

一种 H. 264 残差系数修剪算法

侯兴松, 薛端

(西安交通大学电子与信息工程学院, 710049, 西安)

摘要: 针对 H. 264 残差系数编码中位置信息占用码率较多这一问题,提出了一种 H. 264 残差系数修剪算法. 该算法首先对视频序列进行预测、变换、量化以及扫描,然后对扫描后获得的残差系数进行线性分组,并采用率失真优化准则对残差系数分组进行修减,再在给定量化步长范围内计算每一个宏块在不同量化步长下的修剪结果,最后将率失真代价最小时对应的修剪剩余残差系数编码输出. 实验结果表明,与 H. 264 残差系数编码算法相比,该算法在相同码率下的峰值信噪比最多可以提高 0.9 dB,在相同信噪比下最多可以节省 17% 的码率. 该算法没有改变 H. 264 的码流结构,与标准兼容.

关键词: 系数修剪;量化;率失真优化;H. 264 标准

中图分类号: TN919.8 **文献标志码:** A **文章编号:** 0253-987X(2010)10-0052-05

A Residual Coefficients Pruning Algorithm in H. 264

HOU Xingsong, XUE Duan

(School of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: A pruning algorithm is proposed to reduce the encoding bits for the map information of residual coefficients in H. 264. Video sequences are predicted, transformed, quantized, and scanned. Then residual coefficients obtained from scanning are divided into linear groups, and then the groups are pruned out using the rate-distortion optimization (RDO) sense. Furthermore, the pruning results of each macroblock with different quantization steplengths are calculated in a given range. Finally the remaining coefficients with the minimal rate-distortion cost are encoded and transformed. Experimental results show that the proposed algorithm has better RD performance than the residual coding method in H. 264 at the relatively high bitrates. The proposed algorithm gains an 0.9 dB increase on the peak signal to noise ratio(PSNR) at the same bitrate, or equivalently, a 17% rate saving at the same PSNR. Moreover, the proposed algorithm does not change the bit-stream structure of H. 264, and is compatible with the H. 264 standard.

Keywords: coefficients pruning; quantization; rate-distortion optimization; H. 264 standard

在图像/视频编码中,经过变换以及量化之后的系数在进行熵编码时通常包含两方面信息:幅值信息——待编码系数的模值与符号;位置信息——待编码系数所处的位置. 通常这两方面信息直接进行编码,而不考虑两者之间的相互影响. 该类编码方式所存在的问题是在给定码率下位置信息占用了过多的编码比特. 为了解决该问题,在图像编码领域,

Xiong 和 Velisavljevic 等人^[1-3]提出了一类基于小波变换的空频量化算法,Shoa 等人^[4]则在图像匹配追踪编码中提出了一种兼顾系数幅值编码与位置编码的优化折中编码算法. 这些图像编码算法均实现了待编码系数幅值信息与位置信息之间的优化编码,减少了位置信息的编码代价. 然而,这些算法都有其应用条件限制,它们都不能应用于基于离散余

弦变换(DCT)的视频压缩编码标准中,例如 H. 264 标准. H. 264 作为新一代视频编码标准也同样存在残差系数编码中位置信息占用较多码率这一问题. 针对该问题,本文提出了一种 H. 264 残差系数修剪算法,以减少位置信息的编码比特,提高算法的率失真性能.

1 H. 264 中的残差系数编码

在 H. 264 中,量化之后的残差系数在进行熵编码之前,需要以一个 DCT 变换块为单位进行 zig-zag 扫描,然后熵编码器以一定的顺序对扫描后的残差系数进行编码. 在扫描后的残差系数中,当某个非 0 系数的幅值很小而它之前 0 的个数较多时,该系数在率失真意义下很可能不值得编码. 这是因为

表 1 不值得编码的非 0 系数在总的非 0 系数中的百分比

测试序列	$\delta/\%$						
	$Q_p=12$	$Q_p=15$	$Q_p=18$	$Q_p=21$	$Q_p=24$	$Q_p=27$	$Q_p=30$
Bridge-close	12.08	20.75	18.82	21.79	28.6	34.53	42.15
Bus	9.61	14.17	18.9	23.3	29.19	35.55	40.66
Silent	20.47	21.66	23.32	23.95	23.39	21.59	19.08
Paris	17.29	16.52	16.16	17.02	20.92	25.85	31.84
平均值	14.86	18.28	19.30	21.52	25.53	29.38	33.43

2 残差系数修剪算法设计

为了将不值得编码的残差系数修剪去除,本文首先将扫描后的一维数据定义成一种线性分组结构. 在扫描后的一维数据中,从一个非 0 系数开始直到该一维数据的结尾,其中所有的系数组成一个系数分组. 图 1 给出了一个残差系数线性分组示意图.

