Efficient Motion Vector Composition for H.264/AVC to SVC Video Transcoding

Yung-Hsiang Tang\textsuperscript{1}, Gwo-Long Li\textsuperscript{2}, and Mei-Juan Chen\textsuperscript{1}

\textsuperscript{1}Dept. of Electrical Engineering, National Dong Hwa University, Hualien, Taiwan
\textsuperscript{2}Dept. of Video Coding Core Technology, Industrial Technology Research Institute, Hsinchu, Taiwan

Abstract - To reduce the transcoding overhead between H.264/AVC and Scalable Video Coding, this paper proposes an efficient motion vector composition algorithm combined with adaptive search range decision. In our proposal, the decoded motion vectors of H.264/AVC are reused to derive the motion vectors in compliance with the hierarchical B prediction structure of Scalable Video Coding. Furthermore, the decoded motion vector and residual information are also considered to decide the search range for refining the composed motion vectors so that the composite motion vectors can be as accurate as possible. Simulation results demonstrate that our proposed transcoding algorithm can achieve better rate distortion performance and higher transcoding time reduction compared to previous work.

Keywords: Transcoding, H.264/AVC, SVC, Motion Vector.

1 Introduction

H.264/AVC to Scalable Video Coding (SVC)[1] transcoding is one of the interesting topics since it can convert H.264/AVC bitstream to SVC bitstream for satisfying the application heterogeneities. However, the decoding and re-encoding process consumes lots of computational complexity especially for the motion estimation. In order to reduce computational complexity of transcoding, the method proposed in [2] uses the decoded motion vectors of H.264/AVC to reduce search range for accelerating the motion estimation process. However, the computational complexity would be high if the reference picture far from the current picture. The method proposed in [3] calculates the forward motion vectors of hierarchical B prediction structure by linear translation but lacks of composing backward motion vectors. This paper proposes an efficient motion vector composition algorithm to derive both of forward and backward motion vectors for SVC in compliance with hierarchical B prediction structure. In addition, a computational complexity reduced motion vector refinement algorithm is also proposed to further improve the accuracy of composed motion vectors.

2 Proposed Algorithm

Fig. 1 shows the flowchart of proposed algorithm for H.264/AVC to SVC video transcoding. Our proposed algorithm is mainly composed by two phases. The first phase is the motion vector composition which is used to generate the absent motion vectors just following the hierarchical B prediction structure. The second phase is the motion vector refinement which is used to refine the composed motion vectors so that the transcoding results can be as better as possible.

2.1 Motion Vector Composition

To support temporal scalability, the prediction structure of H.264/AVC has to be changed first to satisfy the hierarchical B prediction structure of SVC. Fig. 2 shows the prediction architecture of H.264/AVC and SVC on the upper and bottom parts, respectively. It should be noticed that only the prediction structure of IPPP is treated in this paper. As shown in Fig. 2, we can find that only the motion vectors pointing to previous frame can be obtained from the decoded H.264/AVC bitstream. Therefore, the main goal of the proposed motion vector composition algorithm is to compose the motion vectors pointing to any reference frame with any frame distance as SVC prediction structure shown. Therefore, based on the smooth moving property between successive frames, the composite motion vectors for SVC can be calculated by

\[ M_{n}^{F} = \sum_{i=0}^{g-1} M_{n-i}^{F} \]  

\[ M_{n}^{B} \] and \[ M_{n}^{F} \] are the H.264/AVC decoded forward motion vectors and SVC forward motion vectors, respectively. \( G \) and \( l \) are the group of picture (GOP) size and the temporal layer index. However, for the backward motion vectors, H.264/AVC decoded motion vectors with IPPP prediction structure don’t contain the motion vectors in backward direction. The proposed motion vector composition algorithm reuses the forward motion vectors decoded from H.264/AVC to compose the backward motion vectors of SVC as shown in following equations.
\[ MV_h^b = -MV_h^f \]  
\[ MV_h^b = \sum_{i=0}^{g} MV_{h+i} \]

