

一种基于统计分析的多视点视频编码中 Direct 模式提前终止算法

王 萍¹, 刘 琪¹, 张 磊²

(1 西安交通大学 电子与信息工程学院, 陕西 西安 710049; 2 北京理工大学 计算机学院, 北京 100081)

摘 要: 提出了一种针对多视点视频编码中 Direct 模式的提前终止算法, 该算法基于当前编码宏块相邻视点、时间及空间方向的已编码邻近宏块的编码信息, 提出了一个描述当前宏块的邻近块代价影响因子, 通过与阈值比较可对多数编码宏块直接选择 Direct 模式进行编码, 从而略过其他模式的相关计算. 实验结果表明, 与原始的多视点视频编码算法相比, 该算法可平均减少约 88.2% 的编码时间, 同时峰值信噪比仅下降约 0.29 dB, 编码比特率增加约 0.68%.

关键词: 多视点视频编码; Direct 模式; 率失真代价; 提前终止

中图分类号: TP919.8

文献标识码: A

文章编号: 1000-7180(2017)01-0095-07

Early Direct Mode Decision for Multi-view Video Coding Based on the Statistical Analysis

WANG Ping¹, LIU Qi¹, ZHANG Lei²

(1 School of Electronic and Information Engineering, Xian Jiaotong University, Xian 710049, China;
2 School of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081, China)

Abstract: In order to reduce the encoder's complexity, an early direct mode decision for MVC is proposed in this paper. Based on the best modes and their rate-distortion costs of encoded neighboring macroblocks which are located in the spatial, temporal and inter-view directions, the influence factor of neighboring macroblocks is used to describe the current macroblock. By comparing this influence factor with the threshold which is obtained by statistical analysis of error selection probability, the proposed algorithm directly select the direct mode for some macroblocks and ignore the calculation of the other remaining modes. The experimental results show that the proposed algorithm can achieve 88.2% of encoding time saving on average with little loss of coding performance with the reference full search encoder.

Key words: multi-view video coding (MVC); direct mode; rate-distortion cost; early decision

1 引言

多视点视频是由位置固定的多个摄像机在不同角度同时对同一场景进行拍摄而得到的一组视频序列, 在带给人们生动的画面感与立体感的同时, 拍摄相机数目的增加也带来了海量的数据, 这对多视点视频的传输与存储带来了较大挑战, 因此对多视点

视频进行有效压缩成为近年来的研究热点. 国际标准组织 JVT(Joint Video Team)公布的多视点视频编码^[1] (Multi-view Video Coding, MVC) 标准, 虽然获取得了良好的编码效率, 但计算复杂度极高, 导致了严重的编码耗时, 这成为了制约其推广应用的主要瓶颈.

为了降低计算复杂度, 针对多视点视频编码的

不同耗时模块,学者们提出了各类优化算法.袁梅冷^[2]等人在运动估计快速搜索算法的基础上,结合信息重用技术,对模式集合进行划分,通过同一集合内部小尺寸块重用大尺寸块的参考帧索引,减少多参考帧循环搜索次数. Zhang^[3]等人利用宏块之间的空间相关性,由已编码大尺寸宏块的参考帧信息推导小尺寸宏块的参考帧. 贺人龙^[4]等人通过分析序列的相关性,将图像序列分割成不同的独立片,通过剔除影响较小的参考帧来优化参考帧选择.

多视点视频编码时采用率失真优化技术结合穷尽搜索模式方法,需要对当前帧的每个编码宏块在时间与视点方向上的所有模式进行遍历搜索,并计算每种模式的率失真代价,然后找出率失真代价最小的模式作为最优编码模式,这个过程需要消耗编码的大量时间^[5-8]. 如果能够提前缩减候选模式的范围,只对部分模式计算率失真代价,就可以降低编码器中模式选择的计算量.

快速模式选择是一种减少编码器计算复杂度的有效方法,本文在分析多视点视频编码最优模式分布的基础上,结合空间、时间及视点三个方向已编码邻近宏块的最优模式信息,提出了一个表征当前编码宏块的邻近块代价影响因子,并通过统计分析了该影响因子与宏块是否采用直接预测模式的关系,得到了可对编码宏块提前终止 Direct 模式的判别阈值.

