

基于 HBP 编码结构的 MVC 快速算法

邓智珩^{*①②} 贾克斌^① 陈锐霖^② 伏长虹^② 萧允治^②

^①(北京工业大学电子信息与控制工程学院 北京 100124)

^②(香港理工大学电子资讯工程系 香港 红磡)

摘要: 针对 MVC 系统中运算量最大的矢量预测部分——单向搜索和双向搜索, 该文提出一种快速算法。根据多视点视频的序列相关性, 采用立体-运动约束模型在运动域和视差域进行联合迭代单向搜索, 设计可信度因子来自适应地调整修正搜索窗口, 同步搜索得到运动和视差矢量。此外, 通过分析双向预测宏块的特性, 提出一种双向搜索提前中止算法。实验结果证明该文算法的有效性, 能在保持编码质量的同时极大减少算法复杂度。

关键词: 多视点视频编码; JMVC; 视差估计; 双向预测

中图分类号: TN919.81

文献标识码: A

文章编号: 1009-5896(2011)08-1955-08

DOI: 10.3724/SP.J.1146.2010.01298

A Fast Algorithm for MVC Using HBP Prediction Structure

Deng Zhi-pin^{①②} Jia Ke-bin^① Chan Yui-lam^② Fu Chang-hong^② Siu Wan-chi^②

^①(Dept. of Electronic Information & Control Engineering, Beijing University of Technology, Beijing 100124, China)

^②(Dept. of Electronic Information Engineering, The Hong Kong Polytechnic University, Hong Kong, China)

Abstract: This paper proposes a fast algorithm particularly for the most computational complexity part in a Multiview Video Coding (MVC) system — uni-directional and bi-directional predictions. Based on the stereo-motion consistency constraint, an iterative search strategy with adaptive refinement search range adjustment technique is proposed for uni-directional prediction, in which motion and disparity vectors can be achieved jointly. Furthermore, an efficient early termination is proposed for bi-directional prediction on the basis of the analysis of multiview video sequences. Experimental results demonstrate the effectiveness of the proposed fast algorithm.

Key words: Multiview Video Coding (MVC); JMVC; Disparity estimation; Bi-directional prediction

1 引言

具有交互性和临场感的多视点视频在自由视点视频, 3D 电视, IMAX 影院, 3D 可视会议等领域展现了广阔的应用前景^[1]。2008 年 7 月, 国际视频小组 JVT 发布了多视点视频编码(Multiview Video Coding, MVC)标准——H.264/MVC, 将其作为 H.264/AVC 标准的多视点高级档次(Multiview High Profile)^[2]。JVT 专门开发了 MVC 的研究平台——JMVC^[3]。JMVC 编码器实现了相邻视点多参考帧选择技术, 利用视差补偿预测来消除相邻视点之间的交叉冗余信息。为了充分消除同一视点间的时间和空间冗余, JMVC 采用目前编码性能最高的分层 B 帧金字塔型预测结构(Hierarchical B Picture, HBP)^[4], 将上一层的 B 帧作为参考帧来预测当前层的 B 帧, 能够节省 20%~30%的比特率。在 JMVC

编码器中, 双向预测 B 帧中的每个宏块都要遍历: 单向搜索(包括前向和后向运动估计, 前向和后向视差估计)和双向搜索, 从中选取使率失真代价(Rate-Distortion Cost, RDCost)值最小的作为当前宏块的最终预测结果。单向搜索和双向搜索过程成为 MVC 系统中复杂度最高的部分^[5]。

国内外众多学者都致力于 MVC 快速算法的研究^[6-13]。文献[6,7]通过分析图像/宏块的特性, 通过判断运动缓慢/平坦区域来设计提前中止算法。文献[8]结合相机的几何特性和对应块的编码信息来设计快速算法减小运动估计(Motion Estimation, ME)或视差估计(Disparity Estimation, DE)的搜索窗口, 达到提高编码速度的目的。文献[9,10]先采用全搜索算法求取视差矢量, 然后利用相邻视点图像帧的关系, 得到一组当前块的候选运动矢量, ME 过程只需要在候选运动矢量附近做小范围的矢量修正, 然而, 这些方法只能加速 ME 或者 DE 过程, 整个 JMVC 编码器的复杂度不能得到大幅度降低。因此, 文献[11,12]针对同时降低 ME 和 DE 的运算复杂度提出快速算法, 以牺牲少量的编码失真度来换取快

2010-11-25 收到, 2011-04-22 改回

国家自然科学基金(30970780)和北京市自然科学基金及教委重点科技项目(KZ200910005005)资助课题

*通信作者: 邓智珩 zhipindeng@gmail.com

速的编码过程,但是,ME 和 DE 被看做两个相对独立的过程,前一步 ME/DE 的结果没有直接被用来提高当前 DE/ME 的预测精度。并且,上述所有 MVC 快速算法都没有考虑 HBP 结构中的双向搜索过程^[14]。在 JMVC 中,除了对当前块进行前向搜索和后向搜索外,还采用双向搜索来寻求更有效的预测结果,并从中选择一个使 RDCost 值最小的作为当前块的最终预测类型。加速双向搜索过程对提升 JMVC 编码器的整体速度起着重大的作用。

鉴于此,本文提出了一种高效的 MVC 快速搜索算法,首先,基于立体-运动约束模型设计适用于单向搜索的运动和视差联合迭代快速算法,通过可信度因子来自适应地调整修正窗口,以较少的迭代次数得到精确的搜索矢量;此外,通过分析多视点视频的特性,提出一种双向搜索提前中止算法,进一步降低编码复杂度。

2 HBP 结构概述

JMVC 编码器采用 HBP 结构作为时序方向的基本预测结构^[4],如图 1 所示,其中,图像组(GOP)的长度为 12。每个 GOP 中除锚定帧之外的所有帧都称为非锚定帧,非锚定帧被分为不同的时序层次(TL),图中用 B_1, B_2, B_3, b_4 表示。当前层的非锚定帧通过参考上层最相邻的两个已编码帧来进行编码。TL 越小,当前帧与参考帧之间的距离越远,如 $TL=2$ 时,当前帧 B_2 帧(T_3 时刻)与参考帧 B_0 帧(T_0 时刻)和 B_1 帧(T_1 时刻)的距离为 3;而当 $TL=4$ 时,当前帧 b_4 帧(T_4 时刻)与参考帧 B_2 帧(T_3 时刻)和 B_3 帧(T_5 时刻)的距离为 1。

