

文章编号: 2095-4980(2020)02-0235-06

基于朴素贝叶斯分类的 DVC-HEVC 快速转码

华玲¹, 唐韬², 卿粼波^{*1}, 何小海¹, 荣松³

(1.四川大学 电子信息学院, 四川 成都 610065; 2.上海交通大学 电子信息与电气工程学院, 上海 200240;

3.北京百度网讯科技有限公司, 北京 100085)

摘要: 分布式视频编码(DVC)与传统视频编码之间的转码为移动终端设备之间的低功耗视频通信提供了一种有效的实现思路。以DVC与HEVC转码为研究对象, 利用DVC解码端信息, 针对高效视频编码(HEVC)中复杂度极高的编码单元(CU)划分过程进行复杂度优化研究。在DVC解码端提取与CU划分相关的纹理复杂度、运动矢量及预测残差3种特征信息; 在HEVC编码端基于朴素贝叶斯原理建立CU快速划分模型, 模型生成后便可以通过输入特征信息对当前CU划分进行快速决策, 避免大量率失真(RD)代价计算过程。实验结果表明, 本方案在编码比特率略有上升的情况下大幅缩短了HEVC编码时间, 平均下降幅度达到58.26%, 且几乎不影响视频质量。

关键词: 分布式视频编码; 高效视频编码; 朴素贝叶斯; 分类模型; CU快速划分

中图分类号: TN919.81

文献标志码: A

doi: 10.11805/TKYDA2018341

Fast transcoding of DVC-HEVC based on naive Bayes classification

HUA Ling¹, TANG Tao², QING Linbo^{*1}, HE Xiaohai¹, RONG Song³

(1.School of Electronic Information Engineering, Sichuan University, Chengdu Sichuan 610065, China;

2.School of Electronic Information and Electrical Engineering, Shanghai Jiaotong University, Shanghai 200240, China;

3.BIDU Co. Ltd, Beijing 100085, China)

Abstract: The transcoding between Distributed Video Coding(DVC) and traditional video provides an effective way to realize low-power video communication between mobile terminal devices. In this paper, DVC and HEVC(High Efficiency Video Coding) transcoding is taken as the research object, and the decoding end information of DVC is utilized to carry out complexity optimization research on the division process of coding unit with high complexity in HEVC. Firstly, texture complexity, motion vector and prediction residuals related to Coding Unit(CU) partition are extracted from DVC decoder. Then, based on naive Bayes principle, a CU fast partition model is established at the HEVC coding end. After the model is generated, current CU division can be quickly determined by inputting feature information, avoiding a cost calculation process of large number of Rate Distortion(RD). Experimental results show that the proposed scheme significantly reduces the HEVC coding time by 58.26% with a slight increase in the encoding bit rate, and hardly affects the quality of video.

Keywords: Distributed Video Coding; High Efficiency Video Coding; naive Bayes; classification model; CU fast partition

随着数字通信、集成电路等技术的飞速发展, 移动终端设备之间的视频通信得到广泛应用, 同时催生了如无线监控视频、视频会议、远程现场指挥等多种新型视频通信服务。然而, 针对诸如山区灾害或军事应用等特定环境, 其设备功耗、存储容量受限, 因而在移动设备之间实现低功耗的视频通信具有较大的应用意义。高效视频编码(HEVC)基于传统的预测加变换的混合编码框架^[1], 采取多种优化处理, 在保证相同视频质量的前提下提高一倍左右的压缩率^[2]。然而与其他的传统编码类似, 其编码端复杂度较高, 与低功耗的视频通信应用要求不符。分布式视频编码(DVC)与传统编码方式不同, 它采用独立编码、联合解码的方式, 将复杂度从编码端移至解码端^[3]。