2 多视点视频编码模式分析

多视点视频编码中,需要将预测编码、多参考帧技术以及视点间的预测相结合. 为了达到良好的编码性能,JVT 采用了由德国 HHI 研究机构提出的分层次 B 帧预测结构(Hierarchical B Prediction, HBP),如图 1 所示.

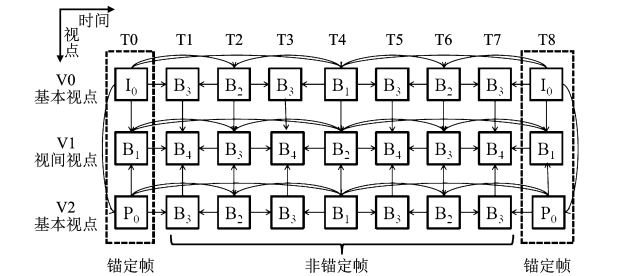


图 1 多视点视频编码中分层次 B 帧预测结构(HBP)

图 1 中, V0, V1 和 V2 为 3 个不同的视点, HBP 将视点分为基本视点和视间视点. 每个视点的图像组(Group of Pictures, GOP)长度取值可以为

8,12,15,本图中 GOP 为 8,对应 T0~T8 这 9 个不同的时刻. HBP 将不同时刻对应的编码帧分为锚定帧与非锚定帧. 对锚定帧而言,帧的重建只需要参考同一时刻对应的参考帧,如 T0 与 T8 时刻所有视点对应的帧. 而对于非锚定帧,其重建还需要参考其他时刻的参考帧. 基本视点 V0, V2 中的非锚定帧仅利用同一视点内时间方向的相邻帧进行预测估计;而视间视点 V1 除了采用同一视点内的运动估计,还需要进行视点间的预测. 由此可见, HBP 结构同时兼顾了时间与视间方向的相关性.

本文针对视间视点的非锚定帧的快速编码展开研究,如图 1 所示,这些帧的所有图像都属于 B 帧预测编码. MVC 编码 B 帧时支持多个变尺寸块的预测模式,其中包括 8 种帧间预测模式(Direct, Inter16×16, Inter16×8, Inter8×16, Inter8×8, Inter8×4, Inter4×8, Inter4×4)和 3 种帧内预测模式(Intra16×16, Intra8×8, Intra4×4). 编码时与 H. 264/AVC 一样采用了率失真优化技术,穷尽搜索所有的预测模式,最终选择具有最小率失真代价的模式作为最优编码模式. 由于编码器需要对每一宏块的所有模式都经过变换、量化、熵编码等过程计算对应的率失真代价,因此全搜索算法的计算复杂度非常高. 如果可以提前缩减候选模式的范围,就可以降低计算复杂度.

此外,预测 B 帧时的 Direct 模式无需进行复杂的运动估计和视差估计,它是直接由先前已编码宏块的运动信息推导当前宏块的运动矢量,因此 Direct 模式无需传输运动矢量残差与参考索引信息,它在所有模式中计算量最小. 如果能提前判定 Direct 模式是最优模式的话,就可以跳过其余所有模式,这样大量的运动估计和视差估计的计算量就可以节省下来,从而降低多视点编码中模式选择的计算量.

3 基于统计分析的 Direct 模式提前终止算法

3.1 Direct 模式分布统计

一般来说,尺寸较大的分块模式通常适合于编码纹理简单均匀或者运动缓慢的区域,而尺寸较小的分块模式较适合编码纹理复杂或快速运动的区域. 在自然场景视频序列中,相邻时间及视点的预测帧之间应绝大多数属于运动缓慢区域,因此可能多数宏块都会选择大尺寸的 Direct 模式或 Inter16×16 模式作为最优编码模式. 如前所述, Direct 模式是所有模式中计算量最小的模式,因此如果能提前

判定 Direct 模式是最优模式的话,将极大地降低编码中模式选择的计算量.为此,我们先去统计多视点视频编码中 Direct 模式是最优模式的概率.

统计实验在通用的多视点视频编码平台 JM-VC8.5 上针对视间视点的非锚定帧进行,测试中选

取了几个具有不同分辨率、纹理复杂度和运动特征的视频序列,采用 MVC 中的原始全搜索算法,统计了在不同量化参数(Quantization Parameter, QP)下 Direct 模式是最优模式的概率,结果如表 1 所示.

