# OPV: BIAS CORRECTION BASED OPTIMAL PROBABILISTIC VIEWPORT-ADAPTIVE STREAMING FOR 360-DEGREE VIDEO

*Weihong Lin , Xinggong Zhang, Zongming Guo* * *, Wei Hu*

Institute of Computer Science & Technology, Peking University, Beijing, China

{linweihong, zhangxg, guozongming, forhuwei}@pku.edu.cn

## ABSTRACT

Nowadays the world is more immersive than ever, but huge bitrate and large Internet delay impede the wide applications of 360-degree video. Viewport adaptive streaming is emerging to transmit quality-variable videos based on user's viewport. To prevent the playback stalling, it is necessary to prefetch some subsequent videos within user's future viewport by head motion prediction. However, long-term prediction is easier biased, which results in videos quality drop and quality oscillation. To alleviate the bias' ill effect, it has to decide whether to download a new segment, or replace old tiles with more accurate viewing probability. To this end, we propose a bias correction based optimal probabilistic viewport-adaptive approach, which selects the best tiles and bitrates to alleviate the ill effect of prediction bias. Besides, viewport quality, playback stalling, and quality oscillation are jointly considered in the optimization model to maximize the Quality of Experience in the viewport. The experimental results confirm that our proposal outperforms the existing viewport adaptive approaches.

***Index Terms***— 360-degree video, tile-based, viewport adaptive streaming, probabilistic, viewport prediction bias

## 1. INTRODUCTION

Nowadays, 360-degree video's applications are more and more popular with people, due to its more immersive experience. But huge bitrate demand and large Internet delay impede the wide applications of 360-degree video. In order to economically efficiently transmit the 360-degree video, the *tile-based* streaming [1–5] is emerging as promising viewport adaptive streaming. But, in 360-degree video streaming, its requirement on network latency is much higher than planer video. Therefore, to better avoid playback stalling and motion-to-photon latency, it is much necessary for client to pre-fetch some amount of subsequent video segments based on user's predicted future viewport. To predicted user's future viewport, there has been a large body of works [1–8] contributed to the viewport prediction. However, for these viewport predictions, they still pose three stringent challenges:

1. The playback stall: To provide more immersive experience for users, the demand for Internet bandwidth is much large for 360-degree video. But, due to harsh bandwidth constraint and motion-to-photon delay, it is easy to induce playback stalling for 360-degree video.

2. The long-term prediction bias: It is difficult to accurately predict the long-term future viewport. Besides, with the increase of prediction time, the bias of viewport prediction will increase significantly.

3. The temporal quality oscillation: due to bandwidth fluctuation and inconsistency of the accuracy of viewport prediction over different temporally video segments, the quality of downloaded video segments are different, which will induce the temporal quality oscillation, and further degrade user's viewing experience.

In order to alleviate the ill effect of long-term prediction bias on viewport quality drop and quality oscillation, while guaranteeing playback continuity, it has to consider a trade-off problem, **whether to download a new segment to avoid potential stalling, or replace old tiles in the buffer by means of more accurate viewing probability?** The trade-off problem is important and meaningful, but so far, there have been not yet existing works to study it. Our paper is **the first work** to contribute to this trade-off problem.

In this paper, we propose a bias correction based optimal probabilistic tile-based streaming, which leverages the viewing probability distribution model and the expected quality utility optimization framework to select the best tiles and bitrates, to alleviate the ill effect of long-term prediction bias. Besides, the viewport quality, playback stalling and quality oscillation are jointly considered in the optimization model to maximize the Quality of Experience in the viewport. Since the optimization problem is NP-hard, we derive its reformulation, which can easily obtain approximate solution with

* Corresponding author (E-mail): guozongming@pku.edu.cn

This work was supported by National R&D project of China under contract No. 2018YFB0803702, Beijing Culture Development Funding under Grant No.2016-288, Ali AIR Funding under Grant No.XT622018001708.

