

一种基于分形小波变换的工业电视 图象编码方法研究

陈力军

(南京大学计算机科学与技术系, 南京 210093)

赵鸿萍

刘富强

徐立中

(中国矿业大学信电学院, 徐州 221008)

(河海大学电气工程学院, 南京 210098)

摘要 通过分析工业电视图象的特点, 提出了一种基于分形小波变换编码的工业电视图象压缩方法, 从而较好地解决了分形编码只能压缩静止图象的不足. 该方法不仅适合于煤矿工业电视图象压缩, 而且还可用于其它背景区图象相对静止的工业电视图象的压缩编码, 实验结果表明, 该法可取得较高压缩比和峰值信噪比, 因此具有实用价值.

关键词 分形编码 小波变换 工业电视图象

中图法分类号: TN919.81 文献标识码: A 文章编号: 1006-8961(2000)12-0994-05

Research on a Method of Fractal-Wavelet Transform Coding in Industrial TV Image

CHEN Li-jun

(Computer Science and Technology Department Nanjing University, Nanjing 210093)

ZHAO Hong-ping, LIU Fu-qiang

(College of Information and Electrical Engineering CUMT, Xuzhou 221008)

XU Li-zhong

(College of Electrical Engineering Hehai University, Nanjing 210098)

Abstract In the multimedia safety monitoring and detecting system of coal mine, the key technology is to compress the video signals which are from the subsystem of optic-fiber industrial TV. Fractal image coding can not be used in real time coding, because this method will have a delay when the code is picked up. After analyzing the feature of industrial TV image, the image is divided into background district and motive district by means of image enhancement and edge measurement. The background district is a static image, which can be compressed by fractal image coding. After that, a method of block-based image coding and wavelet transform coding to compress the image is presented. It is a good method for fractal coding to compress the motive image. It can be used not only in coal mine industrial TV image but also in any industrial TV image which background image is still and has practical value widely. At the end of the paper, a sample of compression is given. Fractals are a modern invention, they have been recognized as useful to natural scientists only over the last 20 years. Euclidean geometry provides concise accurate descriptions of man-made objects but is inappropriate for natural shapes. Fractals provide an excellent description of many natural shapes. In general, fractals are the result of a construction procedure and algorithm that is often recursive and suited to computers. Fractal coding is a new image coding method after subband coding, model-based coding and wavelet transform. It breaks through traditional coding theory frame. In real time coding and automatic coding it will have wide research prospect.

Keywords Fractal coding, Wavelet transform, Industrial TV image

0 引言

分形图象压缩编码是一种完全不同于传统编码方法的图象编码技术, 本文提出了一种将 Jacquin 的分形方块编码^[1]用于工业电视图象背景区图象压缩的方法.

在工业电视图象出现异常情况时, 则可以采用基于小波变换的工业电视图象压缩. 所谓小波变换是指将信号展开成一族小波函数之加权和的变换过程, 本系统则选用了一维正交小波基对图象进行压缩处理.

1 基于小波变换的压缩编码在工业电视图象异常时的编码方法

1.1 煤矿工业电视图象的背景区和运动区分离

通过对大量煤矿现场工业电视图象的分析, 发现大多数图象其运动部分约占全部图象的 30% 左右, 静止部分约占全部图象的 70% 左右, 因此可将工业电视图象的画面划分成背景区和运动区二个区域.

通过一个场景中两幅图象的比较, 可以显示出由于受物体运动的影响所造成的图象区域的不同. 其比较的一种方法是差分两幅图象. 可是对于实际的运动图象, 则可不利用象素间灰度值的差, 而是通

过研究小区域(如 2×2 或 4×6) 间的相似性来进行差分, 下式可作为研究小区域不匹配的尺度

$$\frac{|2(S_C + S_R) + (m_C - m_R)^2|^2}{S_C S_R} > t \quad (1)$$

式中, m_C 和 m_R 是当前画面和比较画面中的小区域灰度平均值, S_C 和 S_R 是各小区域灰度的方差, t 是阈值.

对于连续的画面, 可通过研究其相似度, 将累计不匹配的次數做成一次差分图象, 然后从它的性质中进行运动图象的解析. 若把画面间的一次差分图象再差分, 其所求得的二次差分图象, 则能更有效地检测出运动的对象物. 图 1 为运输皮带的背景区图象.

