

基于稀疏性与同步性的数字文献图像压缩

师黎, 王端, 王治忠

(郑州大学电气工程学院, 郑州 450001)

摘要: 提出基于稀疏性与同步性的数字文献图像压缩方法。基于动物视觉系统的稀疏性, 提取图像基函数。针对响应系数, 利用同步性对基函数进行筛选, 完成对数字文献图像的压缩。实验结果表明, 该方法具有较高压缩比, 能通过少量样本实现对其他数字文献图像的压缩, 且压缩效果理想。

关键词: 数字文献图像; 图像压缩; 稀疏性; 同步性; 基函数

Digital Literature Image Compression Based on Sparsity and Synchronization

SHI Li, WANG Duan, WANG Zhi-zhong

(School of Electrical Engineering, Zhengzhou University, Zhengzhou 450001)

【Abstract】 This paper proposes a method of digital literature image compression based on sparsity and synchronization. The image basic functions are extracted based on the sparsity of animals' vision system. Aiming at the response coefficients, the basic functions are selected by using synchronization, and the digital literature image compression is accomplished. Experimental results demonstrate that this method has high compression rate, can compress other digital literature images by using a few samples, and has satisfactory compression effect.

【Key words】 digital literature image; image compression; sparsity; synchronization; basic function

1 概述

随着现代社会对信息交流与存储要求的不断提高, 图像质量和存储容量之间的矛盾日益突出, 具有庞大数据量的数字文献图像已成为数据通信和共享的瓶颈问题^[1]。图像压缩技术利用图像数据固有的冗余性和相干性, 用尽可能少的字节数表示原始数据, 以减轻图像的存储和传输负担。

第1代基于信息论的变换压缩编码由 Andrews H C 等人于 1968 年提出, 其压缩比较小、图像复原质量差。文献[2]提出嵌入式零树小波算法, 它有效利用了小波系数的特性, 实现了图像的分级编码。以离散余弦变换(Discrete Cosine Transform, DCT)为核心的 JPEG 算法包含有损压缩和无损压缩 2 种模式, 目前网上 80% 的图像都采用 JPEG 压缩标准, 但该算法在图像的细节信息方面损失较大。JPEG2000 算法^[3]采用以离散小波变换(Discrete Wavelet Transform, DWT)为主的多分辨率编码方式, 能实现感兴趣区域压缩, 因此, 在医学图像压缩中被应用广泛, 但该方法各子带间采用相同门限量化, 不能充分利用人眼视觉特性, 限制了图像压缩比的提高。

动物视觉系统的稀疏性、同步性等特点使其能用最少的神经元表示大量自然环境信息。本文借鉴动物视觉系统的优点, 采用稀疏编码方法提取数字文献图像的基函数及其响应系数, 采用同步振荡方法获取最佳响应系数, 并利用少量响应较强的基函数进行图像重构。

2 视觉系统的稀疏性与同步性

2.1 稀疏性

为了表示外界所有输入模式, 视觉系统所需的神经元数量极大, 而实际的视觉系统能用有限的神经元表示出外界的所有输入模式, 当视觉神经系统接收到某一幅自然图像时, 大部分神经元对该幅图像的响应很弱甚至为 0, 只有很少的

神经元会有较强响应。Field D J 将这种特性称为稀疏性^[4]。

神经稀疏编码理论认为自然图像可以看成多个基函数的线性组合:

$$I(x, y) = \sum_i a_i \phi_i(x, y) \quad (1)$$

其中, $I(x, y)$ 为自然图像分割后的图像块; $\phi_i(x, y)$ 为第 i 个基函数; 系数 a_i 为第 i 个神经元的响应。

基函数集合 $\phi(x, y) = \{\phi_i(x, y)\}$ 可以经过多幅自然图像的训练估计出来。由于 a_i 满足稀疏性, 因此每幅参与训练的自然图像都可以通过该基函数集合的稀疏表示得到, 即一幅自然图像可以仅由基函数集合中的少数几个基函数表示。

2.2 同步性

在视觉系统的初级阶段, 大脑中表示不同特征的细胞是通过其活动上的时间相关联系在一起, 神经振荡是实现时间相关的一个自然的方式, 用同步的振荡子群表示同一个目标, 即视觉系统的同步性^[5]。视觉信息处理过程的同步性反映了视觉皮层神经元之间存在的局域联结关系, 即神经元与神经元之间既有激励又有抑制。

