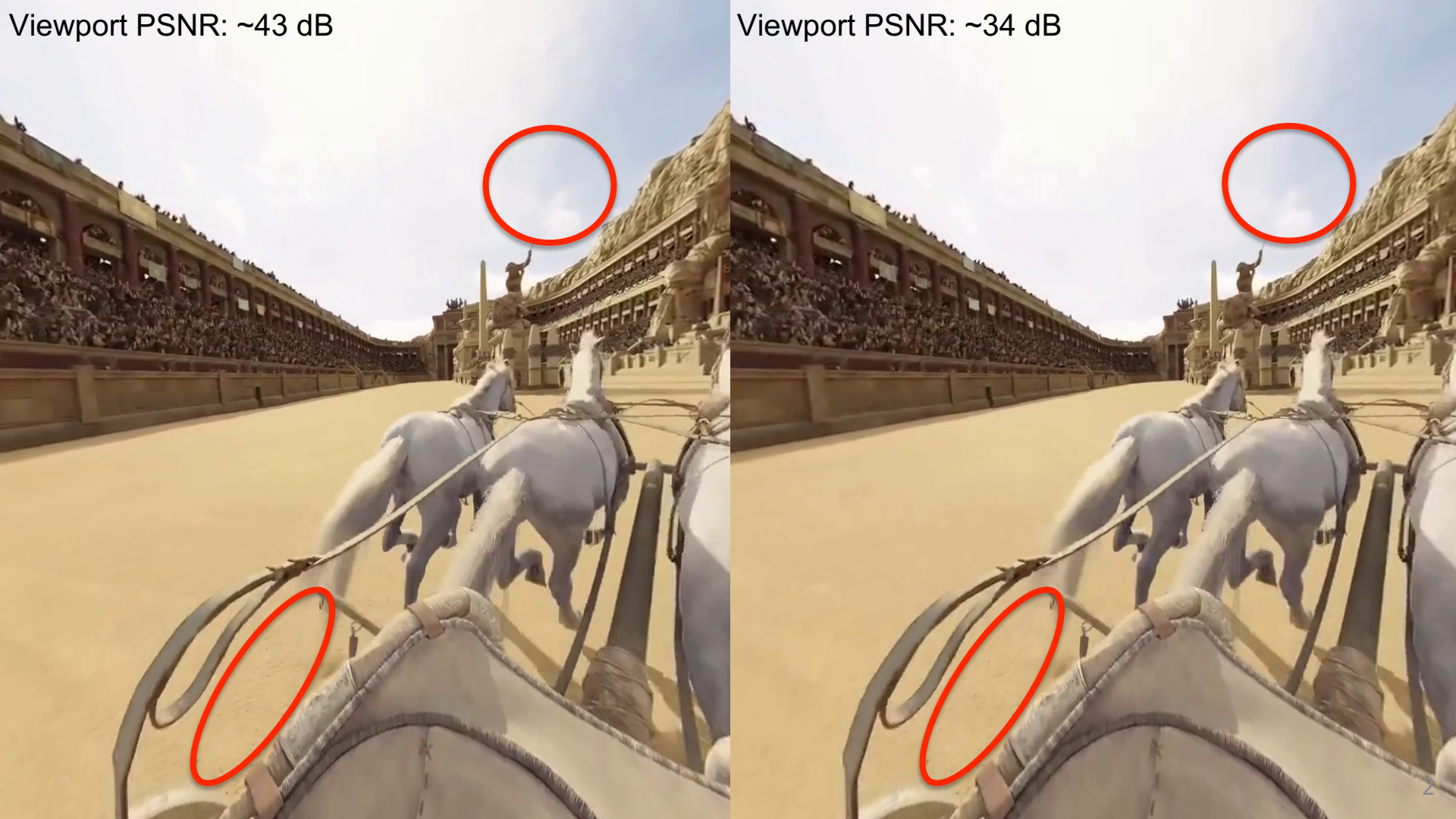
Modeling Quality-of-Experience of 360° Videos in Head-Mounted Virtual Reality

Shun-Huai Yao

Advisor: Cheng-Hsin Hsu

Networking and Multimedia Systems Lab

Institute of Information Systems and Applications, NTHU

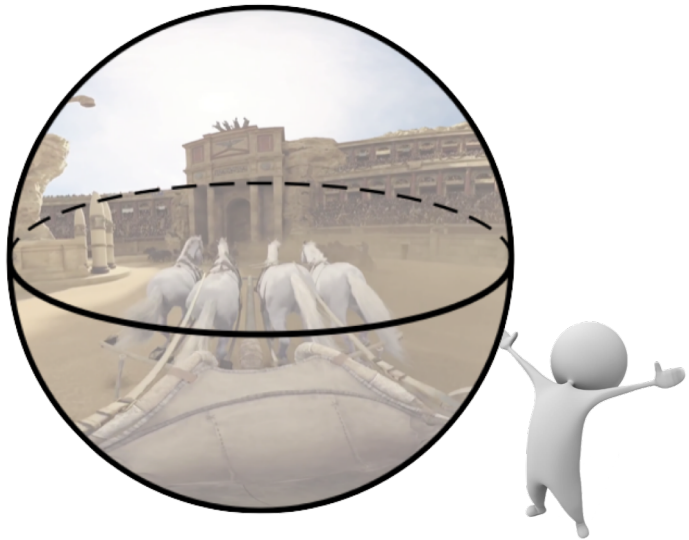

Viewport PSNR: ~43 dB
Viewport PSNR: ~34 dB

Outline

- Introduction
- 360° Video Player
- User Study
- QoE Models
- Conclusions

Virtual Reality (VR) 360° Video is Booming

- 360° video in VR is interactive and immersive.

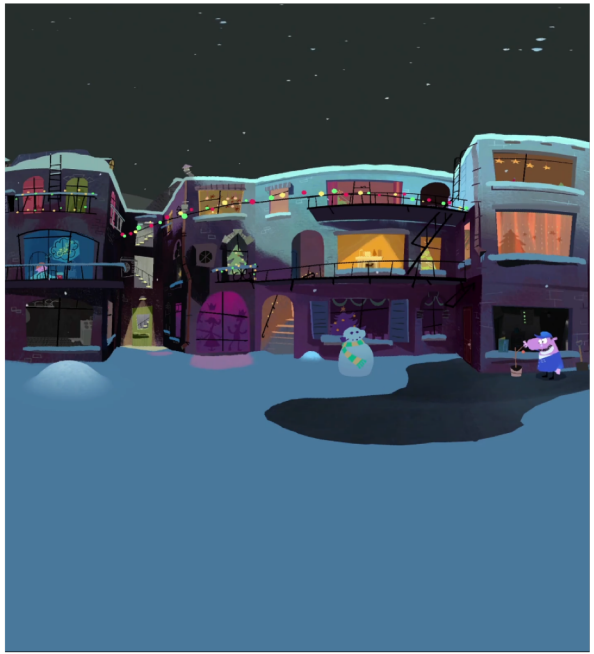


Streaming 360° Video Requires Vast Network Bandwidth

- Acceptable qualities of 360° videos in HMDs require high resolutions and bitrates.

Encoding Quantization Parameters (QPs)

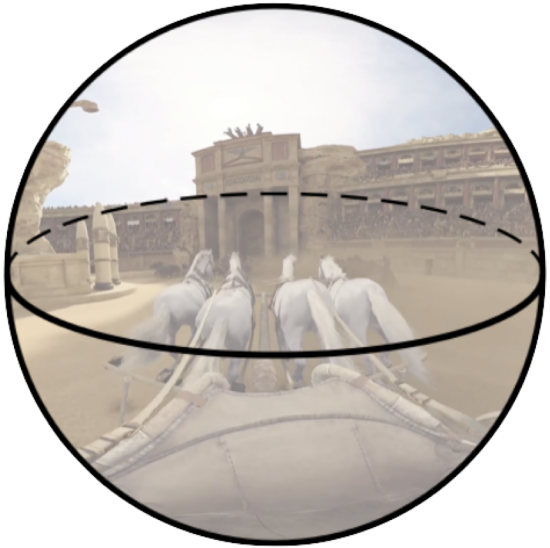
Different Video Genres



Simple, Slow-paced

Complex, Fast-paced

Projection Schemes



360° video

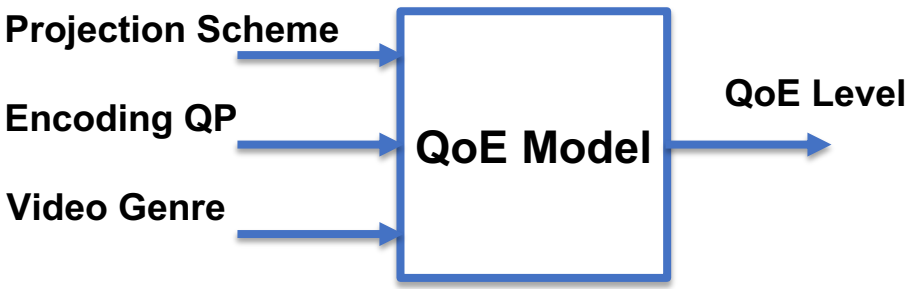
2D rectangular video

Different Projection Schemes Cause Diverse Shape Distortions

Equi-rectangular

Quality-of-Experience (QoE) Models are Needed

- ~~Human ratings are time-consuming.~~
- ~~Objective quality metrics cannot quantify the QoE accurately [1].~~



