

## 基于自适应码率移动增强现实应用的能效优化研究

陈乐<sup>1</sup> 高岭<sup>1,2</sup> 任杰<sup>3</sup> 党鑫<sup>1</sup> 王祎昊<sup>1</sup> 曹瑞<sup>1</sup> 郑杰<sup>1</sup> 王海<sup>1</sup>

1 西北大学信息科学与技术学院/新型网络智能信息服务国家地方联合工程研究中心 西安 710127

2 西安工程大学计算机科学学院/新型网络智能信息服务国家地方联合工程研究中心 西安 710600

3 陕西师范大学计算机科学学院 西安 710119

(chenle@stumail.nwu.edu.cn)

**摘要** 随着移动增强现实(Mobile Augmented Reality, MAR)技术的飞速发展, MAR应用的种类及功能也越来越丰富多样,与此同时用户对MAR应用的视频质量及响应时间也提出了更高的要求。通常来说, MAR应用会将计算密集型任务(目标识别及渲染)卸载到云端或边缘服务器进行处理,并将渲染后的图像下载到移动端。但由于移动网络状态的不稳定性及网络带宽的限制,海量数据的传输将延长MAR应用响应时间,进而增加移动设备的传输能耗开销,严重影响用户使用体验。由此,文中提出了一种基于梯度提升回归(Gradient Boosting Regression, GBR)的自适应码率控制模型。该模型通过感知当前网络环境及拍摄内容,预测用户观感需求并对非关注点部分进行低码率压缩,从而在不影响用户体验的情况下尽可能地降低传输数据量,缩短响应时间。具体来说,通过分析200个热门视频的视频特征,构建视频特征同用户观感需求的内在联系,从而针对不同的用户需求提供合适的视频码率配置,由此达到维持体验、减少时延、节约能耗的目标。实验结果显示,同直接下载渲染后的1080p视频相比,提出的自适应码率控制模型在尽可能维持用户观感体验的前提下,每帧的下载时间平均减少了58%(19.13 ms)。

**关键词:** 移动增强现实; 能效优化; 梯度提升回归; 码率自适应控制

**中图分类号** TP391

## Adaptive Bitrate Streaming for Energy-Efficiency Mobile Augmented Reality

CHEN Le<sup>1</sup>, GAO Ling<sup>1,2</sup>, REN Jie<sup>3</sup>, DANG Xin<sup>1</sup>, WANG Yi-hao<sup>1</sup>, CAO Rui<sup>1</sup>, ZHENG Jie<sup>1</sup> and WANG Hai<sup>1</sup>

1 School of Information Science and Technology, Northwest University, State-Province Joint Engineering and Research Center of Advanced Networking and Intelligent Information Services, Xi'an 710127, China

2 College of Computer Science State-Province, Xi'an Polytechnic University, Joint Engineering and Research Center of Advanced Networking and Intelligent Information Services, Xi'an 710600, China

3 School of Computer Science, Shaanxi Normal University, Xi'an 710119, China

**Abstract** With the development of the mobile augmented reality (MAR), users have higher requirements on video quality and response time on it. MAR applications offload computation-intensive tasks to the cloud or edge servers for processing. In order to provide users with high-quality rendering services, MAR needs to download massive amounts of data from cloud or edge servers. Due to the instability of network condition and the limitation of network bandwidth, data transmission will extend MAR application response time, which increases the energy consumption, and seriously affects the user experience. This paper proposes a bitrate adaptive model based on gradient boosting regression (GBR). The model considers the different needs of users in different network conditions, analyzes the features of the 200 popular videos, finds the connection between the video features and the user requirements, and provides appropriate video bitrate configuration according to different needs, thus to achieve the goal of maintaining experience, reducing latency and saving energy. The results show that compared with the original rendered 1080p video, the proposed bitrate adaptive model can save 58% downloading time latency(19.13 ms) per frame while maintaining the user's viewing experience as much as possible.

收稿日期: 2020-11-16 返修日期: 2021-04-12 本文已加入开放科学计划(OSID), 请扫描上方二维码获取补充信息。

基金项目: 国家重点研发计划(2019YFC1521400); 国家自然科学基金(61902229, 61872294); 陕西省国际科技合作计划项目(2020KW-006); 中央高校基本科研业务费专项资金资助项目(GK202103084)

This work was supported by the National Key R&D Program of China(2019YFC1521400), National Natural Science Foundation of China(61902229, 61872294), International Science and Technology Cooperation Project of Shaanxi(2020KW-006) and Fundamental Research Funds for the Central Universities(GK202103084).

通信作者: 高岭(gl@nwu.edu.cn)

**Keywords** Mobile augmented reality, Performance optimization, Gradient boosting regression, Bitrate adaptive control

## 1 引言

移动增强现实(MAR)技术的飞速发展为用户带来了前所未有的观感体验,并被广泛应用于教育<sup>1)</sup>、医疗<sup>2)</sup>、娱乐<sup>3)</sup>等领域。相比于VR对特定硬件设备的依赖性,AR可应用于一般的智能移动设备中,报告显示,预计在2023年MAR的设备安装量将突破25亿<sup>4)</sup>。

MAR应用利用移动端摄像头及传感器获取用户当前环境实时数据,通过对图像目标物体进行实时检测及动画渲染,从而将真实世界同虚拟图像相结合,最终呈现给用户可以进行交互的虚拟世界。如在线视频会议和网络直播等应用就通过在云端对实时画面进行检测、渲染,为真实画面添加特效等虚拟元素,将带有丰富虚拟图像的视频转播给其他参与用户,为用户带来了更好的使用体验。

但是MAR系统在提供优质服务时的高额能耗问题对于移动设备的电池来说是一个巨大的考验。具体来说,MAR运行过程中的目标识别与三维注册都是计算密集型任务,会给资源受限的移动设备带来严重的能耗负担<sup>[1-2]</sup>。近年来,研究人员提出了MAR任务卸载框架<sup>[3-4]</sup>,将移动端采集到的图像数据上传到服务器端,进而利用计算性能更强的云服务器处理MAR过程中的复杂任务,降低移动端的能耗开销。但网络环境的不稳定以及网络带宽的限制,可能导致服务器将渲染好的高清图像下载到移动端的时延较高,出现卡顿的现象,增加移动设备的能耗开销,严重影响用户体验。因此,设计高效率的MAR数据传输方案至关重要。

本文针对MAR应用从云端到移动端的数据传输过程中的高开销问题,通过视频分割技术和码率控制技术对AR视频码率进行动态控制。同时,考虑到AR视频内容、当前网络环境及用户观感需求,构建了基于梯度提升回归(GBR)的自适应码率控制模型。该模型通过预测最优码率配置,对当前拍摄图像内容进行动态编码,在保证用户体验的同时降低MAR数据传输时延。实验结果表明,相比直接下载渲染后的高质量视频,本文提出的自适应码率控制模型可以使MAR视频每一帧的下载时延平均降低58%(19.13ms)。

## 2 相关工作

随着虚拟现实技术的发展和移动设备性能的提升,工业界和学术界一直在探索如何为移动用户提供高质量的增强现实服务<sup>[5-6]</sup>。但由于AR任务的复杂性<sup>[7]</sup>,为了减轻MAR应用在资源受限的本地设备上的负担,通常将计算密集型任务转移到功能强大的云端和边缘设备上完成<sup>[8]</sup>,但是不稳定的网络环境无法保证用户的实时体验,高延迟的响应又严重增加了移动设备的能耗负担。Liu等<sup>[9]</sup>为满足MAR应用中高