图 1 中,视点 S_i 的所有图像都需要参考相邻视点 S_{i-1} 和 S_{i+1} 中相同时刻的图像帧来进行编码,因此被称为“B 视点”。“B 视点”的每个非锚定帧都

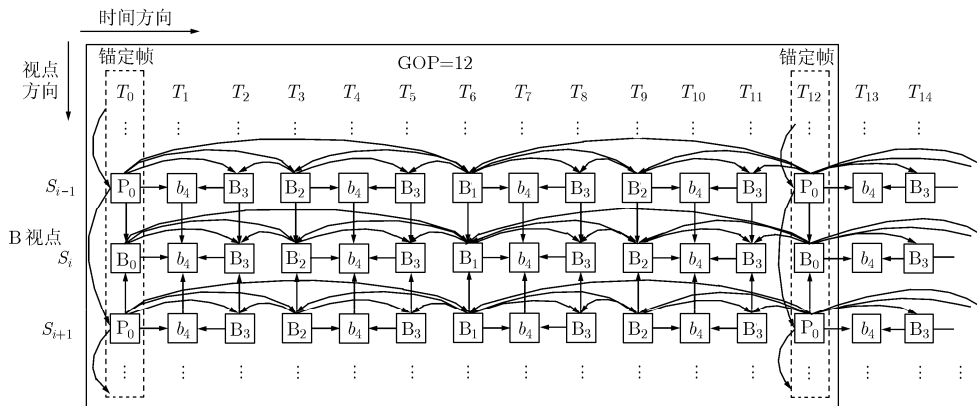


图 1 MVC 中的 HBP 预测结构

有 4 个参考帧,如图 2 所示,假设视点 S_i , t 时刻的帧为当前帧,记作 $f_{i,t}$; $f_{i,t-T}$ 和 $f_{i-1,t}$ 分别是当前帧时间方向和视点方向的前向参考帧,存放在前向参考帧列表 list0 中,分别记作 list0[0]和 list0[1]; $f_{i,t+T}$ 和 $f_{i+1,t}$ 为当前帧时间方向和视点方向的后向参考帧,存在后向参考帧列表 list1 中,分别记作 list1[0]和 list1[1]。

JMVC 中,当前块的帧间编码过程如图 3 所示,首先分别对当前块进行帧间 16×16 预测,帧间 16×8 预测,帧间 8×16 预测,和帧间 8×8 预测,并根据 RDCost 值的大小选择最佳预测模式作为最终预测结果。每种预测模式的计算量都主要来自两个方面:单向搜索(前向搜索和后向搜索)和双向搜索。JMVC 中原有的编码算法存在以下问题:

(1)单向搜索(前向/后向)遍历参考帧列表(list0/list1)中每个参考帧,为了获得较高的编码性能,采用大搜索窗口(如 ± 96 个像素)来分别进行运动/视差估计,运算量巨大;

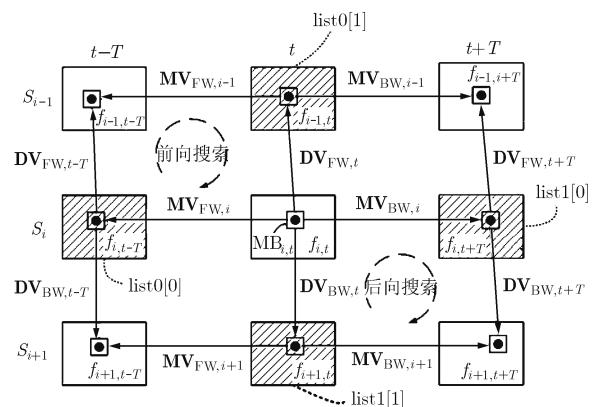


图 2 “B 视点”非锚定帧的运动和视差矢量

(2)单向搜索中,将运动和视差估计分成独立的两个步骤进行,没有利用多视点视频的序列相关性,相邻视点图像的运动矢量和视差矢量之间的高度联系没有得到充分利用;

(3)双向搜索分别从 list0 中选取 1 个前向参考帧,再从 list1 中选取 1 个后向参考帧,穷尽所有参

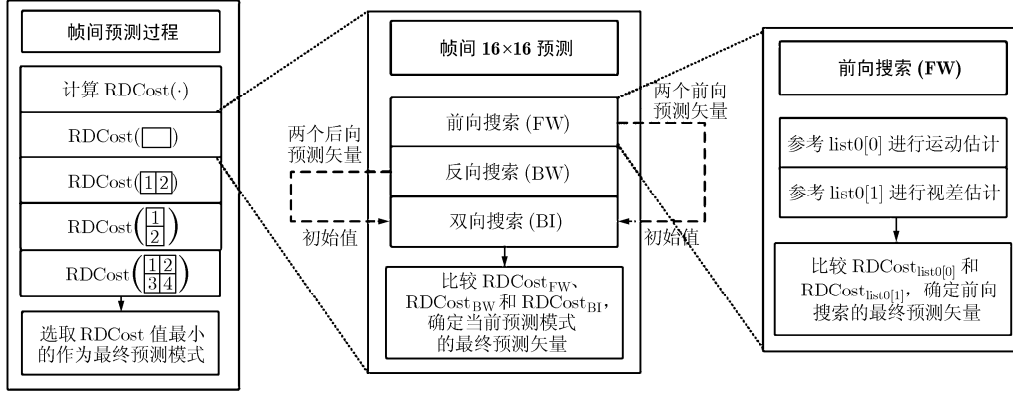


图 3 当前块的帧间预测编码过程示意

考帧反复进行循环迭代搜索，同时编码前向预测矢量差和后向预测矢量差，所需比特数可能更大。

3 基于可信度因子的单向搜索联合估计方法

3.1 基于立体-运动约束模型的迭代搜索策略

立体-运动约束模型的概念早在立体视频编码时就已经被引入^[15]，用来描述相邻视点相邻图像对之间的矢量关系。假设 $MB_{i,t}$ 为当前帧 $f_{i,t}$ 中的当前块，立体-运动约束模型可以被扩展到 MVC 的前向搜索和后向搜索中，如图 2 所示。

