

基于 H. 264 的帧间预测快速算法

范志蔚^{1), 3)} 何加铭^{1), 2)} 郁梅³⁾

¹⁾ (宁波大学通信技术研究, 宁波 315211) ²⁾ (宁波市通信芯片与射频技术重点实验室, 宁波 315040)

³⁾ (宁波大学, 宁波 315211)

摘要 为了快速地进行运动估计, 针对 H. 264 帧间编码全搜索的计算复杂度太高, 不能应用于实时处理的缺点, 本文实现了一种适合于 H. 264 的快速搜索算法。该算法针对 H. 264 多预测模式的特点, 利用其相互间的相关性来确定预测点; 同时利用运动图像的一般规律来确定预测点, 从而使得预测点的确定时间得以减少。在预测点确定后再用钻石算法搜索, 这使得整个帧间预测时间大为缩短。实验结果表明, 该算法能够在码率增加不超过 1%, 信噪比 (signal noise ratio, SNR) 下降不超过 0.1dB 的情况下, 使得运动估计速度提高将近 65%。

关键词 视频编码 H. 264 帧间预测 快速算法

中图法分类号: TN919.8 **文献标识码:** A **文章编号:** 1006-8961(2005)11-1360-04

An Inter Prediction High-speed Algorithm Based on H. 264

FAN Zhi-wei^{1), 2)}, HE Jia-ming^{1), 2)}, YU Mei³⁾

¹⁾ (Technological Research Institute of Communication of Ningbo University, Ningbo 315211)

²⁾ (The Communication Chip and Key Laboratory of Radio Frequency Technology of Ningbo, Ningbo 315040)

³⁾ (Ningbo University, Ningbo 315211)

Abstract In order to make motion estimation quicker, for the computation complexity of full search of inter coding is very costly in H. 264 that can not be used in real-time applying, this paper proposed a quick search algorithm which is adapt to H. 264. For one of the special characteristic of H. 264 is multi-prediction mode, this paper utilize of the correlation of each mode to determinate a prediction; at the time, this paper also consider general regulation of motion picture to determinate a prediction, these prediction can reduce time, then use diamond search algorithm to end up motion estimation, this method can reduce much prediction time. The experimentation result show that the algorithm can raise 65% coding rate with increase no more than 1% bite-rate and descend no more 0.1dB SNR (signal noise ratio).

Keywords video encoding, H. 264, inter prediction, high-speed algorithm

1 引言

现在, 可视功能将成为移动终端的重要组成部分, 而且无论是 3G 手机还是新一代的对讲机, 无不把可视作为一个基本功能。在众多可应用于无线传输的图像处理技术中, 国际电信联盟组织和国际标准化组织成立的联合专家组 (joint video team, JVT) 提出的一项视频编码标准 H. 264^[1-3], 以其极高的编码效率和图像质量引起了业界人士的关注。然

而, 由于 H. 264 编码效率的提升是以计算复杂度的提高为代价的, 因此原始的全搜索算法很难应用于实际应用, 尤其是应用于移动终端等实时处理的场合, 为此本文提出一种基于 H. 264 的帧间预测快速算法, 以便使 H. 264 能面向实际应用。

2 基于 H. 264 运动估计的改进算法

以前已经有过很多快速算法, 如钻石搜索算法、MVFAST (motion vector field adaptive search

基金项目: 国家自然科学基金项目 (60372026); 宁波市工业攻关计划项目 (2003B10012)

收稿日期: 2005-08-29; 改回日期: 2005-09-16

第一作者简介: 范志蔚 (1982 ~), 男, 2005 年获宁波大学学士学位, 现为宁波大学无线通信硕士研究生。研究方向为无线移动通信终端。E-mail: scott623@126.com

technique)^[4]、PMVFAST (predictive motion vector field adaptive search technique)^[5]。虽然这些快速算法(包括 MVFAST 和 PMVFAS)均取得了优异的搜索性能,但是它们并不是非常适合于 H. 264, 因为 H. 264 的运动估计采用了 7 种块大小和形状, 如果每一种方式的运动估计都执行一遍, 则这些快速算法增加的复杂度会变得不可接受。

