

DOI:10.19651/j.cnki.emt.2210889

基于 VVC 的多方向细化 Merge 模式优化算法*

唐健翔 熊淑华 孙伟恒 何小海
(四川大学电子信息学院 成都 610065)

摘要: 针对多功能视频编码标准 VVC 帧间预测中, Merge 模式未充分考虑运动存在的多种方向性从而降低了预测精度的问题, 提出了一种基于 VVC 的多方向细化 Merge 模式优化算法。本算法在分析带运动矢量差的 Merge 技术的基础上, 首先调整了步长选择范围并增加多种搜索方向, 然后根据步长自适应选择色度块的运动补偿方式, 并最终根据率失真代价准则选择最优的运动矢量信息。实验结果表明, 本文算法与 VTM-12.0 参考模型相比, 在低延迟 P 帧配置下 Y、U、V 三分量的 BD-rate 分别平均下降了 0.57%、0.62%、0.25%, 在随机接入配置下的 BD-rate 中 Y 分量下降了 0.27%、U 分量仅提高了 0.11%、V 分量降低了 0.04%, 有效地提升了 Merge 模式的编码性能。

关键词: 多功能视频编码; 帧间预测; 多方向细化; Merge 模式; 编码性能

中图分类号: TN919.81 **文献标识码:** A **国家标准学科分类代码:** 510.5099

Multi-direction refinement Merge mode optimization algorithm based on VVC

Tang Jianxiang Xiong Shuhua Sun Weiheng He Xiaohai

(School of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China)

Abstract: Due to insufficient consideration of various motion orientations in Merge mode, reducing the prediction accuracy of versatile video coding (VVC) interframe prediction, a VVC-based multi-direction refinement Merge mode optimization algorithm is proposed. Based on the analysis of Merge mode with motion vector difference, the algorithm first adjusts the step size selection range and adds multiple search directions, then adaptively selects the motion compensation method of chroma block according to the step size. And finally selects the optimal motion vector information according to the rate-distortion cost criterion. The results show that, comparing with VTM-12.0 reference model, the BD-rate of Y, U and V components decreases by 0.57%, 0.62% and 0.25% in low-delay P-frame configuration, respectively. In the random access configuration, the Y component of BD-rate decreases by 0.27%, the U component only increases by 0.11%, and the V component decreases by 0.04%, which effectively improves the coding performance of Merge mode.

Keywords: versatile video coding; inter prediction; multi-direction refinement; Merge mode; the coding performance

0 引言

2020年7月,新一代视频编码标准发布了,并命名为多功能视频^[1](versatile video coding, VVC)。相对于上一代视频编码标准——高效视频编码(high efficiency video coding, HEVC)^[2],VVC引入了很多新技术,提高编码性能约50%。

VVC^[3]标准的整体编码框架^[4-5]和 HEVC^[6-7]大体一致,仍然是沿用基于块的混合编码结构。但在整个框架中对很多环节都进行了改进或增加了一些新的技术,其中帧

间预测^[8]模块增加了一些重要的改进措施。例如,扩展的 Merge 技术、带运动矢量差的 Merge 技术^[9]、双向光流技术、帧内帧间联合预测等。其中部分新技术对 VVC 的 Merge 模式进行了改进和优化,显著地提高了 VVC 的编码性能。

虽然 Merge 模式能够进一步压缩数据量,但是也会相应地产生一定的误差,仍然有值得完善的地方。虽然直接基于 VVC 中带运动矢量差的 Merge 进行优化的研究还不多,但由于 HEVC 编码标准已用到 Merge 模式,因此针对 Merge 的优化算法较多。为了提高 Merge 模式的预测准

收稿日期:2022-07-29

* 基金项目:国家自然科学基金(61871279, 62211530110)项目资助

确性,近些年涌现了许多相关研究。Hong 等^[10]提出了一种新的 Merge 模式决策算法,该算法基于绝对双向预测差异和模板匹配准则,利用解码端运动矢量(motion vector, MV)候选算法进行 Merge 模式决策。Li 等^[11]提出了一种运动辅助合并模式,通过在特定搜索区域搜索获得的小运动提供更加精准的运动矢量。Li 等^[12]提出了一种低复杂度 Merge 候选决策方案,并利用运动估计过程中得到的真实 MV 和一个帧间预测单元的每个 Merge 候选对象的位置信息,确定最有利的 Merge 候选。Kudo 等^[13]考虑到如果参考块以帧内预测模式进行编码,降低了在 Merge 模式下的运动补偿效率,因此提出了一种考虑预测连续性的 MV 预测方法。Fan 等^[14]发现了 HEVC 中现有的帧内预测的缺点,解释了边界滤波的编码增益原因,并利用广义高斯分布函数建立了边界上的滤波方法,并将所提出的边界滤波器扩展到帧间预测 Merge 过程。Park 等^[15]提出了一种构造 Merge 模式候选列表的有效方法,通过生成一个备选合并候选列表来减少使用零运动矢量作为候选列表的情况。