收稿日期: 2019-01-29; 修回日期: 2019-04-11

作者简介: 华玲(1995-), 女, 在读硕士研究生, 主要研究方向为多媒体通信。email:963214586@qq.com

*通信作者: 卿粼波 email:qing_lb@scu.edu.cn

因此,结合上述两者编码方式的特点,通过转码方式可以为移动终端设备之间的低功耗视频通信提供一种有效的实现思路^[4]。但另一方面,转码过程承担了 DVC 解码与 HEVC 编码 2 个高复杂度的处理过程,计算量极其庞大,严重影响视频传输系统的实际工作效率。综上,为避免过高的通信延迟,如何在保证视频质量基本不变的前提下,降低转码过程的计算复杂度尤为重要。

目前,大多转码框架都基于传统视频编码方式之间,如 H.264 向 HEVC 进行转码;WEI Y 等^[5]利用长短时记忆网络(Long Short Term Memory Network, LSTM)方法对 H.264-HEVC 转码过程的编码单元(CU)划分进行预测;YUAN H 等^[6]根据当前 CU 块运动性质是否均匀对 CU、PU 的划分模式选择不同方式进行提前终止;ZHANG D 等^[7]利用 Fisher 判别式对 CU 划分模式进行快速决策,同时提出了一种在线更新策略提高划分精确度。现有的 DVC 向传统视频编码方式转码的研究大多基于 H.263、H.264 及 SVC 等视频编码框架,利用 DVC 解码码流信息来加速传统视频编码过程。CORRALES 等^[8]利用 DVC 解码过程中生成的运动矢量,通过动态运动估计技术优化了运动矢量的搜索范围,降低了 H.264 编码过程中运动估计的复杂度。此外,CORRALES 等^[9]通过利用图形处理器(Graphics Processing Unit, GPU)将转码服务器中复杂的 H.264 编码过程进行并行处理,从而大幅减少转码过程所需时间。MARTINEZ J 等^[10]通过数据挖掘技术,利用 H.264 中宏块编码模式与 DVC 中边信息残差的高度相关性,加速 H.264 中宏块编码模式的选择过程,实现 DVC 到 H.264 的快速转码。尽管 HEVC 复杂的编码结构导致编码端的计算复杂度大幅增加,但随着高清、超高清视频应用愈加广泛进入人们视野,其突出的压缩性能必将使它成为未来新的主流视频压缩编码标准。针对 DVC 到 HEVC 的转码目前则处于研究初步阶段。RONG S^[4]等基于云计算,提出一种低复杂度端到端的视频通信系统,但并未对 DVC 解码端的相关信息进行充分挖掘和高效利用。卿瀚波^[11]等提出一种基于 Fisher 判别式的 DVC-HEVC 转码方法及其转码器,结合机器学习分类算法对测试序列周期性划分生成快速转码模型。

为充分高效利用解码端信息进行 HEVC 编码端加速,本文结合机器学习方法提出一种 DVC-HEVC 快速转码框架,由 DVC 解码及 HEVC 编码两部分组成,利用 DVC 解码过程中获得的相关信息,针对 HEVC 编码中复杂度较高的编码块划分过程,基于朴素贝叶斯分类生成快速划分模型,避免了迭代遍历编码单元进行 RD 代价计算过程,进而对整个转码过程进行加速。

1 基于朴素贝叶斯分类的 DVC-HEVC 快速转码

DVC-HEVC 转码过程包含了 DVC 解码过程以及 HEVC 编码过程(图 1 虚线框),整体复杂度极高,尤其针对 HEVC 编码端,其引入的多种新的编码技术,如基于四叉树的 CU 划分结构、基于自适应搜索的运动矢量预测技术等都在一定程度上提高了编码复杂度。为降低转码过程的复杂度,本文结合机器学习算法,充分利用 DVC 解码端获得的相关信息,对 HEVC 编码端中复杂度极高的基于四叉树的 CU 划分结构进行复杂度优化。