$$\left. \begin{aligned} \mathbf{MV}_{\text{FW},i} + \mathbf{DV}_{\text{FW},t-T} &= \mathbf{DV}_{\text{FW},t} + \mathbf{MV}_{\text{FW},i-1}, \\ &\text{前向搜索} \\ \mathbf{MV}_{\text{BW},i} + \mathbf{DV}_{\text{BW},t-T} &= \mathbf{DV}_{\text{BW},t} + \mathbf{MV}_{\text{BW},i+1}, \\ &\text{后向搜索} \end{aligned} \right\} (1)$$

其中 $\mathbf{MV}_{\text{FW},i}$ 和 $\mathbf{MV}_{\text{BW},i}$ 分别为当前块的前向和后向运动矢量， $\mathbf{DV}_{\text{FW},t}$ 和 $\mathbf{DV}_{\text{BW},t}$ 分别为当前块的前向和后向视差矢量， $\mathbf{MV}_{\text{FW},i-1}$ 为当前块在 $f_{i-1,t}$ 中视差补偿块的前向运动矢量， $\mathbf{DV}_{\text{FW},t-T}$ 为当前块在 $f_{i,t-T}$ 中运动补偿块的前向视差矢量， $\mathbf{MV}_{\text{BW},i+1}$ 为当前块在 $f_{i+1,t}$ 中视差补偿块的后向运动矢量， $\mathbf{DV}_{\text{BW},t+T}$ 为当前块在 $f_{i,t+T}$ 中运动补偿块的后向视差矢量。

为了解决 JMVC 原有单向搜索算法中存在的问题，本文利用立体-运动约束模型设计一种迭代求解策略，根据式(1)，由相邻图像的 3 个已知运动/视差矢量来求当前块视差/运动矢量的预测初值；再以这个预测初值为中心，设定一个小范围的修正搜索窗口，得到本次迭代修正后的视差/运动矢量。每次迭代都包含以下两个步骤，其中， k 代表当前迭代步数：

步骤 1 假设当前块的运动矢量 ($\mathbf{MV}_{\text{FW},i}^{k-1}$ 或 $\mathbf{MV}_{\text{BW},i}^{k-1}$) 已在 $k-1$ 次迭代过程中得到，利用式(1)求预测初值 ($\mathbf{BDV}_{\text{FW},i}^k$ 或 $\mathbf{BDV}_{\text{BW},i}^k$)；在预测初值周围设

定一个小的修正搜索窗口，并在这个新的搜索窗口内进行全搜索得到本次迭代修正后的视差矢量 ($\mathbf{DV}_{\text{FW},t}^k$ 或 $\mathbf{DV}_{\text{BW},t}^k$)；

步骤 2 利用修正后的视差矢量 ($\mathbf{DV}_{\text{FW},t}^k$ 或 $\mathbf{DV}_{\text{BW},t}^k$)，结合式(1)，计算得到运动矢量的预测初值 ($\mathbf{BMV}_{\text{FW},i}^k$ 或 $\mathbf{BMV}_{\text{BW},i}^k$)；并在这个预测初值附近做小范围的运动矢量修正，得到本次迭代的运动矢量修正值 ($\mathbf{MV}_{\text{FW},i}^k$ 或 $\mathbf{MV}_{\text{BW},i}^k$)。

由于该迭代算法只需要在每个预测初值点附近做小范围的搜索，与 JMVC 编码器中传统的独立搜索算法相比，搜索点数减少，算法复杂度能够大大降低。由于 H.264 中运动/视差估计以块为单位进行，以前向搜索为例，运动补偿块 $\text{MCMB}_{i,t-T}$ 和视差补偿块 $\text{DCMB}_{i-1,t}$ 通常会覆盖多个编码块(如图 4 中 A, B, C, D 所示)，即， $\mathbf{DV}_{\text{FW},t-T}(\text{MCMB}_{i,t-T})$ 和 $\mathbf{MV}_{\text{FW},i-1}(\text{DCMB}_{i-1,t})$ 存在多个可能值。那么， $f_{i-1,t-T}$ 中表示同一个物体在不同 3D 表面的投影块 $\text{MCMB}_{i-1,t-T}$ 和 $\text{DCMB}_{i-1,t-T}$ 并不一定汇聚于同一点，如图 4 所示，这使得式(1)所描述的立体-运动约束模型引入了一个模型误差，用 δ 表示。于是，式(1)转变成，

$$\delta = \left\{ \begin{aligned} &\|\mathbf{DV}_{\text{FW},t}(\text{MB}_{i,t}) + \mathbf{MV}_{\text{FW},t-1}(\text{DCMB}_{i-1,t}) \\ &\quad - \mathbf{DV}_{\text{FW},t-T}(\text{MCMB}_{i,t-T}) - \mathbf{MV}_{\text{FW},i}(\text{MB}_{i,t})\|, \\ &\text{前向搜索} \\ &\|\mathbf{DV}_{\text{BW},t}(\text{MB}_{i,t}) + \mathbf{MV}_{\text{BW},i+1}(\text{DCMB}_{i+1,t}) \\ &\quad - \mathbf{DV}_{\text{BW},t+T}(\text{MCMB}_{i,t+T}) - \mathbf{MV}_{\text{BW},i}(\text{MB}_{i,t})\|, \\ &\text{后向搜索} \end{aligned} \right. (2)$$

其中 $\|\mathbf{v}\|$ 表示矢量 \mathbf{v} 的范数。模型误差 δ 将影响到迭代搜索算法的准确度。当且仅当 $\text{MB}_{i,t}$ ， $\text{MCMB}_{i,t-T}$ ， $\text{DCMB}_{i-1,t}$ ， $\text{MCMB}_{i-1,t-T}$ 和 $\text{DCMB}_{i-1,t-T}$ 为同一个物体在不同图像表面的真实汇聚点时， $\delta = 0$ 。

