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

Minh Nguyen
minh.nguyen@aau.at
Alpen-Adria-Universität (AAU) Klagenfurt
Klagenfurt, Austria

Shivi Vats
shivi.vats@aau.at
Alpen-Adria-Universität (AAU) Klagenfurt
Klagenfurt, Austria

Hermann Hellwagner
hermann.hellwagner@aau.at
Alpen-Adria-Universität (AAU) Klagenfurt
Klagenfurt, Austria

ABSTRACT

Point cloud streaming is becoming increasingly popular due to its ability to provide six degrees of freedom (6DOF) for immersive media. Measuring the quality of experience (QoE) is essential to evaluate the performance of point cloud applications. However, most existing QoE models for point cloud streaming are complicated and/or not open source. Therefore, it is desirable to provide an open-source QoE model for point cloud streaming.

The International Telecommunication Union (ITU) put in a great deal of effort in video quality estimation models, namely ITU-T P.1203 [2]. This P.1203 model was implemented and published on Github[1] [4]. The model's inputs include video characteristics (*i.e.*, bitrate, framerate, codec, and frame size), streaming parameters (*i.e.*, stall events), and viewing conditions (*i.e.*, device type and viewing distance). Point cloud streaming also shares some parameters that can be used in the P.1203 model, such as bitrate, framerate, stall events, and viewing distance. However, as the coefficients in the original P.1203 model were determined from a training phase based on a subjective database for 2D videos [4], they need to be re-trained with a new subjective database for point cloud streaming.

In this work, we provide a fine-tuned ITU-T P.1203 model for dynamic point clouds in Augmented Reality (AR) environments. We re-train the P.1203 model with our dataset published in [3] to get the optimal coefficients in this model that achieves the lowest root mean square error (RMSE). The dataset was collected in a subjective test in which the participants watched dynamic point clouds from the 8i lab database [1] with Microsoft's HoloLens 2 AR glasses. The dynamic point clouds have static qualities or a quality switch in the middle of the sequence. We split this dataset into a training set and a validation set. We train the coefficients of the P.1203 model with the former set and validate its performance with the latter one.

The results show that our fine-tuned P.1203 model outperforms the original model from the ITU. Our model achieves an RMSE of 0.813, compared to 0.887 of the original P.1203 model with the training set. The Pearson Linear Correlation Coefficient (PLCC) and Spearman's Rank Correlation Coefficient (SRCC) of our fine-tuned model are also significantly higher than that of ITU's model (see Table 1). These values are more than 0.9 in our model, compared to less than 0.786 in the standard P.1203 model for the training dataset. Taken into account the validation dataset, it can be seen that our fine-tuned model provides a better RMSE = 0.955, compared with 1.032 of the standard P.1203 model. We also achieved a better correlation with the ground truth with PLCC = 0.958 while this metric of the standard P.1203 model is 0.918. The correlations of the compared models are visualized in Fig. 1. The fine-tuned P.1203 model is published in https://github.com/minhkstn/itu-p1203-point-clouds.

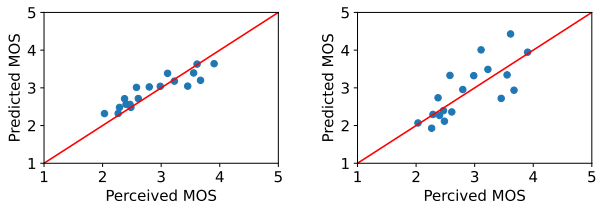
[1]https://github.com/itu-p1203/itu-p1203. Accessed 21 September 2023.

MHV '24, February 11–14, 2024, Denver, CO, USA

ACM ISBN 979-8-4007-0493-2/24/02…$15.00
https://doi.org/10.1145/3638036.3640248

Table 1: Performance of the original P.1203 mode 0 and our fine-tuned P.1203 with training and validation dataset.

	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**



(a) Fine-tuned P.1203 model **(b) Original P.1203 model**

Figure 1: Perceived MOS (from the subjective test in [3]) versus predicted MOS using the fine-tuned P.1203 model and the original P.1203 model.

CCS CONCEPTS

• **Information systems → Multimedia streaming**.

KEYWORDS

QoE model, Point Cloud, Augmented Reality, ITU-T P.1203

ACM Reference Format:
Minh Nguyen, Shivi Vats, and Hermann Hellwagner. 2024. No-Reference Quality of Experience Model for Dynamic Point Clouds in Augmented Reality. In *Mile-High Video Conference (MHV '24), February 11–14, 2024, Denver, CO, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3638036.3640248

ACKNOWLEDGEMENTS

This work has been funded by the European Union (SPIRIT project, Grant Agreement 101070672, https://www.spirit-project.eu/).

REFERENCES

[1] Eugene d'Eon, Bob Harrison, Taos Myers, and Philip A. Chou. 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.*

[2] ITU-T. [n. d.]. Rec. P.1203. Parametric bitstream-based quality assessment of progressive download and adaptive audiovisual streaming services over reliable transport - video quality estimation module. http://handle.itu.int/11.1002/ps/P1203-01 Accessed: 08 July 2021.

[3] Minh Nguyen, Shivi Vats, Sam Van Damme, Jeroen Van Der Hooft, Maria Torres Vega, Tim Wauters, Christian Timmerer, and Hermann Hellwagner. 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.

[4] Werner Robitza, Steve Göring, Alexander Raake, David Lindegren, Gunnar Heikkilä, Jörgen Gustafsson, Peter List, Bernhard Feiten, Ulf Wüstenhagen, Marie-Neige Garcia, Kazuhisa Yamagishi, and Simon Broom. 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