识别率与低传输延迟的要求,将每帧的感兴趣区域上传到边缘服务器进行检测任务,以提高MAR的性能。Zhang等<sup>[10]</sup>利用云端进行目标检测,并在移动设备上部署一套轻量级的跟踪系统,使得MAR可以达到30fps的帧率以保证一定的用户观看体验。虽然强大的云端和边缘端设备可以显著提高MAR性能,降低本地移动设备的CPU成本,但是将高分辨率视频帧卸载到服务器也会大大增加端到端延迟。为了降低视频下载成本,Flare<sup>[11]</sup>通过头部的运动轨迹预测用户将会观看的画面,仅获取观众即将看到的全景场景的一部分。也有其他学者<sup>[12-13]</sup>尝试通过降低视频质量来节省下载流量,虽然这些方法都可以减少卸载开销,但牺牲了用户的感知体验,从而错失了一个提高MAR性能的大好机会。

因此,本文研究通过视频语义分割技术,保留用户关注区域的视频质量,降低非关注区域的码率,以降低传输的数据量和传输延迟。目前主流的语义分割技术包括基于图像语义的分割技术(如Pyramid Scene Parsing Network<sup>[14-15]</sup>)和基于视频语义的分割技术(如Netwarp<sup>[16-17]</sup>)。而本文提出的视频分割方法是以图像的显著性检测为基础<sup>[18-20]</sup>,将识别物体从原始视频中分割出来,形成两个新的视频。其中图像显著性检测通常作为各种计算机程序的第一步(影像分割<sup>[21-22]</sup>、视觉追踪<sup>[23]</sup>、图像检索<sup>[24]</sup>等),得到了广泛的应用,本文将应用于视频分割。

目前,机器学习方法在移动端的能效优化方面具有广泛的应用,如Ren等<sup>[25-26]</sup>针对移动端Web浏览的能耗问题,使用机器学习技术根据用户需求及强度调整硬件工作频率,达到降低能耗的目的;Qin等<sup>[27]</sup>面向移动异构平台,对典型的深度学习模型进行压缩及裁剪,分析不同压缩方案对模型性能的影响,从而针对不同模型特点给出合适的压缩方案。但鲜有利用机器学习方法控制MAR应用中视频的码率来提高MAR能效的研究。

因此本文在不同网络下对MAR应用中的视频进行自适应的码率控制,以实现在保证用户体验的同时对下行链路的端到端传输延迟进行优化。

## 3 研究背景

### 3.1 MAR工作流程

MAR应用的工作流程<sup>[28]</sup>如图1所示。移动端将拍摄画面及传感器数据上传至服务器端,进而利用服务器强大的计算能力对上传图像内容进行物体识别、跟踪、渲染等操作,然后将渲染好的图像下载到目标移动端。图2给出了在普通4G网络(上行带宽:1Mbps,下行带宽:2Mbps,延迟:80ms)下MAR应用在各阶段的时间延迟,其中下载过程占整体端

<sup>1)</sup> <https://thinkmobiles.com/blog/augmented-reality-education/>

<sup>2)</sup> <https://sciencebasedmedicine.org/augmented-reality-in-medicine/>

<sup>3)</sup> <https://jasoren.com/ar-entertainment/>

<sup>4)</sup> <https://www.chinaar.com/ARzx/7878.html>

到端延迟的 21%，因此对 MAR 使用过程中的数据下载阶段进行优化十分重要。

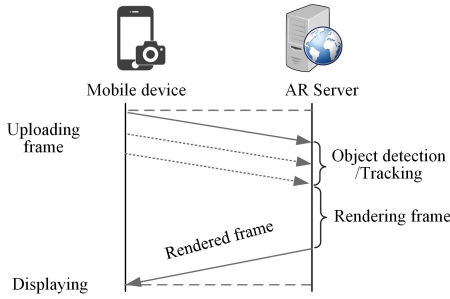


图 1 MAR 的工作流程

Fig. 1 Pipeline of typical MAR systems

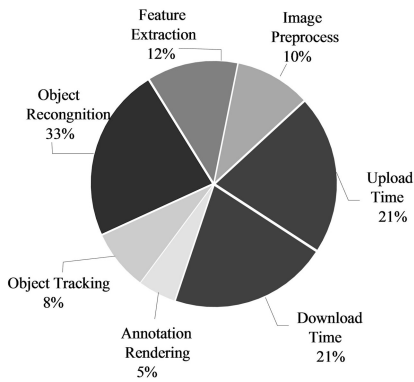


图 2 MAR 各阶段的时延

Fig. 2 Latency detail when running mobile augmented reality

### 3.2 视频语义分割

首先,本文使用基于全卷积神经网络(fully convolutional neural networks)改进的整体嵌套边缘检测(holistically-nested edge detector)图像显著性检测算法<sup>[29]</sup>,对视频帧中的显著目标进行检测和分割,其中显著目标是图像中最具视觉特征的物体或区域,是用户主要观看的区域;然后,分离出仅有图像显著目标的前景帧和包含非显著目标的背景帧。由于前景帧包含了用户关注内容,需尽可能地维持原图图像质量。相反,对于用户非关注部分的背景帧,可以降低这一部分的视频编码质量,在不影响用户整体体验的同时,通过控制编码码率,可大大节省下载流量并保证用户体验。

视频的分割过程如图 3 所示,通过 AR 设备捕获连续的视频帧,对每帧的关注点部分使用图像显著性检测方法进行提取;对提取的三通道灰度图进行二值化操作使每个像素点的值只有 0 和 1,其中关注点部分值为 1,非关注点部分值为 0;使用二值图与原始的 RGB 图像矩阵做点乘运算,如式(1)所示:

$$a_{ij} \cdot b_{ij} = \begin{bmatrix} a_{11}b_{11}, & a_{12}b_{12} & \cdots & a_{1j}b_{1j} \\ a_{21}b_{21}, & a_{22}b_{22} & \cdots & a_{2j}b_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1}b_{i1}, & a_{i2}b_{i2} & \cdots & a_{ij}b_{ij} \end{bmatrix} \quad (1)$$

其中,矩阵 $a_{ij}$ 为二值图,矩阵 $b_{ij}$ 为原始 RGB 图像,如此可得到前景帧;将二值图反相后同理可得到背景帧。对于前景帧和背景帧,通过调用 ffmpeg<sup>1)</sup>的 libavcodec 库来进行最终的视频合成。

#### ROI

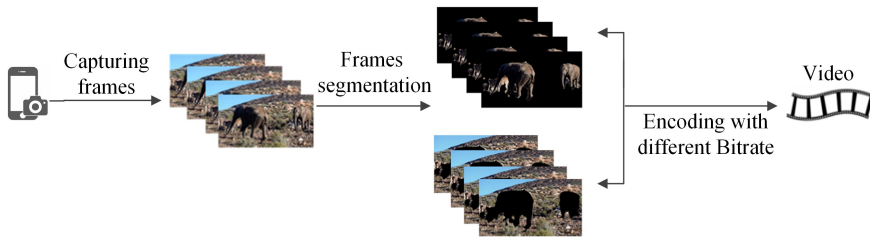


图 3 图像分割流程

Fig. 3 Segmentation workflow

特别地,上述基于语义的视频分割作为预处理操作也是在云端完成的整体操作。通过测量在高性能服务器上进行的视频分割工作,可以得知进行一帧的视频分割仅需 0.03ms,这个时间将在后续的实验作为本文实验预处理耗时加入到整体的时延测量中,本文后续的实验数据均包含此操作的耗时。