978-1-5386-9214-1/19/$31.00 ©2019 IEEE
DOI 10.1109/ICMEW.2019.00072

provable lower and upper bounds. The experimental results confirm that our proposal outperforms the state of the art viewport-adaptive approaches, in terms of viewport quality, playback stalling and quality oscillation.

The main contributions in this paper are as follows:

- The trade-off problem is formulated into novel probabilistic model, which concurrently considers the viewport quality, playback stalling and quality oscillation;
- The experimental results demonstrate that our proposal can better correct the bias of long-term prediction while guaranteeing playback continuity, compared with the existing viewport adaptive streamings.

The rest of paper are organized as follows. Section 2 surveys the related works on tile-based viewport adaptive streaming. In Section 3, we formulate the trade-off problem into optimization model, and then derive reformulation. Performance evaluation is presented in Section 4. Finally, we make a conclusion in Section 5.

## 2. RELATED WORKS

Tile-based streaming [1–5] is considered as key supporting technology for encoding the quality-variable videos, which has been recently extended to meet the demand of 360-degree video streaming. To better avoid playback stalling and motion-to-photon delay, it is very necessary to pre-fetch some subsequent video tiles based on user's predicted future viewport. So far, there have been a large body of works [1–8] contributing to viewport prediction. The simple extrapolation is commonly used in the works [1, 6, 8]. Specifically, Mavlankar et al. [6] adapts autoregressive moving average model to extrapolate user's future viewport, while Quan et al.[1] and Stefano et al. [8] apply Linear Regression to do it. However, the bias of long-term prediction in these ways significantly enlarges with prediction time increases [3], which results that more blank blocks will be rendered in the screen.

To effectively alleviate the blank block problem, the probabilistic viewport predictions [2–5, 7] are emerging. More concretely, Fan et al. [5] leverage the saliency detection and the Recurrent Neural Network to study the region of interest on content. However, this algorithm is very high-computational. Besides, the correlation between image saliency map with user's preference has not been fully investigated, which makes the prediction results not reliable enough. 360ProbDASH[2] proposes a relatively low-complexity way to predict viewing probability through statistically analyzing the Linear Regression's prediction errors. CUB360[4] also estimates viewing probability of tiles by the means of linear regression, and amend its prediction bias through considering cross-users viewing preference on video content. However, the long-term prediction bias still exists in 360ProbDASH and CUB360, since they are both based on the linear regression.

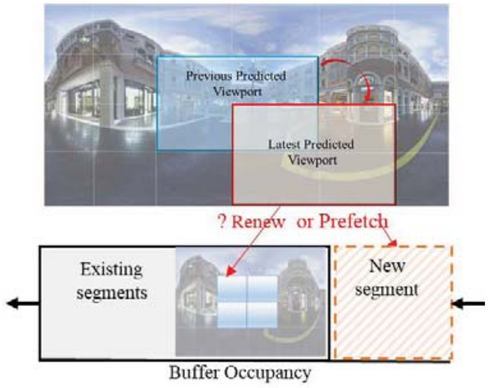



**Fig. 1**. The trade-off problem in the probabilistic tile-based streaming: whether to download a new segment, or replace old tiles of segment in the buffer by means of more accurate viewing probability.

Based on the observation that most users are drawn to similar Region of Interest (ROI) on 360-degree video content, CLS[3] estimates the viewing probability of tiles by analyzing the distribution (over the future tiles) of the viewing frequencies of users in cluster, whose users has the same ROI. However, it is also difficult to accurately predict the long-term future cluster to which the users should belong. Therefore, it is inevitable that long-term prediction bias still exists in CLS.

In summary, the long-term prediction bias still exists in the existing works [1–8]. Moreover, it is also important to avoid playback stalling for 360-degree video streaming. Therefore, to address these challenges, it is of importance to consider the trade-off problem, whether to download a new segment (to avoid potential stalling) or replace the old tiles in the buffer by means of more accurate viewing probability.