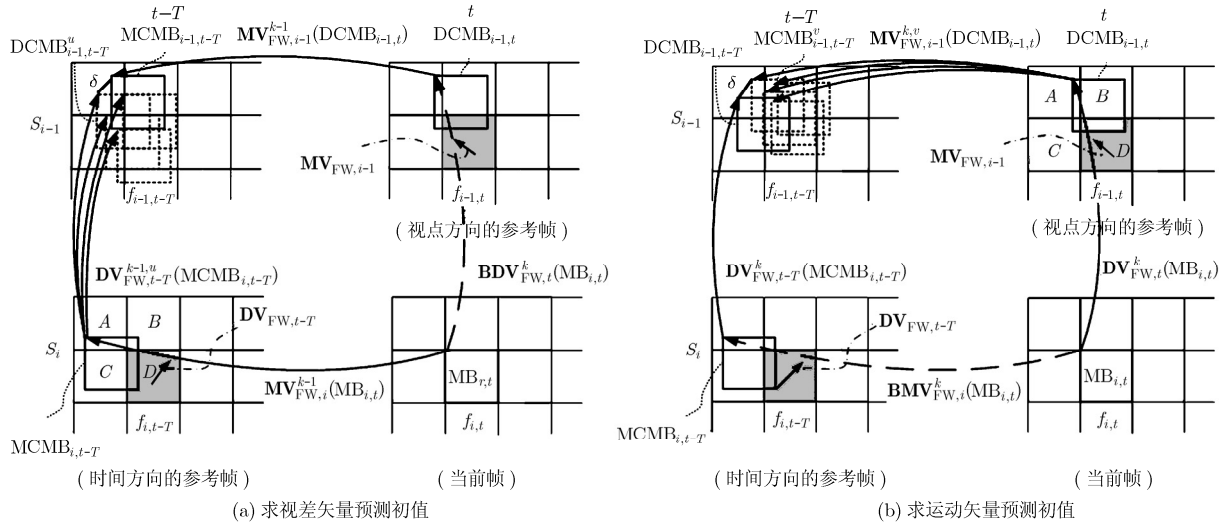


图4 立体-运动约束模型引入的模型误差(前向搜索)

假设当前为第 \$k\$ 次迭代过程, $MV_{FW,t}^{k-1}(MB_{i,t})$ 为上一次迭代得到的当前块的前向运动矢量, 由运动补偿块 $MCMB_{i,t-T}$ 回溯得到的视差矢量将有 4 个可能值 $DV_{FW,t-T}^u(MCMB_{i,t-T})$, 其中 $1 \leq u \leq 4$, 如图 4(a) 所示。当前迭代过程的视差矢量初值可以由式(3)求得

$$BDV_{FW,t}^k(MB_{i,t}) = \text{Min}_{1 \leq u \leq 4} (\text{RDCost}(BDV_{FW,t}^{k,u}(MB_{i,t}))),$$

前向搜索 (3)

其中 $\text{RDCost}(\bullet)$ 表示率失真代价函数^[3], 结合式(1), $BDV_{FW,t}^{k,u}(MB_{i,t})$ 可以由式(4)求得

$$BDV_{FW,t}^{k,u}(MB_{i,t}) = MV_{FW,t}^{k-1}(MB_{i,t}) + DV_{FW,t-T}^{k-1,u}(MCMB_{i,t-T}) - MV_{FW,t-1}^{k-1}(DCMB_{i-1,t}),$$

前向搜索 (4)

其中 $DCMB_{i-1,t}$ 是由上一次迭代的 $DV_{FW,t}^{k-1}(MB_{i,t})$ 回溯得到的视差补偿块。令 $\text{RDCost}(BDV_{FW,t}^{k,u}(MB_{i,t}))$ 最小的 $DV_{FW,t-T}^{k-1,u}(MCMB_{i,t-T})$ 作为 $DV_{FW,t-T}^k(MCMB_{i,t-T})$, 用以计算当前迭代过程下一步的运动矢量预测初值 $BMV_{FW,t}^k(MB_{i,t})$ 。

同理, 当修正后的视差矢量 $DV_{FW,t}^k(MB_{i,t})$ 确定之后, 当前块在视点方向参考帧 $f_{i-1,t}$ 中的视差补偿块 $DCMB_{i-1,t}$ 覆盖的 4 个已编码块的运动矢量, 记作 $MV_{FW,i-1}^v(DCMB_{i-1,t})$, $1 \leq v \leq 4$, 用来作为当前迭代过程中运动矢量预测初值 $BMV_{FW,t}^k(MB_{i,t})$ 的候选值, 如图 4(b) 所示。 $BMV_{FW,t}^k(MB_{i,t})$ 可以通过式(5)求得

$$BMV_{FW,t}^k(MB_{i,t}) = \text{Min}_{1 \leq v \leq 4} (\text{RDCost}(BMV_{FW,t}^{k,v}(MB_{i,t}))),$$

前向搜索 (5)

其中

$$BMV_{FW,t}^k(MB_{i,t}) = DV_{FW,t}^k(MB_{i,t}) + MV_{FW,t-1}^{k,v}(DCMB_{i-1,t}) - DV_{FW,t-T}^k(MCMB_{i,t-T}),$$

前向搜索 (6)

式(6)中 $DV_{FW,t-T}^k(MCMB_{i,t-T})$ 为在第 \$k\$ 次迭代第 1 步中求得。

3.2 基于可信度因子的自适应修正窗口设计

式(2)得到的模型误差 δ 可以作为可信度因子来衡量立体-运动约束模型的准确度, 如果 δ 很小 ($\delta < T_1$), 说明此时的预测矢量候选值非常接近真实值, 用较小的修正窗口(如 ± 2 个像素)就能保证搜索精度; 若 δ 非常大 ($\delta > T_2$), 说明当前块很可能为运动遮挡区域, 立体-运动约束模型不再适用, 因此, 采用一个较大的搜索窗口 (RSR_{MAX}) 来保证编码质量; 反之, 则采用一个自适应的搜索窗口 (RSR),

$$\text{RSR} = \begin{cases} \text{RSR}_{\text{MIN}}, & \delta < T_1 \\ \text{RSR}_{\text{MIN}} + \frac{\delta - T_1}{T_2 - T_1} (\text{RSR}_{\text{MAX}} - \text{RSR}_{\text{MIN}}), & T_1 \leq \delta \leq T_2 \\ \text{RSR}_{\text{MAX}}, & \delta > T_2 \end{cases} \quad (7)$$

