0254-6124/2019/39(6)-847–06 *Chin. J. Space Sci.* 空间科学学报

SUN Jianwei, XUE Changbin, ZHENG Tie, ZHANG Zhongwei. Sequential imagery lossless compression algorithm for space astronomical observation (in Chinese). *Chin. J. Space Sci.*, 2019, **39**(6): 847-852. DOI:10.11728/cjss2019.06.847

面向空间天文观测的序列图像无损压缩算法*

孙建伟[1,2] 薛长斌[1] 郑 铁[1,2] 张忠伟[1]

1(中国科学院国家空间科学中心 北京 100190)

2(中国科学院大学 北京 100049)

摘 要 空间天文观测任务会获得大量天文图像. 对定点天体连续观测得到的序列天文图像具有时间及空间冗余较高的特点. 为了减少序列天文图像的存储与传输的数据量, 保证序列天文图像的完整性, 满足科学目标的任务需求, 需要对其进行无损压缩. 本文提出了一种利用帧内压缩与改进的帧间压缩相结合的无损压缩算法, 将序列天文图像的第一帧进行 JPEG-LS 帧内无损压缩编码, 其余帧进行改进的帧间无损压缩编码, 从而有效去除序列天文图像的时间及空间冗余, 提高序列天文图像的压缩比. 经过试验测试, 改进后的帧间压缩效果优于帧内压缩效果, 改进后的帧间压缩时间少于帧内压缩时间. 结果表明, 该算法简单且高效, 适用于对序列天文图像的无损压缩.

关键词 序列图像, JPEG-LS, 帧内压缩, 帧间压缩

中图分类号 V 524, TP 391

Sequential Imagery Lossless Compression Algorithm for Space Astronomical Observation

SUN Jianwei[1,2] XUE Changbin[1] ZHENG Tie[1,2] ZHANG Zhongwei[1]

1(*National Science Space Center, Chinese Academy of Sciences, Beijing* 100190)

2(*University of Chinese Academy of Sciences, Beijing* 100049)

Abstract In astronomical observation missions, a large number of sequential imagery is obtained. They are produced by fixpoint photography in a period of time, and all have some common features with high resolution and high time-space redundancy. The large volume of such imagery challenges on-board transmission and storage mission. Hence, an algorithm based on inter-frame prediction is proposed in order to reduce the images' bitrates. It includes the intra-frame compression algorithm and an improved inter-frame compression algorithm. The JPEG-LS is adopted to encoding the first image as the intra-frame compression algorithm. Then the other images are compressed by the inter-frame compression algorithm. First, the next image is predicted by its previous image, and then the prediction residuals are transmitted to JPEG-LS predictor for refining. Finally, the refined residuals are encoded by arithmetic encoder. Arithmetic coding does not require encoding each pixel like Huffman coding, and it is closer to the entropy of the image. The experiment result proves that

* 中国科学院空间科学战略性先导科技专项 Ⅱ 期项目资助 (XDA15320100)

2018-11-06 收到原稿, 2019-04-04 收到修定稿

E-mail: sunjianwei729@163.com

the proposed algorithm promotes the compression performance, and it has lower complexity and less compression time comparing with the traditional JPEG-LS. The proposed algorithm is suitable for real-time onboard compression of astronomical observation sequential imagery.

Key words Sequential imagery, JPEG-LS, Intra-frame compression, Inter-frame compression

0 引言

随着天文观测需求的不断发展, 天文图像要求的分辨率越来越高, 这导致其数据量随之增大, 欧洲的 DENIS Sky Survey, 美国的 2MASS Infrared Sky Survey, 法国、加拿大的 MegaCam Survey 和美国的 SDDS (Sloan Digital Sky Survey) 项目, 每天均产生大约 10 TB 的天文图像 [1]. 由于天文学家需要对天文图像进行完整观察和分析, 因而要求天文图像的失真度越小越好. 对天文图像进行无损压缩能很好地解决这一问题, 但由于空间载荷资源有限, 压缩算法需要尽可能简单而高效, 以便于硬件实现.

