

基于 MPC 的音视频同步码率自适应测试

杜宸罡, 李博, 画芊昊

(中北大学 仪器科学与动态测试教育部重点实验室, 太原 030051)

摘要: 随着信息化的发展, 音视频流媒体技术应用面越来越广, 为了使得音视频流媒体技术尤其是在直播方面拥有更好的性能, 得到更多用户的好评, 采用在原本 HTTP 的动态自适应流标准的视频流媒体架构下引入 MPC 控制算法并将 MPC 与码率自适应算法相结合的方法, 进行对 AAC 优化、确定预测模型、测试音视频同步的影响因素以及 PSNR-Y 分量、测试切片时长与跳帧时延, 计算最终的 QoE 用户评价指标来进一步检测音视频流媒体技术的优劣; 经实验仿真测试可知, 相比前人的相关算法, 在不同直播场景下以及不同网络环境下均有更加良好的 QoE 值, 平均 QoE 用户评价指标明显更高, 为 1 237.282 6; 综上分析可知, MPC 的音视频同步码率自适应算法各项性能最好。

关键词: 音视频流媒体; HTTP 协议; 模型预测控制; 码率自适应算法; 音视频同步; QoE 用户评价指标

Code Rate Adaptive Test of Audio and Video Synchronization Based on MPC

DU Chengang, LI Bo, HUA Qianhao

(Key Laboratory of Instrumental Science and Dynamic Testing, Ministry of Education,
North University of China, Taiyuan 030051, China)

Abstract: With the development of information technology, audio and video streaming technology is applied more and more widely. In order to achieve higher performance in audio and video streaming technology and receive more praise from users, especially in live broadcasting, the model predictive control (MPC) control algorithm is introduced under the original hyper text transfer protocol (HTTP) dynamic adaptive streaming standard video streaming media architecture, and the MPC is combined with the code rate adaptive algorithm to optimize the advanced audio coding (AAC), determine the predictive model, test the influencing factors of audio and video synchronization and the Y-component of peak signal to noise ratio (PSNR), and measure the duration and frame hopping delay of the slice. The final evaluation index from users in the quality of experience (QoE) is calculated to further detect the advantages and disadvantages of the audio and video streaming media technology. The experimental simulation results show that compared with previous relevant algorithms, the QoE values are better in different live broadcast scenarios and network environments, and the average evaluation index from users in the QoE is significantly higher, reaching 1 237.282 6. From the above analysis, the code rate adaptive algorithm of the MPC has the best performance.

Keywords: audio and video streaming media; HTTP protocol; MPC; code rate adaptive algorithm; audio and video synchronization; QoE user evaluation indicators

0 引言

随着互联网技术的发展, 高性能的嵌入式流媒体技术应用更加广泛, 具体包括视频直播、远程会议、远程教育 and 视频监控领域等多媒体应用, 教育是一个国家根本的事业, 重要程度不言而喻。

随着我国高速推进“互联网+”在高校领域的推进, 全国应用范围巨大, 改善流媒体的用户体验质量显得愈发关键。码率自适应算法可以根据动态选择进一步改善流媒体的用户体验, 文献 [1] 提出了一种 DASH 标准的基于缓存补偿的码率自适应切换 (BASBC) 算法, 通过提高带宽利用率, 从而提高用户的体验质量 (QoE); 文献 [2] 提出一种基于深度强化学习的码率自适应 (DRLA, deep rein-

forcement learning based ABR) 算法, 用实际网络带宽数据训练神经网络, 通过收集客户端缓冲区占用率和网络吞吐量向视频服务器请求最佳码率的视频, 从而提高了用户的体验质量 (QoE); 文献 [3] 提出了一种名为 NIA (Data-free Network-environmental Imitation-based Rate Adaptation Framework) 的新型无数据蒸馏框架, 生成了具有更好泛化性能的决策树 ABR 算法, NIA 通过网络环境生成模块构建多个人工网络环境, 并在每次迭代训练前使用环境选择模块来选择适合的网络场景, 然后与该场景进行交互, 利用基于学生驱动的模仿学习算法完成决策树的蒸馏过程, 从而提高用户的体验质量 (QoE)^[4]。然而本实验在音视频流媒体直播的应用场景下, MPC (Model Predictive Con-

收稿日期: 2023-10-30; 修回日期: 2023-12-14。

基金项目: 上海交通大学决策咨询课题 (JCXZSJ2020)。

作者简介: 杜宸罡 (1996-), 男, 硕士研究生。

通讯作者: 李博 (1972-), 男, 博士, 副教授。

引用格式: 杜宸罡, 李博, 画芊昊. 基于 MPC 的音视频同步码率自适应测试[J]. 计算机测量与控制, 2024, 32(4): 54-60.

trol)^[5]结合音视频同步算法来提高用户体验质量。通过实验仿真对比分析, 不同网络环境下, 本实验所得的平均 QoE 数据更加良好, 从而使得本实验更有意义。

