Adaptive Prediction based Diamond Search Algorithm (APDS) for Video Encoder of Solar Powered UAV

Nasir Mahmud Khokhar
Department of Aeronautical Eng.,
Faculty of Engineering,
King Abdul-Aziz University,
Jeddah, Kingdom of Saudi Arabia

Wail Harasani
Department of Aeronautical Eng.,
Faculty of Engineering,
King Abdul-Aziz University,
Jeddah, Kingdom of Saudi Arabia

ABSTRACT
Unmanned Aerial Vehicle (UAV) provide bird’s eye view over an intersection or a large area, and provide real-time surveillance of area under observation. UAVs have been playing a vital role in disaster management due to the increased sensing and processing capabilities. This paper proposes a fast adaptive prediction based diamond search Motion Estimation (ME) algorithm for Sun Falcon 2, a solar powered UAV’s video encoder to cope the computational complexity, low power and increased quality of ME process requirement. Results show that the proposed Adaptive Predictive Diamond Search (APDS) ME algorithm performs best in the term of PSNR, MSE and number of Search Points (SP), for approximately all the video sequences. Moreover, performance of APDS is decreased a little bit in term of number of SP when compared to Hexagon search algorithm but its PSNR is still considerably high for those video sequences. The average PSNR improvement rate of APDS is 0.62, 2.67, 0.82, 0.83 and 2.31 for Diamond Search (DS), HexBS, FHS, FSS and MDS respectively, while the average SIR is 25.4404, 6.3374, 48.274 and 205.55 for DS, FHS, FSS and MDS respectively.

General Terms
Video compression, Motion estimation algorithm, Block matching algorithms

Keywords
Motion estimation, Motion vector, MBD, UAV, DS, FHS, MDS, HexBS,

1. INTRODUCTION
Unmanned Aerial Vehicles (UAVs) is an aircraft without the pilot inside. UAVs are famous for the versatile application areas such as: reconnaissance, surveys, military and disaster management. UAVs are equipped with a variety of sensors based on their application area. They can be deployed on various terrains within no time, and communicate with the base station in real-time. They can also be capable of providing bird’s eye view over a large area with the incorporation of cameras, and provide real-time surveillance and monitoring of area under observation.

However, some risks and difficulties are associated with the UAV for using them in different scenarios. For example, due to wind or other environmental factor, they can be unstable and there is a chance of crash, degraded image/video quality or malfunctioning of different sensors used in the UAV technology. Video cameras are the integral part of almost all the UAV’s used for surveillance applications. Video encoder sends the image/video to surveillance system by capturing snapshots through video camera. The quality and speed of transmitted video depends upon video camera as well as video encoder. The video encoder is a compression standard like H.264/AVC, that has a wide range of application from low data rate, like video streaming to High Definition Television (HDTV) broadcast. The video encoder compresses the videos through Motion Estimation (ME) process that reduces temporal redundancy between two or more successive video frames. Many fast ME algorithms are proposed in the literature. ME algorithms are used to compute the displacement between reference and current frame in the video encoder. Normally, previous frame is considered as a reference frame. The luminance of pixel in the current frame and reference frame has some correlation with its neighborhood, which is used to determine the best matching position of pixels intensity values of current frame in reference frame. Where accurate match is found, difference in the positions of current and reference frame is calculated which is called Sum of Absolute Difference (SAD). The displacement vector for this SAD is defined as the Motion Vector (MV) [1–2]. The current frame is divided into non-overlapping blocks like 16x16, 16x8, 8x16, 8x8 etc. to compute MVs [3]. These blocks are called Macro Blocks (MB). One 720p frame consists of 3600 MBs, each having a size of 16x16. One best MV is calculated in the reference frame for every MB of current frame [4]. The best match or MV is calculated by using a cost function in BMME algorithms. This cost function is based on a Minimum Block Distortion (MBD). The cost function may be the SAD or Mean Absolute Differences (MAD) and Mean Squared Error (MSE) [5, 6].