表 1 MVC 在全搜索算法中 Direct 模式是最优模式的宏块比例统计 %

	Ballroom (640×480)	Exit (640×480)	Vassar (640×480)	Balloons (1 024×768)	BookArrival (1 024×768)	Newspaper (1 024×768)
QP=25	74.5	85.2	88.9	83.4	88.1	91.6
QP=29	78.7	88.7	93.9	88.8	90.9	93.8
QP=33	81.7	90.5	95.2	91.7	92.4	95.8
QP=37	84.5	92.3	94.9	94.3	93.1	97.0
平均值	79.9	89.2	93.2	89.6	91.1	94.6

从表 1 可以看出,对各种不同的视频序列,相较于其他所有模式,Direct 模式在全搜索算法中被选作最优模式的可能性都是最高的,平均达到了 79.9%~94.6%.而且随着 QP 的增大,Direct 模式是最优模式的可能性也随之增大,对运动缓慢的“Newspaper”序列,在量化参数为 37 时,有高达 97%的宏块的最优模式都是 Direct 模式.这说明了在对视间视点编码时,大多数宏块的最优模式都可能是 Direct 模式,而 Direct 模式的计算量最小,因此设计有效的提前终止 Direct 模式的方法可以大幅度地降低视间视点的编码计算复杂度.

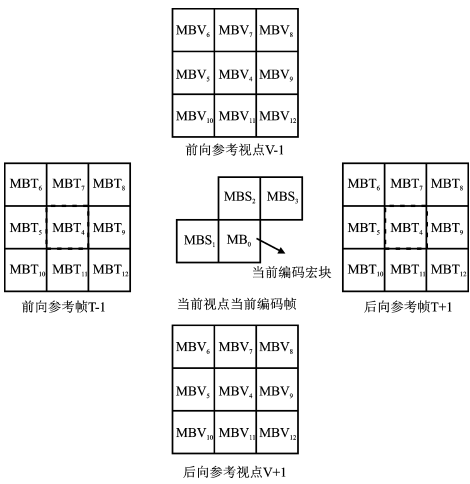
3.2 邻近块代价影响因子

由于多视点视频序列存在着大量空间、时间以及视点间的冗余,因此当前编码宏块应与其对应的空间、时间和视点间三个方向的邻近宏块高度相关.比如在已编码过的邻近宏块中,多数宏块的最优模式都是 Direct 模式,那么当前编码宏块选择 Direct 模式编码的可能性就很高.图 2 为当前编码宏块的邻近宏块示意图. MB_0 是视间视点中当前帧的待编码宏块,其空间已编码邻近宏块为 MBS_1 、 MBS_2 和 MBS_3 . 时间方向的已编码邻近宏块为 MBT_4 及其对应的 8 邻域宏块,其中 MBT_4 为当前宏块在同一视点中前向和后向参考帧中处于同一位置的宏块.视点间方向的已编码邻近宏块为 MBV_4 及其对应的 8 邻域宏块,其中 MBV_4 为当前宏块在两个参考视点中的对应宏块,它通过全局视差向量来定位.

为了充分利用视频序列中的空间、时间和视点间的相关性,我们想通过邻近宏块的已编码信息对当前编码宏块进行描述,这些信息在编码当前宏块时都可直接获得,包括邻近宏块的最优模式的选择及其对应的率失真代价(Rate-distortion cost, RD-

cost)等.最优模式的率失真代价反映了该宏块编码时的综合性能,相邻宏块的率失真代价应该差异不大,因此我们使用邻近宏块最优模式编码时的率失真代价来描述当前宏块.为了充分考虑到各个邻近宏块最优模式之间的差异,我们统计各种最优模式在邻近宏块中所占的比例.如果某种编码模式在已编码邻近宏块中占的比例大,那说明邻近块中这种模式作为最优模式的概率大,此模式对当前编码宏块的影响也应该较大,该模式的信息对当前宏块的参考价值就越大;反之,如果某种编码模式占的比例小,说明该模式对当前编码宏块的影响也较小.因此,对于当前编码宏块,以其邻近宏块中最优模式所占比例为加权系数,通过对邻近宏块最优模式的率失真代价的加权求和构造出邻近块代价影响因子.

