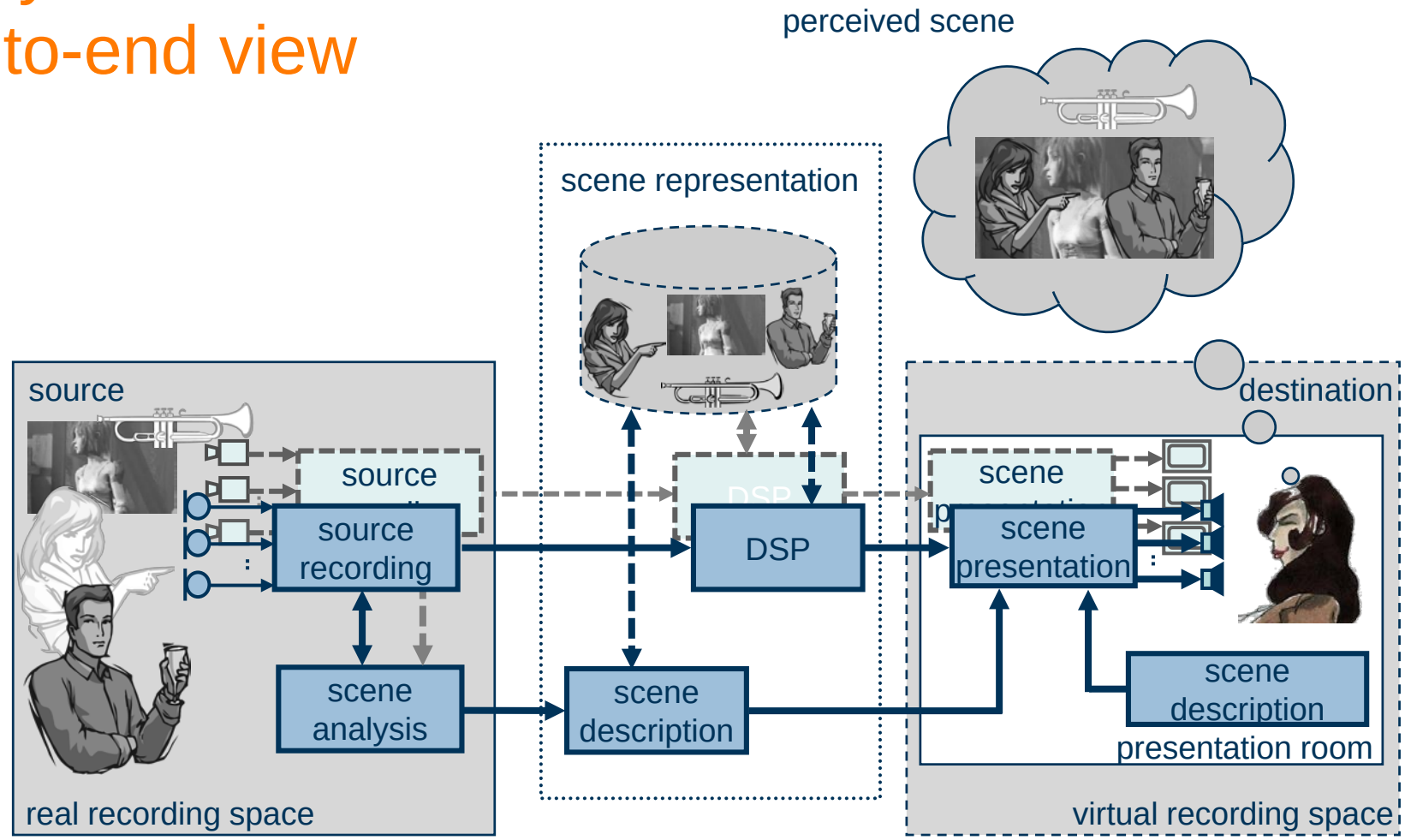
QoE test methods for 360° videos

Ashutosh Singla*, Stephan Fremerey*, Werner Robitza**, Alexander Raake*
*Audiovisual Technology Group, TU Ilmenau, Germany
**Telekom Innovation Laboratories, Deutsche Telekom AG, Germany

