Hardware-Friendly Motion Estimation Algorithms and its Architectures for High Definition Videos

Mário Saldanha, Gabriel Balota, Gustavo Sanchez, Luciano Agostini, Marcelo Porto
{mrdfsaldanha,gmbalota,gfsanchez,agostini,porto}@inf.ufpel.edu.br
Motion Estimation

- Uses a similarity criterion
- SAD: Sum of Absolute Differences

![Diagram of motion estimation process]

- Motion Vector
- Reference Frame
- Search Area
- Co-located block
- Best Match
- Current Frame
- Current Block
Motion Estimation

- Represents 80% of the encoder complexity
- Responsible for achieving high compression ratios
- Encoders do not restrict how the ME is done
- Explore new solutions:
  - High quality video
  - High compression rates
  - Low computational complexity
Diamond Search Algorithm

- Fast Algorithm
  - Reduce computational complexity compared to Full Search (FS) and maintains good quality results

- Uses two patterns
  - LDSP, to perform the search
  - SDSP, to perform the refinement
Hardware-Friendly ME Algorithms

- Objectives
  - High Quality when encoding HD videos
  - Hardware-Friendly Algorithm

- Two techniques are used
  - Multipoint
    - Dynamic Multi-Point Diamond Search (DMPDS)
    - Low Density and Iterative Search (LD&IS)
  - Random
    - Spread and Iterative Search (S&IS)
DMPDS

- The search area is divided into four sectors
  - In the center position and in every sector of the search area is assigned a DS
- Uses the parameter $d$, dynamically controlled
  - Small values for low motion videos
  - Higher values for high motion videos
- Self-adaptive according to the video characteristics
- **Low Density Multipoint Search**
  - Fixes 100 points (25 in each quadrant)
    - 5, 10, 20, 30 and 40 pixels away from the center to each side, in both x and y axis
    - Computes the SAD for each of these points
    - Compares the SADs and select the lowest SAD result

- **Central iterative evaluation**
  - Performs a DS in the central block of the search area
Random step
- The search area is divided into four sectors
  - Spread N candidate blocks (N/4 by sector)
  - Compute the SAD for each position
  - Compare the SADs and select the lowest

Central iterative search
- Performs a DS in the central block of the search area
## Software Results

<table>
<thead>
<tr>
<th>Video</th>
<th>DS PSNR (dB)</th>
<th>#ECB (x10^9)</th>
<th>FS PSNR (dB)</th>
<th>#ECB (x10^9)</th>
<th>DMPDS PSNR (dB)</th>
<th>#ECB (x10^9)</th>
<th>S&amp;IS PSNR (dB)</th>
<th>#ECB (x10^9)</th>
<th>LD&amp;IS PSNR (dB)</th>
<th>#ECB (x10^9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>blue_sky</td>
<td>30.01</td>
<td>0.04</td>
<td>34.43</td>
<td>14.66</td>
<td>33.73</td>
<td>0.24</td>
<td>31.12</td>
<td>0.20</td>
<td>31.74</td>
<td>0.20</td>
</tr>
<tr>
<td>man_in_car</td>
<td>37.80</td>
<td>0.03</td>
<td>39.99</td>
<td>14.66</td>
<td>39.60</td>
<td>0.24</td>
<td>39.31</td>
<td>0.18</td>
<td>39.26</td>
<td>0.19</td>
</tr>
<tr>
<td>pedestrian_area</td>
<td>32.22</td>
<td>0.05</td>
<td>35.97</td>
<td>14.66</td>
<td>35.25</td>
<td>0.34</td>
<td>34.83</td>
<td>0.19</td>
<td>34.79</td>
<td>0.20</td>
</tr>
<tr>
<td>Riverbed</td>
<td>24.42</td>
<td>0.06</td>
<td>27.72</td>
<td>14.66</td>
<td>26.86</td>
<td>0.36</td>
<td>26.47</td>
<td>0.21</td>
<td>26.50</td>
<td>0.21</td>
</tr>
<tr>
<td>rolling_tomatoes</td>
<td>37.38</td>
<td>0.03</td>
<td>38.18</td>
<td>14.66</td>
<td>38.32</td>
<td>0.28</td>
<td>37.87</td>
<td>0.18</td>
<td>37.87</td>
<td>0.19</td>
</tr>
<tr>
<td>rush_hour</td>
<td>36.48</td>
<td>0.03</td>
<td>37.40</td>
<td>14.66</td>
<td>37.28</td>
<td>0.36</td>
<td>36.99</td>
<td>0.18</td>
<td>36.95</td>
<td>0.19</td>
</tr>
<tr>
<td>station2</td>
<td>37.76</td>
<td>0.04</td>
<td>38.64</td>
<td>14.66</td>
<td>38.50</td>
<td>0.22</td>
<td>37.98</td>
<td>0.19</td>
<td>38.03</td>
<td>0.19</td>
</tr>
<tr>
<td>Sunflower</td>
<td>37.11</td>
<td>0.05</td>
<td>39.00</td>
<td>14.66</td>
<td>38.53</td>
<td>0.43</td>
<td>37.90</td>
<td>0.19</td>
<td>37.88</td>
<td>0.21</td>
</tr>
<tr>
<td>Traffic</td>
<td>24.90</td>
<td>0.07</td>
<td>32.45</td>
<td>14.66</td>
<td>28.81</td>
<td>0.39</td>
<td>28.27</td>
<td>0.21</td>
<td>28.03</td>
<td>0.22</td>
</tr>
<tr>
<td>Tractor</td>
<td>29.26</td>
<td>0.06</td>
<td>32.25</td>
<td>14.66</td>
<td>31.85</td>
<td>0.33</td>
<td>30.71</td>
<td>0.21</td>
<td>30.32</td>
<td>0.22</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>32.74</strong></td>
<td><strong>0.05</strong></td>
<td><strong>35.71</strong></td>
<td><strong>14.66</strong></td>
<td><strong>34.87</strong></td>
<td><strong>0.32</strong></td>
<td><strong>34.15</strong></td>
<td><strong>0.20</strong></td>
<td><strong>34.14</strong></td>
<td><strong>0.20</strong></td>
</tr>
</tbody>
</table>
Software Results

