ADAPTIVE SEARCH RANGE SELECTION IN MOTION ESTIMATION

Chung-Cheng Lou, Szu-Wei Lee and C.-C. Jay Kuo

Ming Hsieh Department of Electrical Engineering and Signal and Image Processing Institute
University of Southern California, Los Angeles, CA 90089-2564, USA
E-mails: clou@usc.edu, zuweile@usc.edu, and cckuo@sipi.usc.edu

ABSTRACT

An adaptive search range (SR) selection algorithm in motion estimation for video encoding is proposed in this work with an objective to achieve excellent coding performance with reduced memory access bandwidth. In modern fast motion search algorithms, a motion vector predictor (MVP) is obtained from the motion vectors (MVs) of spatially and temporally neighboring blocks, which is called the predictive MV set. By exploiting the relationship between the variance of the predictive MV set and the SR, we develop an adaptive SR selection algorithm. That is, a larger variance implies lower accuracy of the MVP and, thus, a larger SR. The superior performance of the proposed SR selection algorithm is demonstrated by experimental results.

Index Terms— Motion estimation, motion vector predictor, adaptive search range selection.

1. INTRODUCTION

Motion estimation (ME) is the most computationally intensive module in a video encoding system. It has been reported by researchers [1][2] that 30-40% of total execution time and 90% of total memory access are dedicated to ME. For a target block, ME searches for the block in the reference frame that yields the minimal rate distortion (RD) cost. Since pixels of reference blocks have to be read to compute the RD cost, the computational complexity and the memory access bandwidth requirements are high. They are particularly demanding in resource-limited embedded systems such as mobile devices. In this work, we propose an adaptive search range (SR) selection algorithm to reduce the memory access bandwidth. The objective is to achieve excellent coding performance with reduced memory access bandwidth.

A typical ME algorithm consists of two steps. First, it predicts the initial position of the motion vector (MV) using a motion vector predictor (MVP). Research on fast MV search has been extensively conducted in the last two decades [3], [4]. State-of-the-art algorithms exploit the spatial or temporal correlation between neighboring blocks to obtain the MVP. Second, it searches pixels around the MVP to find the optimal MV that minimizes a certain cost function. The size of the search window is called the SR. Traditionally, a fixed SR is adopted for all macroblocks. More recently, the adaptive SR selection problem begins to receive some attention. The SR was selected as the maximum of spatially neighboring MVs in [5]. The SR was determined by the absolute difference between the spatial and temporal MVPs in [6]. The SR was expressed as a function of the MVP of the target block and the differences between MVs and MVPs of its neighboring blocks in [7]. However, most previous work only considered a squared search region, ignoring the fact that the SR can be rectangular by treating the horizontal SR and the vertical SR separately. Besides, for a given SR, no analysis has been conducted about the probability for the corresponding search window to contain the optimal MV.

Intuitively speaking, the SR is highly dependent on the accuracy of the MVP. If it is accurate, the distance between the MVP and the optimal MV is small, which guarantees a smaller SR. Based on this idea, a preliminary study on adaptive SR selection was reported in [8], which will be briefly reviewed in Sec. 2. As an extension of [8], we address the following important issues in this work. First, we propose an accurate MVP algorithm based on the MVs of spatially and temporally neighboring blocks jointly, which is called the predictive MV set. Second, we compute the variance of MVs in the predictive set and relate it to the SR selection. If the variance is smaller, the accuracy of MVP is higher so that the SR can be smaller. Third, we analyze the probability for a search window of a given SR to contain the optimal MV.

The rest of this paper is organized as follows. Research background is presented in Sec. 2. The accuracy of the temporal and the spatial MVPs is analyzed in Sec. 3, which suggests an approach to consider them jointly. Then, the relationship between the variance of the predictive MV set and the SR is investigated, and an adaptive SR selection algorithm is proposed in Sec. 4. Experimental results are provided in Sec. 5 to demonstrate the superior performance of the adaptive SR selection algorithm. Finally, concluding remarks are given in Sec. 6.