[1] B. Zhang, J. Zhao, S. Yang, Y. Zhang, J. Wang, and Z. Fei, "Subjective and objective quality assessment of panoramic videos in virtual reality environments," in *Proc. of IEEE International Conference on Multimedia and Expo Workshops (ICMEW'17)*, Hong Kong, China, July 2017, pp. 163–168

Contributions

- Realizing a 360° video player supporting several projection schemes
- Conducting a user study to understand the impacts of factors on QoE
- Constructing QoE models to predict QoE levels

Outline

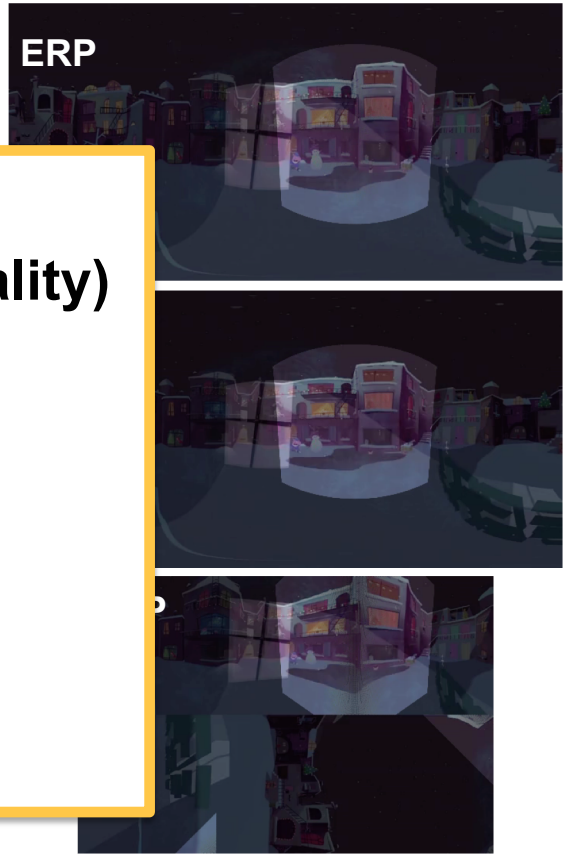
- Introduction
- 360° Video Player
- User Study
- QoE Models
- Conclusions

Testbed Overview

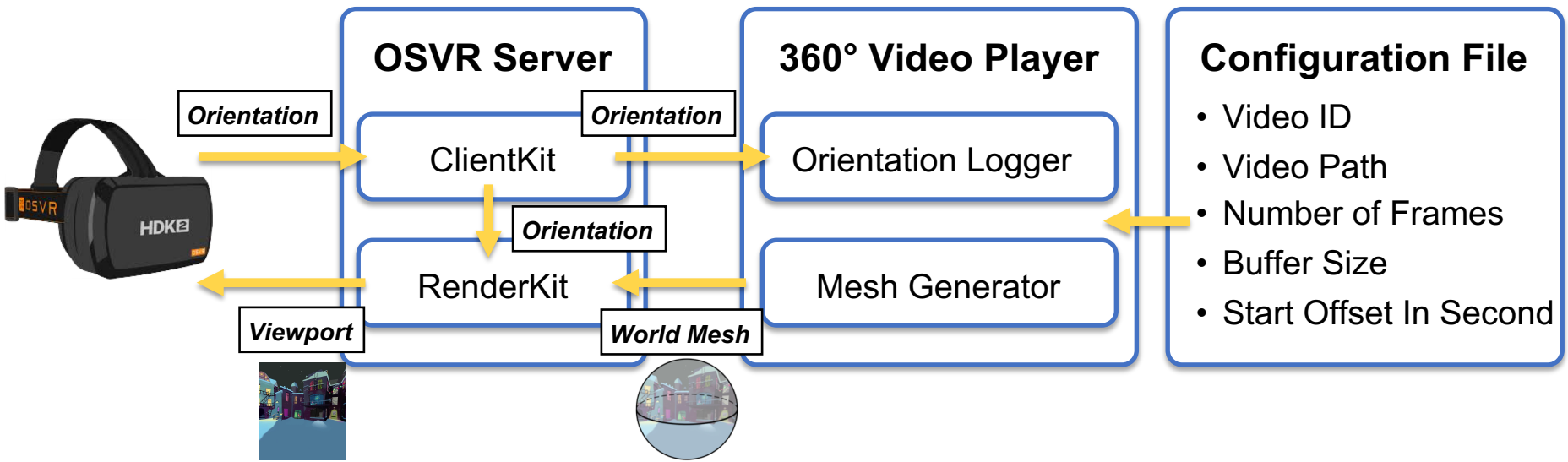
OSVR (Open-Source Virtual Reality)

