

文章编号:1001-5132 (2008) 02-0159-05

Wyner-Ziv 视频编码中 Key 帧快速预测算法

范旭明, 李福翠*, 彭宗举, 陈 恩

(宁波大学 信息科学与工程学院, 浙江 宁波 315211)

摘要: 以提高 Wyner-Ziv 编码性能为目的, 使用 H.264 帧内编码技术对 Wyner-Ziv 编码中的 K 帧进行编码, 利用其优异的编码性能来提高 K 帧编码质量. 同时从帧间相关性利用、亮度块模式之间的选择以及 4×4 亮度块预测模式选择的角度对 H.264 的帧内预测进行了改进, 降低其编码复杂度. 实验表明: 改进后的 K 帧编码算法耗时仅为 H.264 帧内编码算法的 5% 左右.

关键词: Key 帧; Wyner-Ziv 编码; 低复杂度编码

中图分类号: TN919.81

文献标识码: A

低复杂度的实时视频编码能力是目前低功耗网络终端设备的新需求之一, 此类应用要求编码端必需低功耗, 而对解码端功耗限制相对较小, 恰好与主流编解码器复杂度分布特性相反. 当前主流编码技术(如 H.264 编码等)的编码器复杂度一般为解码器的 5~10 倍左右^[1], 因此根本不适合此类应用环境. 而 Wyner-Ziv 编码(WZC)是当前国内外视频编码界和无线通信界的研究热点^[2], 它突破了传统视频编码框架的束缚, 采用帧内编码和帧间联合预测解码技术, 在解码端发掘视频信号的相关性进行帧间预测解码. 因此 WZC 复杂度很低, 可广泛应用于移动可视电话及无线视频监控等低功耗视频编码场合.

由于所应用的范围与场合的不同, 导致 WZC 的编码框架与传统视频编码框架有较大差异. 如图 1 所示, WZC 使用了分级标量量化^[3]、Turbo 编码、边信息重建^[4]等一系列视频编码技术. 在 WZC 中, 待编码的帧可分为 Wyner-Ziv 帧(W 帧)和 Key

帧(K 帧), 这 2 种帧分别在不同的编码器内进行编码, 并在解码端通过生成边信息联合解码^[5]. 由于 W 帧编码是整个 WZC 方案的核心, 目前 WZC 的研究也主要集中在 W 帧上, 而对 K 帧的研究相对较少, 但是 K 帧的编码性能却制约着 W 帧的解码边信息生成质量, 而且 K 帧的编码性能与处理速度也是影响低复杂度的重要因素. 所以研究 K 帧编码对 WZC 有着非常重要的意义. 因此, 本文设计了 K 帧快速预测编码算法, 在提高 K 帧编码性能的同时, 降低了编码复杂度. 实验表明: 本文算法复杂度远远低于 H.264 帧内编码算法, 但编码效率却与其非常接近.

1 Key 帧快速预测编码

当前 K 帧编码通常选用复杂度较低的传统帧内编码方式(如 H.263 编码等), 但是这些技术编码性能却远不及 H.264 标准^[1]. 为此本文选用 H.264

收稿日期: 2007-09-25.

宁波大学学报(理工版)网址: <http://3xb.nbu.edu.cn>

基金项目: 教育部科学技术研究重点项目(2060059); 浙江省自然科学基金(Y105577); 宁波市自然科学基金(2007A610037).

第一作者: 范旭明(1984-), 男, 浙江兰溪人, 在读硕士研究生, 主要研究方向: 多媒体通信. E-mail: wolffan3150@126.com

*通讯作者: 李福翠(1979-), 女, 苗族, 贵州遵义人, 硕士, 主要研究方向: 多媒体及流媒体技术. E-mail: lifucui@nbu.edu.cn

