

可伸缩视频码流的熵编码组合优化方法

赵志杰, 刘增玉, 张立志, 李 云, 单大伟

(哈尔滨商业大学计算机与信息工程学院, 哈尔滨 150028)

摘要: 分析可伸缩视频编码系统中由嵌入式零树编码生成的码流特征, 提出一种利用最优预测估计思想将算术编码和游程编码相结合, 实现熵编码优化的新方法。实验结果表明, 该方法比传统自适应算术编码的效率提高了 10% 左右, 达到了进一步压缩数据的目的。

关键词: 视频码流; 可伸缩; 熵编码; 嵌入式零树编码

Combination Optimized Method of Entropy Coding for Scalable Video Bitstream

ZHAO Zhi-jie, LIU Zeng-yu, ZHANG Li-zhi, LI Yun, SHAN Da-wei

(School of Computer and Information Engineering, Harbin University of Commerce, Harbin 150028)

【Abstract】 This paper analyzes features of bitstream generated by Embedded Zerotree Wavelet(EZW) in scalable video coding system, and presents a new method that combines arithmetic coding and run length coding though optimal predictor to optimize entropy coding. Experimental results show that the coding efficiency average increases about 10% by using the new method than using traditional arithmetic, and it obtains the purposes of further data compression.

【Key words】 video bitstream; scalable; entropy coding; Embedded Zerotree Wavelet(EZW)

1 概述

随着信息和通信技术的发展, 视频编解码技术已成为多媒体研究的重要内容之一。可伸缩视频编码(Scalable Video Coding, SVC)是为解决视频信号在网络异构环境下的高效传输问题而产生的新一代视频编解码技术。可以根据不同用户需求和网络传输条件, 使用简单高效的提取模块对采用 SVC 技术产生的数据流进行编辑, 从而为用户提供完全的个性化服务。

文献[1]提出的嵌入式零树编码(Embedded Zerotree Wavelet, EZW)是可伸缩视频编解码系统中经典的二维离散小波变换系数编码方法。其后出现了一些改进算法, 如SPIHT (Set Partitioning In Hierarchical Trees)^[2], EZBC(Embedded Zero-Block Coding)^[3]等。针对EZW输出码流的熵编码实现, 文献[1]建议使用自适应算术编码, 文献[4]提出游程编码, 文献[5]提出先采用游程编码, 再对游程编码的输出码流进行算术编码的串行编码方式。但分析EZW形成的码流可以发现, 当EZW主扫描码流中零树根节点(Zero Tree Root, ZTR)出现概率偏低时, 游程编码效果较差。而当ZTR出现概率较高时, 游程编码效果优于算术编码。本文基于上述特征, 结合可伸缩视频编码系统生成码流的特点(基于 $t+2D$ 可伸缩视频编码结构), 为嵌入式零树编码后码流优化编码的实现, 提出一种由Astrom最优预报器将算术编码和游程编码自由结合在一起的组合熵编码方法。

2 可伸缩视频编码系统结构

图 1 描述了可伸缩小波视频的编码结构。编码系统先把原始视频序列分解为固定长度的视频帧组, 然后以帧组为单位完成运动估计和运动补偿时域滤波。运动估计生成的运动矢量经运动矢量编码后形成码流输出。完成运动补偿时域滤

波后, 形成的各高、低通帧分别进行空间的二维小波分解, 并以二维小波变换后的子带为单位进行 EZW。EZW 的输出码流采用本文提出的组合优化方法实现熵编码, 最终生成具有可伸缩特性的视频数据流。

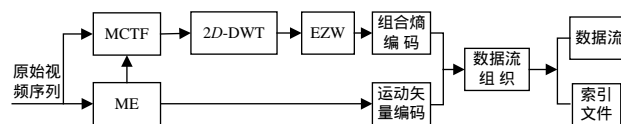


图 1 可伸缩小波视频的编码结构

3 熵编码

熵编码是利用信源概率统计特性进行编码的一种编码方法, 在视频压缩中常用的熵编码方法有霍夫曼编码、算术编码和游程编码。霍夫曼编码是一种形成前缀变长编码的方法。算术编码的思想是用区域划分来表示信源的输出序列。游程编码的基本思想是将具有相同数值、连续出现的信源符号构成的符号串用其数值及串的长度表示, 与前 2 种编码方法相比, 游程编码适合对符号串进行编码。