The ME is the most expensive part of video compression process in term of computational complexity and it takes greater than fifty percent of overall encoding time when consider only one reference frame. Moreover, computational cost increases relatively as the number of reference frames increases gradually. For example, for four reference frames computational cost of ME is about 70% of encoding cost [5]. Due to such high computational cost, this field has gone through a tremendous research activity. There are many type of ME algorithms such as Three Step Search (TSS) [7], New Three Step Search (NTSS) [8], Simple and Efficient Search (SES) [9], Four Step Search (FSS) [10], Diamond Search (DS) [11], Hardware Modified Diamond Search (MDS) [12], Adaptive Rood Pattern Search (ARPS) [13] and Cross Diamond Search (CDS) [14]. Moreover, a merge and split procedures based fast variable block-size motion estimation algorithm is proposed in [15] for MPEG-4/ H.264 and to increase the coding efficacy of latest video encoder like ITU-T H.263and MPEG-4 visual, a multi-frame prediction based algorithm is proposed in [16]. These all algorithms have their own merits and demerits according to different motion.
scenarios like some perform better for low motion activity video sequence or some for high motion with good PSNR. Therefore by keeping in mind the videos from high altitude, low altitude with motion of video acquiring source and power requirement of encoder we propose a new Adaptive Prediction-based Diamond Search (APDS) ME algorithm for UAV encoder. Moreover, we tested the proposed algorithm with HD videos (720p) as well as video acquired from UAVs. The simulation results shows that the proposed APDS algorithm has the highest PSNR for all the video sequence. In some cases, its number of search points has increased a little bit then the HexBS, but, for the last two video sequences having aerial view of UAV, its number of search point is still the lowest than all other algorithms. The rest of the paper is organized as follows: In Section 2, different search patterns and discussed and accuracy analysis of ME process is given whereas, Section 3 describes the proposed ME algorithm. Results are discussed in Section 4 and on the basis of results, conclusion is made in Section 5.

2. SEARCH PATTERN AND ME ACCURACY ANALYSIS
It is observed that the size and shape of any search pattern can increase or decrease the search speed and the quality of a video considerably [2-7].

The size of the pattern can influence the possibility of best matching position, as well as the converging speed of search pattern. Usually, the small search patterns are stuck in the local minimas, particularly for the video sequences having high motion scenarios having high resolution video formats.

In opposition to that, the search pattern having large search format may put the initial search in the wrong direction which may result in the greater time requirement for search procedure or may overlook the best matching position for the videos having low motion scenarios with low video resolutions. Number of search points in search patterns of initial and refinement search have the same impact on the optimal point search speed and quality as size and shape of search pattern. Reducing the number of search points in each search steps also decreases the quality, but increases the search speed. Similarly, by increasing the number of search points will yield to a low distortion measure, but the time required for estimating the motion will be increased.

Fig. 1 shows the different shaped search patterns like diamond, square, hexagon etc. which are used in fast Block Matching ME Algorithms (BMAs).

![Fig. 1. Search patterns: (a) Unity Cross, (b) Flattened Hexagon, (c) Hexagon, (d) Diamond Search](image)

Among fast BMAs, the DS algorithm [5] has the better performance as compared to the search patterns with square shapes [2], due to the fact that it has nine search points and has shape of medium size. Therefore, it has high search speed and low distortion measure for various video resolutions. Furthermore, the continuing search procedure keeps finding the best matching point, hence reduces the distortion measure by avoiding the search trap in local minimas. The hexagon search pattern elaborated in [6] has a circular type shape which covers the large area as compared to DS but has to total seven, reduced number of search points. Due to the reduced number search points, HexBS ME algorithm has greater search speed than the DS algorithm. It has high distortion measure as compared to DS algorithm, because, the search pattern may overlook the best matching point due to its large pattern size. Therefore, the number of search points also affect the speed and accuracy of ME process.

Motion vector distribution analysis is made in the [17] to analyze the distribution of horizontal and vertical probabilities in some most famous test video sequences. From Table 1, it can be observed that the probability of horizontal based motion vectors are quite higher than the probability of vertical based motion vectors.

