

# 高帧频数字图像序列实时压缩算法

隋玉萍<sup>\*1,2</sup>, 何 昕<sup>1</sup>, 魏仲慧<sup>1</sup>

(1. 中国科学院长春光学精密机械与物理研究所, 吉林 长春 130033;  
2. 中国科学院研究生院, 北京 100039)

**摘要:** 针对高帧频数字图像序列的特点, 提出了一种自适应近无损压缩算法, 将连续的  $n$  帧分为一节, 每节的第一帧进行基于帧内感兴趣区域(region of interest, ROI)的压缩, 其他帧在帧间差分 and 阈值处理后进行游程编码. 在第一帧内, 利用形态 Harr 小波提升和连通性分析检测出高频子带中的感兴趣区域. 对感兴趣区域采用 Rice 无损熵编码, 背景区域不编码. 低频子带直接输出. 实验结果表明, 该算法在帧内能有效地划分感兴趣区域, 压缩率在 30 以上, 峰值信噪比在 42 dB 以上, 目标无失真; 帧间压缩的压缩率在 7 左右, 背景区域有误差累积, 但整体图像的峰值信噪比都在 40 dB 以上, 目标无失真. 该算法的计算复杂度较低, 易于硬件实现, 而且还具有自适应性和数据包独立的优点, 适合于高帧频数字图像序列的实时压缩应用.

**关键词:** 高帧频数字图像序列; 实时压缩; 形态 Harr 小波; 感兴趣区域  
**中图分类号:** TP391      **文献标志码:** A

## 0 引言

高速数字图像采集系统(一般指帧频大于 100 f/s)可以对高速运动目标(运动速度为 800~1 000 m/s)进行捕获、记录及实时显示, 因而在靶场测量中得到了广泛的应用<sup>[1]</sup>. 高帧频数字图像采集系统提高测量精度的同时也使得数据输出量猛增, 例如分辨率为  $1\ 536 \times 1\ 024$ 、帧频为 100 f/s、位深度为 8 的高速数字图像采集系统, 每秒获得的数据量为 1 200 Mbits, 这对传输带宽和存储器都提出了很高的要求, 因而有必要对图像序列进行实时压缩.

目前已经具有多种视频压缩编码标准, 如 H. 26x 系列和 MPEG-x 系列. 两者主要应用于低速率视频压缩, 前者应用于实时视频通讯, 后者主要应用于视频数据的存储、广播电视和视频流的网络传输等, 所以这些标准不适合高速率数字图像序列的实时压缩应用. 张太镒等<sup>[2]</sup>对高速电视图像数据采用了帧间差分 and 游程编码相结合的方法, 其压缩率很低. 主要原因是受噪声、背景亮度

的变化、大气抖动等因素的影响, 即使相机不动, 连续两帧的背景同一位置的像素值也很难完全相同, 因而直接对差分后非零系数进行游程编码, 导致压缩率很低.

靶场测量中获得的数字图像一般由运动目标和比较简单的背景构成, 而用户一般只对目标感兴趣. 由于拍摄帧频较高而且相机不动或微动, 背景相对于目标变化缓慢, 邻帧图像的相关性很强. 根据图像序列的这一特点以及图像近无损压缩的要求, 本文采用在帧内进行基于感兴趣区域(region of interest, ROI)的近无损压缩方法, 即对感兴趣区域进行无损压缩, 而对背景区域进行大压缩比的有损压缩方法; 在帧间对图像序列进行阈值处理和差分编码.

## 1 算法描述

由于高速数字图像采集系统的帧频很高, 相邻帧之间的时间间隔仅为几毫秒, 在这么短的时间内很难完成一帧图像的压缩. 本文算法的解决

收稿日期: 2007-08-10; 修回日期: 2009-03-04.  
基金项目: “八六三”国家高技术研究发展计划资助项目.  
作者简介: 隋玉萍\*(1980-), 女, 博士, 大连理工大学 2002 届硕士, E-mail: cindysyp@sina.com; 何 昕(1966-), 男, 研究员, 博士生导师.

思路是在  $n$  帧的时间内完成  $n$  帧的压缩. 具体是将图像序列进行分节, 每节包括连续的  $n$  帧图像. 对每节的第一帧进行帧内压缩, 后面的  $n-1$  帧进行帧间压缩, 如图 1 所示.  $n$  的值是可以选择的, 但  $n$  值较大时, 因误差累积使得重构图像质量下降,  $n$  值较小时, 给帧内压缩留出的时间不充足, 仍不能实时实现. 本文中  $n$  值取 4. 分节有利于并行算法的实现, 而且可以有效地限制误差扩散.