1 音视频流媒体直播架构以及实现模式

1.1 音视频流媒体直播架构

由于本实验涉及音视频同步技术嵌入式流媒体系统平台开发, 因此, 依托 B/S 架构^[6]搭建流媒体服务器, 优点是无需构建大的网络服务器便可形成流媒体应用平台, 通过编制相应的软件确保系统功能的实现并同时保证系统的正常运行。流媒体服务器采用 HTTP/TCP^[7]来传输控制信息, 用实时传输协议/用户数据报协议 (RTP/UDP) 传输实时数据。

1.2 实现模式

嵌入式流媒体系统主要是构建小型的流媒体服务器在保证流媒体通信质量的前提下, 针对流媒体音视频在互联网上质量较低具体包括视频不清晰、卡顿、音视频不同步等问题, 采用最新标准的 H. 265^[8]视频编码标准进行算法优化、实时网络传输技术确保流媒体通信的实时性, 在音视频同步技术提高流媒体音视频通信质量, 采用嵌入式平台开发针对不同的应用要求进行模块化设计, 通过模型预测控制, 实现更多的应用场景需求。

2 基于 MPC 的音视频同步码率自适应测试

2.1 测试流程与介绍

本文算法分为两部分, 第一部分分为 MPC、音视频同步技术、码率自适应技术以及 QoE 指标模型 4 个板块, 第二部分为算法仿真分析。本文算法具体流程如图 1 所示。

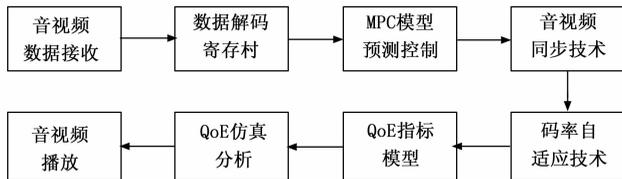


图 1 算法流程

2.2 MPC 的测试

MPC 是一种多变量控制策略, 可以通过控制系统中被控平台所得到的当前系统状态信息, 再加上之后的控制输入变量, 预测到之后的被控平台的状态。预测模型的形式没有固定的标准形式, 状态空间方程、传递函数、阶跃响应模型、脉冲响应模型、模糊模型等都可以作为预测模型的形式。

2.2.1 MPC 对 AAC 的优化测试

为了得到更加良好的数据指标, 将 MPC 应用到 AAC 编码^[9]策略中去, 建立了相对应的目标函数与约束, 具体优化流程如图 2 所示。

具体优化流程为: AAC 编码的状态量和控制量通过被控平台进行状态估计, 利用 MPC 的目标函数与约束以及预测模型得到 AAC 编码的最佳解并传输数据给被控平台, 被控

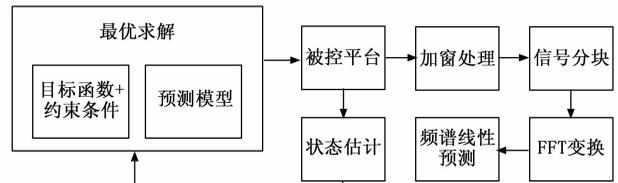


图 2 MPC 对 AAC 的优化流程图

平台再进行音频信号的加窗处理、信号分块^[10]、FFT 变换^[11]、当前频谱线形预测^[12], 分别计算文献 [1-3] 与本文的最大噪声掩蔽功率^[13]、高频部分能量比值^[14]、谱线掩蔽能力^[15]以及感知熵 PE^[16], 如表 1 所示。

表 1 MPC 对 AAC 优化后各参数测试结果

	最大噪声掩蔽功率/dBA	高频部分能量/kHz	谱线掩蔽能力	感知熵 PE/bps
文献[1]	76	10.8	一般	7.6
文献[2]	79	12.9	一般	7.3
文献[3]	87	13.5	一般	6.5
本文算法	102	20.6	较强	4.5

由表 1 可知, 经过 MPC 对 AAC 编码优化改进后, 最大噪声掩蔽功率明显更高, 掩蔽效果更好; 高频部分的能量也明显变高, 高频能量缺失的减少使得音质明显增强, 从而更容易被人耳感知接收, 谱线掩蔽能力相比文献 [1-3] 来说明显提高; 感知熵 PE 的值为 4.5, 介于 3~5 之间, 大大降低了音视频的失真率, 进一步提升了音视频的质量。因此, MPC 预测模型控制对 AAC 编码优化后使得各项指标数据均更加优良。

2.2.2 确定预测模型的测试

往往是通过被控制的对象和即将进行预测的状态来选择合适的预测模型, 对于本文所需的研究而言, 模型预测控制选择状态空间模型^[17]比较合适。状态空间模型是一种以时间为自变量的动态时域模型, 这种状态模型的优点在于不仅能反映系统内部状态, 而且能揭示系统内部状态与外部的输入和输出变量之间的联系; 状态空间模型也可以将多个变量时间序列处理为向量时间序列, 这种从变量到向量的转变更适合解决多输入输出变量情况下的建模问题; 状态空间模型还无需大量的历史资料, 能够用现在和过去的最小信息形式描述系统的状态, 因此既省时又省力。