EZW 包括 2 个内容: 对重要系数的编码, 对零树根及孤立节点的编码。对重要系数的位置编码通过主扫描完成, 生成主表。对重要系数幅值的扫描, 通过副扫描完成, 生成副表。本文在分析 SVC 系统 EZW 生成码流特征的基础上, 对

基金项目: 国家自然科学基金资助项目(60772157); 黑龙江省自然科学基金资助项目(F200501)

作者简介: 赵志杰(1963 -), 男, 教授, 主研方向: 视频编码, 图像处理; 刘增玉, 硕士研究生; 张立志、李 云, 讲师; 单大伟, 硕士研究生

收稿日期: 2008-03-19 **E-mail:** zhijiezz@yahoo.com.cn

EZW 主扫描码流采用本文提出的组合优化熵编码方法；对 EZW 副扫描码流采用传统二进制自适应算术编码。

4 组合优化熵编码方法

本文提出组合优化熵编码方法，通过最优预测估计思想将算术编码和游程编码结合在一起，从而实现组合优化零树编码输出码流的方法。

4.1 组合优化熵编码结构模型

在图 1 所示的可伸缩编码系统中，继运动估计和运动补偿时域滤波、小波变换和嵌入式零树编码之后，零树编码副扫描的输出数据采用二进制自适应算术编码。由于零树编码主扫描码流具有以下特点：当 t 时刻码流的 $y(t)$ (位置映射的熵在总的信息熵中所占比例)较大时，采用游程编码的编码效果优于算术编码。在其他情况下，采用算术编码的效果优于游程编码。基于以上特点，本文提出先对零树编码主扫描采用 Astrom 最优预报器预测 $y(t)$ ，然后根据 $y(t)$ 选择编码方式，最后通过实际采集到的编码效果校正 $y(t)$ 和 Astrom 最优预报器的系数。组合优化熵编码模块的结构如图 2 所示。

