ICME Grand Challenge on Light Field Image Compression

Irene Viola, Martin Rerabek, Tim Bruylants, Touradj Ebrahimi, Fernando Pereira, Peter Schelkens
ICME 2016 Grand Challenge

- Grand Challenge on Light Field image compression
  - Collect new compression solutions for LF images
  - Evaluate proposed compression schemes w.r.t. anchor

  • Objective and subjective evaluations
Light Field photography

- Captures plenoptic information
- Allows change of perspective and refocus in post-processing
Lenslet based LF

- Huge amount of information
- Need for coding techniques to compress the LF structure
Test bed set up

raw sensor data

demosaicing → devignetting → clipping to 8 bits → RGB to YCbCr → 4:4:4 to 4:2:0 → compression 4:2:0 8 bits → bitstream

bitstream

decompression 4:2:0 8 bits → 4:2:0 to 4:4:4 → YCbCr To RGB → lenslet image to LF data structure → rendering → output image (individual view/focus)
Reference and anchor processing chain

- Reference image generated following the same steps
- Anchor generated using legacy JPEG
Coding Conditions

- Twelve contents taken from LF dataset (EPFL)
- Four fixed compression ratios
  - 10:1, 20:1, 40:1, 100:1
  - From 1bpp to 0.1bpp
Light field content
## Coding Conditions

<table>
<thead>
<tr>
<th>Image ID</th>
<th>Image name</th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
<th>R4</th>
</tr>
</thead>
<tbody>
<tr>
<td>I01</td>
<td>Bikes</td>
<td>10:1</td>
<td>20:1</td>
<td>40:1</td>
<td>100:1</td>
</tr>
<tr>
<td>I02</td>
<td>Danger_de_Mort</td>
<td>10:1</td>
<td>20:1</td>
<td>40:1</td>
<td>100:1</td>
</tr>
<tr>
<td>I03</td>
<td>Flowers</td>
<td>10:1</td>
<td>20:1</td>
<td>40:1</td>
<td>100:1</td>
</tr>
<tr>
<td>I04</td>
<td>Stone_Pillars_Outside</td>
<td>10:1</td>
<td>20:1</td>
<td>40:1</td>
<td>100:1</td>
</tr>
<tr>
<td>I05</td>
<td>Vespa</td>
<td>10:1</td>
<td>20:1</td>
<td>40:1</td>
<td>100:1</td>
</tr>
<tr>
<td>I06</td>
<td>Ankylosaurus__Diplodocus_1</td>
<td>10:1</td>
<td>20:1</td>
<td>40:1</td>
<td>100:1</td>
</tr>
<tr>
<td>I07</td>
<td>Desktop</td>
<td>10:1</td>
<td>20:1</td>
<td>40:1</td>
<td>100:1</td>
</tr>
<tr>
<td>I08</td>
<td>Magnets_1</td>
<td>10:1</td>
<td>20:1</td>
<td>40:1</td>
<td>100:1</td>
</tr>
<tr>
<td>I09</td>
<td>Fountain__Vincent_2</td>
<td>10:1</td>
<td>20:1</td>
<td>40:1</td>
<td>100:1</td>
</tr>
<tr>
<td>I10</td>
<td>Friends_1</td>
<td>10:1</td>
<td>20:1</td>
<td>40:1</td>
<td>100:1</td>
</tr>
<tr>
<td>I11</td>
<td>Color_Chart_1</td>
<td>10:1</td>
<td>20:1</td>
<td>40:1</td>
<td>100:1</td>
</tr>
<tr>
<td>I12</td>
<td>ISO_Chart_12</td>
<td>10:1</td>
<td>20:1</td>
<td>40:1</td>
<td>100:1</td>
</tr>
</tbody>
</table>


Quality Assessment Methodology

• **Objective** performance evaluation
  – Objective metrics: PSNR and SSIM

• **Subjective** performance evaluation
  – DSCQS methodology
  – Modified QualityCrowd 2 framework
Objective performance evaluation

- PSNR for Y channel

\[
PSNR_Y(k, l) = 10 \log_{10} \frac{255^2}{MSE(k,l)}
\]

\[
MSE(k, l) = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} [I(i, j) - R(i, j)]^2
\]
Objective performance evaluation

- PSNR for YUV channel

\[
PSNR_{YUV}(k, l) = \frac{6PSNR_Y(k,l) + PSNR_U(k,l) + PSNR_V(k,l)}{8}
\]
Objective performance evaluation

- Mean PSNR

\[ PSNR_{Y,mean} = \frac{1}{(K-2)(L-2)} \sum_{k=2}^{K-1} \sum_{l=2}^{L-1} PSNR_Y (k, l) \]

\[ PSNR_{YUV,mean} = \frac{1}{(K-2)(L-2)} \sum_{k=2}^{K-1} \sum_{l=2}^{L-1} PSNR_{YUV} (k, l) \]
Objective performance evaluation

• SSIM for $Y$

$$SSIM_Y(k, l) = \frac{(2\mu_I\mu_R+c_1)(2\sigma_{IR}+c_2)}{(\mu_I^2 + \mu_R^2+c_1)(\sigma_I^2 + \sigma_R^2+c_2)}$$

24 October 2016
Subjective Performance Evaluation

- 3 rendered perspectives, 2 re-focused points
  - 5 test stimuli per content
- Only 6 contents out of 12 from the dataset
Light field content
Subjective test
Subjective performance evaluation

- DSCQS methodology
  - Double stimulus continuous quality scale
  - Rate quality level of reference and decoded image
  - Side by side configuration on a MacBook Pro Retina
  - 5-scale score from bad to excellent

