

文章编号: 1007-2985(2010)02-0060-04

基于块分割和上下文模型的视频编码算法^{*}

陈炳权^{1,2}, 刘宏立¹

(1. 湖南大学电气与信息工程学院, 湖南 长沙 410082; 2. 吉首大学物理科学与信息工程学院, 湖南 吉首 416000)

摘要: 树结构的构造适合小波图像编码, 但并不适合视频编码, 利用熵编码原理, 提出了视频图像编码算法, 先将每帧的残差补偿到原始序列, 并通过块分割编码和上下文模型进行视频编码, 研究表明: 该算法在视觉质量、算法性能和编码效率上要优于传统的视频编码算法。

关键词: 块分割; 上下文模型; 编码效率; 编码算法; 残差补偿

中图分类号: TN911.73

文献标识码: A

图像编码算法在近年取得了很大的进步, 其中集分割和上下文模型技术显示出了较好的编码性能. 集分割算法的信号大部分能量集中于最低频子带, 而高能量系数主要集中在图像的边缘, 该算法能够快速定位这个集合的高能量系数, 并且只要用少数符号就可以将低能量系数进行编码. 等级树集分割算法 SPIHT 将零树编码思想与集分割方法结合起来, 从而获得了很好的编码效果.^[1-2] 上下文模型利用小波系数间相关性进行编码, 由于决定编码率的条件熵相对于上下文阶数是单调递减的, 所以一个适当的高阶上下文会产生比较好的编码结果.^[3-5] 小波变换能对信号提供一个有效的频域表示, 但时-频平面铺砌并不总是匹配输入图像的频谱, 小波包分解^[6]使底层的小波基自适应地匹配任一给定的信号, 利用小波包进行图像编码可以使重建信号的 PSNR 质量相对于小波图像编码有较大提高.^[7] 由于基于块的分支能够取得与基于树分支相似的编码性能, 于是将集分割算法与上下文模型结合起来而形成一个高效的编码引擎, 通过对每个小波包子带采用块分割方案及上下文模型技术进行编码, 该算法克服了定义系数间父子关系的困难, 但对于视频编码显得无能为力. 利用文献[9]的残差补偿原理设计视频编码方案, 其算法性能、编码效率要优于传统的编码算法。

1 块分割编码算法和上下文模型熵编码

1.1 块分割编码算法

考虑小波包子带 P , 设它的一个像素点 x 的坐标是 (i, j) , 该点的系数值由 m_{ij} 来表示. 设 T 是子带 P 的一个子块, 若(1)式满足, 则称块 T 相对于阈值 n 是重要的, (1)式为

$$\max_{(i,j) \in T} \{ |m_{ij}| \} \geq 2^n, \quad (1)$$

否则称其为不重要的. 块 T 的重要性可表示为

$$\Gamma_n(T) = \begin{cases} 1 & 2^n \leq \max_{(i,j) \in T} \{ |m_{ij}| \}, \\ 0 & \text{其他情况.} \end{cases} \quad (2)$$

块分割编码算法采用位平面编码的方式, 用来定位每个位平面的重要系数, 同时将一组不重要系数仅用 1 个符号来编码, 因此, 该算法先检测一个系数块的重要性状态, 如果这个块在当前位平面 n 是重要的,

* 收稿日期: 2009-11-26

基金项目: 湖南省教育厅科学研究项目(08C714)

作者简介: 陈炳权(1972-), 男(土家族), 湖南桃源人, 湖南大学博士生, 副教授, 主要从事模式识别、信号处理与智能控制研究。

则被分割为 4 个相等的子块以定位其中的重要系数^[8]。

为了编码子带 P , 块分割算法从最高位平面(MSB) 编码至最低位平面(LSB), 在每个位平面有 3 个顺序的编码过程: LIP 过程、LIB 过程和 LSP 过程。LIP 过程扫描 LIP 链表来找到该链表中在当前位平面 n 的重要系数, LIB 过程扫描 LIB 链表以找到在当前位平面 n 重要的块, 并将这些块进行分割来定位其中的重要系数, LSP 过程扫描 LSP 链表输出在以前位平面已经重要的系数在当前位平面的细化比特, 完整的块分割编码算法总结如下:^[8-9]