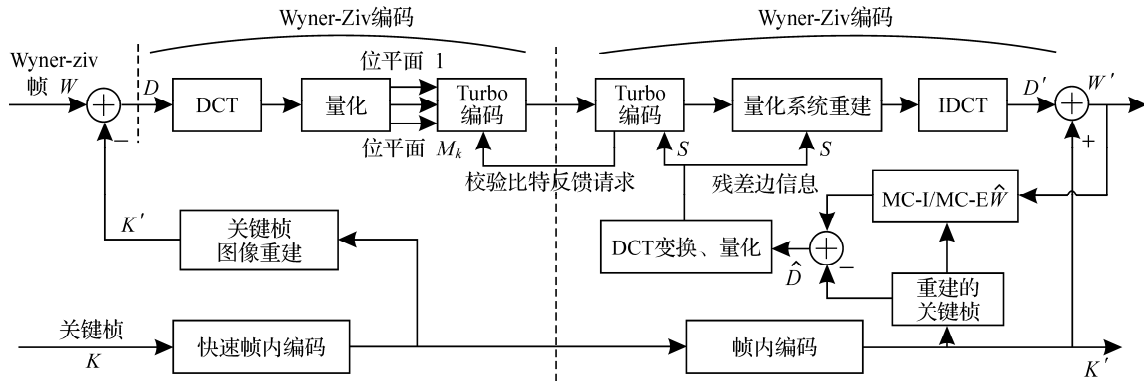


图1 Wyner-Ziv 编码框图

帧内编码作为 K 帧的编码方式. 在 H.264 的帧内预测中, 块模式有 4×4 块(I4MB)和 16×16 块(I16MB)2 种类型. 其中 I4MB 的块尺寸小, 有 9 种预测模式, 包括 DC(直流系数)模式和 8 种方向预测模式, 适合于纹理比较丰富的区域; 而 I16MB 的块尺寸大, 有垂直、水平、DC 和平面 4 种预测模式, 适合于纹理简单且平坦的区域. H.264 在块模式、预测模式选择上都采用完全遍历的算法. 在 JM86 校验模型中先将宏块按 I4MB 方式进行预测并累计编码代价, 然后再按 I16MB 进行预测, 最后比较 I4MB 和 I16MB 的编码代价以选择最优的预测模式^[6].

这样的遍历算法虽然可以得到最优的编码模式, 但计算量极大, 显然不适合直接应用于 K 帧编码. 如果能对 I4MB 和 I16MB 亮度块进行预判, 则能降低帧内预测的计算复杂度. 考虑到视频序列间的高冗余度必然导致视频序列在编码方式上的冗余, 若能有效利用这些冗余信息, 就可以进一步加快亮度块模式和预测模式选择. 因此本文在参考现有的 H.264 快速帧内预测编码算法的基础上^[7-9], 充分利用编码冗余, 提出了一种适合 WZC 的快速 K 帧预测算法. 该算法首先对相邻时刻的两帧做一帧差分析, 找出前后时刻图像得变化区域和非变化区域. 处在变化区域的宏块主要从亮度宏块模式的快速选择、I4MB 中候选预测模式的快速确定两个方面对 H.264 的帧内预测进行改进, 而非变化区域, 并尽可能依据图像帧间的相关性及减少预测次数, 以实现低复杂度 K 帧编码.

1.1 帧差分析

对视频序列中相邻帧作帧差分析, 记帧差宏块内不为零的灰度等级个数为 g , $g < T$, T 为阈值记为未变化区域, 否则为变化区域. T 的取值越小, 判定的条件也越严格, 误判的概率也越小. 如图 2 所示在 $T = 10$ 时, 未发生变化区域宏块编码块模式统计实验表明: 宏块模式在前后时刻相似性极高.

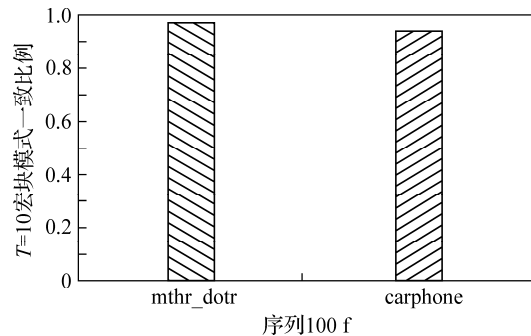


图2 宏块模式选择情况下的前后一致性分析

所以本文算法在 K 帧的 1 个 Group of picture (GOP)内, 只对起始帧的所有宏块和其余帧变化区域宏块作亮度 I4MB 和 I16MB 之间的作快速选择, 剩余宏块编码时沿用前一时间块模式. 为更好利用帧间相关性, 本文算法对未变化区域的帧差求取标准差 σ , 若 $\sigma < Th$ ($Th = 5$ 为阈值), 则当前帧预测模式沿用前帧模式, 否则转入快速预测算法.

