



(12)发明专利申请

(10)申请公布号 CN 110087087 A

(43)申请公布日 2019.08.02

(21)申请号 201910279316.1

(22)申请日 2019.04.09

(71)申请人 同济大学

地址 200092 上海市杨浦区四平路1239号

(72)发明人 张冬冬 钱鹏飞

(74)专利代理机构 上海科律专利代理事务所

(特殊普通合伙) 31290

代理人 叶凤

(51)Int. Cl.

H04N 19/503(2014.01)

H04N 19/147(2014.01)

H04N 19/103(2014.01)

H04N 19/11(2014.01)

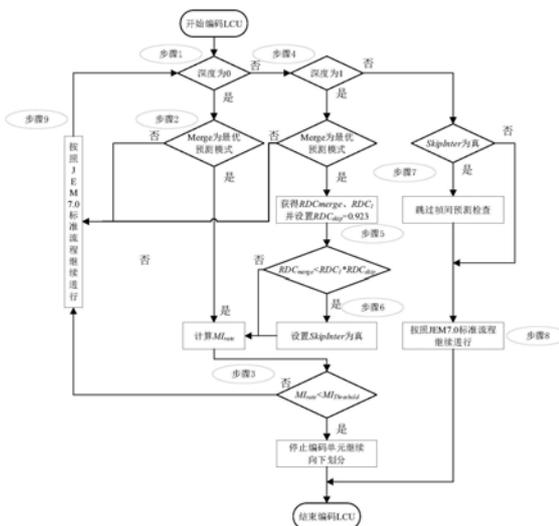
权利要求书1页 说明书4页 附图2页

(54)发明名称

VVC帧间编码单元预测模式提前决策及块划分提前终止方法

(57)摘要

本发明“VVC帧间编码单元预测模式提前决策及块划分提前终止方法”通过分析视频帧之间的运动变化,经过对编码深度为1的帧间编码单元中各个预测模式之间率失真代价的比较以及所选择的最优模式和后续编码深度中帧间预测模式的编码代价的比较,对于符合条件的编码单元直接跳过对帧间预测模式编码代价的检查,从而降低计算复杂度。另外还通过对编码单元当前编码深度和所选择的最优预测模式以及各预测模式之间率失真代价的比值关系来预测当前编码单元的最优划分尺寸,通过对编码单元划分的提前终止,来避免计算复杂度极高的递归划分过程,从而降低其编码器的计算复杂度,在不影响视频编码质量的前提下,可以有效的减少编码器的耗时,提高编码效率。



CN 110087087 A

1. 一种VVC帧间编码单元预测模式提前决策以及块划分提前终止方法,其特征在于包括:

(1) 进行帧间编码单元预测模式选择时,在深度为1的编码单元中,根据所选择的最优预测模式以及当前Merge预测模式与帧间预测模式(Inter_2N×2N)的之间的率失真代价的关系来决策是否跳过后续编码单元中的帧间预测模式;

(2) 进行帧间编码单元块划分决策时,首先统计编码单元的最终编码深度和所选择的最优预测模式之间的关系,统计结果显示两者具有很强的相关性;然后基于对各预测模式的率失真代价的比值 MI_{rate} 来辅助决策是否需要当前编码单元的划分进行终止;

MI_{rate} 的计算如下所示:

$$MI_{rate} = \frac{Merge_{RDC}}{Inter_{RDC}}$$

其中, $Merge_{RDC}$ 表示的是当前编码单元使用Merge预测模式的率失真代价, $Inter_{RDC}$ 则表示是使用帧间预测模式(Inter_2N×2N)的率失真代价。

2. 根据权利要求1所述的一种VVC帧间编码单元预测模式提前决策以及块划分提前终止方法,其特征在于,具体包括如下步骤:

步骤1:基于VVC的通用测试平台JEM7.0,开始编码一个LCU后,首先判断当前的编码深度,初始为0,每递归划分一次编码深度加1,如果编码深度为0则转步骤2,否则转步骤4;

步骤2:判断当前编码单元的最优预测模式是否为Merge,如果是则转步骤3,否则转步骤9;

步骤3:计算 MI_{rate} ,如果 $MI_{rate} < MI_{Threshold}$,编码深度为0和1时, $MI_{Threshold}$ 分别为0.63和0.95,则停止当前编码单元继续向下划分,否则转步骤9;

步骤4:如果当前编码深度为1,则转步骤5,否则转步骤7;