目前应用在天文图像的无损压缩主要有基于离散余弦变换 DCT (Discrete Cosine Transform) 的压缩, 如 JPEG; 基于差分脉冲编码调制 DPCM (Differential Pulse Code Modulation) 的压缩, 如 JPEG-LS[2]; 基于小波变换 WT (Wavelet Transform) 的压缩, 如 JPEG2000 和压缩感知 CS (Compressed Sensing)[3,4] 的方法. 文献 [5] 提出了一种基于零树小波的压缩方式, 其已被应用在詹姆斯 · 韦伯太空望远镜图像网络传输任务中. 文献 [6] 和 [7] 中将 CS 应用到天文图像的压缩工作, 解决了 ESA 发射的赫歇尔空间天文台 PACS 设备的图像传输问题. 文献 [8] 应用 WT 和嵌入零树编码进行天文图像远程传输. 文献 [9] 提出了一种基于分块 DPCM 和整数小波变换相结合的天文图像无损压缩算法. 文献 [10] 提出了一种基于压缩感知分块自适应压缩采样和加权自适应最小权变分法进行压缩. 但是上述工作仅限于对单幅天文图像进行压缩, 没有考虑相邻图像之间的时间相关性, 且 DCT, WT 及 CS 的计算量很大, 算法结构复杂. 基于 DPCM 的方法更适合于硬件实现.

对于空间天文观测, 一般几十秒或者几分钟拍摄一张图片. 由于其观测位置固定, 观测背景不会发生剧烈的变化. 针对这种观测需要的序列天文图像具有分辨率高、数据量大、时间与空间冗余较高等特点, 且科学目标要求为无损压缩传输. 本文提出一种帧内压缩与改进的帧间压缩相结合的无损压缩方法, 提高空间天文图像的压缩比, 解决空间天文图像的存储及传输中数据量大的问题.

1 压缩算法设计

由于星载设备的资源极其有限, 需要采用结构简单、性能稳定、压缩高效的压缩算法实现对天文图像的无损压缩. 本算法将第 1 帧图像称为独立帧 (I 帧), 对其进行 JPEG-LS 帧内无损压缩编码, 其余 n 帧图像称为预测帧 (P 帧), 对其进行改进的帧间无损压缩编码. 本算法的整体结构如图 1 所示.

2 帧内压缩算法

帧内压缩这里采用了 JPEG-LS 算法. Weinberger[11] 等 1996 年提出一种低复杂度、基于上下文的无损图像压缩算法, 即 LOCO_I (Low Complexity

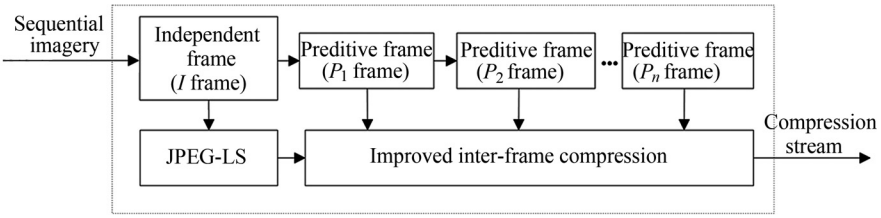


图 1 压缩算法结构

Fig. 1 Structure of the compression algorithm

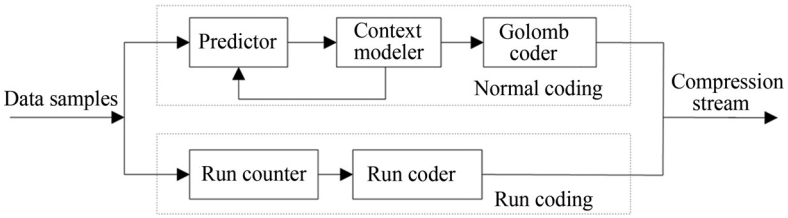


图 2 JPEG-LS 结构

Fig. 2 Structure of JPEG-LS algorithm

Lossless Compression for Image) 算法. 由于该算法的压缩性能很好, ISO/ITU 于 2000 年将其作为 JPEG-LS 的核心算法. JPEG-LS 的整体结构如图 2 所示.

JPEG-LS 算法依据图像样本是否有平坦区域分为正常编码和游程编码两种方式. 该算法先通过对当前像素求解其周围的梯度. 其中 x 为当前像素, a, b, c, d 是其周围的像素, 当前像素及其周围像素分布如图 3 所示.

如果当前像素周围的梯度均为 0, 即为平坦区域, 编码器选择进入游程编码方式, 否则进入正常编码方式. 梯度定义如下:

$$\begin{cases} g_1 = d - b, \\ g_2 = b - c, \\ g_3 = c - a. \end{cases} \tag{1}$$

2.1 正常编码方式

正常编码方式主要包括预测器、上下文建模及 Golomb 编码三部分.