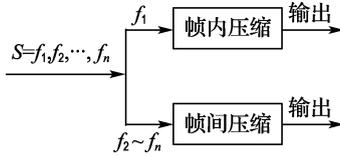


图 1 算法框图

Fig. 1 Diagram of the algorithm

### 1.1 帧内压缩

在帧内实现基于感兴趣区域的近无损压缩算法, 需要自动检测感兴趣区域. 国内外很多学者提出了一些感兴趣区域检测算法<sup>[3,4]</sup>, 这些算法可以很好地检测出感兴趣区域, 但是时间开销较大, 不适合实时应用. 由于图像中目标边缘灰度变化比较剧烈而且背景比较平坦, 可以采用基于边缘检测的方法来进行感兴趣区域分割. 形态 Harr 小波提升在图像边缘检测中有很好的应用效果<sup>[5]</sup>, 其最大提升后的高频子带中很好地保留了局部最大值, 并且不产生任何新的局部最大值<sup>[6,7]</sup>, 因而可以提高边缘检测效率.

帧内压缩流程图如图 2 所示. 首先进行 3 级形态 Harr 小波变换, 然后确定高频子带中的感兴趣区域, 并对其进行 Rice 无损熵编码. LL3 低频子带不编码, 直接输出.

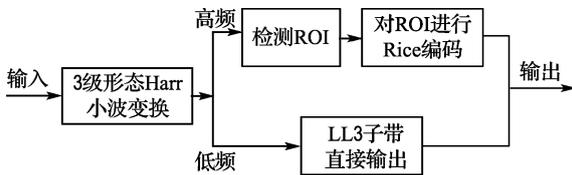


图 2 帧内压缩流程图

Fig. 2 The flow chart of intra-frame compression

1.1.1 形态小波变换 形态小波变换是以数学形态学为基础的一种小波变换, 兼顾了数学形态学的形态滤波特性和小波变换的多分辨率分析特性<sup>[8,9]</sup>. 数学形态学将图像看做是几何物体而不

是 Hilbert 空间的元素, 因而形态算子是非线性的<sup>[6]</sup>. 这种非线性小波具有良好的细节保留特性和抗噪声性能<sup>[10]</sup>.

形态 Harr 小波提升和传统小波提升一样, 包括 3 个步骤: 分裂、预测和更新.

#### (1) 分裂

将原始数据  $x(n)$  进行 lazy 小波分解.

$$x_e(n) = x(2n); x_o(n) = x(2n + 1)$$

#### (2) 预测

利用  $x_e(n)$  来预测  $x_o(n)$ , 得到预测误差  $d(n)$ .

$$d(n) = x_o(n) - (x_e(n) \vee x_e(n + 1))$$

#### (3) 更新

利用  $d(n)$  更新  $x_e(n)$ , 得到原始数据的一个近似  $c(n)$ .

$$c(n) = x_e(n) + (d(n - 1) \vee d(n) \vee 0)$$

式中  $(d(n - 1) \vee d(n) \vee 0) \geq 0$ , 为了使更新后的低频系数的动态范围较小, 本文将上式中的“+”改为“-”. 改进后的更新步骤为

$$c(n) = x_e(n) - (d(n - 1) \vee d(n) \vee 0)$$

1.1.2 高频子带的 ROI 检测 三级小波分解如图 3 所示. 小波分解后的图像感兴趣区域在不同分辨率下具有相似的形状, 而且分辨率低的子带感兴趣区域对分辨率高的子带感兴趣区域有预测全包含关系, 即 HL3 经过简单的  $2 \times 2$  的 Mosaic 放大, 其感兴趣区域可以覆盖整个 HL2 的感兴趣区域并有冗余, HL3 经过简单的  $4 \times 4$  的 Mosaic 放大, 其感兴趣区域也可以覆盖整个 HL1 的感兴趣区域并有冗余, 而且前者的冗余大于后者的<sup>[11]</sup>. 因而只需在第 3 级分解的一个子带中确定感兴趣区域即可. 能量较大的子带中包含绝对值大的系数较多, 不利于提取目标, 因此, 本文选择了在能量最小的 HH3 子带中确定感兴趣区域.