## 3. PROBLEM FORMULATION

### 3.1. Assumption and Definition

According to the long-term prediction bias in probabilistic tile-based streaming [2–5], there are two main observations: *1) for segments $k > l$ in playback sequence, $k$'s accuracy of predicting viewing probability of tiles is lower than that of $l$ when they are predicted at the same time; 2) for one segment $l$, its accuracy of predicting viewing probability of tiles at previous prediction time is lower than that at current prediction time.* Based on these observations, we try to address the trade-off problem based on current more accurate viewing probability. Now, we give some definitions and assumptions.

In tile-based streamings [2–5], a raw 360-degree video is temporally divided into several video segments with same duration $T$, and each segment is characterized by viewing probability over its tiles. We assume that, each video segment is spatially cropped into $N$ tiles, in which each tile is encoded into $M$ bitrate-levels. Let $r_{ij}^l$ denote bitrate of tile $i \in [1, N]$ in segment $l$ with bitrate-level $j \in [1, M]$. Let $\{p_i^l\}$ denote the normalized viewing probability over tile $i$ in a segment $l$ at current prediction time.

We assume the client selects and downloads one segment with some tiles at a time, so that bandwidth can be estimated more accurately. Besides, we further assume that segment $l$ can be selected, only when segment $l-1$ has already been downloaded in the playback buffer. Without loss of generality, we assume there are $L-1$ segments in the buffer, which were downloaded based on previous viewport prediction. The segment $L$ is a new segment which has not yet been downloaded.

For each segment $l \in [1, L]$, there are $N$ (tiles) $\times$ $M$ (bitrate-levels) optional tiles. Let $\vec{\boldsymbol{x}}^l = \{x_{ij}^l\}^{N \times M}$ denote a vector of decision variables, where $x_{ij}^l = 1$ denotes that for segment $l$, its tile $i$ with $j$-th bitrate-level is selected, and $x_{ij}^l = 0$ otherwise. Each tile $i$ in segment $l$ can select at most one bitrate version, i.e.,

$$\sum_{j=1}^{M} x_{ij}^l \leq 1 \quad \forall i \in [1, N], \forall l \in [1, L].$$

Therefore, to address the trade-off problem, our optimization model need to find the optimal decisions $\hat{\vec{\boldsymbol{x}}}^{l^*} = \{\hat{x}_{i^*j^*}^{l^*}\}$, namely, finding a set of optimal tiles $\{i^*\}$ with optimal bitrate-levels $\{j^*\}$ in an optimal segment $l^*$, to maximize the video's quality improvement and minimize the temporal quality oscillation. Obviously, we have $x_{ij}^l = 0$, for $l \neq l^*$.

### 3.2. Quality Improvement Function

In the server side, a quality of tile $i$ with $j$-th bitrate-level is denoted as $q_{ij}^l$ for corresponding segment $l \in [1, L]$. In the client side, for each old segment $l \in [1, L-1]$ (in the buffer), a known quality of its old tile $i$ (which had been allocated with a specific bitrate based on its previous viewing probability) is denoted as $Q_i^l$. Without loss of generality, let $Q_i^l = 0$ when tiles $i$ (in segment $l \in [1, L]$) have not yet been downloaded in the buffer, especially $Q_i^L = 0$ for new segment $L$.

Accordingly, based on current viewing probability $\{p_i^l\}$ of tiles, the Quality Improvement Function can be designed as:

$$F(\vec{\boldsymbol{x}}^l) = \sum_{i=1}^{N} \sum_{j=1}^{M} p_i^l (q_{ij}^l - Q_i^l) x_{ij}^l \qquad \forall l \in [1, L]. \qquad (1)$$

### 3.3. Quality Oscillation Function

For each old segment $l \in [1, L-1]$ in the buffer, when some of its old tiles are replaced with new tiles, the expected quality of this updated segment $l \in [1, L-1]$ can be expressed as: $\sum_{i=1}^{N} p_i^l \left( \sum_{j=1}^{M} q_{ij}^l x_{ij}^l + Q_i^l \left( 1 - \sum_{j=1}^{M} x_{ij}^l \right) \right)$.

