A Flexible Side Information Generation Scheme using Adaptive Search Range and Overlapped Block Motion Compensation

Xiem Hoang Van, Jongbin Park, and Byeungwoo Jeon
School of Information and Computer Engineering, Sungkyunkwan University
300 Chunchun-dong, Jangan-gu, Suwon, 440-746, Korea
xiemhoang@gmail.com, feynmanc@skku.edu, bjeon@skku.edu

ABSTRACT
This paper addresses a problem of side information generation for Wyner-Ziv video coding. A flexible side information generation scheme based on adaptive search range (ASR) and overlapped block motion compensation (OBMC) is proposed. This scheme enables search range adaptation based on the residual energy of two reconstructed frames and overcomes blocking artifact by overlapped block motion compensation. Experimental results demonstrate superiority of the proposed method over existing side information generation techniques.

Categories and Subject Descriptors
I.4.2 [Compression (Coding)]: Approximate methods
General Terms
Algorithms, Measurement, and Performance

Keywords
Wyner-Ziv video coding, side information, adaptive search range, overlapped block motion compensation

1. INTRODUCTION
The popular conventional video coding standards such as MPEG-1/2/4 and H.26x which exploit complex motion estimation at encoder cannot meet the emerging need for low power consumption and low computational complexity of encoder. Therefore, in many new applications where reduction of encoder complexity is utmost important such as in wireless PC cameras, mobile camera phones, sensor networks, or surveillance, a new video coding paradigm with low complex video encoding would be extremely valuable. Distributed video coding (DVC) is one of solutions for such low complexity encoding problems.

Wyner-Ziv video coding is one of the representative DVC schemes [1]. One of the key benefits of DVC is the complexity shift from encoder into decoder. However, this in return makes the decoder computationally extremely heavy. Therefore, it is quite difficult to apply this technique into real time applications.

At decoder side, side information generation is one of the main blocks that cause high complexity and directly affect quality of final reconstructed Wyner-Ziv (WZ) frame. There have been lots of efforts aiming to improve the quality of side information. The work in [2] improves the functionality by introducing a new hierarchical motion estimation technique using local affine motion models and a new motion estimation technique where the smoothness of the motion field is explicitly modeled. In [3], an improved side information generation is presented by applying variable block size to Y, U, and V components motion estimation, and side matching is exploited to improve side information in [4]. However, none of them considers the complexity of motion estimation and the flexibility of DVC codec with different types of sequences because in these literatures, parameters of DVC codec are always fixed. By dealing with these problems, this paper proposes a novel side information scheme with flexibility as well as low complexity. This new scheme describes search range adaptation technique to solve the problem of time consuming and makes the flexibility at side information generation block.

Moreover, the overlapped block motion compensation technique [5] to deal with blocking artifact problem is also applied. In contrast to [3], the OBMC applied in this paper is specified by an overlapped window size (OWS) which can be extended to all neighboring blocks.

The rest of this paper is organized as follows: Section 2 briefly describes side information generation based on motion compensation temporal interpolation (MCTI) schemes. In Section 3, the proposed side information generation scheme is introduced. Experimental results and discussions are presented in Section 4. Finally, in Section 5, some conclusions are drawn.

2. SIDE INFORMATION GENERATION FOR DISTRIBUTED VIDEO CODING
Side information for Distributed Video Coding can be created by several methods, such as frame extrapolation where a forthcoming frame is estimated based on past reference frame(s) [6], or frame interpolation where a frame is estimated based on neighboring frames using both past and future reference (decoded) frames [2]. This paper adopts the interpolation approach since higher rate-distortion performance can be achieved. The technique applied into frame interpolation is called MCTI (Motion
Compensation Temporal Interpolation) (Figure 1). In this technique, a frame is divided into non-overlapped blocks; a block matching estimation and compensation are used to interpolate side information frame. These techniques will be briefly described below:

2.1 Forward motion estimation
In forward motion estimation (ME) algorithm, the motion vectors are computed between the previous frame \( \tilde{X}_{n-1} \) and the following frame \( \tilde{X}_{n+1} \) (see Figure 2 (a)). The parameters that characterize this motion estimation technique are search range, block size, and motion search pixel accuracy. For a given current block, a reference block is obtained by full search method.

2.2 Bidirectional motion estimation with linear trajectory
The bidirectional motion estimation refines the motion vector obtained in the previous step by using linear trajectory assumption (see Figure 2(b)). Here the search range is confined to a small displacement around the initial block position and the motion vectors between the interpolated frame and previous and next key frames are assumed symmetric.

\[
\tilde{X}_{n-1} \xrightarrow{\text{Forward Motion Estimation}} \text{Bidirectional Motion Estimation} \xrightarrow{\text{Spatial Motion Smoothing}} \text{Motion Compensation} \xrightarrow{\text{OBMC}} Y_n
\]

Figure 1. Motion Compensation Temporal Interpolation

\[
\tilde{X}_{n-1} \rightarrow \text{Forward Motion Estimation} \rightarrow \text{Bidirectional Motion Estimation} \rightarrow \text{Spatial Motion Smoothing} \rightarrow \text{Motion Compensation} \rightarrow Y_n
\]

Figure 2. (a) Forward ME; (b) Bidirectional ME

2.3 Spatial motion smoothing
After bidirectional motion estimation, it is observed that the motion vectors have sometimes low spatial coherence; this can be improved by spatial smoothing algorithm. In this paper we use the spatial smoothing using weighted vector median filtering.