其中 RSR_{MIN} 和 RSR_{MAX} 分别表示最小和最大的修正搜索窗口尺寸。 $\text{RSR}_{\text{MAX}} = \lceil \alpha \times \text{SR} \rceil$, $\alpha \leq 1$, SR 为 JMVC 初始搜索窗口, 通过调节 α 的大小可以在预测精度和搜索复杂度之间找到一个平衡点。

3.3 单向搜索快速算法流程图

以前向搜索为例, 假设 $MB_{i,t}$ 为当前块, 本文的快速算法包括初始化、修正窗口大小调整、迭代搜索和中止准则 4 个部分组成, 如图 5 所示, 其中 k 代表迭代步数。运动矢量搜索起始点 ($BMV_{FW,t}^0(MB_{i,t})$) 和视差矢量搜索起始点

($\mathbf{BMV}_{FW,i}^0(\mathbf{MB}_{i,t})$) 在如下的候选矢量集中选取 $\mathbf{BMV}_{FW,i}^0(\mathbf{MB}_{i,t}) : \{\mathbf{MV}_{FW,i-1}, \mathbf{MV}_{med}, \mathbf{MV}_a, \mathbf{MV}_b, \mathbf{MV}_c, \mathbf{0}\}$ 和 $\mathbf{BDV}_{FW,t}^0(\mathbf{MB}_{i,t}) : \{\mathbf{DV}_{FW,t-T}, \mathbf{DV}_{med}, \mathbf{DV}_a, \mathbf{DV}_b, \mathbf{DV}_c, \mathbf{0}\}$ 。其中 $\mathbf{MV}_{FW,i-1}$ 表示参考帧 $f_{i-1,t}$ 中与当前块位置相同的块的运动矢量, $\mathbf{DV}_{FW,t-T}$ 表示 $f_{i,t-T}$ 中与当前块位置相同的块的视差矢量, 如图4所示。 a, b, c 表示当前块相邻的左边块, 上边块和右上块, $\mathbf{MV}_{med} / \mathbf{DV}_{med}$ 为中值矢量, 由于 a, b 和 c 是已编码块, 预测矢量只可能为运动矢量或者视差矢量二者之一^[16,17]。假如 b 是以视差补偿预测方式编码, \mathbf{MV}_b 就不存在, 此时, 我们用零矢量(0,0)来代替 \mathbf{MV}_b 求取中值矢量 \mathbf{MV}_{med} 。分别将上述两组候选矢量集中 RDCost 值最小的作为当前块运动和视差矢量的搜索起始点 $\mathbf{MV}_{FW,i}^0(\mathbf{MB}_{i,t})$ 和 $\mathbf{DV}_{FW,t}^0(\mathbf{MB}_{i,t})$ 。然后, 按照前文描述的方法调整 RSR, 进行迭代搜索求得 $\mathbf{DV}_{FW,t}^k(\mathbf{MB}_{i,t})$ 和 $\mathbf{MV}_{FW,i}^k(\mathbf{MB}_{i,t})$, 当 $\text{RDCost}(\mathbf{MV}_{FW,i}^k(\mathbf{MB}_{i,t})) \geq \text{RDCost}(\mathbf{MV}_{FW,i}^{k-1}(\mathbf{MB}_{i,t}))$ 并且 $\text{RDCost}(\mathbf{DV}_{FW,t}^k(\mathbf{MB}_{i,t})) \geq \text{RDCost}(\mathbf{DV}_{FW,t}^{k-1}(\mathbf{MB}_{i,t}))$ 时, 结束前向搜索, 令 $\mathbf{MV}_{FW,i}^{k-1}(\mathbf{MB}_{i,t})$ 和 $\mathbf{DV}_{FW,t}^{k-1}(\mathbf{MB}_{i,t})$ 作为当前块的前向运动和视差矢量。否则, 更新 δ , 令 $k = k + 1$, 继续下一次迭代过程。

4 双向搜索提前中止算法

双向预测 B 帧同时参考前后向参考帧进行双向预测, 非常适合运动遮挡区域的编码。运动遮挡区域是指由 3 维旋转或变换造成的物体表面某个区域

被遮挡的情况, 比如: 物体甲从物体乙前面经过; 运动物体经过摄像机; 相机镜头拉近/拉远, 物体旋转等。对于运动遮挡区域的宏块, 如果只用一个参考帧从单方向来进行运动/视差估计, 很可能因为找不到对应的像素点而得到错误的预测矢量, 最终导致编码结果变差^[14]。

4.1 双向预测和量化参数 QP 的关系

根据 MVC 中的率失真优化准则, 运动/视差矢量域的分布特性将随着 QP 的大小而变化。率失真代价的计算公式为 $\text{RDCost} = D + \lambda R$, D 指失真度, R 指编码每个宏块所需的比特数, λ 是拉格朗日乘子, λ 等于 $0.85 \times 2^{\frac{\min\{52, \text{QP}\} - 4}{3}}$ 。当 QP 越大时, λ 的值越大, R 对 RDCost 值的影响占的权重更大, 此时, 为了得到较小的 RDCost 值, 运动/视差矢量域的分布更为平滑, 即, 当前块预测矢量的大小更接近于其相邻块的预测矢量。此时, 采用单向搜索得到高预测精度的可能性更大。相反, 在低 QP 时, 运动/视差矢量域的分布没有规律性, 单方向的运动/视差估计远不能满足精确预测的要求, 因此, 需要结合双向搜索同时参考前后向参考帧进行矢量估计来提高预测矢量精度。