Considering temporal quality oscillation caused by inconsistent prediction accuracy over different segments and bandwidth fluctuation, we design a quality oscillation function, whose core idea is computing difference of expected quality between two adjacent video segments, which is expressed as:

$$S(\vec{\boldsymbol{x}}^l) = \left[ \sum_{i=1}^{N} p_i^l \left( \sum_{j=1}^{M} q_{ij}^l x_{ij}^l + Q_i^l \left( 1 - \sum_{j=1}^{M} x_{ij}^l \right) \right) - \sum_{i=1}^{N} p_i^{l-1} Q_i^{l-1} \right]^2 \qquad \forall l \in [1, L] \qquad (2)$$

where $Q_i^0$ denotes the quality of segment $l = 0$ which are playing back currently.

### 3.4. Satisfying the Playback Continuity and Deadline

Let $b(t)$ denote the remaining playback time of video segments in the buffer at time $t$, and $b_{min}$ the minimal buffer level. Besides, let $c(t) \geq 0$ denote bandwidth estimated at time $t$. Thus, the time that the client takes to completely download a segment $l \in [1, L]$, can be expressed as: $\frac{\sum_{i=1}^{N} \sum_{j=1}^{M} x_{ij}^l \cdot r_{ij}^l}{c(t)}$. Notice that the remaining playback time of the buffer would be increased by $T$ when a new segment $L$ is downloaded in the buffer. Therefore, when client selects a segment $l \in [1, L]$, it need to satisfy the constraint of playback continuity:

$$b(t) + T \cdot 1_{\{l=L\}} - \frac{\sum_{i=1}^{N} \sum_{j=1}^{M} x_{ij}^l \cdot r_{ij}^l}{c(t)} \geq b_{min},$$

where $1_{\{l=L\}}$ denotes a binary indicator.

Besides, each segment $l \in [1, L]$ has a playback deadline, which is the remaining time ($l \times T$) before the player starts playing back it. Thus, each selected segment $l \in [1, L]$, it need to be completely downloaded within its deadline $l \times T$, i.e.,

$$\frac{\sum_{i=1}^{N} \sum_{j=1}^{M} x_{ij}^l \cdot r_{ij}^l}{c(t)} \leq l \times T.$$

### 3.5. Maximizing the Quality Utility Function

Concurrently considering both the quality improvement function (1) and quality oscillation function (2), we design a *Quality Utility Function* over per video segment $l \in [1, L]$, which can be expressed as: $G(\vec{\boldsymbol{x}}^l) = F(\vec{\boldsymbol{x}}^l) - \delta \cdot S(\vec{\boldsymbol{x}}^l)$, where $\delta > 0$ denotes a weight of quality oscillation function. Notice that, to indeed improve the expected quality of video in the buffer, the optimal segment $l^*$ have to satisfy $F(\vec{\boldsymbol{x}}^l) > 0$. Therefore, for a specific segment $l \in [1, L]$, a problem of how to select a set of tiles with optimal bitrates to maximize the quality utility function (referred to as **tiles-selection problem**), can be formulated into a binary quadratic programming:

$$\underset{\vec{x}^l}{\text{Maximize}}\ G(\vec{x}^l) = F(\vec{x}^l) - \delta \cdot S(\vec{x}^l) \tag{3a}$$

$$\text{Subj. to:}\quad F(\vec{x}^l) \geq 0 \tag{3b}$$

$$b(t) + T \cdot 1_{\{l=L\}} - \frac{\sum_{i=1}^{N}\sum_{j=1}^{M} x_{ij}^l \cdot r_{ij}^l}{c(t)} \geq b_{min} \tag{3c}$$

$$\frac{\sum_{i=1}^{N}\sum_{j=1}^{M} x_{ij}^l \cdot r_{ij}^l}{c(t)} \leq l \times T \tag{3d}$$

$$\sum_{j=1}^{M} x_{ij}^l \leq 1 \qquad \forall i \in [1, N] \tag{3e}$$

$$x_{ij}^l \in \{0, 1\} \qquad \forall i \in [1, N], \forall j \in [1, M] \tag{3f}$$

Let $\mathcal{D}_1^l$ denote the set of vectors $\vec{x}^l$ satisfying the aforementioned constraints (3b)-(3f).