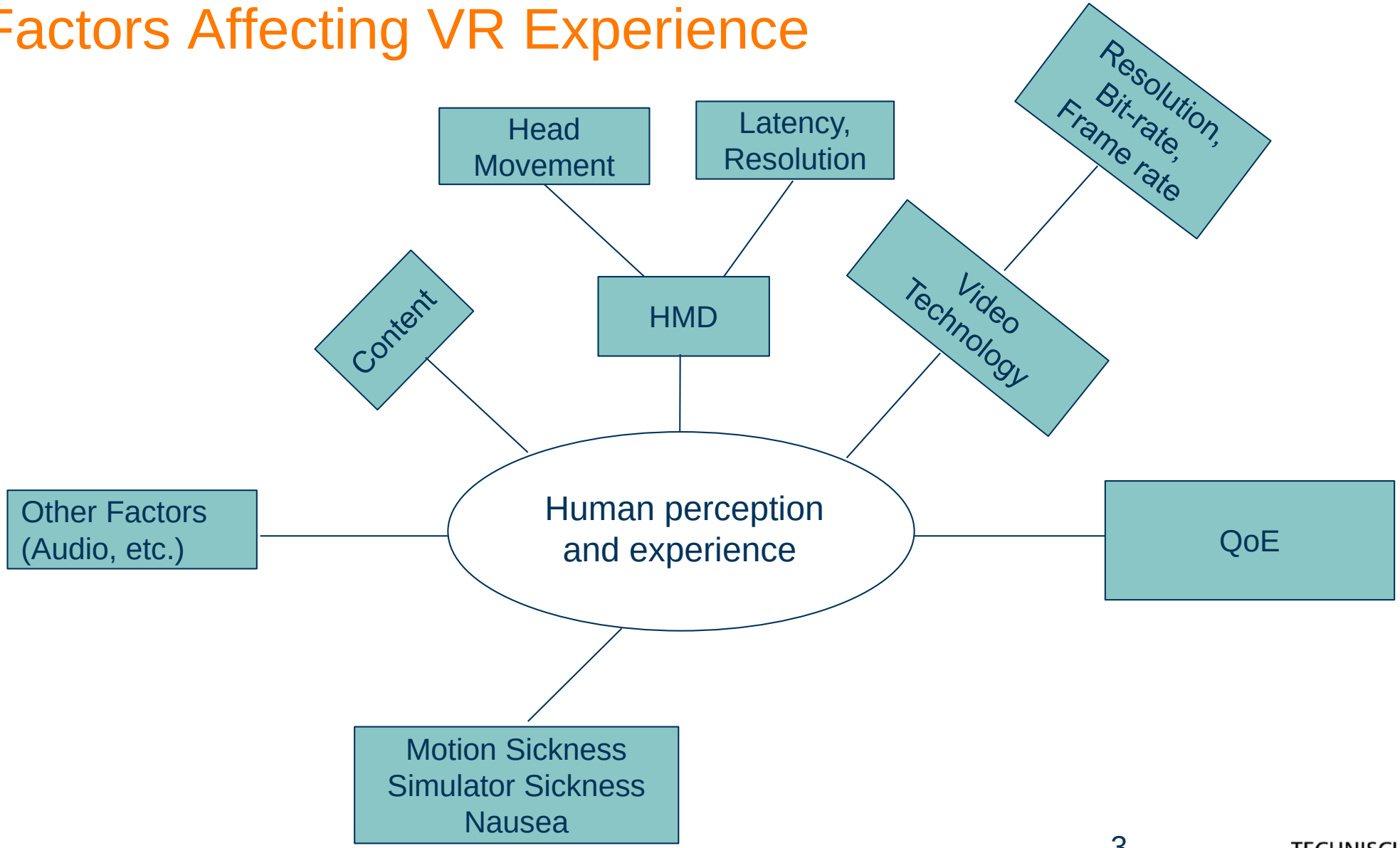
TU Ilmenau Campus

VR System
End-to-end view



(adapted from Raake & Spors 2006,
Spors et al., Proc. IEEE 2013)

Factors Affecting VR Experience



QoE-Relevant Aspects – Examples

- Video: Resolution, Field-of View, motion in different areas, frame rate, color space, etc.
- Audio: Format for recording, representation, playback, device properties, ...
- Audiovisual interaction
- Delivery
  - Bandwidth requirements!
  - Lag
  - specific approaches for delivering 360° video
- QoE and Methodology
  - Scenarios, content
  - Test methods
  - What is being assessed?
    - AV-quality, QoE, Simulator Sickness, Presence, User Behavior
- Recommendations for deployment

What We Have Done So Far?

QoMEX 2017

Singla, A., Fremerey, S., Robitza, W., & Raake, A.: Measuring and Comparing QoE and Simulator Sickness of Omnidirectional Videos in Different Head Mounted Displays.

Research Goals

- Resolution limitation of HMDs: QoE for 4K and FHD?
- Simulator Sickness
  - Content (60s length)
  - Resolution
  - Gender
- Behavioral Analysis
  - Exploration Behavior
- Developed framework for recording head movements
  - Compatible with HTC Vive and Oculus Rift
  - Yaw and pitch values measured every 0.14 s

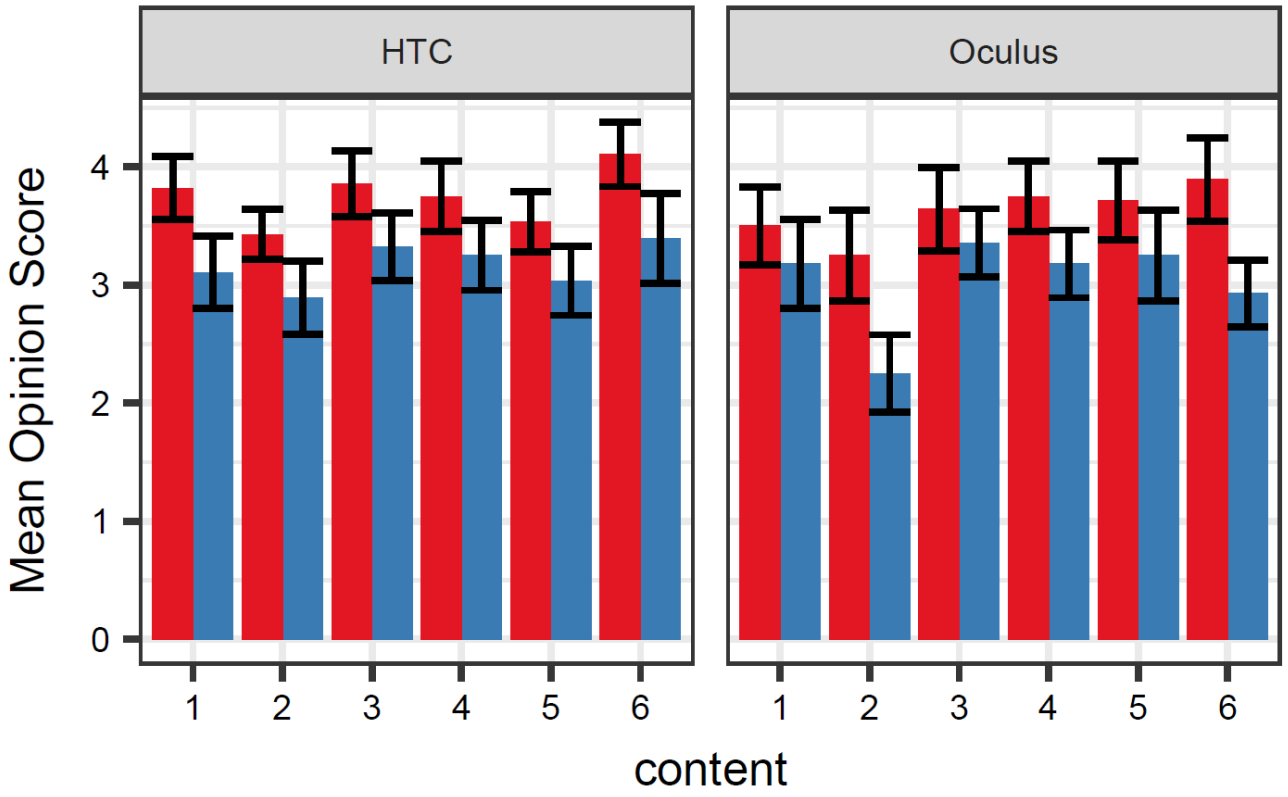
Test Method

- Assessed subjective quality of 360°videos
  - Absolute Category Rating (ACR) scale
- Each content viewed 4 times (FHD + 4K per HMD)
  - HMD wearing duration = 1m
- On paper MOS and Simulator Sickness Questionnaire (SSQ) scale
- Sit on rotating chair, free exploration of 360° video