本文统计了 Ballroom, Exit, Vassar, Flamenco2, Race1 和 Rena 序列采用双向预测的宏块数量随 QP 的变化情况, 得到图6的曲线。可以看出, 当 QP 大于 35 时, 选择双向搜索作为最佳预测类型的宏块仅占 5%。也就是说, 当 QP 较高时, 双向搜索带来的编码增益并不明显, 超过 95% 的宏块通过单向搜索(前向搜索和后向搜索)就已经能够得到足够准确的预测矢量。因此, 本文将 T_{QP} 设为 35, 当 $\text{QP} > T_{QP}$ 时, 不再对当前块做双向搜索; 反之, 当前块需要遍历单向搜索(前向搜索和后向搜索)和双向搜索, 并从中选择最佳预测结果。

4.2 双向预测和时序层次 TL 及模型误差 δ 的关系

HBP 结构中, 当 TL 较小时, 当前帧与参考帧之间的距离越远, 相关性更低, 仅利用单向搜索可能得不到最优的预测结果, 必须结合双向搜索同时从两个方向来进行运动/视差估计, 提高预测精度。而当 TL 较大时, 当前帧与参考帧之间的距离越近, 相关程度越高, 单向搜索得到高精度预测结果的可能性较大。图7给出各个测试序列采用双向预测的宏块数量与 TL 的关系(GOP=12)。可以看出, 当 TL=4 时, 只有少于 10% 的宏块选择双向搜索作为最佳预测类型。这说明, 对于大多数宏块(90%以上), 在时序层数等于 4 时, 仅采用单向搜索就很可能得到足够

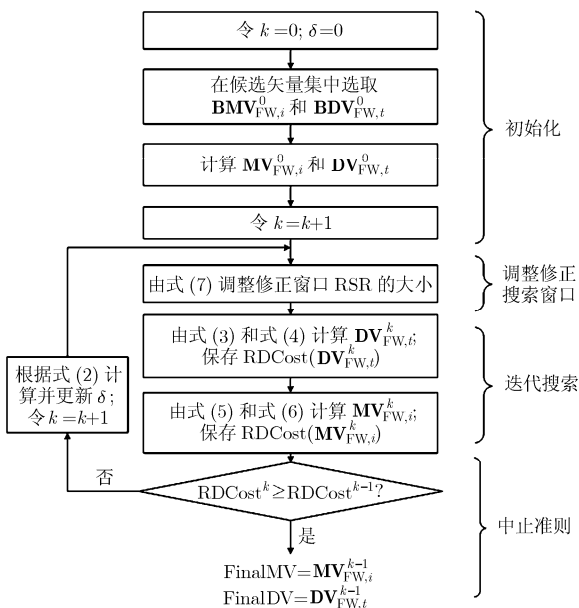


图5 单向搜索快速算法流程图

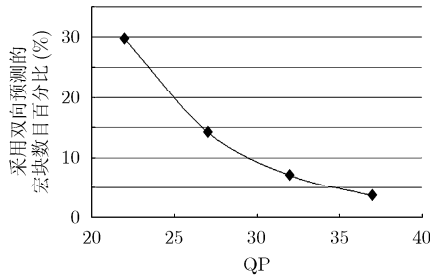


图6 采用双向预测的宏块数目随QP变化的统计图

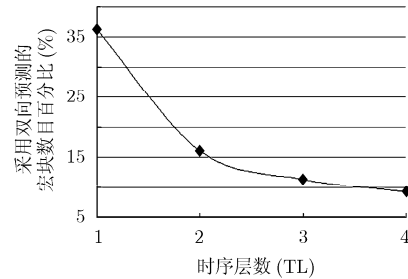


图7 采用双向预测的宏块数目随TL变化的统计图

精确的预测结果；另外，由前文的分析可知，模型误差 δ 的大小可以用来作为判断立体-运动约束模型可靠性的依据，当 δ 不等于0时，当前块可能为被遮挡块，此时，除了采用单向搜索之外，还需要结合双向预测来提高编码结果的准确性。因此，本文令 $T_{TL}=4$ 。当 $TL \geq T_{TL}$ 并且 $\delta = 0$ 时，则跳过双向搜索过程。

5 实验结果

实验测试序列为 Ballroom, Exit, Vassar, Flamenco2, Race1 和 Rena^[18]，用本文提出的 MVC 快速算法对“B 视点”进行测试，实验平台为 JMVC7.2，量化参数 QP 取值为 22, 27, 32, 37，按照 HBP 预测结构进行编码，搜索范围设为 ± 96 ，GOP 的长度为 12。实验参数和实验条件均依照文献[19]公布的 MVC 测试标准进行。 T_1 和 T_2 分别取值为 5 和 20， RSR_{MIN} 为 2， α 等于 0.33。所有实验在配置为 Intel(R) Xeon(R) X5550 2.67 GHz CPU, 12 GB RAM 的 PC 上独立执行。

为了验证本文算法的有效性，将本文算法(Prop)分别与 JMVC 全搜索算法(FS)，JMVC 中的 TZ 快速搜索算法(TZ)，文献[9]中的 MtD 算法(SPIE)，文献[12]的算法(CE)进行比较。

5.1 算法复杂度分析

本文以每个宏块的平均搜索点数以及待比较算法与FS算法相比的加速倍数来衡量算法复杂度，表1给出了不同算法复杂度的比较结果，其中，小括号里的数表示加速倍数，加速倍数等于FS算法的平均搜索点数除以待比较算法的平均搜索点数。从表1可以看出，Prop算法的搜索点数远远低于其它算法，与FS算法相比，Prop算法的加速倍数高达417倍(Vassar序列，QP37)。CE算法根据中值矢量的可信度来调整搜索窗口的大小，能够在一定程度上减少搜索点数，但是对于Race1这种包含快速运动物体的序列，由于传统中值矢量的准确性降低，CE算法的搜索点数远远高于本文算法。SPIE算法的整体加速倍数仅在2左右，这是由于该快速视差估计算法以全