本文针对上述情况, 根据 H. 264 的运动估计的特殊性, 对其作了改进。

3 起始点预测

起始点预测作为一种重要的技术被许多运动搜索快速算法采用, 而将搜索区域设置在整个搜索窗口中最小块失真 (minimum block distortion) 点的周围, 则可以有效地改进运动搜索的性能。在本文的算法中包括了 3 种不同的方法用来预测候选起始点, 并最终选择开销最小的 SAD (sum of absolute differences, 绝对差之和) 作为下一步搜索的起始点。3 种方法描述如下:

3.1 原始点运动矢量^[6]

由于在视频序列中很大一部分宏块 (macro block, MB) 是趋于运动矢量等于 (0, 0) 的低运动矢量, 因此这些 MB 可以直接用 (0, 0) 矢量来计算 SAD 值。这样就可以把原始点 (即运动矢量为 (0, 0) 的点) 作为候选预测。

本文把一个宏块位置的本地运动矢量域定义为这个宏块支持带 (region of support, ROS) 的一组运动矢量。这个宏块的 ROS 包括 n 个相邻宏块。在本算法中 ROS 的 $n=3$, 如图 1 所示。

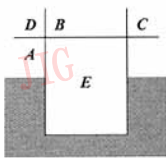


图 1 当前块与相邻块的关系图

Fig. 1 Reference block location for prediction of motion vectors

$$V_{pred}^{motion} = \text{median}(V_A^{motion}, V_B^{motion}, V_C^{motion})$$

这个被预测的运动矢量是 (x_{pred}, y_{pred}) 。如果预测的运动矢量是在边界, 则将这个预测运动矢量定义为: 当 A 块在图片或块组 (Group of blocks, GOB) 边界的外面时, 就用 (0, 0) 点代替; 当 C 块在图片或 GOB 边界的外面时, 则用 D 块的运动矢量代替; 当 B 块和 C 块都在外面时, 就用第 3 块的运动矢量代替。

3.3 上层块运动矢量^[6]

如图 2 所示, 在 H. 264 的中定义了 7 种帧间预测模式。8×8 模式 (模式 4、模式 5、模式 6、模式 7) 是紧接着的 16×16 模式 (模式 3、模式 2、模式 1) 后第 1 个被搜索的。由于这样的策略在应用不同模式的运动相关性时不是很有益, 因此预测模式的搜索次序要根据模式的块大小而改变。本文选择了一个从模式 1 到 7 的分等级的搜索次序作为本算法的预测模式搜索次序, 上层块的运动矢量 (如模式 5、模式 6 是模式 7 的上层块, 模式 4 是模式 5 或模式 6 的上层块等) 将被用作下层的一个候选预测。因此对于模式 1, 中间预测、(0, 0) 矢量、左相邻块运动矢量、上相邻块运动矢量和右上相邻块运动矢量都可以作为它的候选预测; 对于其他模式来说, 中间预测、(0, 0) 矢量和其上层运动矢量可作为它的候选预测。

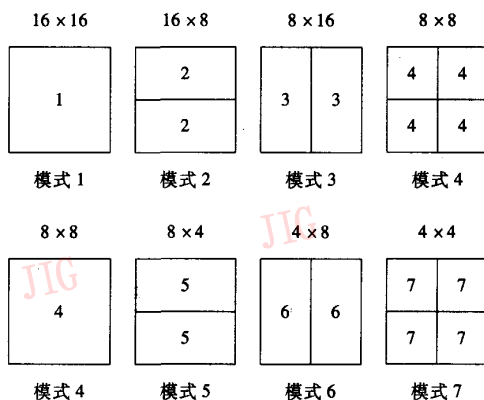


图 2 H. 264 的 7 种帧间预测模式

Fig. 2 Seven prediction modes in H. 264

在这些候选预测中具有最小开销的预测将被用作作下一步的开始搜索位置。

4 搜索策略

先选择大钻石搜索算法作为本文搜索策略的开始搜索算法, 然后再调用小钻石搜索算法。下面是

3.2 中值运动矢量^[6]

如图 1 所示, 中间预测器用于运动矢量的中间预测, 即把当前块的左、上和右上相邻块运动矢量的中值用作预测当前块的运动矢量

两种搜索算法的详细步骤:

4.1 小钻石搜索^[7,8]

小钻石搜索模式 (small diamond search pattern, SDSP) 是以搜索中心为中心, 并且所有的 SDSP 的检查点都已经测试过 (见图 3)。在小钻石搜索中所得到的一个最小失真度点, 就代表最小运动矢量。

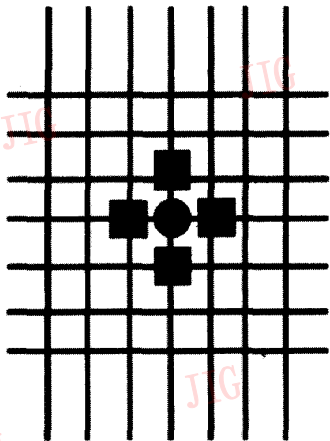


图 3 小钻石搜索
Fig. 3 Small diamond search

4.2 大钻石搜索^[7,8]

大钻石搜索步骤如下:

(1) 大钻石搜索模型 (large diamond search pattern, LDSP) 以搜索中心为中心, 并且 LDSP 所有的检查点都测试过 (见图 4)。如果这个中心位置得到了最小失真度, 则转到第 3 步; 否则转到第 2 步。

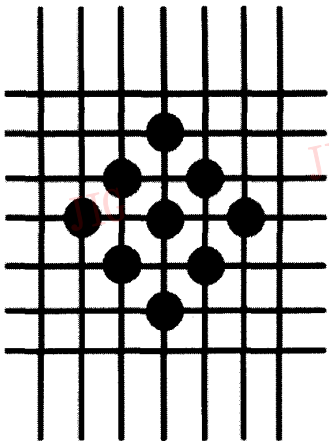


图 4 大钻石搜索初始搜索
Fig. 4 Frst step of large diamond search

(2) LDSP 中心移到上一步来得到最小失真度点, 并且所有的 LDSP 点都测试过 (见图 5)。如果

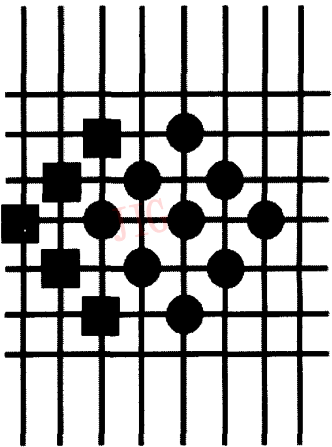


图 5 大钻石搜索再次搜索
Fig. 5 Repeat the large diamond search

中心位置得出了最小失真度, 则转到第 3 步, 否则循环这一步。

(3) 转换 LDSP 到 SDSP。得到的最小失真度之处, 就是最小运动矢量。

5 实验结果分析

实验的测试环境为: CPU 为 P4 1.6G, 内存为 256M, 操作系统为 Microsoft Windows Server 2003。

本算法是在 H. 264 参考模型 JM94 上实现的, 为验证本文算法的效果, 利用了 5 个具有不同运动特征的视频序列进行了测试, 其中 foreman 和 silent 序列的运动范围较大, 而 miss_am、container 和 claire 序列的局部运动则较为剧烈; 量化参数 (Quantity parameter, QP) 分别选择 28, 32, 36 和 40。下面是实验时 JM94 的统一设置:

- (1) 编码时搜索范围为 16pixels;
- (2) 参考帧数为 1;
- (3) I 帧的插入周期为 16;
- (4) 每秒帧率为 30;
- (5) 用了哈德马变换;
- (6) 熵编码方法为 CABAC (Context-based adaptive binary arithmetic coding);
- (7) 用 RD (rate distortion) 优化模式判定。

测试结果如表 1、表 2 和图 6 所示。

测试表明, 本文的快速搜索算法比 JM94 全搜索算法在码率增加不超过 1%, SNR 下降不超过 0.1dB 情况下, 编码速度较 H. 264 传统算法提高将近 65%。

表 1 不同搜索算法下搜索时间和比特率的比较 (量化参数 $Q = 28$)

Tab.1 Comparison of search time and bit rate in different search algorithms (QP: $Q = 28$)