2.1.1 预测器

预测器包括固定预测器和自适应校正. 固定预测器主要通过检测水平或垂直边缘实现预测. 其预测值为

$$P_x = \begin{cases} \min(a,b), & c \geqslant \max(a,b); \\ \max(a,b), & c \leqslant \min(a,b); \\ a+b-c, & \text{其他}. \end{cases} \tag{2}$$

当 $a \leqslant b$ 时, 如果当前像素 x 的左侧有一个垂直边缘, 则 x 的预测值为 a; 如果当前像素 x 的上方有一个水平边缘, 则 x 的预测值为 b; 如果没有检测到边缘, 则 x 的预测值为 $a+b-c$.

自适应校正部分利用预测修正值 $C[Q]$ 进行修正, 预测值修正公式如下:

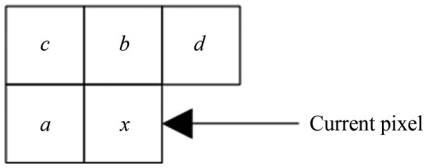


图 3 像素分布关系

Fig. 3 Pixel distribution

$$P_x = \begin{cases} P_x + C[Q], & \text{sign 为 } 1, \\ P_x - C[Q], & \text{sign 为 } -1. \end{cases} \tag{3}$$

预测误差为

$$e_x = x - P_x. \tag{4}$$

需要使预测误差 e_x 控制在 $(-\alpha/2, \alpha/2-1)$ 范围之内, 其中 α 表示图像灰度值的范围. 转换过程如下:

$$e_x = \begin{cases} e_x + \alpha, & e_x < -\alpha/2; \\ e_x - \alpha, & e_x \geqslant \alpha/2. \end{cases} \tag{5}$$

2.1.2 上下文建模

在正常编码方式中, 需要通过 g_1, g_2, g_3 的不同组合进行上下文建模. 但由于组合过多, 需要进行量化, 量化结果由 Q_i 表示. $Q_i = \{q_1, q_2, q_3\}$, 这里 q_1, q_2, q_3 表示 g_1, g_2, g_3 量化后的值. 量化分为 9 个等级, 量化区间分别为

$$\begin{aligned} &\{e|e \leqslant -21\}, \{-7, -8, \cdots, -20\}, \\ &\{-3, -4, -5, -6\}, \{-1, -2\}, \\ &\{0\}, \{1, 2\}, \{3, 4, 5, 6\}, \\ &\{7, 8, \cdots, 20\}, \{e \geqslant 21\}. \end{aligned}$$

考虑到对称性, 合并后上下文数目变为 365, 而编码时需要一个符号标志变量 sign 记录是否改变了符号, 以便解码时恢复.

建模中使用了

$$A[Q], B[Q], C[Q], D[Q], N[Q]$$

用于上下文的误差估计, $A[Q]$ 表示量化误差的绝对累计值, $B[Q]$ 表示修正后预测误差累计值, $C[Q]$ 表示预测修正值, $D[Q]$ 表示固定预测误差的累计值, $N[Q]$ 表示上下文出现的个数. 预测修正值 $C[Q]$ 为

$$C[Q] = \lceil D[Q]/N[Q] \rceil. \tag{6}$$

其中, 式 (6) 中 $\lceil x \rceil$ 为大于等于 x 的最小整数. 为避免出现除法, 式 (6) 可改写为

$$D[Q] = N[Q]C[Q] + B[Q]. \tag{7}$$

其中, 整数 $B[Q]$ 的范围是 $-N[Q] < B[Q] \leqslant 0$, 通过对 $C[Q]$ 和 $B[Q]$ 的存储可以实现对参数的更新. 首先将修正后的预测误差值 e_x 加到 $B[Q]$ 上, 然后通过加减 $N[Q]$ 使 $B[Q]$ 的范围一直满足 $-N[Q] < B[Q] \leqslant 0$, 则 $C[Q]$ 可根据 $N[Q]$ 是否加减来更新.

