Considerations for Immersive Media Services

20171130
The Times, They Are A-Changin’

- Resolution
- Bandwidth
- Formats and Interchange
- The Future is Now
- Looking Forward
## Resolution: Display

<table>
<thead>
<tr>
<th>Typical</th>
<th>Traditional Display</th>
<th>Stereoscopic VR Eye Display</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good</td>
<td>480p “SD”</td>
<td>Effectively unusable</td>
</tr>
<tr>
<td>Better</td>
<td>720p “HD” (1280 x 720)</td>
<td>640x720 (Cardboard)</td>
</tr>
<tr>
<td>Best</td>
<td>1080p “Full HD” (1920 x 1080)</td>
<td>960x1080 (Oculus DK2, Pixel 1&amp;2)</td>
</tr>
<tr>
<td>Unobtainium</td>
<td>1440p “Quad HD” (2560 x 1440)</td>
<td>1280x1440 (GearVR; Pixel 1&amp;2 XL; WinMR) Rift/Vive is 1080x1200</td>
</tr>
<tr>
<td></td>
<td>“4K”</td>
<td>2Kx2K TBD</td>
</tr>
</tbody>
</table>
Resolution: TV and Mobile

<table>
<thead>
<tr>
<th>Year</th>
<th>% of all TVs that are HD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>0%</td>
</tr>
<tr>
<td>2007</td>
<td>11%</td>
</tr>
<tr>
<td>2014</td>
<td>59%</td>
</tr>
<tr>
<td>2017</td>
<td>79%</td>
</tr>
</tbody>
</table>

79% of all TV sets in US households are HDTVs – an increase from 59% in 2014, and 11% in 2007 (Source: Leichtman Research)

Unity Mobile Hardware Stats: Display Resolution

<table>
<thead>
<tr>
<th>Resolution</th>
<th>2015-03</th>
<th>2015-06</th>
<th>2015-09</th>
<th>2015-12</th>
<th>2016-03</th>
<th>2016-06</th>
<th>2016-09</th>
<th>2016-12</th>
</tr>
</thead>
<tbody>
<tr>
<td>2048x1536</td>
<td>2.3%</td>
<td>2.8%</td>
<td>3.0%</td>
<td>3.5%</td>
<td>3.3%</td>
<td>2.6%</td>
<td>2.4%</td>
<td>2.1%</td>
</tr>
<tr>
<td>1920x1080</td>
<td>9.9%</td>
<td>10.3%</td>
<td>12.0%</td>
<td>13.1%</td>
<td>14.8%</td>
<td>17.8%</td>
<td>19.4%</td>
<td>20.0%</td>
</tr>
<tr>
<td>1334x750</td>
<td>1.1%</td>
<td>1.6%</td>
<td>2.5%</td>
<td>2.7%</td>
<td>2.9%</td>
<td>2.4%</td>
<td>2.2%</td>
<td>1.6%</td>
</tr>
<tr>
<td>1280x800</td>
<td>6.2%</td>
<td>5.5%</td>
<td>5.0%</td>
<td>4.5%</td>
<td>4.0%</td>
<td>3.5%</td>
<td>3.9%</td>
<td>4.4%</td>
</tr>
<tr>
<td>1280x720</td>
<td>13.8%</td>
<td>14.9%</td>
<td>15.2%</td>
<td>16.5%</td>
<td>17.4%</td>
<td>20.5%</td>
<td>22.8%</td>
<td>26.0%</td>
</tr>
<tr>
<td>1136x640</td>
<td>10.4%</td>
<td>10.0%</td>
<td>9.9%</td>
<td>8.8%</td>
<td>8.5%</td>
<td>7.3%</td>
<td>5.3%</td>
<td>3.5%</td>
</tr>
<tr>
<td>1024x768</td>
<td>3.2%</td>
<td>3.4%</td>
<td>3.2%</td>
<td>2.9%</td>
<td>2.7%</td>
<td>2.0%</td>
<td>1.9%</td>
<td>1.7%</td>
</tr>
<tr>
<td>1024x600</td>
<td>6.3%</td>
<td>5.8%</td>
<td>6.0%</td>
<td>6.2%</td>
<td>5.7%</td>
<td>5.0%</td>
<td>5.5%</td>
<td>6.2%</td>
</tr>
<tr>
<td>960x540</td>
<td>5.9%</td>
<td>5.8%</td>
<td>5.9%</td>
<td>6.2%</td>
<td>6.7%</td>
<td>7.0%</td>
<td>7.4%</td>
<td>7.7%</td>
</tr>
<tr>
<td>854x480</td>
<td>5.5%</td>
<td>6.3%</td>
<td>6.4%</td>
<td>7.5%</td>
<td>7.4%</td>
<td>7.6%</td>
<td>7.8%</td>
<td>8.2%</td>
</tr>
<tr>
<td>800x480</td>
<td>13.9%</td>
<td>12.7%</td>
<td>11.5%</td>
<td>10.7%</td>
<td>10.1%</td>
<td>9.2%</td>
<td>9.2%</td>
<td>8.1%</td>
</tr>
<tr>
<td>Others</td>
<td>21.4%</td>
<td>21.0%</td>
<td>19.4%</td>
<td>17.3%</td>
<td>16.5%</td>
<td>15.2%</td>
<td>12.2%</td>
<td>10.4%</td>
</tr>
<tr>
<td>2560x1440</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.8%</td>
</tr>
</tbody>
</table>