步骤5:获得Merge预测模式的率失真代价 RDC_{merge} 以及帧间预测模式的率失真代价 RDC_1 ,并设置 $RDC_{skip} = 0.923$,转步骤6;

步骤6:如果 $RDC_{merge} < RDC_1 \times RDC_{skip}$,将变量SkipInter设为真,否则转步骤3;

步骤7:检查变量SkipInter是否为真,如果为真,则跳过当前编码深度中对帧间预测模式的检查,否则转步骤8;

步骤8:继续按照JEM7.0标准流程进行;

步骤9:继续按照JEM7.0标准流程进行直到结束当前LCU的编码。

VVC帧间编码单元预测模式提前决策及块划分提前终止方法

技术领域

[0001] 本发明涉及多功能视频编码 (Versatile Video Coding, VVC) 领域, 特别涉及 VVC 帧间编码单元的预测模式提前决策以及块划分提前终止技术。

背景技术

[0002] 近年来, 随着 VR 技术的发展以及消费者对于观看 360 度全景视频的需求越来越大, 互联网上关于 360 度全景视频的内容也出现爆炸式增长。而现在最通用的 360 度全景视频格式和传统的二维平面视频有较大的不同。以最为流行的等距投影 (EquiRectangular Projection, ERP) 格式为例, 由于特殊投影方式的原因, 所以其视频内容存在一定程度的变形, 越靠近图像的上方以及下方则变形越明显。另外, 为了带给观众更加真实和有沉浸感的观看体验, 一般 360 度全景视频的分辨率和帧率都很高, 这也给现在的视频编码算法带来了极大的挑战。所以为了能够更好的处理这类视频, 国际电信联盟的视频编码专家组 (Video Coding Experts Group, VCEG) 和动态图像专家组 (Moving Picture Experts Group, MPEG) 联合研发了最新的视频编码标准, 多功能视频编码标准 (VVC)。它相对于上一代视频编码标准 HEVC 来说, 有了很多方面的改进。在编码结构方面: 在 HEVC 中的二叉树划分模型基础上引入了四叉-二叉划分模型, 能够适应更加复杂的图像纹理, 提高编码效率。另外在帧间预测方面: 也改进和引入了很多的新技术, 包括局部光照补偿 (Local Illumination Compensation, LIC)、两种新的 Merge 模式: Affine Merge 和 FRUC Merge。同时也加入了一些快速算法, 例如在对帧间视频帧进行帧内预测检查时, 当满足跳过条件时, 则会直接跳过帧内预测检测, 加快编码速度。由于编码结构的改变和在帧内以及帧间编码中新技术的引入, 一方面提高了编码效率, 但是也造成了编码器计算复杂度的上升。因此如果想要将 VVC 应用到实际的业务场景当中, 首先要做的就是对其编码器进行优化。

[0003] VVC 的通用测试平台 JEM7.0 的帧间编码流程为: 首先对当前的最大编码单元, 从深度 0 开始, 逐层的对不同编码深度的编码单元进行预测模式的检查, 选择各个编码深度下每个编码单元最优的预测模式。然后对所有编码深度中进行基于率失真代价的比较来选择出最终的编码单元划分方式。因此需要尝试各种不同的帧间预测模式以及不同的编码单元划分方式, 这就给编码器带来了极高的计算复杂度。而目前也有部分研究人员对 VVC 以及 HEVC 的帧间编码复杂度进行了优化, 并取得了一定的效果。如文献 1 (K. Choi, S. H. Park, and E. S. Jang, Coding tree pruning based CU early termination, JCTVC-F092, 2011.) 中提出了一种编码单元划分快速决策的算法, 该文献经过统计分析发现, 如果在帧间预测编码过程中当前最优预测模式为 SKIP (一种特殊的 Merge 模式), 那么当前编码单元的划分为最优的概率超过 95%。因此该文献以最优预测模式是否为 SKIP 为条件来判断当前编码单元划分是否能提前终止, 从而快速决策出编码单元最优的划分尺寸。文献 2 (X. L. Shen, Y. Lu, CU splitting early termination based on weighted SVM [J], Eurasip Journal on Image & Video Processing, 2013 (1): 1-11, 2013.) 则基于支持向量机 (Support Vector Machine, SVM), 根据编码单元像素梯度信息、率失真代价以及块编码标志等信息构建了分