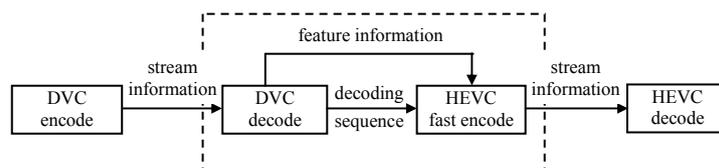


Fig.1 Schematic of DVC-HEVC fast transcoding
图 1 DVC-HEVC 快速转码示意图

1.1 基于四叉树的 HEVC 编码单元划分结构分析

HEVC 中引入了树形编码单元(Coding Tree Unit, CTU)^[12],尺寸为 64×64 ,采用基于四叉树的循环分层结构,可向下划分为 4 个正方形子块,即编码单元 CU, CU 最小尺寸为 8×8 ,如图 2 所示。因此,一个 CTU 可从深度 $d=0$ (CU: 64×64)划分至深度 $d=3$ (CU: 8×8)。

HEVC 定义了上述全新的图像分割模式,目的是更加灵活高效地根据视频内容特性选择编码模式,但也存在着不可忽视的弊端。如图 2 所示,CTU 最优划分方式需要迭代遍历全部有可能的 CU 划分尺寸,且在遍历过程中计算每种 CU 尺寸下的 RD 代价^[1],

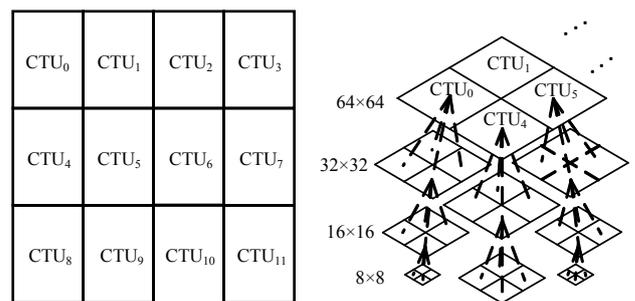


Fig.2 Architecture of CTU partition
图 2 CTU 划分结构

最后取其代价最小的为最优划分模式。整个遍历过程的计算复杂度极高, 占据了编码端复杂度的绝大部分^[12]。

降低 CU 划分复杂度的一种直观思路便是采用复杂度较低的方式来确定 CU 划分方式, 避免大量的 RD 代价计算对比过程, 本文采用基于机器学习的算法来确定 CU 划分。

1.2 基于朴素贝叶斯分类的 CU 快速划分模型

基于机器学习算法提出一种 DVC-HEVC 快速转码方案, 将 HEVC 中复杂的 CU 划分过程转化为可学习的建模分类问题, 利用模型对 CU 划分深度进行快速决策, 避免了大量遍历计算 RD 代价的过程。

具体算法如图 3 所示。首先, 在 DVC 解码端对码流进行解码, 同时提取用于 CU 快速划分模型的相关特征信息一起送入编码端。在 HEVC 编码端将解码重建后的视频序列划分为训练帧与测试帧。训练帧用于生成分类模型, 采用原始 HEVC 编码方式, 同时记录每帧的 CU 划分模式。特征信息作为模型训练的输入, 对应的 CU 划分结果作为模型训练时的监督信息。最后, 基于生成的分类模型, 输入测试帧中提取的相应特征信息, 快速得到 CU 划分结果。