28 Participants
- 15 Females
- 13 Males
- Avg. age = 26.25
- Median age = 25

Absolute Category Rating

Experimental results: Subjective Quality



Resolution

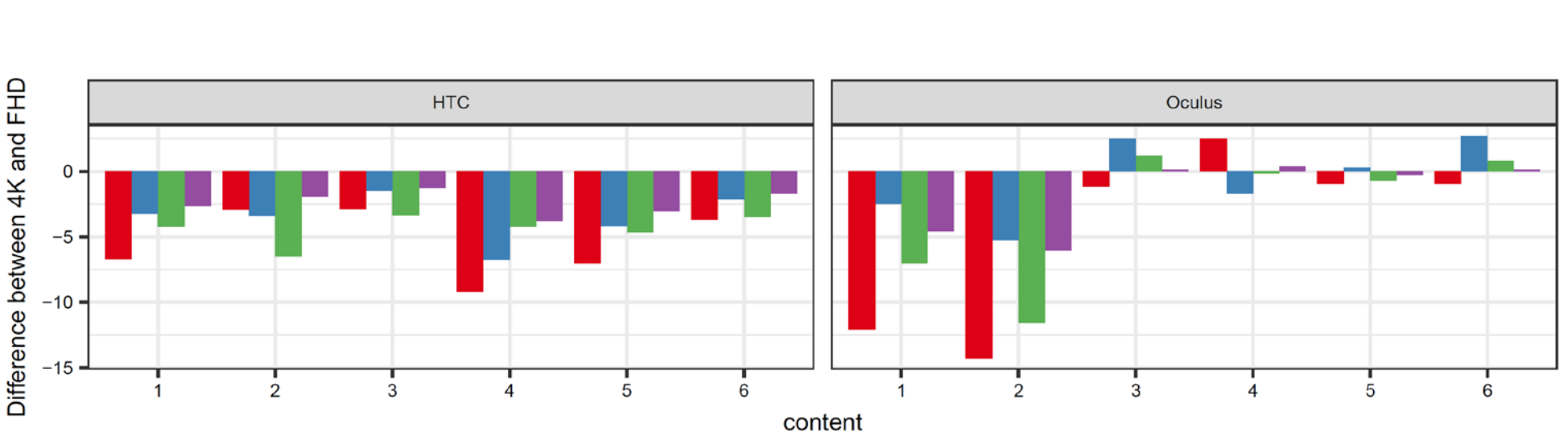
4K

FHD

Contents

1 – Roller Coaster
2 – Project 360
3 – Cockpit View
4 – Sky Diving
5 – Etihad A-380
6 – Elephants

Experimental results: Simulator Sickness



TECHNISCHE UNIVERSITÄT
ILMENAU

Conclusion

- Quality
  - 4K provides significantly better QoE than FHD
  - Significant impact of content on QoE
  - HTC Vive offers slightly better integral quality compared to Oculus Rift

- Simulator Sickness
  - Low motion videos → lowest scores
  - 4K has lower simulator sickness scores as compared to FHD
  - Females more prone to simulator sickness

- Behavioral Analysis
  - Pitch – No visible difference between HMDs
  - Yaw – Noticeable difference between HMDs

ACMMM 2017

Singla, A., Fremerey, S., Robitza, W., Lebreton, P., & Raake, A. (2017): Comparison of Subjective Quality Evaluation for HEVC Encoded Omnidirectional Videos at Different Bit-rates for UHD and FHD Resolution.

Research Goals

- Resolution limitation of HMD
  - Optimal Network Delivery
  - Provide insights into appropriate coding and resolution settings
- Simulator Sickness
  - In the context of the proposed test design
  - Gender
- Behavioral Analysis

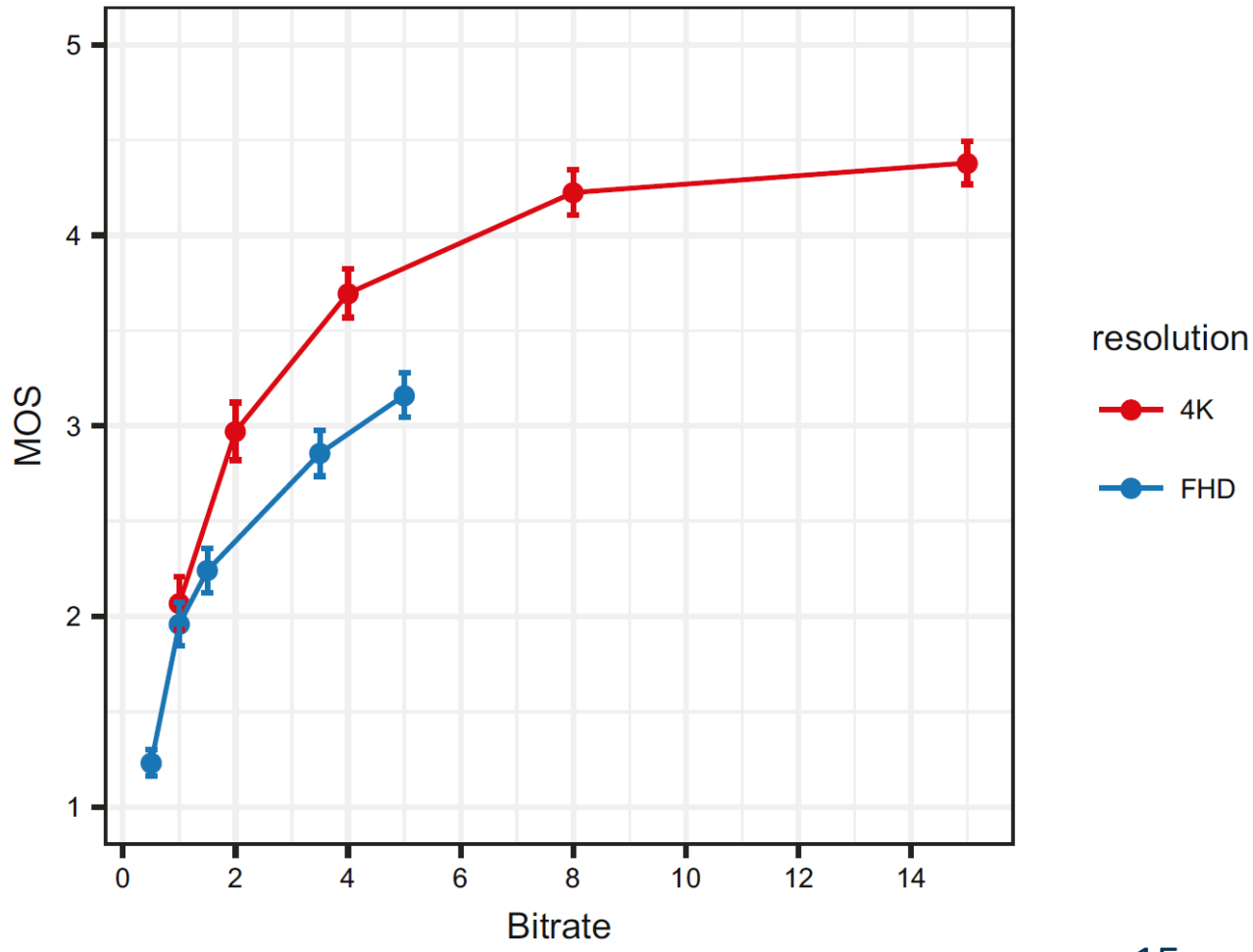
Test Method

- Assessed quality of 360°videos
- Modified Absolute Category Rating (M-ACR) scale
  - Scale is shown on HMD, rating is recorded verbally
    → enables wearing of HMD continuously
- Used six 4K SRCs (by JVET) (10s)
- Videos shown twice (not once as in P.910)
  - More reliable ratings for short sequences (10s)
  - HMD wearing duration = 15m