Moreover, the viewing probabilities of tiles in each segment $l \in [1, L]$ , which are predicted at the same time, are endowed with weight of reliability, denoted by $w^l > 0$. Of course, $w^l$ decreases with the increasing of $l \in [1, L]$. Therefore, the trade-off problem can finally be formulated into:

$$\underset{l \in [1,L]}{\text{Maximize}} \left\{ \underset{\vec{x}^l}{\text{Max}} \quad G(\vec{x}^l) \cdot w^l \right\} \tag{4a}$$

$$\text{Subj. to:}\quad \vec{x}^l \in \mathcal{D}_1^l \qquad \forall l \in [1, L] \tag{4b}$$

### 3.6. Problem Reformulation

Since solution to model (4) depend mainly on solution to model (3), in order to simplify the solution process, we reformulate the tiles-selection problem (3). Let $Q_0^l$ denote a known difference of expected quality between two adjacent segments $l \in [1, L]$, i.e., $Q_0^l = \sum_{i=1}^{N} p_i^{l-1} Q_i^{l-1} - \sum_{i=1}^{N} p_i^l Q_i^l$. Let $a_{ij}^l = p_i^l \cdot (q_{ij}^l - Q_i^l)$ denote a known value of expected quality improvement of each segment $l \in [1, L]$. Therefore, the objective function of model (3) can be reformulated into:

$$\begin{aligned} G(\vec{x}^l) &= \sum_{i=1}^{N}\sum_{j=1}^{M} a_{ij}^l x_{ij}^l - \delta \left( \sum_{i=1}^{N}\sum_{j=1}^{M} a_{ij}^l x_{ij}^l - Q_0^l \right)^2 \\ &= -\delta \left( \sum_{i=1}^{N}\sum_{j=1}^{M} a_{ij}^l x_{ij}^l - \frac{2\delta Q_0^l + 1}{2\delta} \right)^2 + \frac{4\delta Q_0^l + 1}{4\delta} \end{aligned}$$

Accordingly, we can obtain theorem 1 when $Q_0^l \leq -\frac{1}{2\delta}$:

**Theorem 1.** $\hat{\vec{x}}^l = \mathbf{0}$ *is an optimal solution to problem (3) over video segment* $l \in [1, L-1]$*, when* $\sum_{i=1}^{N} p_i^{l-1} Q_i^{l-1} - \sum_{i=1}^{N} p_i^l Q_i^l \leq -\frac{1}{2\delta}$.

Let $R(l) = c(t) \cdot \min\left\{ \left( b(t) - b_{min} + T \cdot 1_{\{l=L\}} \right), l \cdot T \right\}$ and $\vec{a}^l = \{a_{11}^l, ...a_{1M}^l, a_{21}^l, ..., a_{2M}^l, ..., a_{N1}^l, ..., a_{NM}^l\}^T$, $\boldsymbol{A}^l = \vec{a}^l \cdot \vec{a}^{l^T}$ ,thus we obtain Theorem 2 when $Q_0^l > -\frac{1}{2\delta}$:

**Theorem 2.** *When* $Q_0^l > -\frac{1}{2\delta}$ *holds, the optimal solution to problem (3) over segment* $l \in [1, L]$ *is equivalent to the optimal solution to model (5):*

$$\underset{\vec{x}^l}{\text{Minimize}} \quad \vec{x}^{l^T} \boldsymbol{A}^l \vec{x}^l - \frac{2\delta Q_0^l + 1}{\delta} \cdot \vec{a}^{l^T} \cdot \vec{x}^{l^T} + \left( \frac{2\delta Q_0^l + 1}{2\delta} \right)^2$$

$$\text{Subj. to:}\quad \sum_{i=1}^{N}\sum_{j=1}^{M} x_{ij}^l \cdot r_{ij}^l \leq R(l) \tag{5a}$$

$$\sum_{j=1}^{M} x_{ij}^l \leq 1 \qquad \forall i \in [1, N] \tag{5b}$$

$$x_{ij}^l \in \{0, 1\} \qquad \forall i \in [1, N], \forall j \in [1, M] \tag{5c}$$