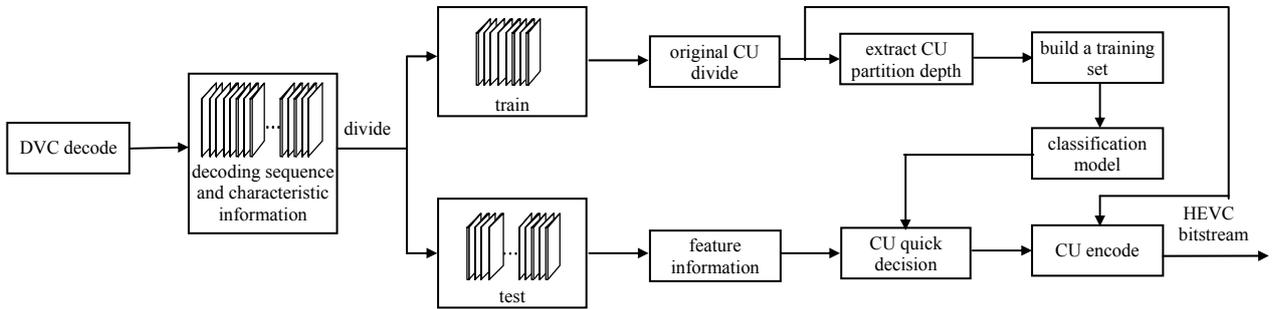


Fig.3 Overall framework of the proposed algorithm
图 3 本文所提算法整体框架

1.2.1 特征信息选取

本文选取运动矢量均值、边信息残差离散度作为 CU 快速划分模型的运动特征, 利用重建帧的纹理复杂度作为纹理特征。在 DVC 解码过程中提取运动矢量信息, 取其均值:

$$V_{ave} = \frac{1}{N} \sum_{i=1}^N \sqrt{V_i(x)^2 + V_i(y)^2} \quad (1)$$

式中: N 为当前 CU 块中所包含 8×8 块的数目; $V_i(x)$ 为 CU 块中第 i 个 8×8 块的水平方向运动矢量; $V_i(y)$ 为 CU 块中第 i 个 8×8 块的垂直方向运动矢量。考虑编码前无法获取残差图像, 利用 DVC 解码端生成的边信息与其重建帧之间的差分图像作为残差图像:

$$Std_{cu} = \sqrt{\frac{\sum_{i=1}^n (P_i - \tilde{P})^2}{n}} \quad (2)$$

式中: n 为当前 CU 中像素总数; $P_i(x,y)$ 为残差图像在当前像素 (x,y) 的差值; \tilde{P} 为残差图像的像素均值。CU 的纹理复杂度计算过程与之类似, 根据该单元内各个像素值与像素均值的绝对误差和来表示其纹理复杂程度:

$$T_{cu} = \sum_{x=0}^{W-1} \sum_{y=0}^{W-1} \left[P(x,y) - \frac{1}{W^2} \left(\sum_{x=0}^{W-1} \sum_{y=0}^{W-1} P(x,y) \right) \right]^2 \quad (3)$$

式中: W 为当前 CU 块的宽度; $P(x,y)$ 为坐标位置 (x,y) 处的像素值。

1.2.2 CU 快速划分模型建立

在 DVC 码流解码过程提取上述选取的特征之后, 基于朴素贝叶斯算法生成 CU 快速划分模型。朴素贝叶斯算法^[12]是一种有监督的分类学习方法, 具有属性条件独立性假设。设有 N 种可能的类别标记, 即 $y = \{C_1, C_2, \dots, C_N\}$, 对每个具有 n 个特征的样本 $x = \{f_1, f_2, \dots, f_n\}$, 选择能使后验概率 $P(C|x)$ 最大的类别作为最优分类类别, 基于属性条件独立性假设, 该分类器描述见式(4)。其中 n 为特征数目, x_i 为 x 在第 i 个特征上的取值。 $P(C)$ 是先验概率, 通过训练集中各类样本出现的频率进行估计。 $P(C|x)$ 是条件概率, 也称为“似然”。

$$H(x) = \arg \max_{C \in y} P(C|x) = \arg \max_{C \in y} P(C) \prod_{i=1}^n P(x_i|C) \quad (4)$$

属性条件独立性假设在实际情况中通常无法满足,但朴素贝叶斯分类仍表现出高精度和高效率^[13],与其他机器学习算法相比,具有算法简单、分类错误率小等特点,在数据挖掘领域具有广泛应用。其简单易行的算法思想也是本文为降低转码复杂度而选择它的主要原因。