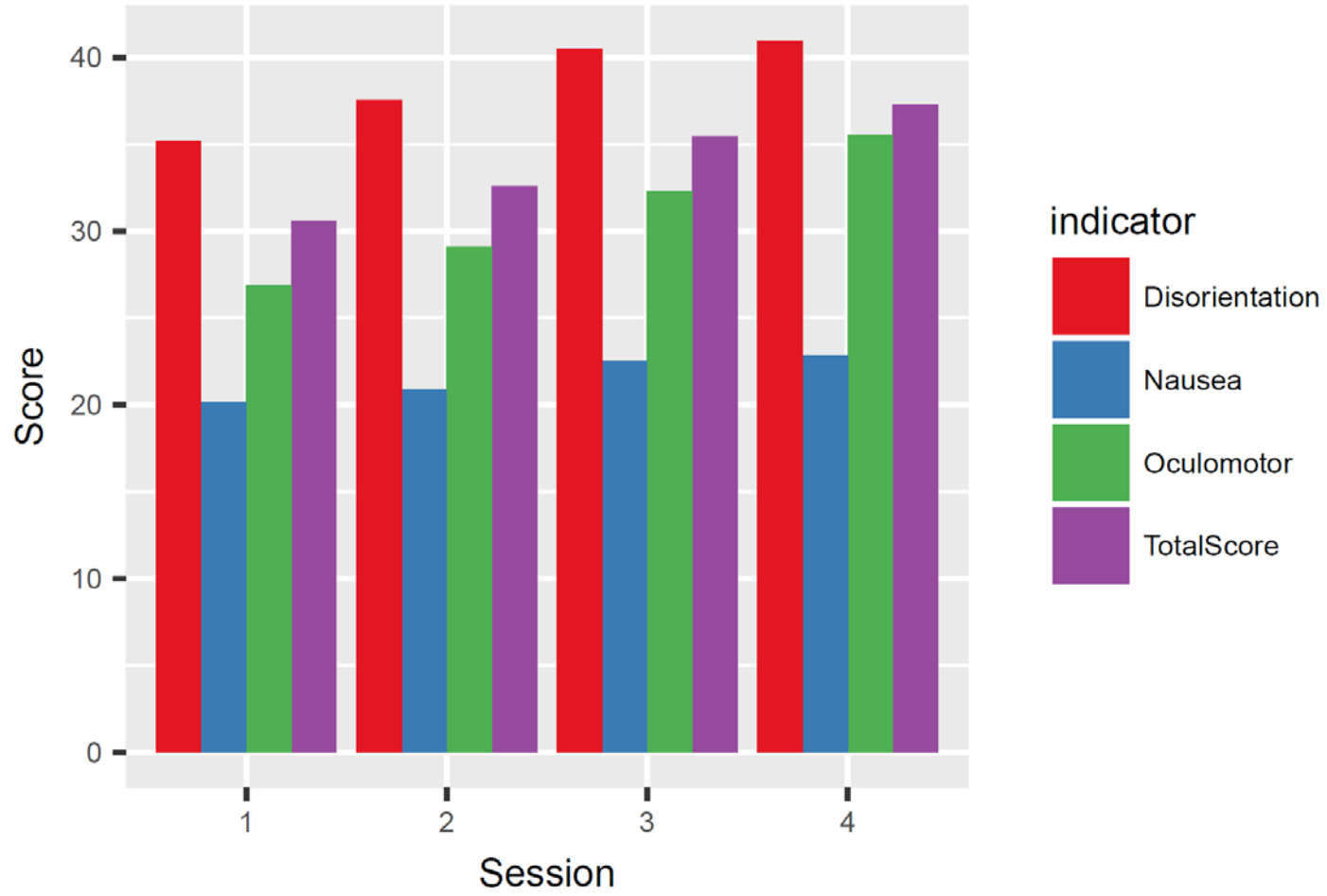
29 Participants
- 14 Females
- 15 Males
- Avg. age = 25.62
- Median age = 25

Modified Absolute Category Rating

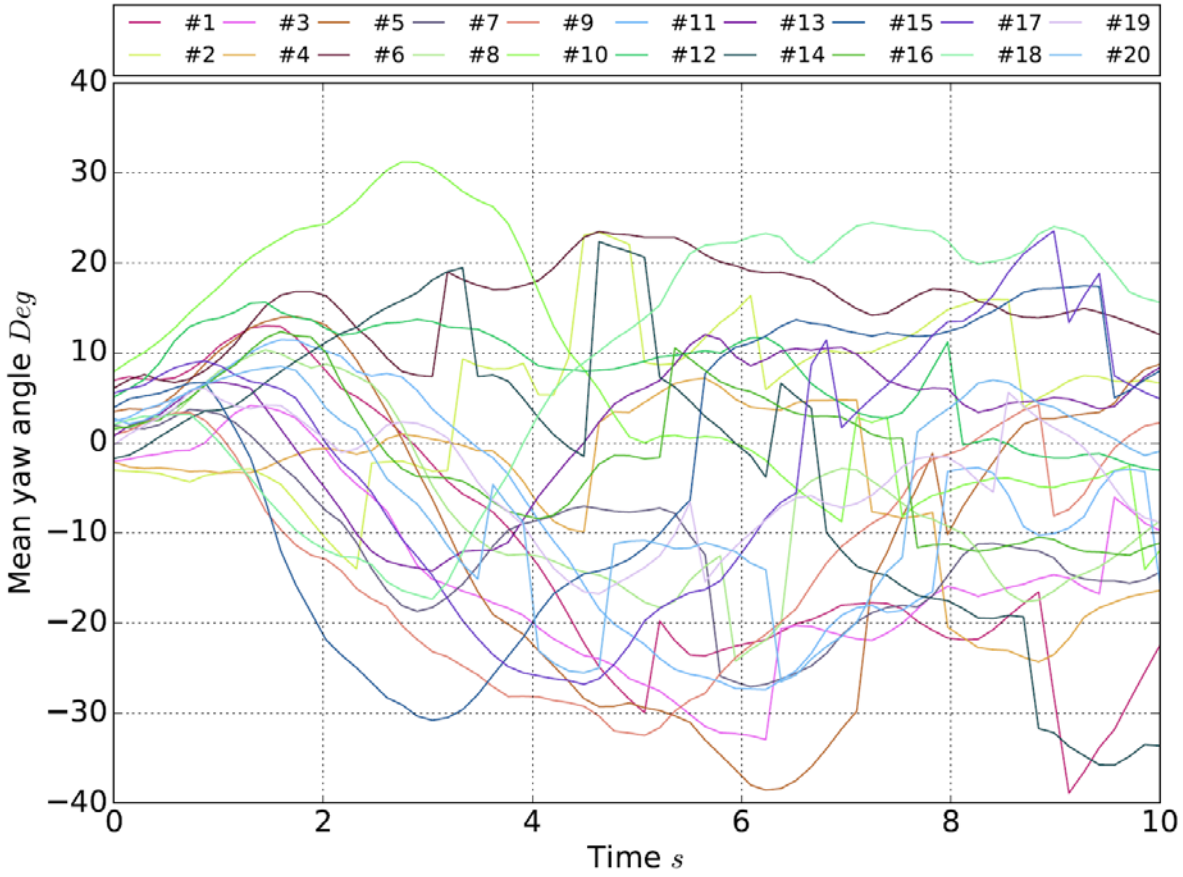
Experimental results: Subjective Quality