搜索的运动估计为代价，尽管视差估计的复杂度得到极大降低，但是运动估计仍然占据了不菲的编码时间。表2给出Prop算法中单向搜索过程的迭代步数 k 、模型误差 δ 和修正搜索窗口RSR的统计结果，其中， $Avg(\bullet)$ 表示取均值。从表中前两行可以看出，95%以上的宏块的迭代步数 k 都不超过5次，并且平均迭代步数 $Avg(k)$ 的值非常小，这说明本文的快速算法能以较小的迭代步数达到收敛。从表2中 $Avg(\delta)$ 和 $Avg(RSR)$ 的值可以看出，对于Race1序列，快速运动物体的存在造成遮挡现象严重，立体-运动约束模型的有效性降低，因此， $Avg(\delta)$ 相对其它序列更大，必须采用大搜索窗口来保证编码质量，所以 $Avg(RSR)$ 也相应增大。尽管如此，Prop算法在Race1序列的搜索点数仍然远远低于其它算法，见表1。

5.2 率失真性能分析

表3给出了不同算法的率失真性能与FS算法比较的结果，其中Delta-dB和%BIT分别表示采用Bjontegaard方法^[20]得到的待比较算法与FS算法之间的绝对PSNR变化量和绝对比特率变化量。从表3的数据可以看出，CE算法和SPIE算法的率失真性能次于FS算法；Prop算法和TZ算法的性能不低于FS算法。在有些序列，TZ算法和Prop算法甚至略优于FS算法，这是因为TZ算法和Prop算法在中值矢量、零矢量和相邻块的预测矢量中选取最佳预测初值，而FS算法仅采用中值矢量作为预测初值。由于传统的中值矢量预测法没有考虑MVC中相邻视点方向的多参考帧技术^[16,17]，例如，当相邻的左边、上边和右上块都为运动补偿预测时，这3个运动矢量的中值仍然会被用来作为当前块视差矢量的预测初值，造成不准确的预测结果，影响整个FS算法的性能。该结果也进一步证明，本文通过模型误差来自适应地调整预测初值的修正窗口，弥补立体-运动约束模型的不足，避免在迭代搜索时陷入局部极小值，最终以较少的搜索点数获得高效的编码性能。

表 1 不同算法的平均搜索点数和加速倍数

	Ballroom					Exit					Vassar				
	FS	TZ	SPIE	CE	Prop	FS	TZ	SPIE	CE	Prop	FS	TZ	SPIE	CE	Prop
QP22	132132	9233	67289	21275	3911	132220	8914	67258	6651	2440	132156	7686	67196	4444	1531
	(14.30)	(2.00)	(6.20)	(33.80)		(14.80)	(2.00)	(19.90)	(54.20)		(17.20)	(2.00)	(29.70)	(86.30)	
QP27	132097	9125	67234	28762	3599	132190	8324	67178	5039	2066	132064	7151	67103	4792	1427
	(14.50)	(2.00)	(4.60)	(36.70)		(15.90)	(2.00)	(26.20)	(64.00)		(18.50)	(2.00)	(27.60)	(92.50)	
QP32	132140	8908	67223	21188	3395	132152	7875	67123	5150	1907	132039	6156	67072	4967	1309
	(14.80)	(2.00)	(6.20)	(38.90)		(16.80)	(2.00)	(25.70)	(69.30)		(21.40)	(2.00)	(26.60)	(100.90)	
QP37	132146	8486	67182	15730	1992	132089	7436	67064	4925	701	132063	5217	67092	4261	316
	(15.60)	(2.00)	(8.40)	(66.40)		(17.80)	(2.00)	(26.80)	(188.40)		(25.30)	(2.00)	(31.00)	(417.40)	

表 1 不同算法的平均搜索点数和加速倍数(续)

	Falmenco2					Race1					Rena				
	FS	TZ	SPIE	CE	Prop	FS	TZ	SPIE	CE	Prop	FS	TZ	SPIE	CE	Prop
QP22	132549	8701	67467	4929	2599	132048	9434	67272	6855	4302	132214	6006	67279	8541	2289
	(15.20)	(2.00)	(26.90)	(51.00)		(14.00)	(2.00)	(19.30)	(30.70)		(22.00)	(2.00)	(15.50)	(57.80)	
QP27	132547	8745	67433	4789	2478	132081	9687	67246	6184	4185	132183	5828	67241	8569	2204
	(15.20)	(2.00)	(27.70)	(53.50)		(13.60)	(2.00)	(21.40)	(31.60)		(22.70)	(2.00)	(15.40)	(60.00)	
QP32	132539	8805	67392	4707	2365	132098	9683	67235	5918	3989	132127	5369	67169	7845	2009
	(15.10)	(2.00)	(28.20)	(56.00)		(13.60)	(2.00)	(22.30)	(33.10)		(24.60)	(2.00)	(16.80)	(65.80)	
QP37	132483	8794	67323	4476	990	132113	9386	67249	5177	2359	132105	4484	67138	6327	651
	(15.10)	(2.00)	(29.60)	(133.90)		(14.10)	(2.00)	(25.50)	(56.00)		(29.50)	(2.00)	(20.90)	(203.10)	

表 2 迭代步数 k , 模型误差 δ 和修正搜索窗口 RSR 统计结果

	Ballroom	Exit	Vassar	Flamenco2	Race1	Rena
$k=1$	84.98%	78.59%	90.51%	70.51%	62.84%	80.70%
$k \leq 5$	98.63%	98.51%	99.68%	98.51%	95.75%	97.78%
Avg(k)	1.32	1.42	1.15	1.54	1.84	1.43
Avg(δ)	2.58	1.88	1.89	2.82	6.70	2.94
Avg(RSR)	2.74	2.15	2.08	2.82	4.03	2.35