对 HH3 子带采用阈值处理<sup>[12]</sup>和基于四邻域的多目标连通性分析方法<sup>[13]</sup>确定感兴趣区域. 图 4 为 HH3 子带的统计直方图, 从直方图分布可以看出其系数分布近似于正态分布. 集中在零附近、绝对值较小的系数主要为图像背景信息, 而绝对值较大系数主要为目标和噪声信息. 目标在整幅图像中所占的百分比较小, 因而可以把绝对值较大的系数近似看成小概率事件, 根据概率统计学中的假设检验  $3\sigma$  法则, 将阈值设为

$$T_1 = \lceil 3\sigma \rceil \quad (1)$$

其中  $\sigma$  为 HH3 子带的均方差,  $\lceil \cdot \rceil$  为向上取整.

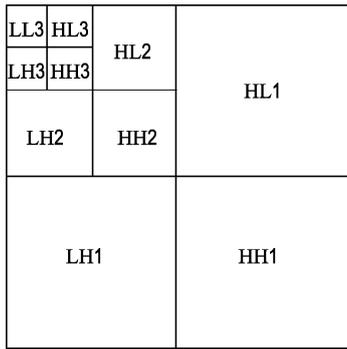


图 3 三级小波分解

Fig. 3 3-Level wavelet decomposition

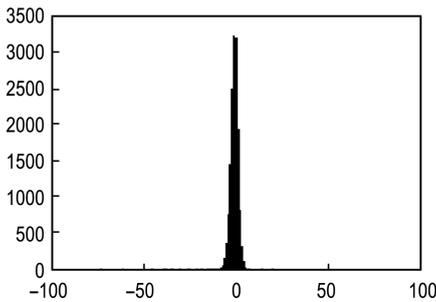


图 4 HH3 子带系数统计直方图

Fig. 4 Histogram of HH3 subband coefficients

在 HH3 子带中提取大值系数的位置信息, 生成一张二值映射表, 绝对值大于阈值的系数映射为 255, 绝对值小于阈值的系数映射为 0, 如图 5 所示.

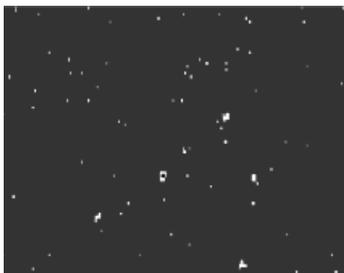


图 5 阈值处理后的 HH3 子带

Fig. 5 HH3 subband processed by threshold

对二值图像进行四邻域的多目标连通性分析, 使用其外接矩形描述 ROI, 将小于等于 3 个点的假目标区域剔除. 为了防止损失目标的边缘, 将矩形的坐标向水平和垂直方向各延长几个像素, 如图 6 所示.

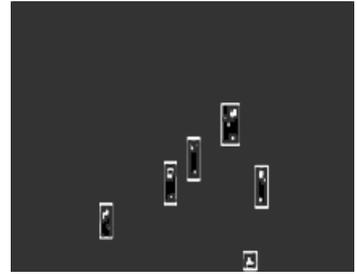


图 6 HH3 子带中的感兴趣区域

Fig. 6 The ROI of HH3 subband

1.1.3 高频子带编码 各高频子带包含感兴趣区域和背景区域. 由于背景区域的系数大多为零或接近于零, 而且对重构图像的质量影响很小, 背景区域不进行编码. 为保证目标的重构质量, 感兴趣区域采用无损压缩算法. 目前无损压缩编码方法有很多, 考虑到实时性要求, 本文选择了简单的 Rice 无损熵编码方法. Rice 编码器<sup>[14]</sup>是由 Rice 等提出来的, 它包括了很多编码选项以适应数据统计特性的变化. 但算法中较多的选项数目导致了表示这些选项的 ID 比特位的增加, 本文对此进行了优化, 优化后的 Rice 熵编码器结构如图 7 所示. 实验发现  $k$  值一般情况下小于 6, 而且第二扩展项和不压缩项效率比较低, 因此这两项和  $k$  值大于 6 的项可以省略.