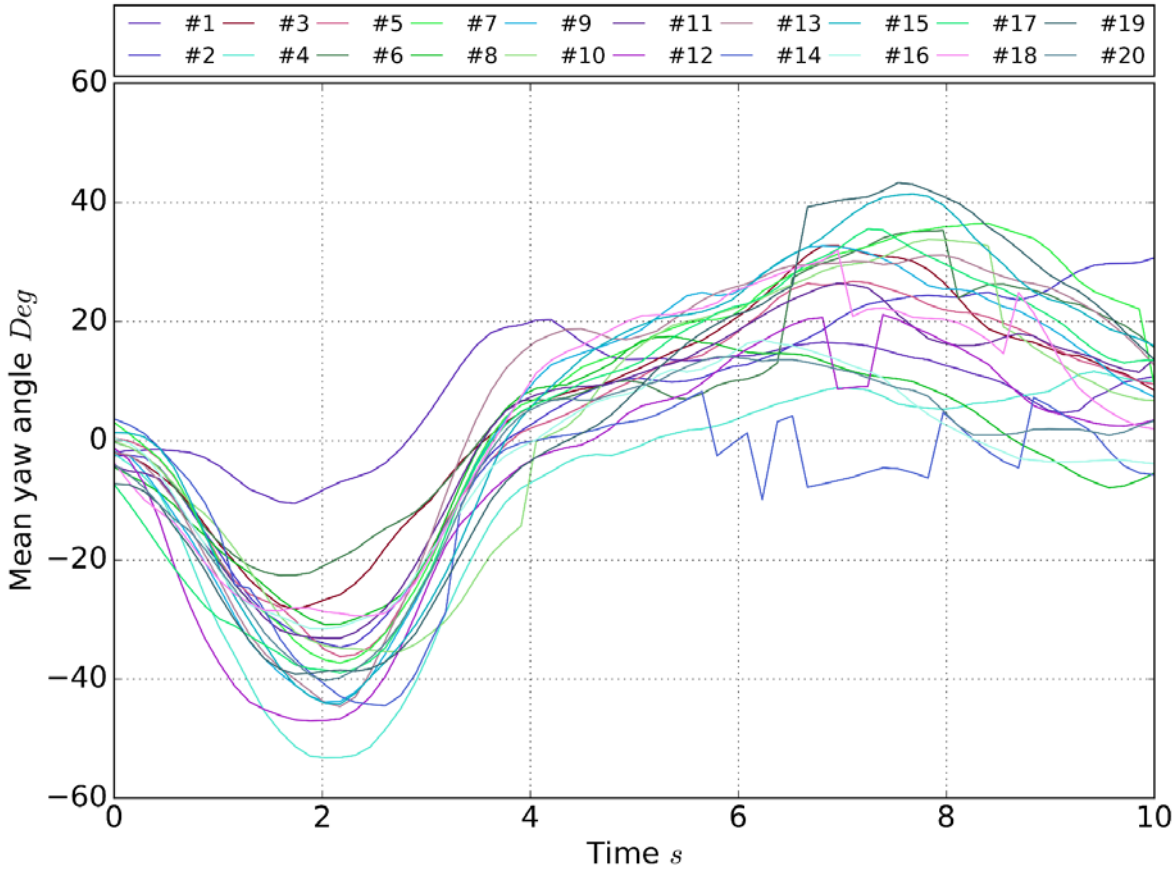
Experimental results: Simulator Sickness



Experimental results: Behavioral Analysis



Content 4



Content 2

Conclusion

- Quality
  - 15 Mbps provides almost same perceived quality as 8 Mbps 4K resolution
  - Perceived quality at FHD 1.5 Mbps is slightly higher as compared to UHD at 1 Mbps (expected due to bitrate – quality curves at different resolutions)

- Simulator Sickness
  - No significant increase in the sickness scores over the test sessions

- Behavioral Analysis
  - Contents 1 and 2 have higher value of pitch movements as compared to yaw
  - Subjects explored the video in a different way except in content 2

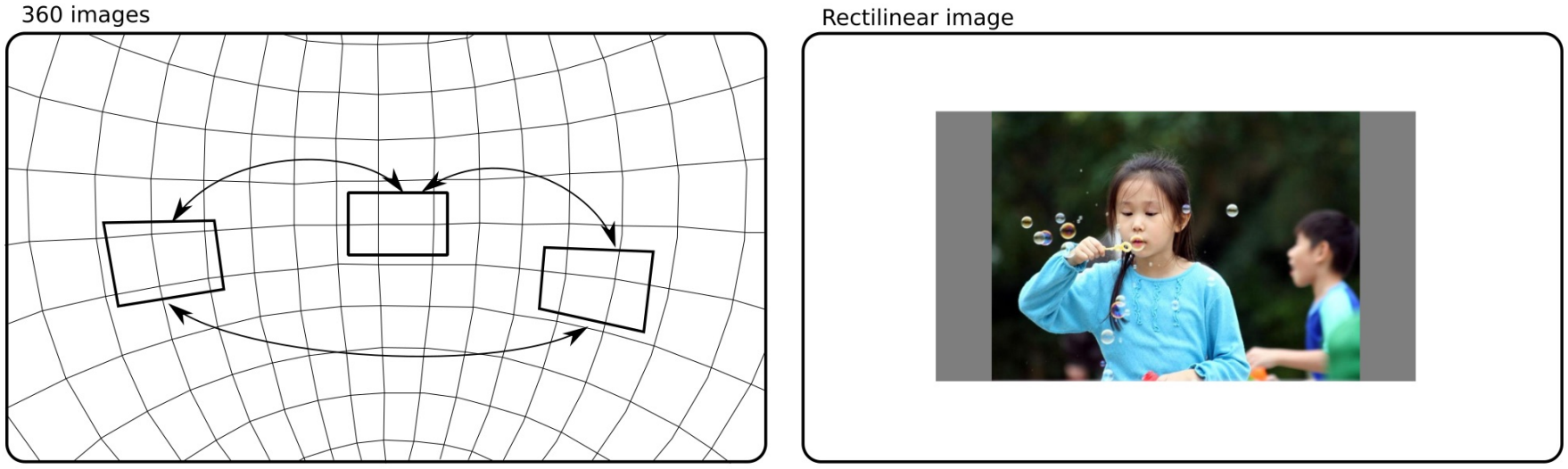
Salient360! IEEE ICME Grand Challenge 2017: GBVS360 – Rectilinear vs Equirectangular Images