本文提出的残差系数修剪算法以该系数分组结构为基本单元进行系数修剪. 如果某个系数分组被选中,那么该系数分组中的所有系数都会被置 0. 针对变换块中哪个系数分组需要被修剪这一问题,本文采用了基于率失真优化准则的竞争机制对需要修

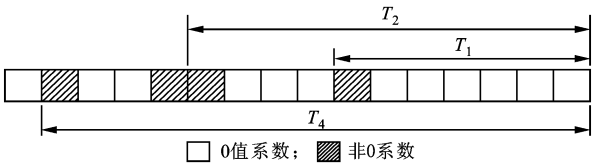
该系数的幅值小,意味着它对重建图像质量的贡献小,紧跟在该系数之前 0 的个数多,意味着它的位置信息编码比特多,编码代价大,所以该系数很可能不值得编码. 表 1 给出了不同序列在利用 H. 264 进行编码时该类非 0 系数在总的非 0 系数中的比例,该比例用 δ 表示. 在统计中,如果一个非 0 系数的幅值小于等于 2,而紧跟在其前面的 0 的个数大于等于 3,那么我们便认为该系数可能不值得编码. 统计实验的参数和所用序列如第 4 节所述. 从表 1 可以看出,在不同量化参数 Q_p 下, δ 的平均值基本都在 15% 以上,随着量化参数的增加, δ 的平均值还会进一步增加. 由此可见,位置信息编码代价大而幅值较小的残差系数的比例较大,因此对 H. 264 中该类残差系数进行修剪是有意义的.

剪的系数分组进行选择. 如图 2 所示,在一个扫描后的系数分组中,计算每一个分组被修剪之后的剩余残差系数的率失真代价. 率失真代价计算公式如下

$$J = D + \lambda R$$

(1)

式中: J 为率失真代价; λ 为拉格朗日乘子; D 为该变换块的失真度,它等于该变换块原始像素与对应重建像素差的平方和; R 为相应系数分组被修剪后剩余系数熵编码所消耗的实际比特. 比较各种修剪情形以及初始残差块的率失真代价大小,并以率失真代价最小的修剪情形作为最终的修剪结果,即与其相对应的系数分组被置 0.



T 表示一个系数分组;下标 i 表示分组中非 0 系数的个数

图 1 残差系数线性分组结构示意图

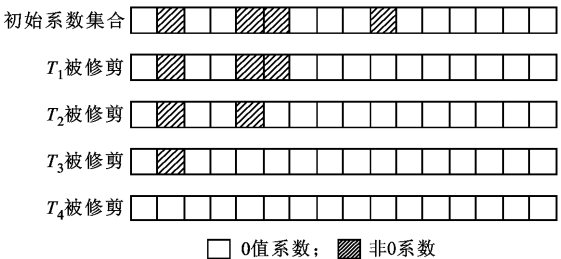


图 2 基于率失真优化的修剪竞争机制

量化后残差系数的幅值和位置信息与标量量化有关. 标量量化步长越大, 变换块中非 0 系数幅值越小, 0 的个数越多. 考虑到标量量化对修剪算法的影响, 本文将标量量化与残差系数修剪相结合.

在 H. 264 中, 宏块是基本的编码单元. 宏块中所有残差系数采用相同的量化参数进行量化. 本文算法采用遍历的方式对每一个宏块的标量量化参数进行调整, 以得到符合残差系数修剪特性的量化残差系数. 若宏块的初始量化参数用 Q_{pinit} 表示, Δ 为对初始量化参数的调整幅度, 于是每一个宏块的量化参数选择范围可以表示成集合 $\{Q_p | Q_p \in [Q_{\text{pinit}} - \Delta, Q_{\text{pinit}} + \Delta]\}$. 遍历的计算宏块在不同量化参数下进行系数修剪的率失真代价, 选择率失真代价最小时对应的标量量化参数作为宏块的最终量化参数, 用 Q_{pmb} 表示. 同时, 宏块初始量化参数与最终量化参数的差值 $Q_{\text{pdelta}} = Q_{\text{pinit}} - Q_{\text{pmb}}$ 需要进行编码并传送到解码器.