2.1.3 Golomb 编码

Golomb 编码是 Golomb[12] 于 1966 年提出的一种具有指数衰减概率分布输入的非负整数编码. 对于给定的一个非负整数 n 和一个正整数除数 m, Golomb 编码 G_m 分为两部分, 即 n/m 的商的一元编码和 $n \bmod m$ 的二进制编码. Golomb 编码在被编码数呈现具有概率质量函数 (Probability Mass Function, PMF) 几何分布时是最优的. PMF 为

$$Q(n) = (1-\rho)\rho^n, \quad 0 < \rho < 1. \tag{8}$$

此时, $m = [\mathrm{lb}(1+\rho)/\mathrm{lb}(\rho^{-1})]$. 当 $m = 2k$ 时, 称其为 Colomb-Rice[13] 编码, 这时计算更为简单, JPEG-LS 中应用了这种编码方式. 编码参数 k 的取值对编码十分重要, 计算过程为

$$k = \lceil \mathrm{lb}E[|e_x|] \rceil. \tag{9}$$

其中, E 为期望, 可以通过 $A[Q]/N[Q]$ 得到. k 的取值为

$$k = \min\{k' | 2^{k'} N[Q] \geqslant A[Q]\}. \tag{10}$$

JPEG-LS 中的编码采用了限长 Golomb 编码, 先将编码码字的最大长度设置为 $L_{\max}$ 比特, $\beta = \lceil \mathrm{lb}\alpha \rceil$, 将待编码数去掉低 k 位得到 q, 如果 $q < L_{\max} - \beta - 1$, 则按照正常编码参数为 k 的编码方式编码, 如果 $q \geqslant L_{\max} - \beta - 1$, 则先将 $L_{\max} - \beta - 1$ 个 0 编入, 再将待编码数 β 位的二进制表示编入比特流中.

2.2 游程编码方式

如果当前像素周围的梯度均为 0, 即 $a = b = c = d$, 则说明当前为平坦区域, 编码器进入游程编码方式. 游程编码方式主要包括计算游程长度、游程长度编码及中断像素编码三个部分. 当编码器进入游程编码后, 持续检测相同的像素直到不同像素出现, 然后利用分块编码的方式进行编码, 首先建立游程入口编码表: 0 0 0 0 1 1 1 1 2 2 2 2 3 3 3 3 4 4 5 5 6 6 7 7 8 9 10 11 12 13 14 15.

编码中使用了两个参数 r_k 和 r_m, 游程入口编码表中的数即为 r_k; 而 $r_m = 2r_k$, 代表代码等级, 编码时对每一个长为 r_m 的码字都用 1 编码. 如果编码被不同的像素中断, 则先编一个码字 0, 再用 r_k 个码字对剩余长度进行编码. 中断像素的编码与正常编码方式基本一致.

3 改进的帧间压缩算法

通常帧间压缩算法只包括帧间预测和编码两个部分. 本算法中改进的帧间压缩算法包括帧间预测、差值帧预测和算术编码三部分. 帧间预测是为了消除相邻两帧图像之间的时间冗余. 差值帧预测是在考虑天文图像分辨率高、空间冗余高的基础上, 将得到的差值帧进行进一步预测以减少空间冗余. 算术编码可以使编码效果更接近图像的熵. 帧间压缩的结构如图 4 所示.

3.1 帧间预测

帧间预测一般用于对视频的压缩. 相邻两帧之

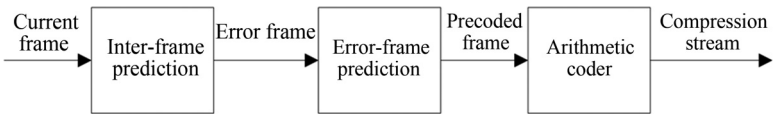


图 4 帧间压缩结构

Fig. 4 Structure of inter-frame compression

间具有较强的时间相关性, 利用帧间预测的方式可以很好地去除时间冗余, 提高整体的压缩比. 帧间预测的结构如图 5 所示.

预测器是将当前帧 $f(x,y,t)$ 的前一帧图像输出作为预测帧 $\hat{f}(x,y,t)$, 其表达式为

$$\hat{f}(x,y,t)=f(x,y,t-1). \tag{11}$$

残差帧 $r(x,y,t)$ 即为当前帧与预测帧的差值, 其表达式为

$$r(x,y,t)=f(x,y,t)-\hat{f}(x,y,t). \tag{12}$$

由于残差帧存在负值, 需要通过映射将负值转换为正值, 变换过程为

$$r(x,y,t)=r(x,y,t)-\min r(x,y,t). \tag{13}$$

将残差帧的像素值变换到 $(0,\alpha-1)$ 范围之内, 即可得到差值帧 $e(x,y,t)$. 其中, α 为像素值的范围, 变换过程如下:

$$e(x,y,t)=\begin{cases} r(x,y,t)-\alpha, & r(x,y,t)\geqslant\alpha; \\ r(x,y,t), & 0\leqslant r(x,y,t)<\alpha. \end{cases} \tag{14}$$