Lebreton, P., Raake, A.

Task & Approach

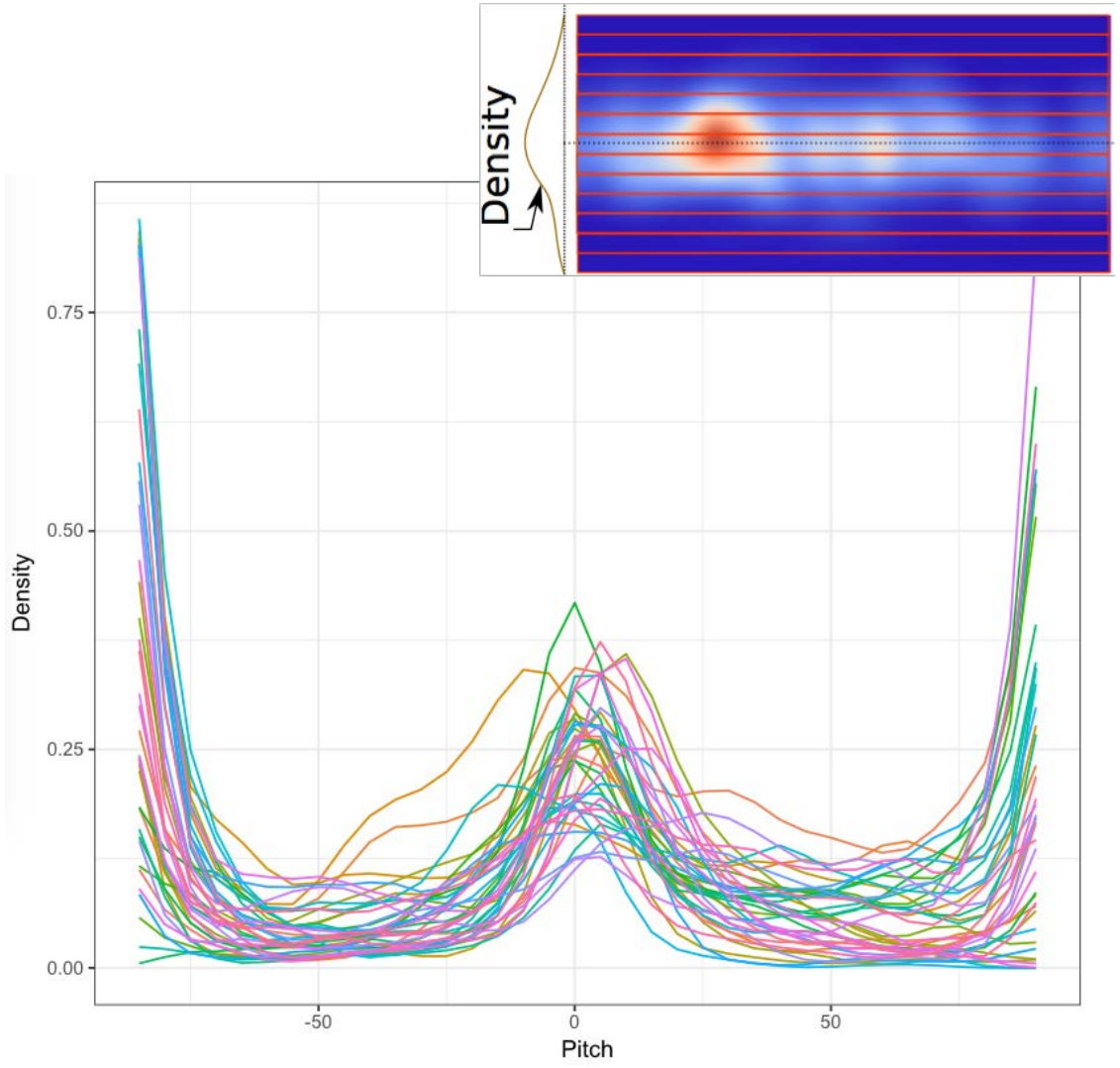
- Given tasks
  - Model type 1: Head motion based saliency model
    - Ground Truth Heat Map (GTHM) derived from head movement
    - output of model = map in equirectangular space
  - Model type 2: Head + eye-motion based saliency model
    - GTHM derived from head movement plus "movement of eye within viewport"
  - Model type 3: Scan-paths of observers in entire 360 panorama
    - Groundtruth scan-path (GTSP) obtained from head and eye-movement data
- Test dataset: 20 360° images viewed by 48 observers on Oculus-DK2 HMD
- Model: Adaptation of different models
  → "GBVS360" (basis: Graph-Based Visual Saliency), "BMS360" (basis: Boolean Map Saliency)
  - Specific features
  - Feature representation adaptation
  - Overall framework