- Modified QualityCrowd 2 framework
Subjective performance evaluation

Image quality

Please rate the "visual" image quality

- Bad
- Poor
- Fair
- Good
- Excellent
Results

• 7 algorithms were received
  – 5 were accepted

• Anonymized through random labels P1 to P5
  – Anchor P0 is legacy JPEG
Results

• Objective evaluation
Results

- Objective evaluation
Results

• Subjective evaluation
Results

- Subjective evaluation

MOS vs bitrate for content l03 (view 5)
Statistical Analysis

• Welch’s t-test of equal means

\[ H_0: MOS_{PA} = MOS_{PB} \]
\[ H_1: MOS_{PA} \neq MOS_{PB} \]
Statistical Analysis

• Examine if the null hypothesis is rejected or not
• For each bitrate, store for how many contents and views one codec performs better than the other
Statistical Analysis

Multiple comparison for compression ratio R1

Multiple comparison for compression ratio R2
### Multiple comparison for compression ratio R3

<table>
<thead>
<tr>
<th></th>
<th>P0</th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>P5</th>
</tr>
</thead>
<tbody>
<tr>
<td>P0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P1</td>
<td>29</td>
<td>10</td>
<td>3</td>
<td>12</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>P2</td>
<td>30</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>P3</td>
<td>30</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>P4</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>P5</td>
<td>30</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

### Multiple comparison for compression ratio R4

<table>
<thead>
<tr>
<th></th>
<th>P0</th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>P5</th>
</tr>
</thead>
<tbody>
<tr>
<td>P0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P1</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P2</td>
<td>30</td>
<td>20</td>
<td>17</td>
<td>27</td>
<td>19</td>
<td>0</td>
</tr>
<tr>
<td>P3</td>
<td>30</td>
<td>7</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P4</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P5</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Statistical Analysis

- Multi-way ANOVA to test interaction among groups
- Inter- and intra-variance among views
### Statistical analysis

<table>
<thead>
<tr>
<th></th>
<th>R1</th>
<th></th>
<th>R2</th>
<th></th>
<th>R3</th>
<th></th>
<th>R4</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>p-value</td>
<td>F</td>
<td>p-value</td>
<td>F</td>
<td>p-value</td>
<td>F</td>
<td>p-value</td>
</tr>
<tr>
<td>contents</td>
<td>86.284</td>
<td>0.000</td>
<td>125.982</td>
<td>0.000</td>
<td>80.109</td>
<td>0.000</td>
<td>122.346</td>
<td>0.000</td>
</tr>
<tr>
<td>proponents</td>
<td>6.715</td>
<td>0.000</td>
<td>184.978</td>
<td>0.000</td>
<td>630.502</td>
<td>0.000</td>
<td>693.525</td>
<td>0.000</td>
</tr>
<tr>
<td>contents*proponents</td>
<td>2.926</td>
<td>0.000</td>
<td>13.710</td>
<td>0.000</td>
<td>3.972</td>
<td>0.000</td>
<td>8.644</td>
<td>0.000</td>
</tr>
<tr>
<td>views</td>
<td>47.664</td>
<td>0.000</td>
<td>29.757</td>
<td>0.000</td>
<td>9.462</td>
<td>0.000</td>
<td>7.199</td>
<td>0.000</td>
</tr>
<tr>
<td>contents*views</td>
<td>2.388</td>
<td>0.002</td>
<td>1.471</td>
<td>0.109</td>
<td>1.684</td>
<td>0.048</td>
<td>1.153</td>
<td>0.311</td>
</tr>
<tr>
<td>proponents*views</td>
<td>1.974</td>
<td>0.015</td>
<td>1.256</td>
<td>0.227</td>
<td>2.385</td>
<td>0.002</td>
<td>1.909</td>
<td>0.019</td>
</tr>
<tr>
<td>focus</td>
<td>0.035</td>
<td>0.854</td>
<td>0.082</td>
<td>0.777</td>
<td>2.241</td>
<td>0.147</td>
<td>3.773</td>
<td>0.063</td>
</tr>
<tr>
<td>contents*focus</td>
<td>2.465</td>
<td>0.060</td>
<td>1.424</td>
<td>0.250</td>
<td>3.992</td>
<td>0.008</td>
<td>1.684</td>
<td>0.175</td>
</tr>
<tr>
<td>proponents*focus</td>
<td>2.697</td>
<td>0.044</td>
<td>0.491</td>
<td>0.780</td>
<td>0.643</td>
<td>0.669</td>
<td>0.927</td>
<td>0.480</td>
</tr>
<tr>
<td>perspective</td>
<td>3.287</td>
<td>0.046</td>
<td>1.113</td>
<td>0.337</td>
<td>0.100</td>
<td>0.905</td>
<td>2.668</td>
<td>0.079</td>
</tr>
<tr>
<td>contents*perspective</td>
<td>3.066</td>
<td>0.004</td>
<td>0.894</td>
<td>0.545</td>
<td>0.324</td>
<td>0.971</td>
<td>0.974</td>
<td>0.478</td>
</tr>
<tr>
<td>proponents*perspective</td>
<td>1.369</td>
<td>0.222</td>
<td>1.249</td>
<td>0.284</td>
<td>2.059</td>
<td>0.046</td>
<td>2.957</td>
<td>0.005</td>
</tr>
</tbody>
</table>
Expert viewers

- Higher confidence interval
- Results are consistent with naïve viewers
Conclusions

• At higher bitrates, proponents and anchor have similar subjective performance
• At lower bitrates, P1 can be identified as performing better for the rendering points that were subjectively assessed,
• However, overall P1 exhibited lower PSNR performance at higher bitrates
• Several lessons were learned and ideas to improve the challenge in the future have been identified
Thank you!