

· 计算机工程与应用 ·

P帧快速模式选择算法

卢光辉

(电子科技大学计算机科学与工程学院 成都 610054)

【摘要】针对H.264编码算法中, 预测帧(P帧)编码模式的选择过程异常复杂, 开销大的问题, 提出了一种新的快速模式选择算法, 该算法利用直流系数能预先判定P帧编码模式的有效范围, 缩小备选编码模式的可选范围, 从而使模式选择的复杂度大为降低。实验结果表明, 该算法可以使编码时间降低54%~74%, 峰值信噪比提高0.05 dB~0.17 dB, 同时码率变化范围为-3.14%~3.97%。应用结果表明, 该算法在基于TI's TMS320DM642视频压缩系统中1路D1视频的压缩速度为25 帧/s。

关键词 H.264; 模式选择; 视频信号处理; 预测帧

中图分类号 TP37

文献标识码 A

doi:10.3969/j.issn.1001-0548.2010.03.016

Fast Mode Checking Algorithm for P-Frame

LU Guang-hui

(School of Computer Science and Engineering, University of Electronic Science and Technology of China Chengdu 610054)

Abstract The procedure of mode checking is very complex. It costs much time for predictive-frame (P-Frame) in the encoding algorithm of H.264. In this paper, a fast mode checking algorithm is presented to reduce the complexity of mode checking procedure based on direct current (DC) coefficient by reducing the number of modes required to be checked. Experimental results show that the proposed algorithm can reduce the encoding time by 54%~74% and increase the peak signal-to-noise ratio (PSNR) by 0.05 dB~0.17 dB while incurring up to 3.14% reduction or up to 3.97% increment on the total bit rate. Application shows that its encoding speed is 25 frames per second in a video encoder with TI's TMS320DM642 for D1 video signal.

Key words H.264; mode checking; predictive-Frame; video signal processing

H.264/AVC^[1]是ITU-T和ISO/IEC联合制定的最新国际视频编码标准, 由于突出的编码性能, 使其可以广泛应用于IPTV(网络电视)、数字电视、视频电话等领域, 尤其适用于无线通信等低速率传输领域。H.264与H.263、MPEG-4视频编码标准一样, 采用了基于宏块的编码架构, 但它比之前的视频编码标准更加先进, 支持宏块多模式(包含有多种尺寸的分块)、多参考帧、16×16/4×4帧内复杂预测、1/4像素精度的运动估计/补偿、4×4整数离散余弦变换、CAVLC/CABAC(context-based adaptive variable-length coding/context-based adaptive binary arithmetic coding)编码以及环路滤波等诸多新的算法, 使其视频编码效率十分显著^[2]。同时, 也由于H.264采用了这些复杂的编码技术使其运算复杂度远高于先前标准, 特别是预测帧中模式选择尤为复杂。根据参考软件JM上的测试发现, 模式选择的开销占整个编码

开销的85%以上, 成为实时编码的主要障碍, 极大地限制了H.264的广泛应用。因此, H.264在实现时首先要考虑的问题是如何降低编码复杂度, 特别是降低宏块模式选择的复杂度。

H.264编码模式的选择主要有两种方法, 一种是用待编码的宏块数据与其预测值的绝对差值之和(SAD)计算出宏块的最佳模式; 另一种则是采用率失真优化(RDO)的方法确定宏块的最佳模式^[3]。通常情况下, 两种方法都需要对每一种可能的模式分别进行检查, 使得编码过程异常复杂, 大大增加了编码所需要的时间。

文献[4]采用率失真优化的方法进行模式选择, 首先根据率失真代价对宏块模式进行分类, 然后按一定的次序对宏块的各个模式类型计算其编码代价, 当某个模式类型的编码代价小于预先设定的阈值, 剩余模式类型的计算就不再进行了。文献[5]通过

收稿日期: 2009-11-04; 修回日期: 2010-01-25

基金项目: 国家863计划(2007AA01Z322、2006AA01Z335); 四川省科技支撑计划(2009GZ0017)