1.2 亮度 I4MB 和 I16MB 之间的快速选择

为快速确定宏块编码块模式, 文献[7]是基于二维直方图的快速预判法, 文献[8]给出基于 SATD (Sum of Absolute Transformed Difference)的预判方法, 文献[9]对一系列常用测试序列进行了宏块亮

度直方图的统计实验. 最后发现, 16×16 亮度块对应直方图中的非零系数个数较少, 而 4×4 亮度块对应直方图中的非零系数个数较多. 为了进一步降低复杂度, 本文综合考虑 2 种算法的优劣, 提出了基于宏块非零系数一维直方图结合 SAD 快速预判的算法. 图 3 给出了在 xmas 序列中, 2 种块方式下的非零系数个数的分布特点. 为了将直方图中的非零系数个数作为判别的特征, 在图 2 中设定了 2 个阈值 T_1 和 T_2 . 另外, 基于亮度块 I16MB 预测编码的计算复杂度要远远小于 I4MB. 因此, 在设法保证率失真性能基本不变的情况下, 适当地增加 16×16 亮度块的选用概率可以进一步降低编码复杂度. 因此本文采用 SAD 准则判断宏块参考像素的平滑性, 并据此来确定宏块模式, 其公式为:

$$SAD = \sum_{i=1}^{33} |C_i - C_{i-1}|,$$

其中, C_i 代表宏块的左边、上边和左上边共 33 个参考像素. SAD 用于衡量参考像素之间的平滑程度和相似度, SAD 越小, 表示这 33 个参考像素之间过渡越平滑. 如果宏块参考像素的亮度 SAD 小于阈值 T_0 , 则直接采用 I16MB 预测编码, 否则采用 I4MB 预测编码.

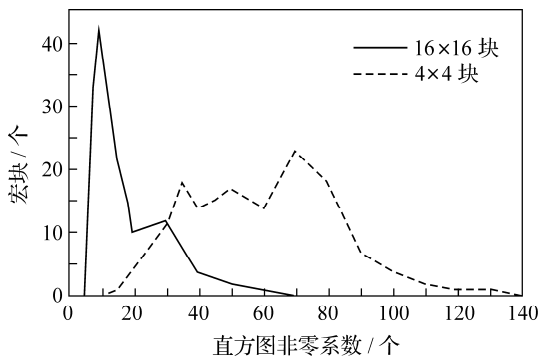


图 3 2 种块模式下的非零系数个数分布特点

判别的规则如下: (1)非零系数个数小于 T_1 的宏块, 采用 16×16 亮度块模式; (2)非零系数个数大于 T_2 的宏块, 采用 4×4 亮度块模式; (3)而非零系数个数趋于 T_1 和 T_2 之间的宏块, 依据 SAD 判别.

由于大部分宏块只需计算一种块方式的率失真代价, 从而降低了帧内预测的计算复杂度.

1.3 I4MB 预测模式快速选择

亮度块 I4MB 的具有 8 种方向预测模式, 鉴于自然场景图像中相邻块的纹理和方向信息的相似性, 相邻亮度 I4MB 块的最佳预测模式在预测方向上也非常的相似. 假设根据左边和上边参考块推导而得当前块最有可能预测模式为 4, 则当前块最佳预测模式极有可能是最有可能模式和两个相邻方向预测模式中的一个, 即预测模式 4、5、6 种的 1 个. 文献[8]中提到了这种相邻块最佳预测方向上的相似性, 且相似度很高. 因此, 如果在方向相邻 3 模式预测结束后, 就能提前判断是否已获得最优预测模式, 可以大大降低编码复杂度, 而不致使率失真性能受到较大影响.

表 1 为当前块最佳预测模式下的预测残差与相邻参考块的预测有着较强的相关性. 在表 1 中, C 表示当前块最佳预测模式下的预测残差, L 表示左边块其最佳预测模式下的预测残差, U 表示上边块其最佳预测模式下的预测残差, 其中 H 为:

$$H = (L + U) / 2.$$

据此, 本文提出了相邻预测方向优先搜索和自适应阈值终止的 I4MB 预测模式快速选择算法:

- (1) 根据预测的最有可能模式, 优先搜索最有可能模式及其方向相邻 2 种模式;
- (2) 在方向相邻 3 模式预测结束后, 比较其最小预测残差的大小;
- (3) 如果其最小预测残差 $< \alpha H$, 则认为已获得最优预测模式, 结束搜索; α 为调节系数; 否则, 搜索 9 种候选预测模式, 并从中选出最优模式.