2.4 Motion compensation
Generally, the bi-directional motion compensation is exploited to interpolate side information \( Y(x, y) \). The equation describes interpolation process as follows:

\[
Y(x, y) = (\tilde{X}_{x, y} + \tilde{X}_{x, y})(x - \tilde{X}_{mx, y}x, y - \tilde{X}_{my, y}y)
\]

where: \( (\tilde{X}_{mx, y}, \tilde{X}_{my, y}) \) is the forward motion vector.

3. PROPOSED METHOD
In the existing methods, the search range in forward motion estimation is always fixed at a high value; ±32 pixels are used for search range in DISCOVER codec [7]. This fixed value of search range contributes to the high complexity in side information generation but not necessarily brings the best quality for side information as verified in next sub-section. Additionally, the block matching algorithm used in motion estimation creates the blocking artifact as marked in Figure 6(a) and Figure 6(c). To overcome these problems, we propose a new scheme having two differences compared to those existing side information generation schemes (see Figure 3). First, an adaptive search range (ASR) is exploited to control the search range in the forward ME; and second, instead of using non-overlapped block motion compensation as [2][7], we apply the overlapped block motion compensation into this scheme that makes the scheme more effective in terms of reducing blocking artifacts. In sub-sections 3.1 and 3.2, we describe these techniques in detail.

3.1 Adaptive search range (ASR)
ASR is one of the key techniques to deal with the computational complexity problem of motion estimation. This technique has been applied into conventional video coding such as H.264 or MPEG-4. It is exploited at encoder side in H.264 to reduce the time complexity while retaining the quality of reconstructed frame. There are three main approaches to make an adaptation of search range. Those are the exploiting the summation of absolute value of previously computed motion vectors and summation of prediction error [8]; using local statistics of motion vectors (MVs) of neighboring blocks in [9]; and exploiting the motion vector difference (MVD) to decide search range as in [10].

However, we cannot directly use the motion vectors from previous interpolated frame to control search range in DVC because there is no original frame available at decoder side. Moreover, in DVC, the search range increment is not necessarily
proportional to the quality of side information. In Figure 4, the PSNR of side information with variable search ranges verifies this assertion. Therefore, the motion vector is not a reliable criterion to use.

To apply ASR into side information scheme for DVC, we only use the sum of absolute difference (SAD) between two reconstructed key frames \( \tilde{X}_{n-1} \) and \( \tilde{X}_{n+1} \) at search range that is set by zero. The reason for this choice is the SAD between \( \tilde{X}_{n-1} \) and \( \tilde{X}_{n+1} \) reflects the degree of motion inside the interpolated frame. Therefore, we can use SAD to evaluate the degree of motion inside a frame.

A frame can be divided into two kinds based on motion: fast motion regions and slow motion regions. We assign a larger search range for the fast motion regions and a smaller search range for the slow motion regions.

![Figure 4. Side information generation with different search ranges](image)

To determine slow motion and fast motion regions, a threshold \( T \) is specified: if SAD is larger than \( T \), we assume fast motion region, otherwise, we assume slow motion region. Here, the candidate search ranges are 4x4 and 8x8 respectively corresponding to the slow and fast motion regions.

### 3.2 Overlapped block motion compensation (OBMC)

In order to construct the side information frame, interpolation is performed using the bi-directional motion compensation as (1). Here, non-overlapped block motion compensation is used and it is very sensitive to blocking artifact because one of this technique’s drawbacks is that the discontinuities at the block boundaries are visible. Moreover, when motion vectors are not correct or motion vectors of neighboring blocks are significantly uncorrelated, overlap pixels and hole regions will be created as shown in Figure 6(a) and Figure 6(c). To overcome these problems, OBMC [5] is applied into the side information generation scheme. The region of overlap is specified by overlapped window size (OWS). When OWS = 12x12, the overlap window can be divided into nine sub-blocks corresponding to nine regions (see Figure 5).

Suppose the top left four neighboring blocks have separate motion vector \((mv_{x,1}, mv_{y,1}) \) \( i \in \{1, 2, 3, 4\} \).

![Figure 5. Block overlapping pattern in OBMC](image)

For pixels in region A:

\[
Y(x, y) = \left( \sum_{i=1}^{4} \{ \tilde{X}_{n-1}, (x + mv_{x,1} + y + mv_{y,1}) + \tilde{X}_{n+1}, (x - mv_{x,1} - y - mv_{y,1}) \} \right) / 8
\]  

(2)

For pixels in region D:

\[
Y(x, y) = \left( \sum_{i=3}^{4} \{ \tilde{X}_{n-1}, (x + mv_{x,3} + y + mv_{y,3}) + \tilde{X}_{n+1}, (x - mv_{x,3} - y - mv_{y,3}) \} \right) / 4
\]  

(3)

For pixels in region E:

\[
Y(x, y) = \left( \tilde{X}_{n-1}, (x + mv_{x,1} + y + mv_{y,1}) + \tilde{X}_{n+1}, (x - mv_{x,1} - y - mv_{y,1}) \right) / 2
\]  

(4)

Pixels in region B, C, F, G, H, and I are calculated in the same way.

