

# 基于神经网络的几种新的数据压缩方案

刘国文, 王志良, 解 仑

( 北京科技大学信息工程学院, 北京 100083 )

**摘 要** 本文首先采用非统计的方法给出了基于神经网络的无损数据压缩新方案。基于神经网络的无损数据压缩新方案的特点是无损的、能压缩一些已用熵编码压缩过的数据, 且在一些情况下压缩比接近 2.58。然后把该新方案嵌入 Huffman 编码系统, 获改进的无损数据压缩方法, 提高了 Huffman 编码系统的压缩比。最后把基于神经网络的无损数据压缩新方案嵌入一小波图像压缩系统得基于小波与神经网络的图像压缩方案, 部分解决了小波图像压缩系统中‘高压缩比’与‘高保真’不能协调统一的矛盾。

**关键词** 数据压缩; 神经网络; 小波; 熵编码

中图分类号: TP311.131 文献标识码: A 文章编号: 1003-7241( 2003 )07-0018-04

## New Data Compressions Methods Based on Neural Network

LIU Guo-wen , WANG Zhi-liang , XIE Lun

( University of Science and Technology , Beijing 100083 , China )

**Abstract** : First , the new scheme of a lossless data compression based on neural network is given by using non-statistical method . It is lossless and can compress some data which are compressed by entropy coding . Sometime its compression ratio is about 2.58 . Second , It is embedded in Huffman coding to construct the improved lossless data compression that heightens the compression ratio of Huffman coding . Final , it is embedded in an image compression system based on wavelet to obtain the image compression based on wavelet and neural network . The image compression partly intercedes the conflict of ‘ high compression ratio ’ and ‘ high lossless degree ’ in image compression systems based on wavelet .

**Key words** : Data compression ; Neural network ; Wavelet ; Entropy coding

### 1 引言

尽管光纤网和光盘设备等宽带信道和大容量存储介质的采用, 使信道带宽及存储量变越来越大, 但信息时代带来了“信息爆炸”, 数据量的增加远远快于带宽和存储量的增长, 因此数据压缩的作用及其社会效益、经济效益将越来越明显。反之, 如果不进行数据压缩, 则无论传输或存储都很难实用化。数据压缩的好处就在于:

①较快地传输各种信源(降低信道占有费用)——时间域的压缩;

②在现有通信干线上开通更多的并行业务(如电视、传真、电话、可视图文等)——频率域的压缩;

③降低发射机功率——能量域的压缩;

④紧缩数据存储容量(降低存储费用)——空间域的压缩。

在计算机系统文件、医疗、保密通讯对数据质量要求苛刻的场合中, 人们开发了哈夫曼编码、算术编码、字典编码、游程编码、预测编码等是目前流行的无损压缩方法<sup>[1-4]</sup>, 但它们都是基于‘统计’的方法, 只消除概率分布冗余(注:许多原始数据经它们压缩之后仍有很高的冗余度, 有很大的压缩‘空间’)。因而对

随机的 0 与 1 的字符串的压缩效果不到 2:1 ,且对已用一编码压缩过的数据几乎无法再用另一压缩方法压缩。

人们要问 ,是否有改进的无损编码方案 ,它能提高以上熵编码的压缩比 ,而且能压缩已用一熵编码压缩过的数据呢 ?

就不同数据特征 ,人们也提出了许多有用的基于小波的有损图像数据压缩方法<sup>[5-6]</sup> ,然而在它们当中存在‘高压缩比’与‘高保真’不能协调统一的矛盾。

是否能改进这些压缩方案 ,使它们能保持原有的保真度 ,而提高压缩比 ,使‘高压缩比’与‘高保真’协调统一呢 ?

本文首先采用非“统计”的方法给出了“基于神经网络的无损数据压缩新方案”。“基于神经网络的无损数据压缩新方案”的特点是无损的、能压缩一些已用熵编码压缩过的数据 ,且在一般情况下压缩比接近 2.58。然后把“基于神经网络的无损数据压缩新方案”嵌入 Huffman 编码系统 ,获“改进的无损数据压缩方法” ,提高了 Huffman 编码系统的压缩比。最后把“基于神经网络的无损数据压缩新方案”嵌入一小波图像编码系统得“基于小波与神经网络的图像压缩方案” ,部分解决了图像压缩中‘高压缩比’与‘高保真’不能协调统一的矛盾。