上述方法都在一定程度上提升了 Merge 模式的编码性能,但在 Merge 模式预测过程中,上述算法并未充分考虑到运动方向的多样性的特点。因此本文通过分析 Merge 模式的运动搜索过程,并基于带运动矢量差的 Merge,提出了一种基于 VVC 的多方向细化 Merge 模式优化算法。

1 VVC 中的 Merge 模式预测技术

1.1 基本流程

如图 1 所示为 VVC 中 Merge 模式预测的整体流程,首先进行预测,建立扩展的 Merge 候选列表、进行帧内帧间联合预测(combined inter and intra prediction, CIIP)候选过程和带运动矢量差的 Merge(Merge mode with MVD, MMVD)候选过程,在此过程中需要更新率失真候选模式列表,并缩短该候选模式列表的长度;然后进行最终的模式选择,迭代进行细选,通过运动补偿和残差编码选择出当前最优 Merge 模式,并最终计算出滤波后的最优代价。

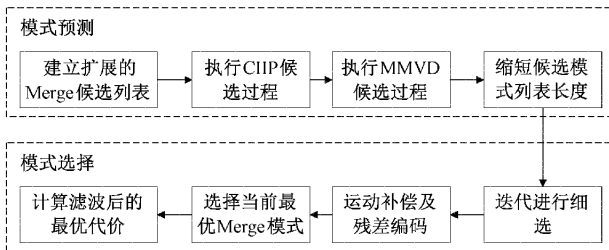


图 1 VVC 中 Merge 模式基本流程

1.2 MMVD 技术简介

MMVD 技术适用于帧间预测^[16]过程中的 Skip 和 Merge 模式,是一种运动矢量(motion vector, MV)的特殊表示形式^[17]。MMVD 技术^[18]实现过程如下:

1)重复利用扩展的 Merge 模式候选列表,选择该列表中的前两个候选 MV 作为初始 MV 信息;

2)主要在运动幅度和运动方向上,分别对 1)中的两个初始 MV 进行扩展,来获取最终的 MV 表示形式,以形成更新的 MV 信息。

具体流程如下:

1)根据 VVC 中扩展的 Merge 候选列表建立过程,得到当前编码单元(coding unit, CU)的 MV 候选列表,检查相邻块 MV 的顺序如图 2 所示。

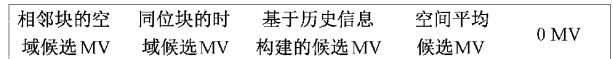


图 2 VVC 的候选列表示意图

值得注意的是,候选列表中最多可有 6 个候选 MV。图 2 所示通过从左到右的顺序依次检查上述的 MV 信息是否有效,若有效则将 MV 信息添加到候选列表中;否则候选列表添加零向量。并形成已初始化的扩展的 Merge 模式候选列表,并包括最多如图 2 所示 5 种 MV 信息。

2)初始化率失真候选模式列表,并设置该候选模式列表的初始长度为 5。对于 1)中初始化得到的扩展的 Merge 模式候选列表的前两个候选 MV,将其作为基准的初始 MV,以该候选 MV 所指向的终止位置为起始点,按照上下左右 4 个方向的方式进行搜索,每个方向分别进行 8 种不同步长的搜索。如图 3 所示,白色圈代表起始点,其他的圈和中心白色圈之间的直线距离代表步长,灰度相同的点步长相同。

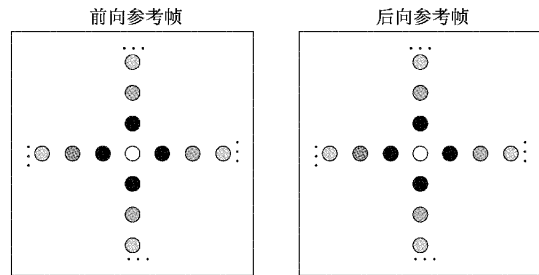


图 3 MMVD 搜索过程示意图

根据不同方向上的不同搜索步长,每一次都将形成一个新的运动矢量,即经过细化的 MV。该 MV 包含 3 个信息,分别是起始点、搜索方向、搜索步长,从而一个初始的 MV 可以扩展出共计 32 个新的 MV 信息。根据每一个 MV 对应的搜索步长索引大小,调整运动补偿和双向光流过程。当步长索引值大于 2 时选择跳过色度块的运动补偿和双向光流技术;否则只选择跳过色度块的运动补偿阶段,然后通过采用运动补偿方式得到当前 CU 预测值。其中,双向光流技术运用于双向预测中,用于修正子块 CU 中亮度分量 MV 及对应的双向预测值,以提高预测精度。

