An Efficient Scheme for Motion Estimation Using Multireference Frames in H.264/AVC

Speaker: Wang Xiao
Content

- Motion Estimation
- Proposed Efficient MRME Algorithm for H.264
- Flow path
- Visual results

Paper Study

Experimental results
Introduction of Motion Estimation

Motion estimation is defined as searching the best motion vector, which is displacement of the coordinate of the best similar block previous frame for the block in current frame.
In H.264, as shown in figure, more than one prior coded frame can be used as reference for motion vector estimation in H.264. The coding efficiency can be improved by using multiple reference frames instead of single frame.

![Diagram of Motion Vector Estimation in H.264](image)

**Fig. 1.** Motion vector estimation using multireference frames in H.264.
Motion Estimation in H.264

The motion vector for the block is estimated as follows:

\[
\begin{align*}
mv^*_{f(t) \rightarrow f(t-k)} &= \arg\min_{mv_{f(t) \rightarrow f(t-k)} \in S_{f(t-k)}} \left\{ \sum_{x=0}^{A-1} \sum_{y=0}^{B-1} |c(x, y) - r_{mv_{f(t) \rightarrow f(t-k)}}(x, y)| \right. \\
& \quad + \lambda_{motion} \times R_{motion}(pmv, mv_{f(t) \rightarrow f(t-k)}) \right\}
\end{align*}
\]
Proposed Efficient MRME Algorithm for H.264

The proposed motion estimation process consists of three phases:

1. Making a vector map between two consecutive frames

2. Composing a temporary motion vector with element Vectors which are in the vector map

3. The temporary predictive motion vector is refined over a narrow search range
Example of the process to make a motion vector map
B. An efficient Motion Estimation

Fig. 3. Proposed process to estimate using motion vector map.
C. Computational Complexity of the Proposed MRME Scheme and conventional MRME

Proposed MRME Scheme:

\[ Q \times M \times (2W_{(t-1)} + 1) \times (2H_{(t-1)} + 1) + Q \times M \times \sum_{k=2}^{N} (2W_{(t-k)}^{\text{new}} + 1) \times (2H_{(t-k)}^{\text{new}} + 1) \]

Conventional Scheme:

\[ Q \times M \times \sum_{k=1}^{N} (2W_{(t-k)} + 1) \times (2H_{(t-k)} + 1) \]

Q is the number of total modes, M is the number of macroblocks in the current frame.

Since \((2W_{(t-k)}^{\text{new}} + 1)\) and \((2H_{(t-k)}^{\text{new}} + 1)\) are much smaller than \((2W_{(t-k)} + 1)\) and \((2H_{(t-k)} + 1)\), respectively, and the procedure is repeated \(Q \times M \times (N-1)\) (N is the number of the total frame) times, the proposed scheme provides a significant gain in the overall complexity.
Experimental results

The particular process is showed as follows:

N=The number of reference frames.
A current frame is f(t).

\( MVMAP_{f(t-k+1)\rightarrow f(t-k)} \), \( k = 2,3,\cdots,N \) are the motion vector map.

Step 0: The mode A × B is 4 × 4, N=4.
Step 1: Estimation of \( mv^*_{f(t)\rightarrow f(t-1)} \) for a frame \( f(t) \) by (1) with \( k=1 \).
Step 2: Make a \( MVMAP_{f(t)\rightarrow f(t-1)} \) from \( mv^*_{f(t)\rightarrow f(t-1)} \), and save the \( MVMAP_{f(t)\rightarrow f(t-1)} \) for a frame \( f(t+1) \) which will be a current frame in the next phase.
Step 3: k=2.
Step 4: \( PMV_{f(t)\rightarrow f(t-k)} = I_{f(t-k+1)\rightarrow f(t-k)} + mv^*_{f(t)\rightarrow f(t-k+1)} \) for all blocks in \( f(t) \).
Step 5: Estimate \( mv^*_{f(t)\rightarrow f(t-k)} \) by refinement of \( PMV_{f(t)\rightarrow f(t-k)} \).
Step 6: If \( k<N \), then k=k+1 and goto Step 4.
Step 7: Among the all \( mv^*_{f(t)\rightarrow f(t-1)} \), \( k =1,2,\cdots,N \), \( MV_{A\times B} \) is selected by minimizing SAD, and stop.
## Experimental results

<table>
<thead>
<tr>
<th>Sequence</th>
<th>The PSNR of a sequence</th>
<th>Conventional MRME</th>
<th>Proposed MRME</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carphone_cif.yuv</td>
<td></td>
<td>35.826</td>
<td>37.339</td>
</tr>
<tr>
<td>Foreman_cif.yuv</td>
<td></td>
<td>37.339</td>
<td>38.588</td>
</tr>
<tr>
<td>Football_cif.yuv</td>
<td></td>
<td>31.318</td>
<td>32.110</td>
</tr>
<tr>
<td>Paris_cif.yuv</td>
<td></td>
<td>34.514</td>
<td>36.090</td>
</tr>
</tbody>
</table>
### Experimental results

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Total MVE time(ms) for a sequence</th>
<th>Reduction ratio (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Conventional MRME</td>
<td>Proposed MRME</td>
</tr>
<tr>
<td>Carphone_cif.yuv</td>
<td>5234</td>
<td>1359</td>
</tr>
<tr>
<td>Foreman_cif.yuv</td>
<td>5125</td>
<td>1344</td>
</tr>
<tr>
<td>Football_cif.yuv</td>
<td>5141</td>
<td>1343</td>
</tr>
<tr>
<td>Paris_cif.yuv</td>
<td>5156</td>
<td>1344</td>
</tr>
</tbody>
</table>
Conclusion

The proposed scheme outperforms the conventional methods. The proposed algorithm enables a faster encoding without any loss of image quality.
Thank you!