通过对哺乳动物大脑视觉皮层中的同步振荡现象进行模拟, Eckhorn R 等人提出了脉冲耦合神经网络(Pulse-Coupled Neural Networks, PCNN), 该模型具有变阈值、强自适应、同步脉冲发放等特性, 更贴近生物实际神经网络。

基金项目: 国家自然科学基金资助项目(60841004)

作者简介: 师黎(1964 -), 女, 教授、博士生导师, 主研方向: 数字图像处理, 智能控制, 故障诊断与容错分析; 王端, 硕士研究生; 王治忠, 博士研究生

收稿日期: 2009-08-07 **E-mail:** wangduanzu@gmail.com

3 基于稀疏性与同步性的数字文献图像压缩算法

3.1 数据采集及预处理

为避免影响像素之间的统计特性,本文采用 $M \times M$ 的滑动空间子窗口对每幅图像进行随机采样,其位置随机设定,这样抽取的子图像块在原始图像中的位置相互有重叠,可以相互弥补子图像块边缘像素间的统计特性。

统一中心,使每个 $I(x, y)$ 和 $\phi(x, y)$ 都成为零均值矢量,令

$$V = I(x, y) - E(I(x, y)) \quad (2)$$

对数据进行白化处理,去除各观测信号间的相关性,加快收敛速度,减少稳态误差,令

$$X = D^{-1/2} E^{-1} V \quad (3)$$

使得

$$E(XX^T) = I \quad (4)$$

成立。其中, X 是经过白化处理过的数据; D 和 E 分别是 V 的协方差矩阵 $R_v = E(VV^T)$ 的特征值矩阵和特征向量矩阵。

3.2 基函数的训练

目前,已有很多证据说明自然图像的高阶统计特性是服从非高斯分布的,本文采用稀疏编码理论来提取数字文献图像的基函数,稀疏编码模型的目标是寻找一个最佳的基函数集合,使由该基函数集合稀疏表示的自然图像最大可能的接近非高斯特性。

定义一个目标函数来衡量基函数的非高斯特性,常用的衡量非高斯特性的标准是负熵:

$$J(y) \approx \{E[G(y_{gauss})] - E[G(y)]\}^2 \quad (5)$$

其中, y_{gauss} 是与 y 具有相同方差的满足高斯分布的变量; G 是非二次函数,本文选择

$$G(u) = \frac{1}{a} \lg \cosh(au) \quad (6)$$

其中, a 取值为 1。

基函数的提取过程可以归结为一个最优化的过程,即寻找使 $\phi(x, y) = WX$ 的负熵 $J(S)$ 最大的转换矩阵 W 。应用梯度法计算转换矩阵 W , 即

$$W_{n+1} = W_n + \eta \frac{\partial J(W_n X)}{\partial W_n} \quad (7)$$

其中, X 为经过预处理之后的训练样本; W_n , W_{n+1} 分别是在 n 和 $n+1$ 时刻得到的转换矩阵; η 为学习速率。

将转换矩阵归一化, 令

$$W_{n+1} = \frac{W_n}{\|W_{n+1}\|} \quad (8)$$

当 W_{n+1} 收敛时, 训练结束。

3.3 响应系数的筛选及图像重构

在得到的稀疏编码模型 $I(x, y) = \sum_i a_i \phi(x, y)$ 中, 响应系数 a_i 表示基函数 $\phi(x, y)$ 的响应程度, 本文采用 PCNN 方法对响应系数进行筛选。

PCNN 的内部活动项 $U_{ij}[n]$ 由反馈输入 $F_{ij}[n]$ 和线性连接输入 $L_{ij}[n]$ 共同构成, 令

$$L_{ij}[n] = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \omega V_{ij} \quad (9)$$

