Robust 360° Video Streaming via Non-Linear Sampling

Mijanur Palash, Voicu Popescu, Amit Sheoran, Sonia Fahmy
Multiple sources of 360° videos

37 million 360° video headsets in US [1]

42.9 million users in US use virtual reality in a month [1]

360° videos require high bandwidth
360° videos require high bandwidth

- Users only watch a small portion of the video
  - User field of view (FoV)
State of the Art

- **Tiling**

- **Rate adaptation over time**
  - e.g., MPC: Yin et al, SIGCOMM 2015.
Observations

- Accurately predicting user FoV is not always possible
  - Errors in prediction can lead to
    - Missing pixels
    - Abrupt changes in quality
  - Re-fetching the video after correcting FoV prediction is difficult
  - Cellular networks can have high uplink/downlink latency \([1]\)

360° video streaming solutions must be robust to view prediction error

Observations

- Bandwidth fluctuation is common
  - Bandwidth disruption during handovers

- When to pre-fetch?
  - Fetching too early can lead to large view prediction errors
  - Fetching too late can lead to stalls

360° video streaming solutions must be robust to bandwidth fluctuation
Objectives

- Robustness to view prediction error
- Robustness to transient bandwidth fluctuation
- Decoding and rendering efficiency
  - Support thinner clients without GPUs
- Compatibility with current protocols (H.264 and DASH)
Compressed Rotated Equirectangular (CoRE) 360° Video Streaming
CoRE Encoding: Rotate to center predicted FoV
CoRE Encoding: Compress frame periphery

Predicted FoV

Rotate

Compress
CoRE Encoding: Compress frame periphery

- Input 360° frame
- Predicted FoV
- CoRE Encoder
- CoRE frame

PURDUE UNIVERSITY
CoRE Frame

Predicted FoV

Central region (uncompressed)

Peripheral region (compressed)
CoRE Encoding

Extension for robustness to bandwidth fluctuation
CoRE Encoding

Extension for robustness to bandwidth fluctuation
- 4s of additional data (1.3x)
- 6s of additional data (1.44x)

TCP RTO-like adaptive prefetching
CoRE Decoding

CoRE frame → CoRE Decoder → Output frame
Experimental Evaluation
How does CoRE compare to other methods?

- 4 bandwidth traces
- 6 videos
- 25 to 60 head movement traces per video

<table>
<thead>
<tr>
<th>Method</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>FoV only</td>
<td>Field of View only (90° x 48°)</td>
</tr>
<tr>
<td>FoV+ 1QL</td>
<td>FoV and padding (20%) of high quality</td>
</tr>
<tr>
<td>FoV+ 2QL</td>
<td>FoV high quality, with padding (20%) in lower quality</td>
</tr>
<tr>
<td>FoV 360</td>
<td>FoV (high quality) and all remaining tiles in lower quality</td>
</tr>
<tr>
<td>FoV+ 360</td>
<td>FoV+padding (high quality) and all remaining tiles in lower quality</td>
</tr>
<tr>
<td>CoRE</td>
<td>4s main part (high quality (90°x48°)) and 6s extension part</td>
</tr>
</tbody>
</table>
Data Transferred

CoRE transfers significantly less data

- Diving
  - FoV only
  - FoV + 1QL
  - FoV + 2QL
  - FoV + 360
  - CoRE

- Rhino
  - 30% increase
  - 52% increase

- Roller
  - 50% increase

Purdue University
Stalls

CoRE has significantly fewer stalls
Average Frame Rate

CoRE has significantly higher frame rate

<table>
<thead>
<tr>
<th>Frame Rate (fps)</th>
<th>Diving</th>
<th>Rhino</th>
<th>Roller</th>
</tr>
</thead>
<tbody>
<tr>
<td>FoV only</td>
<td></td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>FoV + 1QL</td>
<td>231%</td>
<td>87%</td>
<td></td>
</tr>
<tr>
<td>FoV + 2QL</td>
<td></td>
<td></td>
<td>243%</td>
</tr>
<tr>
<td>FoV + 360</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CoRE</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FoV only
FoV + 1QL
FoV + 2QL
FoV + 360
CoRE

22
User Study

- Compare CoRE and FoV+ 1QL tiling with 3 videos
Conclusions

- CoRE is a new approach for 360° video streaming
  - Robust to view prediction errors
  - Robust to bandwidth fluctuation

- CoRE has significantly lower resource requirements
  - Lower energy consumption

- User study shows that CoRE enhances user experience
More in the paper ...

- Results with more videos
- Results with more bandwidth traces
  - AT&T, Verizon, and T-Mobile bandwidth traces
- Additional evaluation metrics such as missing pixels
- Comparison of decoding overhead
  - Energy/Time comparison to tiling
- Impact of view misprediction
- Cost/benefit analysis
Thank you

Questions?