EE 5359
H.264 to VC 1 Transcoding

Vidhya Vijayakumar
Multimedia Processing Lab
MSEE, University of Texas @ Arlington
vidhya.vijayakumar@mavs.uta.edu

Guided by Dr.K.R. Rao
Goals

- The goal of this project is to develop a basic transcoder between H.264 to VC-1 and compare against an open loop transcoder.
- The profile selected for H.264 is baseline (I and P) and profile selected for VC-1 is simple profile (I and P).
What is…

- **H.264**
  - The new industry standard
  - High quality, minimal files
  - Scalable from 3G to HD and beyond

- **VC 1**
  - Informal name of the SMPTE 421M video codec
  - Standard initially developed by Microsoft – WMV 9
  - Supported standard for blu-ray discs and windows media video
What is…

- **Transcoding**
  - Converting a previously compressed video signal into another one with different format
    - Change in bit rate, frame rate, frame size, or even compression standard
  - 2 Ways
    - Decode fully and encode in target standard
    - Change the bit stream format from one standard to another without undergoing the complete decoding and encoding process.
  - Limitations
    - Compression artifacts are cumulative
Why transcode H.264 to VC-1?

- The high definition DVD format Blue ray has mandated MPEG-2, H.264 and VC-1 as the video compression format.
- As H.264 based and VC-1 based content and products become available, transcoding in both directions will be widely used.
- From an end user point of view, any VC-1 decoder can now become twice as powerful as it was earlier.
An Application Scenario

Why VC1?

- Requires less computational power and can be decoded at full 1080i/p resolution on today’s off-the-shelf PC
- Delivers HD content at bit rates as low as 6-8 Mbps
- Advanced profile delivers compression efficiencies far superior to MPEG-2
- Better visual quality* when compared with H.264 (before the fidelity range extensions) and MPEG-2 demonstrated in independent tests.
- Prevalent codec in Microsoft’s ASF files

* VC-1 codecs have performed well in independent subjective quality tests
More of VC1…

- DCT-based video codec design
- Coding tools for interlaced video sequences as well as progressive encoding
- 8-bit, 4:2:0 format
- Uses block based transform and motion compensation with quantization and entropy coding.
## VC-1 Profiles

<table>
<thead>
<tr>
<th>Feature</th>
<th>Simple</th>
<th>Main</th>
<th>Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline intra frame compression</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Variable-sized transform</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>16-bit transform</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Overlapped transform</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>4 motion vectors per macroblock</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>¼ pixel luminance motion compensation</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>¼ pixel chrominance motion compensation</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Start codes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Extended motion vectors</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Feature</td>
<td>Simple</td>
<td>Main</td>
<td>Advanced</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>--------</td>
<td>-------</td>
<td>----------</td>
</tr>
<tr>
<td>Loop filter</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Dynamic resolution change</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Adaptive macroblock quantisation</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>B frames</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Intensity compensation</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Range adjustment</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Field and frame coding modes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>GOP Layer</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Display metadata</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Comparison of H.264 and VC-1

<table>
<thead>
<tr>
<th>Feature</th>
<th>VC-1</th>
<th>H.264</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partition sizes</td>
<td>16x16 and 8x8</td>
<td>16x16, 16x8, 8x16, 8x8, 8x4, 4x8, and 4x4</td>
</tr>
<tr>
<td>Color space</td>
<td>4:2:0</td>
<td>4:2:0, 4:2:2, 4:4:4</td>
</tr>
<tr>
<td>Integer transform</td>
<td>8x8, 4x8, 8x4, and 4x4</td>
<td>4x4; 8x8 available in high profile only</td>
</tr>
<tr>
<td>CABAC</td>
<td>No</td>
<td>In main and high profiles</td>
</tr>
<tr>
<td>DBF</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Macroblock sizes</td>
<td>16x16 only</td>
<td></td>
</tr>
<tr>
<td>Motion vector</td>
<td>Two dimensional vector offset from current position to reference frame</td>
<td></td>
</tr>
<tr>
<td>Picture</td>
<td>A field or frame</td>
<td></td>
</tr>
<tr>
<td>Skipped MB</td>
<td>No data is encoded for macroblock</td>
<td></td>
</tr>
</tbody>
</table>
H.264 Vs VC-1

H.264 ME/MC sizes

VC-1 ME/MC sizes
**H.264 Vs VC-1**

<table>
<thead>
<tr>
<th>H.264 transform sizes</th>
<th>VC-1 transform sizes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>8x8</strong></td>
<td><strong>8x8</strong></td>
</tr>
<tr>
<td>0 1</td>
<td>0 1</td>
</tr>
<tr>
<td>2 3</td>
<td>2 3</td>
</tr>
<tr>
<td><strong>4x4</strong></td>
<td><strong>4x4</strong></td>
</tr>
<tr>
<td>0 1</td>
<td>0 1</td>
</tr>
<tr>
<td>2 3</td>
<td>2 3</td>
</tr>
</tbody>
</table>
Mapping the differences between H.264 and VC-1
Intra MB Mode Mapping

- VC-1 has no spatial intra prediction
- Always uses 8x8 transform block size

<table>
<thead>
<tr>
<th>H.264 Intra MB</th>
<th>VC-1 Intra MB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Intra 16x16 (any mode)</td>
<td>4 Intra MB 8x8</td>
</tr>
<tr>
<td>4 Intra 4x4 (any mode)</td>
<td>1 Intra MB 8x8</td>
</tr>
</tbody>
</table>
## Inter MB Mode Mapping