在空间方向上有 3 个已编码的邻近宏块,基于它们的最优模式及其对应的率失真代价,通过式(1)可以得到当前宏块在空间方向的邻近块代价影响因



子 $Rdfactor_s$.

$$Rdfactor_s = \frac{1}{3} \sum_{i=1}^3 \alpha(\text{mode}(\text{MBS}_i)) \cdot \text{Rdcost}(\text{MBS}_i) \tag{1}$$

式中, $\text{MBS}_i (i = 1, 2, 3)$ 为空间相邻的 3 个宏块, Rdcost 是邻近宏块最优模式编码时的率失真代价值, mode 是邻近宏块的最优模式, 率失真代价的加权系数是最优模式在邻近宏块中所占的比重, 由下式计算:

$$\alpha(\text{mode}_i) = \frac{N_{\text{mode}}}{N_s + N_t + N_v} \tag{2}$$

式中, N_{mode} 为在所有邻近宏块中最优模式是 mode 的宏块总数, N_s, N_t, N_v 分别为空间、时间及视点方向的邻近宏块总数, 根据图 2 可得, $N_s = 3, N_v = N_t = 18$.

在时间方向上, 因为当前宏块与前向和后向时间相邻参考帧中同一位置及其相邻的宏块之间有很强的时间相关性, 所以对这些邻近宏块的最优模式的率失真代价进行加权求和, 可以得到当前宏块在时间方向的邻近块代价影响因子 $Rdfactor_t$, 如式 (3) 所示. 式中, $T-1$ 和 $T+1$ 表示前向和后向的两个时间参考帧, $\text{MBT}_i (i = 4 \cdots 12)$ 表示参考帧中对应位置的宏块及其 8 邻域宏块.

$$Rdfactor_t = \frac{1}{N_t} \sum_{i=4}^{12} \left[\alpha(\text{mode}(\text{MBT}_i)) \cdot \text{Rdcost}(\text{MBT}_i) \Big|_{T-1} + \alpha(\text{mode}(\text{MBT}_i)) \cdot \text{Rdcost}(\text{MBT}_i) \Big|_{T+1} \right] \tag{3}$$

同理, 多视点视频中处于同一时刻的几个相邻视点的图像通常具有相同的内容, 因此, 当前宏块与其对应的视点间的相邻宏块之间也具有很强的视点间相关性, 据此得到当前宏块在视间方向的邻近块代价影响因子 $Rdfactor_v$, 如式 (4) 所示. 式中, $V-1$ 和 $V+1$ 表示两个参考视点, $\text{MBV}_i (i = 4 \cdots 12)$ 表示参考视点中对应的相邻宏块及其 8 邻域宏块.

$$Rdfactor_v = \frac{1}{N_v} \sum_{i=4}^{12} \left[\alpha(\text{mode}(\text{MBV}_i)) \cdot \text{Rdcost}(\text{MBV}_i) \Big|_{V-1} + \alpha(\text{mode}(\text{MBV}_i)) \cdot \text{Rdcost}(\text{MBV}_i) \Big|_{V+1} \right] \tag{4}$$

以上考虑到多视点视频中存在的空间、时间及视点间的冗余, 分别计算了 3 个方向的邻近块代价影响因子. 位于不同运动特征区域的待编码宏块, 其各个方向的相关性各不相同. 例如, 对于包含剧烈运动的宏块, 其空间相关性可能相对较高; 包含在明显场景切换中的宏块, 其视点间相关性应该较高; 而对

于处在背景或运动平缓的宏块, 其时间相关性更高. 这里为了对具有不同运动特征的宏块进行统一的描述, 我们将当前编码宏块的总的邻近块代价影响因子定义为三个方向影响因子的平均值, 如下式:

$$Rdfactor = \text{mean}(Rdfactor_s, Rdfactor_t, Rdfactor_v) \tag{5}$$

综上, 对于当前待编码宏块, 通过统计其空间、时间和视点三个方向已编码的邻近宏块的最优模式, 并对其最优模式的率失真代价进行加权求和, 得到了描述当前编码宏块的邻近块代价影响因子.

