

# 基于 H.264 帧间预测的快速算法

宋 彬, 常义林, 周宁兆

(西安电子科技大学 8N 国家重点实验室, 陕西西安 710071)

**摘 要:** 本文在 H.264 编码算法的基础上, 提出了改变图像存储结构的 1/4 像素精度插值, 快速搜索位移矢量的四步法, 以及快速准确选择预测编码模式等多种新颖的帧间预测快速算法. 实验结果表明, 与未使用快速算法相比, 综合使用本文提出的多种快速算法, 在码率增加不到 1%, 恢复图像平均 PSNR(峰值信噪比)下降 0.1 dB 的情况下, 编码速度提高 13~17 倍, 满足实时视频通信的要求. 算法与 H.264 标准兼容, 可用于实际产品.

**关键词:** 视频通信; H.264; 1/4 像素精度; 位移矢量估计; 预测模式

**中图分类号:** TN919.8 **文献标识码:** A **文章编号:** 0372-2112(2006)01-0031-04

## Fast Inter Prediction Method Based H.264

SONG Bin, CHANG Yi-lin, ZHOU Ning-zhao

(National Key Lab on Integrated Services Network, Xidian University, Xi'an, Shaanxi 710071, China)

**Abstract** This paper proposes fast inter prediction methods for real-time video coding based on H.264. These methods include storing structure of 1/4-pel resolution interpolating image, four-step search algorithm for fast block motion estimation and fast decision of inter prediction mode. Simulation results show that inter macroblock encoding is obviously speeded up by using above methods. Furthermore, these algorithms are compatible with H.264 standard and can be used in practice.

**Key words** video communication; H.264; 1/4-pel resolution; estimation of motion vector; prediction mode

### 1 引言

随着第三代移动通信系统的出现和 IP(Internet Protocol)网络的迅速发展, 视频通信已经成为通信的主要业务之一. 继制定了 H.261, H.263, H.263+<sup>[1]</sup> 等视频压缩标准后, ITU-T 和 ISO/IEC 联合制定新的视频压缩编码标准 H.264<sup>[2]</sup>, 它同时也是 MPEG-4 第 10 部分的主要内容. 与其它视频编码标准相比, H.264 标准的目的在于更加有效地提高视频编码效率和对网络的适配性, 它可广泛应用于数字电视、无线视频通信和 IP 视频会议及其它多媒体业务.

然而, 由于 H.264 编码算法效率高, 导致算法复杂、运算量大, 给其用于实时视频通信带来困难. 特别是 H.264 使用 1/4 像素精度位移估值(包括 P 帧和 B 帧)、位移估值块大小和形状可变, 以及多种预测编码模式选择等提高编码效率的新方法, 使得 H.264 中预测帧的编码复杂度显著增加, 成为实时编码的瓶颈. 因此, 必须采用有效、实用的快速算法以提高 H.264 中预测帧的编码速度, 满足实时视频通信的要求.

本文的工作是基于 H.264 编码算法的特点, 针对帧间预测提出了新的参考图像存储方法, 位移矢量的快速搜索和预测编码模式判决等多种新颖的快速算法. 实验结果表明, 综合使用本文提出的多种帧间预测快速算法, 在基本

没有降低 H.264 编码效率的前提下, 大大提高了视频编码的速度, 满足实时视频通信的要求. 算法与 H.264 标准兼容, 可用于实际视频通信产品中.

### 2 H.264 的高效编码算法

与目前广泛使用的 H.263+ 标准相比, H.264 视频编码算法有许多新颖之处, 下面首先给出 H.264 编码算法的新特点<sup>[3]</sup>.