作者简介: 卢光辉(1971-), 男, 博士, 副教授, 主要从事图形图像、计算机网络等方面的研究。

计算视频对象的纹理特征以及图像序列的时间平稳性预先判断宏块的可能模式,从而简化宏块模式的选择过程。文献[6]根据当前帧和前一帧中已编码的相邻宏块模式预测当前宏块的可能模式,以减少宏块模式的判定过程。文献[7]使用图像序列时/空相关性比较准确地选择预测模式及基于量化因子,使模式选择参数自适应更新以提高模式选择的速度。文献[8]首先计算P帧中宏块在SKIP模式下的率失真代价,当率失真代价小于某个最低阈值时,宏块就被判断为SKIP模式;当率失真代价大于某个最高阈值时,宏块就只检测它的帧内模式,否则根据宏块的运动特性把可能的模式分为三类,再在每一类中对其最终模式进行判定。文献[9]利用率失真优化方法和动态阈值技术对P帧中SKIP模式和P16×16模式进行预判,以减少对后继模式的判定计算。

以上方法都是对宏块的一些不必要的模式预先剔除,然后再在余下的可能模式中计算出最终的模式。它们中有的方法预先剔除的模式太少,导致余下的可能模式太多,从而使模式判断的开销较大;而有的则额外运算后预先剔除的模式太多,余下的可能模式太少,从而使模式判定不够准确,重建的图像质量损失较大。

本文主要研究P帧宏块模式的快速判定问题。首先计算宏块在SKIP模式下的当前值与预测值的残差形成的直流系数,然后结合P16×16模式下的率失真代价,对宏块的一些不必要的可能模式进行剔除,达到快速选择宏块模式的目的。该方法由于计算直流系数开销小,因此带来的额外计算少;此外,该方法对宏块余下的可能模式预判准确,使重建的图像质量损失小。

1 H.264 P帧编码算法

在讨论P帧宏块模式的快速选择算法前,先对H.264的相关算法进行简单介绍。

1.1 编码模式

H.264的编码模式主要分为帧内编码和帧间编码两大类。帧内编码利用图像的空间相关性进行预测编码。亮度信号的帧内编码模式又分为I16×16和I4×4两类。16×16区域的帧内预测有4种I16×16,4×4区域的帧内预测有9种I4×4。P帧帧间编码则利用视频序列时间上的相关性,把经过运动补偿的上一帧(几帧)图像的数据作为预测值。帧间编码按照分块的大小首先分为P16×16、P16×8、P8×16、P8×8 4种模式,其中,对于P8×8模式,宏块被分为4个8×8的子

块,每一个子块中又包含PS8×8、PS8×4、PS4×8、PS4×4子模式。

P帧宏块模式有SKIP、P16×16、P16×8、P8×16、P8×8(包含PS8×8、PS8×4、PS4×8、PS4×4)、I4×4、I16×16等7种模式,其中SKIP模式是P16×16的一种特殊模式,其运动矢量和DCT(discrete cosine transform)信息不需要编码。

1.2 多参考帧编码

在H.264中,支持多参考帧的帧间预测编码,该技术在场景交替切换的情形下,可以使编码效率比较有效地得以提高。

1.3 1/4像素精度的运动估计/补偿

为了进一步提高视频编码质量,H.264使用了1/4像素精度的运动估计/补偿算法。该算法先使用一个6阶滤波器(1,-5,20,20,-5,1)计算出1/2精度像素位置的值,再使用整像素和1/2精度像素的线性插值得到1/4精度像素位置的预测值。

1.4 整数离散余弦变换与量化

设 X 为残差矩阵, Y 为变换后的系数矩阵,H.264的4×4整数离散余弦变换定义为:

$$Y = D \cdot X \cdot D^T \otimes E \quad (1)$$

式中 $D = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix}$; \otimes 运算表示矩阵对应