本文使用 ffmpeg 库中 ABR(Average Bit Rate)<sup>2)</sup>方法进行码率控制,不仅可以在为简单场景分配较少码率的同时将节约的码率额外再提供给复杂场景,即在保证视频质量的情况下,还可以控制视频大小。

本文主要在两个方面使用了码率控制,一是以低码率对用户非关注部分的背景帧进行编码,以降低数据传输量;二是寻找合适的合成码率将背景帧同用户关注的前景帧进行合

成,在保证一定视频观看质量的同时降低视频的大小。

### 3.3 H. 265

H. 265<sup>[30]</sup>也称 HEVC(High Efficiency Video Coding),是一种新型的视频压缩技术,其目标是在有限的带宽下传输更高质量的视频。相比 H. 264 技术,H. 265 由于添加了更多的帧内预测模式且拥有更好的运动补偿处理与矢量预测方法以提供更强有力的编码方式,从而具有更高的压缩率。本文所有实验均基于 H. 265 编解码器对图像、视频进行处理。

### 3.4 MAR 拍摄内容复杂度

移动设备在相同时间段内因拍摄内容的不同而产生的数据量也是不同的,其根本原因是图像复杂度的差异性,细节越多的图像会产生更多的数据量。由此本文使用协议 ITU-R BT. 1788<sup>[31-32]</sup>提供的 SI 与 TI 来衡量 MAR 拍摄过程中的场

<sup>1)</sup> <https://www.ffmpeg.org/>

<sup>2)</sup> <https://docs.microsoft.com/en-us/windows/win32/medfound/peak-constrained-variable-bit-rate--vbr--encoding>

景差异性导致传输数据量的不同。其中空间观感信息(Spatial Perceptual Information, SI)表征一帧图像中空间细节的复杂度,空间细节越多,SI值越大;时间观感信息(Temporal Perceptual Information, TI)表示时间维度上相邻帧的差异性,如图像中物体高速运动,则TI值较大。因此本文使用SI和TI来判断MAR拍摄内容的复杂度。

#### 4 梯度提升回归

梯度提升回归<sup>[33]</sup>(GBR)是一种集成学习模型<sup>[34-35]</sup>,通过组合多个有偏好的弱学习器从而得到一个表现更好的强学习器。

GBR通过使用分类与回归树<sup>[36]</sup>(Classification and Regression Trees, CART)来进行决策的弱学习器累加,得到最终进行决策的强学习器,其中CART树的决策过程如算法1所示,其希望划分的分支之间的误差越小越好。GBR利用加法模型,每进行一轮迭代就找到一棵使本轮的损失函数最小化的CART树,最终得到多棵CART树,即得到了多个弱学习器,多个弱学习器累加得到最终所需的强学习器。

##### 算法1 CART树的决策过程

输入:训练数据

输出:CART树

1. while 结果不满足停止条件 do
2.     选择特征J和切分点s
3.     对于选定的(J,s)划分区域得到区域J1,J2
4.     确定该区域的预测结果
5. end while
6. return 生成树

与一般的boosting算法利用样本不同权重进行训练不同,GBR是对每次训练时的拟合目标进行操作,GBR训练时将已有弱学习器(CART树)在本轮学习产生的残差作为下一轮的输入,下一轮学习时尽可能地拟合残差值(此处将损失函数的负梯度在当前模型的值作为残差的近似值),为之前的学习器添加一个新的弱学习器(CART树),进而使下一轮输出的残差不断变小。因此GBR可以做到每一轮的学习结果向损失函数减小的梯度方向变化。

GBR共有4种损失函数,具体如下。

LAD:最小绝对值偏差,可以将预测值与真实值的绝对差值的总和最小化。

$$S = \sum_{i=1}^n |y_i - f(x_i)| \quad (2)$$

LS:最小平方误差,可以将预测值与真实值的差值的平方和最小化。

$$S = \sum_{i=1}^n (y_i - f(x_i))^2 \quad (3)$$

Huber:平均绝对误差,是LS与LAD方法的综合,可增强平方误差MSE对于离群点的鲁棒性。

$$L_{\delta}(y_i, f(x_i)) = \begin{cases} \frac{1}{2}(y_i - f(x_i))^2, & \text{for } |y_i - f(x_i)| \leq \delta \\ \delta|y_i - f(x_i)| - \frac{1}{2}\delta^2, & \text{otherwise} \end{cases} \quad (4)$$

Quantile:分位损失,具有不同的权重,对于预测值大于或

小于真实值的惩罚是不同的,一般在需要分段预测时采用。

$$L\gamma(y, f(x)) = \sum_{i \in \{i | y_i < f(x_i)\}} (1-\gamma)|y_i - f(x_i)| + \sum_{i \in \{i | y_i \geq f(x_i)\}} \gamma|y_i - f(x_i)| \quad (5)$$

除此之外,GBR还需要对学习率、弱学习器个数、最小样本数及树深度进行调节,以达到模型最优。

#### 5 MAR应用中的能效问题

由于网络状态的不稳定性及带宽的限制,从云端下载高质量视频(如HD,4K,8K)会带来较高的响应延迟。本节以两个内容复杂度不同的视频为例,通过调节视频码率配置,记录在不同配置下的传输时延和用户观感,旨在找到最优码率配置,即在保证用户体验的前提下,尽可能地降低响应延迟,从而优化MAR能效。

图4(a)、图4(b)为实验选取的两个视频片段,帧率为30帧/s,时长为3s。图5(a)、图5(b)分别是视频的SI和TI,在视频1中,SI的平均值是9,TI的平均值是3;在视频2中,SI的平均值是72,TI的平均值是36。由此可知,选取的两个视频在SI和TI上都具有较大差异。图4(c)、图4(d)给出了视频1和视频2经过服务器处理后需要下载到移动端的视频画面。最后在3种网络环境(见表1)中进行测试,实验结果如图6所示,可以看出通过调节视频码率来降低传输延迟是可行的。

进一步,本文使用PI(Perceptual Index)<sup>[37]</sup>对视频的感知质量进行量化。

$$PI = \frac{1}{2}((10 - Ma) + NIQE) \quad (6)$$

其中,Ma<sup>[38]</sup>和NIQE<sup>[39]</sup>是两个无参考图像质量评价指标。与SSIM,PSNR等通过计算图像像素之间相似性来衡量图像失真程度的客观评价指标不同,PI将感知质量定义为处理后图像的视觉质量,它依据处理后的图像是否像一个有效的自然图像,即与原始图像的人眼观看感受的相似程度来进行评分,PI值越小,处理后的图像与原始图像的观看感受越接近,图像的感知质量越好。