**算法的分层结构** H.264 编码算法总体上分为两层: 视频编码层完成对视频内容的有效描述, 它采用多种新颖的编码算法来提高编码效率; 网络适配层主要根据网络类型和网络的质量(QoS(服务质量)), 有效保证视频流和网络的适配.

**运动描述** H.264 支持不同形状和大小的块的运动补偿编码, 运动预测的精度也有所改进, 使用 1/4 像素精度. 另外, H.264 还允许使用多个参考帧来进行运动补偿.

**帧内编码** H.264 在帧内编码中使用了空间预测技术, 用已编码区域的边缘像素对帧内编码区域像素做空间预测, 这一技术有效提高了帧内编码效率.

**整数变换编码**在进行  $4 \times 4$  块变换时, H.264 使用整数变换编码代替传统的 DCT(离散余弦变换). 与 H.263+ 中  $8 \times 8$  块 DCT 相比, 这种  $4 \times 4$  块整数变换减小了方块效应, 且运算

收稿日期: 2004-07-21; 修回日期: 2005-12-01

基金项目: 国家自然科学基金(N060172030); 陕西省自然科学基金基础研究计划(2004F07); 华为基金; 综合业务网理论与关键技术国家重点实验室资助项目

速度快、结果精度高。

系数量化和熵编码 H. 264对变换系数做尺度变换归一化处理,然后量化,保证在 32位运算系统中计算结果精度高且不会溢出。在 H. 264中,可以使用基于内容的可变长编码和自适应二进制算术编码两种熵编码方法。

与以往的标准相比,H. 264可以提高 50%左右的编码效率<sup>[3]</sup>。但是,H. 264中新颖的编码算法运算量大、复杂度高,不利于低延时实时视频通信系统的实现。因此,本文针对 H. 264帧间预测中的主要耗时模块进行算法优化,意图满足实时视频通信的要求。

### 3 帧间预测的快速算法

在 H.264编码过程中,帧间预测占整个编码时间的 73%以上,因此本文重点优化帧间预测中主要耗时的步骤,包括:1/4像素精度插值的优化,快速搜索位移矢量和 H.264中多种预测编码模式的快速选择等。

#### 3.1 1/4像素插值的快速算法

H.264使用 1/4像素精度完成运动估计/补偿,与 H.263+的 1/2像素精度相比可得到 2dB的编码增益<sup>[4]</sup>。1/4像素插值的具体算法是:首先使用一个 6阶滤波器(1, -5, 20, 20, -5, 1)计算得到 1/2像素的值,然后使用整像素和半像素的线性插值得到 1/4像素位置的亮度预测值。而色度信号在编码端不参与运动估计,因此编码端不需要对它进行插值;在解码端进行运动补偿时,精确到 1/8像素,各亚像素的值通过双线性插值得到。显然 1/4像素精度运动矢量预测能够提高编码效率,但也增加了编解码器的运算复杂度和内存的读写调用,因此,必须对 1/4像素插值算法进行优化。

由前面的算法描述可知,SMD(单指令多数据流)技术非常适合插值过程的优化<sup>[5]</sup>,如采用 MMX(MultiMedia Extension)或 SSE(Stream SMD Extension)/SSE2等并行处理指令。但是,传统的插值方法使 1/2像素和 1/4像素在内存中不能连续存储,很难使用 SMD技术。因此,本文提出一种新的参考图像的数据组织方法,如图 1所示。



图 1 比较 1/4像素插值的图像组织方法

在图 1(c)中,整像素点存储在第 0幅图像中,1/2像素点存储在第 4 & 12幅图像中,其余图像存储 1/4像素点。由图 1可知,新的图像存储方法可以实现每个 1/2或 1/4像素位置对应的像素值在内存中连续存储,便于 SMD技术优化。如果将图 1(c)中 16幅图像单独存储,就会导致运动估计/补偿时在内存中跳跃式读取数据。因此,本文将这

16幅图像按编号顺序拼接成一幅大的图像存储在内存中,来减少往复访问内存对编码速度的影响。

另外,计算 1/4像素插值的中间过程中,在保证对最终结果没有任何影响的前提下,本文对多个临时计算结果限幅,从而不是使用 32位操作数,而是完全使用 16位数据操作计算,有利于使用 SMD并行计算。

#### 3.2 位移矢量的快速搜索算法

H.264支持不同形状和大小的编码块,其中包括:16×16,16×8,8×16,8×8,8×4,4×8,4×4。在搜索位移矢量时,所用的匹配函数是 SAD(绝对差值和)。如果采用全搜索方法估值编码块的位移矢量(MV),那么求每一个假定位移矢量的匹配函数值,并比较它们,其运算量非常大。因此,针对 H.264中不同形状和大小编码块,本文提出一种新的四步法快速搜索位移矢量,具体算法如下:

步骤 1 确定位移矢量的初始值和终止门限

MV的初始值是由一组相关块的已知位移估值和特殊点的位移估值组成,主要是为运动搜索过程提供更接近最佳 MV的搜索起点,减少搜索次数,下面给出 MV初始值集合  $S_{MVP}$ :

$$S_{MVP} = \{MVP_0, MVP_{MED}, MVP_L, MVP_{UL}, MVP_U, MVP_{UR}, MVP_{LAST}\} \quad (1)$$

式(1)中  $MVP_0$ 表示零运动矢量, $MVP_L, MVP_{UL}, MVP_U, MVP_{UR}$ 分别表示当前匹配块左方、左上方、上方和右上方的 MV, $MVP_{LAST}$ 是前一帧同位置匹配块的 MV, $MVP_{MED}$ 是  $MVP_L, MVP_{UL}, MVP_U$ 的中值。终止门限  $T_1, T_2$ 的取值:

$$T_1 = 2 \times W_{block} \times H_{block} \quad T_2 = T_1 / 2 \quad (2)$$

式(2)中  $W_{block}$ 和  $H_{block}$ 分别表示匹配块的宽度和高度。

步骤 2 确定搜索中心和运动强度因子

分别搜索集合  $S_{MVP}$ 中指定的 7个位置,计算得到最小 SAD:

$$SAD_{min} = \underset{i \in S_{MVP}}{Min} \left( \sum_{y=0}^{H_{block}-1} \sum_{x=0}^{W_{block}-1} MB_0(x, y) - MB_i(x, y) \right) \quad (3)$$

式(3)中  $MB_0(x, y)$ 表示当前块中坐标  $(x, y)$ 的像素值, $MB_i(x, y)$ 表示集合  $S_{MVP}$ 中指定位置块坐标  $(x, y)$ 的像素值。经过式(3)计算,将  $SAD_{min}$ 对应位置作为下一步的搜索中心。

如果式(3)在计算过程中,SAD值小于门限  $T_1$ ,则认为该位置对应的 MV为最佳 MV,并结束整个搜索过程;否则,根据式(3)得到的最佳 MV预测值  $(mvx', mvy')$ 来计算运动强度因子  $M = |mvx'| + |mvy'|$ 。若  $M > M_c$  ( $M_c$ 为运动强度门限,与视频序列的内容相关),搜索执行下面的步骤 3;否则执行步骤 4。

步骤 3 六边形搜索

本文在 ChunHo Cheung快速搜索算法<sup>[6]</sup>的基础上,提出一种新的 8点六边形搜索算法,如图 2(a)所示,图中搜索中心点由步骤 2得到。本文根据文献[6]中的运动统计结果和六边形搜索算法,在搜索中心点正上方和正下方加入两个搜索位置(图 2(a)中方形点),在基本没有增加运

算复杂度的前提下,大大提高了位移估值的准确度.

计算图 2(a) 9 个位置点中 SAD 最小的一个位置  $L_{SAD}$ . 如果  $L_{SAD}$  是搜索中心, 就进行步骤 4 的搜索; 否则, 以  $L_{SAD}$  为新的搜索中心, 重复步骤 3 另外, 如果在搜索过程中某个位置 SAD 值小于门限  $T_2$ , 则认为该位置对应的 MV 为最佳 MV, 并结束整个搜索过程.

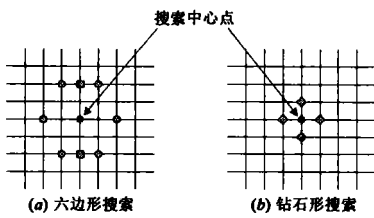


图 2 搜索位置示意图

步骤 4 钻石形精细搜索

计算图 2(b) 中 5 个位置点的 SAD, 最小 SAD 对应位置  $L'_{SAD}$ . 如果  $L'_{SAD}$  是搜索中心, 就认为该位置对应的 MV 为最佳 MV, 并结束搜索; 否则, 以  $L'_{SAD}$  为新的搜索中心, 重复步骤 4 另外, 如果在搜索过程中某个位置 SAD 值小于门限  $T_2$ , 则认为该位置对应的 MV 为最佳 MV, 并结束搜索.