测试序列	搜索算法	平均运动估计时间 (ms)	运动估计时间提高率 (%)	比特率 (kbit/s)
foreman (400)	全搜索算法	1344	64.21	166.37
	快速算法	481		166.93
miss_am (150)	全搜索算法	1328	65.66	44.66
	快速算法	456		44.55
container (300)	全搜索算法	1341	66.74	74.97
	快速算法	446		74.82
silent (300)	全搜索算法	1369	65.08	119.38
	快速算法	478		119.84
claire (494)	全搜索算法	1272	65.57	47.34
	快速算法	438		47.32

表 2 不同搜索算法下的信噪比比比较 (量化参数: $Q = 28$)

Tab.2 Comparison of SNR in different search algorithms (QP: $Q = 28$)

测试序列	搜索算法	SNR_Y (dB)	SNR_U (dB)	SNR_V (dB)
foreman (400)	全搜索算法	35.52	39.07	40.38
	快速算法	35.49	39.06	40.37
miss_am (150)	全搜索算法	40.35	39.3	39.5
	快速算法	40.33	39.31	39.52
container (300)	全搜索算法	36.41	41.32	41.11
	快速算法	36.4	41.32	41.1
silent (300)	全搜索算法	36.06	38.82	39.9
	快速算法	36.06	38.8	39.9
claire (494)	全搜索算法	40.25	40.29	42.71
	快速算法	40.25	40.28	42.69

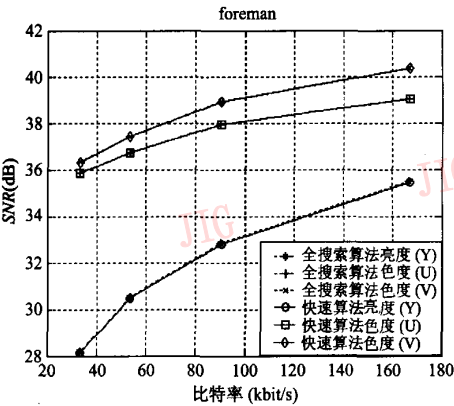


图 6 比特率和信噪比的关系图 (foreman)
Fig. 6 Comparison of bit rate and SNR (foreman)

参考文献 (References)

1 YAN De-cong, CHEN Geng, YANG Zong-kai. A new generation video coding standard H. 264/AVC [J]. Telecommunications Information, 2003, 4(4): 22 ~ 24. [严德聪, 程庚, 杨宗凯. 新一代运动图像压缩标准 H. 264/AVC [J]. 电信快报, 2003, 4(4): 22 ~ 24.]

2 SHEN Lan-sun, ZHOU Li. Video coding and low rate transmission [M]. Beijing: Publishing house of electronics industry, 2001. [沈兰荪, 卓力等. 视频编码与低速率传输 [M]. 北京: 电子工业出版社, 2001.]

3 Clark R J. Transform coding of images [M]. London: Academic Press, 1985.

4 Prabhudev Irappa Hosur, Kai-Kuang Ma. Motion vector field adaptive fast motion estimation [A]. In: Second International Conference on Information, Communications and Signal Processing (ICICS '99) [C], Singapore, 1999.

5 Alexis M Tourapis, Oscar C Au, Ming L Liou. Fast block-matching motion estimation using predictive motion vector field adaptive search technique [D]. Hon Kong: Department of Electrical and Electronic Engineering, The Hong Kong University of Science and Technology, 2000.

6 CHEN Z B, ZHOU P, HE Y. Fast integer pel and fractional pel motion estimation for JVT [R]. Beijing: Department of Electronic Engineering, Tsinghua University, 2002.

7 GONG Tao, DING R T. A modified motion estimation algorithm—Pseudo-diamond search method [J]. Journal of Image and Graphics, 2002, 7(11): 1181 ~ 1186. [龚涛, 丁润涛. 一种改进的有效运动估计方法——准菱形搜索法 [J]. 中国图象图形学报, 2002, 7(11): 1181 ~ 1186.]

8 Cheung C H, Po L M. A novel cross-diamond search algorithm for fast block motion estimation [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2002, 12(12): 1168 ~ 1177.