### 4. EXPERIMENTAL RESULTS

To evaluate efficiency of the proposed scheme, sequences Foreman, Hall Monitor, Stefan, and Soccer are test with each of additional blocks under test conditions that: 74 frames, spatial resolution QCIF, and temporal resolution 15Hz. Group of Picture (GOP) is 2 and block size is 8x8. In the experiment, different quantization [7] is applied to obtain eight rate-distortion points as depicted in Table 1. The threshold \( T \) is empirically set to 5800.

In the case of ASR block, Table 2 and Table 3 describe time savings and PSNR of side information for the case of Foreman and Soccer sequences.

![Table 1. Quantization parameter setting for key frame](image)

*Qm: quantization matrix number for WZ frame
Table 2. PSNR of Side Information (PSNR_SI) and Time saving (TS) with ASR (Foreman, QCIF, @15Hz)

<table>
<thead>
<tr>
<th>Qp</th>
<th>PSNR_SI [dB]</th>
<th>Time (s)</th>
<th>PSNR_SI [dB]</th>
<th>Time(s)</th>
<th>TS (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>29.06</td>
<td>128</td>
<td>29.00</td>
<td>53</td>
<td>59</td>
</tr>
<tr>
<td>32</td>
<td>28.11</td>
<td>128</td>
<td>28.08</td>
<td>53</td>
<td>59</td>
</tr>
<tr>
<td>38</td>
<td>26.76</td>
<td>128</td>
<td>26.74</td>
<td>53</td>
<td>59</td>
</tr>
<tr>
<td>40</td>
<td>26.10</td>
<td>126</td>
<td>26.16</td>
<td>52</td>
<td>58.7</td>
</tr>
</tbody>
</table>

*Qp: quantization parameter for key frame

The time saving is calculated as follows:

\[ TS = \frac{Time[S\ R = 12 \times 12] - Time[proposed]}{Time[S\ R = 12 \times 12]} \times 100 \% \quad (5) \]

The time at Search range (SR = 12x12) is chosen as anchor to calculate time saving because it gives the highest PSNR for all quantization parameters (see Figure 4).

Table 3. PSNR of Side Information (PSNR_SI) and Time saving (TS) with ASR (Soccer, QCIF, @15Hz)

<table>
<thead>
<tr>
<th>Qp</th>
<th>PSNR_SI [dB]</th>
<th>Time (s)</th>
<th>PSNR_SI [dB]</th>
<th>Time(s)</th>
<th>TS (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>21.76</td>
<td>118</td>
<td>21.57</td>
<td>53</td>
<td>55</td>
</tr>
<tr>
<td>34</td>
<td>21.56</td>
<td>119</td>
<td>21.44</td>
<td>51</td>
<td>57</td>
</tr>
<tr>
<td>41</td>
<td>21.37</td>
<td>117</td>
<td>21.29</td>
<td>50</td>
<td>57</td>
</tr>
<tr>
<td>44</td>
<td>21.32</td>
<td>117</td>
<td>21.22</td>
<td>50</td>
<td>57</td>
</tr>
</tbody>
</table>

When OBMC is applied in our scheme, the blocking artifact is significantly reduced as depicted in Figure 6(b) and Figure 6(d). Figure 7 draws the Rate-Distortion (R-D) performance of DVC codec with and without OBMC in the case the ASR function is turn off. The PSNR of reconstructed WZ frames can gain up to 0.34 dB in the Soccer sequence. However, in Hall Monitor sequence, OBMC seems not improve R – D performance. Based on these above results, two important conclusions are taken:

- OBMC is very efficient in term of subjective quality. The reason is due to a lot of blocking artifacts have solved when applied OBMC.
- OBMC is more suitable for high motion sequences such as Stefan or Soccer than slow motion sequence like Hall monitor. Because there are more blocking artifact in high motion regions than that in slow motion regions. Therefore, when applied OBMC into side information generation, we can significantly remove blocking artifact and make the side information better.

5. CONCLUSIONS

In this paper, a flexible side information generation is introduced for DVC. It overcomes the problem computational complexity at decoder. This is obtained by using an adaptive search range block. Moreover, this scheme also solves the problem of blocking artifact by using overlapped block motion compensation. Experimental results show that the proposed scheme not only saves processing time at side information generation up to 59% but also gains by 0.34 dB in PSNR for the WZ frames and efficiently reduces blocking artifact when OBMC is applied. For the future works, we will create side information with another technique to remove blocking artifact such as de-blocking filter and compare to OBMC in term of coding and objective efficiency.

6. ACKNOWLEDGMENTS

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (No. 20100008030).

7. REFERENCES


Figure 7. R – D performance with and without OBMC: Only WZ frames