ACR360: A Dataset on Subjective 360° Video Quality Assessment Using ACR Methods

Majed Elwardy, Hans-Jürgen Zepernick,
Yan Hu, and Thi My Chinh Chu

Blekinge Institute of Technology, Sweden

International Conference on Signal Processing and Communication Systems
Bydgoszcz, Poland, 6 - 8 September 2023

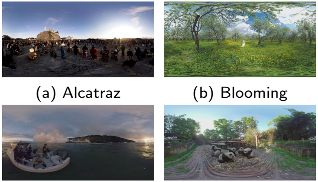
Motivation

- Immersive media has received significant attention across various applications in recent years.
- Subjective tests play a crucial role in the development of immersive media systems.
- **ACR360:** A dataset on subjective 360° video quality assessment using ACR methods.

Stimuli

- Four natural scenes of 10 s duration each with different resolutions.
- Frame rate of 29.97 frames per second (fps).
- 120 (360° videos) with different quality levels.

Note: Optimal resolution consider the projection and resolution limitations of HMD to garante a maximized per-pixel display (HTC VIVE: 3600×1800)



(a) Alcatraz (b) Blooming (c) Formation (d) Panda

Figure 1: Sample frames of the 360° video scenes [VQA, 2017],[Li et al., 2018].

Table 1: Summary of Stimuli

Ref. stimuli	Resolution: 8K, 6K, 4K, optimal [Zhang et al., 2018], 2K
Test stimuli	Resolution: 8K, 6K, 4K, optimal [Zhang et al., 2018], 2K QP: 22, 27, 32, 37, 42

Test methods and procedure

ACR method

- **ACR method:** Each stimulus is shown once followed by rating its quality [Recommendation ITU-T P.919, 2020].

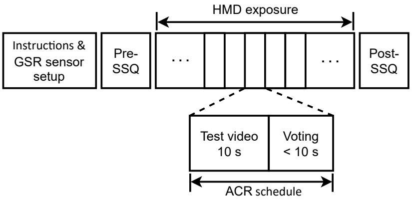


Figure 2: Procedure of the subjective tests using the ACR method ©[2019]IEEE. Reprint, with permission, from [Elwardy et al., 2019].

Test methods and procedure

MACR method

- **MACR method:** Each stimulus is shown twice with a 3-second mid-grey screen between [Singla et al., 2018].

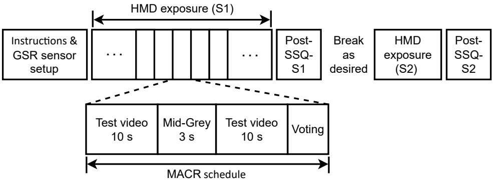


Figure 3: Procedure of the subjective tests using the MACR method which is split into Session 1 (S1) and Session 2 (S2).

Human-machine interface

- Head-mounted display (HMD):
  - HTC Vive Pro with integrated eye-tracker
- HTC Vive controller:
  - (1) Execute calibration instructions
  - (2) Cast quality ratings
- Shimmer GSR biosensor:
  - (1) Galvanic skin response
  - (2) Heart rate

Software suites and computing platform

- Test/development platform:
  - Unity 3D Version 2018.3.11f1
  - Visual Studio 2017
- iMotion Software Version 7.1:
  - (1) Bio recordings
  - (2) Simulator sickness questionnaire (SSQ)
- Corsair One i160 Gaming PC:
  - (1) Intel I9-9900K processor
  - (2) NVIDIA GeForce RTX 2080 TI graphics card

Participants

- Pilot tests with ACR and MACR:
  - 5 participants (2 females, 3 males).
  - 30-60 years, average age of 38.2 years
  - Experts familiar with immersive media
- Subjective tests with ACR:
  - 30 participants (7 females, 23 males)
  - 20-36 years, average age of 25.37 years
  - Often used: 0, Sometimes used: 17, and Never used: 13
- Subjective tests with MACR:
  - 30 participants (9 females, 21 males)
  - 23-46 years, average age of 29.53 years
  - Often used: 0, Sometimes used: 13, and Never used: 17

Dataset Structure

Directory structure

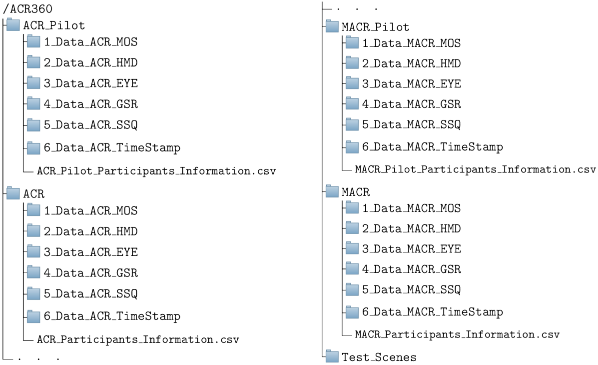


Figure 4: ACR360 dataset directory structure [Elwardy, 2023]

https://github.com/MajedElwardy/ACR360.