3.3 阈值的设定

在得到当前编码宏块的邻近块代价影响因子后, 如果可以知道该因子与宏块最终是否判定为 Direct 模式的对应关系, 我们就可以在宏块进行模式遍历前, 提前决定宏块是否仅需要采用 Direct 模式进行编码, 就可以节省大量计算量.

当 Direct 模式是最优编码模式时, 编码器不需要传输运动矢量残差信息和参考索引信息, 因此其率失真代价值较小. 如果邻近块中采用 Direct 模式的宏块数较多, 通过加权求和得到的当前编码宏块的邻近块代价影响因子应该也较小. 因此, 邻近块代价影响因子越小, 当前宏块应该越有可能采用 Direct 模式编码. 为此, 我们通过对测试样本的统计去观察邻近块代价影响因子与宏块最优编码模式是否是 Direct 模式的关系. 测试样本通过采用 JMVC8.5 中的全搜索算法对 6 个具有不同运动特性与纹理特征的视频序列编码获得, 编码参数如表 2 所示, 在编码过程中保存非锚定帧中每个宏块的邻近块代价影响因子 $Rdfactor$ 的值和其对应的是否为 Direct 模式的标志作为样本数据. 表 2 中的图像组长度设为 12, 对每个视频序列编码 11 个非锚定帧, 6 个视频序列包括 3 个分辨率为 640×480 的序列和 3 个 1024×768 的序列, 因此在每个量化参数下得到的宏块样本数为 $3 \times 40 \times 30 \times 11 + 3 \times 64 \times 48 \times 11 = 140\,976$ 个. 6 个量化参数下的总样本数是个 845 856 个.

表 2 获取测试样本的编码器参数设置

BasicQP	17, 21, 25, 29, 33, 37
IntraPeriod	12
NumberReferenceFrames	2
SearchRange	64
SymbolMode	CABAC
SearchMode	TZsearch

图 3 给出了在不同量化参数下宏块的邻近块代价影响因子与全搜索算法中宏块最优编码模式是否是 Direct 模式的概率直方图统计结果,图中横坐标为影响因子取值,纵坐标为在该影响因子下选择 Direct 模式与非 Direct 模式的宏块比例.从统计结果可以看出,邻近块代价影响因子与宏块的最优模式选择有较强的相关性.当 Rdfactor 值较小时,绝大多数宏块的最优模式都是 Direct 模式,而只有较小比例的宏块选取了其他模式编码.随着 Rdfactor 值的增大,选择非 Direct 模式编码的宏块比例也随之增大.据此我们想要找到一个阈值,当影响因子低

于该阈值时,将宏块的编码模式直接判定为 Direct 模式,而忽略其他模式的计算,从而节省模式选择过程的计算量;而当影响因子高于阈值时,非 Direct 模式被选择的可能性较大,因此不能直接终止 Direct 模式,此时需要对所有候选模式进行编码.

为了确定合适的阈值,我们首先选取最优模式为 Direct 模式的宏块数所占比例为 50%时所对应的 Rdfactor 值为初始阈值 Th_0 .这是因为比该阈值小的邻近块代价影响因子值所对应的宏块,其 Direct 模式的可能性都高于非 Direct 模式.