(1) 初始化. 输出 $n_{\max} = \log_2(\max_{(i,j) \in P} |m_{i,j}|)$; 令 $n = n_{\max}$, 置 LIP 和 LSP 为空链表, 并将 P 加入到 LIB 链表中。

(2) 运行 LIP 编码. 对于 LIP(所有以前位平面不重要的系数链表) 中的每个像素 $x(i, j)$, 执行输出 $\Gamma_n(x)$, 如果 $\Gamma_n(x) = 1$, 则输出系数 $m_{i,j}$ 的正负号比特, 从 LIP 中删除 x , 并将其加入到 LSP 链表中。

(3) 运行 LIB 编码. 对于 LIB(所有以前位平面不重要的系数块的链表) 中的每个块 T , 执行输出 $\Gamma_n(T)$, 如果 $\Gamma_n(T) = 1$, 则从 LIB 中删掉 T ; 如果块 T 的系数个数大于 MinSize (MinSize 是 1 个预先给定的整数), 则将 T 分割为 4 个子块, 并将这 4 个子块顺序地加入 LIB 链表末尾, 否则运行 $\text{CodeMinBlock}(T)$ 进程。

(4) 运行 LSP 编码. 对于 LSP(在以前位平面已经重要的系数) 中的每个像素 $x(i, j)$, 但不包括由上 2 个过程在本次位平面 n 刚加入的像素, 执行输出 $|m_{i,j}|$ 在当前位平面 n 的细化比特。

(5) $\text{CodeMinBlock}(T)$ 进程. 按行扫描块 T , 对其中的每个像素 $x(i, j)$, 先输出 $\Gamma_n(x)$, 如果 $\Gamma_n(x) = 1$, 输出系数 $m_{i,j}$ 的正负号比特, 并将 x 加入到 LSP 链表中, 否则将 x 加入到 LIP 链表中. 参数 MinSize 用于控制块分割的粒度, 当 1 个块 T 的系数个数不大于 MinSize 时, 块 T 就不再继续分割了, 而是用按照行扫描的方式进行编码。

从上面的编码过程可以看到: 对于子带 P , 从最高位平面到最低位平面总共有 $3n_{\max} + 1$ 个编码过程, 其中 $3n_{\max} + 1$ 对应 P 的最高比特平面 MSB。^[8]

1.2 上下文模型熵编码

上下文模型熵编码原理利用相邻系数间的相关性而对变换图像进行压缩编码, 文献[8]根据细致的上下文模型, 静止图像压缩标准 JPEG2000^[5] 获得了很高的编码效率. 采用 JPEG2000 上下文模型作为块分割算法后的熵编码模型, 即每个块分割输出符号按照 JPEG2000 上下文模型进行熵编码. 在 JPEG2000 中, 1 个待编码系数的上下文由其 8 邻域系数的状态来得到, 其中第 0~ 8 个上下文是根据这 8 个邻域系数的重要性状态而得到, 并用于编码当前系数的重要性状态. 当前重要系数的正负号比特是由第 9~ 13 个上下文来编码, 这 5 个上下文是根据 8 邻域系数的正负号状态来产生. 已经重要系数的细化比特由第 14~ 16 个上下文进行编码, 而这 3 个上下文也是根据这个重要系数的 8 邻域系数重要性状态来产生. JPEG2000 的第 17 个上下文的产生不依赖于系数状态, 其上下文用于游程编码, 如表 1 所示. 块分割编码算法的每个编码过程所输出的比特符号连同它所对应的上下文标签一同送入到 1 个自适应算术编码器中^[9] 进行上下文算术编码。

表 1 块分割算法输出符号和相应的上下文标签

过程	比特输出	上下文标签
LIP	系数的重要性状态	0~ 8
	重要系数的正负号	9~ 13
LIB	系数块的重要性状态	17
	系数的重要性状态	0~ 8
	重要系数的正负号	9~ 13
LSP	已经重要系数的细化比特	14~ 16

在编码中若一个小块是重要的, 那么由于小波系数的能量聚集性, 这个小块的大多数系数更有可能是重要的, 所以基于上下文的熵编码会比继续分割这个小块取得更好的编码性能. 笔者的研究中, 最小块被限制为 4×4 , 即如果一个重要块 T 的大小不大于 4×4 , T 就不再进行分割, 而是按照行来扫描, 并把每个系数的重要性及可能的正负号比特进行基于上下文的熵编码。

2 残差补偿的视频编码原理

图 1 为最简单的多描述编码分帧框图, 信源 $\{x_k\}$ 经过编码器后生成 2 路描述, 分别送上 2 个猎的逻辑

信道. 在接收端有 3 个解码器, 分别处理收到一路、描述和二路同时收到的情况.