建立状态空间方程式:

$$\frac{dx}{dt} = Ax + Bu \quad (1)$$

$$Y = Cx \quad (2)$$

其中: u 为输入变量, x 为状态变量。

假设在 t 时刻预测 $t+i$ 时刻的输入值为 $u(t+ilt)$, 在预测区间 W 内。

$$u(t) = \begin{pmatrix} u(t | t) \\ u(t+1 | t) \\ \vdots \\ u(t+W-1 | t) \end{pmatrix} \quad (3)$$

在 t 时刻预测第 $t+i$ 时刻的系统状态为 $x(t+ilt)$ 。

$$x(t) = \begin{pmatrix} x(t | t) \\ x(t+1 | t) \\ \vdots \\ x(t+W | t) \end{pmatrix} \quad (4)$$

MPC 预测控制的优化采用的是滚动式有限时域，凭借每时每刻的优化性能指标，得到每时每刻的有限时段的最优控制率，优化过程也不是一次离线完成的，而是反复在线完成，优化性能指标每一采样点只涉及这一时刻到未来有限时间内的下一个采样时刻点。预测控制求解属于一种开环优化，在反馈校正环节，不断根据系统的实际输出对预测输出做出校正利用反馈信息，让滚动优化不仅停留在模型，更构成闭环优化控制。据状态空间方程式进行进一步滚动优化及反馈矫正，假设预测时域为 Nf ，控制时域为 Nv ， $Nf \geq Nv$ ，则：

$$G(t+1) = \alpha G(t) + \beta u(t | t) \quad (5)$$

$$G(t+2) = \alpha^2 G(t+1) + \alpha\beta G(t | t) + \beta G(t+1 | t) \quad (6)$$

$$G(t+3) = \alpha^3 G(t) + \alpha^2\beta G(t | t) + \alpha\beta G(t+1 | t) + \beta G(t+2 | t) \quad (7)$$

⋮

依次类推：

$$G(t+N_f) = \alpha^{N_f} G(t) + \alpha^{N_f-1}\beta G(t | t) + \alpha^{N_f-2}\beta G(t+1 | t) + \dots + \beta G(t+N_f-1) \quad (8)$$

$$G(t+N_v) = \alpha^{N_v} G(t) + \alpha^{N_v-1}\beta G(t | t) + \alpha^{N_v-2}\beta G(t+1 | t) + \dots + \beta G(t+N_v-1) \quad (9)$$

在 Simulink 和 Carsim 平台上搭建仿真，实现预测控制优化，MPC-AAC 在实现仿真测试过程中，通过 AAC 的预测时域 Nf 以及控制时域 Nv 进行迭代从而得到最少迭代次数和个体最优值。迭代曲线如图 3 所示。

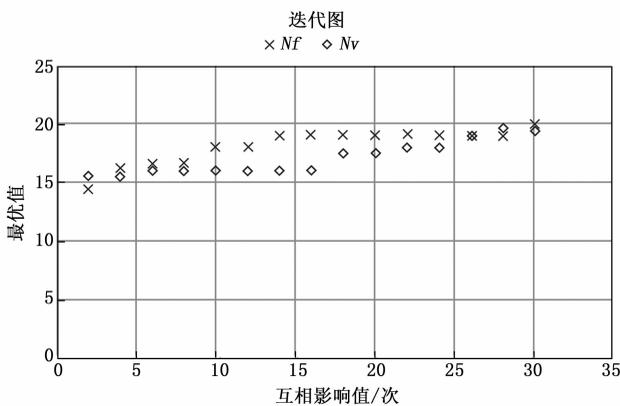


图 3 迭代曲线

通过图 3 可知，AAC 的预测时域 Nf 和控制时域 Nv 在 MPC 迭代过程中，预测时域 Nf 和控制时域 Nv 均取 26 时有最优解，最佳解为 19。

2.3 音视频同步技术的测试

流媒体通信质量好坏的一个重要标准之一就是音视频是否同步。音视频同步^[18]是多媒体系统服务质量 (QoS

Quality of Service) 研究中的一项重要内容。音视频同步可以保证音视频文件的时间连续性，在视频会议、可视电话、视频点播等多媒体应用中，由于网络传输中的延迟、抖动、时间偏差、网络传输条件的变化以及发送端的发送速度与接收端的接收速度的不匹配等问题，使得接收端的媒体存在异步现象。通过媒体同步技术可以解决这些问题，其中音视频的同步是一个关键性技术。

由于处理器对音视频编码的差异，对视频处理的时间相比音频较长，针对音视频分开采集之间没有关联性，直接传输会造成视频滞后于音频的问题，对此本文进行了优化，采取动态缓存技术在服务器端分别设置音视频动态缓存区，传输时进行同步控制保证音视频交替传输，再从服务器输出时打上时间戳建立关联性，在接收端设置动态缓存区，同时解析音视频码流，通过对比时间戳实现音视频同步。为了更好地应对复杂网络环境下确保音视频同步效果，另外建立通信反馈机制，实现服务器端自动通过降低视频码率、视频分辨率、调整动态缓冲区等措施确保音视频同步效果。