表 3 采用 Bjontegaard 方法得到的率失真性能比较结果

	FS		TZ		CE		SPIE		Prop	
	PSNR-Y (dB)	Bitrate (kbps)	Delta-dB (dB)	%BIT (%)	Delta-dB (dB)	%BIT (%)	Delta-dB (dB)	%BIT (%)	Delta-dB (dB)	%BIT (%)
Ballroom	770.21	35.53	0.02	-0.55	-0.16	4.33	-0.18	4.67	0.00	0.16
Exit	390.10	37.74	0.00	0.08	-0.20	8.12	-0.05	2.06	0.00	0.37
Vassar	480.22	35.83	0.01	-0.32	-0.01	0.74	-0.01	0.36	0.00	0.02
Falmenco2	759.14	37.82	0.04	-0.67	-0.35	6.69	-0.37	7.26	0.02	-0.36
Race1	814.69	36.53	0.05	-1.13	-0.48	11.62	-0.13	3.19	0.04	-0.98
Rena	476.46	40.13	0.09	-1.62	0.04	-0.83	0.01	-0.20	0.08	-1.48

6 结论

本文旨在结合多视点视频的序列相关性，解决 JMVC 现有算法运动估计和视差估计相对独立，编码复杂度高的难点。针对 HBP 结构中编码最为复杂的“B 视点”提出并实现了 MVC 快速算法，首先，采用基于可信度因子的迭代搜索策略对当前块进行快速前向搜索和后向搜索，结合立体-运动约束模

型，在运动域和视差域多参考帧中自由切换，轮换搜索；并根据多视点视频的宏块特性开发一种双向搜索提前中止算法，在保证编码质量的同时进一步提高编码速度。

参考文献

- [1] Chen Y, Wang Y K, Ugur K, *et al.* The emerging MVC standard for 3D video services [J]. *EURASIP Journal on*

- Advances in Signal Processing*, 2009(1), doi: 10.1155/2009/786015.
- [2] ISO/IEC JTC1/SC29/WG11. Editors' draft revision to ITU-T Rec. H.264 | ISO/IEC 14496-10 Advanced Video Coding - in preparation for ITU-T SG 16 AAP Consent (in integrated form) [C]. JVT-AD007, Geneva, CH, 2009.
- [3] ISO/IEC JTC1/SC29/WG11. WD 1 Reference software for MVC [C]. JVT-AA212, Geneva, CH, 2008.
- [4] Merkle P, Smolic A, Muller K, *et al.* Efficient prediction structures for multiview video coding [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2007, 17(11): 1461-1473.
- [5] Wang R and Wang Y. Multiview video sequence analysis, compression, and virtual viewpoint synthesis [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2000, 10(3): 397-410.
- [6] Shen L, Liu Z, Liu S, *et al.* Selective disparity estimation and variable size motion estimation based on motion homogeneity for multi-view coding [J]. *IEEE Transactions on Broadcasting*, 2009, 55(4): 761-766.
- [7] Lin J P and Tang A C. A fast direction predictor of inter frame prediction for multi-view video coding [C]. Proceeding of International Conference on Acoustics, Speech, and Signal Processing, Taipei, 2009: 2589-2592.
- [8] Xu X and He Y. Fast disparity motion estimation in MVC based on range prediction [C]. Proceeding of International Conference on Image Processing, San Diego, CA, 2008: 2000-2003.
- [9] Lai P and Ortega A. Predictive fast motion/disparity search for multiview video coding [C]. Proceeding of SPIE Visual Communications and Image Processing, San Jose, USA, 2006: 1-11.
- [10] Tsung P K, Chen W Y, Ding L F, *et al.* Cache-based integer motion/disparity estimation for quad-HD H.264/AVC and HD multiview video coding [C]. Proceeding of International Conference on Acoustics, Speech, and Signal Processing, Taipei, 2009: 2013-2016.
- [11] Li X, Zhao D, Ma S, *et al.* Fast disparity and motion estimation based on correlations for multiview video coding [J]. *IEEE Transactions on Consumer Electronics*, 2008, 55(4): 2037-2044.
- [12] Kim Y, Kim J, and Sohn K. Fast disparity and motion estimation for multi-view video coding [J]. *IEEE Transactions on Consumer Electronics*, 2007, 53(2): 712-719.
- [13] Peng Z, Jiang G, Yu M, *et al.* Fast macroblock mode selection algorithm for multiview video coding [J]. *EURASIP Journal on Image and Video Processing*, 2008, doi:10.1155/2008/393727.
- [14] Lin H C and Hang H M. Fast algorithm on selecting bi-directional prediction type in H.264/AVC scalable video coding [C]. Proceeding of International Conference on Acoustics, Speech, and Signal Processing, Paris, 2010: 113-116.
- [15] 邓智玘, 贾克斌, 陈锐霖, 等. 基于立体-运动约束模型的视差-运动同步联立预测方法[J]. *计算机辅助设计与图形学学报*, 2010, 22(10): 1709-1714.
Deng Zhi-pin, Jia Ke-bin, Chan Yui-lam, *et al.* A simultaneous motion and disparity estimation for stereoscopic video coding [J]. *Journal of Computer-Aided Design & Computer Graphic*, 2010, 22(10): 1709-1714.
- [16] ISO/IEC JTC1/SC29/WG11. Disparity vector prediction methods in MVC [C]. JVT-U040, Hangzhou, China, 2006.
- [17] ISO/IEC JTC1/SC29/WG11. Regional disparity based motion and disparity prediction for MVC [C]. JVT-V071, Marrakech, Morocco, 2007.
- [18] ISO/IEC JTC1/SC29/WG11. Multiview video test sequences from MERL [C]. M12077, Busan, Korea, 2005.
- [19] ISO/IEC JTC1/SC29/WG11. Common Test Conditions for Multiview Video Coding [C]. JVT-T207, Klagenfurt, Austria, 2006.
- [20] ISO/IEC JTC1/SC29/WG11. Calculation of average PSNR differences between RD-curves [C]. M33, Austin, USA, 2001.
- 邓智玘: 女, 1983年生, 博士生, 研究方向为多视点视频编码、视频检索等.
- 贾克斌: 男, 1961年生, 博士, 教授, 博士生导师, 研究方向为多媒体信息处理等.
- 陈锐霖: 男, 1971年生, 博士, 教授, 博士生导师, 研究方向为信息处理、视频编码等.
- 伏长虹: 男, 1981年生, 博士, 研究方向为视频技术等.
- 萧允治: 男, 1954年生, 博士, 教授, 博士生导师, 研究方向为图像和视频技术等.