2. RESEARCH BACKGROUND

The relationship among the MVP, the optimal MV, and the SR is shown in Fig. 1. The difference between the optimal MV and the MVP is called the motion vector difference (MVD). Since pixels of reference blocks centered at each search position have to be loaded to compute the RD cost, a larger SR implies higher memory access bandwidth and latency. On the other hand, if the SR is too small and the optimal MV falls outside the SR, we may either sacrifice the coding efficiency by selecting a sub-optimal MV within the SR or enlarge the SR again. Neither is desirable. Here, we would like to consider an adaptive SR selection algorithm, which depends on the quality of the MVP. That is, if the MVP is likely to be accurate, we can use a small SR.

Two MVPs were examined separately in [8]; namely, the spatial MVP ($MVP_s$) and the temporal MVP ($MVP_t$), where $MVP_s$ is the median MV of spatially neighboring blocks while $MVP_t$ is the collocated MV of the previous frame. Then, the difference of these two MVPs is calculated. If their difference is small, it implies that $MVP_s$ and $MVP_t$ are consistent. Consequently, any of them should be accurate and the MVD is expected to be small, which al-
An accurate MVP is critical to the SR reduction. In this section, we first analyze two common MVPs (i.e., \( MVP_s \) and \( MVP_t \)) and show that each of them is not able to provide a very accurate estimate of the optimal MV (\( MVP_{op} \)) alone. Then, we propose a joint spatial-temporal MVP.

To illustrate the shortcoming of \( MVP_s \) and \( MVP_t \), we use the CIF Foreman sequence (300 frames, 30 fps and \( QP = 28 \)) as an example. We compute their MVDs as

\[
MVD_s = MVP_{op} - MVP_s \\
MVD_t = MVP_{op} - MVP_t,
\]

and plot the 2D histograms of their x-component and the y-component in Figs. 3(a) and (b), respectively. We see that the probability density is very much concentrated at the origin, which means that either \( MVP_s \) or \( MVP_t \) is accurate in these cases. For a bin that is closer to the y-axis than the x-axis, it means that the component of \( MVP_s \) is smaller so that it provides a better estimate than the corresponding component of \( MVP_t \). For bins along the 45-degree line, \( MVP_s \) and \( MVP_t \) have the same prediction performance. The prediction \( MVP_t \) is accurate for bins on the x-axis while the prediction \( MVP_s \) is accurate for bins on the y-axis.

![Fig. 3: The 2D histograms of MVD_s and MVD_t: (a) the x-component and (b) the y-component.](image)

To cover all of the above three cases (i.e., the x-axis, the y-axis and the 45-degree line), it is natural to conclude that a better MVP should take both the spatial and temporal predictions into account. This motivates us to examine a joint MV predictive set. As shown in Fig. 4, the spatial MV candidate set contains four neighboring MVs, denoted by \( N_s = \{mv_{x_0}, mv_{x_1}, mv_{x_2}, mv_{x_3} \} \), and the temporal MV candidate set contains nine MVs of the previous frame denoted by \( N_t = \{mv_{t_0} \sim mv_{t_8} \} \). Then, the joint spatial-temporal MV predictive set can be defined as

\[
N_{st} = N_s \cup N_t,
\]

Then, we can select the joint spatial-temporal MVP via

\[
MVP_{st} = \text{Mean}(N_{st}).
\]

It is observed that \( MVP_{st} \) is more accurate than \( MVP_s \) and \( MVP_t \). Besides, variances of MVDs for \( MVP_{st} \) are much smaller than those of \( MVP_s \) and \( MVP_t \). The performance of the new MVP predictor, \( MVP_{st} \), will be shown in Sec. 5.

### 3. JOINT SPATIAL-TEMPORAL MVP

An accurate MVP is critical to the SR reduction. In this section, we first analyze two common MVPs (i.e., \( MVP_s \) and \( MVP_t \)) and show that each of them is not able to provide a very accurate estimate of the optimal MV (\( MVP_{op} \)) alone. Then, we propose a joint spatial-temporal MVP.