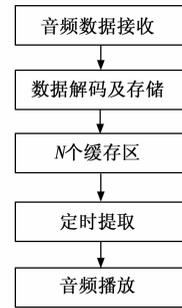


图 4 接收端对音频的处理框图

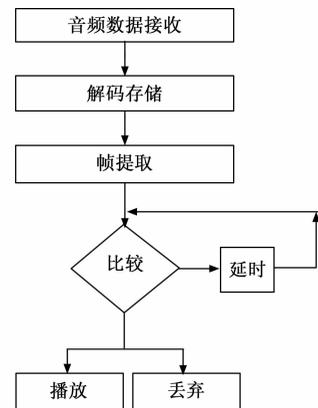


图 5 接收端对视频的处理框图

2.3.1 检测影响因素

时延抖动^[19]和时延偏移^[20]是影响音视频是否同步的重要因素，表 2 通过与文献 [1-3] 对比测试分析了音视频文件、压缩视频文件、数据文本的时延抖动情况；表 3 通过与文献 [1-3] 对比测试分析了视频文件、图像文件、紧密型音频文件、宽松型音频文件、文本文件的时延偏移情况。

表 2 时延抖动测试数据结果 ms

时延抖动	音视频文件	压缩视频文件	数据文本
文献[1]	12	1.2	30
文献[2]	11	1.1	25
文献[3]	10	1.0	20
本文算法	4	0.5	6

表 3 时延偏移测试数据结果 ms

时延偏移	视频文件	图像文件	紧密型音频文件	宽松型音频文件	文本文件
文献[1]	116	243	97	530	237
文献[2]	102	240	94	445	226
文献[3]	95	234	92	428	221
本文算法	22	85	15	107	82

由表 2 和表 3 测试结果可知, 经过音视频同步方案的优化, 相比文献 [1-3] 的方案, 在时延抖动和时延偏移两方面均在数值上得到了大幅度的降低, 从而使得音视频同步效果得到了质的飞跃。

2.3.2 音视频同步压缩的测试

音视频同步压缩^[21]是基于 H.264 或者 AVC 压缩算法利用 Zig-zag 扫描^[22]方法实现的, 系统采集到音视频数据后经过压缩处理传输到服务器, 由于这种技术本身具有先低频系数后跟随高频系数的特点且人们对听觉比视觉相对更加敏感, 因此在原有的音视频同步技术基础上, 可以用 DCT 变化系数^[23]与音频相匹配。

设码率流为 W , 拖尾系数为 H , 解码时最后一位拖尾系数为 Y , 解析音频流为 Z 。在嵌入编码时, 当 $W=1$ 时, $H=Y=1$; 当 $W=0$ 时, $H=Y=-1$; 在解码时, 当 $H \neq 0$ 时, $H=Y=1$ 时, $Z=1$; 当 $H=Y=-1$ 时, $Z=0$; 当 $H=0$ 时, 当 $H=Y>0$ 时, $Z=1$; 当 $H=Y<0$ 时, $Z=0$ 。

由于在图像压缩过程中有亮度块需要嵌入, 需要引入音视频图像峰值信噪比 $PSNR$ ^[26]。

$$PSNR = 10 \log_{10} \frac{255^2}{MSE} \text{ (dB)} \quad (10)$$

其中: MSE 为压缩图像和之前图像的误差, 图 6 为峰值信噪比 $PSNR$ 的 Y 分量测量对比分析图, 纵坐标单位为 dB , 横坐标单位为个数。

根据图 6 原视频、文献 [1-3]、本文嵌入后音频的 20 帧对比测试结果可知, 文献 [1-3] 的 20 帧 $PSNR$ - Y 分量数据均介于原视频和本文嵌入后音频的数据。因此, 本文嵌入后视频图像 $PSNR$ - Y 分量影响最小, 即本文算法嵌入后视频的同步压缩过程最为良好。

2.4 码率自适应测试

码率自适应测试^[24] (ABR, adaptive bitrate streaming) 可以分为 3 种类型。Rate-based 类型是通过历史分片下载时的网络情况来预测之后的网络带宽, 从而驱动视频码率决策, 这种类型相对简单, 但一般单个噪音的干扰信号非常强波动也非常明显, 探测数据既会浪费带宽, 历史数据预判性也非常低; Buffer-based 类型是根据客户端的播放缓冲