- GitHub link:

Opinion Scores

- Opinion scores: Measures the participants subjective perception of the stimuli.
- Mean opinion scores (MOS): Opinion scores averaged over the number of participants.
- Average MOS: Mean opinion scores averaged over the four scenes.

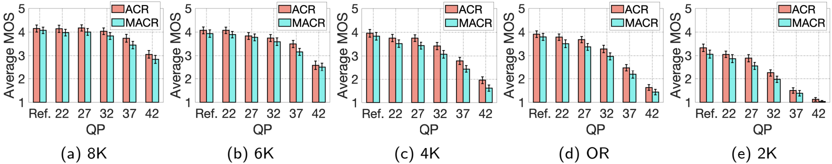


Figure 5: Average MOS over the four scenes versus quantization parameter for each resolution and 95% confidence interval.

Rating Times

- Rating times provide insights into the difficulty of given quality score to stimuli.

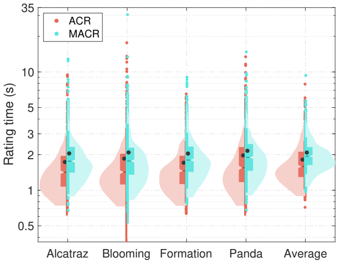


Figure 6: Violin plots of the rating times for each video scene and average rating times over the four video scenes for the ACR and MACR methods.

Head Movements

Cumulative Distribution Function (CDF)

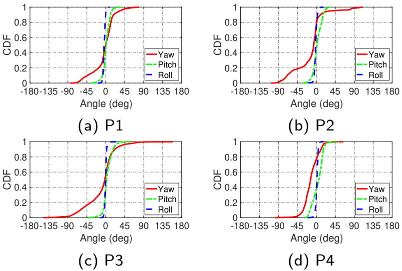


Figure 7: CDFs of yaw, pitch, and roll angles for ACR.

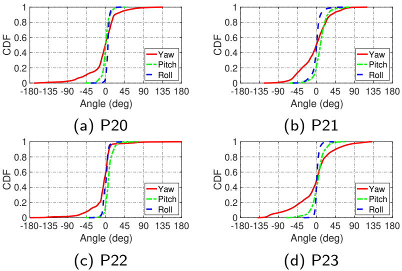
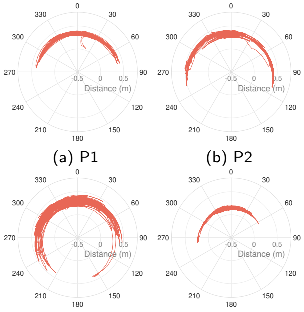


Figure 8: CDFs of yaw, pitch, and roll angles for MACR.

- **Note:** Exploration behavior may vary significantly among participants and sessions.

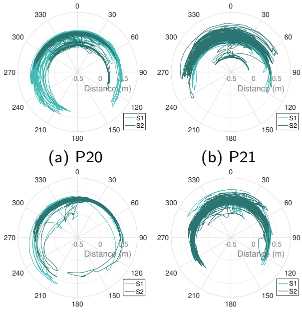
Head Movements

Head Trajectories



(a) P1 (b) P2
(c) P3 (d) P4

Figure 9: Samples of participants' head trajectories for ACR.



(a) P20 (b) P21
(c) P22 (d) P23

Figure 10: Samples of participants' head trajectories for MACR.

- **Note:** Exploration behavior may vary significantly among participants and sessions.

Pupil Dilation

- Provide objective measure of participants arousal or cognitive load in response to stimuli.

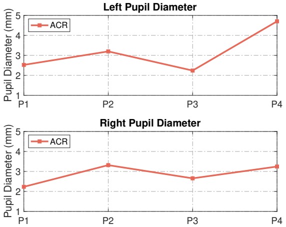


Figure 11: Eye pupil diameters for participants P1-P4 for ACR.