It's easy to prove the two above theorems. Although both original model (3) and model (5) are binary quadratic programming, which are NP-hard, the objective function of model (5) is semidefinite function. Thus, by the means of Semidefinite Programming Relaxation [9], it's easy to obtain approximate solutions to model (5), which has provable lower and upper bounds on the optimal value [9].

## 4. PERFORMANCE EVALUATION

To validate the efficiency of our proposal (OPV), three typical 360 video streaming methods are selected as the comparisons.

- ERP: The Equirectangular Projection (ERP) transmits the whole 360-degree video. We select it as a *baseline method*, since it is widely developed on Internet.
- Tile-LR: Linear Regression (LR) is a commonly used extrapolation in viewport prediction. In Tile-LR [1], bitrate of tiles in predicted viewport are allocated equally.
- 360ProbDASH [2]: The 360ProbDASH is a representative probabilistic viewport adaptation with relatively low computation complexity. It predicts the viewing probability by analyzing prediction errors of LR model.

Besides, we mainly consider the following metrics:

- Viewport PSNR (V-PSNR): The metric directly indicates the video's quality in the user's viewport, which is commonly used in viewport predictions [2–4].
- Temporal Quality Oscillation: To quantify this metric, we calculate the coefficient of variation of the video's quality in the viewport over time-sequence video.

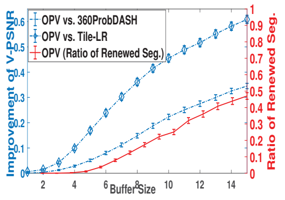



**Fig. 2**. V-PSNR vs. Ratio of Renewed Segments in buffer

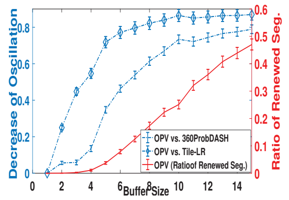
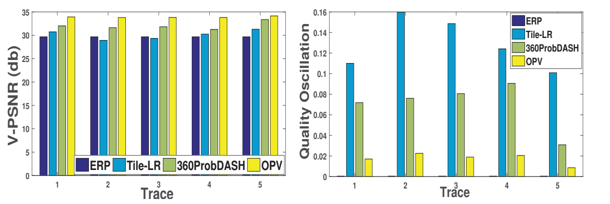
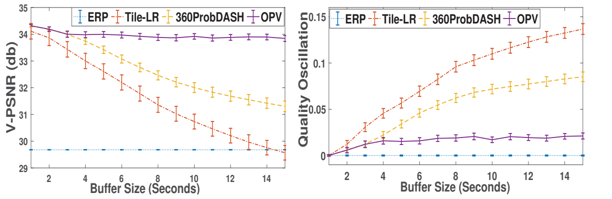



**Fig. 3**. Oscillation vs. Ratio of Renewed Segments



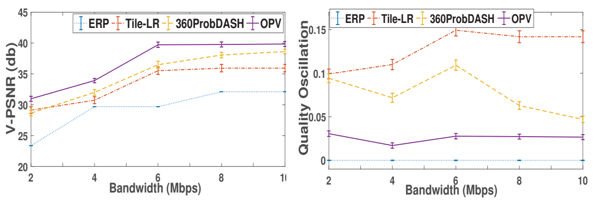

(a) Viewport PSNR (b) Temporal Quality Oscillation

**Fig. 4**. Under 5 traces (bandwidth=4Mbps, buffer size=10s)




(a) Viewport PSNR (b) Temporal Quality Oscillation

**Fig. 5**. Under various buffer sizes(bandwidth=4Mbps, trace1)




(a) Viewport PSNR (b) Temporal Quality Oscillation

**Fig. 6**. Under various bandwidths (buffer size=10s, trace 1)

- Stall Ratio [2]: This metric calculates the percentage of the duration of stall over the total video streaming time. We measure it under real-world Internet bandwidth.