类器来决策当前编码单元是否继续递归向下划分。文献3 (M.Xu, T.Li, Z.Wang, X.Deng, R.Yang and Z.Guan, Reducing Complexity of HEVC: A Deep Learning Approach, in IEEE Transactions on Image Processing, vol.27, no.10, pp.5044-5059, Oct.2018.) 则基于HEVC使用卷积神经网络 (Convolutional Neural Networks, CNN) 和长短期记忆网络 (Long Short-Term Memory, LSTM) 通过分析编码单元划分深度在时域上的相关性。因为时域上越接近, 则其编码单元划分深度的相似性就可能越高。该文献通过构建帧间训练集, 将编码单元残差输入CNN, 利用LSTM学习前后帧在编码单元划分深度上的相关性来对当前编码单元的划分进行提前预测。

发明内容

[0004] 本发明的目的在于提供一种VVC帧间编码单元预测模式提前决策以及块划分提前终止方法, 通过分析视频帧之间的运动变化, 经过对编码深度为1的帧间编码单元中各个预测模式之间率失真代价的比较以及所选择的最优模式和后续编码深度中帧间预测模式 (Inter_2N×2N) 的编码代价的比较, 对于符合条件的编码单元直接跳过对帧间预测模式编码代价的检查, 从而降低计算复杂度。另外还通过对编码单元当前编码深度和所选择的最优预测模式以及各预测模式之间率失真代价的比值关系来预测当前编码单元的最优划分尺寸, 通过对编码单元划分的提前终止, 来避免计算复杂度极高的递归划分过程, 从而降低其编码器的计算复杂度, 在几乎不影响视频编码质量的前提下, 可以有效的减少编码器的耗时, 提高编码效率。

[0005] 本发明技术方案:

[0006] (1) 进行帧间编码单元预测模式选择时, 在深度为1的编码单元中, 根据所选择的最优预测模式以及当前Merge预测模式与帧间预测模式 (Inter_2N×2N) 的之间的率失真代价的关系来决策是否跳过后续编码单元中的帧间预测模式。

[0007] (2) 进行帧间编码单元块划分决策时, 首先统计了编码单元的最终编码深度和所选择的最优预测模式之间的关系, 统计结果显示两者具有很强的相关性。因此我们又基于对各预测模式的率失真代价的比值 MI_{rate} 来辅助决策是否需要当前编码单元的划分进行终止。 MI_{rate} 的计算如公式 (1) 所示, 其中 $Merge_{RDC}$ 表示的是当前编码单元使用Merge预测模式的率失真代价, $Inter_{RDC}$ 则表示是使用帧间预测模式 (Inter_2N×2N) 的率失真代价。

$$[0008] \quad MI_{rate} = \frac{Merge_{RDC}}{Inter_{RDC}} \quad \text{公式 (1)}$$

[0009] 采用上述方案, 本发明的有益效果是:

[0010] 1. 本发明通过对不同编码深度下帧间预测模式率失真代价大小的比较, 在满足条件的编码单元中将帧间预测模式的检查跳过, 从而去除模式搜索中的冗余计算, 降低其计算复杂度。

[0011] 2. 本发明通过对不同类型的视频序列中编码深度和其最优预测模式之间存在的相关性进行分析, 发现其和不同预测模式之间率失真代价的比值存在一定的相关性, 从而快速的终止当前编码单元的递归划分。

[0012] 3. 本发明综合考虑了360度全景视频的特点, 可以有效的降低在预测模式搜索和编码单元最优划分方式决策中存在的冗余计算, 从而可以在几乎不损失编码质量的前提

下,提高VVC编码器的编码效率。

附图说明

- [0013] 图1是在不同的编码深度下各层帧间预测模式率失真代价比较的情况。
[0014] 图2是在不同的编码深度下编码单元选择Merge为最优预测模式的百分比。
[0015] 图3是最终的实验结果并和标准算法进行对比。
[0016] 图4是本发明VVC帧间快速算法的整体流程图。

具体实施方式

[0017] 本发明的主要思想是利用视频帧之间运动变化关系,通过对不同编码深度下帧间预测模式的编码代价的比较来决定是否跳过后续编码深度下对帧间预测模式的检查。另外通过对编码单元中所选择的最优预测模式和不同预测模式率失真代价的比值关系来对当前编码单元的划分进行提前终止。通过这两方面的优化来降低VVC帧间编码的计算复杂度。本发明实施例所使用的测试视频序列均为4K视频(分辨率:3840×1920),原始视频均为mp4格式,处理为YUV420格式后使用,视频来自于(<http://zy.dmgeek.com/c/vrshipin>)。