Global vs. Local Saliency Maps Computation

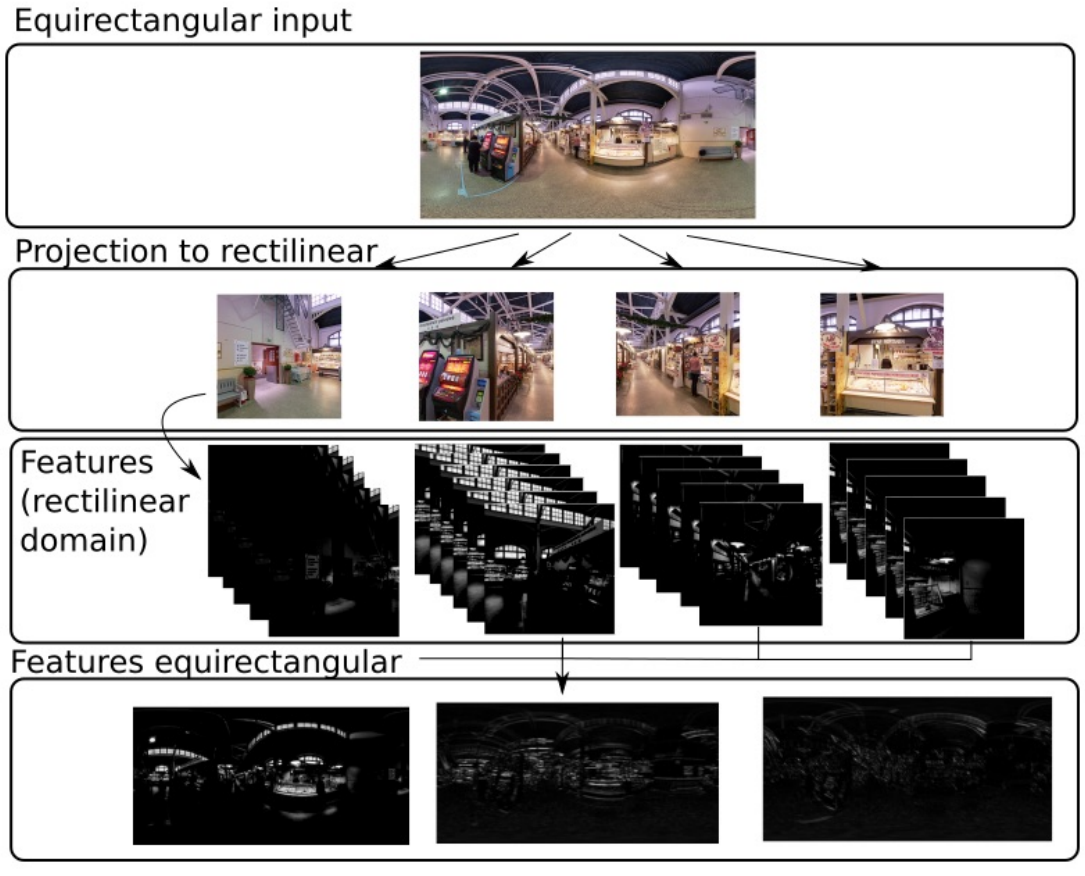


- Novelty with regard to context
- What structural changes are needed in existing models?

Adaptive equatorial Prior

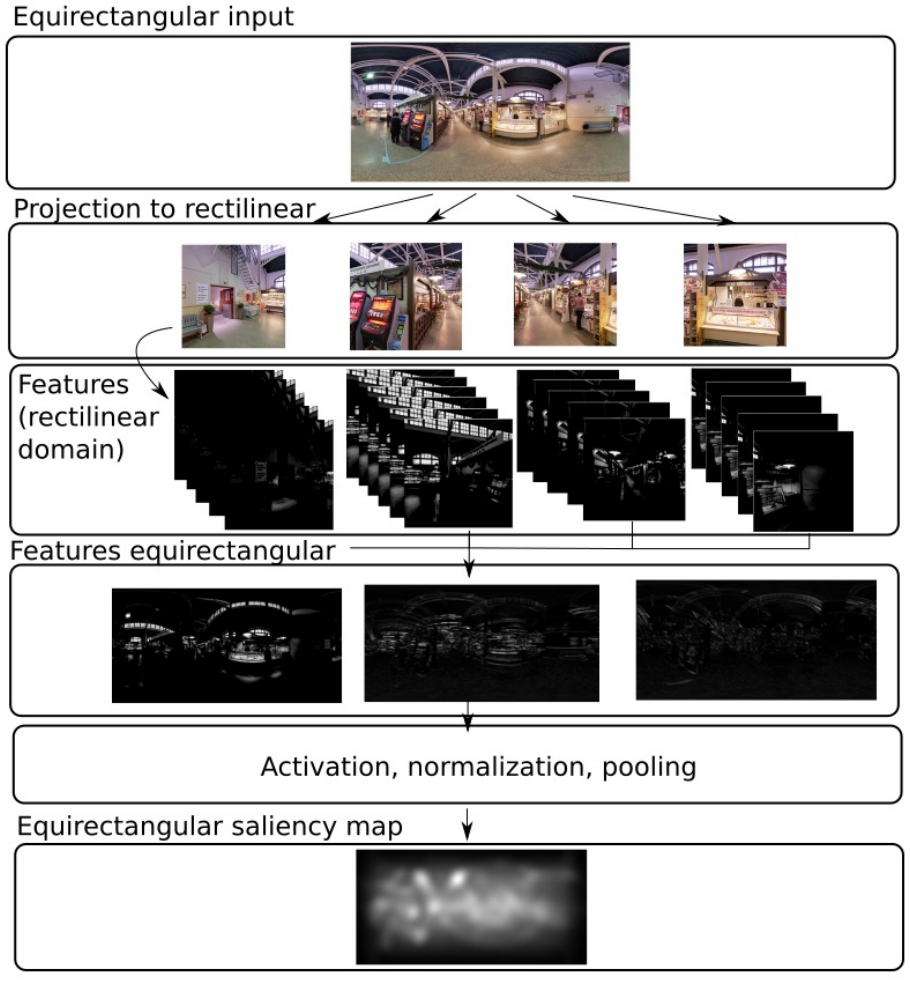


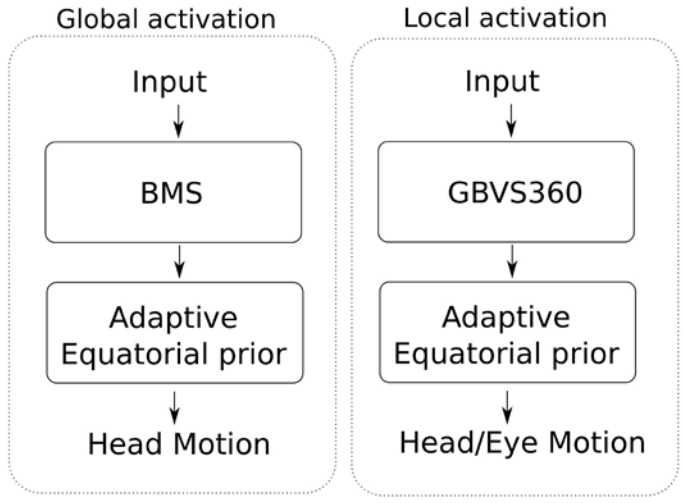
GBVS360 Framework – the right Features

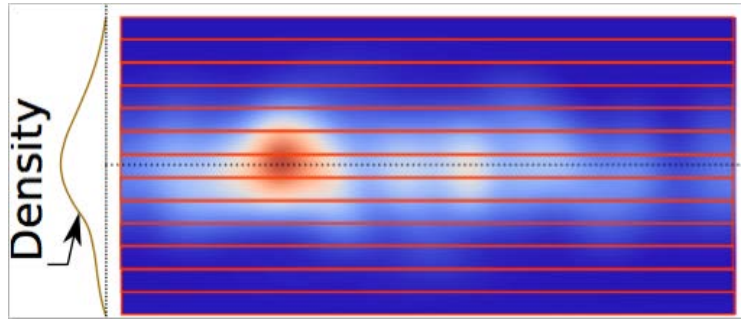


- Computation of features in the right domain
- Representation in a "global" format