编码之前进行前处理可以降低二路视频码流编码时所引入的噪声的相关性. 令 x 为原始信号, x^1 , x^2 为编码后的信号, n_1 , n_2 为引入的编码噪声, 且其相关系数为 ρ 并令 n_1 , n_2 的方差为 σ_1^2 , σ_2^2 , 并设 α 为加权系数, 则可以证明噪声 n_m 的方差为

$$\sigma_m^2 = [1 - \frac{(1 - \beta)^2}{(\beta - \rho)^2 + (1 - \rho^2)}] \cdot \sigma_2^2, \quad (3)$$

其中 $\beta = \sigma_1 / \sigma_2$. 所以适当选择 α 经过加权平均后的图像质量有所提高^[9].

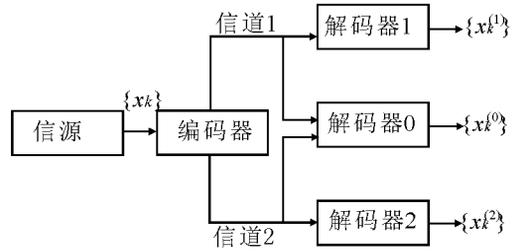


图 1 多描述编码的框图

3 算法设计

编码算法中, 块分割编码算法和上下文模型熵编码可以设置成 1 个编码器, 用于对视频图像. 块分割算法在编码 1 个子带 P 时, 从其 MSB 到 LSB 需要 $3n_{\max} + 1$ 个编码过程, 其中 n_{\max} 对应子带 P 的最高位平面, 可以对 P 产生 $3n_{\max} + 2$ 个率失真对, 第 1 个率失真对对应编码开始点, 而其余的对应每个编码过程的结束点. 在每一个率失真对中, 失真由子带 P 的原始系数与重建系数的误差平方和得到, 而相应的编码率由各编码符号在相应上下文下的自信息量之和来计算, 图 1 是完整视频编码算法的框图.

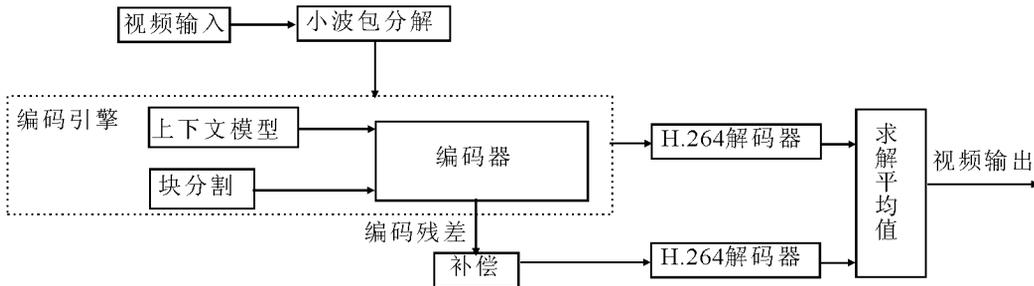


图 2 视频编码解码算法框图

完整的编码算法如下:

- (1) 将输入图像完全分解至 1 个给定分解层数 N , 从而得到 1 个完全分解二叉树;
- (2) 以块分割算法及基于上下文模型的编码器作为一个编码引擎, 对完全树上的每个节点产生 $3n_{\max} + 2$ 个率失真对, 其中 n_{\max} 对应子带 P 的最高位平面;
- (3) 利用子带编码引擎, 编码最优分解结构的每个小波包子带至其最优编码精度, 输出结果比特流.
- (4) 利用 H. 264/ AVC 的 JM10.0 视频编解码器进行仿真, 在编码中将每帧的残差补偿到原始序列, 之后进行编码.
- (5) 若二路能正确解码, 则进行合理平均, 可以得到较好的视频质量.