- Windows and Linux
- Any HMDs

Projections

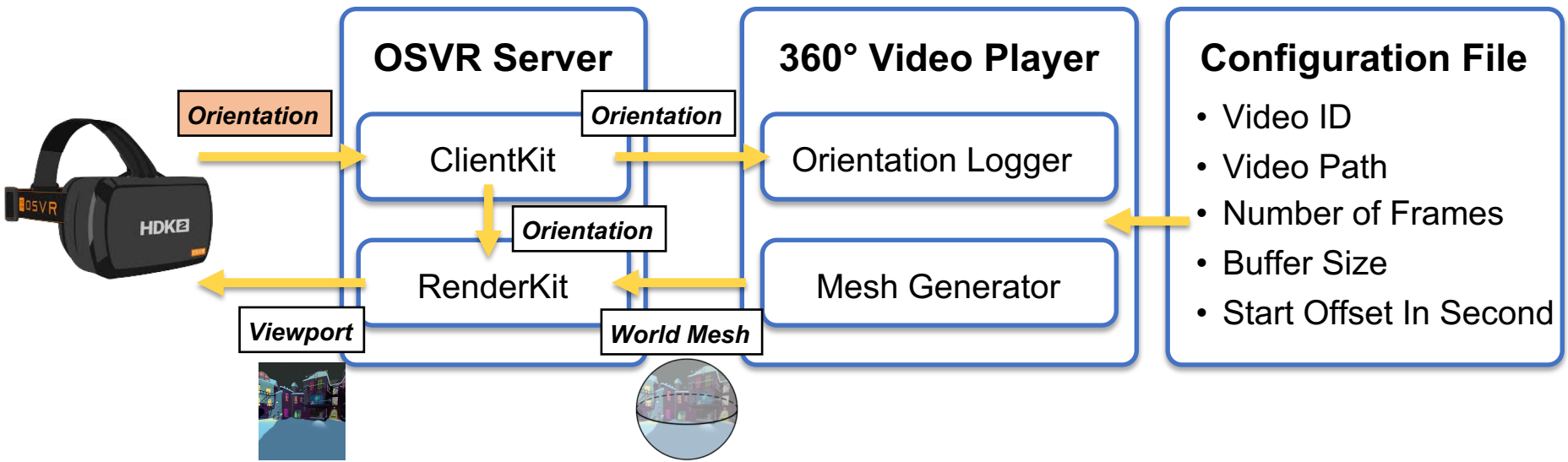


OSVR-Based 360° Video Player



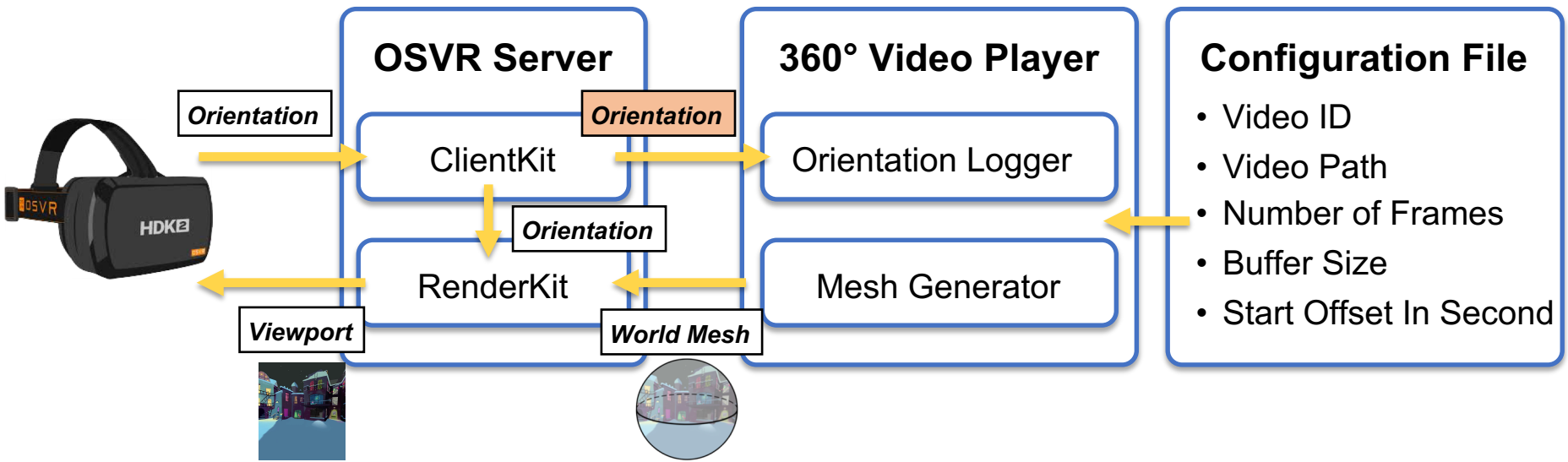
[2] X. Corbillon, F. De Simone, and G. Simon. 360-degree video head movement dataset. In *Proc. of the ACM on Multimedia Systems Conference (MMSys'17)*, pages 199–204, Taipei, Taiwan, June 2017.

OSVR-Based 360° Video Player



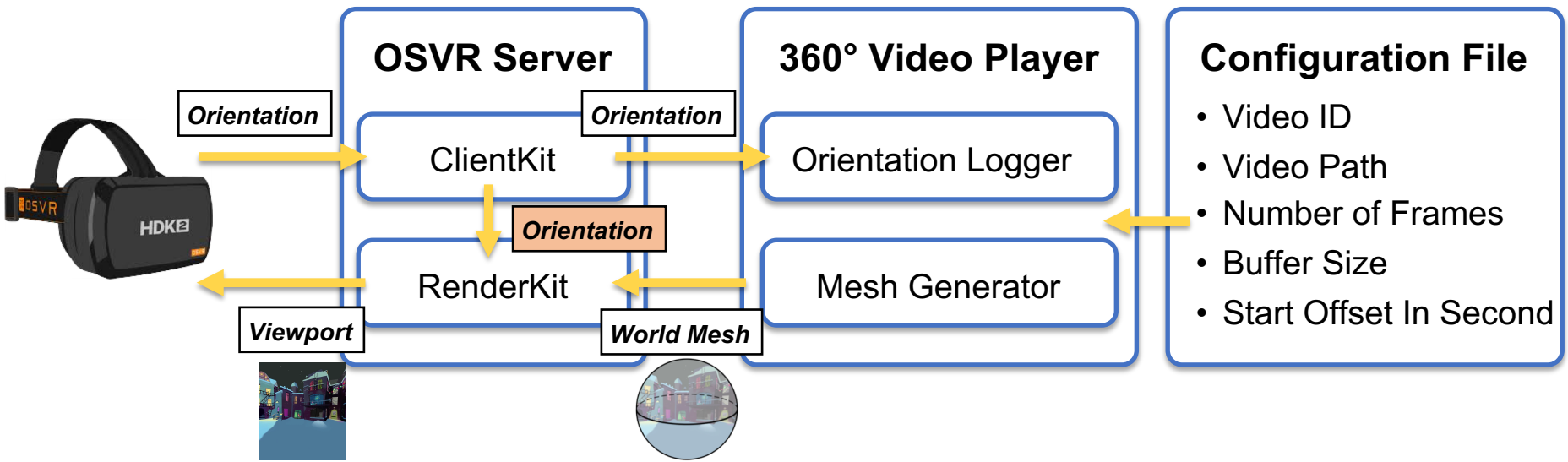
[2] X. Corbillon, F. De Simone, and G. Simon. 360-degree video head movement dataset. In *Proc. of the ACM on Multimedia Systems Conference (MMSys'17)*, pages 199–204, Taipei, Taiwan, June 2017.

OSVR-Based 360° Video Player



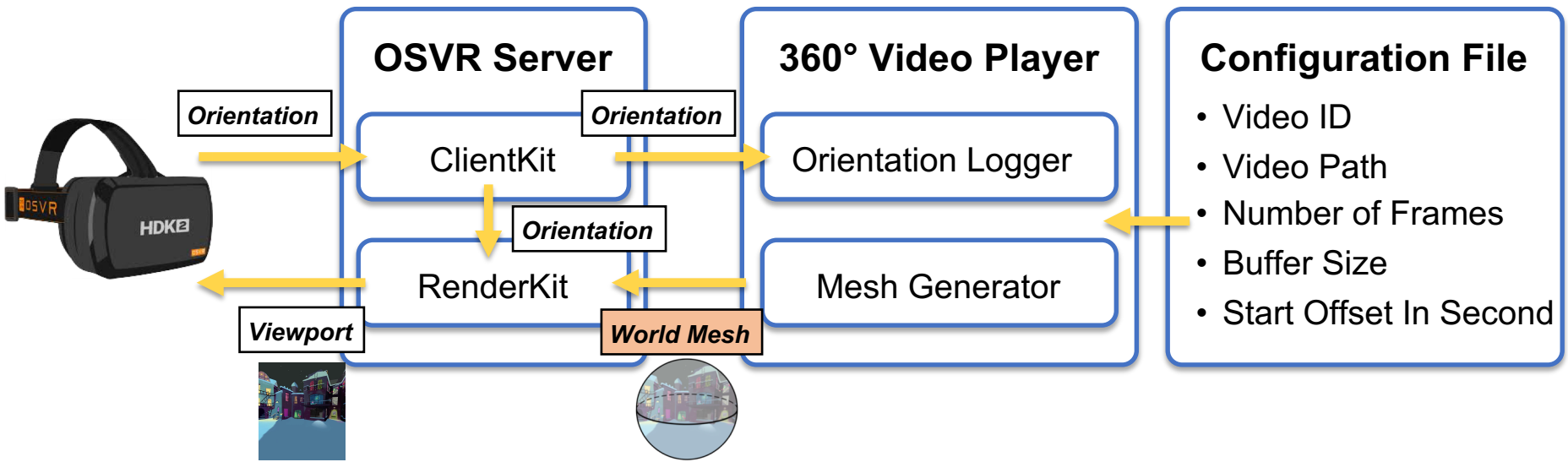
[2] X. Corbillon, F. De Simone, and G. Simon. 360-degree video head movement dataset. In *Proc. of the ACM on Multimedia Systems Conference (MMSys'17)*, pages 199–204, Taipei, Taiwan, June 2017.

OSVR-Based 360° Video Player



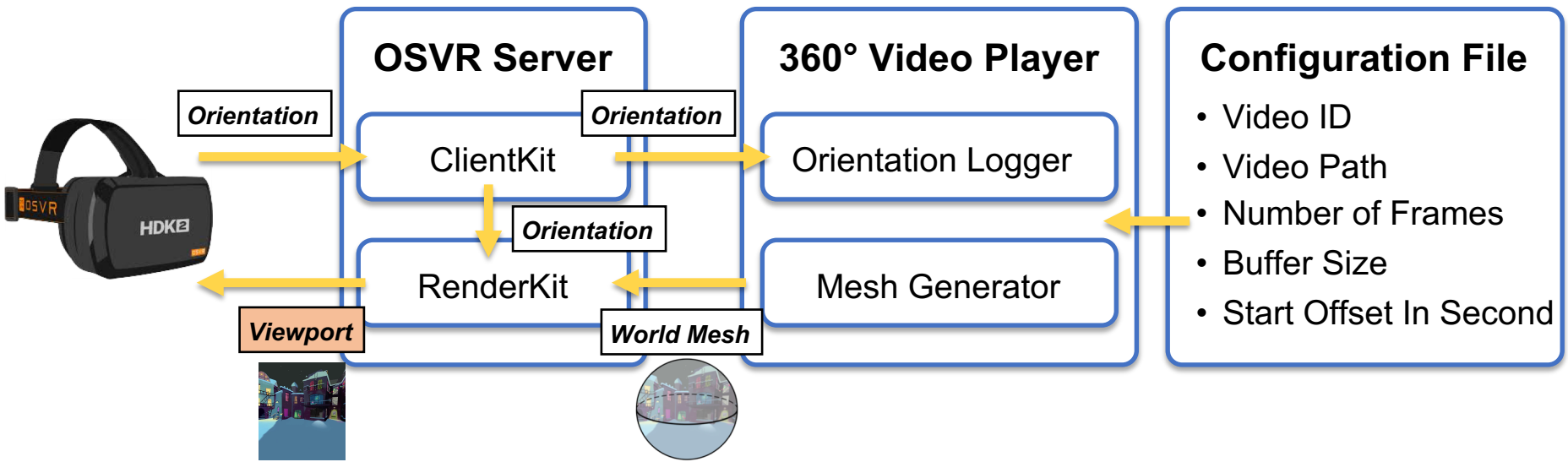
[2] X. Corbillon, F. De Simone, and G. Simon. 360-degree video head movement dataset. In *Proc. of the ACM on Multimedia Systems Conference (MMSys'17)*, pages 199–204, Taipei, Taiwan, June 2017.