<table>
<thead>
<tr>
<th>H.264 Inter MB</th>
<th>VC-1 Inter MB</th>
<th>Transform size in VC-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inter 16x16</td>
<td>Inter 16x16</td>
<td>8x8</td>
</tr>
<tr>
<td>2 Inter 16x8</td>
<td>Inter 16x16</td>
<td>8x4</td>
</tr>
<tr>
<td>2 Inter 8x16</td>
<td>Inter 16x16</td>
<td>4x8</td>
</tr>
<tr>
<td>Inter 8x8</td>
<td>Inter 8x8</td>
<td>8x8</td>
</tr>
<tr>
<td>2 Inter 4x8</td>
<td>Inter 8x8</td>
<td>4x8</td>
</tr>
<tr>
<td>2 Inter 8x4</td>
<td>Inter 8x8</td>
<td>8x4</td>
</tr>
<tr>
<td>4 Inter 4x4</td>
<td>Inter 8x8</td>
<td>4x4</td>
</tr>
</tbody>
</table>
# Motion vector mapping

<table>
<thead>
<tr>
<th>H.264 Inter MB</th>
<th>VC-1 Inter MB</th>
<th>Motion Vector Re-use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inter 16x16</td>
<td>Inter 16x16</td>
<td>Same motion vectors</td>
</tr>
<tr>
<td>2 Inter 16x8</td>
<td>Inter 16x16</td>
<td>Average of motion vectors</td>
</tr>
<tr>
<td>2 Inter 8x16</td>
<td>Inter 16x16</td>
<td>Average of motion vectors</td>
</tr>
<tr>
<td>Inter 8x8</td>
<td>Inter 8x8</td>
<td>Same motion vectors</td>
</tr>
<tr>
<td>2 Inter 4x8</td>
<td>Inter 8x8</td>
<td>Average of motion vectors</td>
</tr>
<tr>
<td>2 Inter 8x4</td>
<td>Inter 8x8</td>
<td>Average of motion vectors</td>
</tr>
<tr>
<td>4 Inter 4x4</td>
<td>Inter 8x8</td>
<td>Average of motion vectors</td>
</tr>
</tbody>
</table>
Reference Pictures

H.264

\[ \downarrow \]

VC-1
Progress - Open loop transcoder
Progress - Open loop performance

MSE for Akiyo of size QCIF

Mean Square Error vs Frame number
Progress - Open loop performance
Progress - Open loop performance

SSIM for Akiyo of size QCIF

- SSIM

- Frame number

Values range from 0.952 to 0.964
Progress – Extract information from bitstream

- H.264 bitstream contains information about
  - Macroblock type
    - P16x16
    - P16x8
    - P8x16
    - P8x8
    - I4MB
    - I16MB
  - Macroblock sub block type
    - SMB8x8
    - SMB8x4
    - SMB4x8
    - SMB4x4
  - Reference picture index
  - Motion vector x, y
<table>
<thead>
<tr>
<th>mb_no 0</th>
<th>mb_no 1</th>
<th>mb_no 2</th>
<th>mb_no 3</th>
<th>mb_no 4</th>
<th>mb_no 5</th>
<th>mb_no 6</th>
<th>mb_no 7</th>
<th>mb_no 8</th>
<th>mb_no 9</th>
<th>mb_no 10</th>
<th>mb_no 11</th>
<th>mb_no 12</th>
<th>mb_no 13</th>
<th>mb_no 14</th>
<th>mb_no 15</th>
</tr>
</thead>
<tbody>
<tr>
<td>mb_type 9</td>
<td>mb_type 9</td>
<td>mb_type 9</td>
<td>mb_type 9</td>
<td>mb_type 9</td>
<td>mb_type 9</td>
<td>mb_type 9</td>
<td>mb_type 9</td>
<td>mb_type 9</td>
<td>mb_type 9</td>
<td>mb_type 9</td>
<td>mb_type 9</td>
<td>mb_type 9</td>
<td>mb_type 9</td>
<td>mb_type 9</td>
<td>mb_type 9</td>
</tr>
</tbody>
</table>

```plaintext
mb_no 0
mb_type 2
mb_ref_pic_idx 0
mb_motion_vector_x -1
mb_motion_vector_y -1
mb_ref_pic_idx 0
mb_motion_vector_x -3
mb_motion_vector_y -3
mb_no 1
mb_type 8
sub_mb_mode 4
mb_ref_pic_idx 0
mb_motion_vector_x 59
mb_motion_vector_y 59
sub_mb_mode 4
mb_ref_pic_idx 0
mb_motion_vector_x 91
mb_motion_vector_y 91
sub_mb_mode 6
mb_ref_pic_idx 0
mb_motion_vector_x 59
mb_motion_vector_y 59
mb_ref_pic_idx 0
mb_motion_vector_x 77
mb_motion_vector_y 77
sub_mb_mode 4
mb_ref_pic_idx 0
mb_motion_vector_x 93
mb_motion_vector_y 93
mb_no 2
mb_type 2
mb_ref_pic_idx 0
mb_motion_vector_x 123
```
Stepping forward...

- Use extracted information to encode in VC-1 using the mapping table (currently working on this)
- Compare results with the open loop transcoder
Future Work

Future work

- Extend the transcoder for main profile in H.264 which includes B pictures.
- Improve motion vector accuracy to improve the visual quality
Thank You!