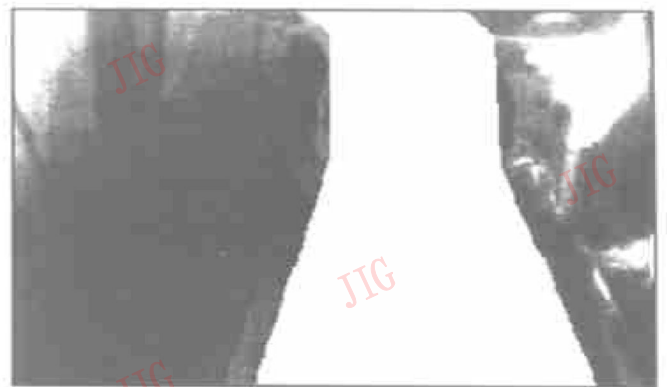


图 1 运输皮带的背景区图象

1.2 编码解码框图

编码、解码系统框图^[2]如图 2 所示.

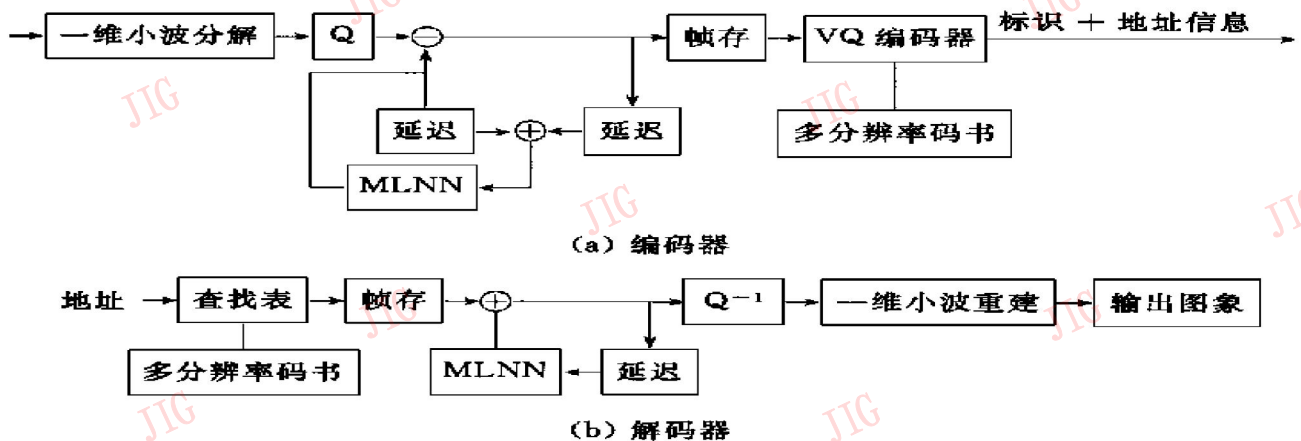


图 2 工业电视图象异常时的编码、解码系统框图

1.3 编码步骤

步骤 1 背景区图象的选取. 选皮带机图象作为背景区图象帧存, 并传至目的地帧存;

步骤 2 图象增强、边缘检测与运动区识别;

步骤 3 运动区图象压缩.

运动区图象压缩的具体做法如下:

(1) 一维小波分解及量化

在对图象进行小波分解前, 可利用小波包算法选取最佳基. 具体做法是: 先令 H 为使编码比特率最低的目标函数, 同时取 3~5 幅以往记忆的图象进行统计, 并选取最常成为最佳基的那一组正交小波基作为小波分解的基; 然后基于此基, 再应用 Mallat 算法对

每一图象行进行4层(即 $M=4$)小波分解展开。

对于小波分解后的多分辨率子图象,则根据人类视觉系统对不同频带内的信号失真所具有的敏感度不同,分别采用相匹配的间隔量化和门限抽样,即,对 $A_{M,n}$ 采用量化间隔为1的均匀量化;在分辨率 2^j 层,采用下式量化

$$D_{j,n}^q = \begin{cases} \Delta_j \cdot \text{INT}\{(D_{j,n} - T_j)/\Delta_j + 0.5\} & (D_{j,n} > T_j) \\ \Delta_j \cdot \text{INT}\{(D_{j,n} + T_j)/\Delta_j - 0.5\} & (D_{j,n} < -T_j) \\ 0 & (|D_{j,n}| \leq T_j) \end{cases} \quad (2)$$

其中, $D_{j,n}^q$ 为与 $D_{j,n}$ 对应的量化值, Δ_j 为量化间隔, T_j 为量化门限。在实验时,迭代次数的取值完全由算法何时收敛决定。