首先基于不同的 CU 深度各自建立模型,分别为深度 $d=0(64\times 64)$, $d=1(32\times 32)$, $d=2(16\times 16)$, 仅输入数据不同,模型建立过程相同。如建立深度 $d=0$ 的模型,首先将训练帧划分为 64×64 的单元,输入数据为每一个 64×64 单元的特征信息,该单元是否继续向下划分的标志作为监督信息。对于深度 $d=1$ 的模型,则是将训练帧划分为 32×32 的单元,提取每一个 32×32 单元的特征信息以及划分标志。 $d=2$ 的模型同理。3 个分类模型的类别都为两类,分别为 C_N 及 C_S , 用于判断当前 CU 深度是否继续向下划分,其中 C_S 表示当前 CU 继续向下“划分”, C_N 表示当前 CU “不划分”。快速划分过程具体如图 4 所示。首先,将测试帧划分为初始单元,即 64×64 的尺寸,将该 CU 单元对应的特征信息输入 $d=0(64\times 64)$ 的分类模型,得到针对 2 个类别的后验概率 $P(C_S|x)$ 与 $P(C_N|x)$, 若 $P(C_S|x)$ 大于 $P(C_N|x)$, 将其分为“划分”类别,继续向下输入 $d=1(32\times 32)$ 的分类模型进行划分判别。若当前 CU 特征通过模型后得到的 $P(C_N|x)$ 大于 $P(C_S|x)$, 则将其分为“不划分”类别,取当前深度为 CU 的最优划分方式。本文所提 CU 快速划分过程仍是一种迭代遍历过程,与 HEVC 原始 CU 划分过程一致,不同之处在于本文通过快速划分模型进行 CU 划分决策,减少了大量 RD 代价计算,明显降低了计算复杂度。

2 实验结果与分析

2.1 实验参数设置及性能评价指标

为评估本文所提方案的性能及其有效性,在基于相同的实验配置环境下,分别对本文方案以及文献[9]方案,以传统级联转码方案作为基准,进行性能测试。实验平台配置为:操作系统 Windows7 64 位,处理器 Intel(R) Core(TM) i5-3470 CPU@3.20 GHz,内存 8 GB,开发工具 Microsoft Visual Studio 2010。HEVC 参考软件为 HM16.5。视频序列采用了 HEVC 标准序列 BasketballDrill(832×480)、BQMall(832×480)、FourPeople($1\ 280\times 720$)、Johnny($1\ 280\times 720$),帧率为 30 fps,测试帧数取 100 帧,HEVC 编码量化参数分别取值 22,26,30,34。DVC 编码基于小波变换方式,量化参数设为 15。分别从峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)、比特率(Bit Rate, BR)以及编码时间(T) 3 个性能指标进行对比分析^[5],最后利用 RD 曲线图给出编码性能对比。

2.2 实验结果分析

表 1 为本文所提方案与文献[11]在 PSNR, BR 及 T 三个方面与传统级联转码的对比测试结果。从表中可清晰看出,在时间复杂度方面,两个方案较之级联转码都大幅下降, CU 划分时间平均下降 50% 以上;在编码比特率方面均略有提升;在峰值信噪比方面则平均仅减少 0.04 dB 左右,与原始 HEVC 编码对比几乎不影响其视频质量。但本文方案在视频质量损失与文献[11]相比略低的情况下,编码时间下降幅度更大。此外,对比 Johnny 序列结果,本文方案对于运动较为剧烈的视频序列的复杂度降低程度更大。从不同尺寸的视频序列对比结果可看出,对于大尺寸视频序列,本文所提方案的时间复杂度降低效果更明显,这是由于视频序列尺寸越大,其需要划分的 CU 数目越多,迭代遍历过程计算量更大,利用本文方案跳过的 RD 代价计算过程则更多。此外,基于同等计算机配置,在 Matlab2014a 平台针对基于朴素贝叶斯分类算法的划分模型建立过程进行分析,各尺寸的测试序列的划分模型建立时间均在 1 s 以内,该时长相对于编码时长可忽略不计,进一步验证本文分类模型算法的简单有效。