To illustrate the shortcoming of \( MVP_s \) and \( MVP_t \), we use the CIF Foreman sequence (300 frames, 30 fps and \( QP = 28 \)) as an example. We compute their MVDs as

\[
MVD_s = MVP_{op} - MVP_s, \\
MVD_t = MVP_{op} - MVP_t,
\]

and plot the 2D histograms of their x-component and the y-component in Figs. 3(a) and (b), respectively. We see that the probability density is very much concentrated at the origin, which means that either \( MVP_s \) or \( MVP_t \) is accurate in these cases. For a bin that is closer to the y-axis than the x-axis, it means that the component of \( MVP_s \) is smaller so that it provides a better estimate than the corresponding component of \( MVP_t \). For bins along the 45-degree line, \( MVP_s \) and \( MVP_t \) have the same prediction performance. The prediction \( MVP_t \) is accurate for bins on the x-axis while the prediction \( MVP_s \) is accurate for bins on the y-axis.

![Fig. 3: The 2D histograms of MVD_s and MVD_t: (a) the x-component and (b) the y-component.](image)

To cover all of the above three cases (i.e., the x-axis, the y-axis and the 45-degree line), it is natural to conclude that a better MVP should take both the spatial and temporal predictions into account. This motivates us to examine a joint MV predictive set. As shown in Fig. 4, the spatial MV candidate set contains four neighboring MVs, denoted by \( N_s = \{mv_{x_0}, mv_{x_1}, mv_{x_2}, mv_{x_3} \} \), and the temporal MV candidate set contains nine MVs of the previous frame denoted by \( N_t = \{mv_{t_0} \sim mv_{t_8} \} \). Then, the joint spatial-temporal MV predictive set can be defined as

\[
N_{st} = N_s \cup N_t,
\]

Then, we can select the joint spatial-temporal MVP via

\[
MVP_{st} = \text{Mean}(N_{st}).
\]

It is observed that \( MVP_{st} \) is more accurate than \( MVP_s \) and \( MVP_t \). Besides, variances of MVDs for \( MVP_{st} \) are much smaller than those of \( MVP_s \) and \( MVP_t \). The performance of the new MV predictor, \( MVP_{st} \), will be shown in Sec. 5.

### 4. ADAPTIVE SR SELECTION ALGORITHM

In this section, we first analyze the probability for the optimal MV to fall in the search window for fixed horizontal and vertical search ranges (SRs) in Sec. 4.1. Then, we study how to select SRs to meet a performance criterion.
4.1. Probability of Optimal MV within Search Window

The variance $\sigma_{st}^2$ of the predictive MV set, $N_{st}$, consists of the consistency of MV prediction. When it is small, all MVP candidates are similar so it is likely that MV_opt is close to MVP_st. Thus, a smaller SR can be used. To support this claim, we plot the variance of $MVD_{st}$ as a function of $\sigma_{st}$ in Fig 5 for the CIF Foreman sequence (30fps). We see from the figures that the variance of $MVD_{st}$ is proportional to $\sigma_{st}$. As a result, a larger $\sigma_{st}$ implies a larger variance of $MVD_{st}$, and hence, a larger SR value.

![Fig. 5](image-url)  
Fig. 5: The variance of $MVD_{st}$ as a function of $\sigma_{st}$: (a) the x-component and (b) the y-component.

Furthermore, for a given SR, we would like to quantify the probability for the optimal MV to fall within a search window. By using the modified-Cauchy function in Eq. (1) as the probability density function of the $MVD_{st}$, then, the probability of the MV_opt within $(SR_x, SR_y)$ can be found via

$$F_{MC}(SR_x, SR_y) = F_{MC,X}(SR_x) \cdot F_{MC,Y}(SR_y),$$

$$F_{MC,T}(SR_t) = \int_{-SR_t-0.5}^{SR_t+0.5} f_{MC,T}(t) dt, \ t \in \{x, y\}.$$  

Model parameters $(\zeta_x, \zeta_y)$ can be estimated from $\sigma_{st}$. Once these parameters are determined, Eq. (5) can be used to decide the SR.