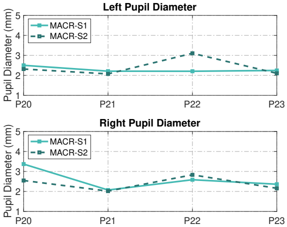


Figure 12: Eye pupil diameters for participants P20-P23 for MACR.

Galvanic Skin Response (GSR)

- GSR measures changes in the electrical conductance of the skin which reflects emotional arousal or stress level.

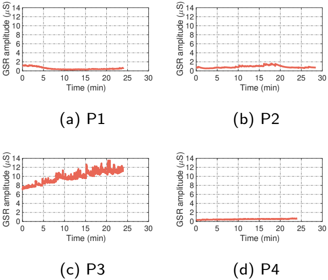


Figure 13: Samples of GSR amplitudes for the ACR method during HMD exposure.

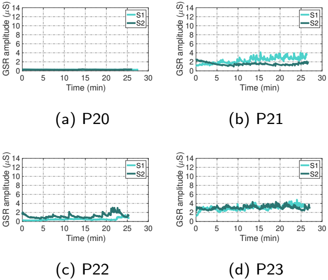


Figure 14: Samples of GSR amplitudes for the MACR method during HMD exposure.

Simulator Sickness Questionnaire (SSQ)

- Self reported symptoms of simulator sickness experienced by participant.

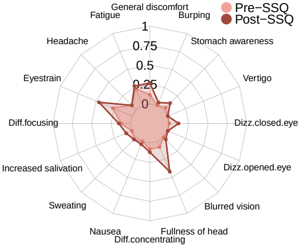


Figure 15: Mean Pre-SSQ scores and mean Post-SSQ scores for each symptom when using the ACR.

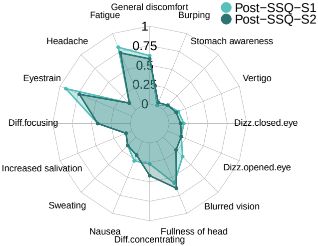


Figure 16: Mean Post-SSQ scores for Session 1 and Session 2 for each symptom when using the MACR.

Conclusions

- ACR360 dataset: provides the psychophysical and psychophysiological data has been made publicly available on GitHub.
- Using the ACR360 dataset as a ground truth for designing:
  - 360° video systems.
  - Benchmarking algorithms of 360° video processing chains.
  - Conducting meta-analysis.
- ACR360 dataset opens up opportunities for further research in understanding human behavior in immersive media
  - Developing new objective models and metrics.
  - Expanding subjective tests to other types of immersive media.

Acknowledgments

- This work was supported in part by the Knowledge Foundation, Sweden, through the ViaTecH and HINTS projects under contract numbers 20170056 and 20220068, respectively.
- We thank all volunteers who generously shared their time to participate in this study.

References

(2017). *VQA-ODV. 2017*. Beihang University, School of Electronic and Information Engineering, Beijing, China. Available online: https://github.com/Archer-Tatsu/VQA-ODV (accessed on 13. March 2023).

Elwardy, M. (2023). *ACR360 Dataset. 2023*. Blekinge Institute of Technology, Karlskrona, Sweden. Available online: https://github.com/MajedElwardy/ACR360 (accessed on 23. May 2022).

Elwardy, M., Zepernick, H.-J., Sundstedt, V., and Hu, Y. (2019). Impact of Participants' Experiences with Immersive Multimedia on 360° Video Quality Assessment. In *Proc. Int. Conf. on Signal Process. and Commun. Systems*, pages 40–49, Gold Coast, Australia.

Li, C., Xu, M., and Wang, Z. (2018). Bridge the Gap Between VQA and Human Behavior on Omnidirectional Video: A Large-Scale Dataset and a Deep Learning Model. In *Proc. ACM Int. Conf. on Multimedia*, pages 932–940, Seoul, Republic of Korea.

Recommendation ITU-T P.919 (2020). *Subjective Test Methodologies for 360° Video on Head-Mounted Displays*. International Telecommunication Union, Geneva, Switzerland.

Singla, A., Robitza, W., and Raake, A. (2018). Comparison of Subjective Quality Evaluation Methods for Omnidirectional Videos with DSIS and Modified ACR. In *Proc. Int. Symp. of Human Vision and Electronic Imaging*, pages 1–6, Burlingame, CA, USA.

Zhang, Y., Wang, Y., Liu, F., Liu, Z., Li, Y., Yang, D., and Chen, Z. (2018). Subjective Panoramic Video Quality Assessment Database for Coding Applications. *IEEE Trans. on Broadcasting*, 64(2):461–473.

Thank you!

Questions?