## 2 基于神经网络的无损数据压缩新方案

### A. 映射 M 及 BP 神经网络逼近

计算机中所有数据都可最终表示为 0 与 1 的字符串 ,因此不失一般性我们假定要压缩的数据是 0 与 1 的字符长串。由于压缩太短的数据不会有好的压缩效果且为了考虑问题方便 ,因此本文只直接考察长度为  $2 \times 1024 \times 11 = 22528$  的 0 与 1 的长串。一般的长串可切分或补充为  $2 \times 1024 \times 11$  长度的 0 与 1 的长串再予以压缩。

设有待压缩的长串  $\gamma = c_1 \dots c_{22528}$  ,  $c_j = 0$  or  $1$  ,其中  $\gamma$  的长度为 22528。下面类似<sup>[7]</sup>定义 11 维单位超立方体顶点和中心的集合为原像集 ,像落在 11 维单位超立方体表面和中心的映射  $M_0$ 。

映射  $M$  :设  $A = \{ (x_1, \dots, x_i, \dots, x_{11}) \mid x_i = 0, 1 \}$  把  $(x_1, \dots, x_i, \dots, x_{11})$  按二进制数  $x_1 x_2 \dots x_i \dots x_{11}$  的大小倒序排列 ,即排为有序点  $a_1 = (1, 1, \dots, 1, 1)$  ,  $a_2 = (1, \dots, 1, 0)$  ,  $a_3 = (1, \dots, 1, 0, 1)$  ,  $\dots$  ,  $a_{2047} = (0, \dots, 0, 1)$  ,  $a_{2048} = (0, \dots, 0, 0)$ 。假定待压缩的 0 与 1 的字符串  $\gamma = c_1 \dots c_{22528}$  ,  $c_j = 0$  or  $1$  ,其中  $\gamma$  的长度为 2048 ,记  $B = \{b_j \mid b_j = (c_{11(j-1)+1}, \dots, c_{11 \times j}) , 1 \leq j \leq 2048\}$ 。注  $b_j$  中的  $c_i$  要转换为 0 与

1 数的形式。定义  $M$  为  $A$  到  $B$  的映射

$$M : A \rightarrow B$$

$$M : a_i \rightarrow b_j , i = 1, \dots, 2048 , a_i \in A , b_j \in B$$

即  $b_j = M(a_i) , i = 1, \dots, 2048$ 。

易见 ,顺序输入固定向量  $a_i \in A , i = 1, \dots, 2048$  ,可顺序产生  $b_j \in B , j = 1, \dots, 2048$  ,从而产生  $\gamma = c_1 \dots c_{22528}$  ,  $c_j = 0$  or  $1$ 。因此  $\gamma$  确实是产生  $\gamma = c_1 \dots c_{22528}$  ,  $c_j = 0$  or  $1$  的虚拟信源。是 0 与 1 的字符串的通用虚拟信源。

神经网络有神奇的非线性映射能力 ,1989 年 ,Cybenko 证明了带隐含层的神经网络可实现任意非线性函数关系的映射<sup>[8]</sup>。这些惊人的结果启示我们用适当的神经网络去模拟产生 0、1 字符串的虚拟信源  $M_0$ 。

大量实验结果也表明 ,对于固定的  $n$  维距离空间到  $n$  维距离空间有界连续映射  $G$  ,可找到输入层  $n$  个神经元 ,隐含层有适当个、输出层有  $n$  个神经元及适当非线性输出函数的 3 层 BP 网络充分逼近  $G_0$ 。

[9] 给出了选择隐函数神经元个数  $n_2$  的经验方法 :

$$n_2 = \lceil \log_2 n \rceil + k , 1 \leq k \leq n \quad (n \text{ 为输入层神经元个数} ) \quad (k \text{ 为经验参数} )$$

或

$$n_2 = \sqrt{n + m} + a \quad (a = 1, \dots, 11 , n \text{ 为输入层神经元个数} , m \text{ 为输出层神经元个数} )$$

由于虚拟信源  $M$  是有限点集  $A$  到有限点集  $B$  的有界映射 ,它一定可以扩充为球  $\bar{A} = \{ (x_1, \dots, x_i, \dots, x_{11}) \mid 0 \leq |x_i| \leq 1 , i = 1, \dots, 11 \}$  上的有界连续映射  $\bar{M}$ 。由前面引用结果可知 ,存在 3 层 BP 网络充分逼近  $\bar{M}$  ,自然该网络充分逼近  $M_0$ 。(  $G$  为有界连续映射 ,  $M$  为虚拟信源 )