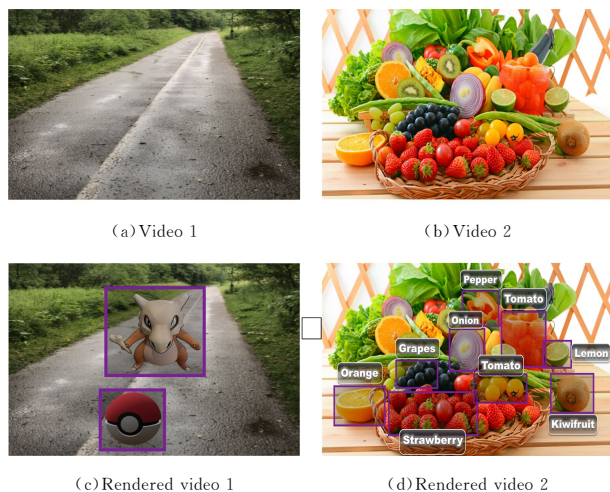


图4 原始视频和渲染后的视频

Fig. 4 Original testing videos and rendered videos

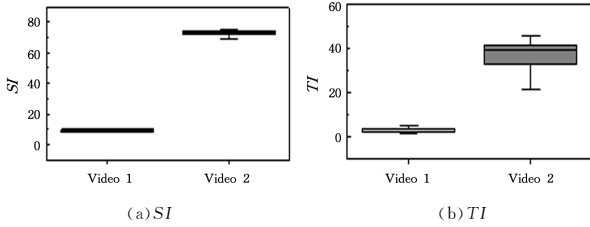


图 5 视频的 SI 和 TI  
Fig. 5 SI and TI for videos

表 1 网络情况

Table 1 Network condition

网络状况	上行速度/Mbp	下行速度/Mbp	延迟/ms
普通 4G	1	2	80
良好 4G	8	15	50
WiFi	15	30	5

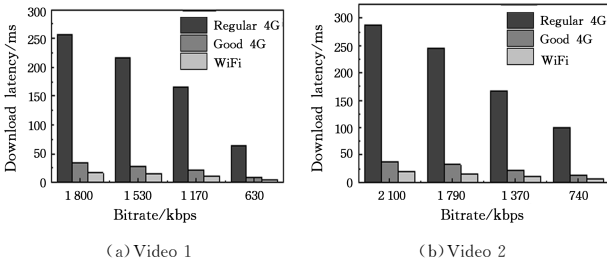


图 6 3 种网络下的下载时延  
Fig. 6 Download latency in three network conditions

本文使用表 2 所列的码率配置对视频进行处理,并根据不同的优化目标(最优的用户观看感受、最低的延迟、可接受观看感受下的最低延迟)对当前视频进行编码。表 2 中 ROI 表示感兴趣区域部分即视频的前景帧码率, NROI 表示非感兴趣区域即视频的背景帧码率, Cmp 表示前景帧与背景帧的合成码率(例如,  $\langle \text{NROI}:1187, \text{Cmp}:1955 \rangle$  表示背景帧码率为 1187 kbps, 合成码率为 1955 kbps)。图 7 给出了两个视频在不同配置下用来量化不同优化目标的 3 个性能指标。

表 2 视频的码率配置

Table 2 Bitrate configurations for videos

码率配置(NROI:背景帧码率, Cmp:前景帧与背景帧的合成码率, ROI:前景帧码率)				
视频 1 ROI:1400	$\langle \text{NROI}:1187, \text{Cmp}:1955 \rangle$	$\langle \text{NROI}:1187, \text{Cmp}:1662 \rangle$	$\langle \text{NROI}:1187, \text{Cmp}:1271 \rangle$	$\langle \text{NROI}:1187, \text{Cmp}:685 \rangle$
	$\langle \text{NROI}:907, \text{Cmp}:1760 \rangle$	$\langle \text{NROI}:907, \text{Cmp}:1466 \rangle$	$\langle \text{NROI}:1145 \rangle$	$\langle \text{NROI}:907, \text{Cmp}:616 \rangle$
	$\langle \text{NROI}:488, \text{Cmp}:1466 \rangle$	$\langle \text{NROI}:488, \text{Cmp}:1245 \rangle$	$\langle \text{NROI}:488, \text{Cmp}:955 \rangle$	$\langle \text{NROI}:488, \text{Cmp}:515 \rangle$
视频 2 ROI:1500	$\langle \text{NROI}:1248, \text{Cmp}:2055 \rangle$	$\langle \text{NROI}:1248, \text{Cmp}:1746 \rangle$	$\langle \text{NROI}:1248, \text{Cmp}:1335 \rangle$	$\langle \text{NROI}:1248, \text{Cmp}:720 \rangle$
	$\langle \text{NROI}:955, \text{Cmp}:1850 \rangle$	$\langle \text{NROI}:955, \text{Cmp}:1572 \rangle$	$\langle \text{NROI}:955, \text{Cmp}:1200 \rangle$	$\langle \text{NROI}:855, \text{Cmp}:650 \rangle$
	$\langle \text{NROI}:515, \text{Cmp}:1540 \rangle$	$\langle \text{NROI}:515, \text{Cmp}:1310 \rangle$	$\langle \text{NROI}:515, \text{Cmp}:1000 \rangle$	$\langle \text{NROI}:515, \text{Cmp}:540 \rangle$

当用户在 WiFi 网络条件下,由于此时网络条件较好,有足够的带宽来支持高质量视频的传输,因此选择 PI 值最小的配置(视频 1 编码配置  $\langle \text{NROI}:488, \text{Cmp}:1466 \rangle$ , 视频 2 编码配置  $\langle \text{NROI}:1248, \text{Cmp}:2055 \rangle$ )。其中 PI 值用来量化视频的感知质量, PI 值越小, 用户观看体验越好。

在良好的 4G 环境中,当前网络状态一般,故需要综合考虑 MAR 延迟和用户观看感受,去除相对较高的 PI 值后选择  $PI * \text{视频大小}$  最小的配置为最优配置(视频 1 选择配置  $\langle \text{NROI}:488, \text{Cmp}:955 \rangle$ , 视频 2 选择配置  $\langle \text{NROI}:955, \text{Cmp}:1200 \rangle$ ), 该配置可以在保证用户观看感受的前提下,令 MAR

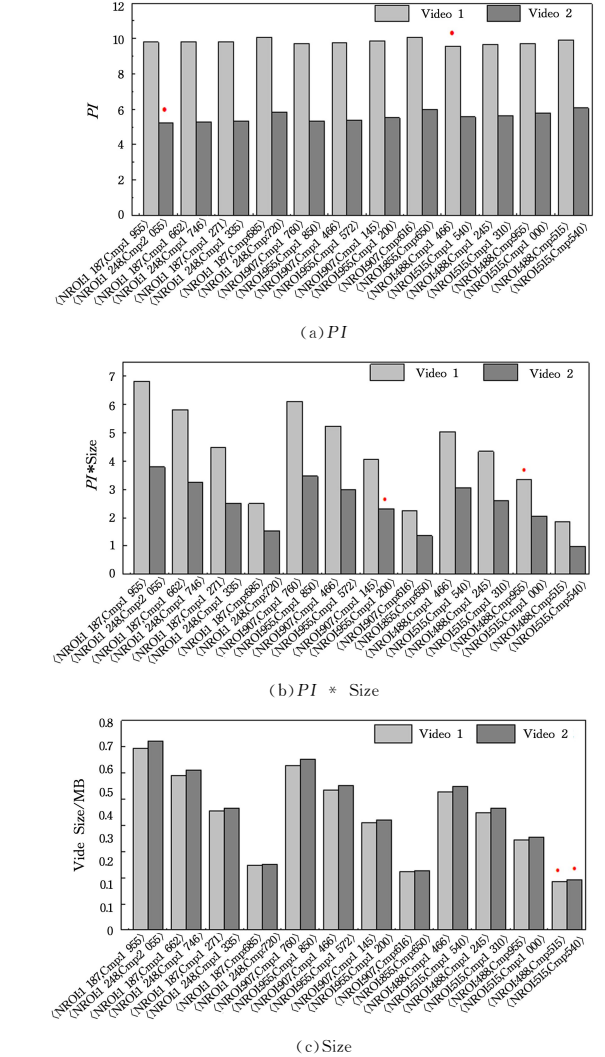


图 7 3 种不同的度量指标

Fig. 7 Three evaluation metrics for videos with different encoding configurations