4.2. Performance-Controlled SR Selection

It is often the case that we would like to define the probability of the optimal MV to fall into the desired window and use it to select the proper $(SR_x, SR_y)$. This problem can be formulated as the following constrained optimization problem:

$$\min J(SR_x, SR_y) = (2SR_x + 1)(2SR_y + 1)$$

subject to $F_{MC}(SR_x, SR_y) = T_{prob}$

where $T_{prob}$ is the user-defined probability constrain. This problem can be converted into an unconstrained optimization problem by adding a Lagrange multiplier; namely,

$$J'(SR_x, SR_y) = (2SR_x + 1)(2SR_y + 1) + \lambda (F_{MC}(SR_x, SR_y) - T_{prob})$$  

To minimize Eq. (6), one can derive the following condition:

$$F_{MC,X}(SR_x) = F_{MC,Y}(SR_y) = \sqrt{T_{prob}}$$  

To save the complexity, the computation of $F_{MC}(SR_x, SR_y)$ is implemented by table look-up and interpolation in our experiment. Besides, since the x- and y-components of $MVD_{st}$ may be different, we may get different $SR_x$ and $SR_y$ so that the search window may not be a squared region. Please note that $\sigma_{st}$ could be zero if all candidate MVs in $\{N_{St}, N_{T}\}$ are the same. To compute model parameters $(\zeta_x, \zeta_y)$ in this case, we set the minimal variance $\sigma_{min}$ to 0.07, which comes from the smallest possible variance when only one MV differs from other MVs in the candidate set.

5. EXPERIMENTAL RESULTS

The proposed adaptive SR selection algorithm was implemented in the baseline profile of the H.264/AVC JM15.1 reference codes to test its performance. We compare the prediction accuracy of $MVP_{st}$, $MVP_{avg}$, and $MVP_{T}$ in Table 1, where the Foreman CIEF sequence was encoded with 300 frames, 30 fps, $Q_p = 28$ and different MVPs were chosen as the center of the search window. The maximum search window is set to $[\pm32, \pm32]$. Table 1 shows that the variance of $MVD$ is the lowest when we select $MVP_{st}$ as the search center. Besides, the variance along the horizontal direction is higher than the one along the vertical direction, which justifies the selection of a rectangular search window.

<table>
<thead>
<tr>
<th></th>
<th>$MVP_{st}$</th>
<th>$MVP_{avg}$</th>
<th>$MVP_{T}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>x-component</td>
<td>20.609</td>
<td>21.282</td>
<td>34.629</td>
</tr>
<tr>
<td>y-component</td>
<td>15.298</td>
<td>16.226</td>
<td>23.759</td>
</tr>
</tbody>
</table>

Table 1: Comparison of the variance of $MVD$ with different MVPs as the search window center.

We compare the RD performance and the average search window size of the fixed SR $\{[\pm32, \pm32]\}$ and the adaptive SR selection algorithms in Figs. 6a and 6b, respectively, where the CIEF Foreman sequence was encoded by $Q_p = \{20, 24, 28, 32, 36, 40, 44\}$. The user-defined probability constraint $T_{prob}$ was set as 94%. We see that the adaptive SR algorithm provides almost the same RD performance as the fixed SR algorithm. The average SR size of the adaptive SR algorithm is however only 1.5% of the fixed SR algorithm. As a result, the adaptive SR algorithm can save the memory access bandwidth significantly with no R-D performance degradation.

Three adaptive SR decision algorithms (Algorithms A [5], Algorithm B [6], and Algorithm C [7]) were implemented for performance benchmarking. Their RD performance and average search window sizes are compared in Figs. 7a and 7b, respectively. The proposed SR selection algorithm and Algorithm B have similar RD performance. The proposed algorithm outperforms Algorithms A by 1.2dB and Algorithm C by 1.5-2dB, respectively. The average search window size of the proposed algorithm is about 58.6% and 12.1% of those of Algorithm A and Algorithm B, respectively. It is however larger than that of Algorithm C. For a given input sequence and a fixed SR decision algorithm, we observe that the search window size becomes larger when the bit-rate is lower. This is due to the
Fig. 6: Performance comparison of the adaptive and fixed SR selection algorithms: (a) the R-D performance and (b) the search window size.

fact that the MVP accuracy is affected by the quantization parameter and the search window has to be larger if the MVP accuracy is lower.

6. CONCLUSION

An adaptive SR selection algorithm was proposed to reduce the search window size. It was shown to outperform several state-of-the-art benchmarking algorithms by keeping excellent RD performance while reducing the memory access bandwidth significantly.

7. REFERENCES


Fig. 7: Performance comparison of the purposed algorithm with benchmarking algorithms: (a) the R-D performance and (b) the search window size.