3)根据率失真代价准则更新率失真候选模式列表。在 Merge 模式下,两个初始 MV 会得到 64 个预测值,并计算

相应的率失真代价,选择代价小的模式依次插入到上述列表中,形成一个率失真代价依次增加的候选模式列表。实际过程中为了降低计算复杂度,会循环遍历该候选列表,当第*i*个模式的代价大于率失真候选模式列表中的第1个模式(即代价最小的那个模式)的1.25倍,则后面的模式不再进行变换域的绝对误差和计算(sum of absolute transformed difference, SATD),记缩短后的列表长度为*i*,率失真代价计算公式为:

$$J = SAD + \lambda \times B \quad (1)$$

式中:*SAD*表示绝对误差和, λ 为拉格朗日算子,*B*为实际编码的比特数。

4)遍历缩短后的率失真候选模式列表。对于率失真代价候选列表的前*i*个逐个进行遍历,如果选择的是MMVD模式,会根据该模式对应的Merge候选列表中的索引,计算出搜索方向和搜索步长,进行运动补偿得到当前CU的预测值。使用残差编码^[19]将预测值与当前CU的像素值进行相减,得到当前CU像素域残差。再将残差进行哈达玛变换并通过计算得到SATD,用于率失真代价选择,并最终选择出代价最小的那个Merge模式。

5)对于3)中选择的代价最小的那个Merge模式进行熵编码,熵编码^[20]方案沿用计算率失真代价时计算得到的比特数,具体计算方法如下:

在计算率失真代价的时候需要根据索引计算出编码当前CU所需要的比特数,索引包含3种信息——初始运动矢量索引、搜索步长索引、方向索引。

表1对应初始运动矢量索引,采用一位二进制编码,表2对应搜索步长索引及相应编码位数。表3对应方向索引及对应码字,包含4种方向,即上下左右,需要采用两位二进制编码。最终编码当前CU所需要的比特数等于初始运动矢量码长、搜索步长对应码长、方向对应码长之和。

表1 初始运动矢量索引

初始运动矢量索引	0	1
编码位数	1	1

表2 搜索步长索引及编码位数

步长索引	0	1	2	3	4	5	6	7
步长距离	1/4	1/2	1	2	4	8	16	32
编码位数	1	2	3	4	5	6	7	8

表3 移动方向索引

方向索引	0	1	2	3
方向码字	00	01	10	11
X轴	+	-	N/A	N/A
Y轴	N/A	N/A	+	-

2 VVC Merge 模式优化

2.1 MMVD 技术的分析

本文统计了VVC标准测试序列中,低延迟P(Low delay P, LDP)配置下搜索步长的选中概率。由于所有的序列均具有差不多相似的特性,因此这里只展示如表4所示的5个序列。这些序列来自于不同的分辨率,每一种分辨率选择4种常用的量化参数(quantization parameter, QP),即22、27、32、37。为了方便叙述,每个序列小步长(包括1/4、1/2、1、2)的选择概率分别为上述4种QP下小步长选择概率的平均值。经过统计计算,这5个序列中选择小步长的概率,分别是82.20%、85.00%、86.41%、88.32%、77.53%,平均选中概率为83.39%。若定义超过80%占比为绝大多数,那么这里可以得出,在进行步长搜索过程中,较小像素精度步长的选择占据了绝大多数比例。

不可忽略的是,从表4观察到,随着QP的增加,搜索步长的选取也会随之增加,从而出现大步长(4、8、16、32)选择概率出现略微上升的情况。基于以上统计分析,可以定义一种新的步长选择方式,尽可能地减少步长选择范围,并提高步长选择精度,即可以压缩搜索范围,以实现使用较少比特数对平移运动进行更高的精细化表示。

2.2 Merge 模式优化算法

通过2.1节的统计分析可以得出,随着量化参数QP的增加,小步长的选择概率出现了一定波动。但是在不同分辨率尺寸序列的不同QP条件下,小步长的选择基本上仍然占据了绝大部分。同时,在选择较大步长时,会增加所需要表示运动信息的比特数,从而造成了预测性能的下降。同时,考虑到实际运动存在的多方向性,本节提出了一种基于VVC的多方向细化Merge模式优化算法。

