Content Classification and Non-Square Partition for Fast Inter Prediction of Screen Content Coding

Background

Screen Content Coding
Encode videos generated by computer

Problem

Inter prediction: takes 78% complexity

1. 4 Reference frames (RF)
2. 4 Coding Unit (CU) Depth
3. 8 Prediction Unit (PU)
4. Motion Estimation (ME)

Solution

Reduce complexity by reducing RF and skipping CU & PU

Proposal 1: RF reduction
Proposal 2: CU & PU skipping

Conventional (ME simplification)

Proposals

1. Content classification based PU level RF reduction

   - Different motion leads to different RF selection
     Need to classify contents

   1. Natural contents
   2. Screen contents

   - Different block selects different RF
     Need different RF reduction process

   S1: Content classification
   By number of color

   Natural bck 1 ... Natural bck n
   Screen bck 1 ... Screen bck n

   RF reduction for Natural
   RF reduction for Screen

2. Machine learning based non-square PU skipping

   - Use SVM to classify and find simple region with small motion
   - Skip non-square PUs in CU depth 1, 2, 3

   Separate contents by machine learning

S2: Block level process

   - Use SVM to classify and find simple region with small motion
   - Skip non-square PUs in CU depth 1, 2, 3

   Merit: few miss RF & PU skipping, reasonable complexity reduction with not obvious efficiency loss

Experiment result

<table>
<thead>
<tr>
<th>Categories</th>
<th>Sequences</th>
<th>Xiao’s BD-rate [%]</th>
<th>Xiao’s TS [%]</th>
<th>Proposal 1 &amp; 2 BD-rate [%]</th>
<th>Proposal 1 &amp; 2 TS [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text &amp; Graphics</td>
<td>Flying Gra.</td>
<td>-0.03</td>
<td>4.40</td>
<td>0.11</td>
<td>36.80</td>
</tr>
<tr>
<td></td>
<td>Desktop</td>
<td>0.00</td>
<td>25.40</td>
<td>0.64</td>
<td>41.44</td>
</tr>
<tr>
<td></td>
<td>Console</td>
<td>0.09</td>
<td>10.03</td>
<td>0.79</td>
<td>43.62</td>
</tr>
<tr>
<td></td>
<td>Web Brow.</td>
<td>0.00</td>
<td>34.83</td>
<td>0.03</td>
<td>53.21</td>
</tr>
<tr>
<td></td>
<td>Word Edi.</td>
<td>-</td>
<td></td>
<td>0.87</td>
<td>32.72</td>
</tr>
<tr>
<td></td>
<td>Video Conf.</td>
<td>-</td>
<td>-</td>
<td>-0.43</td>
<td>32.81</td>
</tr>
<tr>
<td></td>
<td>Map</td>
<td>0.00</td>
<td>18.86</td>
<td>0.56</td>
<td>44.62</td>
</tr>
<tr>
<td></td>
<td>Programming</td>
<td>-0.06</td>
<td>8.39</td>
<td>0.89</td>
<td>37.06</td>
</tr>
<tr>
<td></td>
<td>Slide Show</td>
<td>0.00</td>
<td>12.35</td>
<td>1.43</td>
<td>42.51</td>
</tr>
<tr>
<td>Mixed content</td>
<td>Mission Con.</td>
<td>0.00</td>
<td>18.53</td>
<td>0.62</td>
<td>37.01</td>
</tr>
<tr>
<td>Animation</td>
<td>Robot</td>
<td>0.01</td>
<td>4.69</td>
<td>2.27</td>
<td>47.31</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>0.00</td>
<td>15.27</td>
<td>0.71</td>
<td>40.83</td>
</tr>
</tbody>
</table>

Original SCC Proposal
Subjective result of worst case
Observe information loss by enlarging

Conclusion

Proposed methods can achieve 29.41% more time saving compared with conventional work with 0.71% BD-rate degradation