3.2 差值帧预测

由于天文图像的细节较丰富, 在帧间预测后仍存在大量空间冗余. 为了减少帧间预测后得到的差值帧包含的空间冗余, 将其做进一步预测. 差值帧预测采用 JPEG-LS 的固定预测器部分, 预测部分如式 (2) 所示, 预测误差 e_x 的值如式 (4) 所示. 然后, 使预测误差的值在 $(-\alpha/2,\alpha/2-1)$ 范围之内. 其中 α 表示图像的灰度值范围. 转换后即可得到预编码帧, 转换过程如式 (5).

3.3 算术编码

算术编码是将得到的预编码帧进行编码. 首先根据图像中所有灰度值出现的概率分成不同的区域, 然后按照被编码的第一个灰度值的概率区间扩展为编码的全部概率区间, 之后将这个缩窄的概率区间按照原始灰度值概率进行细分, 并继续对下一个灰度值进行同样的分配, 直到最后的灰度值, 最终得到的概率区间就可以表示这个编码. 这样的编码方式不需要像 Huffman 编码 [14] 一样每次将一个灰度值进行编码, 其比 Huffman 编码更接近图像的熵值 [15].

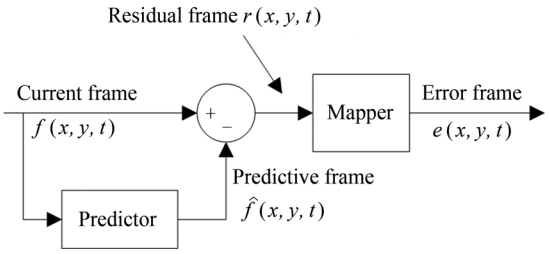


图 5 帧间预测结构

Fig. 5 Structure of Inter-frame prediction

4 实验测试与分析

这里采用两组连续采集的 7 帧日冕仪图像及 4 帧全日面 360 nm 波段的图像进行实验测试. 日冕仪图像均为 16 位 512×512 像素的灰度图像, 每帧之间的采集间隔为 5 min. 全日面 360 nm 波段图像均为 16 位 2048×2072 像素的灰度图像, 每帧之间的采集间隔为 1 min[16]. 分别将两组图像的第 1 帧进行 JPEG-LS 帧内压缩, 其余 n 帧进行改进的帧间压缩. 压缩得到的帧内与改进的帧间压缩效果对比列于表 1.

从表 1 可以看出, 这两种图像改进的帧间压缩效果均优于帧内压缩效果, 即改进的帧间压缩比大于帧内压缩比, 改进的帧间压缩很好地去除了图像的时间冗余和空间冗余. 全日面 360 nm 波段图像比日冕仪图像的分辨率更高, 细节变化更剧烈, 这导致其帧间

表 1 帧内与改进的帧间压缩效果对比 (bit·pixel^{-1})

Table 1 Compression performance of the intraframe and the improved interframe (bit·pixel^{-1})

图像类别	帧内压缩	改进的帧间压缩
日冕仪图像 1	5.88	—
日冕仪图像 2	5.88	4.27
日冕仪图像 3	5.88	4.17
日冕仪图像 4	5.90	5.12
日冕仪图像 5	5.93	5.45
日冕仪图像 6	5.94	4.59
日冕仪图像 7	5.94	4.34
全日面 360 nm 波段图像 1	5.33	—
全日面 360 nm 波段图像 2	5.41	5.21
全日面 360 nm 波段图像 3	5.39	5.16
全日面 360 nm 波段图像 4	5.42	5.34

表 2 帧内与改进的帧间压缩时间对比

Table 2 Comparison of the compression time between the intra-frame and the improved inter-frame

图像类别	帧内压缩/s	改进的帧间压缩/s
日冕仪图像 1	10.292	—
日冕仪图像 2	10.209	0.407
日冕仪图像 3	10.382	0.387
日冕仪图像 4	10.042	0.428
日冕仪图像 5	10.246	0.419
日冕仪图像 6	10.071	0.408
日冕仪图像 7	10.224	0.401
全日面 360 nm 波段图像 1	155.676	—
全日面 360 nm 波段图像 2	158.007	4.359
全日面 360 nm 波段图像 3	161.364	4.375
全日面 360 nm 波段图像 4	155.654	4.412

压缩效果略高于帧内压缩. 通过对残差帧计算熵值可知, 全日面 360 nm 波段图像目前达到的压缩效果已接近熵值.