图4是本优化算法的整体流程。

本方法新增加了一种预测模式,并命名为MMVD_PRO。考虑到较大步长的选择会增加表示运动信息的比特数,因此缩减了步长选择范围。考虑到实际运动的多方向性,因此从原来的4个方向(上、下、左、右)扩展到八个方向,这8个方向分别是上、下、左、右、东北、东南、西南、西北。同时,改进了在倾斜方向(非上下左右4个方向)上的搜索步长选择,以提高在倾斜方向上的选择精度。为了进一步提高预测精度,将原有的步长索引大于2时,调整运动补偿和双向光流过程,改为步长索引大于1时进行调整。

步骤1)执行扩展的Merge候选过程,初始化扩展的Merge模式候选列表和率失真候选模式列表,并确定率失真候选模式列表的长度。

步骤2)执行CIIP候选过程和标准的MMVD候选过程,更新步骤1)中的候选模式列表。

步骤3)进行MMVD_PRO候选,选择步骤1)中扩展的Merge候选列表中前两个MV信息作为初始MV,并从4种步长、8个方向进行拓展,判断当前的步长索引是否大

表 4 LDP 条件下 MMVD 不同分辨率序列的搜索步长选中概率

序列名称	量化参数	1/4 pixels	1/2 pixels	1 pixels	2 pixels	4 pixels	8 pixels	16 pixels	32 pixels
BasketballPass 416×240	22	36.55%	23.59%	14.12%	8.4%	5.68%	4.81%	3.43%	3.43%
	27	28.30%	26.71%	16.46%	9.83%	6.40%	5.23%	3.54%	3.54%
	32	21.30%	30.12%	19.13%	11.58%	6.94%	5.16%	2.88%	2.88%
	37	14.54%	32.18%	21.84%	14.12%	8.05%	5.26%	2.00%	2.00%
BasketballDrill 832×480	22	43.89%	19.67%	12.67%	8.58%	5.96%	4.71%	2.26%	2.26%
	27	36.37%	23.55%	14.87%	10.00%	6.55%	4.78%	1.94%	1.94%
	32	27.68%	28.72%	17.32%	11.75%	7.29%	4.74%	1.25%	1.25%
KristenAndSara 1 280×720	22	21.21%	31.31%	18.74%	13.69%	8.66%	5.15%	0.62%	0.62%
	27	31.36%	28.51%	16.96%	9.30%	5.78%	4.15%	1.97%	1.97%
	32	29.01%	30.01%	18.39%	10.57%	6.24%	4.04%	0.87%	0.87%
BQTerrace 1 920×1 080	22	25.25%	29.61%	19.34%	12.04%	7.50%	4.75%	0.75%	0.75%
	27	21.10%	30.55%	20.02%	13.61%	8.54%	5.55%	0.31%	0.31%
	32	54.45%	17.48%	10.00%	5.54%	3.87%	3.31%	2.68%	2.68%
SlideShow 1 280×720	22	53.96%	18.20%	9.73%	5.63%	3.92%	3.40%	2.59%	2.59%
	27	50.96%	19.81%	10.85%	6.66%	4.37%	3.35%	2.00%	2.00%
	32	52.16%	19.91%	10.88%	7.05%	4.70%	3.20%	1.04%	1.04%
	22	29.16%	24.10%	16.05%	10.29%	8.00%	7.15%	2.63%	2.63%
	27	22.73%	24.97%	17.98%	12.18%	9.42%	8.60%	2.06%	2.06%
	32	17.93%	27.88%	18.45%	13.13%	10.48%	9.71%	1.21%	1.21%
	37	15.56%	28.02%	17.66%	14.04%	11.83%	11.32%	0.79%	0.79%