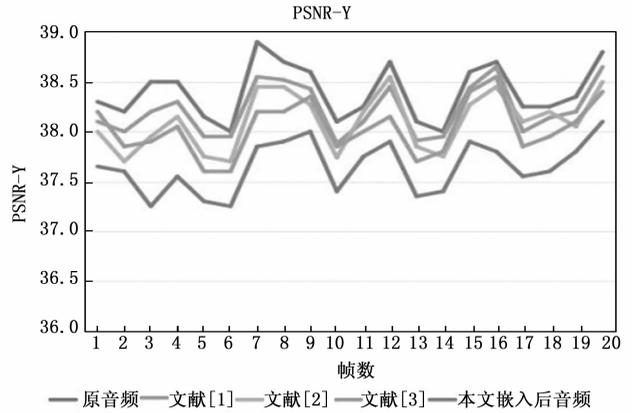


图 6 $PSNR$ - Y 测试结果图

区的 buffer 情况决定下一片段的码率档位, 这种类型的缺点是一般在低延迟环境下不适合更多数据量的缓存, buffer 量的大小也非常难控制, 而且不适合 UDP 环境; Hybrid 类型属于一种混合模式, 既考虑 buffer 的信息决策下一片段的码率档位, 也考虑预测吞吐量情况, 本文的算法即是这种混合模式类型, 在原本 HTTP 的动态自适应流标准的视频流媒体架构下引入了 MPC 控制算法。

2.4.1 切片约束条件的对比测试

设 x 为第 x 个视频切片编码和处理, $S_i(x_i)$ 为第 x 个视频切片所对应 x_i 码率的视频大小, t_x 为视频编码和切片处理前所下载信息的吞吐量, T_x 为第 x 个视频被下载所用的时间, 服务器所需等待时间为 W_i , 视频切片时长为 l 。当 x 为 0 时数据开始采集、编码和切片处理, 切片约束条件如下:

$$T_x = \frac{S_i(x_i)}{t_x} \quad (11)$$

当第 $x+1$ 个视频切片开始下载时,

$$x+1(1) \leq T_x^* \quad (12)$$

$$W_i = 0 \quad (13)$$

当第 $x+1$ 个视频切片还没开始下载时,

$$x+1(1) > T_x^* \quad (14)$$

对于直播场景而言, 视频切片所需时长 l 越短, 视频更稳定更流畅, 表 4 对比了文献 [1-3] 以及本文文献视频切片 l 所需时长。

表 4 切片时长测试结果 s

切片时长	第一次测试	第二次测试	第三次测试	平均时长
文献[1]	9.2	8.5	9.3	9.0
文献[2]	8.3	8.6	8.6	8.5
文献[3]	7.8	7.3	7.4	7.5
本文算法	3.2	2.7	4.3	3.4

通过表 4 可知, 与文献 [1-3] 相比, 本文在原本架构下引入 MPC 后使得切片时长明显变短, 进一步提升了视频流畅度、清晰度和稳定度。

2.4.2 跳帧时延的对比测试

为了进一步降低时延执行跳帧操作，设定了两个两组缓冲区阈值，一组为 (T 最小 1, T 目标 1, T 最大 1) 和 (T 最小 2, T 目标 2, T 最大 2)，一组为 (T 最小 3, T 目标 3, T 最大 3) 和 (T 最小 4, T 目标 4, T 最大 4)，其中满足 $T_{\text{最小}1} < T_{\text{最小}2} < T_{\text{最小}3} < T_{\text{最小}4} < T_{\text{目标}1} < T_{\text{目标}2} < T_{\text{目标}3} < T_{\text{目标}4} < T_{\text{最大}1} < T_{\text{最大}2} < T_{\text{最大}3} < T_{\text{最大}4}$ ，决策变量为 $H_x = (1, 2, 3, 4)$ ，即第 x 个下载视频的目标缓冲水平。

表 5 为分别测试对比了第一组和第二组缓冲区阈值的帧率以及码率情况。

表 5 缓冲区阈值帧率及码率测试结果

缓冲区阈值分组和各组编号		帧率/fps	码率/Mbit
第一组	缓冲区阈值 1	23	0.5
	缓冲区阈值 2	26	0.8
第二组	缓冲区阈值 3	28	0.9
	缓冲区阈值 4	30	1.2

由表 5 可知，当缓冲区阈值为 1 时，帧率和码率均最低，即播放速度相对最慢，视频会产生严重卡顿情况；当缓冲区阈值为 4 时，播放速度相对较快，视频最为清晰流畅。

2.4.3 视频跳帧的对比测试

视频跳帧^[25]又叫视频丢帧，当显示器显示的次数和刷新的次数不一致时，显示器刷新率达不到视频 FPS 的要求，往往会自动舍弃一部分画质，这种现象被称为视频跳帧。为了进一步优化流媒体视频的画质，解决跳帧现象，假定显示器刷新率均为 90 Hz 且 CPU 内存相同情况下，对比原始视频和降低视频分辨率后经过选取更为合适的跳帧阈值后视频的情况，进一步测试十组播放器缓存时间，测试结果如图 7 所示。