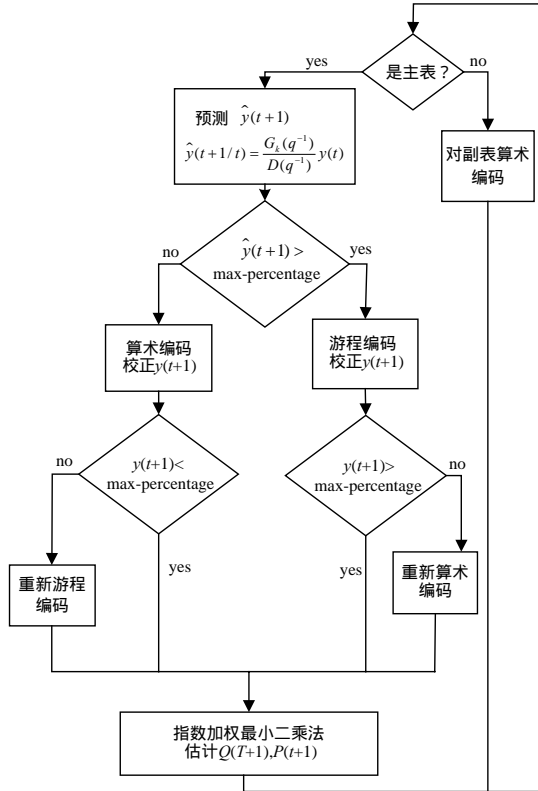


图 2 组合优化熵编码模块的结构

4.2 熵编码结构模型和 Astrom 最优预报器

本文对系统中的原子粒(最小编码单元)的主表数据部分采用组合熵编码(即游程编码和算术编码的混合编码)。具体选择何种编码方式根据 $y(t)$ 来确定。而 $y(t)$ 可通过以下 2 种方法获得：(1)编码之前遍历整个主表，计算出主表的 $y(t)$ ，从而确定选择游程编码或算术编码。(2)根据相邻原子粒相似性，通过已编码的原子粒进行预测。第(1)种方法通过遍历整个原子粒的主表来确定当前原子粒的数据组成，虽然简单易实现，但增加了编码系统的复杂度。为了降低编码计算复杂度，本文基于相邻原子粒之间的相似性，采用预测方法确定 $y(t)$ 。预测模型采用 ARMA(Autoregressive Moving Average)模型^[6]，即

$$A(q^{-1})y(t) = D(q^{-1})\varepsilon(t)$$

ARMA 过程超前一步 Astrom 最优预报器为

$$\hat{y}(t+1/t) = \frac{G_k(q^{-1})}{D(q^{-1})}y(t)$$

其中， $A(q^{-1}) = 1 - a_1q^{-1} - \dots - a_nq^{-n}$ ， $D(q^{-1}) = 1 - d_1q^{-1} - \dots - d_qq^{-q}$ ， $a_1, a_2, \dots, a_n, d_1, d_2, \dots, d_n$ 是模型参数； $q^{-1}, q^{-2}, \dots, q^{-n}$ 为滞后因子； $\varepsilon(t)$ 是零均值，方差为 δ^2 的白噪声。 $F_k(q^{-1})$ 为 $k-1$ 阶多项式， $G_k(q^{-1})$ 为 $\max(p-1, q-k)$ 阶多项式， $G_k(q^{-1})$ 和 $F_k(q^{-1})$ 由 Diophantine 方程确定，即 $D(q^{-1}) = A(q^{-1})F_k(q^{-1}) + q^{-k}G_k(q^{-1})$ 。

4.3 基于递推增广最小二乘法的 ARMA 模型系数估计

带时变参数的指数加权 ARMA 系数递推估计公式如下：

$$\hat{\theta}(t+1) = \hat{\theta}(t) + \frac{P(t)\varphi(t+1)[y(t+1) - \varphi(t+1)\hat{\theta}(t)]}{\lambda + \varphi^T(t+1)P(t)\varphi(t+1)}$$

$$P(t+1) = \frac{1}{\lambda} \left\{ P(t) - \frac{[P(t)\varphi(t+1)][P(t)\varphi(t+1)]^T}{\lambda + \varphi^T(t+1)P(t)\varphi(t+1)} \right\}$$

其中， $\hat{\theta}(t+1) = [a_1, a_2, \dots, a_n, d_1, d_2, \dots, d_q]$ 是第 $t+1$ 时刻 ARMA

模型系数组成的行向量； λ 是遗忘因子； $P(t) = [\Phi^T(t)\Phi(t)]^{-1}$ ；

$\Phi(t) = [\varphi^T(1), \varphi^T(2), \dots, \varphi^T(t)]^T$ ； $\varphi(t) = [-y(t-1), -y(t-2), \dots, -y(t-p),$

$\varepsilon(t-1), \varepsilon(t-2), \dots, \varepsilon(t-q)]^T$ ； $y(t)$ 的意义与前文定义相同。

$\varphi(t+1)$ 中的 $\varepsilon(t) = y(t) - \hat{y}(t/t-1)$ 。

$$\hat{Q}(t) = P(t)\Phi^T(t)Y(t)$$

$\hat{Q}(t)$ 是 t 时刻 ARMA 模型系数组成的行向量。

$$Y(t) = [y(1), y(2), \dots, y(t), \varepsilon(1), \varepsilon(2), \dots, \varepsilon(t)]^T$$

4.4 衡量熵编码效率的准则函数

零树根游程编码可分解为 2 个序列，第 1 个序列记录零树根长度和非零树根的位置信息序列(位置映射)；第 2 个序列用来记录依次出现的非零树根符号，记为非零树根信息^[7]。