Smartphones evolve faster than TVs.
Resolution: Mobile and VR Options

(Source: DeviceAtlas)

Most used smartphone screen resolutions 2016 part 1

Most used smartphone screen resolutions 2016 part 2

Germany

<table>
<thead>
<tr>
<th>Resolution</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>640x1136</td>
<td>14.8</td>
</tr>
<tr>
<td>720x1280</td>
<td>15.4</td>
</tr>
<tr>
<td>750x1334</td>
<td>17</td>
</tr>
<tr>
<td>1080x1920</td>
<td>23.4</td>
</tr>
<tr>
<td>1440x2560</td>
<td>12.2</td>
</tr>
</tbody>
</table>

No official VR option

USA

<table>
<thead>
<tr>
<th>Resolution</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>640x1136</td>
<td>19.5</td>
</tr>
<tr>
<td>720x1280</td>
<td>5.7</td>
</tr>
<tr>
<td>750x1334</td>
<td>26.8</td>
</tr>
<tr>
<td>1080x1920</td>
<td>20.8</td>
</tr>
<tr>
<td>1440x2560</td>
<td>17.2</td>
</tr>
</tbody>
</table>

No official VR option

Standalone VR HMDs for the other half

(Source: DeviceAtlas)
Resolution: Traditional vs. 360 Video

<table>
<thead>
<tr>
<th>Traditional Video</th>
<th>360 Video Apparent Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>480p “SD”</td>
<td>Effectively unusable</td>
</tr>
<tr>
<td>720p “HD” (1280 x 720)</td>
<td>Effectively unusable</td>
</tr>
<tr>
<td>1080p “Full HD” (1920 x 1080)</td>
<td>480p “SD”</td>
</tr>
<tr>
<td>1440p “Quad HD” (2560 x 1440)</td>
<td>720p “HD”</td>
</tr>
<tr>
<td>“4K”</td>
<td>“Full HD”</td>
</tr>
</tbody>
</table>

The apparent resolution of 360 video is roughly “two steps down” from traditional.
## Resolution: Display vs. Video

<table>
<thead>
<tr>
<th>Typical</th>
<th>Traditional Display</th>
<th>Stereoscopic VR Eye Display</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>480p “SD”</td>
<td>Effectively unusable</td>
</tr>
</tbody>
</table>

| Good     | 720p “HD” (1280 x 720) | 640x720 (Cardboard) |

| Better   | 1080p “Full HD” (1920 x 1080) | 960x1080 (Oculus DK2, Pixel 1&2) |

| Best     | 1440p “Quad HD” (2560 x 1440) | 1280x1440 (GearVR; Pixel 1&2 XL; WinMR) Rift/Vive is 1080x1200 |

| Unobtainium | “4K” | 2Kx2K TBD |

### 360 Video Apparent Resolution

<table>
<thead>
<tr>
<th>Traditional Video</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effectively unusable</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>720p “HD” (1280 x 720)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>480p “SD”</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>1080p “Full HD” (1920 x 1080)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>720p “HD”</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>1440p “Quad HD” (2560 x 1440)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>“Full HD”</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>“Quad HD” / “4K”</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>“6K” / “8K” +</th>
</tr>
</thead>
</table>
# Bandwidth: Resolutions and Rates

