Deep Reinforcement Learning-based Rate Adaptation for Adaptive 360-degree Video Streaming

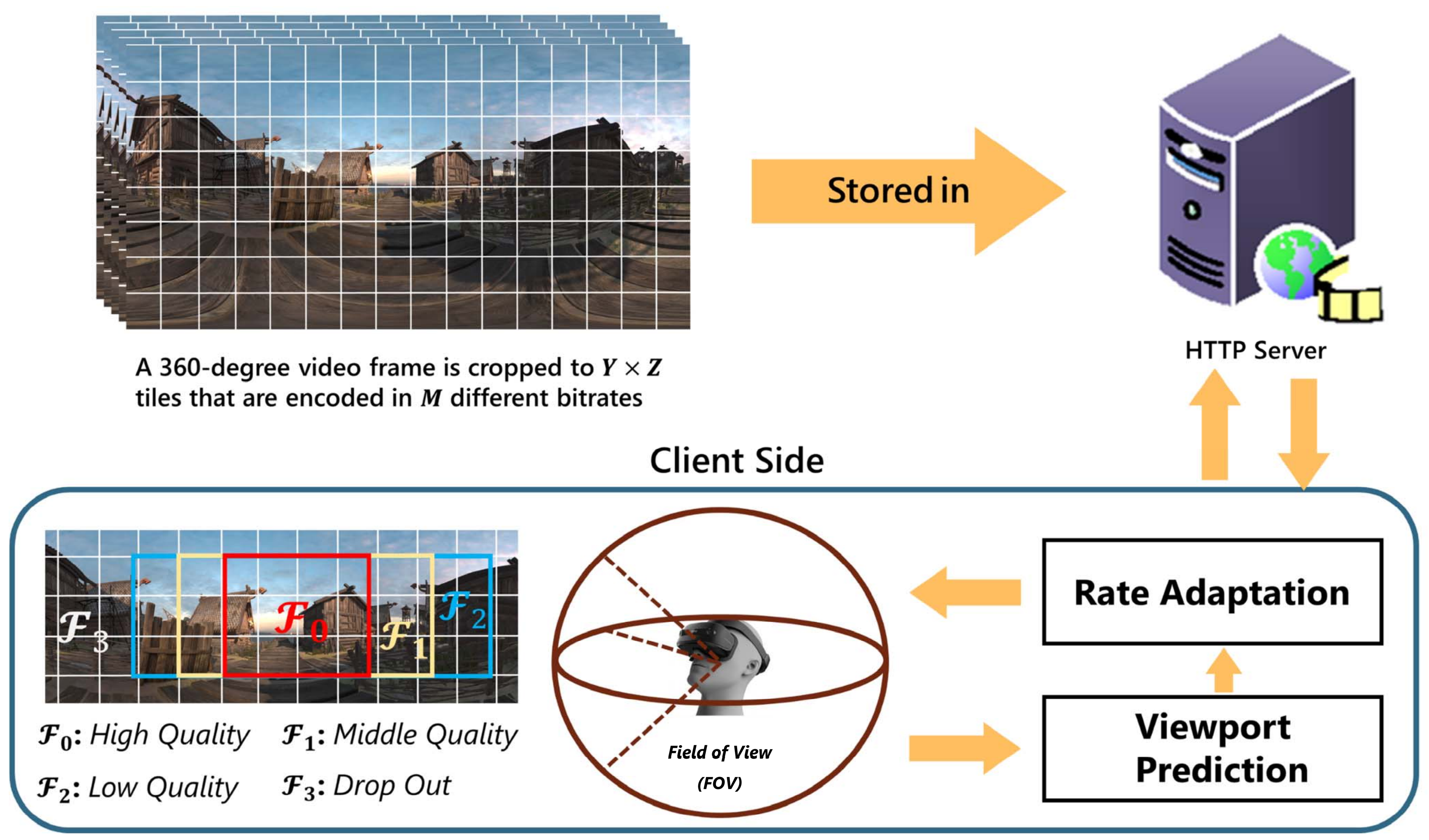
SHANGHAI JIAO TONG UNIVERSITY
上海交通大學
Institute of Media, Information, and Network