综上所述, 本文提出的基于宏块的残差系数修剪算法可以归纳为

$$J = \min_{q \in Q} \min_{S \leq T}^{(a)} \{D(q, S) + \lambda R(q, S)\} \quad (2)$$

式中: a 为内层优化参数; b 为外层优化参数; q 为每一个宏块的最终量化步长; Q 为量化参数集合; S 为修剪后的残差系数集合, 是内层优化参数 a 在固定标量量化步长 q 下对残差系数分组 T 进行修剪得到的. 需要注意的是, 在每一个宏块的率失真优化过程中拉格朗日乘子 λ 设置为 $\lambda = 0.85 \times 2^{(Q_{\text{pinit}} - 12)/3 [5]}$.

3 快速算法设计

本文算法包含系数修剪模块和标量量化步长选择 2 个编码模块. 下面针对这 2 个编码模块的复杂度进行分析, 并给出相应的快速编码算法.

在系数修剪模块中, 为了选择变换块中需要被修剪的系数分组, 需要计算每一种修剪情形的实际率失真代价. 根据修剪算法的设计方法, 变换块中修剪情形的数量等于该变换块中非 0 系数的数量, 所以当变换块中的非 0 系数较多时, 本文算法的复杂度将大大增加.

为减少系数修剪模块的计算复杂度, 本文采用一种简单的估计算法来判断每种修剪情形的有效性. 在修剪算法的竞争机制中, 当一种修剪情形的率失真代价小于初始变换块的率失真代价时, 这种修剪情形才被认为是有效的, 可以表示为 $J_1 - J_0 <$

0, 其中 J_1 与 J_0 分别表示当前修剪情形与初始变换块的率失真代价. 将式(1)代入 $J_1 - J_0 < 0$ 中并做相应的变换, 可以得到

$$(R_0 - R_1)/(D_1 - D_0) > 1/\lambda \quad (3)$$

式中: R_0 与 R_1 、 D_0 与 D_1 分别是初始变换块和该修剪情形下的实际编码比特和失真; $R_0 - R_1$ 表示修剪相应系数后节省的编码比特, 它等于被修剪系数的幅值与位置信息所需的编码比特; $D_1 - D_0$ 则表示修剪相应系数后增加的失真度. 为了避免实际的比特消耗与失真度计算, 本文采用变换块中的残差系数的相关信息来估计式(3)中的各个参数值. 在该估计算法中, 被修剪非 0 系数的位置的编码比特用紧跟在这些系数之前 0 的个数来估计, 被修剪非 0 系数幅值的编码比特用这些非 0 系数幅值的二进制位数来估计, 而失真则利用它们的幅值的平方和来估计. 这样本文的估计算法便可以表示为

$$\left(N_{\text{zeros}} + \sum_{i=1}^N (\text{lb} |c_i| + 1)\right) / \sum_{i=1}^N c_i^2 > T_h \quad (4)$$

式中: N_{zeros} 表示所有被修剪的非 0 系数之前的 0 的个数之和; N 表示被修剪的非 0 系数的个数; c_i 表示被修剪非 0 系数的幅值; T_h 为本文所设置的经验阈值. 当某种修剪情形不满足不等式(4)时, 那么该修剪情形在修剪算法的竞争机制中便被排除在外. 这样也就避免了不必要的率失真计算.

在标量量化步长选择模块中, 为了找出与修剪算法相匹配的标量量化步长, 每一个宏块需要在不同的量化步长下编码 $2\Delta + 1$ 次, 因此该编码模块的计算复杂度取决于参数 Δ 的大小. 为了减少该模块的计算复杂度, 可以为参数 Δ 选择一个较小的值, 例如将 Δ 的值设置为 1, 当然, 这将给算法的性能带来一定的损失.

4 实验结果及分析

采用 JM12.4 参考软件对本文算法进行实现及验证. 为了检验算法的有效性, 我们对典型序列进行了测试, 并给出其中 4 个序列的测试结果, 序列信息如表 2 所示. 实验的参数设置如下: H. 264 编码层次选择基本层 Baseline; 最大参考帧长度为 5; 运动估计采用全搜索策略, 最大搜索范围为 32; 帧率为 30 帧/s; 测试序列帧中只有第一帧为 I 帧, 其余帧均为 P 帧, 即帧结构为 IPPP; 率失真优化模式采用高复杂度模式; 编码器不采用码率控制模式. 测试序列在不同码率下的率失真性能通过为测试序列设置不同的初始量化步长得到.