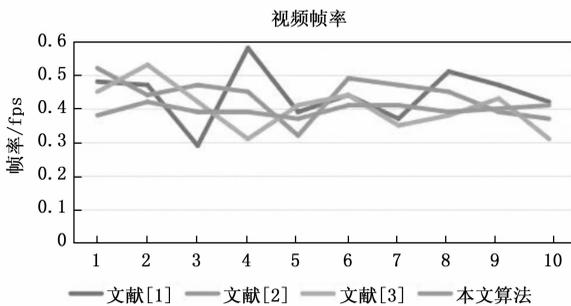


图 7 视频跳帧测试结果图

由图 7 可知，本文算法通过降低视频分辨率选取合适的跳帧阈值后视频的缓存时间曲线明显更趋近于稳定，既不会出现跳帧情况，也不会影响视频的实时性，视频效果最好。

2.4.4 QoE 指标模型的测试

QoE 指标可以很好地反映该种码率自适应算法的优劣性，影响 QoE 指标的因素有很多，比如端与端之间的时延、

卡顿以及跳帧时长、码率切换频次、视频质量等等。

QoE 目标函数为：

$$QoE_{\text{目标}} = \sum_{x=1}^X (\sigma_a f(A_x) - \sigma_b D_x - \sigma_c C_x - \sigma_d B_x) - \sum_{x=1}^{X-1} \sigma_e |A_{x+1} - A_x| \quad (15)$$

其中： $QoE_{\text{目标}}$ 为全部 X 个视频切片的 QoE 之和， A_x 为第 x 个视频切片的码率， B_x 为 x 个视频切片跳帧时间总长， C_x 为视频切片端与端之间的时延， D_x 为卡顿时间总长， $f(A_x)$ 为判断视频好坏的函数， $|A_{x+1} - A_x|$ 为视频切片码率^[26-27]的幅度值， $\sigma_a, \sigma_b, \sigma_c, \sigma_d, \sigma_e$ 分别为各参数指标权重。

QoE 指标的好坏可以充分反应出用户对服务质量的评价，尽可能地解决跳帧、时延、卡顿等情况。

3 QoE 指标仿真结果

本实验在音视频流媒体直播竞赛的 ACM Multimedia 的仿真开发环境下完成，利用公式 (15) 可以计算出 QoE 指标的大小，再求出其平均值。为了很好地验证本文算法的性能，与文献 [1-3] 进行各性能上的仿真测试与比较。

3.1 几种算法的平均 QoE 的 CDF 曲线

本文在原有的码率自适应算法基础上引入了 MPC 与音视频同步算法，通过前文的测试发现，本文算法各项相关数据均得到了大幅度的提升，文献 [1-3] 与本文算法的平均 QoE 的 CDF 曲线如图 8 所示。

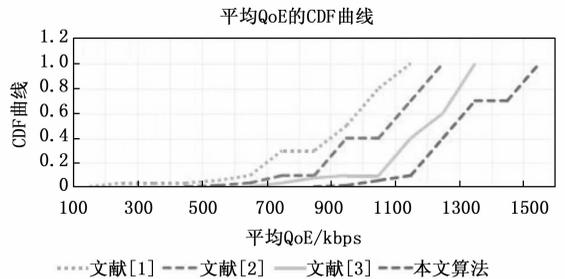


图 8 平均 QoE 的 CDF 曲线

通过本文算法与文献 [1-3] 的平均 QoE 的 CDF 曲线的测试对比分析可知，本文算法的 QoE 指标明显更高，即本文算法的综合性能更好。

3.2 不同直播场下的几种算法的平均 QoE 指标的测试

网红带货、才艺表演、体育比赛、风景打卡是当今互联网直播环境中最受欢迎的其中几个板块，因此分别用文献 [1-3] 以及本文算法对这几个板块分别进行 QoE 指标仿真测试，测试结果如图 9 所示。

图 9 中，曲线从上到下依次为本文算法、文献 [1-3]，由图中测试结果可知，网络直播环境下在网络带货、才艺表演、体育比赛、风景打卡等方面，本文所采用的算法的平均 QoE 指标最大，即用户评价指标最高，因此本文算法相对更加优良。

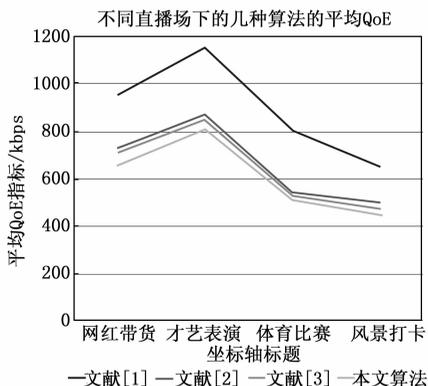


图 9 不同直播场下的几种算法的平均 QoE

3.3 不同网络环境下几种算法的平均 QoE 指标的测试

直播环境中是否卡顿是影响直播感受非常重要的因素之一, 因此对视频卡顿、微卡顿、适中、流畅 4 种情况分别对比文献 [1-3] 以及本文算法进行仿真测试分析, 测试结果如图 10 所示。

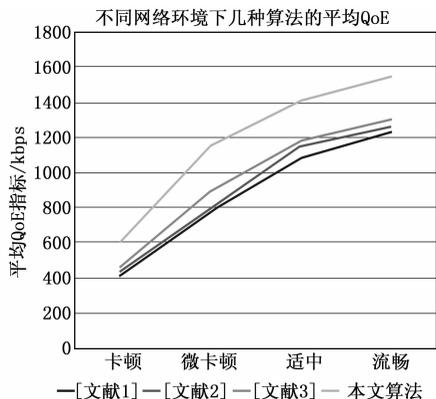


图 10 不同网络环境下几种算法的平均 QoE

图 10 中, 曲线从上到下依次为本文算法、文献 [1-3], 根据图中测试结果可知, 对于卡顿、微卡顿、适中、流畅情况下, 本文算法的 QoE 指标明显比文献 [1-3] 的 QoE 指标高很多, 意味着本文算法使得不同网络环境下的 QoE 指标都得到了很大的提高。