Nuowen Kan, Junni Zou, Kexin Tang, Chenglin Li, Ning Liu, Hongkai Xiong
School of Electronic Information & Electrical Engineering, Shanghai Jiao Tong University, China

I. ABSTRACT

Compared to traditional videos, 360-degree videos have much higher resolution and thus require more bandwidth to deliver the entire scene, which may result in rebuffering under the bandwidth-constrained environment. In this paper, we propose a deep reinforcement learning (DRL)-based rate adaptation algorithm for adaptive 360-degree video streaming, which is able to maximize the quality of experience (QoE) of viewers by adapting the transmitted video quality to the time-varying network conditions. Specifically, to reduce the possible switching latency of the field of view (FoV), we design a new QoE metric by introducing a penalty term for the large buffer occupancy. A scalable FoV method is further proposed to alleviate the combinatorial explosion of the action space in the DRL formulation. Then, we model the rate adaptation logic as a Markov decision process and employ the DRL-based algorithm to dynamically learn the optimal video transmission rate. Simulation results show the superior performance of the proposed algorithm compared to the existing algorithms.

II. 360-DEGREE VIDEO DELIVERY



III. OPTIMAZATION PROBLEM

$$\max_{R_{y,z,k}\in\mathcal{R}} \sum_{k=\tau}^{\infty} \gamma^{k-\tau}\mathrm{QoE}_k$$
$$\text{s.t} \quad R_{y,z,k} = R_{y',z',k}, \forall\, T_{y,z}, T_{y',z'} \in \mathcal{F}_i,$$

QoE Metric: $\mathrm{QoE}_k = Q_k - \beta(B_k - d_k)_+ - \lambda\big(d_k - B_k\big)_+ - \mu\sum_{i=0}^{I}\Delta_i,$

Visual Quality: $Q_k = \sum_{i=0}^{I}\mathcal{P}_i \sum_{T_{y,z}\in\mathcal{F}_i} q_{y,z,k},$

Rebuffering time: $(B_k - d_k)_+$

Quality Variance: $\Delta_i = \sum_{T_{y,z}\in\mathcal{F}_i} q_{y,z,k} - q_{y,z,k-1}$

FoV Switching latency: $\big(d_k - B_k\big)_+$

Symbol	Definition
$\{\mathcal{F}_0, \mathcal{F}_1, \ldots, \mathcal{F}_I\}$	Scalable FoV set, which is sorted in a decreasing order of the probability that each FoV is watched by the viewer.
$T_{y,z}$	The y-th row and z-th column tile.
$R_{y,z,k}$	The bitrate for $T_{y,z}$ in k-th chunk.
$q_{y,z,k}$	The quality of $T_{y,z}$ in k-th chunk.
P_i	The average viewing probability of tiles in $\mathcal{F}_i$.
B_k	The buffer occupancy when the client starts to download k-th chunk.
d_k	The download time of k-th chunk.
$(x)_+=\max\{x,0\}$	Ensures that the term is nonnegative.