位置上的元素相乘; $E = \begin{bmatrix} a^2 & \frac{ab}{2} & a^2 & \frac{ab}{2} \\ \frac{ab}{2} & \frac{b^2}{4} & \frac{ab}{2} & \frac{b^2}{4} \\ a^2 & \frac{ab}{2} & a^2 & \frac{ab}{2} \\ \frac{ab}{2} & \frac{b^2}{4} & \frac{ab}{2} & \frac{b^2}{4} \end{bmatrix}$,

$$a = \frac{1}{2}, \quad b = \sqrt{\frac{2}{5}}.$$

为了有利于实现,在H.264的量化表设计中,将缩放矩阵 E 融合到量化中,在量化阶段就可以近似地通过乘法和移位来实现,避免了除法。实际上,H.264的整数离散余弦变换与通常的离散余弦变换DCT近似,但避免了浮点运算,并且大大减小了计算量。

H.264采用分级量化,共有52级量化,量化参数每增加6,量化步长就增加一倍,量化参数越小,量化越精细。

2 快速模式选择算法

从式(1)可以看出,求残差的直流系数十分简单,即是求残差的平均值,这也与信号处理方面的理论吻合。同时,从信号处理与图像处理理论方面可以知道,残差的直流系数对应于信号(残差)中不随时间变换的稳定分量,而且很多情况下,直流系数是所有DCT系数中绝对值最大的^[10]。因此本文利用直流系数快速选择P帧宏块模式。率失真优化方法(RDO)是H.264 P帧编码中提高重建视频图像的质量的有效方法,其主要思想是把预测误差和编码率结合起来进行编码模式的选择^[11]。

$$J_{mode}(MB_i, P_i | \lambda_{mode}, q) = D_{rec}(MB_i, P_i | q) + \lambda_{mode} R_{rec}(MB_i, P_i | q) \quad (2)$$

式中 J_{mode} 为率失真代价; MB_i 为当前图像帧中的第 i 个宏块; P_i 为宏块的可选模式; Lagrangian参数 λ_{mode} 的值为 $0.85 \cdot 2^{(q-12)/3}$; D_{rec} 为对应预测模式下的重建图像的失真,即重建图像和原始图像的差值均方和; R_{rec} 为对应预测模式下的编码比特数; q 为量化系数。

率失真优化方法通常先用穷举法计算出所有可选模式的率失真代价 J_{mode} , 然后比较所有模式的率失真代价, 代价最小的模式即为最终的宏块模式。

从式(2)可以看出, 计算每种模式的率失真代价 J_{mode} 时, 运动估计、帧内预测、整数离散余弦变换、量化等编码过程以及相应的逆过程都需要一次进行, RDO模式判决方法的运算量庞大。因此, 减少候选模式的数量是降低RDO模式判决运算量的有效途径。

由前面的讨论可以看出, 在P帧编码时, 如果每个宏块都要对7种模式运算一遍得到对应的率失真代价 J_{mode} , 然后比较所有模式的率失真代价, 选取代价最小的模式为最佳模式, 运算量将会非常大。本文利用SKIP模式下残差的直流系数, 结合P16×16模式下的率失真代价, 提出一种P帧快速模式选择算法, 该算法的过程如图1所示, 其主要步骤如下。

1) 计算SKIP模式下的直流系数。

设当前宏块的一个4×4块为 $(c_{ij})_{4 \times 4}$, 简记为 C , 在SKIP模式下对应的预测值为 $(p_{ij})_{4 \times 4}$, 简记为 P , 残差为 $(x_{ij})_{4 \times 4}$, 简记为 X , 直流系数为 d , 量化参数为 q , 则:

$$X = C - P \quad (3)$$

直流系数 d 就是 $|\sum_{i,j} (c_{ij} - p_{ij})|$ 在量化参数 q 下量化后产生的计算结果。

2) 计算P16×16模式下的率失真代价。

首先计算P16×16模式的最佳匹配运动向量, 然后根据式(2)计算其对应的率失真代价 $J_{P16 \times 16}$ (整像素/子像素运动估计、整数离散余弦变换、量化等编码过程以及对应的逆过程都要完成, 同时还要计算出量化后的DCT系数对应的系数代价 $coeff_cost$ ^[12])。