OSVR-Based 360° Video Player



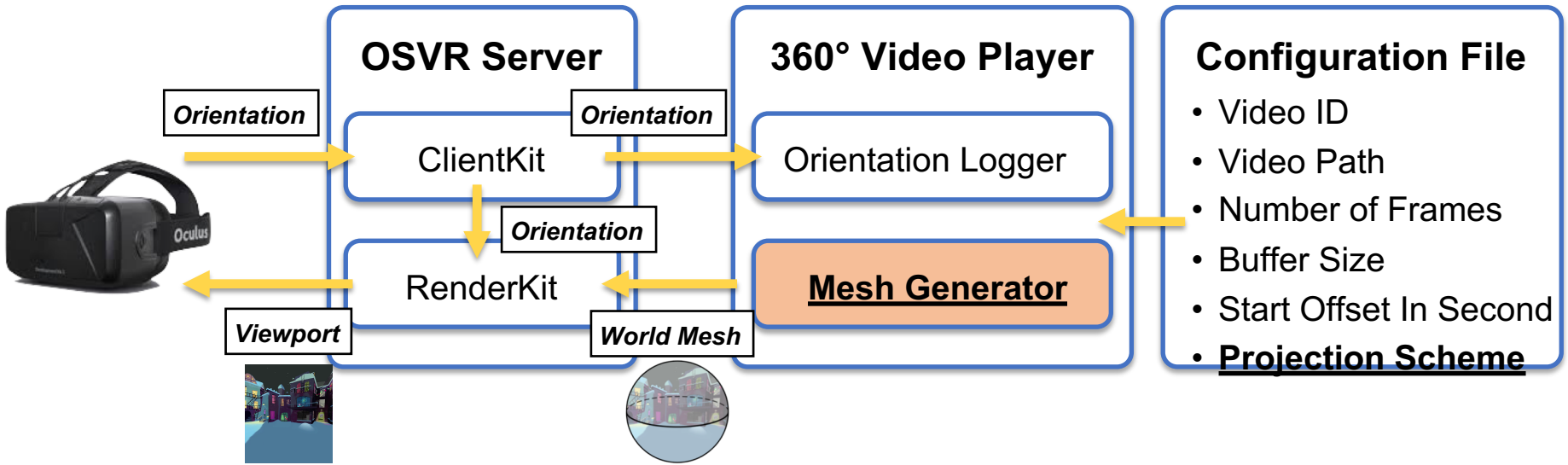
[2] X. Corbillon, F. De Simone, and G. Simon. 360-degree video head movement dataset. In *Proc. of the ACM on Multimedia Systems Conference (MMSys'17)*, pages 199–204, Taipei, Taiwan, June 2017.

OSVR-Based 360° Video Player



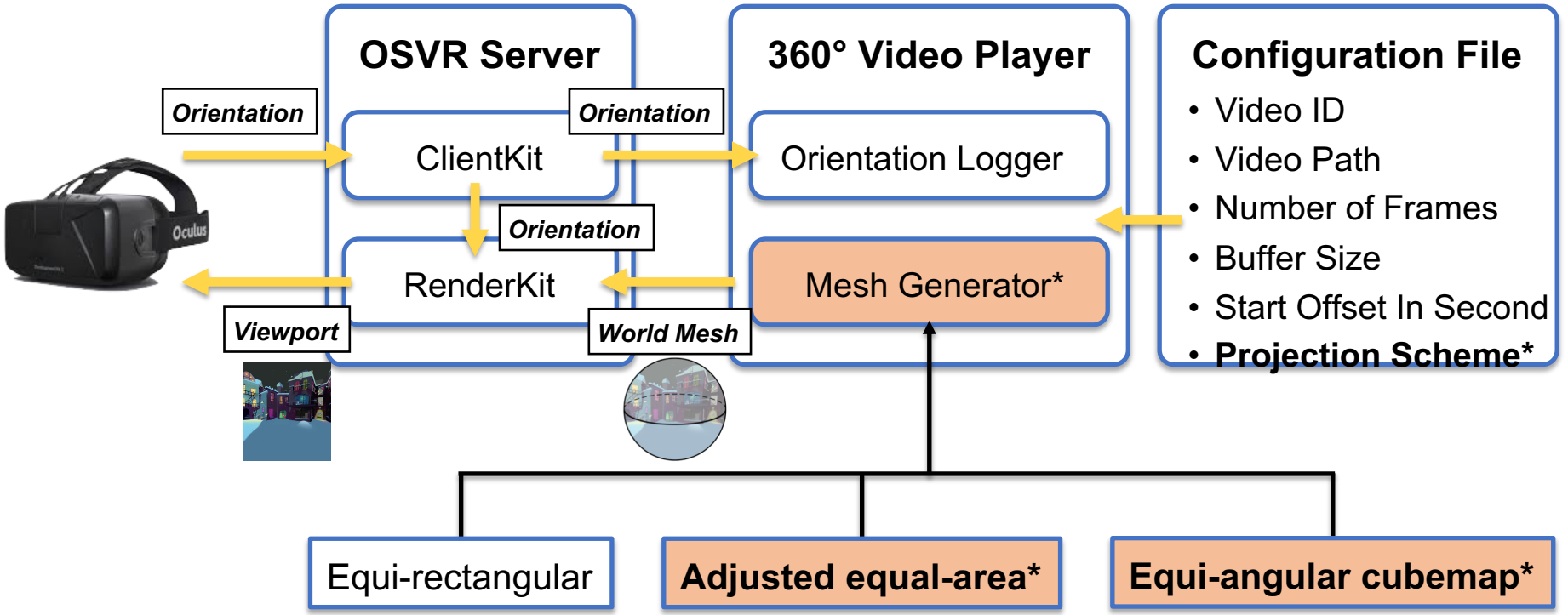
[2] X. Corbillon, F. De Simone, and G. Simon. 360-degree video head movement dataset. In *Proc. of the ACM on Multimedia Systems Conference (MMSys'17)*, pages 199–204, Taipei, Taiwan, June 2017.

Diverse Projection Schemes



[3] S. Yao, "A 360 degree video player and head movement logger for an HMD using the OSVR API," 2018, https://github.com/shunhuaiyao/360Degree_Head_Movement_Dataset/tree/win32

Our 360° Video Player



Why Choose These Projection Schemes?

- Equi-Rectangular Projection (ERP)
- Adjusted Equal-area Projection (AEP)
- Equi-angular Cubemap Projection (ECP)

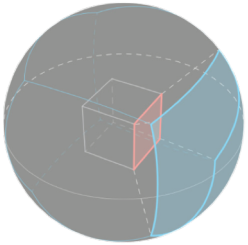
Equi-Rectangular Projection (ERP)

- The most common projection scheme

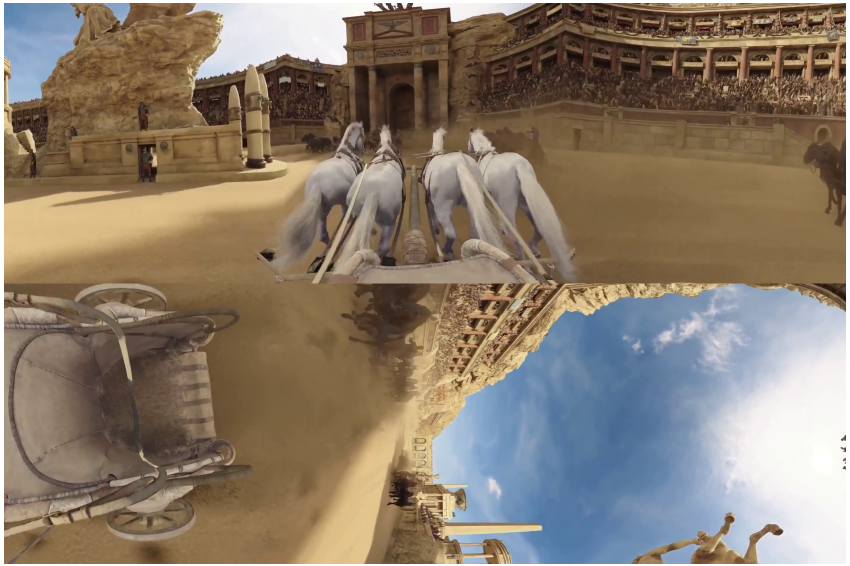
Adjusted Equal-area Projection (AEP)

- This compensates high horizontal sampling density closed to poles.