从表 2 可以看出, 改进的帧间压缩时间比帧内 JPEG-LS 压缩时间缩短了很多. 这说明帧间压缩算法可以有效减少压缩时间.

5 结论

针对空间天体定点连续观测形成的序列天文图像分辨率高、时间及空间冗余高、数据量大的特点, 提出了一种帧内与改进的帧间相结合的无损压缩算法. 试验测试结果表明, 本文提出的算法有效地去除了帧间的时间及空间冗余, 提高了帧间压缩比, 缩短了帧间压缩时间. 该算法结构简单, 易于硬件实现, 对序列天文图像的无损压缩具有一定应用价值. 后续工作中将考虑采用滤波器的方式进一步减少帧间的时间及空间冗余.

参考文献

[1] STARCK J L, MURTAGH F. Astronomical image and data analysis [M]. New York: Springer Science & Business Media, 2007

[2] WEINBERGER M J, SEROUSSI G, SAPIRO G. The LOCO-I lossless image compression algorithm: principles and standardization into JPEG-LS [J]. *IEEE Trans. Image Process.*, 2000, **9**(8): 1309-1324

[3] CANDES E J, ROMBERG J, TAO T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information [J]. *IEEE Trans. Inf. Theory*, 2006, **52**(2): 489-509

[4] DONOHO D L. Compressed sensing [J]. *IEEE Trans. Inf. Theory*, 2006, **52**(4): 1289-1306

[5] THIENPHRAPA P, DONG J. A robust transmission system for astronomical images over error-prone links [C]// Proceedings SPIE, 2005, **6015**: 60150J

[6] BOBIN J, STARCK J L, OTTENSAMER R. Compressed sensing in astronomy [J]. *IEEE J. Sel. Top. Sign. Process.*, 2008, **2**(5): 718-726

[7] STARCK J L, BOBIN J. Astronomical data analysis and sparsity: From wavelets to compressed sensing [J]. *Proc. IEEE*, 2010, **98**(6): 1021-1030

[8] LI Long, DAI Hongbing, XU Jun. Application research of image compression based on wavelet transform in astronomical remote observation [J]. *Astron. Res. Technol.-Natl.* (*Astron. Observ.*), 2008, **5**(4): 380-385

[9] ZHU Fugui. Research and Implementation of Lossless Compression Algorithm for Astronomical Images [D]. Kunming: Kunming University of Science and Technology, 2009 (朱富贵. 天文图像无损压缩算法研究与实现 [D]. 昆明: 昆明理工大学, 2009)

[10] LI Yang. Research of Compressed Sensing in Astronomy Images [D]. Nanjing: Nanjing University of Information Science & Technology, 2014 (李洋. 压缩感知在天文图像中的应用研究 [D]. 南京: 南京信息工程大学, 2014)

[11] WEINBERGER M J, SEROUSSI G, SAPIRO G. LOCO-I: A low complexity, context-based, lossless image compression algorithm [C]//Data Compression Conference, 1996. DCC'96 Proceedings. Snowbird: IEEE, 1996: 140-149

[12] GOLOMB S. Run-length encodings (Corresp.) [J]. *IEEE Trans. Inf. Theory*, 1966, **12**(3): 399-401

[13] RICE R F. Some practical universal noiseless coding techniques [R]. California: Jet Propulsion Laboratory, 1991

[14] HUFFMAN D A. A method for the construction of minimum-redundancy codes [J]. *Proc. IRE*, 1952, **40**(9): 1098-1101

[15] GONZALEZ R C, WOODS R E. Digital image processing [M]. 3rd ed. Beijing: Publishing House of Electronics Industry, 2001

[16] SUN Jianwei, ZHANG Zhongwei, ZHENG Tie, XUE Changbin, CHEN Cong. Design of lossless compression system for CCSDS on-board data based on FPGA [J]. *Chin. J. Space Sci.*, 2019, **39**(5): 694-700 (孙建伟, 张忠伟, 郑铁, 薛长斌, 陈聪. 基于 FPGA 的 CCSDS 星载数据无损压缩系统设计. 空间科学学报, 2019, **39**(5): 694-700)