GBVS360 Framework – the right Features







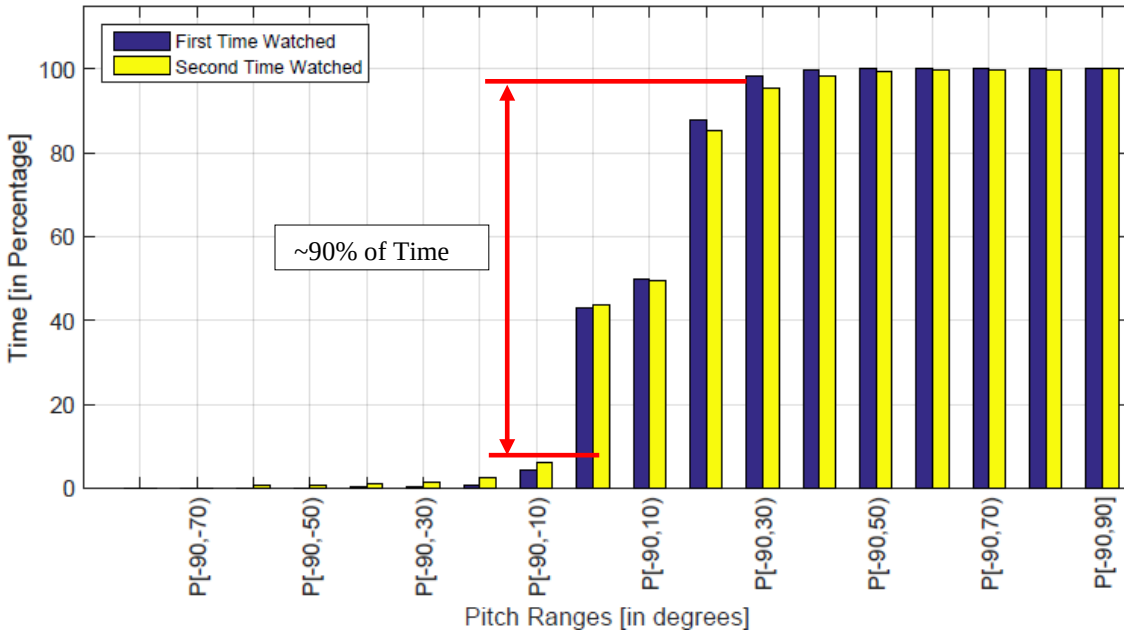
8th JVET Meeting
China 2017

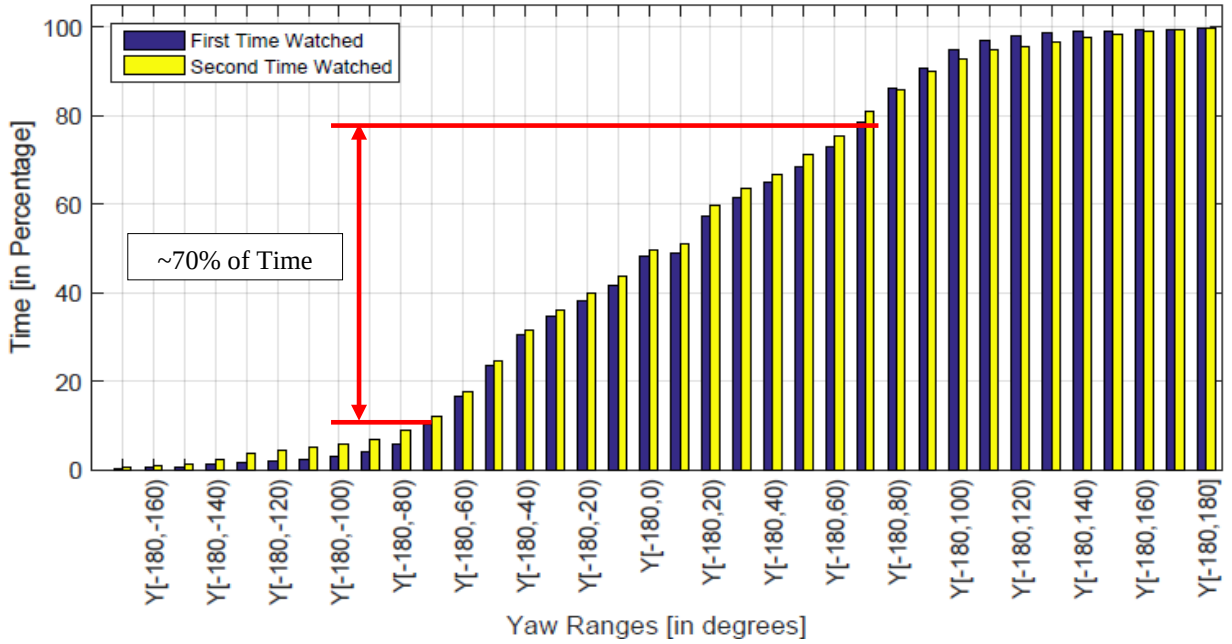
Singla, A., Fremerey, S., Raake, A., List, P. & Feiten, B. (2017).
AhG8: Measurement of User Exploration Behavior for Omnidirectional (360°) Videos with a Head Mounted Display

Research goals

- Exploration behavior of users watching 360° videos
- Similar behavior between subjects regarding yaw and pitch direction?
- Percentage of time spent on different viewports?
- Appropriate viewport to be shown in video tests on classical 2D screens?

Experimental Results: User Behavior





Cumulative Histogram of yaw and pitch interval

Experimental Results: User Behavior

Conclusion

- Subjects explored content quite equally

- Within all contents & quality levels: subjects don‘t explore entire pitch/yaw ranges
  - For pitch direction: > 80% of time spent between [-10, 30]
  - For yaw direction: > 60% of time spent between [-80, 60]

- Subjects almost don‘t move their head up-/downwards
  - Almost no exploration in „extreme ranges“
    - For pitch: [-90, -50], [50, 90]
    - For yaw: [-180, -140], [140, 180]

- 45% and 20% of time users are not moving their head in pitch and yaw direction

Questions/Suggestions