Although Eq.(1) to Eq.(3) can be used to compose the corresponding motion vectors both in forward and backward directions to satisfy the prediction structure of SVC, the problem will be faced during the motion vectors selection process due to the variable block size issue as shown in Fig. 3. When composing the corresponding motion vectors, our proposed algorithm needs to pick one motion vector from reference frame in order to derive the composed motion vectors. However, for the motion vector in frame \( n \) pointing to previous frame \( n-1 \), it might cover several blocks which have their own motion vectors. Therefore, our proposed algorithm needs to decide which motion vector will be selected to be used for motion vector composition. In our notation, the motion vector in frame \( n \) is denoted as \( MV_c \) and the motion vectors covered by \( MV_c \) are denoted as \( MV_c \). In our proposed algorithm, the motion vector selection mechanism is based on the minimum motion vector distance between the \( MV_c \) and \( MV_f \) as shown in Eq.(4).

\[ MV_f^c = MV_c \min ||MV_c-MV_f||, MV_c \in \{ \text{covered}\ MVs\} \]  

![Fig. 2 The prediction structure of H.264/AVC on the upper part and SVC on the bottom part](image)

![Fig. 3 The relationship between \( MV_f \) and \( MV_c \)](image)

### 2.2 Motion Vector Refinement

In the proposed motion vector composition algorithm, the absent motion vectors will be composed to derive the motion vectors required for supporting temporal scalability in SVC. However, the composed motion vectors could not be the best motion vectors for SVC. Therefore, a motion vector refinement algorithm is further proposed to improve the accuracy of composed motion vectors. Fig. 1 exhibits the detailed operations of our proposed motion vector refinement algorithm. In our proposal, the motion vector intensity (MVI), averaged motion vector (\( MV_{Avg} \)), and the sum of absolute residuals (R) are calculated to determine the motion behavior of the current macroblock. If both conditions of \( R \geq 255 \) and \( MVI > MV_{Avg} \) are satisfied, the full search motion vector refinement process will be applied to refine the composed motion vectors around the motion vector predictor (MVP) of the current macroblock. Otherwise, a new search range shown in below will be decided to refine the composed motion vectors so that the overhead of motion vector refinement can be lighten when compared to full search motion vector refinement.

\[ SR_{Adaptive} = \begin{cases} ||MV_P-MV_f||, & \text{forward} \\ ||MV_P-MV_f||, & \text{backward} \end{cases} \]

For spatial scalability, the motion vectors composed by our proposed algorithm in base layer will be directly used in spatial enhancement layer by simply using the upsampling mechanism.

### 3 Simulation Results

The proposed algorithm is evaluated by test sequences (200 frames for each sequence) with CIF and 4CIF resolutions for GOP with 8 frames compressed by the H.264/AVC reference software JM 18.0 and JSVM 9.18. Table 1 shows the comparison for the proposed algorithm with previous work [2]. From this table, we can find that our proposed algorithm can achieve better rate distortion performance as well as coding time reduction. On average, our proposed algorithm can further reduce the coding time 6.92% on average when compared to [2].

<table>
<thead>
<tr>
<th>Sequence</th>
<th>BD-PSNR(dB)</th>
<th>BD-Bitrate(%)</th>
<th>TR(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Akiyo</td>
<td>-0.06</td>
<td>1.61</td>
<td>1.52</td>
</tr>
<tr>
<td>Container</td>
<td>-0.12</td>
<td>4.91</td>
<td>13.6</td>
</tr>
<tr>
<td>Foreman</td>
<td>-0.05</td>
<td>1.23</td>
<td>1.41</td>
</tr>
<tr>
<td>Stefan</td>
<td>-0.24</td>
<td>4.74</td>
<td>0.63</td>
</tr>
<tr>
<td>Harbour</td>
<td>-0.01</td>
<td>0.34</td>
<td>0.64</td>
</tr>
<tr>
<td>Crew</td>
<td>-0.03</td>
<td>0.96</td>
<td>0.98</td>
</tr>
<tr>
<td>Soccer</td>
<td>-0.04</td>
<td>1.37</td>
<td>1.10</td>
</tr>
<tr>
<td>Average</td>
<td>-0.08</td>
<td>2.17</td>
<td>1.04</td>
</tr>
</tbody>
</table>

### 4 Conclusions

This paper proposes a fast H.264/AVC to SVC transcoding algorithm by using the motion vector composition approach. In addition, a low computational complexity motion vector refinement algorithm is also proposed to further improve the accuracy of composite motion vectors. Through the proposed algorithm, the transcoding time can have better saving with high rate-distortion performance compared to previous work [2].

### 5 References