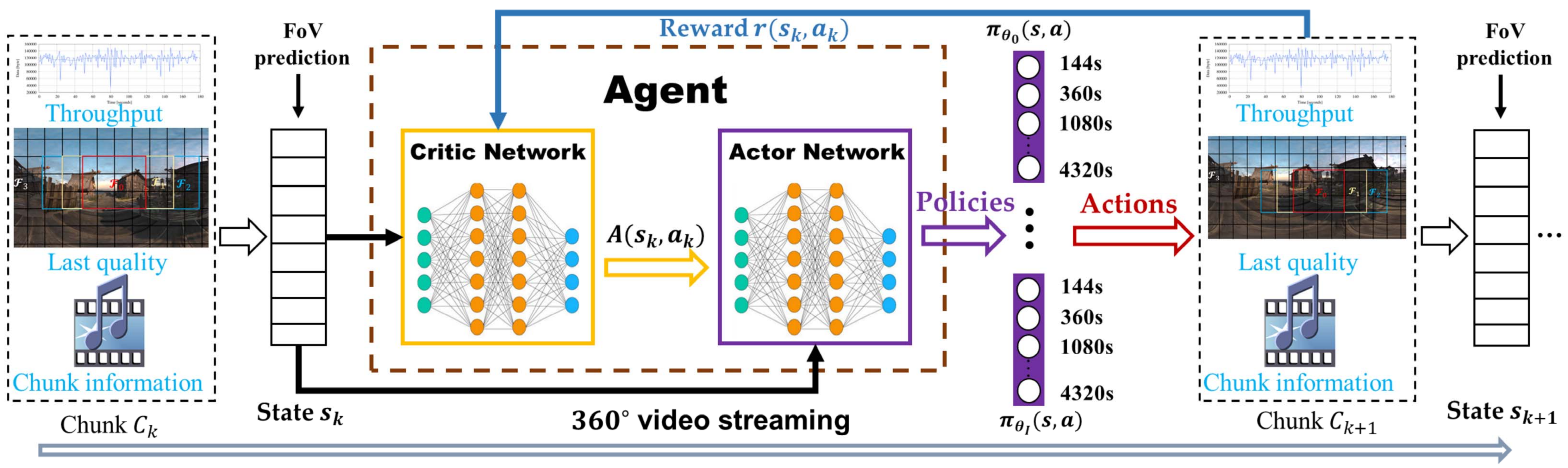
IV. RATE ADAPTATION WITH DRL

Markov Decision Process:

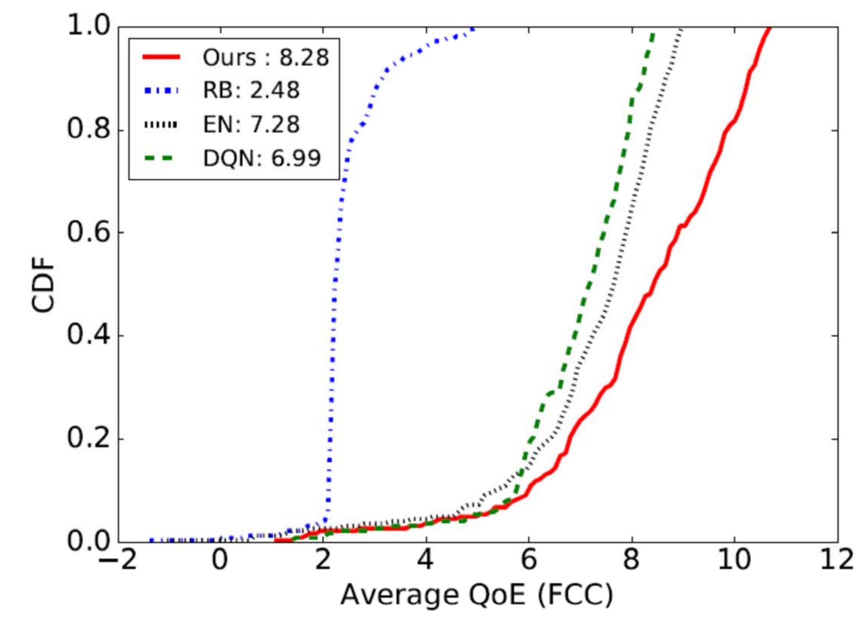
- **State:** the environment, including network throughput, buffer occupancy, etc.
- **Action:** bitrate for each $\mathcal{F}_i$
- **Reward:** QoE_k

Agent Policy:

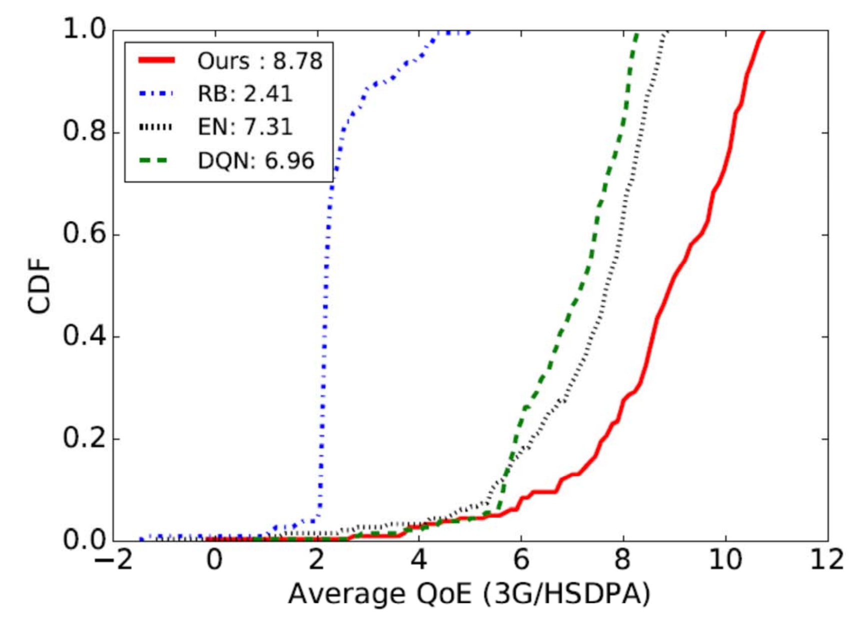
$$\pi(a_k|s_k) = \prod_{i=0}^{I}\pi^i(a_k^i|s_k).$$