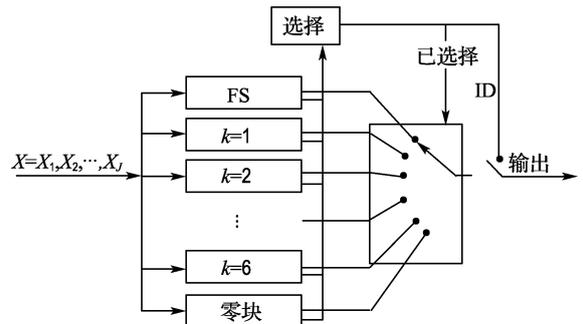


图 7 Rice 编码器的结构框图

Fig. 7 The architecture of Rice coder

Rice 编码器一次可以处理  $J$  个数据,  $J$  为任意值. 实验结果表明, 小于 16 的  $J$  值导致较低的压缩比, 大于 16 的  $J$  值不能很好地适应数据源统计特性的变化. 因此本文中参数  $J$  选为 16.  $k$  值是满足式 (2) 的小于等于 6 的最大非负整数<sup>[15]</sup>,

$$J \cdot 2^{k+7} \leq 128 \cdot \Delta + 49 \cdot J \quad (2)$$

其中  $\Delta$  是一个块中所有数据的和,

$$\Delta = \sum_{i=1}^l X_i \quad (3)$$

1.1.4 低频子带的编码 LL3 低频子带包含了图像的绝大部分能量,重要信息很多,因此需要对该子带进行无损或近无损压缩.由于 LL3 子带大小仅为原图像的 1/64,而且保持了很好的动态范围,本文算法对该子带采用不编码、直接输出的方法.

## 2 帧间压缩

图像序列的相邻帧之间存在很大的时间相关性,因而通过减少时间冗余,可以提高压缩编码的效率.一种有效的方法是基于块匹配的运动估计,这种方法已经被许多视频编码标准采用.但是由于在靶场测量应用中,电视系统的帧频较高,基于块匹配的运动估计算法很难实现实时性.因此,本文采用了传统的帧间差分方法.

$$d_k(x, y) = f_k(x, y) - f_{k-1}(x, y) \quad (4)$$

式中:  $f_k(x, y)$ 、 $f_{k-1}(x, y)$  分别表示第  $k$  帧和第  $k-1$  帧图像;  $d_k(x, y)$  表示差分图像.由于背景亮度的变化、大气抖动等因素的影响,即使相机不动,连续两帧的背景中同一位置的像素值也很难完全相同,差分后的图像中非零系数很多,仍不利于压缩.为了提高压缩比,本文采用了对差分后的图像进行阈值处理的方法.

$$d_k(x, y) = \begin{cases} d_k(x, y); & |d_k(x, y)| \geq T_2 \\ 0; & |d_k(x, y)| < T_2 \end{cases}$$

差分图像  $d_k(x, y)$  中背景的差分点近似服从均值为 0、方差为  $\sigma$  的高斯分布.  $d_k(x, y)$  的方差为  $\sigma'$ ,通过实验发现,  $\sigma = \alpha\sigma'$ ,  $\alpha$  取 0.2 ~ 0.6. 本文  $\alpha$  取 0.3. 由于相机不动,背景的差分图像中大值系数出现的概率很小,可近似看做小概率事件,根据概率统计学中的假设检验  $3\sigma$  法则,阈值为

$$T_2 = 3\alpha\sigma' \quad (5)$$

由于目标中心灰度和边缘的不同,即使前后两帧的运动目标有重叠,阈值处理后也不会出现空洞,这样保证了目标区域无损压缩.

对处理后的非零系数进行游程编码.码字包括两部分:一部分是非零系数前连续零系数的个数,另一部分是非零系数.为了限制误差扩散,每行系数独立编码.

## 3 实验结果

本文实验中,图像大小为  $1536 \times 1024$ ,帧频为 100 f/s,位深为 8 位.表 1 给出了图像序列中两节( $G_1$ 、 $G_2$ )的压缩比( $R_c$ )和峰值信噪比( $R_{psn}$ ).从表 1 可以看出,每节的第一帧获得了较大的  $R_c$ ,都在 30 以上, $R_{psn}$  都在 42 dB 以上;每节其他帧的  $R_c$  在 7 左右,由于误差累积,其  $R_{psn}$  逐渐降低,但都在 40 dB 以上.每节的平均  $R_c$  在 13 以上,平均  $R_{psn}$  在 42 dB 以上.

表 1 压缩比和峰值信噪比

Tab. 1  $R_c$  and  $R_{psn}$

	$n$	$R_c$	$R_{psn}/\text{dB}$
$G_1$	1	33.86	42.55
	2	7.25	42.46
	3	7.80	42.21
	4	7.35	41.90
$G_2$	1	30.82	42.58
	2	7.50	42.50
	3	6.99	42.37
	4	7.73	41.38

## 4 结 论

根据靶场测量中高帧频数字图像序列的特点,本文提出了一种将帧内近无损压缩和帧间差分相结合的算法.该算法有利于并行实现,能适应数据统计特性的变化,而且图像重构质量较好,目标无失真.压缩后的数据可以分组传输,能有效地限制误差扩散.算法复杂度低,易于硬件实现,适合于高帧频图像实时压缩的应用.

