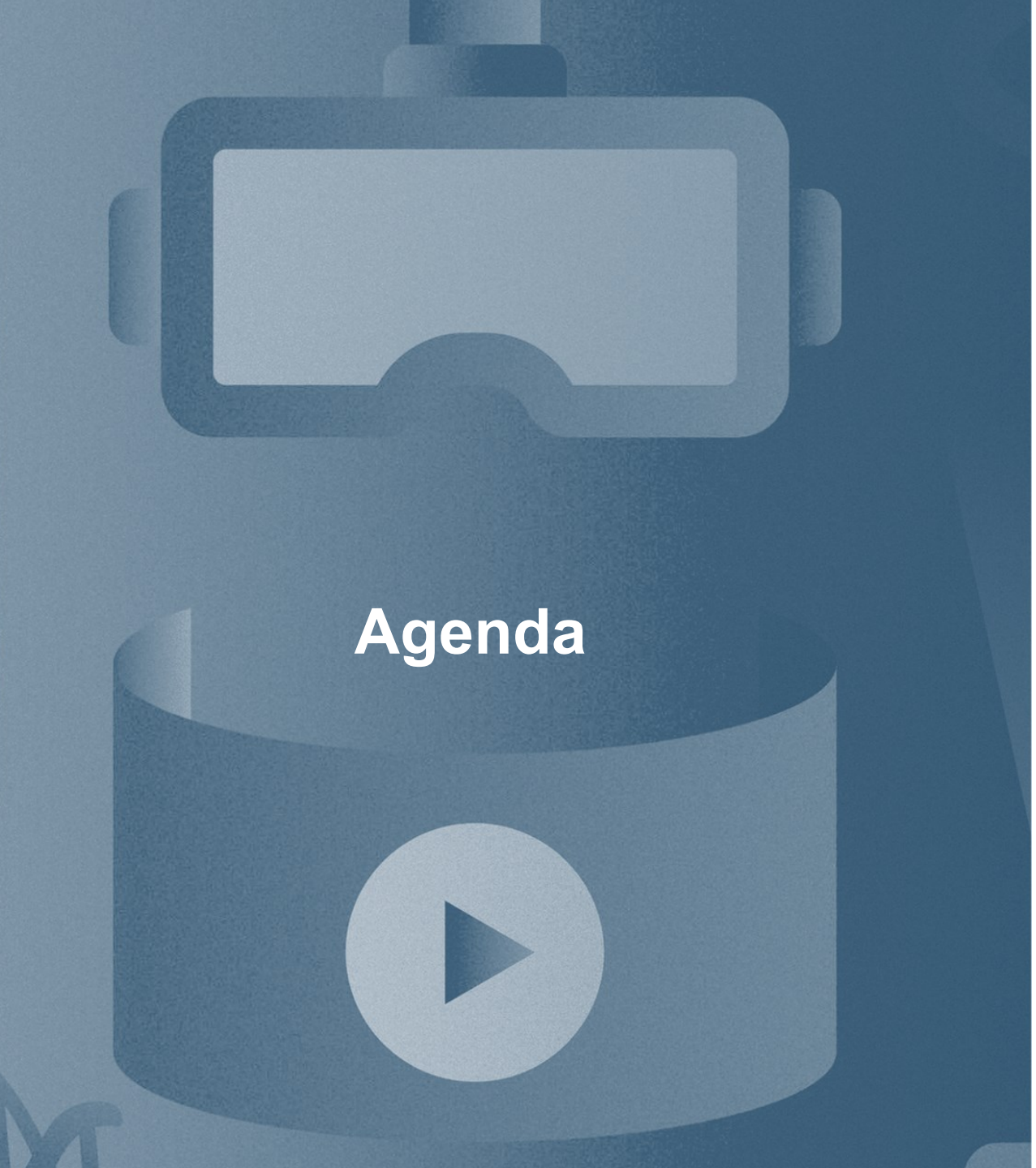
No-Reference Quality of Experience Model for Dynamic Point Clouds in Augmented Reality

Minh Nguyen, **Shivi Vats**, Hermann Hellwagner
University of Klagenfurt, Austria
shivi.vats@aau.at