4 实验结果

根据图 2 的流程, 算法对 Barbara 图像进行了实验如图 3 所示. 视频中的每一帧图像都是 1024×1024 宽高, 且像素灰度值为 16 bit, 在实验中, 用来比较和评价各算法性能的客观指标用峰值信噪比 (PSNR) 编码效率及冗余度. 从图 3 所示的重建图像可以看出本文算法很容易地保持了图像中桌布和围巾的纹理信息, 而 SPIHT 部分地丢失了这些细节信息, 与其他的小波包编码算法相比, WP-BPO 比 CWP^[10] 和 WP-SFQ^[5] 分别都有所提高

仿真结果中第 1 路和第 2 路的平均 PSNR ≈ 36 dB, 但 2 路的平均 PSNR = 41 dB. 效率达 88.1%, 冗余度为 9%.



图3 Barbara 的原始图像和重建图像的比较

5 结论

将块分割编码算法和上下文模型有机结合,并通过残差补偿进行视频编码.该算法不仅在主观质量上,而且在客观度量上,都一致地优于其他视频编码算法,PSNR 和编码效率得到了较大的提高.

参考文献:

- [1] SAID A, PEARLMAN W A. A New, Fast and Efficient Image Coded Based on Set Partitioning in Hierarchical Trees [J]. IEEE Trans. Circ Syst Video Technology, 1996, 6(6): 243-250.
- [2] SHAPIRO J M. Embedded Image Coding Using Zerotrees of Wavelet Coefficients [J]. IEEE Trans. Signal Process, 1993, 41(10): 3445-3462.
- [3] WU X. High Order Context Modeling and Embedded Conditional Entropy Coding of Wavelet Coefficients for Image Compression. In: Thirty First Asilomar Conf On Signal, and Computers, 1997, 2: 1378-1382.
- [4] TAUBMAN D. High Performance Scalable Image Compression with EBCOT [J]. IEEE Trans Image Process, 2000, 9(7): 1158-1170.
- [5] TAUBMAN D, MARCELLIN M W. JPEG2000 Image Compression Fundamentals, Standards and Practice [M]. [s. l.]: Kluwer Academic Publishers, 2002.
- [6] COIFMAN R R, WICKERHAUSER M V. Entropy Based Algorithms for Best basis Selection [J]. IEEE Trans. Inform Theory, Special Issue on Wavelet Transforms and Multires [J]. Signal Anal, 1992, 38(3): 713-718.
- [7] XIONG Z, RAMCHANDRAN K, ORCHARD M T. Wavelet Packets Coding Using Space Frequency Quantization [J]. IEEE Trans. Image Process, 1998, 7(6): 892-898.
- [8] 杨永明, 许超. 一个基于小波包分解的率失真最优化块分割图像编码算法 [J]. 中国科学 E 辑, 2008, 38(8): 1204-1219.
- [9] 赵安邦, 余赟, 孙同仓, 等. 基于残差补偿的多描述视频编码 [J]. 中国科学 F 辑, 2009, 39(7): 751-756.
- [10] MEYER F G, AVERBUCH A Z, STRMBERG J O. Fast Adaptive Wavelet Packet Image Compression [J]. IEEE Trans. Imag Process, 2000, 9(5): 792-800.

Video Coding Algorithm Based on Block Segmentation and Context Model

CHEN Bing-quan^{1,2}, LIU Hong-li¹

(1. College of Electricity and Information Engineering, Hunan University, Changsha 410082, China; 2. College of Physical Science and Information Engineering, Jishou University, Jishou 416000, Hunan China)

Abstract: The tree structure quantification suits the wavelet image coding but doesn't suit the video coding. By using entropy coding principle, the authors put forward the video coding algorithm based on block segmentation and context model, which firstly compensates the initial sequence for the residual of each frame and then carries out the normal coding. Results show that the visual quality, performance, and coding efficiency of the new algorithm are superior to those of the traditional video coding algorithm.

Key words: block segmentation; context model; coding efficiency; coding algorithm; residual compensation

(责任编辑 陈炳权)