

H.264 帧间预测模式的快速选择算法

宋 彬, 常义林, 李春林

(西安电子科技大学 ISN 国家重点实验室, 陕西西安 710071)

摘 要: 根据 H.264 预测帧的编码算法, 本文研究了利用图像序列的时/空相关性, 正确选择预测编码模式的快速算法, 并提出了模式选择参数的自适应更新算法. 实验结果表明, 与未使用快速模式选择算法相比, 在码率增加 2%, 恢复图像平均 PSNR(峰值信噪比)下降不到 0.2 dB 的情况下, 使用快速算法的编码速度提高 2~3 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; 预测模式; 时/空相关性

中图分类号: TN919.8 **文献标识码:** A **文章编号:** 0372-2112(2007)04-0697-04

Novel Fast Selecting Inter Prediction Mode Algorithm Based H.264

SONG Bin, CHANG Yi-lin, LI Chun-lin

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

Abstract: Based on H.264, this paper proposes a novel fast selecting inter prediction mode algorithm by using temporal/spatial correlation of video sequence. Two key parameters by which inter prediction mode is determined is adaptive updated. 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; prediction mode; temporal/spatial correlation

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+^[1] 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^[2], 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它对网络的适配性, 其编码算法可广泛应用于数字电视、无线视频通信和 IP 视频会议等.

然而, 由于 H.264 编码算法效率高, 导致算法复杂、运算量大, 给实时视频通信带来困难. 特别是 H.264 预测帧(包括 P 帧和 B 帧)中编码块大小和形状可变, 有多种预测模式可选, 成为实时编码的主要瓶颈. 因此, 需要在保证帧间预测的编码效率情况下, 采用多种有效、实用的算法提高 H.264 预测帧的编码速度, 满足实时视频通信的要求.

在给出 H.264 预测帧编码算法的基础上, 本文首先利用图像序列时/空相关性提出快速、准确选择预测模

式的算法; 然后基于量化因子提出模式选择参数的自适应更新算法. 实验结果表明, 使用本文提出的预测模式快速选择算法, 在 H.264 编码效率降低很少的情况下, 大大提高了帧间预测的速度, 满足实时视频通信的要求. 且算法与 H.264 标准兼容, 可用于实际视频通信产品中.

2 H.264 预测帧的编码算法

与目前使用最广泛的 H.263+ 标准相比, H.264 视频编码算法有许多新颖之处, 特别是在编码预测帧时, H.264 提出编码块大小和形状可变、多参考帧预测、1/4 像素精度的位移估值等多种提高编码效率的新算法^[3]. 为了讨论帧间预测模式快速选择算法, 下面简要给出 H.264 的有关算法.

编码块大小和形状可变 H.264 支持 7 种大小和形状的块: 16×16 、 16×8 、 8×16 、 8×8 、 8×4 、 4×8 、 4×4 , 如图 1 所示. 根据采用块大小的不同, 宏块的编码模式分为 4 种: $P_{16 \times 16}$ 、 $P_{16 \times 8}$ 、 $P_{8 \times 16}$ 和 $P_{8 \times 8}$. 其中, 在 $P_{8 \times 8}$ 模式下, 每个宏块又被分为 4 个 8×8 子块, 而每个子块又有 4 种可能的子模式: $PS_{8 \times 8}$ 、 $PS_{8 \times 4}$ 、 $PS_{4 \times 8}$ 和

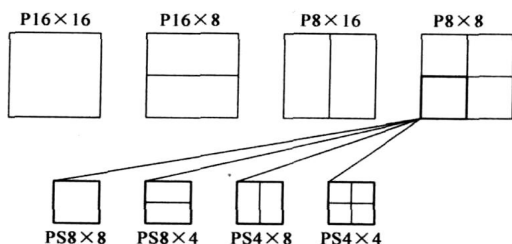


图1 H.264中大小和形状可变的编码块

PS4×4. H.264 采用尺寸可变块编码, 大大提高宏块运动估值的准确性, 从而提高编码效率.

1/4 像素精度 H.264 使用了 1/4 像素精度的运动估值/补偿算法, 与 H.263+ 的 1/2 像素精度相比可得 2dB 编码增益^[4], 具体算法是: 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, 再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.

多参考帧编码 在 H.264 中, 支持采用多个参考帧进行帧间预测, 这在场景交替切换等情况下, 可以提高编码效率.

帧内预测模式 H.264 预测帧中每个宏块都要做帧内预测编码^[3], 具体算法是: 将图像内已编码区域的边缘像素用于待编码区域的空间预测, 对亮度信号既要 4×4 预测 (I4×4) 还要做 16×16 预测 (I16×16), 这一技术有效提高了帧内预测的编码效率.

在 H.264 预测帧中, 每个宏块的预测模式主要有 6 种: 帧间预测的 P16×16、P16×8、P8×16 和 P8×8, 以及帧内预测的 I4×4 和 I16×16. 而 P8×8 模式又包括四种子预测模式: PS8×8、PS8×4、PS4×8、PS4×4. 由此可见, H.264 预测帧的编码算法运算量大、复杂度高, 不利于低延时视频通信系统的实现. 因此, 本文针对 H.264 预测帧的多种预测模式选择进行优化, 来满足实时视频通信的要求.