- Best Quality Results: DMPDS
  - Best trade-off between quality and computational complexity

- Compared with Full Search (FS)
  - Lost only 0.84 dB of PSNR
  - The computational complexity was 45 times lower

- Reason
  - As mentioned above has a parameter dynamically controlled which adapts according to video characteristics
ME Architecture

- Developed in VHDL
- Synthesized to an FPGA Stratix Altera 4
- Block size: 16x16
- Sub-sampling ratio 4:1
- Limit of iterations of the Diamond Search: 5
## Synthesis Results

<table>
<thead>
<tr>
<th>Architecture</th>
<th>Technology</th>
<th>Frequency (MHz)</th>
<th>Area</th>
<th>Memory (Kbits)</th>
<th>Cycles per Block</th>
<th>HD fps</th>
<th>QFHD fps</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMPDS</td>
<td>Stratix 4</td>
<td>187.58</td>
<td>34.5 KALUTs</td>
<td>46.2</td>
<td>170</td>
<td>136</td>
<td>34</td>
</tr>
<tr>
<td>LD&amp;IS</td>
<td>Stratix 4</td>
<td>256.84</td>
<td>18.5 KALUTs</td>
<td>46</td>
<td>174</td>
<td>180</td>
<td>45</td>
</tr>
<tr>
<td>S&amp;IS</td>
<td>Stratix 4</td>
<td>210.5</td>
<td>18.5 KALUTs</td>
<td>37</td>
<td>174</td>
<td>149</td>
<td>37.3</td>
</tr>
<tr>
<td>S&amp;IS</td>
<td>90nm</td>
<td>169</td>
<td>84.32 KGates</td>
<td>55.9</td>
<td>174</td>
<td>119.9</td>
<td>30</td>
</tr>
</tbody>
</table>
Best match: LD&IS

- The architecture achieved a better processing rate
- Using low resources available in hardware
- Processing up to 45 QFHD frames per second
Conclusion

- **Best Algorithm: DMPDS**
  - Higher trade-off between video quality and computational complexity
  - Compared to FS:
    - 0.84 dB less PSNR and 45 times less computational complexity

- **Best Architecture: LD&IS**
  - Best processing rate
  - Lower utilization of hardware resources
  - Able to process up to 45 QFHD frames per second
Thanks!

Mário Saldanha, Gabriel Balota, Gustavo Sanchez, Luciano Agostini, Marcelo Porto
{mrdfsaldanha,gmbalota,gfsanchez,agostini,porto}@inf.ufpel.edu.br