为了便于更加清晰地展现本文方案在大幅降低转码的时间复杂度的同时保持了良好的视频质量,以级联转码的 RD 性能作为对比的基准,图 5 分别给出了所有测试序列的 RD 曲线图。由于本文方案在码率以及 PSNR 两

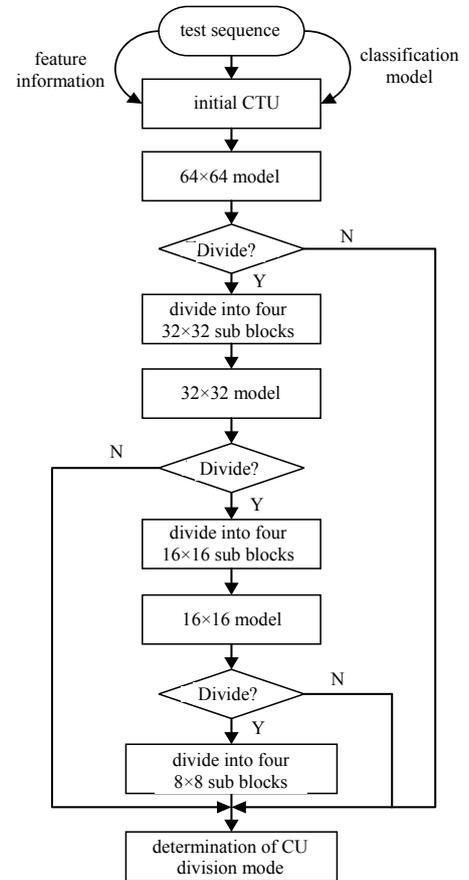


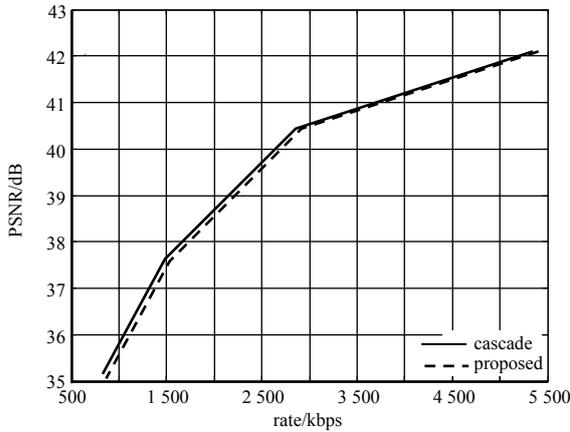
Fig.4 Flow of proposed CU fast partition algorithm
图 4 本文所提 CU 快速划分算法流程图

方面与文献[11]都极为接近,且略高于文献[11],因此为了避免线条叠加造成视觉混乱,仅展示本文方案的 RD 曲线与级联转码的 RD 曲线,以此说明本文方案对视频质量几乎不产生影响。图中曲线的横纵坐标分别表示编码比特率及峰值信噪比,在同一横坐标值处,曲线的纵坐标值越高代表其 RD 性能越高。从图 5 中可观察出,本文所提方案与级联转码方案相比性能略低,但整体上两者的 RD 曲线基本重合,说明本文方案在大幅降低编码复杂度的同时对其原有性能影响较小,保持了与级联转码基本一致的视频质量。