$$F_{ij}[n] = I_{ij} \quad (10)$$

$$U_{ij}[n] = F_{ij}[n](1 + \beta L_{ij}[n]) \quad (11)$$

在式(9)中, V_{ij} 为响应系数的一个 $M \times M$ 滑动子窗口; ω 为

响应系数的加权矩阵。在式(10)中, I_{ij} 为外部输入刺激信号, 对应 V_{ij} 范围内中心点的响应系数。在式(11)中, β 为连接强度系数。

动态阈值 $\theta_{ij}[n]$ 随时间指数衰减, 即

$$\theta_{ij}[n] = \exp(-\alpha_\theta) \theta_{ij}[n-1] \quad (12)$$

其中, α_θ 为动态门限的衰减时间常数。

当 $U_{ij}[n]$ 大于动态门限 $\theta_{ij}[n-1]$ 时, PCNN 产生输出时序脉冲序列 $Y_{ij}[n]$, 被激活的神经元置 1, 对于没有被选中的基函数, 将其响应系数置 0, 即

$$Y_{ij}[n] = \begin{cases} 1, & U_{ij}[n] > \theta_{ij}[n] \\ 0, & U_{ij}[n] \leq \theta_{ij}[n] \end{cases} \quad (13)$$

由动物视觉系统的生理特性可知, 通过 PCNN 筛选出的少量的基函数完全可以重构出原始图像。本文采用图像压缩比 CR 和峰值信噪比 $PSNR$ 对图像恢复效果进行评价。令 $f(x, y)$ 表示一幅 $M \times N$ 的原始图像, $\hat{f}(x, y)$ 表示压缩后的重构图像, 则

$$CR = P_{f(x,y)} / P_{\hat{f}(x,y)} \quad (14)$$

$$PSNR = 10 \lg \left[\frac{255^2 MN}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(x, y) - \hat{f}(x, y)]^2} \right] \quad (15)$$

其中, $P_{f(x,y)}$ 和 $P_{\hat{f}(x,y)}$ 分别代表原始图像和重构图像的大小。

本文算法流程如图 1 所示。

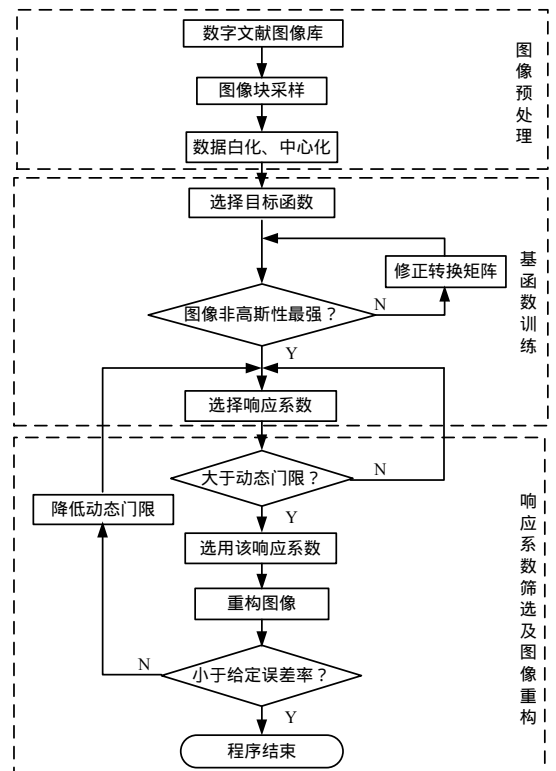


图 1 本文算法流程

4 实验结果及分析

本文选用 Eckhart Tolle 著的《A New Earth》中 50 页上的 512×720 大小的数字文献图像作为样本图像进行训练, 在每幅图像中选取 1 000 个 16×16 像素的图像块作为样本来训练基函数。对于基函数的响应系数, PCNN 采用 3×3 的子窗口进行逐个筛选, 窗口中心点作为 PCNN 外部输入刺激信号,

对周围 8 个神经元产生激励或抑制。经过训练得到的基函数如图 2 所示。

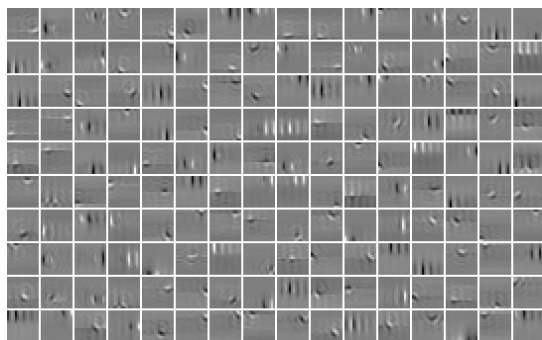


图 2 由 50 幅文字图像训练得到的基函数

为了说明上述算法的压缩效果，本文分别采用上述算法和独立分量分析 (Independent Component Analysis, ICA) 方法进行比较。本文对书中任选的 30 页文字图像进行压缩，其中一页的压缩效果如图 3 所示。



