Contextual Bandit Learning–Based Viewport Prediction for 360 Video

4 authors, including:

Joris Heyse
Ghent University
7 PUBLICATIONS   **8** CITATIONS

Maria Torres Vega
Ghent University
31 PUBLICATIONS   **191** CITATIONS

Femke De Backere
Ghent University
43 PUBLICATIONS   **134** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project   Fallrisk View project

Project   OCareClouds View project

Contextual Bandit Learning-Based Viewport Prediction for 360 Video

Joris Heyse*   Maria Torres Vega   Femke De Backere   Filip De Turck

Ghent University - imec, IDLab, Department of Information Technology
Technologiepark-Zwijnaarde 126, B-9052 Ghent, Belgium

ABSTRACT

Accurately predicting where the user of a Virtual Reality (VR) application will be looking at in the near future improves the perceive quality of services, such as adaptive tile-based streaming or personalized online training. However, because of the unpredictability and dissimilarity of user behavior it is still a big challenge. In this work, we propose to use reinforcement learning, in particular contextual bandits, to solve this problem. The proposed solution tackles the prediction in two stages: (1) detection of movement; (2) prediction of direction. In order to prove its potential for VR services, the method was deployed on an adaptive tile-based VR streaming testbed, for benchmarking against a 3D trajectory extrapolation approach. Our results showed a significant improvement in terms of prediction error compared to the benchmark. This reduced prediction error also resulted in an enhancement on the perceived video quality.

Keywords: Adaptive 360 video streaming, contextual bandit, VR.

Index Terms: Information system—Information systems applications—Multimedia information systems—Multimedia streaming; Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Virtual reality

1 INTRODUCTION

With the increase in cheap Head Mounted Displays (HMD), the demand for immersive content and Virtual Reality (VR) applications is also rising. In such applications, accurate predictions of the behavior of users within the omnidirectional environment allows to improve the perceived quality. One prominent example is adaptive tile-based 360° video streaming. Adaptive tile-based streaming exploits the fact that at any given moment the user can only observe a small portion of the entire panoramic video, *i.e.*, the viewport. It encodes the videos at different quality levels, splits the streams into temporal segments and divides these segments into spatial tiles. During streaming, tiles within the viewport are delivered in high quality while the others are in low quality [4]. In order to ensure a good experience, the system needs accurate predictions of the viewed viewport for the next segment. However, because of the unpredictability of head movement, this is currently a big challenge.

Existing solutions are either content-aware or content agnostic. Most content-aware approaches require preprocessing of the video content and thus, are too computationally intensive for real-time predictions and unfit for live VR video streaming (*e.g.*, Pano2vid [3]). Content agnostic solutions, in contrast, are much more lightweight as they only require the HMD sensor data. These approaches tend to assume linear motion for the viewing behavior, such as the linear regression approach of Qian et al. [2]. Recently, Xie et al. [5] presented a Machine Learning (ML) approach based on pre-clustering users and multi-user regions of interest. However, their solution requires a lot of training data, while each video needs its individually trained model.

*e-mail: joris.heyse@ugent.be

As many users do not move very often, existing general solutions are often biased towards predicting that the user never moves at all. However, other users, in contrast, might move very much and erratically, resulting in bad predictions for these users. An ML model, that first predicts the likelihood of the user moving significantly and next predicts how they move, could improve the performance, while avoiding per-video preprocessing. Furthermore, since each user produces a unique trajectory by watching a video, the model needs to adapt to new data and become more personalized over time.

Reinforcement Learning (RL) models can provide the required personalization and per-stage learning. Their general principle is that the agents selects an action according to some trained policy based on some state that is presented to it (*i.e.* the context). In response to choosing an action, the agent receives a reward to reinforce good decisions and penalize bad ones. A well known subclass of RL are the Contextual Bandits (CB) agents, in which the chosen action has no influence on the next state. This characteristic makes CB agents the best fit for viewport prediction as the predicted location has no influence on where the user will actually look next.

In this paper, we present a novel CB learning-based viewport prediction approach for 360° video, which by only using the HMD sensor data, is able to accurately predict the trajectories of different types of users. The method has been evaluated on an adaptive tile-based streaming setup for 360° video, and has been benchmarked against a 3-dimensional extrapolation method.