[0018] 本发明首先统计了当在编码深度为1的编码单元中选择了Merge为最优预测模式,后续各编码深度下使用帧间预测模式时,各编码层次之间率失真代价的差异性。统计结果参加图1,其中 $RDC_{depth, depth=0, \dots, 5}$ 表示在当前最大编码单元(Largest Coding Unit, LCU)中各编码深度下每一层所有编码单元使用帧间预测模式的率失真代价总合。 RDC_{skip} 为是否跳过后续帧间预测模式的阈值。可以看出,当编码深度为1的编码单元选择了Merge为最优预测模式,对于四个不同的测试序列,后续的帧间预测模式平均只有0.72%的比例会优于 $RDC_1 \times RDC_{skip}$ (0.923)。也就是说,如果在深度为1的编码单元中,如果当前的Merge预测模式的率失真代价小于帧间预测模式的0.923倍,那么在后续的编码单元中进行更加精细的运动搜索过程也几乎不会超过当前Merge预测模式的编码效果。

[0019] 同时我们也统计了在不同的编码深度下,所选择的最优预测模式和编码单元的划分深度之间的关系,统计结果如图2。从统计结果中可以看出:当编码深度为0时,平均约有98.04%的帧间编码单元选择了Merge为最优预测模式。而当编码深度为1时,这一比例为77.07%。这说明了帧间编码单元所选择的最优预测模式和其否继续递归划分有较大的关系,因此可以据此来提前终止帧间编码单元的继续递归划分。

[0020] 基于本发明的技术方案思想,并结合附图4所示流程图对本发明做进一步的说明。

[0021] 步骤1:基于VVC的通用测试平台JEM7.0。开始编码一个LCU后,首先判断当前的编码深度(初始为0,每递归划分一次编码深度加1)。如果编码深度为0则转步骤2,否则转步骤4。

[0022] 步骤2:判断当前编码单元的最优预测模式是否为Merge,如果是则转步骤3,否则转步骤9。

[0023] 步骤3:计算 MI_{rate} ,如果 $MI_{rate} < MI_{Threshold}$ (编码深度为0和1时, $MI_{Threshold}$ 分别为0.63和0.95),则停止当前编码单元继续向下划分,否则转步骤9。

[0024] 步骤4:如果当前编码深度为1,则转步骤5,否则转步骤7。

[0025] 步骤5:获得Merge预测模式的率失真代价 RDC_{merge} 以及帧间预测模式的率失真代价

RDC₁,并设置RDC_{skip}=0.923。转步骤6。

[0026] 步骤6:如果RDC_{merge}<RDC₁×RDC_{skip},将变量SkipInter设为真。否则转步骤3。

[0027] 步骤7:检查变量SkipInter是否为真。如果为真,则跳过当前编码深度中对帧间预测模式的检查,否则转步骤8。

[0028] 步骤8:继续按照JEM7.0标准流程进行。

[0029] 步骤9:继续按照JEM7.0标准流程进行直到结束当前LCU的编码。

[0030] 最后,基于JEM7.0平台对本发明中提出的快速算法进行了验证,并和标准算法进行了对比。可以看出在基本未损失编码效率的情况下,平均可以节省30.34%的编码时间,统计结果参见图3。

[0031] 上述描述仅是对本发明较佳实施例的描述,并非是对本发明范围的任何限定。任何熟悉该领域的普通技术人员根据上述揭示的技术内容做出的任何变更或修饰均应当视为等同的有效实施例,均属于本发明技术方案保护的范围。

序列	$RDC_{depth} < RDC_1 \times (RDC_{skip} = 0.923)$	$RDC_{depth} < RDC_1 \times (RDC_{skip} = 1)$
Cat_4K	0.06	23.42
Dog_4K	0.33	27.62
Field_4K	1.93	36.31
Robot_4K	0.54	5.37
平均	0.72	23.09

图1

序列	深度	选择 Merge	当前深度占总体百分比
Cat_4K	0	99.38	95.67
	1	65.70	2.99
	2	32.46	1.17
	3	18.86	0.15
Dog_4K	0	99.23	89.89
	1	84.55	5.88
	2	67.23	2.95
	3	70.67	1.08
Field_4K	0	93.56	83.30
	1	82.55	13.07
	2	63.24	3.10
	3	20.90	0.48
Robot_4K	0	100	92.70
	1	75.46	3.80
	2	72.73	1.97
	3	46.76	1.53
平均	0	98.04	90.39
	1	77.07	6.44
	2	58.92	2.30
	3	39.30	0.81