<table>
<thead>
<tr>
<th>Display Resolution</th>
<th>Uncompressed Pixels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Typical</td>
<td></td>
</tr>
<tr>
<td>480p “SD”</td>
<td>~ 0.35 M pixel</td>
</tr>
<tr>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>720p “HD” (1280 x 720)</td>
<td>~ 0.9 M pixels</td>
</tr>
<tr>
<td>Better</td>
<td></td>
</tr>
<tr>
<td>1080p “Full HD” (1920 x 1080)</td>
<td>~ 2.1 M pixels</td>
</tr>
<tr>
<td>Best</td>
<td></td>
</tr>
<tr>
<td>1440p “Quad HD” (2560 x 1440)</td>
<td>~ 3.7 M pixels</td>
</tr>
<tr>
<td>Unobtainium</td>
<td></td>
</tr>
<tr>
<td>“4K”</td>
<td>~ 7 to 8 M pixels</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>360 Video Apparent Resolution</th>
<th>Traditional Video</th>
<th>Typical Bit Rate, Mezzanine</th>
<th>Typical Bit Rate, Download</th>
<th>Typical Bit Rate, Stream</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effectively unusable</td>
<td>720p “HD” (1280 x 720)</td>
<td>12 to 25+ Mbps</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>480p “SD”</td>
<td>1080p “Full HD” (1920 x 1080)</td>
<td>30 to 40+ Mbps</td>
<td>—</td>
<td>&lt; 5 Mbps</td>
</tr>
<tr>
<td>720p “HD”</td>
<td>1440p “Quad HD” (2560 x 1440)</td>
<td>—</td>
<td>—</td>
<td>5 to 8 Mbps</td>
</tr>
<tr>
<td>“Full HD”</td>
<td>“4K”</td>
<td>100 Mbps+</td>
<td>10 to 30+ Mbps</td>
<td>8 to 10 Mbps</td>
</tr>
<tr>
<td>“Quad HD” / “4K”</td>
<td>“6K” / “8K” +</td>
<td>TBD</td>
<td>TBD</td>
<td>TBD</td>
</tr>
</tbody>
</table>

- 360 Video Apparent Resolution:
- Traditional Video
- Typical Bit Rate, Mezzanine
- Typical Bit Rate, Download
- Typical Bit Rate, Stream
Several / many cameras are often stitched together into a few streams.

“Delivery” often means multiple hierarchical CDNs with many edge nodes.

Popular content has many viewers.

MANY INSTANCES

SEVERAL / MANY
Formats and Interchange: Projections

- Multiple rectangular / fisheye images
  - Cubemap
    - Cubemap (compact)
    - Cylindrical
  - Equirectangular

(Rendered image)
Formats and Interchange: System Perspective

Source: “Evolution of immersive media” slide 9, MPEG120 m41901

Several / many cameras are often stitched together into a few streams

Capture e.g. rectangular / fisheye, etc.

Format(s)

Stage

Capture

Rectangular

Fisheye

Cylindrical

Interchange

Equirectangular

Cubemap

Cylindrical

Distribution

(see above)

(see above)

Interchange e.g. equirectangular

One per instance; latency critical

Distribution e.g. equirectangular; cubemap

Distribution e.g. equirectangular; cubemap

Distribution e.g. equirectangular; cubemap

“Delivery” often means multiple hierarchical CDNs with many edge nodes

A FEW

A BUNCH

MANY

Popular content has many viewers

MANY INSTANCES

SEVERAL / MANY

Many instances

One per instance; latency critical

Format(s)

Texture

Shader renders onto playback surface

Eye view transform, etc.

Display

Audio rendering

Audio decoding

Audio Metadata

File/segment decapsulation

Metadata

Head/eye tracking

Orientation/viewport metadata

VR service platform

VR Application

VR Service Provider

Audio preprocessing

Audio encoding

Audio Metadata

File/segment encapsulation

VR Service Provider

Visual preprocessing

Video encoding

Video Metadata

File/segment encapsulation

VR Service Provider

Original Content (Footage)

Original Content (Footage)

Original Content (Footage)

Original Content (Footage)

Original Content (Footage)

Original Content (Footage)

VR Content Provider

Visual stitching

Bin 1

Bin 2

Bin 3

Bin 4

Bin 5

VR Service Provider

Audio Metadata

File/segment encapsulation

VR Service Provider

Visual Metadata

File/segment encapsulation

VR Service Provider

Texture

Shader

Displays

Audio/visual/other

Eye view transform, etc.