3.4 多种算法每种指标的比较测试

视频质量、卡顿时长、平均时延均严重影响着直播环境中 QoE 指标的大小, 分别用文献 [1-3]、本文算法对这几个方面进行测试对比分析, 数据结果如表 6 所示。

表 6 平均 QoE 指标测试结果

算法	平均 QoE	视频质量/kbps	卡顿时长/s	平均时延/s
文献[1]	954.276 3	515.353 6	95.157 3	0.803 7
文献[2]	987.326 5	520.276 3	87.902 3	0.771 2
文献[3]	996.135 7	527.838 5	81.258 4	0.764 5
本文算法	1237.282 6	897.371 3	41.171 7	0.448 2

由表 6 可知, 本文算法很好地降低了时延与卡顿时长,

大幅度优化了视频质量, 极大地提升了平均 QoE 用户评价指标值的大小。

4 结束语

本文在原本码率自适应算法基础上引入了 MPC 与音视频同步技术并与之相结合。在 HTTP 的动态自适应流标准的视频流媒体架构下, 首先通过对 AAC 编码的优化, 将 MPC 应用到 AAC 编码中, 从而得到 AAC 编码的最佳解。

经过被控平台后续实验与前人的测试对比分析, 掩蔽能力得到了大幅度提升, 可接收到更多的高频能量被人耳感知, 感知熵 PE 也得到了大幅提升, 大大降低了失真率; 随后进行了预测模型的确定, 并通过预测时域与控制时域进行迭代从而得到最少的迭代次数和个体最优值; 再随后对音视频同步的影响因素进行了探索与仿真测试对比分析, 时延抖动和时延偏移是最大的两个影响因素, 时延抖动分别从音视频文件、压缩视频文件、数据文本 3 个方面与前人的算法进行仿真对比测试, 时延偏移分别从视频文件、图像文件、紧密型音频文件、宽松型音频文件、文本文件这几个方面与前人的算法进行仿真对比测试, 经对比分析, 本文算法在数据上均有了大幅度的优化体现; 再随后利用 zig-zag 扫描方法进行了音视频同步压缩, 将前人的算法与本实验算法分别对 20 帧 PSNR-Y 分量进行对比测试, 由仿真测试图可知本文算法在嵌入后视频压缩效果最佳; 再随后测试切片时长与跳帧时延, 与前人的算法进行测试对比分析, 本文算法具有更低的切片时长、相对较快的播放速度、更为稳定的视频缓存时间; 最后获取最终的 QoE 用户评价指标来进一步检测音视频流媒体技术的优劣, 并在音视频流媒体直播竞赛的 ACM Multimedia 的仿真开发环境下分别从 QoE 的 CDF 曲线、不同直播场下平均 QoE 指标、不同网络环境下的平均 QoE 指标、各指标对比的平均 QoE 四个方面与前人算法进行仿真测试对比分析, 测试结果表明本文算法相较而言具有最高的平均 QoE 指标, 即本文算法各项性能最好。

在当今互联网高速发展的时代, 音视频流媒体技术具有无限的发展潜力, 特别是本实验着重侧重于音视频流媒体直播的应用场景, 网红带货、才艺表演、体育比赛、风景打卡等等应用场景, 均与我们的工作和生活息息相关, 非常需要我们深入研究与探索, 除此之外, 腾讯会议网络授课、专家作报告等方面即教育行业、事业单位开会也是音视频流媒体直播非常广泛的应用场景, 因此, 视频加密技术显得极其重要, 下一步将本文中的算法尝试应用到更多领域进行仿真测试分析进一步验证该算法的优越性以及进一步试图探索更加优良的音视频流媒体算法来获取更高的 QoE 用户评价价值, 从而得到更大的社会价值。

参考文献:

[1] 吉爱国, 栾云哲. 基于缓存补偿的视频码率自适应算法 [J]. 计算机应用, 2022, 42 (9): 2816-2822.