的延迟最低。

当网络为普通 4G 时,由于网络条件较差,高质量的视频传输会带来较高的延迟,影响 MAR 的使用体验,因此应尽量保证 MAR 响应延迟最小,故选择视频大小最小的配置作为最优配置(视频 1 选择配置  $\langle \text{NROI}:488, \text{Cmp}:515 \rangle$ , 视频 2 选择配置  $\langle \text{NROI}:515, \text{Cmp}:540 \rangle$ ), 此时 MAR 传输延迟最低。

相比直接传输 2900 kbps 码率的原视频,处理后的视频 1、视频 2 平均节约时间 65% 和 57%。如图 8 所示,在 WiFi 情况下,视频 1、视频 2 的 PI 分别为 9.57 和 5.25, 同原视频相比,使用最优码率处理的视频传输时延分别优化了 48% 和

32%;在良好4G的情况下,视频1、视频2的 $PI$ 分别为9.72和5.52,时延分别优化了66%和60%;在普通4G的情况下,此时视频1和视频2的 $PI$ 分别是9.92和6.00,而时延优化达到了81%和80%。

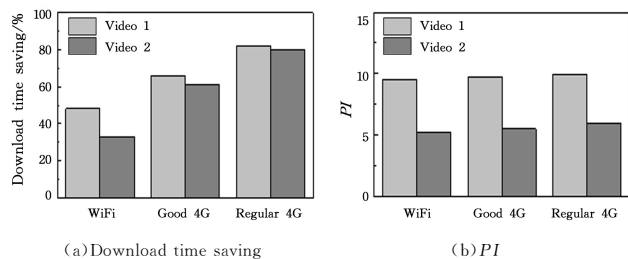


图8 3种网络下视频下载时间节约比例和 $PI$

Fig. 8 Download time saving and  $PI$  when download 2 videos in 3 network conditions

上述结果表明,可以通过调节码率来降低视频的体积,从而降低下行链路的传输时延,且不同的视频拥有不同的最优码率配置。相较而言,简单视频由于图像信息简单,可以分配更低的码率进行压缩,而对于复杂视频来说,由于视频信息复杂运动较多,需要分配更多的码率来维持观看体验。因此需要一种高效的码率配置策略动态地调节背景码率和合成码率,以满足不同视频的最优码率配置需求。

## 6 自适应码率控制模型

图9为本文整体的算法框架图,对视频进行语义分割,使用本节描述的自适应码率控制模型对不同部分的视频进行码率控制。

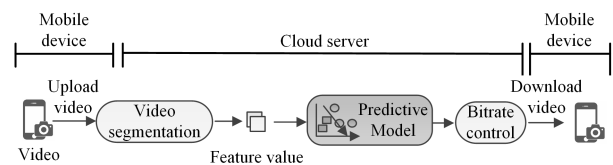


图9 本文方法的整体架构

Fig. 9 Framework of the approach

本文针对MAR在不同网络环境下的视频内容码率配置需求提出一种自适应码率控制模型,其工作流程如图10所示。该模型通过分析拍摄内容特性(如时间复杂度、空间复杂性等),并基于当前网络环境,预测满足时延及用户观感体验需求的码率配置。将其同SVM、随机森林、决策树、神经网络等方法进行对比,选择预测精度最高的GBR方法进行建模,具体如6.3节所述。

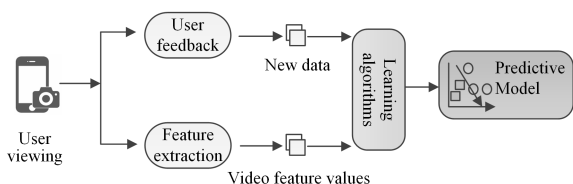


图10 自适应码率控制模型的工作流程

Fig. 10 Workflow of adaptive bitrate control model

### 6.1 特征提取

本文从MAR拍摄内容中共抽取了27个特征,如表3所列。具体来说,实验通过提取视频关键帧中的灰度共生矩阵

分别获取了用于描述视频的纹理信息特征,使用 $TI$ 特征和 $SI$ 特征说明视频的复杂程度,再结合原视频码率这一特征,共同构成了本文模型的特征部分,共计27个特征。

表3 提取的原始特征

Table 3 Extracted raw features		
灰度共生矩阵	对比度、差异性、逆差分矩、能量、相关性、熵 (生成角度为 $0^\circ, 45^\circ, 90^\circ, 135^\circ$ )	
视频复杂度	$SI$	$TI$
其他	原视频的码率	

灰度共生矩阵是一个通过计算两点灰度之间的特性来反映图像纹理信息的矩阵函数。本文使用 $P(i, j)$ 代表灰度共生矩阵,其中 $P(i, j)$ 定义为从 $i$ 点出发到某个固定位置(距离为 $d$ ,角度为 $\alpha$ )上灰度值为 $j$ 的概率。灰度共生矩阵主要由4个角度生成: $0^\circ$ 代表方向为从左到右; $45^\circ$ 代表方向为从左下到右上; $90^\circ$ 代表方向为从下到上; $135^\circ$ 代表方向为从右下到左上。图像的纹理特征可以从4个角度不同的共生矩阵中得到,灰度共生矩阵中常用的特征如式(7)~式(12)所示。

$$\text{对比度: 对比度越低, 纹理沟纹越浅, 视觉效果越不清晰.} \\ \text{Con} = \sum_i \sum_j (i-j)^2 P(i-j) \quad (7)$$

$$\text{Dis} = \sum_i \sum_j P(i-j)^2 |i-j| \quad (8)$$

逆差分矩:逆差分矩较高时,纹理区域之间缺少变化,局部会变得均匀。

$$\text{Hom} = \sum_i \sum_j \frac{P(i-j)}{1+(i-j)^2} \quad (9)$$

能量:能量越高,纹理模式越均一与规则。

$$\text{ASM} = \sum_i \sum_j P(i-j)^2 \quad (10)$$

相关性:相关性高则相邻像素之间的相关性高。

$$\text{Corr} = \frac{[\sum_i \sum_j ((ij)P(i,j) - \mu_x \mu_y)]}{\sigma_x \sigma_y} \quad (11)$$

其中, $\mu_x$ 和 $\mu_y$ 是像素的均值, $\sigma_x$ 和 $\sigma_y$ 是像素的方差。

熵:熵越低则图像复杂度越低,纹理越粗糙。

$$\text{Ent} = - \sum_i \sum_j P(i,j) \log P(i,j) \quad (12)$$

为了进一步降低处理开销,本文删除强相关特征,通过计算特征间的皮尔逊系数获得特征相关性(皮尔逊系数是用来衡量两个特征之间相关性的指标,该值越靠近1表示两个特征的相关性越强),并删除相关性大于0.75的特征。最终删除皮尔逊系数的强相关性特征,保留了如下6个特征: $0^\circ$ 共生矩阵的能量特征、 $45^\circ$ 共生矩阵的相关性特征、 $135^\circ$ 共生矩阵的对比度特征、原视频码率,以及视频的 $SI$ 和 $TI$ ,这6个特征的皮尔逊相关性热力图如图11所示。