如果 ZTR, IZ, POS, NEG 的概率分别记为 $P(ZTR), P(IZ), P(POS), P(NEG)$ ，则信源的熵为

$$H(x) = -P(ZTR) \times \text{lb}P(ZTR) - P(IZ) \times \text{lb}P(IZ) - P(POS) \times \text{lb}P(POS) - P(NEG) \times \text{lb}P(NEG)$$

位置映射的熵为

$$H_p = -P(ZTR) \times \text{lb}P(ZTR) - (1 - P(ZTR)) \times \text{lb}(1 - P(ZTR))$$

非零树根信息的熵为

$$H_v = -\left(\frac{P(IZ)}{(1 - P(ZTR))}\right) \times \text{lb}\left(\frac{P(IZ)}{(1 - P(ZTR))}\right) -$$

$$\left(\frac{P(POS)}{(1 - P(ZTR))}\right) \times \text{lb}\left(\frac{P(POS)}{(1 - P(ZTR))}\right) -$$

$$\left(\frac{P(NEG)}{(1 - P(ZTR))}\right) \times \text{lb}\left(\frac{P(NEG)}{(1 - P(ZTR))}\right)$$

零树编码主扫描输出序列的熵 $H(X)$ 可表示为

$$H(x) = H_p + (1 - P(ZTR)) \times H_v$$

上述信源公式表明了原始数据流的熵中各种信息占的比重。由于游程编码是一种有效压缩位置映射的编码，因此当位置映射在原始信源中占较大比例时，游程编码的效果优于算术编码。反之，则算术编码的效果优于游程编码。本文选择 t 时刻 $y(t) = H_p/H(x) \times 100\%$ 作为衡量在 t 时刻的码流中使用游程编码优劣的评价准则。

4.5 熵编码模块算法

熵编码模块算法步骤如下：

(1) $\hat{Q}(0)$ 和 $P(0)$ 初始化。

(2) 由 ARMA 过程超前一步 Astrom 最优预报器预测 t 时刻原子粒主表中的位置熵在原子粒主表总的熵中占的百分比 $\hat{y}(t)$ 。

(3) 根据 $y(t)$ 选择算术编码或游程编码。

(4) 对原子粒进行熵编码，并校正 $y(t)$ 。

(5) 由 $y(t)$ 校正 $\hat{y}(t)$ ，若 $\hat{y}(t)$ 预测的编码方法正确，则转第(7)步。

(6) 对当前原子粒重新进行熵编码。

(7) 由递推增广最小二乘法校正 ARMA 模型参数，转第(2)步。

5 实验结果及讨论

为了降低计算复杂度，ARMA 模型阶数不应该选择太高。本文实验采用 ARMA(2,1)模型。使用 QCIF 格式视频序列 mother_daughter, foreman, Suzie, News 作为实验样本。图 3~图 6 是采用本文方法进行不同层数 EZW 后的压缩比相对于单独使用算术编码的压缩比提高的百分比及其变化趋势。由图 3~图 6 可见，算术编码和游程编码自由结合后与单独用算术编码的方式相比，其编码效率平均提高 10% 左右，在低码率段压缩比提高更明显。