## 参考文献:

- [1] 高 昕,王 颖,苏增立,等. 实现高速大面阵图像摄取处理存储一体系统的技术途径[J]. 飞行器测控学报, 2002, 21(2):28-31
- [2] 张太镒,胡 飞. 高速电视图象数据实时压缩技术的研究[J]. 光子学报, 1997, 26(11):1015-1019
- [3] ITTI L, KOCH C, NIEBUR E. A model of saliency-based visual attention for rapid scene analysis [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998, 20(11):1254-1259
- [4] KADIR T, BRADY M. Saliency, scale and image description [J]. *International Journal of Computer Vision*, 2001, 45(2):83-105

- [5] 康志伟, 廖剑利, 何怡刚. 基于提升算法的不可分离小波图像边缘检测[J]. 华中科技大学学报(自然科学版), 2006, **34**(4): 56-62
- [6] HEIJMANS H J A M, GOUTSIAS J. Morphology-based perfect reconstruction filter banks [C] // **Proceedings of the IEEE-SP International Symposium on Time-frequency and Time-Scale Analysis**. Piscataway: IEEE, 1998
- [7] 段 汕, 梅键新, 秦前清. 形态小波变换在图像处理中的应用[J]. 中南民族大学学报(自然科学版), 2003, **22**(2): 81-85
- [8] 余 莉, 王润生, 韩方剑. 多分辨率形态学目标检测[J]. 计算机辅助设计与图形学学报, 2006, **18**(6): 849-853
- [9] 刘 循, 游志胜. 形态小波变换及其在图象编码中的应用[J]. 四川大学学报(自然科学版), 2003, **40**(1): 69-73
- [10] 张光新, 蔡晋辉, 周泽魁. 基于模板的二维多通道形态小波的构造[J]. 光电子·激光, 2005, **16**(2): 217-221
- [11] 毋立芳, 沈兰荪, 张晓玲. 基于感兴趣区的图象近无损压缩[J]. 中国图象图形学报, 2001, **6A**(6): 528-532
- [12] 田宝凤, 徐抒岩, 孙荣春, 等. 一种适合星上应用的遥感图像无损压缩算法[J]. 光学精密工程, 2006, **14**(4): 725-730
- [13] 魏仲慧, 何 昕, 郝志航. 多目标的实时连通性分析方法[J]. 光学精密工程, 1995, **3**(4): 23-26
- [14] RICE R F, YEH P S, MILLER W. Algorithm for a very high speed universal noiseless coding module [R]. Pasadena: Jet Propulsion Laboratory, 1991
- [15] CCSDS. Image data compression [CCSDS122.0-B-1]. Reston: CCSDS, 2005

## A real-time compression algorithm of high frame-frequency digital image sequence

SUI Yu-ping<sup>\*1,2</sup>, HE Xin<sup>1</sup>, WEI Zhong-hui<sup>1</sup>

( 1. Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China;  
2. Graduate School of the Chinese Academy of Sciences, Beijing 100039, China )

**Abstract:** According to the characteristics of high frame-frequency digital image sequence, an adaptive near-lossless compression algorithm is proposed.  $n$  consecutive frames were defined as a segment, its first frame was compressed based on ROI (region of interest) intra-frames, and the others were run-length coded after inter-frame difference and threshold processing. In the first frame, ROI of high-frequency subbands was detected by morphological Harr wavelet lifting and connectivity analysis. Rice lossless entropy encoder was used for ROI and non-coded for background. Low subband was directly output. Experimental results show that the algorithm can segmentize ROI effectively, CR (compression ratio) is higher than 30, PSNR (peak signal to noise ratio) of the whole image is higher than 42 dB, and the object keeps original fidelity. Inter-frame CR is about 7, background has cumulative error and PSNR is higher than 40 dB. The algorithm possesses low complexity and is easy to be implemented in hardware. Moreover, it has the merits of adaptability and independent data packet. Therefore the algorithm is suitable for high frame-frequency digital image sequence real-time compression.

**Key words:** high frame-frequency digital image sequence; real-time compression; morphological Harr wavelet; region of interest