3.3 编码模式的快速判决算法

由前面的讨论可知, 在帧间预测时, 由于 H. 264 支持不同形状和大小的块, 因此编码一个宏块 (MB) 需要判决 9 种预测模式, 其中 NTRA\_ 4x4, NTRA\_ 16x16 分别表示亮度信号 4x4 和 16x16 帧内预测模式, 而 NTER\_ 16x16, NTER\_ 16x8, NTER\_ 8x16, NTER\_ 8x8, NTER\_ 8x4, NTER\_ 4x8, NTER\_ 4x4 分别表示 16x16, 16x8, 8x16, 8x8, 8x4, 4x8, 4x4 帧间预测模式. 如果每个宏块都要对这 9 种预测编码模式预运行, 然后根据编码效率比较和取舍, 则编码器的运算量非常大, 同样影响实时视频通信. 本文利用视频序列的时间-空间相关性提出一种判决预测模式的快速算法.

首先将视频图像划分为背景区 ( $S_{BC}$ )、亚背景区 ( $S_{SBC}$ ) 和运动区 ( $S_{MC}$ ) 3 类集合, 其中  $S_{BC}$  中的宏块运动很小,  $S_{SBC}$  中宏块的运动属于中等, 而  $S_{MC}$  中的宏块运动较大. 下面以第  $N$

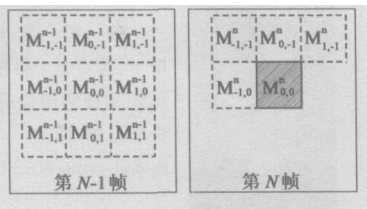


图 3 对当前宏块的运动划分

帧图像的宏块  $M_{0,0}^n$  (图 3 中的阴影宏块) 为例, 给出这三类区域的划分方法, 其中,  $M_{0,0}^n$  上标  $n$  表示宏块所在帧, 下标  $(0,0)$  表示宏块在图像中的行、列坐标 (本文将当前宏块位置定为  $(0,0)$ ). 具体划分过程如下:

(1) 在 H. 264 基本框架中, 整像素精度 MV 的搜索范围在  $\pm 16$  之间, 因此如果宏块  $M_{0,0}^n$  在第  $N-1$  帧中对应位置宏块  $M_{0,0}^{n-1}$  及其周围相邻 8 个宏块都属于  $S_{BC}$ , 则第  $N$  帧宏块  $M_{0,0}^n$  也属于  $S_{BC}$ .

(2) 如果第  $N-1$  帧中 9 个宏块没有一个属于  $S_{BC}$ , 则宏块  $M_{0,0}^n$  属于  $S_{MC}$ .

(3) 如果宏块  $M_{0,0}^{n-1}$  不属于  $S_{BC}$ , 但其周围相邻 8 个宏块中有属于  $S_{BC}$  的情况下, 就要利用图像序列的空间相关性来做进一步判断, 具体过程分二步:

第一步 计算与  $M_{0,0}^n$  相邻宏块 SAD 的平均值  $SAD_{avg}$ . 由于 H. 264 编码过程的限制, 只有与当前宏块左方、左上方、上方和右上方四个相邻宏块的 SAD 可用, 如图 3 第  $N$  中的宏块  $M_{-1,0}^n, M_{-1,-1}^n, M_{0,-1}^n$  和  $M_{1,-1}^n$ . 因此平均 SAD 的计算如下:

$$SAD_{avg} = (SAD_{mb}(M_{-1,0}^n) + SAD_{mb}(M_{-1,-1}^n) + SAD_{mb}(M_{0,-1}^n) + SAD_{mb}(M_{1,-1}^n)) / 4 \quad (4)$$

式中函数  $SAD_{mb}(\cdot)$  是计算宏块与参考图像最佳匹配区域的 SAD 值.