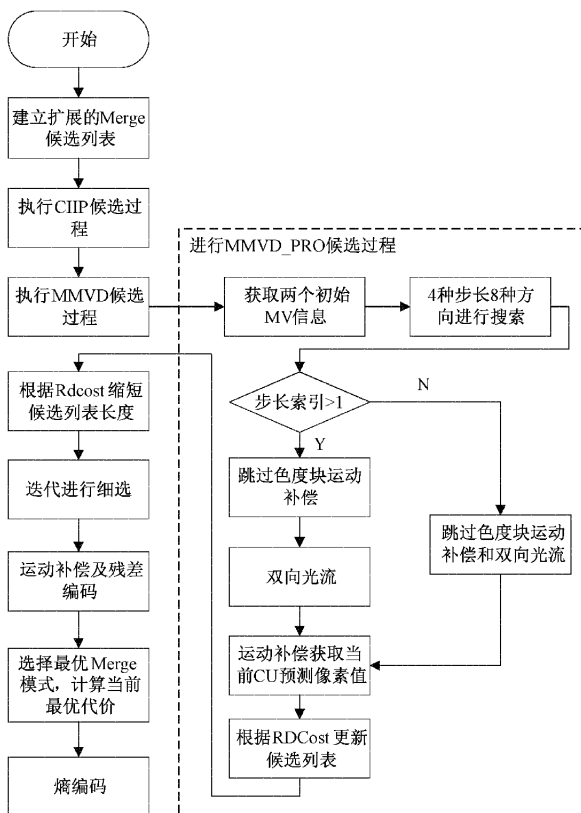


图 4 优化算法的整体流程

于 1,若是,则跳过色度块运动补偿和双向光流预测;否则,跳过色度块运动补偿。

步骤 4) 通过运动补偿获取当前 CU 的预测像素值,通过计算不同预测值对应的率失真代价,更新步骤 2)中的候选模式列表。

步骤 5) 根据率失真代价缩减步骤 4)中的候选模式列表长度以进行细选,再次通过运动补偿计算预测值,并进行残差编码,最终得到当前最优的 Merge 模式及其代价。

步骤 6) 对于最终选择得到的最优模式进行熵编码,若选择的是 MMVD_PRO 模式,则编码 MMVD_PRO 标志位,否则编码其他模式的标志位。

图 5 中展示了本文算法优化后的方向,其搜索方向在上下左右的基础上增加了对角方向,其中实线代表原始方

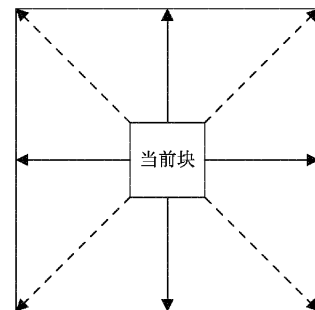


图 5 优化后的搜索方向

向,新增方向采用虚线表示。利用对角方向,增加覆盖可能存在的细化方向,对多种运动情况进行多方向细化表示。

表5展示了优化后8种方向对应的搜索步长及编码位数,其中倾斜方向上的搜索步长选择的像素精度相对于东南西北更高,这是因为较小精度的选择占据绝大部分,提高倾斜方向上的步长选择精度有利于预测性能。

表5 优化后的搜索步长及对应编码位数

步长索引	0	1	2	3
上下左右像素距离	1/4	1/2	1	2
对角方向像素距离	1/8	1/4	1/2	1
编码位数	1	2	3	4

表6给出了优化后的搜索方向表,表示8种不同方向及其对应的索引及码字。

表6 优化后的搜索方向表

方向索引	0	1	2	3	4	5	6	7
方向码字	000	001	010	011	100	101	110	111
X轴	+	-	N/A	N/A	+	+	-	-
Y轴	N/A	N/A	+	-	+	-	+	-

图6展示了优化后的搜索过程示意图,其中灰色块和黑色块分别表示对角方向的搜索过程和上下左右4种方向的搜索过程,适应了多方向下的运动情况。因此,本优化算法适合运动不剧烈,且方向变化较多的运动场景。由于优化后的搜索步长索引码字减少了1/2,但方向索引的码字仅增加了一个比特,因此可以一定程度上节省编码比特数。同时,本文算法和原有的Merge模式算法叠加使用,利用率失真代价准则选择出当前最优Merge模式。

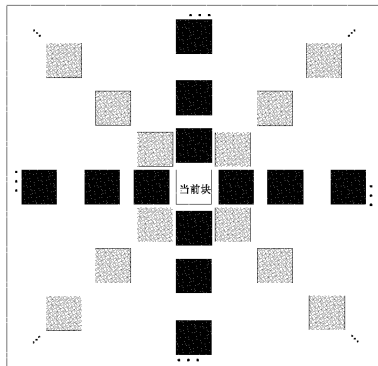


图6 修改后的搜索过程示意图

3 实验结果及分析

为了检验本文算法的性能,因此采用了标准测试环境对于H.266/VVC标准的视频序列进行测试。测试序列

的具体实验平台配置为16 GB内存,CPU为11th Gen Intel(R)Core(TM)i5-11500@2.70 GHz,64位Windows系统的PC机。使用H.266/VVC标准参考软件VTM-12.0作为测试条件,编译环境是Microsoft Visual Studio 2019。测试时的量化参数QP选择22,27,32,37。根据不同QP对应峰值信噪比和码率变化关系,会计算得到每种QP所对应的BD-rate(bjontegaard-delta rate),并在表7和8中进行展示。分别采用低延迟P(low delay P,LDP)和随机接入(random access,RA)预测结构进行测试,图像组(group of picture,GOP)的尺寸分别设置为8和16,并对不同视频序列的前100帧进行测试。