In the tests, we choose a video sequence, which is about 3 minutes long, and 5 users' head motion traces on this sequence, which both are provided by AT&T [1]. According to test settings in [2], the sequence is cropped into segments with 1s duration ($T = 1s$), and then each segment is cropped into $6 \times 12$ tiles. For each tile, the bitrate levels are set as $\{20kbps, 50kbps, 100kbps, 200kbps, 300kbps\}$. Besides, the weight of temporal quality oscillation in our objective function (3a) is empirically set to $\delta = 0.0015$. The viewing probabilities over tiles are predicted by statistically analyzing prediction errors of Linear Regression [1] prediction. The weight value of prediction's reliability $w^l$ over segment $l$ ($l \in [1, L]$) is set as: $w^l = L - l + 1$. The results are averaged with the $95\%$ confidence intervals in Fig.2-3 and Fig.5-6.

### 4.1. Ratio of Renewed video segments in buffer

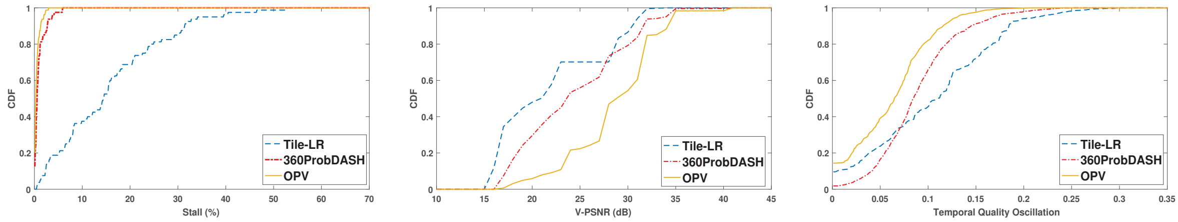
Our optimization model is devoted to addressing the trade-off problem, i.e., whether to download a new segment or replace old tiles in the buffer. In order to have a deep insight into our approach (OPV), we analyze the ratio of old segments (in the buffer), which were already partially renewed, to the total downloaded segments, under various buffer sizes (measured by seconds). Then we further evaluate the effect of this ratio on viewport quality and temporal quality oscillation. The experiments of Fig.2-3 are tested on trace of user1 when bandwidth is fixed as 4 Mbps. As shown in Fig.2-3, when buffer size $\leq 4s$, the ratio of renewed old segments is close to 0. This is because the short-term prediction bias is relatively small, our proposal (OPV) make a better decision to download new segment to avoid potential stalling. With buffer size increases, the ratio of renewed segments significantly increases, in order to alleviate the long-term prediction bias. Accordingly, with the ratio of renewed segments increases, the improvements of V-PSNR of OPV compared with other two tile-based methods (360ProbDASH and Tile-LR) are significantly improved (Fig.2), and the decreases of temporal quality oscillation of OPV compared with the two methods are also increased (Fig.3). Therefore, OPV can better alleviate the negative effect of long-term prediction bias on video's quality.

### 4.2. Viewport PSNR under different conditions

Regardless of buffer sizes, as shown in Fig.5(a), OPV always outperforms the three methods in terms of V-PSNR. With buffer size increases, OPV's V-PSNR just decreases slightly, while both 360ProbDASH and Tile-LR drop significantly. Besides, whenever under various fixed bandwidths (Fig.6(a)) and different users' head motion traces(Fig.4(a)), OPV always has the best performance than other three methods, in terms of V-PSNR.