<table>
<thead>
<tr>
<th>Video Sequences</th>
<th>Vertical Motion Probabilities (%):</th>
<th>Horizontal Motion Probabilities (%):</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forman</td>
<td>22.61</td>
<td>30.17</td>
</tr>
<tr>
<td>Tennis</td>
<td>20.50</td>
<td>26.07</td>
</tr>
<tr>
<td>Football</td>
<td>17.05</td>
<td>27.14</td>
</tr>
<tr>
<td>Garden</td>
<td>12.12</td>
<td>92.78</td>
</tr>
</tbody>
</table>

The probability distribution is computed by calculating MV angles. If the MV angle lies between 330° to 30° and 150° to 210° then the MV is considered as horizontally biased motion, whereas, if the MV angle lies between 60° to 120° and 240° to 300°, then the MV considered as vertically biased. We can summarize that maximum motion along x and y-axis by analysing the different video sequences. Therefore, it can be concluded that the shape and size of search pattern should be selected for putting the search quickly into the corresponding direction of horizontal or vertical axis to find the best matching point rapidly. The analysis made above strengthens the argument that, if the Plus Shaped Adaptive Predictor (PSAP) is used before implication of DS pattern, the performance and speed may be improved significantly, because the PSAP will initialize the search process on the horizontal or vertical axis and DS ME algorithm has better speed and low distortion measure as compared to the other algorithms, when used as refinement search. Moreover, the adaptive prediction may also be helpful to speed-up the ME process.

3. PROPOSED ME ALGORITHM
DS is very famous algorithm and accepted in the MPEG-4 Video Verification Model version 14.0 [18]. The original DS algorithm uses two different types of fixed diamond shaped search patterns, one is Large Diamond Search Pattern (LDSP) and the other is Small Diamond Search Pattern (SDSP). We have incorporated the adaptivity in the existing DS algorithm and proposed a new fast adaptive prediction based Diamond Search algorithm for Sun Falcon 2, solar powered UAV’s video encoder. The incorporation of adaptivity is based on the assumption that most commonly, camera is installed underneath the UAV and motion in the videos is along y-axis and x-axis during UAV flight. The speed and accuracy is based on shape and size of ME search pattern, therefore a PSAP is best for motion prediction that directly puts the
search into the desired area where there are high chances of best matching block. We also used previous MV to predict the MV of next neighbored MB. For example, if the MV of the previous MB is located at (-4, 2), then the same location of reference MB is also searched as a predictor which is the sixth location of PSAP. The size of PSAP is adjusted according to previous MV. Therefore it adapts the search according to previously calculated MV. Since PSAP quickly puts the initial search in the region having the highest probability of best matching position, which speed-ups the ME process. After prediction search, normal DS algorithm is applied to refine the search procedure.

The APDS starts the search by initializing search at \(C_x, C_y = (0,0)\) and calculates SAD(0,0) for left most MB of each video frame. Where \(C_x, C_y\) are search center position. Then SAD \((C_x + 1, C_y), SAD (C_x - 1, C_y), SAD (C_x, C_y + 1)\) and SAD \((C_x, C_y - 1)\), are calculated. SAD of all five search points are compared and LDSP is applied by keeping the minimum SAD point as the new origin. Then normal DS algorithm is applied by computing SAD of eight search points of LDSP which are: SAD \((C_x, C_y - 2), SAD (C_x - 1, C_y - 1), SAD (C_x + 1, C_y - 1), SAD (C_x - 2, C_y), SAD (C_x + 2, C_y), SAD (C_x - 1, C_y + 1), SAD (C_x + 1, C_y + 1)\) and SAD\((C_x, C_y + 2)\). The LDSP is applied until the minimum SAD is found at the center of LDSP. When minimum SAD found at center of LDSP, then SDSP is applied for, \(SAD (C_x, C_y - 1)\), \(SAD (C_x - 1, C_y), SAD (C_x + 1, C_y)\) and \(SAD (C_x, C_y + 1)\) search points. The minimum SAD point is considered as best MV for first MB. To compute the MV for the next MB, PSAP is applied and the size of PSAP is determined from the MV of the previous MB by using following relation:

\[
\text{SizeOfPSAP} = \text{Max} \left( \text{MV} \left( |X|, |Y| \right) \right)
\]

where \(X\) and \(Y\) are the coordinates of MV of previous MB.