Equi-angular Cubemap Projection (ECP)



- This projects to 6 faces of its circumscribed cube.

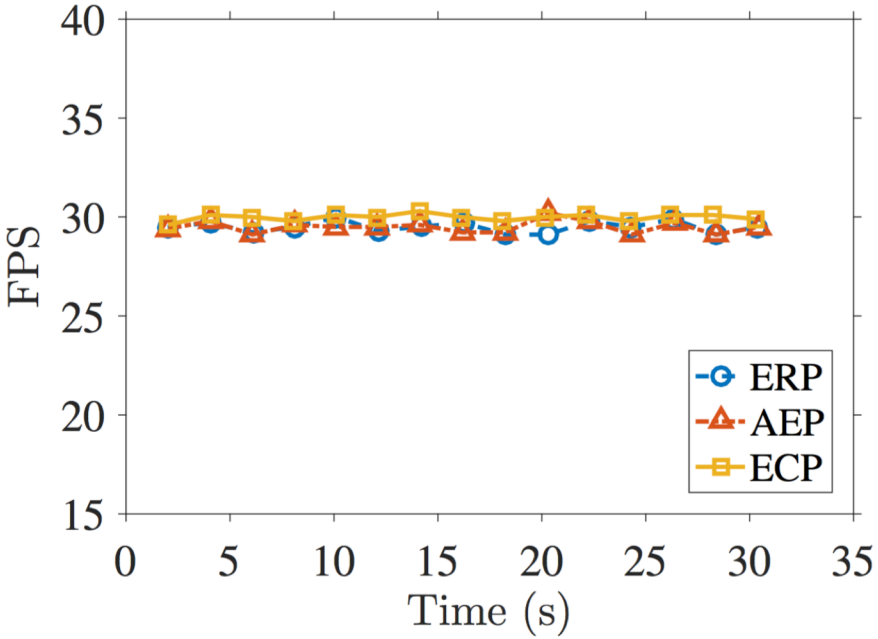


The Average Resource Consumption at 30/45 FPS

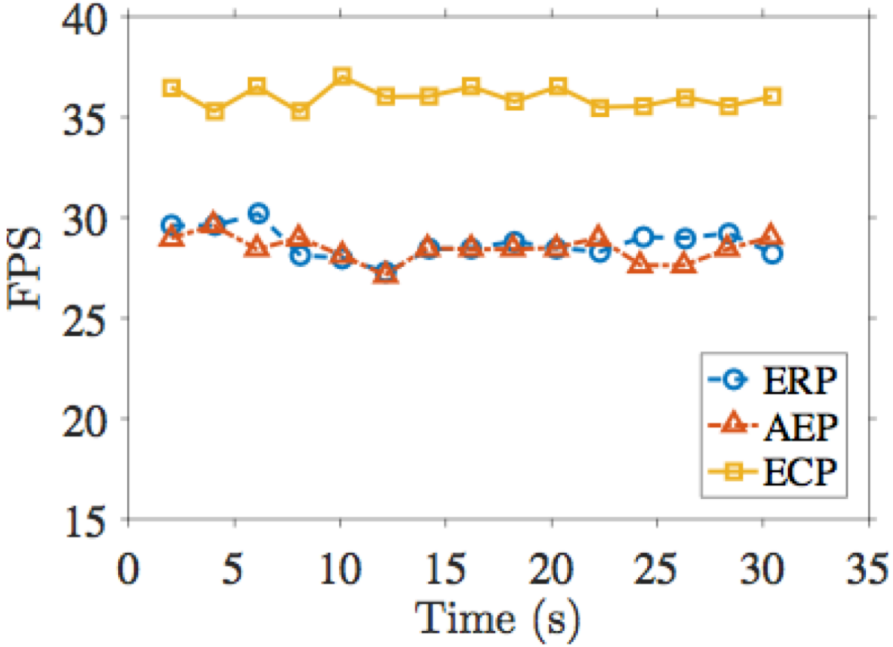
- ECP achieves slightly lower resource consumption.

Projection	CPU Load (%)	RAM (MB)	GPU Load (%)
ERP	22.81/23.54	1271.82/1271.56	15.91/15.47
AEP	21.80/23.35	1267.82/1271.58	16.66/15.81
ECP	20.38/22.56	967.65/971.57	12.72/14.81

Achieved Frame Rates of Different Projection Schemes



30 FPS



45 FPS

4K 360° Videos at 30 FPS

OSVR Server

ClientKit

RenderKit

360° Video Player

Orientation Handler

Mesh Generator

Projections

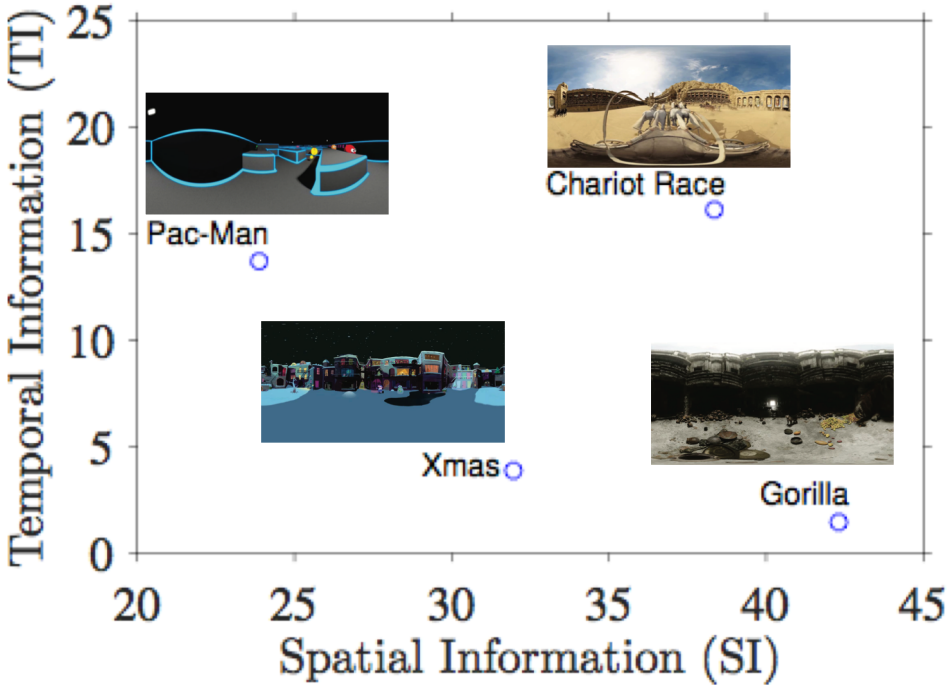
Outline

- Introduction
- 360° Video Player
- User Study
- QoE Models
- Conclusions

User Study Design

- Dependent variable
  - **Subjective opinion score**, ranged from 1 to 9
- Independent variables
  - **Projection scheme** (ERP, AEP, and ECP) - *streaming system design*
  - **Encoding QP** (22, 30, and 38) - *video codec*
  - **Temporal video genre** (Slow- versus Fast-paced) - *video genre*
  - **Spatial video genre** (Simple versus Complex) - *video genre*

4 Videos Categorized by Spatial and Temporal Complexity



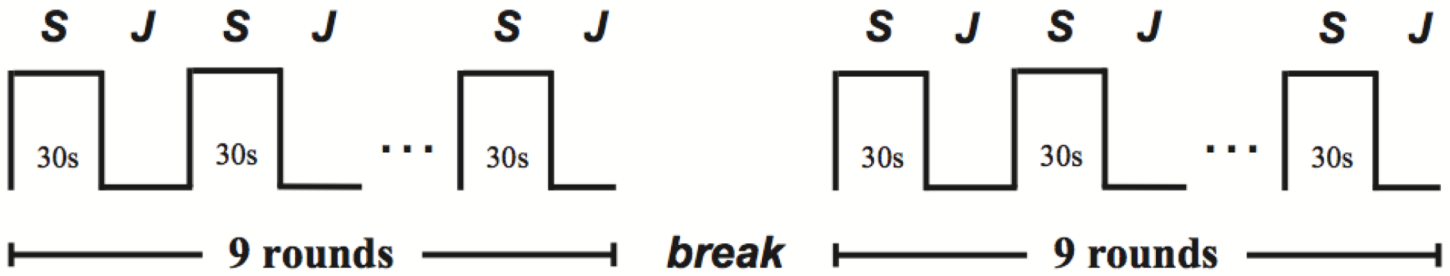
[4] ITU-T, "P.910 Subjective video quality assessment methods for multimedia applications," Tech. Rep., April 2008.