(2) 行间预测

采用3层前馈神经网络对相邻行象素值进行行间预测,其目的是对运动区图象进行编码。对于多层网络,它的学习时间一般很长,而本文采用的是SOFM算法,由于其在码书较大和训练序列较短时,学习速度快,并可事先控制训练时间,因而编码时间短,且能逼近全局最优。

本步骤需设计4个神经网络非线性预测器,分别对应于4级分辨率的小波图象。在确定了神经网络的拓扑结构后,就需要给出合适的学习准则,并通过对样本的学习来提取出样本集内隐含的规律,从而完成信息的相关存储和非线性映射,其信息存储方式采用分布式,而非线性映射是并行完成的。本文采用反向(BP)学习算法,该算法是有教师的学习算法,其学习过程由正向传播和反向传播组成,具体方法如下:

① 置所有网络权值 $\{W_{i,j}\}$ 与门限值 $\{\theta\}$ 为较小的随机数,并令 $t=0$,这里, $W_{i,j}$ 表示第 i 个神经元与下一层第 j 个神经元的连接权; θ 表示下一层第 j 个神经元的门限值;

② 输入一个 K 维训练样本;

③ 用下式计算第 $k(k=2,3)$ 层第 j 个神经元的输出 V_j^k

$$V_j^k = \frac{1}{1 + e^{-u_j^k}} \quad (3)$$

其中,

$$u_j^k = \sum_i W_{i,j} V_i^{k-1} + \theta \quad (4)$$

④ 用下式调整连接权值

$$\Delta W_{i,j}(t+1) = -\epsilon d_j^k V_i^{k-1} + \alpha \Delta W_{i,j}(t) \quad (5)$$

其中, d_j^k 表示第 k 层第 j 个神经元的误差, ϵ 为

学习率参数,用其控制网络收敛的速度, α 为惯性因子,一般取 $0 < \alpha < 1$ 。

⑤ 置 $t=t+1$,回到第②步,直到均方误差小于预先给定的允许误差为止。

(3) 运动图象的矢量量化编码

鉴于8维空间的SOFM算法具有很强的抗信道误码能力,因此,采用8维空间的SOFM算法来设计差值图象的码书,具体算法如下:

设网络有 N 个输出节点,各节点对应 K 维权矢量 $W_{ij}, j \in [1, N]$;输出节点定义在 M 维空间中,坐标为 $\{U_j(m), m \in [1, M]\}$;输出节点 j 的邻域定义为 NE_j ,则8维空间的相应定义如下:

节点 j 的坐标 $U_j = (U_j(1), U_j(2), \dots, U_j(8))$,其邻域定义为 j 相邻点的集合,即

$$NE_j = \{U_{j'} \mid d(U_j, U_{j'}) < \text{门限} D\} \quad (6)$$

其中,

$$d(U_j, U_{j'}) = \sum (U_j(m) - U_{j'}(m))^2 \quad (7)$$

节点 j 的坐标

$$U_j = (j \bmod Z, [j \text{ div } Z] \bmod Z, [j \text{ div } Z^2] \bmod Z, \dots, [j \text{ div } Z^7] \bmod Z), Z = [\lg M N] \quad (8)$$

实验时, Z 取2。式中 $[\cdot]$ 表示取整运算,mod是整数模运算,div是整数除运算。

2 工业电视图象背景区分形编码方法

本文根据工业电视图象背景区在正常情况下相对静止的特点,即视该区域为静止区,因而采用由Jacquin提出的分形方块编码方法来进行压缩编码。

2.1 基于PIFS的背景区分块编码方法

基于PIFS (Partitioned Iterated Function System)的分块编码方法^[1,3]是利用图象部分与部分之间的相似性,先将图象分割成不相互交叠的小方块,然后从某一小方块开始,搜索与其相匹配的更大一些的方块,整个过程全部由计算机完成。

2.1.1 几何变换^[4]

首先定义大方块与小方块的尺度比为尺度压缩因子,一般来说,尺度压缩因子小一些,更容易找到匹配,但解码时误差较大;反之,则较难找到很好地匹配,但解码时误差较小。这里,匹配的过程包括几何变换与亮度变换,几何变换又包括空间尺度的压缩和空间位置的转换,此时的三维仿射变换可表示为如下形式

$$\omega \begin{vmatrix} x \\ y \\ z \end{vmatrix} = \begin{vmatrix} a_i & b_i & 0 \\ c_i & c_i & 0 \\ 0 & 0 & s_i \end{vmatrix} \begin{vmatrix} x \\ y \\ z \end{vmatrix} + \begin{vmatrix} e_i \\ f_i \\ o_i \end{vmatrix} \quad (9)$$