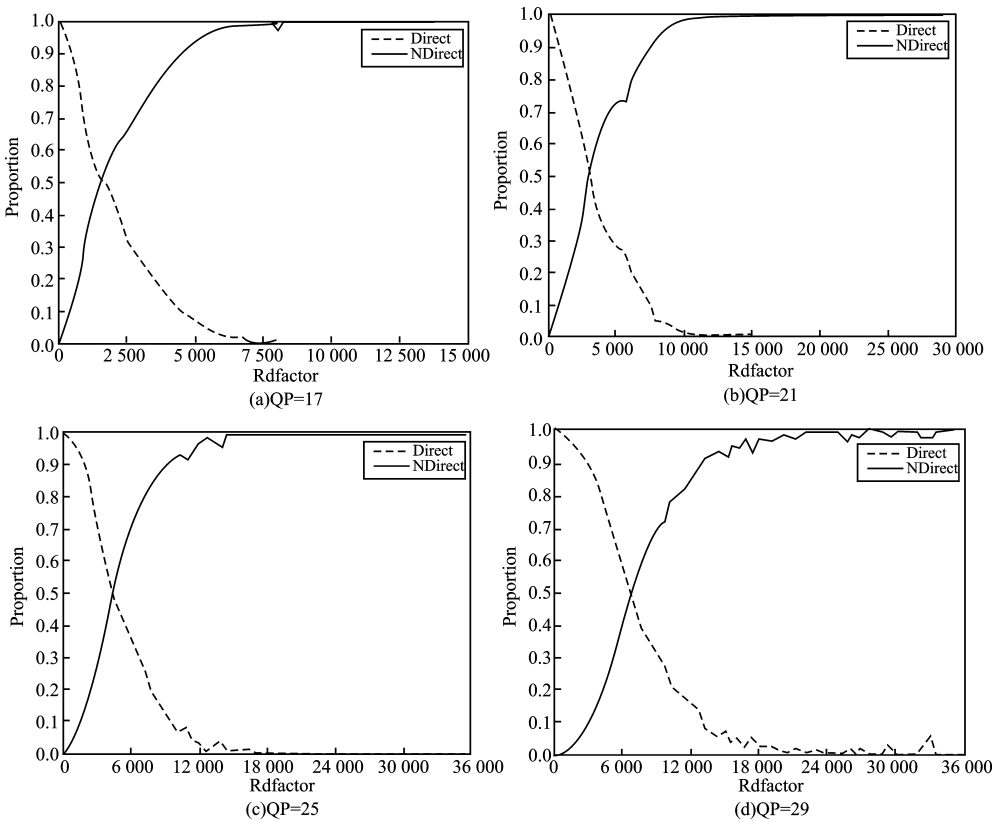


图 3 不同 Rdfactor 取值下宏块是 Direct 模式与非 Direct 模式的分布统计直方图

比较邻近块代价影响因子与初始阈值,当待编码宏块的 Rdfactor 值小于设定的初始阈值时,宏块模式被提前判定为 Direct 模式,此时无需再进行其他模式的遍历;而当待编码宏块的 Rdfactor 值高于设定初始阈值时,宏块不能被提前判定为 Direct 模式,而是需要遍历所有的候选编码模式.值得注意的是,初始阈值只是要求当 Rdfactor 值低于该阈值时,Direct 模式是最优模式的可能性高于 50%,但因为每个 Rdfactor 值对应的宏块数目不同,仅采用此阈值有可能造成非 Direct 模式被误判为 Direct 模式的宏块数目较多,从而导致编码的信噪比性能

下降过大.因此接下来我们通过对错误概率的分析进一步确定阈值.与阈值相关的错判概率如式(6)所示,如果 Rdfactor 值小于阈值,但全搜索算法中最优模式却不是 Direct 模式时,统计这些宏块的比例,计算出错误概率.

$$P(\text{error} | Th_0) = P(\text{mode} \neq \text{Direct} | \text{Rdfactor} \leq Th_0) \tag{6}$$

当改变阈值时,错误概率也随之改变.由图 3 可知,当阈值越大时,此时被判为 Direct 模式的宏块数目越多,节省的计算量也越大,但同时错误概率也越大,被误判为 Direct 模式的宏块数目越多,从而

影响了编码的率失真性能. 反之, 设定的阈值越小时, 被提前终止模式选择的宏块数目也越少, 节省的编码时间有限, 但被错判的宏块数目也越少, 对编码性能的影响也越小. 因此阈值的大小将直接影响计算复杂度和编码效率之间的平衡. 本文实验中选定错误概率低于 3% 对应的 Rdfactor 值为最终阈值 Th, 如表 3 所示.

表 3 根据错误概率为 3% 确定的提前终止阈值

QP	17	21	25	29	33	37	41
Th	279	634	1 749	4 663	10 477	16 199	25 016

从表 3 可以看到, 阈值随着量化参数的增大而增大, 这说明量化参数增大时, 会有更多的宏块被提前终止为 Direct 模式. 最后我们采用最小二乘法多项式拟合方法得到判别阈值与量化参数间的关系, 如图 4 所示, 定量关系如式(7)所示.