2 APPROACH

In general, the system (Figure 1) takes in a set of previous orientations of the HMD and feeds this to a 2D CB agent bank, which predicts, in two stages, if the user is bound to move and in which direction. The outcome of the agents results in the final prediction.

The 3D data $[x_t, y_t, z_t]$ from the HMD orientation is first processed and projected onto the 2D space, providing the longitudinal θ_t and the latitudinal ϕ_t coordinates on the 2D equirectangular projection of the sphere. Because of the wraparound of the projection in the longitudinal dimension, *i.e.,* values *pi* and $-pi$ correspond to the same position on the sphere, ϕ_t is first converted to the unit vector (cosine and sine of the angle) before being put in the CB bank. A sliding window provides the last k samples to the CB bank. In order to simplify the prediction of the next position, the state space is discretized. Specifically, each position sample $[\theta_t, \phi_t]$ on the video is mapped to a cell $[\theta_t^i, \phi_t^j]$ of a grid with dimensions $n \times m$, where i and j are the row and column of the cell closest to the actual position. The CB bank consists of four Contextual bandits, one per dimension and decision (either motion or direction). The predicted actions of the four CB agents $[A_{t+1}^{mot_\theta}, A_{t+1}^{dir_\theta}, A_{t+1}^{mot_\phi}, A_{t+1}^{dir_\phi}]$ are subsequently used in the following algorithmic block to obtain the discretized predicted decisions $[\hat{\theta}_{t+1}^i, \hat{\phi}_{t+1}^i]$.

3 RESULTS ON AN ADAPTIVE TILE-BASED VR VIDEO STREAMING SYSTEM

A testbed was developed for the adaptive tile-based streaming use case of 360° video [1]. A mobile HMD is wirelessly connected to a VR content server that performs all predictions, stitching and encoding of the tiles before streaming the video to the HMD.
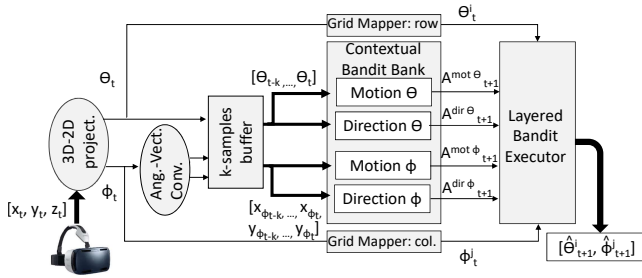
Figure 1: Block diagram of the contextual bandit learning based viewport prediction algorithm.

		user 1	user 2	*user 3*	user 4	user 5
Error	Bandit	*0.454*	**0.308**	*0.383*	**0.696**	**0.342**
	3D-Extr.	*0.621*	0.393	*0.552*	0.992	0.568
Relative quality	Bandit	*0.842*	**0.835**	*0.845*	**0.855**	**0.858**
	3D-Extr.	*0.741*	0.781	*0.763*	0.647	0.804

Table 1: Top: The average prediction error of the algorithms expressed in radians (lower is better). Bottom: The percentage of overlap between the viewport and the predicted tile, giving the quality of the viewport (higher is better).

The study of the performance of the CB learning system was done with 5 users, while watching the Spotlight Stories Video: HELP, produced by Google[1]. The first three users were given no instructions except for watching the video as they normally would for the first time. The last two users were given instructions to move very erratic and to move continuously at a constant speed, respectively. The video was encoded with two qualities (full quality and low quality) and a 4×4 tiling scheme. Video segments were 1000ms so predictions were also done over 1000ms, while sensor samples were collected every 200ms. Regarding the CB approach, the sliding window was set to hold 5 samples and the entire input space was discretized in 16×32 cells for the prediction. The system was trained on the data from users 1 and 3, and the evaluation was performed on all users. The proposed approach was compared with a 3-dimensional extrapolation method consisting of extrapolating the current sample based on the speed estimate obtained from the last two samples in the 3D Cartesian space. Both prediction algorithms were evaluated according to two metrics, namely the prediction error and the relative quality. The error was defined as the angular difference between the prediction and the actual label sample for each time step. The quality was calculated as the percentage of the viewport that overlaps with the surface of the predicted tile. Since the predicted tile is streamed at the highest quality and the rest at the lowest quality, this metric gives an indication of the relative quality of the perceived video.