3 选择预测模式的快速算法

在 H.264 编码过程中, 帧间预测所需时间占整个编码时间的 75% 以上, 因此本文重点优化帧间预测中耗时的预测模式选择模块. 然后, 权衡量化因子 (QP) 和最佳预测模式的关系, 提出模式选择参数的自适应更新算法, 进一步降低帧间预测的运算复杂度.

3.1 利用时/空相关性的快速算法

由前面的讨论可知, 在帧间预测时, 如果每个宏块都要对 6 种预测模式和 4 种子预测模式预运行一遍, 然后选择编码效率最佳的模式, 则运算量非常大, 影响实时视频通信. 本文利用视频序列的时/空相关性提出一种选择预测模式的快速算法.

首先将视频图像划分为背景区 (S_{BG})、准背景区

(S_{SBC}) 和运动区 (S_{MC}) 三类集合, 其中 S_{BG} 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, S_{MC} 又根据运动程度进一步划分为 S_{MC-A} 和 S_{MC-B} 两个子集. 本文提出的视频图像划分算法是基于图像采用的最佳编码模式与其内容的关系, 如静止区域和运动简单的区域多采用匹配块较大的帧间编码模式, 而运动复杂的区域多采用匹配块较小的帧间编码模式^[5].

下面以第 n 帧

图像的宏块 $M_{0,0}^n$ (图 2 中的阴影宏块) 为例, 给出运动

区域的划分方法, 其中, $M_{0,0}^n$ 上标 n

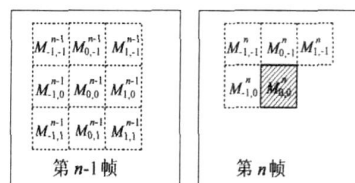


图2 对当前宏块的运动划分

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 ± 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_{BG} , 则第 n 帧宏块 $M_{0,0}^n$ 也属于 S_{BG} .

(2) 如果第 $n-1$ 帧中 9 个宏块没有一个属于 S_{BG} , 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_{BG} , 但其周围相邻 8 个宏块中有属于 S_{BG} 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} . 在搜索位移矢量时, 所用的匹配函数是 SAD (绝对差值和). 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 2 第 n 帧中的宏块 $M_{-1,0}^n$, $M_{-1,-1}^n$, $M_{0,-1}^n$ 和 $M_{1,-1}^n$, 因此平均 SAD 的计算如下:

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

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, α 是与 QP 相关的常数.

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_{BG} 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) 在其他情况下, $M_{0,0}^n$ 属于 S_{SBC} .

(5) 如果 $M_{0,0}^n$ 属于 S_{MC} , 则重复 (3) 中的两步算法进一步选择预测模式, 其中使用表 1 中 S_{SBC} 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} 的子集 S_{MC-B} , 否则

宏块 $M_{0,0}^n$ 属于子集 S_{MC-A} .

表 1 不同运动集合中的可选编码模式

运动集合	可选编码模式
S_{BG}	$116 \times 16, P16 \times 16$
S_{SBG}	$116 \times 16, P16 \times 16, P16 \times 8, P8 \times 16$
S_{MC-A}	$14 \times 4, 116 \times 16, P16 \times 16, P16 \times 8, P8 \times 16, P8 \times 8 (PS8 \times 8)$
S_{MC-B}	$14 \times 4, 116 \times 16, P16 \times 16, P16 \times 8, P8 \times 16, P8 \times 8 (PS8 \times 8, PS8 \times 4, PS4 \times 8, PS4 \times 4)$

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,降低编码器的运算复杂度.

为防止快速模式判决算法引入的预测误差,在每帧后对预测帧进行预测模式更新,即该帧内所有宏块都属于 S_{MC-B} ,都要完成表 1 中对应的可选编码模式.这样,随着周期性的更新,就可以减小预测误差,提高帧间预测的准确性.下面详细讨论 SAD_{avg}' 中 α 的取值.

3.2 帧间预测模式的自适应更新算法

在帧间预测模式中,可使用率失真优化方法(RDO)对预测模式进行判决,即将预测误差和信息编码码率作为判决依据[6].

$$J(MB_k, I_k | QP) = D(MB_k, I_k | QP) + \lambda_{mode} R_{REC}(MB_k, I_k | QP) \quad (2)$$

式中, $D(MB_k, I_k | QP)$ 表示宏块 I_k 表示宏块 I_k 的失真, λ_{mode} 为模式选择的拉格朗日乘子.

样会有更多的宏块属于子集 S_{MC-A} ,从而降低帧间预测的复杂度.同时, α 值也越来越小,基本不会影响预测的准确性.