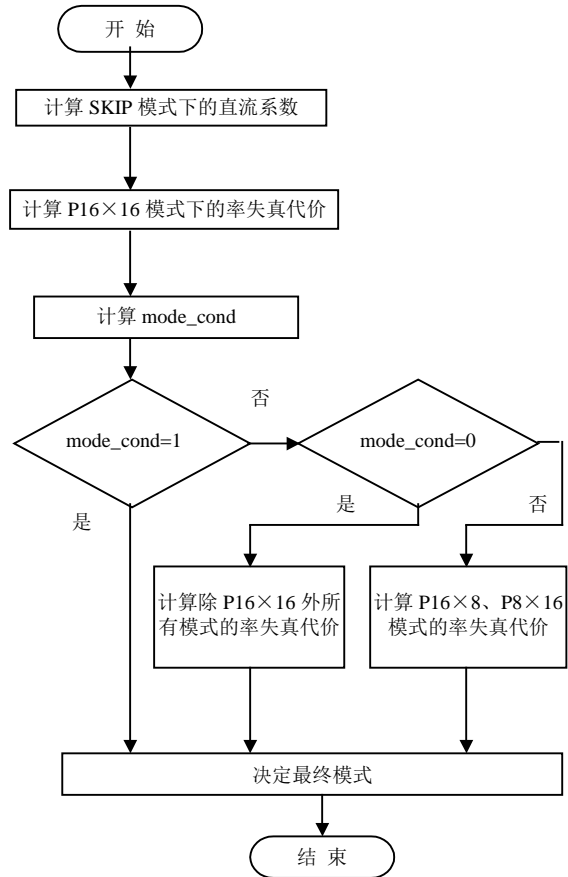


图1 本文的快速模式选择算法流程图

3) 计算mode_cond。

(1) 首先令 $mode_cond=0$ 。

(2) 如果步骤1)计算出的所有4×4块的直流系数 d 均为0, 则令 $mode_cond=2$ 。

(3) 计算条件1: 步骤2)计算时产生的所有DCT系数均为0, 并且当前宏块的最佳参考帧为第0参考帧, 最佳匹配运动向量等于运动向量的预测值(即运动矢量残差MVD为0); 计算条件2: 步骤2)计算时产生的所有DCT系数均为0或者 $coeff_cost < 2$, 并且 $mode_cond=2$ 。

如果条件1或者条件2为真, 则令 $mode_cond=1$; 否则, 如果步骤2)计算时产生的所有DCT系数均为0, 则令 $mode_cond=2$ 。

4) 如果 $mode_cond=1$, 则当前宏块的最佳模式为SKIP模式, 模式判定结束; 否则进入下一步。

5) 如果 $mode_cond=2$, 则根据式(2)分别计算

P16×8模式和P8×16模式下对应的率失真代价 $J_{P16 \times 8}$ 和 $J_{P8 \times 16}$, 然后比较 $J_{P16 \times 16}$ 、 $J_{P16 \times 8}$ 和 $J_{P8 \times 16}$ 的大小, 最小者对应的模式即为最佳模式, 模式判定结束; 否则进入下一步。

6) 根据式(2)分别计算除P16×16模式外余下所有模式的率失真代价, 然后比较所有模式对应的率失真代价大小, 率失真代价最小者对应的模式即为最佳模式, 模式判定结束。

3 实验结果及应用

3.1 实验结果

本文把提出的宏块模式快速选择算法融入H.264编码参考软件JM15.1^[12], 并与其进行性能对比实验。同时, 本文还把该算法与Jo算法^[9]进行性能对比实验。参考软件JM15.1的具体配置和实验环境如表1所示, Jo算法以及本文的算法的测试均以该配置和环境为参考。