Moreover, the SAD \((X, Y)\) is also computed as sixth location of PSAP for this MB. After this step, the original DS is applied for refinement search, as explained above. Furthermore, the early termination scheme is incorporated to increase the speed of ME process. In the early termination scheme, if the best matching position is at the center of unity sized PSAP, algorithm is terminated and the MB is declared as static.

The search procedure of APDS is illustrated in Fig 2, which shows that initially PSAP is applied with six search points and then two iterations of LDSP are applied and at second iteration of LDSP, the MDB is obtained at center of SDSP.

Then SDSP is applied and final minimum SAD point is declared as the best MV. So, the whole process takes sixteen search points within three iterations to compute the final MV.

4. RESULTS AND DISCUSSION

The compilation and simulation is done by using MATLAB 2014a software tool. The specifications of the computer system which is used for the compilation and simulation are: Intel® Core™ i3 Dual Processor 3050 @ 2.20G processor, 2 MB L2 Cache, 800 MT/s FSB, 4 GB DDR3 RAM, video card of 512 MB memory and 32 bit Windows® 7 operating system.

Seven different video sequences [19], each consisting of 100 video frames were used during the experiment for the comparison of APDS with other five ME algorithms. Three video sequences are captured from the different UAVs on different altitudes and variable UAV speed. Moreover, remaining video sequences are having different kind of motion scenarios like low motion with unidirectional and multidirectional motion. For example, InToTree and OldTownCross videos are having unidirectional and low motion contents. Similarly, the DucksTakeOff and Parkjoy having moderate unidirectional motion and high bidirectional motion respectively. Various parameters were used for simulation like video format is “yuv”, MB size is 16. Moreover four performance measuring parameters like MSE, PSNR, PSNR improvement, number of Search Points (SP), ME time, and Speed Improvement Rate (SIR) of proposed APDS ME algorithm with five famous fast BMME algorithms like DS, HexBS, FHS, FSS and MDS.

MSE tells about the accuracy of any ME algorithm. Lower MSE means the better accuracy of the algorithm. The average MSE of 100 frames from different test video sequences is calculated for each algorithm and illustrated in Fig. 3 which clearly shows that the proposed APDS calculates minimum MSE, therefore, it computes more accurate MV vectors when compared with all other algorithms considered.
PSNR depicts the quality measuring information and it is a function of MSE as given in equation “2" which shows that lower the MSE means greater the PSNR which leads to increase in accuracy of calculated MVs.

\[
PSNR = 10 \log_{10} \left( \frac{\text{peak to peak value of original data}}{\text{MSE}} \right)
\]  

(2)

The average PSNR taken by each algorithm to calculate the MVs of 100 videos frames of different test video sequences is given in Table 2.

<table>
<thead>
<tr>
<th>Video Sequence</th>
<th>DS</th>
<th>HexBS</th>
<th>FHS</th>
<th>FSS</th>
<th>MDS</th>
<th>APD S</th>
</tr>
</thead>
<tbody>
<tr>
<td>DucksTakeOff</td>
<td>30.95</td>
<td>30.90</td>
<td>30.95</td>
<td>30.92</td>
<td>30.63</td>
<td>30.94</td>
</tr>
<tr>
<td>InToTree</td>
<td>35.40</td>
<td>35.11</td>
<td>35.38</td>
<td>35.34</td>
<td>34.97</td>
<td>35.37</td>
</tr>
<tr>
<td>OldTownCross</td>
<td>34.75</td>
<td>34.49</td>
<td>34.75</td>
<td>34.70</td>
<td>34.57</td>
<td>34.73</td>
</tr>
<tr>
<td>ParkJoy</td>
<td>25.75</td>
<td>24.95</td>
<td>25.68</td>
<td>25.65</td>
<td>25.66</td>
<td>25.83</td>
</tr>
<tr>
<td>FlyingDown</td>
<td>36.95</td>
<td>36.57</td>
<td>36.95</td>
<td>36.81</td>
<td>36.33</td>
<td>36.86</td>
</tr>
<tr>
<td>HeageWMill</td>
<td>34.58</td>
<td>32.73</td>
<td>34.24</td>
<td>34.74</td>
<td>33.94</td>
<td>36.17</td>
</tr>
<tr>
<td>DntLook</td>
<td>37.98</td>
<td>36.68</td>
<td>37.94</td>
<td>37.70</td>
<td>35.99</td>
<td>38.01</td>
</tr>
</tbody>
</table>