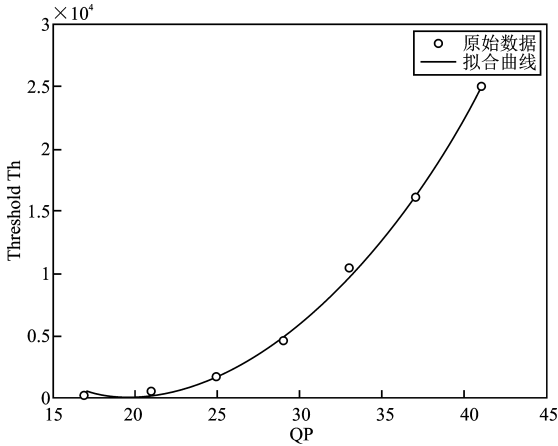


图 4 阈值 TH 与量化参数 (QP) 之间的关系

$$\begin{aligned} Th &= 52.94QP^2 + 2\,051.8QP + 20\,026.0 \\ 17 &\leq QP \leq 45 \end{aligned} \tag{7}$$

综上, 提出的 Direct 模式提前终止算法描述如下所示.

第 1 步: 检查当前宏块是否位于锚定帧上, 若是, 则执行全搜索算法; 否则, 执行第 2 步;

第 2 步: 确定当前编码宏块的空间邻近宏块、时间与视点方向的前后邻近宏块及其 8 邻域宏块 (如图 2 所示), 并根据这些邻近宏块的率失真代价值与最优编码模式信息, 计算当前编码宏块的邻近块代价影响因子 Rdfactor;

第 3 步: 根据当前编码量化参数 QP 的取值, 根据式(7)计算提前终止的判别阈值 Th;

第 4 步: 将第 2 步中得到的当前编码块的邻近块代价影响因子 Rdfactor 与判别阈值 Th 相比较, 如果 $Rdfactor \leq Th$, 则当前宏块模式直接判定为

Direct 模式, 终止模式搜索过程; 否则, 执行全搜索算法确定最优编码模式.

4 实验结果

为了测试本文的快速视间视点编码算法的性能, 我们在 JMVC8.5 参考编码器中实现了本文及对比算法 (Wang^[5]). 仿真实验中测试序列均使用了 3 路视点, 其中第 1 路和第 3 路用做基本视点, 第 2 路采用视间视点编码. 编码器的主要参数设置如下: GOP 取值为 12; 编码两个 GOP (即 24 帧); 量化参数选取 25, 29, 33 和 37; 运动搜索采用 TZSearch 搜索算法, 搜索范围设置为 64. 实验中选用了 4 个不同分辨率的测试视频序列 Vassar (640×480)、Ballroom (640×480)、Doorflower (1024×768) 和 Pozon Hall (1920×1088), 这些序列中同时包含了缓慢运动、剧烈运动等不同的运动特征以及简单纹理、复杂纹理等不同纹理特征.

评价快速编码算法的性能时, 平均峰值信噪比和比特率两个指标用来衡量算法的率失真性能, 总编码时间用来衡量算法的计算复杂度. 所有算法均与 JMVC8.5 的原始全搜索算法相比较, 得到平均峰值信噪比的减少量 ($\Delta PSNR$)、平均比特率的增加量 (ΔBR) 和平均编码时间的减少量 (ΔT), 其定义如下:

$$\begin{cases} \Delta PSNR = PSNR_{\text{proposed}} - PSNR_{\text{JMVC}} \\ \Delta BR = \frac{\text{Bitrate}_{\text{proposed}} - \text{Bitrate}_{\text{JMVC}}}{\text{Bitrate}_{\text{JMVC}}} \times 100\% \\ \Delta T = \frac{T_{\text{proposed}} - T_{\text{JMVC}}}{T_{\text{JMVC}}} \times 100\% \end{cases} \tag{8}$$

式中, $PSNR_{\text{JMVC}}$ 、 $\text{Bitrate}_{\text{JMVC}}$ 和 T_{JMVC} 分别表示 JMVC 中全搜索算法的峰值信噪比、比特率和编码时间, $PSNR_{\text{proposed}}$ 、 $\text{Bitrate}_{\text{proposed}}$ 和 T_{proposed} 分别表示快速算法 (本文算法及对比算法) 的峰值信噪比、比特率和编码时间.