表1 编码算法测试条件

项目	测试条件
	Windows XP专业版
计算机配置	1 GB内存
	Intel Pentium-M1.73 GHz处理器
框架	基本框架
率失真优化	高复杂度模式
GOP类型	IPPP
帧尺寸	352×288
搜索范围	±32
搜索模式	增强型预测区域搜索
总帧数	100
参考帧数	5

为了测试本文算法在不同运动场景下的有效性, 选用剧烈运动视频序列“foreman”、中等运动视频序列“paris”、较小运动视频序列“hall monitor”进行实验, 并用编码时间、重建图像亮度信号的平均PSNR以及编码码率作为H.264编码算法性能和效率的衡量指标。

在性能和效率对比时, 令 $T_{JM15.1}$ 、 $PSNR_{JM15.1}$ 和 $Bitrate_{JM15.1}$ 分别为编码参考软件JM15.1的运行时间、平均峰值信噪比和编码码率; 令 T_N 、 $PSNR_N$ 和 $Bitrate_N$ 分别为本文算法的运行时间、平均峰值信噪比和编码码率。则本文算法与编码参考软件JM15.1的时间差百分比 ΔT 、峰值信噪比差 $\Delta PSNR$ 和码率差百分比 $\Delta Bitrate$ 分别为:

$$\Delta T = \frac{T_N - T_{JM15.1}}{T_{JM15.1}} \times 100\% \quad (4)$$

$$\Delta PSNR = PSNR_N - PSNR_{JM15.1} \quad (5)$$

$$\Delta Bitrate = \frac{Bitrate_N - Bitrate_{JM15.1}}{Bitrate_{JM15.1}} \times 100\% \quad (6)$$

式中 正数表示增加; 负数表示降低或者减少。对于其他算法有类似定义。

视频序列“foreman”、“paris”、“hall monitor”的各种编码算法的性能和效率分别如表2~表4所示(q 为量化参数)。

表2 “foreman”序列各种编码算法性能对比

q	JM15.1	算法	ΔT /(%)	$\Delta PSNR$ /dB	$\Delta Bitrate$ /(%)
28	115.845 s	Jo	-30.40	-0.06	-0.88
	37.29 dB	本文算法	-54.31	0.10	1.89
	416.20 kb/s				
32	111.973 s	Jo	-38.99	-0.08	-0.58
	34.69 dB	本文算法	-64.02	0.05	-3.14
	234.07 kb/s				

表3 “paris”序列各种编码算法性能对比

q	JM15.1	算法	ΔT /(%)	$\Delta PSNR$ /dB	$\Delta Bitrate$ /(%)
28	106.825 s	Jo	-50.27	-0.06	-0.33
	36.06 dB	本文算法	-54.65	0.17	3.82
	650.80 kb/s				
32	103.792 s	Jo	-54.37	-0.09	-0.68
	32.81 dB	本文算法	-62.12	0.09	3.97
	358.65 kb/s				

表4 “hall monitor”序列各种编码算法性能对比

q	JM15.1	算法	ΔT /(%)	$\Delta PSNR$ /dB	$\Delta Bitrate$ /(%)
28	90.739 s	Jo	-58.89	-0.10	-3.75
	38.06 dB	本文算法	-67.51	0.11	2.65
	248.36 kb/s				
32	87.912 s	Jo	-68.26	-0.11	-4.52
	35.70 dB	本文算法	-74.64	0.10	-2.57
	116.66 kb/s				

从实验结果可以看出: 本文的宏块模式快速选择算法与编码参考软件JM15.1相比, 可以明显降低编码时间, 达到54%~74%; 显著提高峰值信噪比PSNR, 达到0.05 dB~0.17 dB; 码率可以降低3.14%, 而增加不超过3.97%; 同时, 本文算法与Jo算法^[9]相比, 在码率相差不明显的情况下, 编码时间降低和峰值信噪比提高十分明显。

3.2 应用

本文把提出的算法应用于一实时H.264视频压缩系统。该系统采用TI's TMS320DM642 DSP, 主频为600 MHz, 主要软件代码采用线性汇编语言实现。