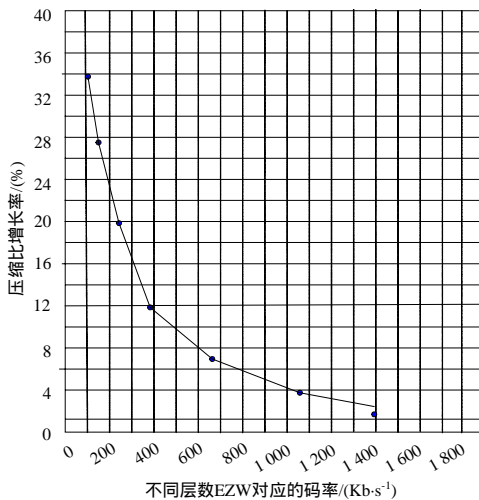


图 3 视频序列 mother_daughter 的实验结果

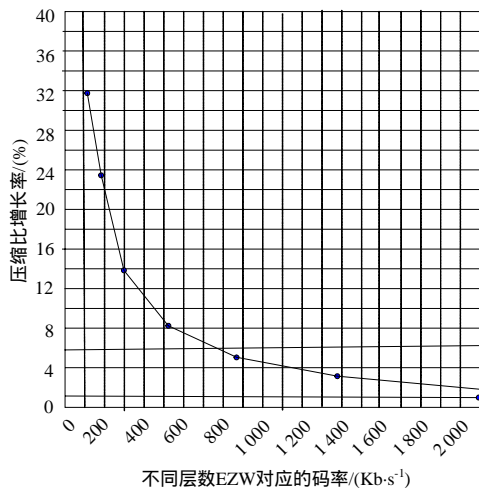


图 4 视频序列 foreman 的实验结果

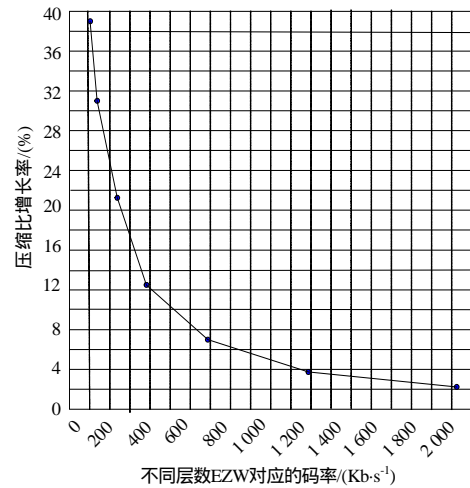


图 5 视频序列 Suzi 的实验结果

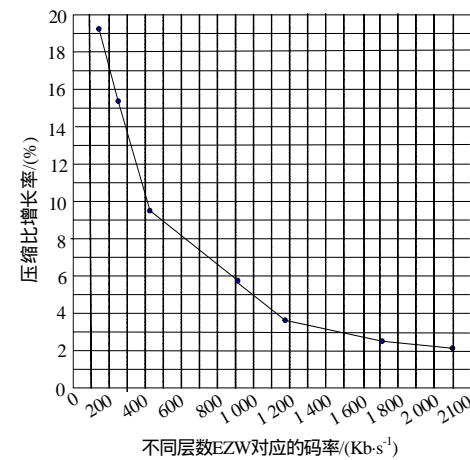


图 6 视频序列 News 的实验结果

6 结束语

本文方法进一步提高了系统的熵编码效率，且由于此方法在解码端只须增加一个判断语句，因此对系统的解码时间几乎没有影响。

参考文献

- [1] Shapiro J M. Embedded Image Coding Using Zerotrees of Wavelet Coefficients[J]. IEEE Transactions on Signal Processing, 1993, 41(12): 3445-3462.
- [2] Said A, Pearlman W A. A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees[J]. IEEE Trans. on Circuits Syst. Video Technology, 1996, 6(3): 243-250.
- [3] Hsiang S T, Woods J W. Embedded Image Coding Using Zero Blocks of Subband/Wavelet Coefficients and Context Modeling[C]//Proceedings of the IEEE International Symposium on Circuits and Systems. Geneva, Switzerland: [s. n.], 2000: 662-665.
- [4] 姜 威, 李文明, 张英和. 一种利用游程编码优化 EZW 的方法[J]. 山东大学学报: 工学版, 2004, 34(3): 63-67.
- [5] 黄卓君, 马争鸣. 一种零树与游程相结合的小波图像编码方法[J]. 中国图象图形学报, 2001, 6(11): 1118-1124.
- [6] 邓自立. 最优估计理论及其应用[M]. 哈尔滨: 哈尔滨工业大学出版社, 2005.
- [7] 张旭东, 卢国栋, 冯 健. 图像编码基础和小波压缩技术[M]. 北京: 清华大学出版社, 2004.