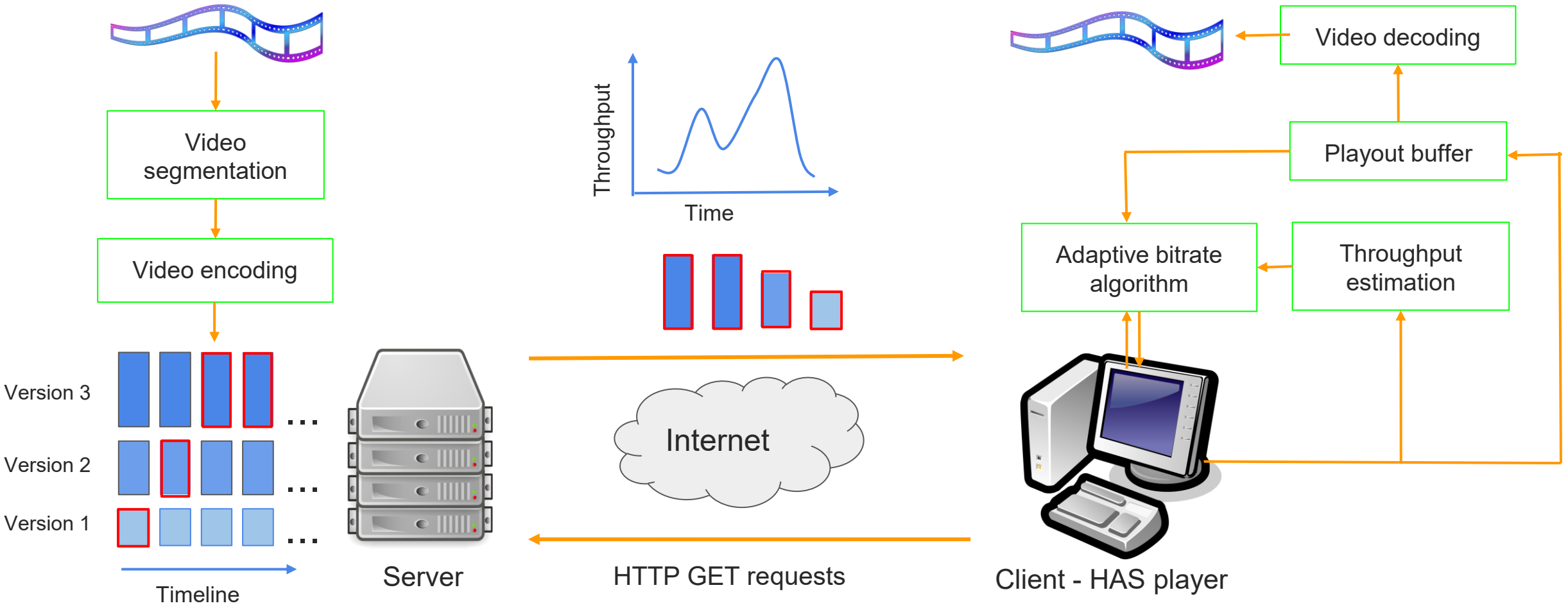
Extended Abstract Paper
ACM Mile High Video 2024
11-14 February 2024



Agenda

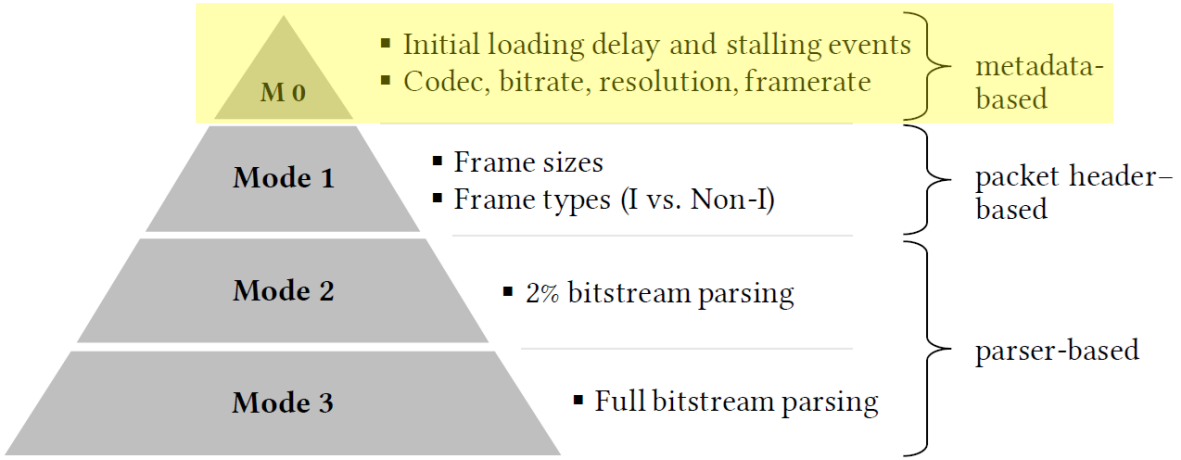
- HTTP Adaptive Streaming
- ITU-T P.1203
- Point Cloud Streaming
- Subjective QoE Tests
- No-Reference QoE Model for PC Streaming

HTTP Adaptive Streaming (HAS)



ITU-T P.1203

- Quality of Experience (QoE) prediction model for 2D videos[1]
- Trained using a subjective database for 2D videos
- Up to FHD resolution
- Includes switching between different representations



[1] Werner Robitza et al. 2018. HTTP Adaptive Streaming QoE Estimation with ITU-T Rec. P.1203 – Open Databases and Software. In *9th ACM Multimedia Systems Conference.* Amsterdam. https://doi.org/10.1145/3204949.3208124