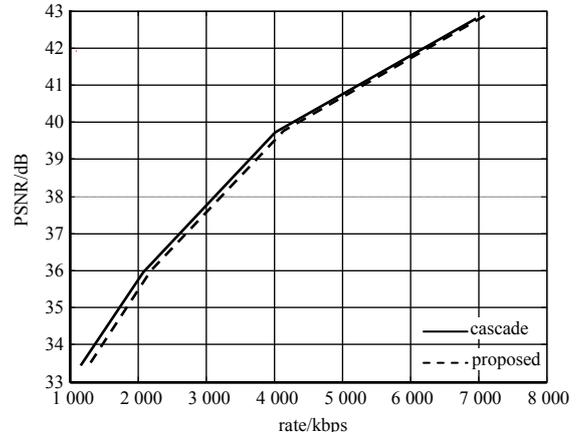
表 1 转码性能对比

Table1 Transcoding performance comparison

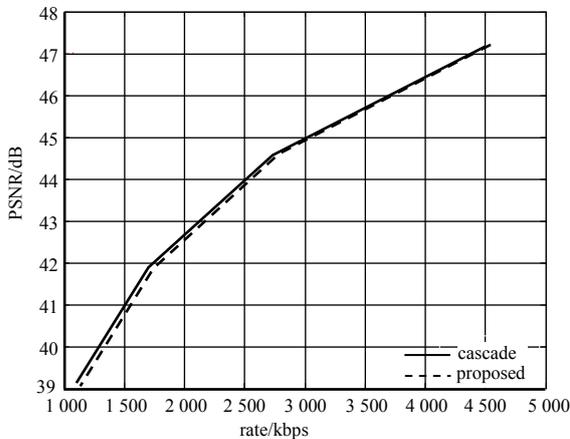
video sequence	QP	$\Delta PSNR/dB$		$\Delta BR/\%$		$\Delta T/\%$	
		proposed	Ref.[11]	proposed	Ref.[11]	proposed	Ref.[11]
BasketballDrill 832×480	22	-0.03	-0.02	0.94	0.94	50.14	45.23
	26	-0.05	-0.02	1.75	1.58	53.98	42.40
	30	-0.06	-0.05	2.36	2.36	57.89	42.39
	34	-0.07	-0.09	2.29	3.65	61.03	42.69
BQMall 832×480	22	-0.05	-0.01	1.42	0.87	48.29	42.17
	26	-0.04	-0.02	2.36	1.40	51.52	42.27
	30	-0.04	-0.03	3.77	1.96	54.26	42.10
	34	-0.10	-0.09	2.74	2.22	57.31	42.62
FourPeople 1280×720	22	-0.07	-0.06	0.20	0.39	60.38	61.34
	26	-0.02	-0.03	0.44	0.59	58.70	62.60
	30	-0.02	-0.02	0.52	0.72	61.51	62.57
	34	-0.05	-0.03	0.40	0.48	64.86	63.50
Johnny 1280×720	22	-0.03	-0.01	0.41	0.90	57.73	52.22
	26	-0.02	-0.01	0.54	1.57	61.47	50.79
	30	-0.02	-0.07	0.65	1.63	64.64	53.22
	34	-0.05	-0.17	0.54	1.07	68.51	54.61
average		-0.044	-0.046	1.33	1.40	58.26	50.14