表 2 实验所用视频序列

序列	序列帧数	格式
Bridge-close	300	QCIF
Silent	300	CIF
Bus	75	QCIF
Paris	300	CIF

4.1 本文算法的性能

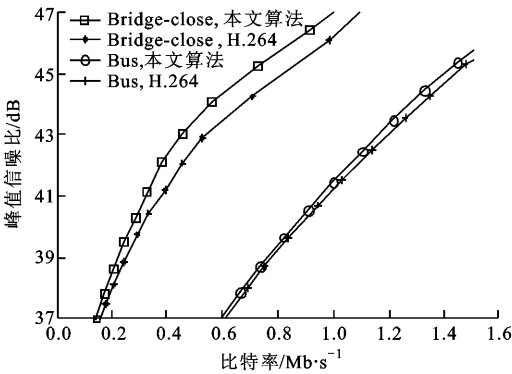
图 3 给出了不同测试序列在参数 $\Delta=3$ 时本文算法与 H. 264 的率失真性能比较. 由图 3 可见, 对于给定的测试序列, 本文算法相比 H. 264 在中高码率下可以获得更好的率失真性能. 对于 Bridge-close 序列, 在相同码率下, 本文算法与 H. 264 相比最多可以获得 0.9 dB 的峰值信噪比 (PSNR) 提高, 或者在相同的 PSNR 下, 本文算法与 H. 264 相比最多可以获得 17% 的码率节省. 对于其他运动相对缓慢的序列, 本文算法在特定码率下也能够获得 0.5~0.7 dB 的 PSNR 提高, 而对于运动相对剧烈的 Bus 序列, 本文算法可以获得 0.1~0.3 dB 的 PSNR 提高,

性能提高较小. 这是因为对于运动剧烈的视频序列, 运动估计往往不够精确, 从而造成预测误差较大. 这使得残差系数经量化后, 非 0 系数个数较多, 幅值较大. 本文提出的系数修剪算法对该类特性的残差系数往往不做修剪, 因此算法性能提高有限.

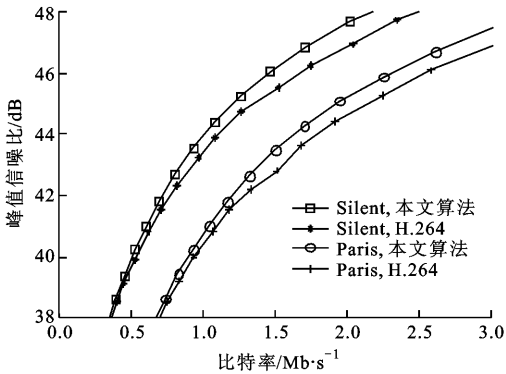
为了获得最佳的率失真性能, 本文算法在量化步长选择模块对每一个宏块的量化参数进行调整, 以得到符合修剪算法特性的量化残差系数. 该宏块级的量化参数调整会在一定程度上影响 H. 264 基本单元层码率控制的精度. 然而, 由于本文算法不会对帧层和图像组 (GOP) 层的码率分配策略产生影响, 因此本文算法并不会对码率控制的最终精度产生明显的影响.

4.2 幅值与位置信息的优化折中

通过对待编码系数进行修剪, 本文算法可以减少码流中残差系数位置信息的编码比特, 从而能够分配更多的比特用于幅值信息的编码. 这样, 相比于 H. 264, 本文算法的重建图像在一个给定码率下可以获得更小的失真. 图 4 给出了本文算法与 H. 264 在上述实验条件下残差系数幅值与位置信息在总码

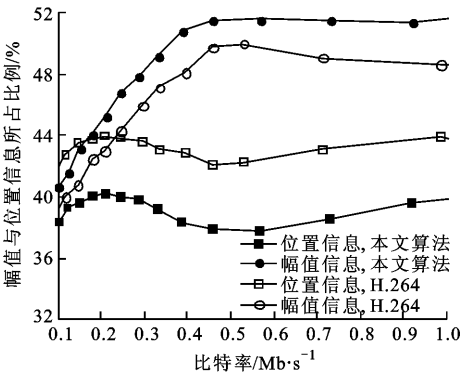


(a) Bridge-close 与 Bus

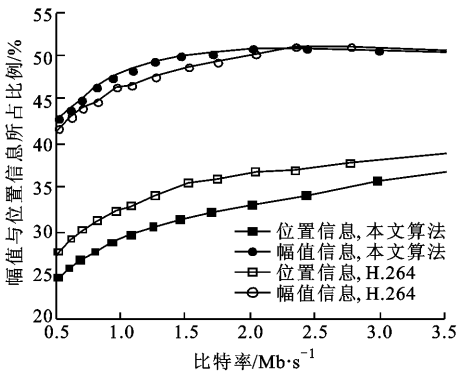


(b) Silent 与 Paris

图 3 不同测试序列的率失真性能比较



(a) Bridge-close



(b) Silent

图 4 本文算法与 H. 264 中残差系数幅值与位置信息在总码率中的比例

率中的比例. H. 264 采用基于上下文自适应可变长编码(CAVLC)方式为残差系数编码时共有 5 个语法元素^[6-7]. 本文采用 TotalZeros 与 RunBefore 这 2 个语法元素的编码比特作为位置信息的编码比特, 而其他 3 个语法元素的编码比特作为幅值信息的编码比特. 由图 4 可以看出, 对于 Bridge-close 与 Silent 序列, 本文算法可使码流中的位置信息比例下降 4% 左右, 并使幅值信息比例提高 2% 左右.