Mixed-Design Alleviating Subjects' Fatigues

- Within-subject variables
  - Projection scheme, Encoding QP, and Temporal video genre
- Between-subject variable
  - Spatial video genre
- Each subject only scores either simple or complex spatial video genre (18 test videos).

User Study Procedure

- Question: How is your overall experience about this 360° video? (9-point scale)



S: Stimulation phase, J: Judgement phase

Subjects in User Study

- 60 recruited subjects
  - 30 subjects in either simple or complex video group
- 34 males v.s. 26 females
- Age range: 19-36 years old

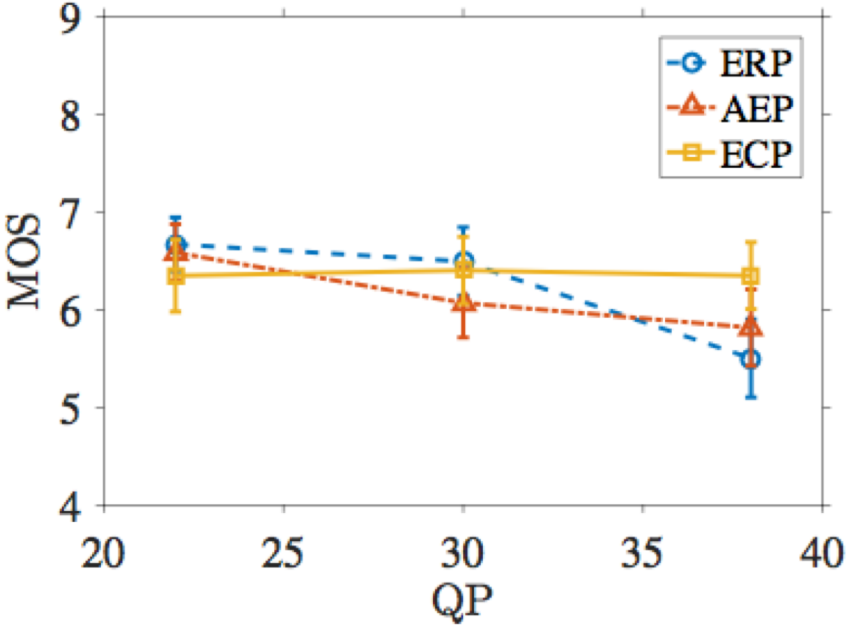
Finding 1: Projection Schemes Alone Have no Impact on QoE

- ECP provides a better experience at low quality videos

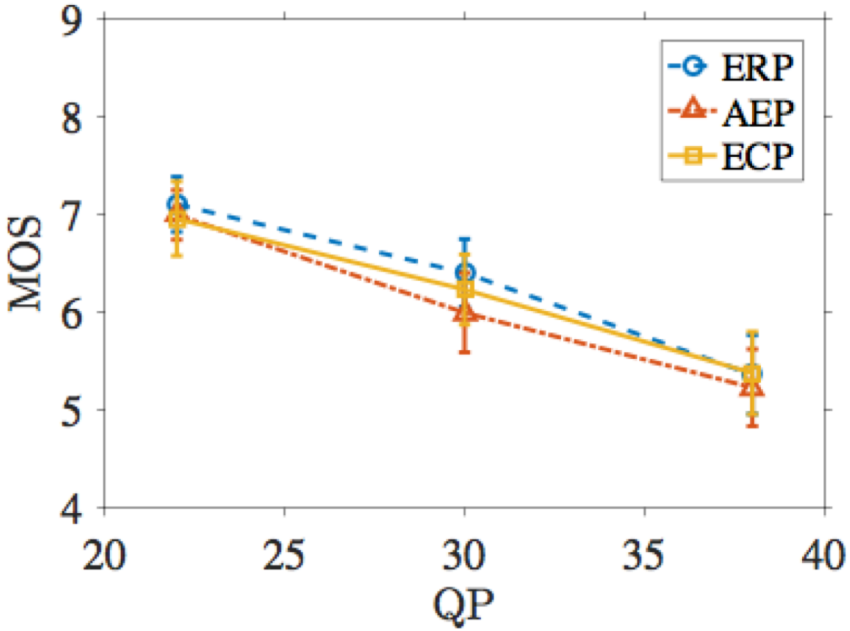
MOS Score (Standard Deviation)

Projection	Encoding QP		
	22	30	38
ERP	**6.884 (0.204)**	**6.451 (0.250)**	5.433 (0.289)
AEP	6.789 (0.200)	6.033 (0.273)	5.521 (0.288)
ECP	6.640 (0.276)	6.319 (0.253)	**5.844 (0.293)**