为了衡量本优化算法的时间复杂度,因此采用相对编码时间来进行比较,计算公式定义为:

$$T = \frac{T_{Proposed}}{T_{VVC}} \times 100\% \quad (2)$$

式中: $T_{Proposed}$ 是所提算法的总体编码时间, T_{VVC} 是VVC标准算法总体编码时间,为上述两者相除的结果,表示相对编码时间。

为了衡量本优化算法中MMVD_PRO模式的总体使用情况,采用算法使用率来进行测试,计算方法定义为:

$$P = \frac{N_{Proposed}}{N_{Total}} \quad (3)$$

式中: $N_{Proposed}$ 是所提模式的使用次数, N_{Total} 是所有模式的使用次数。

BD-rate表示在同一视频客观质量的情况下,优化后算法相较于标准算法编码性能的变化量。BD-rate为负则表示优化后算法的编码性能得到了提高,反之算法性能下降。如表7所示,给出了在LDP配置下的算法实验结果。如表8所示,给出在RA配置下的算法实验结果。

如表7所示,在LDP配置下,包含了5种不同类型的视频,B、C、D、E对应4种不同分辨率^[21]类型,F为面向屏幕内容的视频类型。对于较小分辨率视频序列,由于一个CU包含的内容较为丰富,因此可能在较小的步长范围内找到较为匹配的内容;反之,对于较大分辨率视频序列,可能在较大步长范围内找到匹配内容。同时,本优化算法相较于VTM-12.0标准算法所有序列在亮度分量Y下的BD-rate均为负数,这表明编码性能均具有一定的提升,且Y、U、V三分量的BD-rate分别平均下降了0.57%、0.62%、0.25%。其中序列BQSquare中Y分量的BD-rate下降了1.33%,是所有序列中下降程度最高的,这是因为该序列中存在着大量的平动区域,而且部分平动区域的运动是沿着倾斜方向的,因此本算法针对斜向微小的运动区域上进行了更加精细化的运动表示。因此针对存在倾斜方向上的平动运动,本算法可以提高预测的准确性。

如表8所示,在RA配置下所有的序列在亮度分量Y的BD-rate均为负数,性能有一定提升,Y、U、V三分量的BD-rate分别是-0.27%、0.11%、-0.04%。其中视频序

表 7 在 LDP 配置下的算法实验结果

类别	序列	BD-rate/%			T	P
		Y	U	V		
B 1 920×1 080	Cactus	-0.13	0.22	0.13	120.73%	20.04%
	BasketballDrive	-0.31	-0.01	0.29	115.23%	21.64%
	BQTerrace	-0.88	0.12	-0.06	110.56%	23.50%
C 832×480	BasketballDrill	-0.63	-0.74	-0.22	129.35%	24.65%
	BQMall	-0.30	-1.18	-0.24	117.02%	28.97%
	PartyScene	-0.81	-0.88	-0.35	136.79%	22.68%
D 416×240	RaceHorsesC	-0.64	-1.27	-0.90	124.01%	25.71%
	BasketballPass	-0.56	-0.49	-0.45	111.54%	26.97%
	BQSquare	-1.33	-3.18	-0.86	112.81%	25.58%
E 1 280×720	BlowingBubbles	-0.57	-0.73	-0.62	116.19%	23.87%
	RaceHorses	-0.80	-1.26	-1.35	111.44%	26.87%
	FourPeople	-0.20	0.32	-0.06	111.02%	26.90%
F	Johnny	-0.55	-0.18	0.59	125.51%	29.24%
	KristenAndSara	-0.69	0	0.39	116.65%	27.98%
	BasketballDrillText	-0.48	-0.48	0.03	112.48%	27.61%
平均值	SlideEditing	-0.34	0.17	0.01	120.26%	12.49%
	SlideShow	-0.50	-0.96	-0.55	118.60%	26.56%
平均值	—	-0.57	-0.62	-0.25	118.35%	24.78%

表 8 在 RA 配置下的算法实验结果

类别	序列	BD-rate/%			T	P
		Y	U	V		
B 1 920×1 080	Cactus	-0.03	0.29	0.25	122.28%	19.46%
	BasketballDrive	-0.26	0.32	-0.08	120.21%	19.68%
	BQTerrace	-0.38	0.38	-0.27	129.90%	25.54%
C 832×480	BasketballDrill	-0.06	0.09	0.13	110.78%	22.40%
	BQMall	-0.21	0.24	-0.06	106.71%	23.77%
	PartyScene	-0.20	-0.24	0.24	134.71%	23.06%
D 416×240	RaceHorses	-0.28	-0.32	-0.16	108.34%	23.93%
	BasketballPass	-0.26	-0.18	-0.58	110.23%	23.80%
	BQSquare	-0.63	-0.32	-0.31	125.12%	26.96%
F	BlowingBubbles	-0.48	0.10	-0.41	134.04%	23.58%
	RaceHorses	-0.59	0.48	-0.98	107.49%	24.60%
	BasketballDrillText	-0.16	0.60	0.44	111.52%	21.80%
平均值	SlideEditing	-0.06	0.03	0.03	110.31%	14.14%
	SlideShow	-0.18	0.07	1.14	112.76%	23.92%
平均值	—	-0.27	0.11	-0.04	117.46%	22.62%