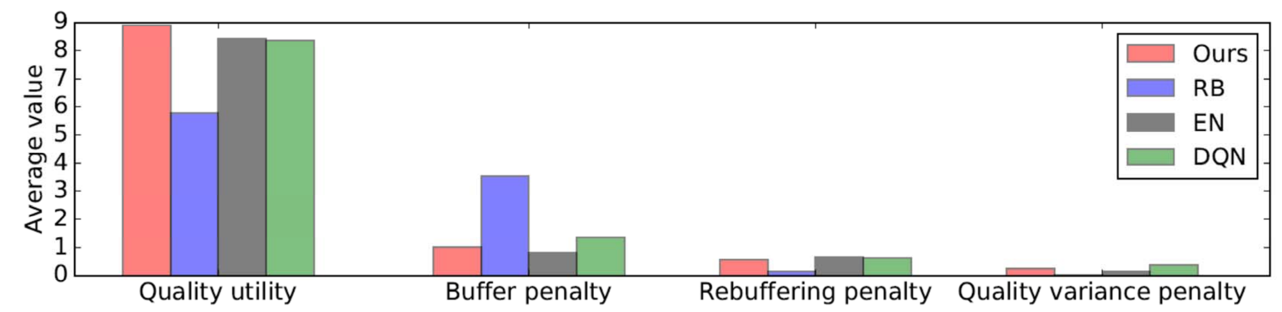
V. EXPERIMENT RESULTS



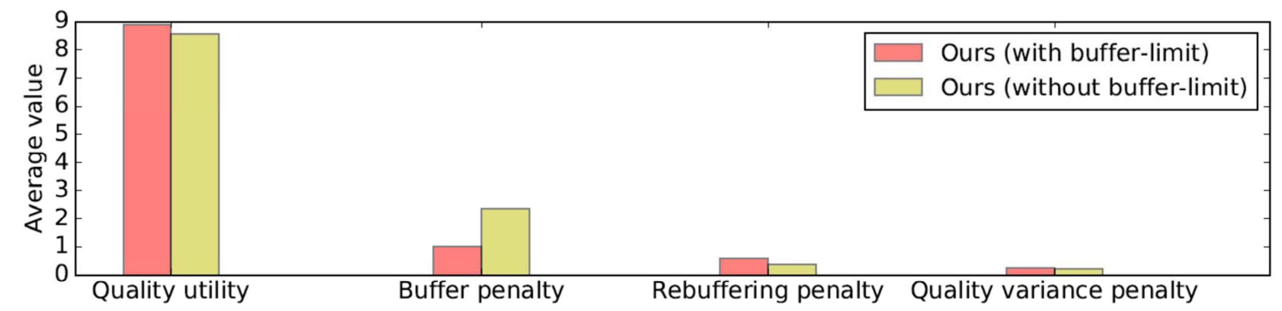
(a)



(b)



(a)



(b)

Reference:

1. C. Liu, N. Kan, J. Zou, and et.al, "Server-side rate adaptation for multi-user 360-degree video streaming," ICIP'18
2. H. Mao, R. Netravali, and M. Alizadeh, "Neural adaptive video streaming with pensieve," SIGCOMM '17
3. L. Xie, Z. Xu, Y. Ban, and et.al, "360Prob-DASH: Improving QoE of 360 video streaming using tile-based HTTP adaptive streaming," ACM MM '17