The top of Table 1 shows the average prediction errors for all 5 users. The average error was calculated for both the bandit model predictions and the extrapolations. From this Table, it is clear that the bandit learning approach results in much better performance in terms of prediction error. As it could be expected, the extrapolation method performs worse for user 4, because the extrapolation is based on the heuristic that motion in one direction will continue in that direction at the same speed. However, the bandit approach, which was trained on users that did not show abnormal viewing behavior, performs significantly better in this case. This suggests that the bandit agent is able to learn to anticipate certain motion.

Prediction error only gives half of the view on the performance and effectiveness of using a bandit agent to predict viewing behavior for adaptive 360° video streaming. The quality of the video within
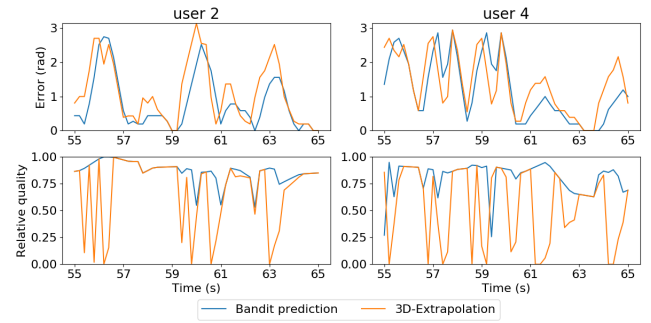

Figure 2: Plots for the prediction error (top) and relative viewport quality (bottom) of users 2 (left) and 4 (right) over a time window of 10 seconds.

the viewport is the most interesting outcome. The bottom of Table 1 shows the average relative quality within the viewport over the entire video for each user. The Table shows that the bandit agent approach performs better than the extrapolation method for all users. This means that even if there are errors in the prediction, due to the fact that the CB learning model is able to learn patterns, they are less noticeable for the quality.

To further illustrate, Figure 2 plots out the prediction error and perceived quality of users 2 and 4 over a 10 seconds time window. These plots show that indeed, our approach does not only perform better on average, but persistently outperforms the extrapolation method. This is clear from the two rightmost plots, where the CB method avoids the quality to drop to 0%.

4 CONCLUSION

This paper presents a novel CB learning based approach to viewport prediction for 360° video. The proposed solution tackles the problem with layered self-learning components. The algorithm was applied on an adaptive tile-based streaming use case. The results have shown that our method results in better viewport prediction in comparison with a 3-dimensional extrapolation method. This reduced prediction error also results in improved, perceived video quality by the users. Future research will investigate the possibility of personalized and individual models for improved predictions, based on specific user behavior.

REFERENCES

[1] J. Heyse, M. Torres Vega, T. Wauters, F. De Backere, and F. De Turck. Effects of Adaptive Streaming Optimizations on the Perception of 360 Virtual Reality Video. In *Proceedings of ITC 30*, pp. 89–92, 2018.

[2] F. Qian, L. Ji, B. Han, and V. Gopalakrishnan. Optimizing 360 video delivery over cellular networks. In *Proceedings of the 5th Workshop on All Things Cellular: Operations, Applications and Challenges*, ATC '16, pp. 1–6. ACM, 2016.

[3] Y.-C. Su, D. Jayaraman, and K. Grauman. Pano2vid: Automatic cinematography for watching 360 videos. In *Eurographics Workshop on Intelligent Cinematography and Editing*, pp. 154–171. Springer, 2017.

[4] J. van der Hooft, M. Torres Vega, S. Petrangeli, T. Wauters, and F. De Turck. Optimizing Adaptive Tile-Based Virtual Reality Video Streaming. In *Proceedings of IFIP/IEEE IM 2019*, 2019.

[5] L. Xie, X. Zhang, and Z. Guo. Cls: A cross-user learning based system for improving qoe in 360-degree video adaptive streaming. In *Proceedings of ACM MM 2018*, pp. 564–572. ACM, 2018.

---

[1]https://www.youtube.com/watch?v=G-XZhKqQAHU