现在我们根据实践经验和 [9] 的建议 ,考虑到有 2048 个样本要处理 ,选用 3 层的有  $550 + 47 = 597$  个自由度的 BP 模型来逼近  $M_0$ 。选定神经网络某一满足条件的状态作为虚拟信源  $M$  的模型 ,取模型参数编码字符串  $\gamma = c_1 \dots c_{22528}$  ,  $c_j = 0$  or  $1$  ,以便压缩字符串长串  $\gamma$ 。本文用 3 层 47 个神经元 (输入、输出层各 11 个 ,隐含层 25 个)和各层输出函数为 Sigmoid 函数 :

$$\sigma_j(u_j) = \frac{1}{1 + e^{-u_j}}$$

的 BP 模型来作为特定映射  $M$  的逼近映射。其中各层神经元的联接权 (550 个) 矩阵分别为

$$w_{ij}^{(k)} , i, j = 1, \dots, k , k = 1, 2 \text{ 和 } 47 \text{ 个阈值 } w_j^{(k)} , j = 1, \dots, k = 1, 2 ,$$

3. (K 为经验参数)

# B. 基于神经网络的无损数据压缩原理

本文基于神经网络的无损数据压缩具体原理由压缩原理和解压缩原理两部分组成。

压缩原理 :如前设定待压缩的长串  $\gamma = c_1 \dots c_{22528}$ ,  $c_j = 0$  or  $1$ , 其中  $\gamma$  的长度为 22528,  $a_k, b_k, k = 1 \dots 2048$  及映射  $M_{\cdot}(i)$  用顺序样本  $(a_k, b_k)$  训练所选 BP 网络,  $k = 1 \dots 2048$ , 且设定  $|w_{ij}^{(k)}| < 100, |w_j^{(k)}| < 100$ , 训练时所有权值  $w_{ij}^{(k)}$  和阈值  $w_j^{(k)}$  都精确到小数点后 2 位 (i i) 选定 BP 网络的一状态 (称为稳态) :该状态下输出  $b_k'$  与期望输出  $b_k$  的分量差的绝对值小于 0.45, 即  $|b_{ki}' - b_{ki}| < 0.45, i = 1 \dots 11$  (i i i) 在稳态时, 顺序取 550 个神经元联接权值  $w_{ij}^{(k)}, i, j = 1 \dots k = 1, 2$  和 47 个阈值  $w_j^{(k)}, j = 1 \dots k = 1, 2, 3$ , 作为字符串  $\gamma$  的编码 (iv) 顺序把 550 个神经元联接权值和 47 个阈值依次 2 进制编码, 方案是用 2 个字节 16 位为每个  $w_{ij}^{(k)}, w_j^{(k)}$  编码, 其中 1 位是符号位, 1 位是小数点记号位, 7 位用于小数点后 2 位编码, 7 位用于小数点前 2 位编码, 此编码可确保当  $|w_{ij}^{(k)}| < 100, |w_j^{(k)}| < 100, w_{ij}^{(k)}, w_j^{(k)}$  精确到小数点后 2 位时, 整个压缩方法是无损压缩方案, 严格数学证明见 [7] 中推导, 这里从略。

由上述原理可知, 若 (i) (i i) 能顺利进行, 则该压缩方法的压缩比为

$$\frac{2 \times 1024 \times 11}{550 \text{ 个权值的二进制码长度} + 47 \text{ 个阈值的二进制码长度}} = \frac{1024 \times 22}{597 \times 16} = 2.58$$

解压缩原理 :BP 取整函数解压缩模型是由所选 BP 网络构成的映射复合 (或外套) 取整函数而成。 (i) 恢复稳态时的  $w_{ij}^{(k)}, i, j = 1 \dots k = 1, 2$  值和稳态时的阈值  $w_j^{(k)}, j = 1 \dots k = 1, 2, 3$ , 恢复 BP 模型稳定状态 (i i) 把恢复的网络视为一固定映射  $Y^1$ , 这时  $Y^1$  是  $M$  的近似映射, 且逼近度量值小于 0.45 (i i i) 把  $Y^1$  与取整函数  $[b_{ki}' + 0.5], i = 1 \dots 11$ , 复合得  $[b_{ki}'(X) + 0.5], i = 1 \dots 11, X = (x_1 \dots x_i \dots x_{11})$  (i v) 把  $a_k, k = 1 \dots 2048$ , 顺序输入复合函数  $[b_{ki}'(X) + 0.5], i = 1 \dots 11$ , 中, 顺序获得  $b_k, k = 1, \dots 2048$  (v) 由  $b_k, k = 1 \dots 2048$  完全恢复长串  $\gamma = c_1 \dots c_{22528}$ 。