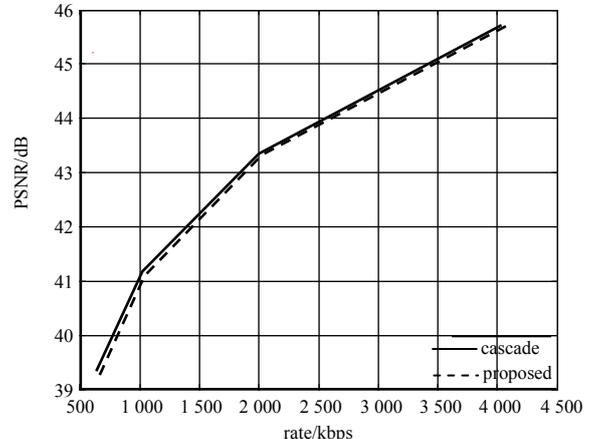
(a) BasketballDrill



(b) BQMall



(c) FourPeople



(d) Johnny

Fig.5 RD performance comparison

图 5 RD 性能比较

3 结论

针对移动设备之间实现低功耗的视频通信用应用问题,结合分布式视频编码及 HEVC 的各自特点,即分布式视频编码方式编码端简单,解码端复杂;HEVC 的编码端复杂,解码端简单,通过转码形式实现一种低功耗的视频通信。为降低转码部分的高复杂度,本文通过提取 DVC 解码过程中的相关特征信息,利用朴素贝叶斯算法生成 CU 快速划分模型,避免了 HEVC 编码过程中复杂度极高的 CU 递归遍历的计算过程。在相同的实验配置环境及 HEVC 参考软件下,对本文所提方案与级联转码方案进行了测性对比。通过结果分析对比,本文方案在编码比特率略有上升的情况下大幅缩短了 HEVC 编码时间,且几乎不影响视频质量。

参考文献:

- [1] SULLIVAN G,OHM J,HAN W,et al. Overview of the High Efficiency Video Coding(HEVC) standard[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2012,22(12):1649–1668.
- [2] CORREA G,ASSUNCAO P,AGOSTINI L V,et al. Performance and computational complexity assessment of High-Efficiency Video Encoders[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2012,22(12):1899–1909.
- [3] GUO X,LU Y,WU F,et al. Distributed video coding using wavelet[C]// IEEE International Symposium on Circuits and Systems. Kos,Greece:IEEE, 2006:5427–5430.
- [4] RONG S,QING L,XU Y,et al. Mobile video communications based on cloud transcoding[J]. Proceedings, 2017,1(3):143.
- [5] WEI Y,WANG Z,XU M,et al. An LSTM method for predicting CU splitting in H.264 to HEVC transcoding[C]// Visual Communications and Image Processing. Petersburg,Florida,USA:[s.n.], 2017:1–4.
- [6] YUAN H,GUO C,LIU J,et al. Motion-homogeneous-based fast transcoding method from H.264/AVC to HEVC[J]. IEEE Transactions on Multimedia, 2017,19(7):1416–1430.
- [7] ZHANG D,TONG J,ZANG D. Fast CU partition for H.264/AVC to HEVC transcoding based on Fisher discriminant analysis[C]// Visual Communications and Image Processing. Chengdu,Sichuan,China:IEEE, 2016:1–4.
- [8] CORRALES–GARCÍA A,RODRÍGUEZ–SÁNCHEZ R,MARTINEZ J L,et al. Multimedia communications using a fast and flexible DVC to H.264/AVC/SVC transcoder[J]. Journal of Signal Processing Systems, 2015,79(3):211–232.
- [9] CORRALES–GARCÍA A,RODRÍGUEZ–SÁNCHEZ R,MARTINEZ J L,et al. A GPU-based DVC to H.264/AVC transcoder[C]// International Conference on Hybrid Artificial Intelligence Systems. San Sebastian,Spain:Spring, 2010:233–240.
- [10] MARTINEZ J,GARCIA–CORRALES G,CUENCA P,et al. Applying data mining techniques in a Wyner–Ziv to H.264 video transcoder[J]. Advances in Computational Intelligence, 2011(6691):497–504.
- [11] 卿粼波,华玲,何小海,等. 一种基于 Fisher 判别式的 DVC–HEVC 视频转码方法:中国,201810573870.6[P]. 2018. (QING Linbo,HUA Ling,HE Xiaohai,et al. A DVC–HEVC video transcoding method based on Fisher discriminant: China, 201810573870.6[P]. 2018.)
- [12] LI X,WIEN M,OHM J R. Rate–complexity–distortion optimization for hybrid video coding[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2011,21(7):957–970.
- [13] RISH I. An empirical study of the naïve Bayes classifier[J]. Workshop on Empirical Methods in Artificial Intelligence, 2001(1):41–46.