Finding 2: ECP Achieves the Highest QoE Level with Simple Videos

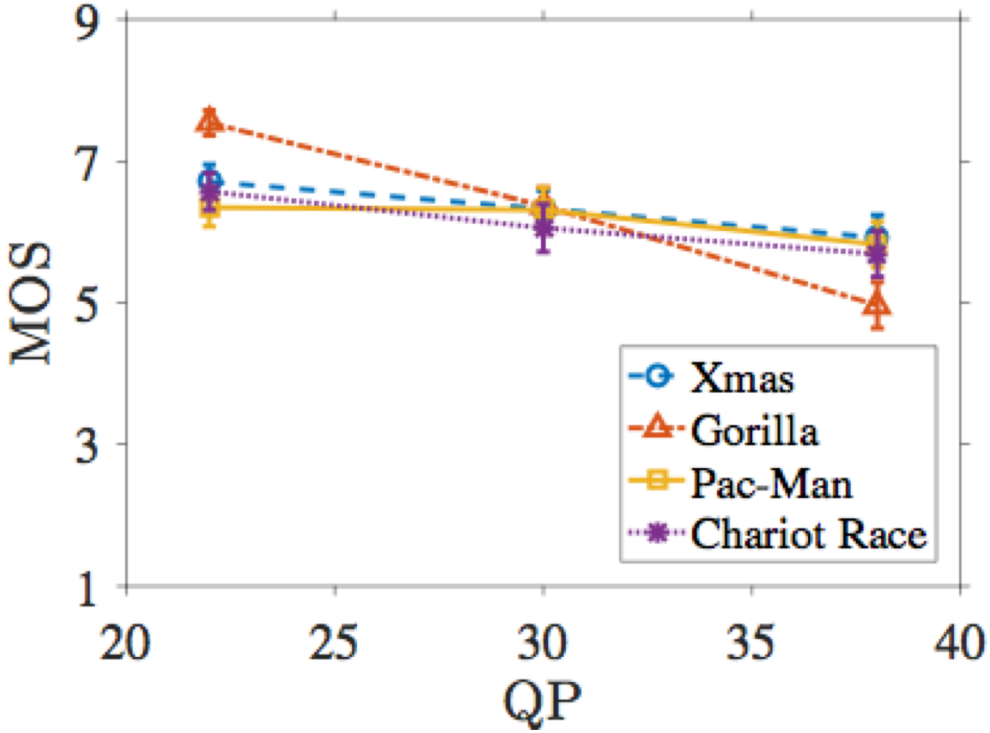


Simple

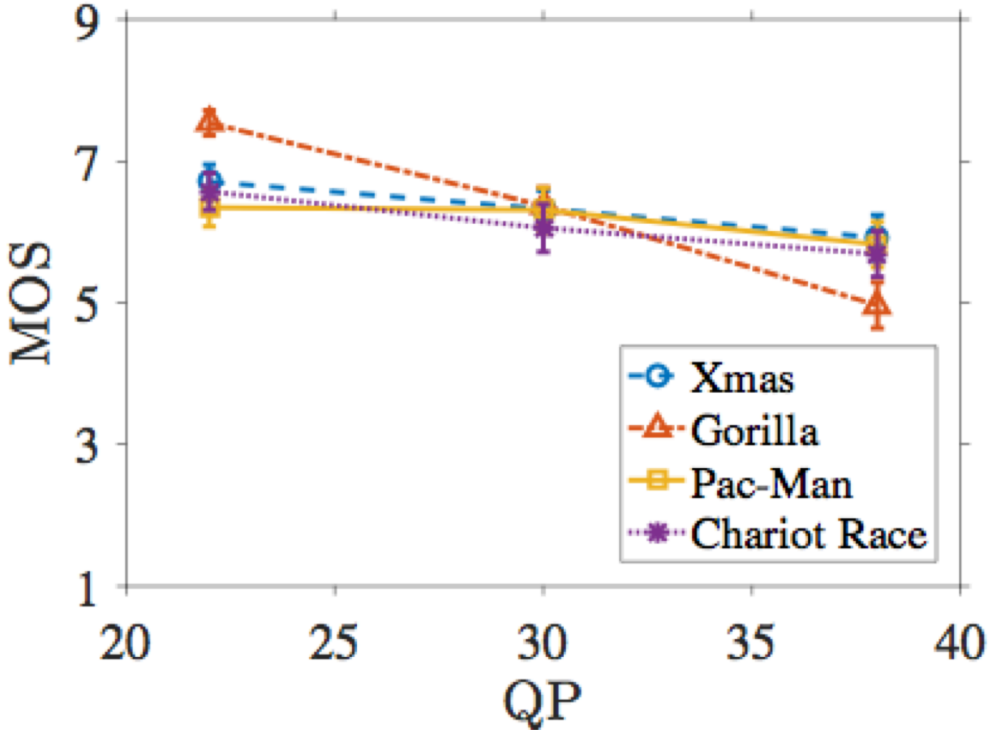


Complex

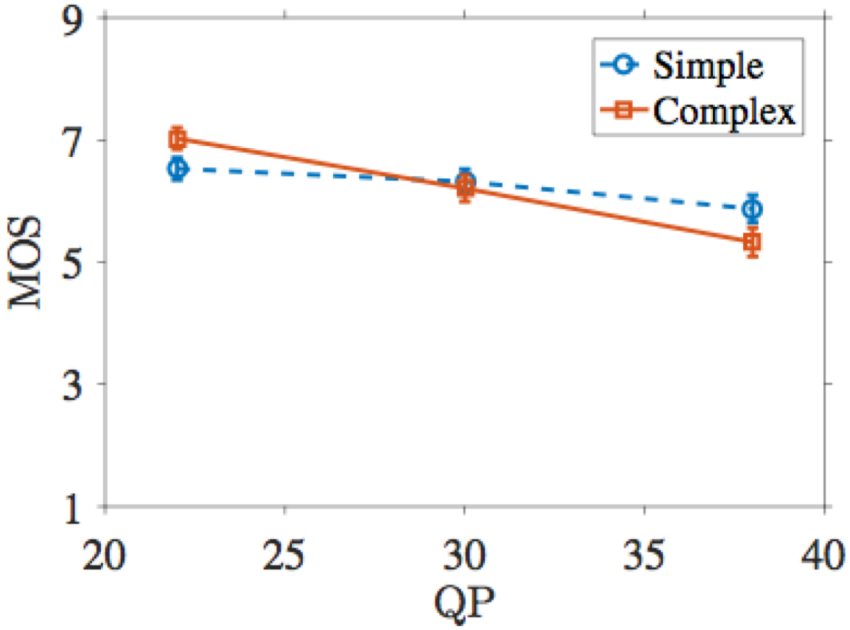
Finding 3: QoE of Videos Decreases as Encoding QP Rises



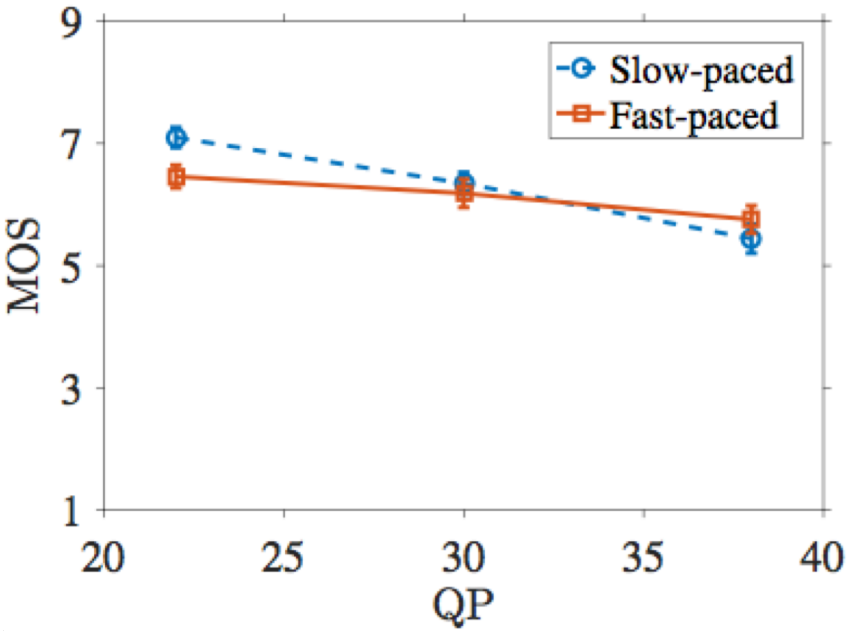
Finding 4: Encoding QP Has Different Effects on Different Videos



Finding 5: QoE of Complex and Slow-Paced Videos are More Sensitive to QP Values



Spatial Video Genre



Temporal Video Genre

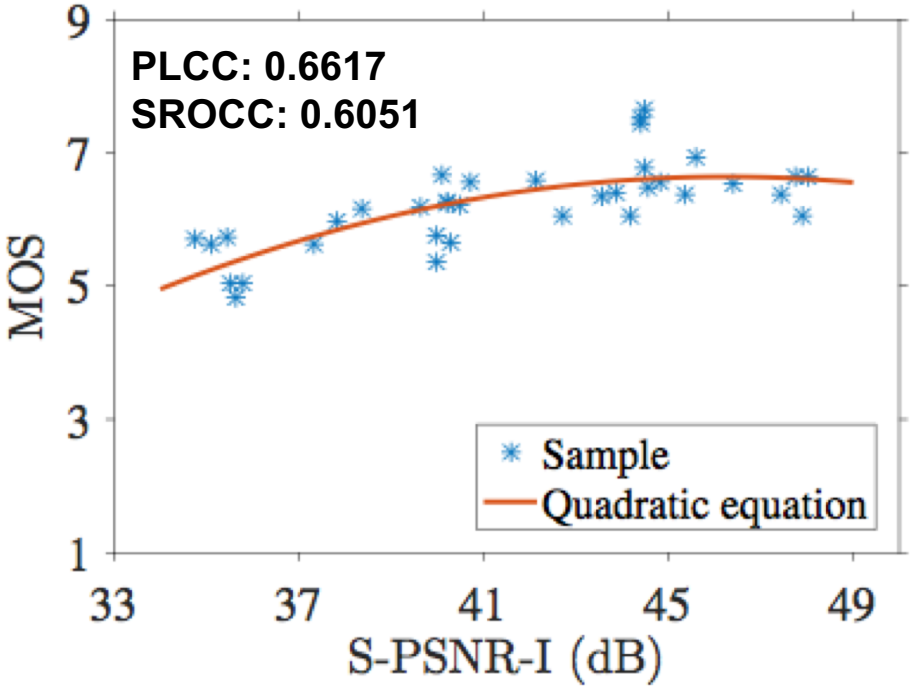
New Findings

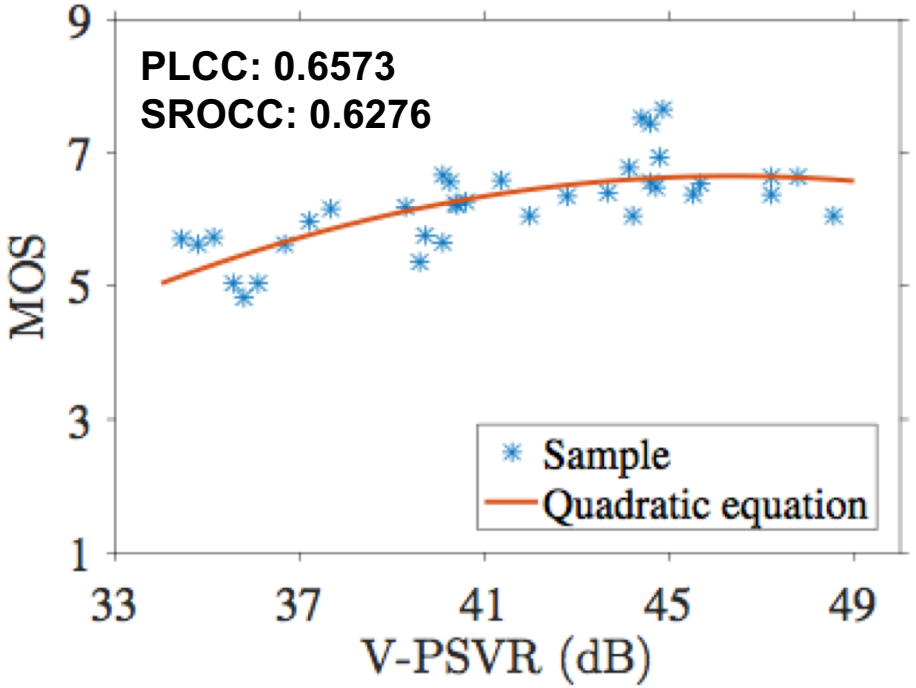
- Projection schemes alone have no impact on QoE.
- ECP achieves the highest QoE level with simple videos.
- QoE of complex and slow-paced videos are more sensitive to QP values.