列 BQSquare 的 Y 分量的 BD-rate 下降了 0.63%，仍然是所有序列中最高的，也再次印证了本算法对于斜向微小运动区域的精细化运动表示，即多方向运动的细化表示。

综合表 7 和 8，本算法对于沿倾斜方向上的平动区域进行了更加精细化的运动表示，提高了预测的准确性，所

有序列的 Y 分量上 BD-rate 都有了一定程度的下降。然而，部分序列的色度分量 U、V 的 BD-rate 存在一定的上升，这是因为在进行 Merge 模式预测时，为了较大程度地压缩数据量，不会对色度分量进行单独划分，而直接采用亮度块的运动矢量信息作为当前色度块的运动信息。若

色度块内容表现相对平坦,则不太适应于色度块,因此出现了色度分量上的性能下降。同时在平均编码时间上,本算法与标准算法相比变化不大,但是率失真性能具有一定程度的提升。本算法的平均使用率在22%~24%,进一步印证了本算法的有效性。

表9给出了部分VVC标准算法和本文算法的码率和PSNR实验对比结果,并绘制了如图7所示的率失真曲线。

表9 本文所提算法与VVC标准算法的部分实验结果对比

视频序列	QP	码率/Kbps		PSNR/dB	
		VVC	本文	VVC	本文
BQ	22	1 864.632	1 832.126 4	38.554 2	38.560 9
	27	588.216	581.832	34.359 7	34.383 3
	32	251.860 8	250.219 2	31.293	31.289 7
Square	37	120.393 6	120.638 4	28.507 1	28.55
	22	8 789.012	8 722.328	38.151 3	38.162
Party Scene	27	3 404.896	3 383.376	34.043 2	34.054 1
	32	1 488.2	1 482.924	30.734 7	30.744 4
	37	660.276	659.296	27.798 8	27.821 1

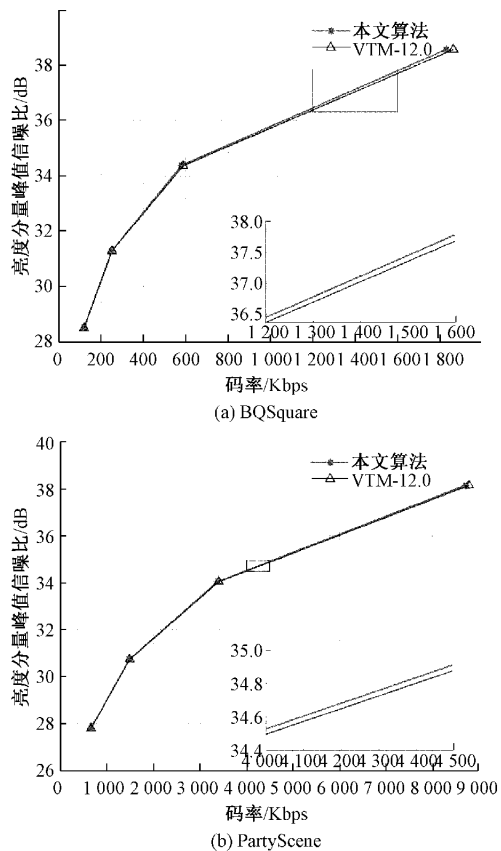


图7 部分序列率失真曲线

PSNR为峰值信噪比,单位dB,表示重建视频的质量。纵坐标表示其亮度分量的峰值信噪比。横坐标为码率,单

位Kbps,表示单位时间传送的数据位数。如图7(a)和(b)所示,本文算法的率失真曲线均位于VVC标准之上,进一步证明了本文算法相对于VVC标准算法的优越性。

4 结 论

针对VVC帧间预测中Merge模式存在预测不精准,未充分考虑到倾斜方向上的运动情形的问题,提出了一种基于VVC的多方向细化Merge模式优化算法。本文从Merge模式中的MMVD技术出发,分析了MMVD技术搜索过程及最终编码方式。由于实际中较大步长的选择概率较低,同时考虑到运动可能存在的多方向性,因此本算法同时调整搜索步长,并增加了搜索方向,自适应调节运动搜索过程。实验结果表明,在不明显增加时间复杂度的情况下,所提算法相较于VVC标准算法有效地提高了视频序列的编码性能。其中低延迟P配置下Y分量BD-rate平均下降了0.57%;随机接入配置下Y分量BD-rate平均下降了0.27%。然而值得注意的是,同一个序列下不同QP的搜索步长的分布存在差异,如何理解这种差异以进一步提升该算法的编码性能,还需要进行进一步的研究。