表 1 相邻块预测残差间的关系

	$C < L$	$C < U$	$C < H$	$C < 9/8 \times H$
概率/%	49.89	52.82	55.64	65.08
	$C < 10/8 \times H$ $C < 11/8 \times H$ $C < 12/8 \times H$			
概率/%	72.61	78.40	82.75	

1.4 本文快速帧内预测模式选择的算法步骤

本文对帧内预测算法的改进充分利用了序列间前后时刻编码冗余, 对 4×4 亮度块和 16×16 亮

度块之间的快速选择算法,利用宏块亮度直方图的非零个数系数对2种块方式进行预判;而基于参考像素亮度值相似性的快速帧内预测算法是针对4×4亮度块的模式选择所提出的快速算法.由此可见,以上3种算法是从不同的角度降低了帧内预测的计算复杂度,存在互补性.所以,本文将以上3种算法结合,提出了一种高效的快速帧内预测模式选择算法(以下简称本文方案).本文方案的算法步骤具体如下:

- (1) 首先对相邻时刻图像进行帧差分析,区分出图像变化区域和未变化区域.
- (2) 未发生变化区域宏块编码块模式直接使用前一时刻对应块模式,若当前编码块与前一时刻对应块的标准差 $\sigma < Th$,则当前块预测模式直接使用前一时刻对应预测模式.否则转(4).
- (3) 对变化区域宏块,根据亮度直方图的非零系数个数,确定候选的亮度块方式.若个数小于阈值 T_1 ,则仅采用16×16亮度块;若个数小于阈值 T_2 ,则仅采用4×4亮度块方式;若个数趋于 T_1 和 T_2 之间,则选用SAD判别.
- (4) 对4×4亮度块进行模式选择时,优先搜索最有可能模式及其方向相邻两模式,如果预测残差小于事先设定阈值 αH ,那么该4×4块就采用most_probable_mode作为最佳预测模式,否则,还是采用RDO计算选择最佳的预测模式.

2 实验结果与分析

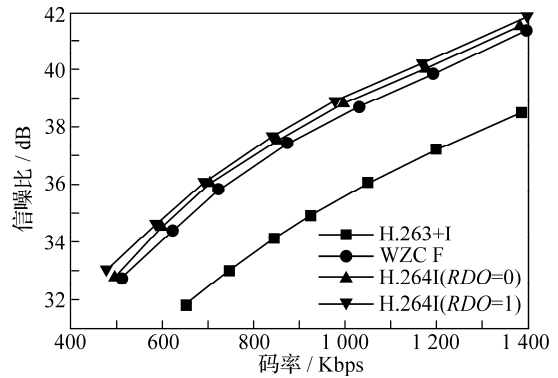
本文采用‘Xmas’序列作为实验序列.实验中随机抽取‘Xmas’第51视点,12帧做了测试,并从率失真性能和编码时间等方面与H.264校验模型JM86的帧内、帧间编码进行了比较.实验所用PC机操作系统为Windows 2000,CPU为Pentium IV 3G,512M内存,安全模式状态.编码帧率为10 fps, GOP=4(即Wyner-Zi视点编码帧类型为K-W-W-W-K,H.264帧间编码帧类型为I-P-P-P-I).Wyner-

Ziv视点编码采用基于4×4块的生成DCT变换,采用可扩展量化器.Turbo码的矩阵为:

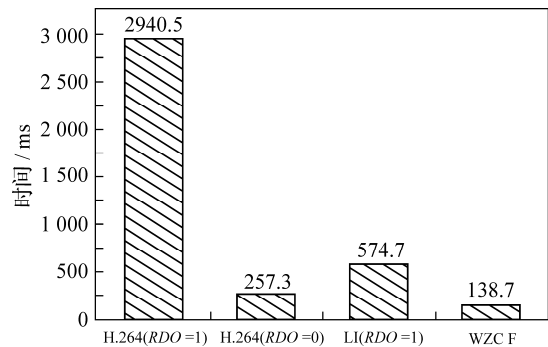
$$\begin{bmatrix} 1 & \frac{1+D+D^3+D^4}{1+D^3+D^4} \end{bmatrix}$$

并仅向解码端传送经1/4删余后的校验码.