### 4.3. Quality Oscillation under different conditions

Whenever under various buffer sizes (Fig.5(b)), OPV always has best performance than the two tile-based methods (360ProbDASH, Tile-LR) in terms of temporal quality oscillation. Moreover, whenever under various fixed bandwidths (Fig.6(b)) and traces of different users (Fig.4(b)), OPV always demonstrates its significant superiority over Tile-LR and 360ProbDASH, in terms of quality oscillation. ERP has




(a) CDF of Stall Ratio (%)   (b) CDF of Viewport PSNR   (c) CDF of Temporal Quality Oscillation

**Fig. 7**. Under the bandwidth traces of real-world Internet

the smoothest quality oscillation under these three conditions, since ERP always transmits the whole video with the same bitrate under the fixed bandwidth.

### 4.4. Under the bandwidth traces of real-world Internet

Since OPV, 360ProbDASH and Tile-LR are tile-based approaches, which are more adaptable to the time-varying bandwidth, we evaluate their performances under real-world Internet's bandwidths. Fig.7(a) shows that, in terms of stalling ratio, 360ProbDASH is very closer to OPV while both of them are better than Tile-LR. What's more, OPV almost has the best performance than other two methods, in terms of V-PSNR (Fig.7(b)) and quality oscillation (Fig.7(c)), under time-varying bandwidths. Therefore, our proposal can better alleviate the negative effect of long-term prediction bias on video's quality, while guaranteeing playback continuity.

## 5. CONCLUSION

The long-term prediction bias is a common problem in existing viewport prediction approaches. It is of great importance to alleviate the long-term prediction bias while guaranteeing playback continuity. Therefore, it is meaningful and necessary to consider the trade-off problem, i.e., whether to download a new segment or replace old tiles by means of more accurate viewing probability. Our paper is the first work to propose a probabilistic optimization model to address this trade-off problem. Extensive experiments confirm our proposal's significant superiority over the existing works, in terms of video's quality, temporal quality oscillation and stalling.

## References

[1] Feng Qian, Lusheng Ji, Bo Han, and Vijay Gopalakrishnan, "Optimizing 360 video delivery over cellular networks," in *The Workshop on All Things Cellular: Operations*, 2016, pp. 1–6.

[2] Lan Xie, Zhimin Xu, Yixuan Ban, Xinggong Zhang, and Zongming Guo, "360probdash: Improving qoe of 360 video streaming using tile-based http adaptive streaming," in *ACM on Multimedia Conference*, 2017, pp. 315–323.

[3] Lan Xie, Xinggong Zhang, and Zongming Guo, "Cls: A cross-user learning based system for improving qoe in 360-degree video adaptive streaming," in *ACM on Multimedia Conference*, 2018.

[4] Yixuan Ban, Lan Xie, Xu Zhimin, Xinggong Zhang, Zongming Guo, Shengbin Meng, and Yue Wang, "Cub360: Exploiting cross-users behaviors for viewport prediction in 360 video adaptive streaming," in *IEEE International Conference on Multimedia and Expo (ICME)*, 2018.

[5] Ching Ling Fan, Jean Lee, Chun Ying Huang, Kuan Ta Chen, and Cheng Hsin Hsu, "Fixation prediction for 360 video streaming in head-mounted virtual reality," in *The Workshop on Network and Operating Systems Support for Digital Audio and Video*, 2017, pp. 67–72.

[6] A. Mavlankar and B. Girod, "Pre-fetching based on video analysis for interactive region-of-interest streaming of soccer sequences," in *IEEE International Conference on Image Processing*, 2009, pp. 3025–3028.

[7] Zhimin Xu, Xinggong Zhang, Kai Zhang, and Zongming Guo, "Probabilistic viewport adaptive streaming for 360-degree videos," in *IEEE International Symposium on Circuits and Systems (ISCAS)*, 2018.

[8] Stefano Petrangeli, Viswanathan Swaminathan, Mohammad Hosseini, and Filip De Turck, "An http/2-based adaptive streaming framework for 360 virtual reality videos," in *ACM on Multimedia Conference*, 2017, pp. 306–314.

[9] Jaehyun Park and Stephen Boyd, "A semidefinite programming method for integer convex quadratic minimization," *Optimization Letters*, vol. 63, no. 2, pp. 1–20, 2017.