# C. 实验结果

实验表明, 对很多  $\gamma = c_1 \dots c_{22528}$ , 相应的 BP 网络的训练很快达到要求, 说明该压缩方法有效可行。而对有的  $\gamma = c^1 \dots c_{22529}$ , 相应的 BP 网络训练很长时间不收敛, 说明 BP 网络训练算法有待改进和计算机的运行速度有待提高。为做到实时或压缩顺利

进行, 我们采用限制神经网络训练时间的策略。对在限制时间内神经网络训练达不到稳态的情况, 引入中断机制, 相应的串  $\gamma = c_1 \dots c_{22528}$  不予以压缩。

20 次实验结果表明, 有 6 次压缩比是接近 2.58, 其它次压缩比是 1 到 2.58 之间, 该压缩方法的压缩比大约为 2.58。由于它的思想方法与熵编码压缩方法的思想不同, 考虑问题的角度不同, 且 3 层的有 597 个自由度的 BP 模型有自学能力, 因此有些熵编码没有找到的规律; 基于神经网络的无损数据压缩新方法“能认识, 从而它能有效压缩已用熵编码压缩过的数据, 这与我们的实验相吻合。

# 3 一种改进数据压缩方案及实验结果

根据以上所获“基于神经网络的无损数据压缩新方法”能压缩已用熵编码压缩过的数据的特点, 我们把“基于神经网络的无损数据压缩新方案”嵌入 Huffman 编码系统, 获“改进的无损数据压缩方案”, 把“基于神经网络的无损数据压缩新方案”嵌入一小波图像编码系统得“基于小波与神经网络的图像压缩方案”。

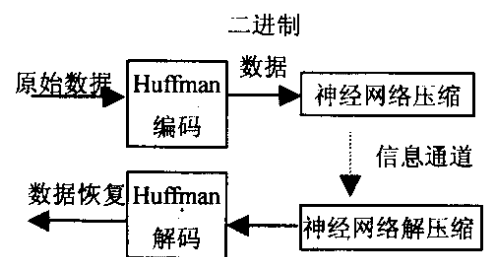


图 1 改进的无损数据压缩方案

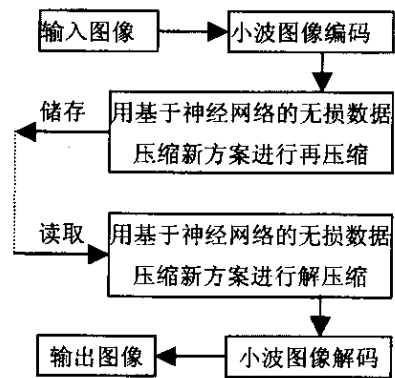


图 2 基于小波与神经网络的图像压缩方案

实验表明, “改进的无损数据压缩方法”和“基于小波与神经网络的图像压缩方案”分别提高了 Huffman 编码系统和原小波图像编码系统的压缩比。但以上压缩方法在实时性方面有待改进。

## 4 参考文献：

- [1] D. A. HUFFMAN, A method for the construction of minimum - redundancy codes[ C ]. Proceeding of the IRE 40 ,1952 :1098 - 1101
- [2] I. H. WITTEN ,R. M. NEAL and J. G. CLEARY ,Arithmetic coding for data compression ,CACM[ J ]. 1987 ,30( 6 ) :520 - 541
- [3] ZIV and A LEMPEL ,Compression of individual sequences via variable rate coding[ J ]. IEEE Transaction on Information Theory , 1978 ,24 :530 - 536
- [4] T. A. WELCH ,A technique for high - performance data compression[ J ]. IEEE computer ,1984 ,17( 6 ) :8 - 19
- [5] 吴乐南.数据压缩[ M ].电子工业出版社 ,东南大学出版社 2000