Point Cloud (PC) Streaming

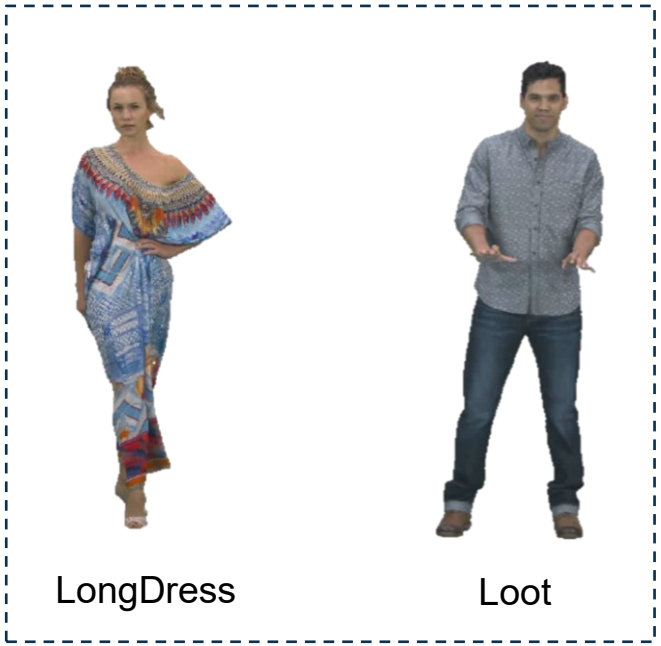
- Similar characteristics: bitrates, frame size, frame rate, stall events, viewing distance
- Differences:
  - Viewing environment
  - PCs can provide 6DoF
  - Bandwidth requirements

Video	Quality		
	Q1	Q2	Q3
Loot	2.28	5.63	16.68
LongDress	4.64	14.05	46.78
RedAndBlack	3.39	7.55	22.90
Soldier	4.38	11.58	35.29

Bandwidth in Mbit/s

Subjective QoE Tests

- Subjective tests for QoE of PCs[2]
  - MR environment with 4 life-sized digital humans from 8i lab database[3]
  - 32 participants asked to rate dynamic PC sequences



[2] Minh Nguyen et al. 2023. Impact of Quality and Distance on the Perception of Point Clouds in Mixed Reality. In *2023 15th International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 87–90.

[3] Eugene d'Eon et al. 2017. 8i Voxelized Full Bodies, version 2 – A Voxelized Point Cloud Dataset. In *ISO/IEC JTC1/SC29 Joint WG11/WG1 (MPEG/JPEG) Input Document M40059/M74006*.

Subjective QoE Tests

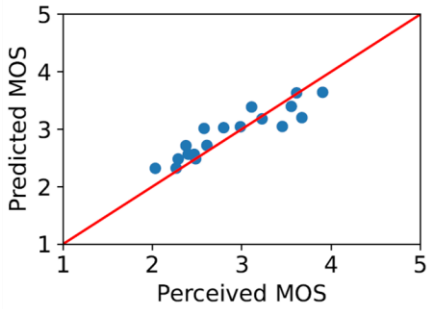
- 3 quality levels with static quality and quality switching – LongDress and Loot
- 3 viewing distances with static quality – RedAndBlack and Soldier
- Participants on a fixed location
- Conclusions:
  - Switching to a lower QoE decreases quality, and vice-versa
  - Closer viewing distance decreases quality, and vice-versa
  - Participants are less sensitive to quality distortions in less contrasted content
- Test data made public containing: PC qualities, distance, and rated QoE (1-10)

minhkstn/QoE-and-Immersion-of-Dynamic-Point-Cloud

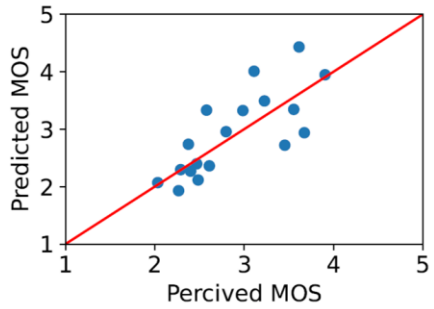
No-Reference QoE Model for PC Streaming

- Re-training ITU-T P.1203 mode 0 coefficients for dynamic PC videos using our dataset
  - Static qualities, quality switching,
  - *LongDress* and *Loot* for training
  - *Soldier* and *RedAndBlack* for validation
- Aimed for lowest RMSE between predicted and perceived MOS

	Training			Validation		
	PLCC ↑	SRCC ↑	RMSE ↓	PLCC ↑	SRCC ↑	RMSE ↓
ITU-T P.1203	0.766	0.785	0.887	0.918	0.829	1.032
Fine-tuned P.1203	**0.919**	**0.953**	**0.813**	**0.958**	**0.828**	**0.955**



Fine-tuned P.1203 model



Original P.1203 model

Conclusion

- Trained a no-reference model for QoE estimation of point cloud videos
  - Fine-tuned ITU.T P.1203 weights
- Better results than original P.1203 model in PLCC, SRCC, and RMSE
- The model is available at *github.com/minhkstn/itu-p1203-point-clouds*

minhkstn/itu-p1203-point-clouds

SPIRIT

THANKS

Contact:
Shivi Vats | shivi.vats@aau.at

spirit-project.eu

@SPIRIT_eu

@spirit-eu-project

SPIRIT project is funded by the EU's Horizon Europe programme under Grant Agreement number 101070672