- [2] 易 令, 李泽平. 基于深度强化学习的码率自适应算法研究 [J]. 电子学报, 2022, 50 (5): 1192-1200.
- [3] 黄天驰, 李朝阳, 张睿霄, 等. 决策树码率自适应算法的无数数据蒸馏框架 [J]. 计算机学报, 2023 (3): 1-18.
- [4] ROLANDO H. Multifactor QoE scores in IoT topologies [J]. Internet of Things, 2023, 24 (27): 36-23.
- [5] ETTORE Z, DONGHUN K, DAVID B, et al. Performance comparison of quadratic, nonlinear, and mixed integer nonlinear MPC formulations and solvers on an air source heat pump hydronic floor heating system [J]. Journal of Building Performance Simulation, 2023, 16 (2): 21-22.
- [6] YUAN H, YOU Y, JINGFU T, et al. Relay protection setting calculation system for nuclear power plant based on B/S architecture and cloud computing [J]. Energies, 2022, 15 (24): 23-45.
- [7] SDG M, RC S, MR G. A protection system against HTTP flood attacks using software defined networking [J]. Journal of Network and Systems Management, 2022, 31 (1): 11-43.
- [8] WENYING W, RONGXIN T, YUSHU Z, et al. A multi-level approach with visual information for encrypted H.265/HEVC videos [J]. Multimedia Systems, 2022, 29 (3): 13-31.
- [9] THONGTHA A, MANEEWAN S, FAZLIAN A. Enhancing thermal performance of autoclaved aerated concrete (AAC) incorporating sugar sediment waste and recycled AAC with phase change material-coated applications for sustainable energy conservation in building [J]. Sustainability, 2023, 15 (19): 11-52.
- [10] YANGLONG L, YAN W. A physics-constrained dictionary learning approach for compression of vibration signals [J]. Mechanical Systems and Signal Processing, 2021, 153 (1): 136-124.
- [11] ABBAS S P, JASEM J, HOSSEIN M D F, et al. Low complexity super-resolution wideband DOA estimation for LFM signals using FFT dechirp algorithm with a few snapshots [J]. Circuits, Systems, and Signal Processing, 2023, 42 (11): 23-25.
- [12] YUE S, RUI X, CHENXU W, et al. Deep neural network based battery impedance spectrum prediction using only impedance at characteristic frequencies [J]. Journal of Power Sources, 2023, 580 (3): 123-256.
- [13] SERTOL A K, SAITÜ. Multiband absorber design and optimization with lorenz chaotic attractor [J]. Physica Scripta, 2023, 98 (11): 11-12.
- [14] BOBROV V B. On the high-frequency energy distribution of equilibrium radiation in an ideal electron plasma [J]. High Temperature, 2023, 60 (s1): 11-14.
- [15] KOJI H, MINAMI A, AGATA C, et al. Accurate determination of line intensity of H [formula omitted] O near 7 181 cm⁻¹: SI-traceable measurement using primary trace-moisture standard in N [formula omitted] gas [J]. Journal of Quantitative Spectroscopy and Radiative Transfer, 2023, 311 (2): 121-247.
- [16] THOMAS M, MARIO B, VICTOR T. Quantifying risk perception; the entropy decision risk model utility (EDRM-U) [J]. Systems, 2020, 8 (4): 13-14.
- [17] ELISE G B. The neighborhood deprivation gradient and child physical abuse and neglect: a bayesian spatial model [J]. Child Abuse & Neglect, 2023, 146 (2): 221-231.
- [18] THENMOZHI A, KANNAN P. Performance analysis of audio and video synchronization using spreaded code delay measurement technique [J]. ICTACT Journal on Image and Video Processing, 2019, 9 (1): 7-12.
- [19] GABOR L, KEIICHI S, KENJIRO C. Benchmarking methodology for stateful nat64 gateways [J]. Computer Communications, 2023, 210 (3): 123-241.
- [20] FAZAL B, RAHMATULLAH, MUHAMMAD Y Z, et al. Phase time delay caused by quantum effects in nearby plasmonic nanostructures of a one-dimensional photonic crystal [J]. Physica Scripta, 2023, 98 (11): 24-34.
- [21] ZHIWEN G, SHENG F, CHANGCHANG M, et al. Urban river ammonia nitrogen prediction model based on improved whale optimization support vector regression mixed synchronous compression wavelet transform [J]. Chemometrics and Intelligent Laboratory Systems, 2023, 240 (1): 233-242.
- [22] Information Technology-Information and Data Encoding and Encryption. Research data from national institute of technology update understanding of information and data encoding and encryption [J]. Journal of Mathematics, 2020, 135 (1): 113-145.
- [23] JIMSON L P. Video summarization using deep convolutional neural networks and mutual probability-based K-nearest neighbour [J]. Journal of Experimental & Theoretical Artificial Intelligence, 2023, 35 (8): 213-243.
- [24] ZHILIN Z, LONGXINBING C, KUNYU L. Supervision adaptation balancing in-distribution generalization and out-of-distribution detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 352 (1): 212-245.
- [25] MASATOSGI T, AKIHIRO H, KRISTINE B, et al. Bone marrow failure and immunodeficiency associated with human rad50 variants. [J]. Journal of Clinical Immunology, 2023, 23 (2): 13-18.
- [26] SAUMMIT K, KUMAR R G. Link quality improvement analysis of relay-assisted hybrid RF/FSO systems in challenging environments [J]. Optik, 2023, 294 (1): 112-321.
- [27] ABHAYNARASIMHA K, VENKATA V M, MATHINI S. Mitigating nonlinear distortions of high-powered LEDs for VLC using deep neural networks [J]. Optics Communications, 2024, 550 (2): 143-322.