使用式(4)计算 α ,当 QP 增大时,帧间预测的速度.然而,这一算法也使得快速预测模式判决的准确性降低.下式来计算周期间隔 N_{prof} ,自适应地减少 N_{prof} 对 H.264 编码效率的影响:

$$N_{prof} = N_I - \left(\frac{QP + \beta}{6} \right) \cdot \frac{Num_I}{N_I}$$

式中 N_I 为帧内(I)帧的刷新周期, Num_I 为帧内(I)帧的个数.由式(5)可知,随着 QP 值的增大, N_{prof} 减小,即更新的周期间隔越来越小,从而带来的预测模式判决误差也越来越小.

另外,为了提高预测精度,本文还对像素精度插值、运动矢量搜索等模块进行优化,综合满足实时视频通信的要求.

4 实验结果

本文使用 H.264 编码器(剧烈)、“重帧” (重帧)进行实验.

的 α :

$$\alpha \in [0, 51]$$

计算, α 是量化因子 QP

可以

$$D_{avg} = \frac{1}{N} \sum_{k=1}^N D_k$$

更多的宏块

式(5)中

0.2%	146.6
------	-------

从表 2 可以看出, 使用 FPMD 快速算法后, 在恢复图像亮度信号平均 PSNR 下降不到 0.2dB, 码率增加小于 2% 的情况下, 编码速度提高 2~3 倍, 可以满足以计算机为平台的实时视频通信要求。

FPMD 快速算法准确性的实验是以“foreman”序列(QCIF 格式)的第 9 帧为例, 实验结果如图 3 所示, 图中实线给出了宏块的最佳预测模式。

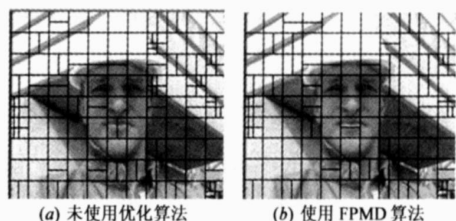


图 3 第 9 帧图像的预测模式准确性比较

从图 3 可以看出, 使用 FPMD 快速算法, 可以较为准确地划分图像的 S_{BG} 区域, 而对区域 S_{SBC} 和 S_{MGC} 的划分, 算法存在判决误差。

以运动剧烈的图像序列“foreman”(QCIF 格式)为例, 进一步给出了在 FPMD 算法的基础上, 使用模式选择参数的自适应更新算法(AFPMD)的实验结果, 如表 3 所示, 表中 PSNR 和编码比特数的变化都是与未使用模式选择优化算法相比的结果。

表 3 AFPMD 算法的实验结果

编码方法 量化因子	FPMD 算法			AFPMD 算法		
	平均 PSNR 的变化/dB	编码比特 数的变化	编码速度 (帧/秒)	平均 PSNR 的变化/dB	编码比特 数的变化	编码速度 (帧/秒)
18	-0.12	+1.0%	58.9	-0.13	+1.2%	60.9
28	-0.15	+1.9%	87.3	-0.17	+2.1%	92.2
38	-0.20	+2.2%	135.6	-0.20	+2.5%	149.2

从表 3 可以看出, 与 FPMD 算法相比, AFPMD 算法在 PSNR 和编码比特数变化很小的情况下, 又将编码速度提高了 3%~10%。而且实验结果表明, QP 越大, 效果越显著。因此, AFPMD 算法在保证编码效率的前提下, 进一步提高了编码速度。

5 结论及今后工作

本文利用图像序列的时/空相关性, 提出了 H.264 中预测帧编码算法的快速预测模式选择算法, 并根据量化因子自适应更新模式选择参数。实验结果表明, 与未使用快速算法相比, 使用本文提出的预测模式选择快速算法, 在对 H.264 编码效率影响较小的前提下, 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。

今后, 要研究快速而更加准确的运动区域(包括

S_{SBC} 和 S_{MGC}) 跟踪算法, 从而进一步提高 H.264 中预测帧的编码速度。

参考文献:

- [1] ITU Telecom. Standardization Sector of ITU. Video Coding for Low Bit Rate Communication[S]. ITU-T Recommendation H.263 Version2. 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 Transaction on Circuits and Systems for Video Technology, 2003, 13(7): 560-576.
- [4] Thomas Wedi, Hans Georg Musmann. Motion and aliasing compensated prediction for hybrid video coding[J]. IEEE Transaction on Circuits and Systems for Video Technology, 2003, 13(7): 577-586.
- [5] M Chan, Y Yu, A Constantinides. Variable size block matching motion compensation with applications to video coding[A]. IEE Proceedings[C]. 1990, 137(4): 205-212.
- [6] Thomas Wiegand, Heiko Schwarz, Anthony Joch, et al. Rate constrained coder control and comparison of video coding standards[J]. IEEE Transaction on Circuits and Systems for Video Technology, 2003, 13(7): 688-703.

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 20 多篇。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: ylchang@xidian.edu.cn

李春林男, 1980 年生于山西太原, 西安电子科技大学硕士研究生, 研究方向为多媒体通信。