The PSNR results in Table 2 shows that the proposed algorithm overcome all discussed algorithm and computes more accurate MVs but in some cases it has negligible PSNR degradation than DS and FHS which can be described in term of PSNR improvement rate. The PSNR improvement of proposed APDS is given in Table 3. It can be observed that PSNR improvement rate varies from –0.25 to 4.41 for DS and –0.24 to 5.35 for FHS but the average improvement rate is 0.62, 2.67, 0.82, 0.83 and 2.31 for DS, HexBS, FHS, FSS and MDS respectively.

<table>
<thead>
<tr>
<th>Video Sequence</th>
<th>DS</th>
<th>HexBS</th>
<th>FHS</th>
<th>FSS</th>
<th>MDS</th>
<th>APD S</th>
</tr>
</thead>
<tbody>
<tr>
<td>DucksTakeOff</td>
<td>-0.03</td>
<td>0.11</td>
<td>-0.03</td>
<td>0.05</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>InToTree</td>
<td>-0.09</td>
<td>0.72</td>
<td>-0.05</td>
<td>0.08</td>
<td>1.12</td>
<td></td>
</tr>
<tr>
<td>OldTownCross</td>
<td>-0.06</td>
<td>0.69</td>
<td>-0.05</td>
<td>0.10</td>
<td>0.46</td>
<td></td>
</tr>
<tr>
<td>ParkJoy</td>
<td>0.30</td>
<td>3.41</td>
<td>0.57</td>
<td>0.70</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>FlyingDown</td>
<td>-0.25</td>
<td>0.78</td>
<td>-0.24</td>
<td>0.13</td>
<td>1.42</td>
<td></td>
</tr>
<tr>
<td>HeageWMill</td>
<td>4.41</td>
<td>9.51</td>
<td>5.35</td>
<td>3.95</td>
<td>6.16</td>
<td></td>
</tr>
<tr>
<td>DntLook</td>
<td>0.07</td>
<td>3.49</td>
<td>0.18</td>
<td>0.82</td>
<td>5.32</td>
<td></td>
</tr>
<tr>
<td>Avg.</td>
<td>0.62</td>
<td>2.67</td>
<td>0.82</td>
<td>0.83</td>
<td>2.31</td>
<td></td>
</tr>
</tbody>
</table>

ME time and computational complexity are directly affected by number of SP for computing MV by any algorithm. Greater number of SPs increase the computations of ME algorithm, which causes to increase ME time for MV calculation. Therefore, number SP is an important parameter to evaluate the performance of proposed ME algorithm. The average number of SP of one MB for 100 video frames from different test video sequences to calculate the MVs by each algorithm is shown in Fig. 4. It can be observed that the proposed algorithm takes minimum number of search points for all sequences except three sequences like ParkJoy, FlyingDown and DntLook where HexBS perform well but APDS has good PSNR for these video sequences than the all other algorithms.

The literature review shows that the parameter SIR purely depends on number of SPs [20]. Table 4 gives the speed improvement rate of proposed algorithm from other algorithm.

