</td>
<td>-2.37</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>0.25</td>
<td>-1.96</td>
</tr>
</tbody>
</table>

The proposed algorithm without adaptive mode selection shows higher coding gain than Ref. [7]. Compared with Ref. [6], it shows 0.1% more BDBR reduction with 2.16 times in encoding complexity. In the proposed algorithm, it results in 0.08 dB more gain in terms of BDPSNR compared with [7]. In terms of the encoding complexity, the proposed algorithm shows 1.72 times, while the corresponding complexity of [7] is 5.0 times. In [6], the average gains are 0.35 dB in BDPSNR and a 2.40% BDBR reduction, respectively. However, [6] requires a lot of encoding complexity of about 36.56 times because of its exhaustive mode searching. It is because the proposed algorithm always provides 17 intra prediction modes using adaptive mode selection which efficiently control the complexity of encoding. The rate distortion curves for the two sequences are shown in Fig. 10. The rate distortion curves show that the proposed method outperforms H.264/AVC.
IV. Conclusion

In this paper, a novel intra prediction with adaptive mode selection method is proposed in order to improve the coding efficiency. To consider the refined directional feature and gradation region, combined prediction modes and simplified gradient modes are introduced. Moreover, it efficiently prevents high computational complexity by using an adaptive mode selection method. The simulation results demonstrate that the proposed method improves the coding gain in terms of both the BDBR and BDPSNR with about 1.72 times computational complexity compared to H.264/AVC. In addition, the proposed method has high compatibility with the existing prediction algorithm. It can be adaptively extended to other intra prediction schemes. In future research, it will be applied to the intra prediction of High Efficiency Video Coding (HEVC). As HEVC standards have developed for high video quality and a data compression ratio compared to H.264/AVC, it is expected that this research can provide further coding performance improvements.

Table 3  The Average Performances and Computational Complexity compared with [7] and [6] in QCIF/CIF

<table>
<thead>
<tr>
<th></th>
<th>BDPSNR [dB]</th>
<th>BDBR [%]</th>
<th>△Time [times]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed Algorithm</td>
<td>0.28</td>
<td>-2.05</td>
<td>1.72</td>
</tr>
<tr>
<td>Proposed Algorithm without Adaptive Mode Selection</td>
<td>0.34</td>
<td>-2.50</td>
<td>2.16</td>
</tr>
<tr>
<td>Ref. [7]</td>
<td>0.20</td>
<td>-2.35</td>
<td>5.00</td>
</tr>
<tr>
<td>Ref. [6] MRL with gradient (delta=3)</td>
<td>0.35</td>
<td>-2.40</td>
<td>36.56</td>
</tr>
</tbody>
</table>

Fig. 10. The RD curves comparing JM11.0 with proposed algorithm (a) foreman (QCIF) and (b) BlowingBubbles (WQVGA).

References


박 대현 (Daehyun Pak)

2005년 3월 연세대학교 전기전자공학과 학사
2009년 3월~현재 연세대학교 전기전자공학과 통합과정
<관심분야> H.264/AVC, HEVC, Mobile IPTV system

이상윤 (Sangyoun Lee)

1987년 2월 연세대학교 전자공학과 학사
1989년 2월 연세대학교 전자공학과 석사
1999년 2월 Georgia Tech. 전기 및 컴퓨터공학과 박사
1989년~2004년 KT선임연구원
2004년~현재 연세대학교 전기전자공학부 부교수
<관심분야> 생체인식, 컴퓨터비전, 영상부호화