它可进一步分解为

$$\omega \begin{vmatrix} x \\ y \end{vmatrix} = \begin{vmatrix} a_i & b_i \\ c_i & d_i \end{vmatrix} \begin{vmatrix} x \\ y \end{vmatrix} + \begin{vmatrix} e_i \\ f_i \end{vmatrix} \quad (10)$$

$$\omega'(z) = s_i z + o_i \quad (11)$$

式(10)为几何变换,其中 e_i, f_i 是大尺度方块的起始位置,而 ω 则决定了 Domain Block 是如何映射到 Range Block 上的,它反映了两者间的位置、比例等几何关系,由 Domain Block 与 Range Block 的分割可知,其压缩因子为 1/4,式中, a_i, b_i, c_i, d_i 是几何位置变换矩阵的系数,按 Jacquin 的方案取 8 种对称变换。

2.1.2 灰度比例因子与灰度平移量的确定^[5]

灰度比例因子 s_i 与灰度平移量 o_i 是灰度的线性变换式(11)中的两个常系数,它们可按如下方法确定:

其中,灰度比例因子 s_i 定义为

$$s_i = \frac{\left| n \left| \sum_{j=1}^n u_j v_j \right| - \left| \sum_{j=1}^n u_j \right| \left| \sum_{j=1}^n v_j \right| \right|}{\left| n \left| \sum_{j=1}^n u_j \right|^2 - \left| \sum_{j=1}^n u_j \right|^2 \right|} \quad (12)$$

式中, u_j 与 v_j 分别为变换后的 Domain Block 与 Range Block 的象素值。

为简化论述, s_i 通常定义为 Range Block 中象素之间的最大灰度差去除以 Domain Block 中象素之间的最大灰度差。

而灰度平移量 o_i 定义为

$$o_i = \left| \sum_{j=1}^n v_j - s_i \sum_{j=1}^n u_j \right| / n \quad (13)$$

为简化论述, o_i 通常定义为 Range Block 中象素灰度均值与 Domain Block 中象素灰度均值之差。

2.1.3 相似度的确定

首先将 $2N \times 2N$ 的 Domain Block 经过压缩采样,具体做法可简单地用一个象素来代替,即将图象转变成 $N \times N$ 的大小,再经过矩阵变换和灰度的线性变换,将其变换为 $\{D_{i,j}\}_{N \times N}$,将 Range Block 变换成为 $\{R_{i,j}\}_{N \times N}$,若定义它们之间的方差为

$$E = \sum_{i=1}^N \sum_{j=1}^N (s \times D_{i,j} + o + R_{i,j})^2 \quad (14)$$

则它们之间的相似度就是 $\min E$,即 Domain Block 经过几何变换和灰度变换后与 Range Block 之间的最小方差。

2.1.4 编码过程

原始图象经上述方法分割后,还要对每一个 Range Block 找到一个 Domain Block,并把所有 Domain Block 的集合称为 Domain Pool,再从 Domain Pool 中选择一个和它最相似的 Domain Block,然后将该 Range Block 和对应的 Domain Block 的信息记录下来,也即记录 Domain Block 的起始位置 e_i, f_i 和几何变换矩阵编号,以及灰度线性变换的两个常系数 s_i, o_i ,从而即得到所求的编码。

根据上述算法,本文对煤矿运输皮带的背景区图象进行了压缩实验(如图 3 所示)。其图象处理均是在视霸卡 VIDEO BLASTER RT 300 环境下进行,其实验图象为 320×240 象素的 8bit 灰度图象。

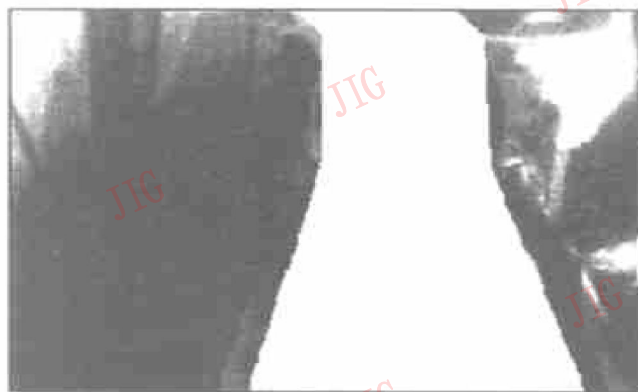


图 3 煤矿运输皮带背景区图象分块编码解码图
(压缩比= 26.51, PSNR= 31.3dB)