Display

Audio rendering

Audio decoding

Audio Metadata

File/segment decapsulation

Metadata

Head/eye tracking

Orientation/viewport metadata

VR service platform

VR Application

VR Service Provider

Original Content (Footage)

Original Content (Footage)

Original Content (Footage)

Original Content (Footage)

Original Content (Footage)

Original Content (Footage)
Formats and Interchange: System Perspective

Stage | Format(s)
--- | ---
Capture | Rectangular, Fisheye, Cylindrical
Interchange | Equirectangular, Cubemap, Cylindrical
Distribution | (see above)
Texture | (see above)

Capture: Multiple rectangular / fisheye images
Interchange: Single equirectangular image
Distribution: Texture

Spherical playback surface:
- Expensive surface w/simple shader,
- or simple surface w/expensive shader

Editing
Processing
Production
Archival

Wide tools support for equirectangular
Untouched; avoid quality loss
Formats and Interchange: System Perspective

- **Capture**
  - Multiple rectangular / fisheye images

- **Interchange**
  - Single equirectangular image
  - Single cubemap image
  - Editing
  - Processing
  - Production
  - Archival

- **Distribution**
  - Propagation
  - Delivery

- **Texture**
  - Rendering

**Stage**
- **Capture**
  - Rectangular
  - Fisheye
  - Cylindrical
- **Interchange**
  - Equirectangular
  - Cubemap
  - Cylindrical
- **Distribution**
  - (see above)
- **Texture**
  - (see above)

**Format(s)**
- **Capture**
  - Rectangular
  - Fisheye
  - Cylindrical
- **Interchange**
  - Equirectangular
  - Cubemap
  - Cylindrical
- **Distribution**
  - (see above)
- **Texture**
  - (see above)

- **Wide tools support for equirectangular**
- **Transform required; quality loss**
- **Cubic playback surface:** Simple surface w/simple shader
Formats and Interchange: System Perspective

Multiple rectangular / fisheye images | Single cubemap image | (same) | (same)

Capture | Stitch | Interchange | Distribution | Texture

Editing | Processing | Production | Archival | Propagation | Delivery | Rendering

Cubemap less intuitive for editing | Untouched; avoid quality loss | Cubic playback surface: Simple surface w/simple shader

Stage | Format(s)
Capture | Rectangular | Fisheye | Cylindrical
Interchange | Equirectangular | Cubemap | Cylindrical
Distribution | (see above) | (see above)
Texture

Q: Can interchange steps leverage cubemap?
Formats and Interchange: System Perspective

Capture
- Single cylindrical image
- Cylindrical still intuitive for editing
- Cylindrical vertical FOV reduction

Interchange
- Editing
- Processing
- Production
- Archival

Distribution
- Propagation
- Delivery
- Untouched; avoid quality loss
- Cylindrical playback surface: Less expensive surface w/simple shader, or simple surface w/less expensive shader

Texture
- Rendering

Stage | Format(s)
--- | ---
Capture | Rectangular, Fisheye, Cylindrical
Interchange | Equirectangular, Cubemap, Cylindrical
Distribution | (see above)
Texture | (see above)
The future is now: cloud rendering

Examples… on Xfinity’s X1 service today

<table>
<thead>
<tr>
<th>360 video on X1, via the cloud</th>
<th>3D rendered games on X1, via the cloud</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fifty Shades Darker: Masquerade Ball</td>
<td>The Walking Dead; Free Games</td>
</tr>
</tbody>
</table>

- Why? Simple… Smartphones >> STBs
  - Sure, on high end. But what about lesser models?
Looking forward: Not your parents’ TV

• Lightfield displays
  • Projecting light rays to myriad orientations
    • Myriad orientations = myriad bandwidth multiples
  • Are raster representations feasible? If not, what gets sent for display?
    • Hybrid natural / synthetic scene (HNSS) container
🎶 For the times they are a-changin’ 🎶

–Bob Dylan