Downscaling Decoder
A Low Complexity Approach to Decoding MPEG Video

SHI Shu
shishu@mprc.pku.edu.cn

Multimedia Group
Microprocessor Research & Development Center
Peking University
Oct. 26, 2006
What’s all about?

- New techniques to implement arbitrary downscaling in decoding MPEG video with low complexity.
  - DCT block recomposition and sub-band approximation in frequency domain
  - New motion compensation based on the block group motion evaluation
Where does it come out?

- How to decode the MPEG-4 video at real-time in an 100MHz RISC processor without any multimedia acceleration from ASIC or multimedia extension instructions?
  - Software optimization
  - Drop frame
  - Reduce spatial resolution
What can it do?

- Set Box (since 1993)
  SDTV decoder with HDTV capability
- Mobile Device
  Display large resolution video in a small device.
- ......
How to do it?

- Previous Approach
  - Early HD to SD decoder focused on memory rather than complexity reduction
  - Downscaling factor restricted to 2 or $2^n$

- Related Research
  - Scaling Static Image
  - DCT Block Recomposition
  - Sub-band Approximation
  - Transcoding
  - AMVR -- Adaptive Motion Vector Resampling
  - Motion Alignment
How to do it?

Focus on the most computation step.

- VLD
- IQ
- IDCT
- MC
- Frame Buffer

Uncompressed Stream → VLD → IQ → IDCT → MC → + → Reconstructed Frame
How it performs?

- **IDCT**

<table>
<thead>
<tr>
<th></th>
<th>MATLAB IDCT</th>
<th>Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>L=3, M=3, N=8, recompose 9 8×8 blocks to 1 block</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orig (9 IDCT)</td>
<td>33930</td>
<td>--</td>
</tr>
<tr>
<td>Our (R, 1 IDCT)</td>
<td>7754</td>
<td>87%</td>
</tr>
<tr>
<td><strong>L=2, M=2, N=8, recompose 4 8×8 blocks to 1 block</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orig (4 IDCT)</td>
<td>10580</td>
<td>--</td>
</tr>
<tr>
<td>Our (R, 1 IDCT)</td>
<td>5474</td>
<td>48%</td>
</tr>
<tr>
<td><strong>L=3/2, M=3/2, N=8, recompose 9 8×8 blocks to 4 blocks</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orig (9 IDCT)</td>
<td>33930</td>
<td>--</td>
</tr>
<tr>
<td>Our (D, R, 4IDCT)</td>
<td>21168</td>
<td>38%</td>
</tr>
</tbody>
</table>
How it performs?

- MC
  - Well-aligned motion: reduce the computation complexity to approximate $\frac{1}{L*M}$
  - Not-well-aligned motion: more complicated. depends on the MC methods in downscaled reference frame.
How it performs?

- Quality

![PSNR Graph](image-url)
Thank You!