对比算法的测试结果可以看到, 与原始的 JMVC 全搜索算法相比, 对于不同分辨率的视频序列, 本文算法带来的编码性能的损失都是较小的, 其中峰值信噪比平均下降 0.29 dB, 编码比特率平均增加 0.68%. 同时编码时间平均减少了约 88.2%, 这说明所提算法有效地减少了计算量. 对于 Ballroom 序列, 其背景慢速运动、没有旋转与缩放, 但前景中速运动, 纹理复杂, 算法对此类序列直接选中 Direct 模式的宏块数稍少, 因此编码时间的减少量相对最少, 但其平均时间节省也可达 80% 左右. 对于 Door-

flower 序列,该序列背景静止,前景运动平缓,视点间与空间的相关性都较强,直接预测模式所占比例较高,因此快速算法获得的时间节省最多. Wang^[5]的算法,其峰值信噪比的下降与比特率的增加相对较低,与本文算法相比,虽然其率失真性能更优,但其时间节省率平均为 65.34%,比本文算法低了约 22.9%,因此本文算法对编码时间减少的优势是很明显的.

综上,本文提出的 Direct 模式提前终止算法,能够在保持编码性能较小损失的前提下,有效地减少视间视点编码的计算复杂度,缩短编码时间. 尤其对于运动相对平缓的序列,时间、空间和视点间的相关性较强,可以提前终止 Direct 模式的宏块数较多,本文算法会获得更多的编码时间节省.

5 结束语

针对多视点视频编码中模式选择的高复杂度,本文提出了一种基于统计分析的 Direct 模式提前终止算法. 根据时间、空间和视点三个方向相邻已编码宏块的编码信息,得到当前宏块的邻近块代价影响因子,并利用测试样本集中影响因子与 Direct 模式选择的对应关系,设定了与量化参数相关的提前终止门限值. 提出的算法能对多数宏块提前选择 Direct 模式,而略过其余模式的计算,从而节省了编码过程中的大量计算量. 实验结果表明该提前终止算法能够在保持编码性能基本不变的情况下,有效地减少编码时间.

参考文献:

[1] Vetro A, Wiegand T, Sullivan G J. Overview of the stereo and multiview video coding extensions of the H. 264/MPEG-4 AVC standard [J]. Proceedings of the IEEE, 2011, 99(4): 626-642.

[2] 袁梅冷, 杨张, 雷海军. 多视点视频编码模式决策快速算法[J]. 小型微型计算机系统, 2015, 36(3): 582-585.

[3] Zhang Yun, Sam Kwong, Jiang Cangyi, et al. Efficient multi-reference frame selection algorithm for hierarchical B pictures in Multi-view video coding [J]. IEEE Transactions on Broadcasting, 2011, 57(1): 15-23.

[4] 贺人龙, 蒋刚毅, 符冉迪, 等. 基于自适应参考帧模式选择的多视点视频编码方法[J]. 计算机辅助设计与图形学学报, 2007, 40(9): 1133-1137.

[5] Wang Fengsui, Zeng Huangqiang, Shen Qinghong, et al. Efficient early direct mode decision for multi-view video coding[J]. Signal Processing: Image Communication, 2013, 28(7): 736-744.

[6] Ding Cong, Sang Xingzhu, Zhao Tianqi, et al. A fast mode decision algorithm for multiview auto-stereoscopic 3D video coding based on mode and disparity statistic analysis[J]. Proceedings of SPIE on Optoelectronic Imaging and Multimedia Technology, 2012, 8558: 85580P-1-85580P-12.

[7] Zhu Wei, Tian Xiang, Zhou Fan, et al. Fast inter mode decision based on textural segmentation and correlations for multiview video coding [J]. IEEE Transactions on Consumer Electronics, 2010, 56(3): 1696-1704.

[8] Shen Liquan, Liu Zhi, An Ping, et al. Low-complexity mode decision for MVC [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2010, 21(6): 837-847.

作者简介:

王萍女, (1976-), 博士, 副教授. 研究方向为视频编码.
E-mail: ping.fu@xjtu.edu.cn.