4.3 快速算法性能

第 2 节给出了相应的快速算法, 这里来验证快速算法对本文算法性能的影响. 在快速算法中, 式 (4) 中的参数 T_h 设置为 3, 联合优化参数 Δ 设置为 1, 其他参数与上述实验参数相同. 图 5 给出了本文算法的快速算法模式与非快速算法模式的性能比较. 由图可见, 本文算法的快速算法模式与非快速算法模式相比率失真性能仅下降 0.1 dB 左右. 因此,

本文的快速算法在减少算法复杂度的同时也能很好地保持算法的性能.

5 结 论

本文提出的残差系数修剪算法将不重要的非 0 系数去除, 这样便不必为这些系数输出额外的位置信息, 这种修剪系数的思想与空频量化算法相似. 实验结果表明, 本文算法针对运动相对缓慢的视频测试序列在中高码率下获得了很好的编码效果, 此外, 本文算法增加了修剪算法和量化步长选择 2 个模块, 这 2 个编码模块非常适合并行计算, 易于多核实现.

参考文献:

- [1] XIONG Zixiang, RAMCHANDRAN K, ORCHARD M. Space-frequency quantization for wavelet image coding [J]. IEEE Transactions on Image Processing, 1997, 6(5): 677-693.
- [2] GLEICH D, PLANINSIC P, GERGIC B, et al. Progressive space frequency quantization for SAR data compression [J]. IEEE Transactions on Geoscience and Remote Sensing, 2002, 40(1): 3-10.
- [3] VELISAVLJEVIC V, BEFERULL-LOZANO B, VETTERLI M. Space-frequency quantization for image compression with directionlets [J]. IEEE Transactions on Image Processing, 2007, 16(7): 1761-1773.
- [4] SHOA A, SHIRANI S. Optimized atom position and coefficient coding for matching pursuit-based image compression [J]. IEEE Transactions on Image Processing, 2009, 18(12): 677-693.
- [5] WIEGAND T, SCHWARZ H, JOCH A, et al. Rate-constrained coder control and comparison of video coding standards [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7): 688-703.
- [6] 毕厚杰. 新一代视频压缩编码标准——H. 264/AVC [M]. 北京: 人民邮电出版社, 2005.
- [7] WIEGAND T, SULLIVAN G, BJONTEGAARD G, et al. Overview of the H. 264/AVC video coding standard [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7): 560-576.

(编辑 刘杨)

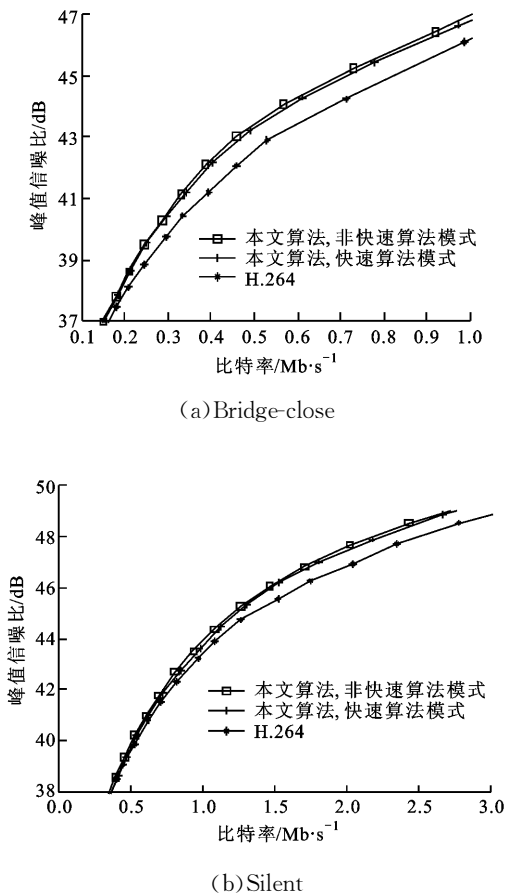


图5 本文算法的快速算法模式与非快速算法模式的性能比较