参考文献

- [1] 杨嘉棋,任立成,欧阳益民,等.超高清显示系统设计[J].国外电子测量技术,2020,39(11):107-112.
- [2] SULLIVAN G J, OHM J R, HAN W J, 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.
- [3] 施金诚,杨静.基于深度学习的VVC快速帧内模式决策[J].电子测量技术,2022,45(3):104-111.
- [4] 周芸,郭晓强,潘晓菲.H.266/VVC视频编码帧间预测关键技术研究[J].广播与电视技术,2020,47(7):36-43.
- [5] BROSS B, CHEN J, OHM J R, et al. Developments in international video coding standardization after AVC, with an overview of versatile video coding (VVC) [J]. Proceedings of the IEEE, 2021, 109(9): 1463-1493.
- [6] 夏正鹏,李开宇.高效视频编码(HEVC)帧内预测的硬件实现方案[J].电子测量技术,2017,40(9):201-205.
- [7] 第一个端到端HEVC测试解决方案实现到4K的无缝迁移[J].电子测量技术,2016,39(4):108.
- [8] CHIEN W J, ZHANG L, WINKEN M, et al. Motion vector coding and block merging in the versatile video coding standard [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2021, 31(10): 3848-3861.
- [9] JEONG S, PIAO Y, PARK M W, et al. Merge mode with motion vector difference [C]. 2020 IEEE

- International Conference on Image Processing (ICIP), IEEE, 2020: 1157-1160.
- [10] HONG C Y, HSIEH H J, LIN Y. New merge mode decision in high efficiency video coding (HEVC) [C]. 2014 International Computer Science and Engineering Conference (ICSEC), IEEE, 2014: 103-107.
- [11] LI H, FAN K, WANG R, et al. A motion aided merge mode for HEVC [C]. 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2018: 1773-1776.
- [12] LI M, CHONO K, GOTO S. Low-complexity merge candidate decision for fast HEVC encoding [C]. 2013 IEEE International Conference on Multimedia and Expo Workshops (ICMEW), IEEE, 2013: 1-6.
- [13] KUDO S, KITAHARA M, SHIMIZU A. Motion vector prediction methods considering prediction continuity in HEVC [C]. 2016 Picture Coding Symposium (PCS), IEEE, 2016: 1-5.
- [14] FAN K, WANG R, LI G, et al. Efficient prediction methods with enhanced spatial-temporal correlation for HEVC [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2018, 29 (12): 3716-3728.
- [15] PARK D H, YOON Y U, DO J H, et al. A method of merge candidate list construction using an alternative merge candidate [J]. Journal of Broadcast Engineering, 2019, 24(1): 41-47.
- [16] VENUGOPAL G, MÜLLER K, MARPE D, et al. A unified region-based template matching approach for intra and inter prediction in VVC [C]. 2019 IEEE International Conference on Image Processing (ICIP), IEEE, 2019: 4115-4119.
- [17] 郭泽. H. 266/VVC 高效帧间预测编码算法研究 [D]. 西安: 西安电子科技大学, 2019.
- [18] 毛珏. 视频编码中帧间预测技术研究 [D]. 杭州: 浙江大学, 2020.
- [19] 宋世昶. 下一代视频编码标准的帧间预测及变换编码技术研究 [D]. 成都: 电子科技大学, 2018.
- [20] 王译堃. AVS3 熵编码算法优化 [D]. 北京: 北京邮电大学, 2021.
- [21] 王瑞, 史天运, 包云. 一种基于视频的铁路周界入侵检测智能综合识别技术研究 [J]. 仪器仪表学报, 2020, 41(9): 188-195.

作者简介

唐健翔, 硕士研究生, 主要研究方向为图像与视频编解码。

E-mail: 1506454026@qq.com

熊淑华 (通信作者), 副教授, 主要研究方向为多媒体通信。

E-mail: xiongsh@scu.edu.cn

孙伟恒, 博士研究生, 主要研究方向为深度学习、视频编解码、图像视频后处理。

E-mail: 273375696@qq.com

何小海, 教授, 博士生导师, 主要研究方向为图像处理与网络通信、人工智能与大数据分析、图像通信。

E-mail: hxh@scu.edu