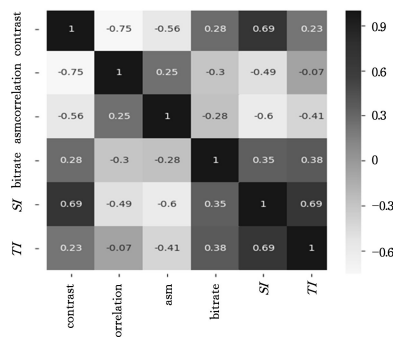


图11 特征之间相关性的热力图

Fig. 11 Feature correlation heatmap

## 6.2 GBR 建模

本文选取 Pexels<sup>1)</sup> 网站中最热门的 200 个视频作为实验数据集,每个视频的时长为 3 s,分辨率为  $1920 \times 1080$ 。图 12 为数据集中视频 *SI* 和 *TI* 的分布情况,数据集包括从简单静态的视频到复杂动态的视频,更多关于数据集及其预处理的内容详见第 7.2 节。

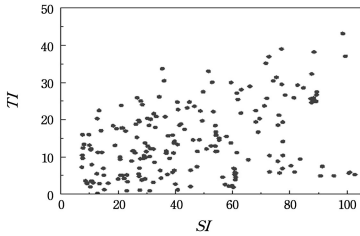


图 12 200 个视频的 *TI* 和 *SI* 分布

Fig. 12 Distribution of *SI* and *TI* of 200 popular videos

进一步,将包含 200 个视频样本的数据集通过不同的码率配置(见表 3)进行处理,并采集所有配置下视频的 *PI* 值(用户观感体验优先)、*PI* \* 视频大小(综合考虑用户观感体验与传输延迟)、视频大小(延迟优先)3 种指标,标记对应的网络状况下视频的最优配置。表 3 中的背景码率是由背景帧自适应生成的码率,3 种合成码率是由恒定的前景码率与不同的背景码率合成时所用的码率,其前景码率是由感兴趣区域即前景帧通过 ffmpeg 自适应生成的。

当网络环境较好时(如 WiFi),MAR 传输延迟较低,因此以 *PI* 值最小为依据选择配置得到最优码率(*PI* 值量化了观看视频的感知质量的好坏);当网络环境状态一般时(如良好 4G),综合考虑视频的观看质量与传输延迟,在 *PI* 指标可接

受的范围内使用 *PI*\* 视频大小的最小值为依据选择最优码率配置;当网络环境较差时(如普通 4G),响应延迟优先级最高,故以视频最小值为判断依据,选择该码率配置为最优配置,此时拥有最低的延迟。在好的网络下偏向于保持视频观看感受以保证用户观看体验;在网络较差时,更偏向于降低延迟、减少能耗,以此来为不同网络下的视频标注标签。

本文通过 3 种不同的网络环境下 MAR 应用的不同需求,对 3 种网络情况分别进行建模,找到特征值与最优配置之间的联系,并输出 GBR 模型。GBR 预测模型工作流程如图 13 所示。

(1)对 200 个视频进行视频特征提取;

(2)遍历执行码率配置,标记每个视频在不同网络下的最优码率配置;

(3)将步骤(1)、步骤(2)的信息组对放入 GBR 回归模型进行学习,并对 200 个视频做 10 折交叉验证;

(4)模型输出合成码率的回归预测结果。

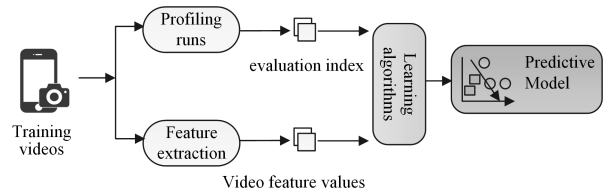


图 13 构建 GBR 回归模型的流程图

Fig. 13 Process of building GBR model

根据预测的合成码率和表 4 的配置,使用算法 2 所示的配置搜索算法,确定该预测的合成码率所匹配的背景码率值,最终得到码率配置并进行视频处理。

表 4 码率配置

Table 4 Bitrate configurations

(单位:kbps)

码率配置(背景码率,合成码率)		
(0.85 * 背景码率,1.00 * 合成码率 1)	(0.65 * 背景码率,1.00 * 合成码率 2)	(0.35 * 背景码率,1.00 * 合成码率 3)
(0.85 * 背景码率,0.85 * 合成码率 1)	(0.65 * 背景码率,0.85 * 合成码率 2)	(0.35 * 背景码率,0.85 * 合成码率 3)
(0.85 * 背景码率,0.65 * 合成码率 1)	(0.65 * 背景码率,0.65 * 合成码率 2)	(0.35 * 背景码率,0.65 * 合成码率 3)
(0.85 * 背景码率,0.35 * 合成码率 1)	(0.65 * 背景码率,0.35 * 合成码率 2)	(0.35 * 背景码率,0.35 * 合成码率 3)

### 算法 2 配置搜索算法

输入:合成码率的预测值(Pre);配置中的合成码率(A[0, ..., 11]);配置中对应的背景码率(B[0, ..., 11])

输出:背景码率

```

1. i=0;
2. min=Pre;
3. # 初始化最小差值
4. res=0;
5. while i < 12 do
6.   if |Pre-A[i]| < min then
7.     min=|Pre-A[i]|;
8.     res=i;
9. end if
10.  i++;
11. end while

```

12. # 遍历找到差值最小的配置

13. return B[res];

14. # 返回对应的背景码率

算法 2 描述的是根据预测到的合成码率在已有配置(即表 4 所列的 12 种配置)中逐一进行搜索。每次让预测值与每种配置中的合成码率进行比较,找到差值最小的配置,从而获得该配置下的背景码率。

### 6.3 预测准确率

本文分别使用了 GBR、SVM、随机森林、决策树和神经网络 5 种方法来进行建模,如图 14 所示,当模型所处网络环境状态一般时(如良好 4G 网络),GBR 方法的准确率为 87%,随机森林的准确率为 85%,神经网络的准确率为 83%,SVM 的准确率为 81%,决策树的准确率为 84%,因此本文选择其中精度最高的 GBR 模型为预测模型。

<sup>1)</sup> <https://www.Pexels.com>

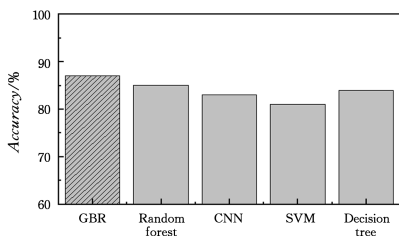


图 14 不同机器学习方法的准确率

Fig. 14 Accuracy of different methods

在 GBR 模型中,损失函数用于估量预测值与真实值的差异,有最小平方误差 LS、最小绝对值偏差 LAD、平均绝对误差 Huber 和分位误差 Quantile 4 种。当预测误差偏小时采用 LS,当预测误差偏大时采用 LAD,而 Huber 结合了二者的优点,以增强模型的鲁棒性,因此在本文的模型中损失函数选择 Huber。然后通过网格搜索方法确定模型的其他参数:模型一共建立 280 个弱学习器即回归树的个数为 280,每棵回归树的最大深度为 20,最小叶子节点为 2。而对于“分裂条件” criterion 则选择 mae。通过实验测试不同学习率对准确率的影响,如图 15 所示,本文选择准确率最高时的学习率(0.04)为模型的学习率。