第二步 比较  $SAD_{avg}$  与  $\min SAD_{mb}(M_{0,0}^n)$  的大小, 其中使用表 1 中  $S_{BC}$  对应的可选编码模式来计算  $\min SAD_{mb}(M_{0,0}^n)$ . 如果  $\min SAD_{mb}(M_{0,0}^n)$  大于  $SAD_{avg}$ , 则宏块  $M_{0,0}^n$  属于  $S_{MC}$ , 否则宏块  $M_{0,0}^n$  属于  $S_{SBC}$ .

(4) 在其它情况下,  $M_{0,0}^n$  属于  $S_{SBC}$ .

表 1 不同运动集合中的可选编码模式

不同运动集合	可选编码模式
$S_{BC}$	NTRA_ 16x16, NTER_ 16x16
$S_{SBC}$	NTRA_ 4x4, NTRA_ 16x16, NTER_ 16x16, NTER_ 16x8, NTER_ 8x16
$S_{MC}$	NTRA_ 4x4, NTRA_ 16x16 以及所有帧间预测模式

对宏块进行运动划分后, 使用表 1 中对应的可选编码模式, 可以大大简化 H. 264 中预测模式的判决过程, 从而降低编码器的运算复杂度.

另外, 为防止快速模式判决算法引入的预测误差, 本文在帧内编码 ( $I$ ) 帧后的第一个预测帧  $F_{m,r}$  进行预测模式刷新, 即  $F_{m,r}$  中所有宏块都属于  $S_{MC}$ , 都要完成表 1 中对应的所有可选编码模式. 这样, 随着周期性  $I$  帧刷新, 就可以彻底去除预测误差积累, 保证帧间预测的准确性.

4 实验结果

本文使用三个典型的标准图像序列“foreman”(运动剧烈)、“news”(中等运动)、“claire”(运动较小)共 300 帧 (重复 20 次) 进行实验研究, 它们是 QCIF 格式, Y:U:V 是 4:2:0 实验平台是 PIV 2GH z 计算机和自主开发的 H. 264 软件编解码器, 下面实验结果中的数据均以此开发平台上的时间为准. 并用恢复图像亮度信号的平均 PSNR 和编码比特数/帧作为衡量 H. 264 编码性能和效率的测度. 在 H. 264 编码器中, 设定量化因子 QP 为 28 可选择使用表 1 中  $S_{MC}$  对应的所有编码模式, 满足 H. 264 基本框架 (Baseline Profile) 要求, 3.1 节中运动强度门限  $M_c$  取值为 1.

首先, 给出 3.1 节中 1/4 像素插值快速算法 (FPIA) 的实验结果, 如表 2 所示. 表中给出未改变存储结构 (无优化), 和改变存储结构后使用 SMD 技术优化插值计算的比较结果.

表 2 1/4像素插值的实现算法比较

图像序列 (6000帧)	无优化	MMX优化	SSE2优化
Claire	33 88秒	7 81秒	7 55秒
News	33 51秒	7 64秒	7 38秒
Foreman	33 42秒	7 43秒	7 27秒

由表 2 的实验结果可以看出,在本文提出的图像存储结构基础上,使用 SMD 技术进行优化,1/4像素插值速度提高 3 倍以上,且没有改变预测精度,不会影响 H.264 编码效率。

然后,在 FPIA 算法的基础上,加入运动矢量快速搜索算法 (FMVSA) 和预测模式快速判决方法 (FPM SA),实验结果见表 3 和图 4 图 5 其中,表 3 给出 H.264 编码器的编码速度,图 4 图 5 以运动剧烈的图像序列“foreman”为例给出 H.264 编码效率的变化。

表 3 多种优化算法的编码速度比较 (单位: 帧/秒)

图像序列	未使用优化算法	FPIA 算法	FPIA 算法 + FMVSA 算法	FPIA 算法 + FMVSA 算法 + FPM SA 算法
claire	4.215	5.238	50.531	75.501
news	4.039	5.098	46.375	66.930
foreman	3.656	4.791	37.212	51.835