Outline

- Introduction
- 360° Video Player
- User Study
- QoE Models
- Conclusions

Objective quality metrics alone are not good indicators for QoE





Several Potential Factors are Considered

- Projection scheme (P)
- Encoding QP (Q)
- Spatial video genre (S)
- Temporal video genre (T)
- S-PSNR-I (q_{spsnr})
- V-PSNR (q_{vpsnr})

Factor	DF	Sum Square	F Ratio	p-value
P	2	3.4110	0.722	0.4859
Q	1	151.7627	70.784	$< .0001^*$
S	1	1.7654	0.748	0.3874
T	1	2.2804	0.967	0.3259
$P \times Q$	2	33.2184	7.166	0.0008^*
$P \times S$	2	2.2222	0.470	0.6250
$P \times T$	2	0.9990	0.211	0.8096
$Q \times S$	1	28.4052	12.236	0.0005^*
$Q \times T$	1	20.3240	8.711	0.0033^*
$S \times T$	1	0.0718	0.030	0.8616
$P \times Q \times S$	2	7.5008	1.593	0.2042
$P \times Q \times T$	2	11.8743	2.528	0.0806
$P \times S \times T$	2	3.0190	0.639	0.5280
$Q \times S \times T$	1	15.5842	6.660	0.0101^*
$P \times Q \times S \times T$	2	0.2476	0.052	0.9490
q_{spsnr}	1	133.9045	61.715	$< .0001^*$
q_{vpsnr}	1	120.6699	55.132	$< .0001^*$

Stepwise Linear Regression

Algorithm 1 Stepwise Regression Algorithm

```
1: Initialize a linear regression model
2: Examine p-values for each factor
3: while available factors not included in the model yet have
   p-values smaller than the entering threshold do
4:     Add factor with smallest p-value into the model
5:     Re-calculate p-values for individual factors
6:     if all available factors included in the model have p-
   values smaller than the stopping threshold then
7:         Output the QoE model
8:     else
9:         Remove the factor with the worst p-value
10:        Re-calculate p-values for individual factors
11:    end if
12: end while
```

Adds the most significant factor into the model in each iteration

QoE Models Applied to Different Streaming Systems

(i) Video genres only

(ii) Projection schemes

(iii) Objective quality metrics

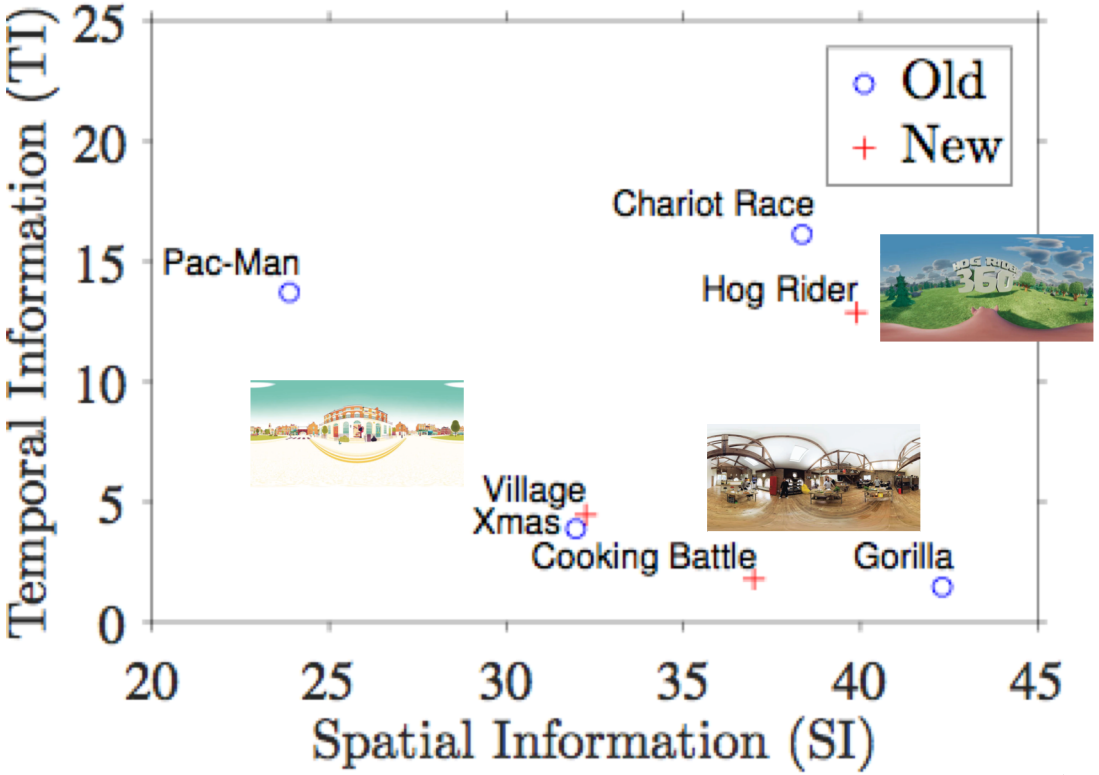
Model	Parameters
① : $\alpha_{1,1} + \alpha_{1,2}Q$	$6.22, -3.77$
② : ① + $\alpha_{2,3}QS$	$6.22, -3.77, 1.59$
③ : ② + $\alpha_{3,4}QT$	$6.22, -3.77, 1.59, -1.74$
④ : ③ + $\alpha_{4,5}QST$	$6.22, -3.77, 1.59, -1.74, 1.12$
⑤ : ④ + $\alpha_{5,6}PQ$	$6.22, -3.77, 1.59, -1.74, 1.12, 1.03$
⑥ : ⑤ + $\alpha_{6,7}q_{spsnr}$	$6.22, -3.58, 1.58, -1.73, 1.13, 1.04, 0.34$
⑦ : ⑤ + $\alpha_{7,7}q_{vpsnr}$	$6.22, -3.66, 1.59, -1.73, 1.13, 1.04, 0.19$

Model 5 Achieves the Highest PLCC and SROCC Scores

- 3-fold cross validation

Model	Training Set			Validation Set	
	PLCC	SROCC	p-value	PLCC	SROCC
①	0.7570	0.7767	$< .0001^*$	0.6769	0.6861
②	0.7998	0.7959	$< .0001^*$	0.7190	0.7158
③	0.8630	0.8353	$< .0001^*$	0.7476	0.7205
④	0.8877	0.8353	$< .0001^*$	0.7723	0.7205
⑤	0.9046	0.8570	$< .0001^*$	**0.7905**	**0.7430**
⑥	**0.9190**	0.8901	$< .0001^*$	0.7395	0.6414
⑦	0.9159	**0.8913**	$< .0001^*$	0.7455	0.6560

More Evaluations on New Videos and 10 Subjects



Model 5 is Better

- The evaluations on new videos and subjects confirm the robustness of our derived QoE models.

Model	Testing Set	
	PLCC	SROCC
④	0.6880	0.7319
⑤	0.7099	0.7664

Outline

- Introduction
- 360° Video Player
- User Study
- QoE Models
- Conclusions

Conclusions

- Realizing an open-source 360° video player supporting several projection schemes
- New findings of our user study
  - Projection scheme alone has no significant impact.
  - The QoE levels of complex or slow-paced 360° videos are more sensitive to QP values.
- The QoE model with projection schemes achieves up to 0.71 in PLCC and 0.77 in SROCC scores.

[3] S. Yao, “A 360 degree video player and head movement logger for an HMD using the OSVR API,” 2018, https://github.com/shunhuaiyao/360Degree_Head_Movement_Dataset/tree/win32

Limitations and Future Directions

- Human viewing behavior
  - Individual QoE model
- The degree of sickness
  - How dizzy to watch the 360° video?
- Integration with 360° video streaming systems
  - The impact of transmission bandwidth on QoE

Questions?

Comparisons with Other Studies

Study	Zhang et al. [3]	Tran et al. [18, 19]	Tran et al. [15]	Singla et al. [16, 17]	This paper
Method	Subjective, Objective	Subjective, Objective	Subjective	Subjective	Subjective, Objective
Encoder	H.264, H.265, VP9	H.264	H.264	H.265	H.264
Encoding Bitrates/QPs	0.3-10 Mbps	22-40 QP	22-40 QP	0.5-15 Mbps	22-38 QP
Encoding Resolutions	4K	720p-4K	720p-4K	1080p, 4K	4K
Projection Schemes	ERP	ERP	ERP	ERP	ERP, AEP, ECP
Video Genres	None		They reveal that different 360° videos affect QoE levels.		Spatial, Temporal
QoE Model	Not Developed				The First QoE Models