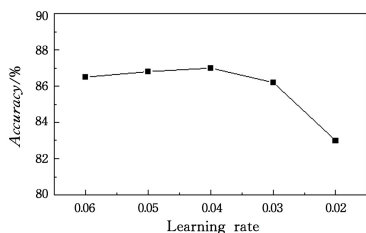


图 15 学习率与准确率的关系

Fig. 15 Relationship between learning rate and accuracy

## 7 实验评估

### 7.1 实验平台

本文使用 NVIDIA Jetson TX2 开发板作为用户端的移动设备,使用搭载了 E5-2683 v4 CPU 和 Nvidia Tesla P40 GPU 的远程服务器作为云服务器端。同时在服务器端搭载 Ubuntu 16.04 系统,使用开源的 ffmpeg 库来进行视频分割与码率控制,使用 Python scikit-learn<sup>1)</sup> 机器学习包构建回归预测模型。

### 7.2 数据集

本文从公开的视频网站 Pexels 中选取了在不同类别(人物、动物、风景等)中点击量最高的视频共 200 个,并对其进行统一的预处理工作,以保证实验的准确性。首先,将所有视频切分为时长为 3s 的视频,保证视频时长的一致性;其次,通过开源的 ffmpeg 工具的 libavformat 库获取视频中的音频流(AVMEDIA\_TYPE\_AUDIO)和视频流(AVMEDIA\_TYPE\_VIDEO),并仅提取视频流为输出流而摒弃音频流,因为音频信息在视频中占有一定比重,需要去除其对实验的影响;最后,对所有视频由 H.264 编码转为 H.265 编码,这是由于 H.265 编码比 H.264 编码更加高效、所占空间更小、更符合

实验高效压缩需求。由此,实验所用的数据集是由 200 个时长为 3s 的无音频信息的 H.265 编码的纯视频流组成,数据集中视频的分辨率为  $1920 \times 1080$ ,视频的码率分布从 100~10000 kbps 不等。

### 7.3 回归模型预测结果

回归预测模型主要是根据当前的网络状况以及输入视频的不同,获取一个合适的合成码率,再通过算法 2 进行配置搜索,得到相应的背景码率,组成最优码率配置,以便生成理想的视频。本文使用准确率指标来评判模型的表现,准确率表示预测结果的准确程度,本文中准确率的计算如式(13)所示:

$$Acc = 1 - \frac{|y_{pre} - y_{true}|}{y_{true}} \quad (13)$$

其中,Acc 表示预测的准确率, $y_{pre}$  表示通过预测模型预测得到的码率值, $y_{true}$  表示标签中码率的真实值。模型使用 10 折交叉验证计算平均准确率,将数据集划分为 10 组,每组包含 20 个视频,每次使用一组作为验证集,其余 9 组作为训练集,重复以上过程 10 次。最终的结果取 10 次准确率的平均值,如图 16 所示,其中白色的点代表平均值,竖线为其中 50% 数据的分布情况。

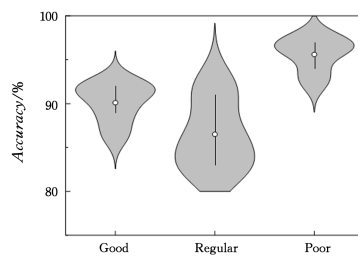


图 16 模型在 3 种网络下的预测准确率

Fig. 16 Accuracy of our model in 3 networking conditions

当网络环境较好时,模型的预测准确率为 90%;当网络环境一般时,预测准确率为 87%;当网络环境较差时,预测准确率为 96%。

### 7.4 性能分析

本文对比了 3 种网络情况下,200 个视频在最优配置下与未作处理时的原始视频,下载过程中的时延优化比例和视频观看质量的情况如图 17 所示。图 17(a)给出了与原始视频相比,最优配置下处理后视频的下载传输时间节约的比例;图 17(b)给出了处理后视频的 PI 值分布的变化情况。

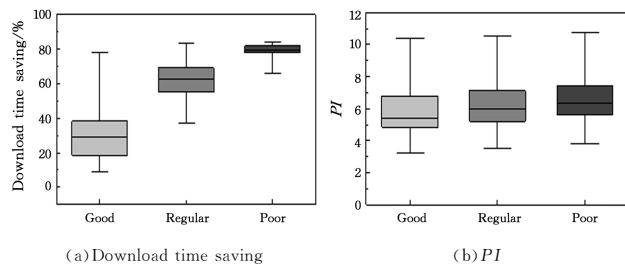


图 17 3 种网络下 200 个视频的下载时间节约比例和 PI 值分布情况

Fig. 17 Download time saving and PI for 200 videos with their optimal configuration under three network conditions

<sup>1)</sup> <https://scikit-learn.org/>

当网络环境较好时,模型的配置拥有最好的视频感知质量,此时 200 个视频的时延平均可以节约 32%(最高可达 77%),同时  $PI$  的平均值为 5.84;在网络环境一般的情况下,综合考量了下载延迟和视频的感知质量,此时时延平均可节约 63%(最高可达 82%), $PI$  平均值是 6.16;在网络环境较差时,有最低的下載延迟,但牺牲了视频的感知质量,此时时延平均可以节约 80%(最高可达 83%),而  $PI$  平均值为 6.55。

由此可知,通过机器学习的方法,对视频进行码率控制之后,平均节约时延可达 58%,有效减少了下载过程中的资源浪费,同时提升了 MAR 的使用性能。

本文为了证明所提方法的先进性,以良好网络状态下的情况为例,以本文实验所用的 200 个视频为对象,将其与 Qian 等<sup>[1]</sup>提出的 Flare 方法的优化情况进行对比,即通过收集 Flare 方法和本文方法的下载时间节约比例以及  $PI$  值的分布情况进行对比,结果如图 18 所示。可以看出 Flare 平均下载时间节约为 40%,但是  $PI$  平均值为 6.78。虽然 Flare 在时延优化效果上略优于本文方法,但是在视频观看体验上远不如本文方法,因此本文方法不仅在时延优化上具有一定的效果,而且保证了用户观感,具有一定的先进性。

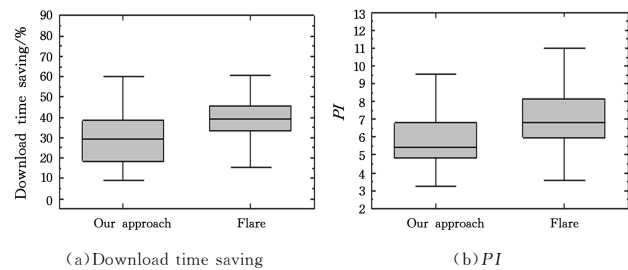


图 18 200 个视频下本文方法与 Flare 方法的下载时间节约比例和  $PI$  值分布情况

Fig. 18 Download time saving and  $PI$  for 200 videos in our approach and Flare

**结束语** 针对需在云端进行渲染操作而产生下载能效问题的 MAR 应用如在线视频会议等,本文提出了基于 GBR 的自适应码率控制模型,该模型通过抽取当前移动设备拍摄内容特征、网络状态及用户观看需求,对用户关注区域及非关注区域进行自适应编码,以保证用户体验的同时降低 MAR 应用的数据传输时延。在移动开发板 Jetson TX2 上进行实验,结果表明,相比直接下载渲染后的高质量视频,本文提出的自适应码率控制模型平均降低了 58%(19.13ms)的传输时延。