实验发现, 使用传统的模式选择算法(JM参考软件), 即使是运动量很小的“hall monitor”序列(分辨率为352×288), 该系统的视频压缩速度也平均仅为

10 帧/s, 远远达不到实时压缩(25 帧/s)的要求。3.1 节的实验结果可以看出, 本文提出的算法具有明显的优势, 因此把其应用在该系统中。该算法在系统上的结果为: 对于分辨率为D1(720×576), 运动剧烈程度如“foreman”和“paris”等的一路普通视频信号, 系统的平均压缩速度为25 帧/s。

4 结 论

本文提出了一种新的针对P帧的快速模式选择算法。该算法利用SKIP模式下残差的直流系数和P16×16模式下的率失真代价, 对宏块的一些不必要的可能模式进行剔除, 使模式选择的过程得到简化, 显著提高了编码速度。实验结果表明, 该算法与其他最新算法相比, 在码率变化很小的情况下, 不仅能大大提高P帧编码的速度, 而且重建图像的质量增强明显; 对某些视频图像, 本文算法不仅可大大提高编码速度, 而且能显著增强重建图像的质量, 降低编码码率, 具有明显的优势。应用结果表明, 该算法在视频压缩系统中能够实现实时压缩。

参 考 文 献

- [1] ITU Telecom. Standardization Sector of ITU. Advanced video coding for generic audiovisual services. ITU-T Recommendation H.264[S]. 2003.
- [2] KAMACI N, ALTUNBASAK Y. Performance comparison of the emerging H.264 video coding standard with the existing standards[C]//Proc IEEE International Conference on Multimedia and Expo (ICME2003). Baltimore: IEEE, 2003: 345-348.
- [3] SULLIVAN G J, WIEGAND T. Rate-distortion optimization for video compression[J]. IEEE Signal Processing Magazine, 1998, 15(11): 74-90.
- [4] LEE J, JEON B. Fast mode decision for H.264 with variable motion block sizes[C]//ISCIS 2003. Antalya: ISCIS, 2003: 723-730.
- [5] WU D, WU S, LIM K P, et al. Block inter mode decision for fast encoding of H.264[C]//IEEE ICASSP. Quebec: ICASSP, 2004: 181-184.
- [6] AHMAD A, KHAN N, MASUDET S, et al. Selection of variable block sizes in H. 264[C]//IEEE ICASSP. Quebec: ICASSP, 2004: 319-321.
- [7] 宋 彬, 常义林, 李春林. H.264帧间预测模式的快速选择算法[J]. 电子学报, 2007, 35(4): 697-700.
SONG Bin, CHANG Yi-lin, LI Chun-lin. Novel fast selecting inter prediction mode algorithm based H.264[J]. Acta Electronica Sinica, 2007, 35(4): 697-700.
- [8] ZENG Huan-qiang, CAI Can-hui, MA Kai-kuang . Fast mode decision for H.264/AVC based on macroblock motion activity[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2009, 19(4): 491-499.
- [9] JO Young-sub, JIN Soon-jong, YOU Jong-min, et al. Adaptive fast inter mode decision algorithm using dynamic threshold technique for H.264/AVC P slices[C]//2009 IEEE/ACS International Conference on Computer Systems and Applications. Rabat, Morocco: AICCSA, 2009: 913-918.
- [10] YU A, LEE R, FLYNN M. Early detection of all-zero coefficients in h.263[C]//Proceedings of the Picture Coding Symposium (PCS 1997). Berlin, Germany: VDE-Verlag, 1997: 159-164.
- [11] WIEGAND T, SCHWARZ H, ANTHONY J, 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.
- [12] Joint Video Team. Reference software JM15.1[EB/OL]. [2009-07-07]. [http://iphome. hhi. de/suehring/tml /download/old_jm/](http://iphome.hhi.de/suehring/tml/download/old_jm/).

编 辑 漆 蓉