3 结束语

本文提出的基于分形-小波变换的混合编码方法,克服了分形图象压缩编码在 IFS 代码提取过程中有延迟的缺点。同时,对于运动图象和工业电视图象出现异常时,则可采用小波变换编码。该方法不仅适合于煤矿工业电视图象压缩,而且还可用于其它背景区图象相对静止的工业电视图象的压缩编码,并具有实用价值。

参 考 文 献

- 1 Jackquin A E. Image coding based on a fractal theory of iterated contractile image transformation. IEEE Trans on image Processing, 1992, IP-1(1): 18~ 30.
- 2 于洪珍,赵鸿萍,王正友. 基于小波理论的工业电视图象压缩研

究. 东南大学学报, 1997, 27(5B): 114~ 117.

- 3 Barnsley M F, Sloan A D. A better way to compress images. BYTE Magazine, 1988, 13(1): 215~ 223.
- 4 谢和平, 薛秀谦. 分形应用中的数字基础与方法. 北京: 科学出版社, 1997.
- 5 叶航军, 杨士强. 基于四叉树分割的分形图象压缩算法. 见: 第六届全国多媒体技术学术会议论文集, NCMT '97, 西安交通大学, 1997.



刘富强 博士, 博士后, 中国矿业大学信电学院副教授. 感兴趣的研究领域为图象压缩编码, 多媒体技术等.



陈力军 1961年生, 博士. 现在南京大学计算机科学与技术系做博士后. 感兴趣的研究领域为图象压缩编码, 模式识别, 智能CAD等. 现已发表论文10余篇. 曾获省部级二等奖2次.



徐立中 博士, 博士后, 河海大学电气工程学院副教授. 感兴趣的研究领域为图象压缩编码, 多媒体技术, 自动化系统理论及系统综合与集成技术等.

赵鸿萍 硕士, 中国矿业大学信电学院. 感兴趣的研究领域为图象压缩编码.

第六届国际青年计算机会议征文通知

(The Sixth International Conference for Young Computer Scientists)

国际青年计算机会议(ICYCS)是由中国计算机学会主办的系列国际会议, 每两年召开一次, 至今已成功举办了5届. 第六届国际青年计算机会议(ICYCS' 2001)由中国计算机学会主办、浙江大学承办, 将于2001年10月23~25日在杭州浙江大学召开.

会议主题: 新世纪信息技术与应用, 主要涉及: e-business, e-man, e-engineering, e-learning等.

届时将有众多中外著名计算机专家及后起之秀出席会议. 会议将为与会者提供一个学术论坛, 供与会者与专家进行学术交流, 介绍和探讨计算机理论及应用领域的最新进展. 会议论文集将由国内著名出版社出版, 会议收录论文将进入EI检索.

征文范围主要有以下领域(但不限于): 软件理论与实践、计算机系统结构与芯片设计技术、网络通信与信息安全技术、智能信息处理技术、理论与计算、分布式及并行处理技术、图形图象技术、生物信息理论与技术、商业领域中的信息技术及应用、工业领域中的信息技术及应用、教育领域中的信息技术及应用.

论文要求:

- (1) 会议语言是英语, 提交的论文也必须用英语写作;
- (2) 第一作者必须是45岁以下;
- (3) 论文必须有题目、作者全名、单位全称、通信地址、电话传真、Email及关键词. A4幅面5页左右的全文或2~3页包含有主要结论的摘要均可接收. 论文必须是没有公开发表的.
- (4) 附相应的中文摘要(包括题目、作者、单位、通信地址), 以便于联系;
- (5) 论文格式: 用PDF格式或MS-word格式输入排版, 通过Email发送或拷入软盘邮寄到ICYCS' 2001组委会;
- (6) 提请注意: 只接收电子稿.

关于征文的更详细情况以及会议的最新消息请访问会议网站: <http://www.icycs2001.org>

重要日期:

接收论文截止日期: 2001年5月20日; 论文录用通知: 2001年6月25日; 正式打印稿: 2001年8月15日.

论文请用Email发到: icycs@cs.zju.edu.cn

大会组委会: 杭州玉泉浙江大学计算机系 张三元 魏宝刚 邮编: 310027

联系电话: (0571) 7951909-204, 7951904-204 唐路 毛浙西