- [6] JYH - LIANG JOU ,et al ,Image coding from wavelet transform maxima and projection onto convex sets[ J ]. Image and Vision Computing ,1999 ,17 :295 - 298
- [7] 杨国为 ,涂序彦 ,庞杰.基于虚拟信源的无损数据压缩方法研究[ J ].电子学报 ,2003 ,31( 5 ) :728 - 731
- [8] G. . CYBENKO ,Continuous value neural networks with two hidden layers are sufficient[ C ]. Math. , Contr. , Sign. & Sys. ,1989 , ( 2 ) 303 - 314
- [9] 靳蕃.神经计算智能基础[ M ].成都 :西南交通大学出版社 2001

作者简介 刘国文( 1958 - ) ,北京科技大学博士研究生 ,主要从事控制工程与控制理论方面的研究。

( 上接第 7 页 )

$$\det H(V(x, c) + P) \Big|_{c \neq 0} \quad (12)$$

即  $V(x, c) + P$  也在  $c$  有一个非退化的临界点。由此 ,我们就可以说脆性势函数  $V(x, c)$  在临界点  $c$  是结构稳定的。

把退化的脆性域  $C_T$  中的临界点利用上述提供的方法 ,将其进一步分类 ,分成稳定的和不稳定的两类临界点。至此 ,已经将所有临界点分类完毕。

研究的范围确定在不稳定的临界点集中 ,临界点集可以组成系统处于稳定与不稳定状态的边界线。当系统处于由不稳定临界点集组成的区域时 ,系统的脆性被激发的可能性是很大的 ,但非是一定的。所以 ,在这个区域里 ,可以将其分成几个脆性被激发的可能性大小的等级 ,这样可以将其与定义的脆性度联系起来 ,来判断系统在哪容易发生崩溃 ,也是系统的脆性最易被激发的地方。这样 ,可以将不稳定的临界点集中的使系统不能正常工作的临界点集记为  $C_{T_{21}}$  ,那么就可以有针对性的采取一定的控制措施来防止系统运动到它的不稳定的工作域  $C_{T_{21}}$  。至此 ,将一个子系统的脆性分析进行完了。那么对于整个复杂系统来说 ,有针对性的进行对子系统的脆性分析 ,然后根据其与其它子系统的关系 ,研究整个复杂系统的脆性问题。

系统的脆性是与系统的非线性联系在一起的。当一个子系统受到攻击而崩溃时 ,由于子系统之间的联系呈现非线性特征 ,子系统之间呈现非对称性 ,崩溃的子系统的脆性则同化了其它子系统特性。脆性在各个子系统中成为支配着其它特性的变化的主要矛盾。此时 ,子系统之间的协同效应与放大效应便呈现出来。协同效应促使各个子系统都以脆性为主的协同一致的运动。放大效应使被脆性所同化的子系统的数量越来越多 ,最终 ,

整个复杂巨系统都呈现出以脆性为主要矛盾的运行状态。

## 7 结论

本文在对复杂系统所存在的问题——脆性问题进行描述的同时 ,采用了脆性势函数作为分析系统的脆性基础 ,在此基础上 ,对其临界点进行了分类 ,从而将系统的脆性分析完了。本文中所涉及到的理论在实际应用中是具有应用价值的 ,比如说在交通流中、网络安全中等等都可以应用到本文中所涉及到的问题与方法。

## 8 参考文献：

- [1] 颜泽贤 陈忠 胡皓编.复杂系统演化论[ M ].北京 :人民出版社 ,1993 年 12 月。
- [2] 凌复华.突变理论及其应用[ M ].上海 :上海交通大学出版社 ,1987 年 12 月。
- [3] 项静恬 史久恩.非线性系统中数据处理的统计方法[ M ].北京 :科学出版社 ,1997 年 8 月
- [4] 于景元.钱学森关于开放的复杂巨系统的研究[ J ].系统工程理论与实践 ,1992 年 9 月 8 - 12
- [5] 桑博得 ,著 ,凌复华 ,译.突变理论入门[ M ].上海 :上海科学技术文献出版社 ,1983 年 7 月
- [6] 曾廷全.复杂系统的一般数学框架( I ) [ J ].应用数学和力学 ,1993 ,14( 4 ) :367 - 374 页

作者简介 郭健 1977 - 女 哈尔滨工程大学控制理论与控制工程专业 ,博士研究生 ,主要研究方向 ,复杂系统的脆性研究。