Wyner-Ziv K帧编解码实验中 $H=(L+U)/2$, $\alpha=1$.实验结果如图4所示,图4(a)中曲线WZCF为本文K帧编码算法实验结果,曲线H.264 RDO=1和H.264 RDO=0分别表示H.264在有码率优化和无码率优化条件下的K帧编码性能曲线,H.263+I表示K帧选用H.263编码的编码性能曲线.各帧内编码算法的平均耗时如图4(b)所示,H.264 RDO=1,H.264 RDO=0分别表示H.264在有码率优化和无码率优化条件下的K帧平均编码时间,LI RDO=1表示文献[8]提出的H.264帧内编码快速算法所耗编码时间,WZCF表示本文算法K帧的平均编码时间.本文算法在无率失真优化的情况下,K帧平均编码时间为137.8 ms,而H.264 RDO=1时K帧



(a) 关键帧视点编解码 PSNR 曲线



(b) 关键帧平均编码时间

图4 K帧视点编解码实验结果

平均编码时间分别为 2 940.5 ms. 因此本文算法耗时, 仅为“H.264 $RDO=1$ ”耗时的 5%左右, 但本文 K 帧算法编码性能与“H.264 $RDO=1$ ”相比降低了 0.3 dB 左右.

3 结论

本文选用 H.264 帧内编码技术对 WZC 的 K 帧进行编码, 并将帧间相关性引入到 WZC 的 K 帧帧内编码中, 同时为了降低 H.264 的复杂度, 设计了相应快速预测算法, 从 I4MB 和 I16MB 的提前判定、I4MB 中 9 种候选预测模式的快速选择等方面对 H.264 帧内预测编码进行了改进. 实验结果表明: 本文关键帧编码算法和 JM86 相比, K 帧编码时间降低 44.7%左右, 同时保持较好的率失真性能.

参考文献:

- [1] Wiegand T, Sullivan G J, Bjontegaard G. Overview of the H.264/AVC video coding standard[J]. IEEE Trans Circuits and Systems for Video Technology, 2003, 13(7):560-576.
- [2] Xu Q, Xiong Z. Layered Wyner-Ziv video coding[J]. Signal Processing, 2006, 86(11):3 212-3 225.
- [3] Sheinin V, Jagmohan A, He D. Uniform scalar quantization based Wyner-Ziv coding of Laplace-Markov source: ICASSP, Hawaii, 2007[C]. USA: Honolulu, 2007.
- [4] Weissmann T, ElGamal A. Source coding with limited-look-ahead side information at the decoder[J]. IEEE Trans on Information Theory, 2006, 52(12):5 218- 5 239.
- [5] Wei L, Zhao Y, Wang A. Improved side-information in distributed video coding: on innovative computing, information and control, 2006[C]. China: Beijing, 2006.
- [6] TU-TRec. H.264|ISO/IEC11496-10AVC.JM8.6[S]. 2004.
- [7] Jing X, Kebin J. A fast intra-frame prediction algorithm based on two-dimensional histogram for H.264/AVC[J]. Journal of Electronics and Information Technology, 2005 (7):1 053-1 057.
- [8] Zhang Kun, Yuan Chun, Li Qiang. A fast block type decision method for H.264/AVC intra prediction: ICACT 2007, 2007[C]. Korea: Phoenix Park, 2007.
- [9] Li Jiang Yu. New fast mode selection for intra prediction [J]. Acta Electronica Sinica, 2006(1):141-146.

Fast Key Frame Prediction Algorithm for Wyner-Ziv Video Coding

FAN Xu-ming, LI Fu-cui*, PENG Zong-ju, CHEN Ken

(Faculty of Information Science and Technology, Ningbo University, Ningbo 315211, China)

Abstract: Wyner-Ziv coding scheme is a low complexity encoding method for video compression applications with low power consumption. In this paper, K-frame is encoded by H.264 intra coder to improve Wyner-Ziv coding performance. The intra prediction of H.264 is improved in three aspects, i.e., the use of correlation between frames, block mode selection, and prediction mode selection for 4×4 luminance block to decrease the coding complexity. Experimental results suggest that the coding time of new K-frame coding framework takes up only 5% of time that H.264 intra coding consumes.

Key words: Key frame; Wyner-Ziv coding; low encoding complexity

CLC number: TN919.81

Document code: A

(责任编辑 章践立)