## 参考文献

- [1] CHEN H X, DAI Y T, MENG H, et al. Understanding the characteristics of mobile augmented reality applications[C]// 2018 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS). Piscataway NJ: IEEE, 2018: 128-138.
- [2] ZHOU Y Q, SUN B L, QI Y L, et al. Mobile AR/VR in 5G based on convergence of communication and computing[J]. Telecommunications Science, 2018, 34(8): 19-33.
- [3] LIU G Q. Research on offloading strategy based on mobile edge computing[D]. Harbin: Harbin Institute of Technology, 2018.
- [4] SHAUKAT N. Minimization of energy consumption for mobile edge computing-based augmented reality application[D]. Beijing: Beijing University of Post and Telecommunication, 2019.
- [5] GALIH S, IANSYAH S, YUDISTIRA H. Location and orientation based augmented reality mobile application for enhancing heritage landmark history and information[J]. Universal Journal of Electrical and Electronic Engineering, 2019, 6(2A): 39-45.
- [6] KICO I, LIAROKAPIS F. Investigating the learning process of folk dances using mobile augmented reality[J]. Applied Sciences, 2020, 10(2): 599.
- [7] CHEN Y H, RAVINDRANATH L, DENG S, et al. Glimpse: Continuous, real-time object recognition on mobile devices[C]// Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems (Sensys). New York NY: ACM, 2015: 155-168.
- [8] REN J, GAO L, YU J L, et al. Energy-efficient Deep Learning Task Scheduling Strategy for Edge Device[J]. Chinese Journal of Computers, 2020, 43(3): 440-452.
- [9] LIU L Y, LI H Y, GRUTESER M. Edge assisted real-time object detection for mobile augmented reality[C]// Proc of the the 25th Annual International Conference on Mobile Computing and Networking (MobiCom'19). New York NY: ACM, 2019: 1-16.
- [10] ZHANG W X, LIN S K, BIJARBOONEH F H. CloudAR: A Cloud-based Framework for Mobile Augmented Reality[C]// Proceedings of the on Thematic Workshops of ACM Multimedia (Workshops '17). 2017: 194-200.
- [11] QIAN F, HAN B, XIAO Q Y. Flare: practical viewport-adaptive 360-degree video streaming for mobile devices[C]// Proc of the 24th Annual International Conference on Mobile Computing and Networking (MobiCom'18). New York NY: ACM, India, 2018: 99-114.
- [12] LENG Y, CHEN C C, SUN Q Y, et al. Energy-efficient video processing for virtual reality[C]// Proceedings of the 46th International Symposium on Computer Architecture (ISCA). New York NY: ACM, 2019: 91-103.
- [13] RUAN X K, CHEN H L, ZHU X D, et al. Deep decision: a mobile deep learning framework for edge video analytics[C]// INFOCOM, IEEE Conference on Computer Communications (INFOCOM). Piscataway NJ: IEEE, 2018: 1421-1429.
- [14] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network [C]// The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway NJ: IEEE, 2017: 2881-2890.
- [15] LI Y, SHI J, LIN D. Low-latency video semantic segmentation [C]// The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway NJ: IEEE, 2018: 5997-6005.
- [16] GADDE R, JAMPANI V, GEHLER P V. Semantic video cnns through representation warping [C]// The IEEE International Conference on Computer Vision (ICCV). Piscataway NJ: IEEE, 2017: 4453-4462.

- [17] ZHU X, XIONG Y, DAI J, et al. Deep feature flow for video recognition[C]// The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway NJ: IEEE, 2017: 2349-2358.
- [18] LIU N, HAN J. Deep hierarchical saliency network for salient object detection[C]// The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway NJ: IEEE, 2016: 678-686.
- [19] WANG L, LU H, RUAN X, et al. Deep networks for saliency detection via local estimation and global search[C]// The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway NJ: IEEE, 2015: 3183-3192.
- [20] ZHAO R, OUYANG W, LI H, et al. Saliency detection by multi-context deep learning[C]// The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway NJ: IEEE, 2015: 1265-1274.
- [21] DONOSER M, URSCHLER M, HIRZER M, et al. Saliency driven total variation segmentation[C]// The IEEE International Conference on Computer Vision (ICCV). Piscataway NJ: IEEE, 2009: 817-824.
- [22] BAI X, WANG W, LIANG J. An active contour model based on region saliency for image segmentation[J]. Journal of Computer Research and Development, 2012; 49(12): 2686-2695.
- [23] BORJI A, FRINTROP S, SIHITE D L, et al. Adaptive object tracking by learning background context[C]// 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. Piscataway NJ: IEEE, 2012: 23-30.
- [24] GAO Y, WANG M, TAO D, et al. 3-D object retrieval and recognition with hypergraph analysis[J]. IEEE Transactions on Image Processing, 2012, 21(9): 4290-4303.
- [25] REN J, WANG X, FANG J. Proteus: network-aware web browsing on heterogeneous mobile systems[C]// Proc of the 14th International Conference on Emerging Networking Experiments and Technologies. New York NY: ACM, 2018: 379-392.
- [26] REN J, YUAN L, NURMI P, et al. Camel: smart, adaptive energy optimization for mobile web interactions[C]// IEEE Conference on Computer Communications (INFOCOM). Piscataway NJ: IEEE, 2020: 119-128.
- [27] QIN Q, REN J, GAO L, et al. To compress, or not to compress: Characterizing deep learning model compression for embedded inference[C]// 2018 IEEE Intl Symp on Parallel and Distributed Processing with Applications (ISPA). Piscataway NJ: IEEE, 2019: 729-736.
- [28] MARK B, ADRIAN C, GUN L. A survey of augmented reality [M]. Boston, Amercia; Now Publishers Foundations and Trends, 2015: 126-146.
- [29] HOU Q I, CHEN M M, HU X, et al. Deeply supervised salient object detection with short connections [C]// The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway NJ: IEEE, 2017: 3203-3212.
- [30] WAN S, YANG Z F. A new generation of efficient video coding H. 265/HEVC: principle, standard and implementation[M]. Beijing Publishing House of Electronics Industry, 2014: 2-15.
- [31] BT R I. Methodology for the subjective assessment of video quality in multimedia applications[S]. Geneva: The International Telecommunication Union (ITU), 2007.
- [32] SEVCIK L, BEHAN L, FRNDA J, et al. Prediction of subjective video quality based on objective assessment[C]// 2018 26th Telecommunications Forum (TELFOR). Piscataway NJ: IEEE, 2018: 1-4.
- [33] JEROME H F. Greedy Function approximation: a gradient boosting machine[J]. Annals of Statistics, 2001, 29(5): 1189-1232.
- [34] LEO B. Bagging predictors[J]. Machine Learning, 1996(24): 123-140.
- [35] FREUND Y, SCHAPIRE R E. A decision-theoretic generalization of on-line learning and an application to boosting[J]. Journal of Computer and System Sciences, 1997, 55(1): 119-139.
- [36] BREIMAN L, FRIEDMAN J H, OLSHEN R A, et al. Classification and regression trees[J]. International Biometric Society, 1984, 40(3): 874-874.
- [37] BLAU Y, MICHAELI T. The perception-distortion Tradeoff [C]// The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway NJ: IEEE, 2018: 6228-6237.
- [38] ANISH M, RAJIV S, ALAN B. Making a "Completely Blind" Image Quality Analyzer[J]. IEEE Signal Processing Letters, 2013, 20(3): 209-212.
- [39] MA C, YANG C, YANG X. Learning a No-Reference Quality Metric for Single-Image Super-Resolution[J]. Computer Vision and Image Understanding, 2016, 158: 1-16.



**CHEN Le**, born in 1996, postgraduate. Her main research interests include mobile computing and deep learning.



**GAO Ling**, born in 1964, Ph.D, professor, is a member of China Computer Federation. His main research interests include cyber security, network management and embedded system.