从表 3 的实验结果可知,使用 FPIA 算法后整个编码帧率提高 24% ~ 31%,而在 FPIA 算法上加入 FMVSA 算法,与全搜索算法相比编码速度可提高 7~9 倍,最后加入 FPM SA 算法, H.264 编码器速度可再提高 40% ~ 50%。由表 3 和图 4 图 5 可以看出,与未使用优化算法相比,综合使用本文提出的多种快速算法,在码率增加不到 1%,恢复图像亮度信号的平均 PSNR 下降 0.1dB 的情况下,编码速度提高 13~17 倍,可以满足以计算机为平台的实时视频通信要求。

## 5 结论

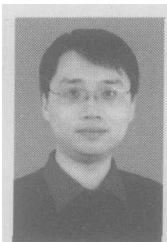
本文首先给出 H.264 编码算法的特点,然后提出了改变图像存储结构的 1/4 像素精度插值,快速搜索位移矢量

的四步法,以及快速准确选择预测编码模式等多种新颖的帧间预测快速算法。实验结果表明,与未使用快速算法相比,综合使用本文提出的多种快速算法,在基本没有影响 H.264 编码效率的前提下,编码速度得到显著提高,可以满足实时视频通信的要求。算法与 H.264 标准兼容,可用于实际视频通信产品中。

## 参考文献:

- [1] ITU Telecom. Standardization Sector of ITU. Video coding for low bit rate communication [S]. ITU-T Recommendation H.263 Version 2, 1998.
- [2] ITU Telecom. Standardization Sector of ITU. Advanced video coding for generic audiovisual services [S]. ITU-T Recommendation H.264, 2003.
- [3] Thomas Wiegand, Gary J. Sullivan, Gisle Bjontegaard et al. Overview of the H.264/AVC video coding standard [J]. IEEE Trans on Circuits and Systems for Video Tech, 2003, 13(7): 560-576.
- [4] Thomas Wedi, Hans Georg Musmann. Motion and aliasing-compensated prediction for hybrid video coding [J]. IEEE Trans on Circuits and Systems for Video Tech, 2003, 13(7): 577-586.
- [5] 宋彬, 常义林. 实时 H.263+ 视频编码器的 DSP 实现 [J]. 通信学报, 2003, 24(8): 88-94.
- [6] Chun Ho Cheung, Lai Man Po. A novel cross-dimension search algorithm for fast block motion estimation [J]. IEEE Trans on Circuits and Systems for Video Tech, 2002, 12(12): 1168-1177.

## 作者简介:



宋彬 男, 1973 年 11 月生于河南郑州, 博士, 西安电子科技大学副教授, 主要从事图像压缩和视频通信等领域的研究工作, 发表学术论文近 20 篇。

E-mail: bsong@mail.xidian.edu.cn



常义林 男, 1944 年 12 月生于江苏江都, 西安电子科技大学教授, 博士生导师, 主要研究方向为多媒体通信和网络管理等, 多次获得省部级科技进步奖励, 发表学术论文 100 多篇。

E-mail: yldhang@xidian.edu.cn

周宁兆 男, 1978 年 11 月生于浙江宁波, 西安电子科技大学硕士研究生, 研究方向为多媒体通信。

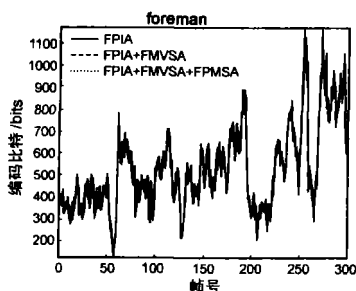


图 4 每帧图像的编码比特数

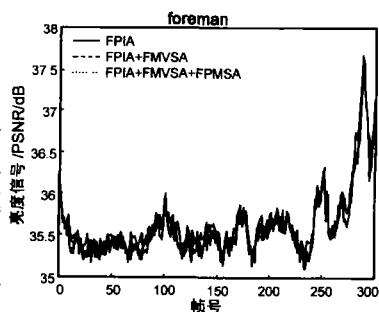


图 5 恢复图像亮度信号的 PSNR