图2

序列	∇T	$\nabla Bitrate$	$BD-Rate$
Beach_4K	18.60	-0.34	0.34
Cat_4K	28.03	0.96	0.79
Dog_4K	32.04	0.68	0.91
Field_4K	18.81	-0.26	1.46
Robot_4K	42.35	0.84	0.25
Sea_4K	36.27	2.00	1.60
Surfing_4K	36.26	2.14	1.07
平均	30.34	0.86	0.92

图3

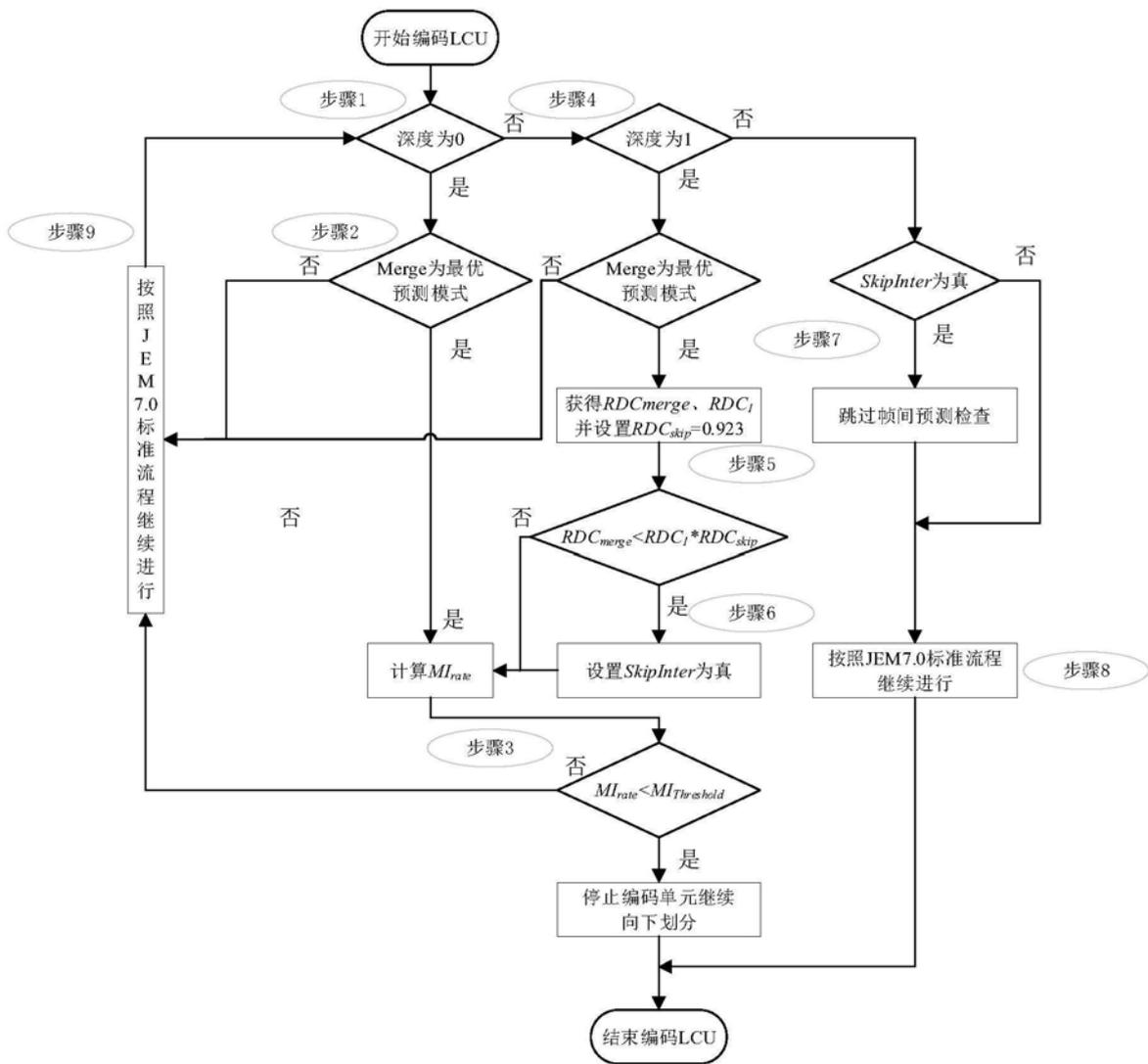


图4