<table>
<thead>
<tr>
<th>Video Sequence</th>
<th>SIR</th>
<th>DS</th>
<th>HexBS</th>
<th>FHS</th>
<th>FSS</th>
<th>MDS</th>
<th>APD S</th>
</tr>
</thead>
<tbody>
<tr>
<td>DucksTakeOff</td>
<td>32.7312</td>
<td>10.8911</td>
<td>11.537</td>
<td>66.058</td>
<td>278.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>InToTree</td>
<td>25.4434</td>
<td>-1.0751</td>
<td>6.2574</td>
<td>54.809</td>
<td>232.41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OldTownCross</td>
<td>29.6422</td>
<td>1.99428</td>
<td>10.048</td>
<td>61.917</td>
<td>244.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ParkJoy</td>
<td>9.37884</td>
<td>-25.441</td>
<td>-10.51</td>
<td>19.41</td>
<td>108.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FlyingDown</td>
<td>42.4804</td>
<td>9.55703</td>
<td>20.972</td>
<td>71.52</td>
<td>276.41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HeageWMill</td>
<td>22.25809</td>
<td>-26.98</td>
<td>8.1249</td>
<td>33.998</td>
<td>141.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DntLook</td>
<td>16.1387</td>
<td>-21.942</td>
<td>-2.071</td>
<td>30.208</td>
<td>156.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg.</td>
<td>25.4404</td>
<td>-7.6983</td>
<td>6.3374</td>
<td>48.274</td>
<td>205.55</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The average SIR of proposed APDS is 25.4404, -7.6983, 6.3374, 48.274 and 205.55 for DS, HexBS, FHS, FSS and MDS respectively. The above results show that there is a tradeoff between SP and PSNR, we can observe HexBS has negative SIR but it has low PSNR than our proposed APDS.

Fig. 5 shows the result for average ME time of 100 frames of different test video sequences taken by each algorithm to calculate the MVs. It can be observed that the proposed APDS algorithm takes minimum ME time as compared to the other algorithms.

![Fig. 5. Average ME time of 100 HD and CIF video frames](image-url)
5. CONCLUSION
This paper proposes an adaptive prediction based Diamond Search ME algorithm specially designed for Sun Falcon 2, solar powered UAV. APDS ME algorithm is proposed by considering the low power requirements and increased video quality of the Sun Falcon 2 UAV. The results show that the proposed APDS has better PSNR and less number of search points for all video sequences when compared with the other ME algorithms. Moreover, the performance of APDS is lower than HexBS algorithm in term of number of SPs for few videos sequences but PSNR is quite high for the same videos sequences. In addition with, results show that the proposed APDS algorithm has positive average SIR and PSNR improvement rate over all other algorithms which indicates its fitness over other algorithms. Therefore, higher PSNR and less number of search points makes the proposed APDS algorithm the best choice for incorporation into the video encoder of Sun Falcon 2, solar powered UAV. APDS ME algorithm can also be used for the video sequences having the maximum motion contents along the vertical axis.

6. FUTURE WORK
After going through the thorough implementation and testing, some future work is proposed. Early search process termination scheme can be incorporated in the APDS algorithm to make the search fast. Moreover, to achieve the parallelism available in hardware, hardware design of the same algorithm can also be proposed.

7. ACKNOWLEDGEMENT
This project was funded by Deanship of Scientific Research (DSR) King Abdulaziz University, Jeddah, under the grant no. (431/009), the authors therefore, acknowledge with thanks DSR technical and financial support, furthermore, the authors would like to express their gratitude and appreciation to Tokai University for their technical help and support.

8. ACKNOWLEDGEMENT
This project was funded by Deanship of Scientific Research (DSR) King Abdulaziz University, Jeddah, under the grant no. (431/009), the authors therefore, acknowledge with thanks DSR technical and financial support, furthermore, the authors would like to express their gratitude and appreciation to Tokai University for their technical help and support.

9. REFERENCES
[19] Available online at: https://media.xiph.org/video/derf/

10. AUTHOR PROFILE

Nasir Mahmud Khokhar is a Ph.D. student in Electrical Engineering Department at University of Engineering and Technology Taxila, Pakistan. He received his Master degree in Computer Engineering from University of Engineering and Technology Taxila, Pakistan in 2010. Currently he is working as a Research Associate at King Abdulaziz University Jeddah, Saudi Arabia. His research focuses on Formal Modelling, Security, Intrusion Detection and Prevention in Wireless Sensor Networks and Motion Estimation (ME) Algorithms for increased quality of ME process requirement.

Wail I. Harasani is an Associate Professor at King Abdulaziz University KAU, Saudi Arabia, MS in Aerospace Engineering from Kansas University (U.S.A), and Ph.D in Aerospace Engineering from Cranfield University (U.K) (2005). Specialized in aircraft design, unmanned air-vehicles design and airline fleet planning.

IJCA™: www.ijcaonline.org