(a)原始图像



(b)ICA 算法重构的压缩图像



(c)本文算法重构的压缩图像

图 3 实验结果对比

从图 3 可以看出，由于本文应用了 PCNN 同步振荡的特性，压缩后的图像更细腻。与 ICA 方法相比，本文算法能在利用尽可能少的基函数的前提下恢复出更清晰、压缩比更高的图像。

本文对 30 幅数字文献图像压缩后的统计结果如表 1 所示，可以看出，本文方法可以更有效地对图像进行压缩处理，

本文算法处理后的图片在保持分辨率不变的情况下，具有更高压缩比，且图像的峰值信噪比基本不变。

表 1 30 幅数字文献图像压缩后的统计结果

图像	图像大小/KB	图像压缩比	峰值信噪比
原始图像	1 523.77	-	-
ICA 方法所得	158.16	9.634	7.45
本文方法所得	113.94	13.373	7.69

为了进一步验证本文算法的优越性，本文分别随机抽取《A New Earth》中的 30 幅样本图像和 2 组样本外图像进行压缩和图像重构，其对比结果见表 2。

表 2 样本和样本外数字文献图像压缩后的统计结果

图像	原始图像大小/KB	压缩后图像大小/KB	图像压缩比	峰值信噪比
样本图像	1 523.77	113.94	13.373	7.69
样本外图像组 1	1 545.48	111.07	13.914	6.97
样本外图像组 2	1 520.89	113.65	13.382	7.62

由表 2 可以看出，2 组样本外图像由于图像内容的差异，因此压缩比和峰值信噪比略有不同，但随着样本量的增加两者能基本趋于一致。可见，本文算法可以利用少量样本训练的基函数恢复出样本之外的大量图像，且压缩比和峰值信噪比与样本图像基本相同，这是本文算法的优点之一。

5 结束语

与原有同类算法相比，本文算法的处理效果有了很大改进，但它的压缩速度比 ICA 略慢，因此，需要对其进行进一步研究并改善。

参考文献

- [1] Taşkın C, Sarikoz S K. An Overview of Image Compression Approaches[C]//Proc. of the 3rd International Conference on Digital Telecommunications. Bucharest, Romania: [s. n.], 2008.
- [2] Shapiro J M. Embedded Image Coding Using Zerotrees of Wavelet Coefficients[J]. Signal Processing, 1993, 41(12): 3445-3462.
- [3] 尹平, 汪宇飞, 胡迎新. JPEG2000 格式图像压缩算法的研究[J]. 计算机工程, 2007, 33(5): 205-207.
- [4] Nguyen T V, Patra J C. A Simple ICA-based Digital Image Watermarking Scheme[J]. Digital Signal Processing, 2008, 18(5): 762-776.
- [5] Henrik B, Roland O, Thomas L. Automatic Design of Pulse Coupled Neurons for Image Segmentation[J]. Neurocomputing, 2008, 71(10-12): 1980-1993.

编辑 陈 晖

(上接第 206 页)

参考文献

- [1] Barnsley M F, Sloan A D. A Better Way to Compress Images[J]. Byte Magazine, 1988, 13(1): 215-223.
- [2] Jacquin A E. A Fractal Theory of Iterated Markov Operators with Applications to Digital Image Coding[D]. Atlanta, GA, USA: Georgia Institute of Technology, 1989.
- [3] Jacquin A E. Image Coding Based on a Fractal Theory of Iterated Contractive Image Transformations[J]. IEEE Transactions on Image

Processing, 1992, 1(1): 18-30.

- [4] Hartenstein H, Saupe D. Lossless Acceleration of Fractal Image Encoding via the Fast Fourier Transform[J]. Signal Processing: Image Communication, 2000, 16(4): 383-394.
- [5] 何传江, 李高平. 分形图像编码的改进算法[J]. 计算机仿真, 2004, 21(8): 66-69.
- [6] 林洪文, 杨绍清. 基于分形图像编码的海空目标检测方法研究[J]. 计算机工程, 2007, 33(11): 199-200.

编辑 顾姣健