

在小波变换域内应用BP网络 进行图象压缩的方法研究

武 斌

(海军电子工程学院, 南京 210029)

李象霖

(中国科学技术大学研究生院, 北京 100039)

摘 要 主要讨论了经过小波变换后较难压缩的低频子带系数特性, 并据此利用基于BP算法的多层前向神经网络高度非线性映射能力, 构造了一个非线性预测器, 有效地去除了低频子带系数间的冗余度, 在较高的保真条件下获得了较满意的压缩比。

关键词 小波变换 非线性预测 BP网络

0 引 言

当前, 我们正处在一个高速发展的信息时代, 在众多的信息形式中, 图象信息最具有直观性和生动性, 因而越来越成为人们需求的主要信息形式。然而由于图象信息的数据量太大, 直接制约着图象信息的存储和传送。这就出现了视频图象作数字存储时占用空间特别大, 作数字传输时占用的信道频带又非常宽的问题。所以, 为了解决上述问题, 作为上述应用的关键技术——高效率图象数据压缩技术就倍受人们的关注。

从另一方面来看, 作为图象数据的压缩编码是完全可行的。这主要因为: 第一、规则的图象必然存在相关性, 亦即存在较大的冗余度; 第二、用户的视觉效果可以允许图象有一定的失真而不影响人的主观感觉; 第三、根据信道的分辨率要求, 适当降低原始图象的分辨率对于传输图象的分辨率影响不大; 第四、在不同的场合下, 用户有时并不是对原始图象的所有信息都感兴趣。

1989年S. Mallat将小波变换用于信号处理, 提出了多分辨率分析的概念, 给出了图象信号分解为不同频率通道的算法, 开创了小波变换在图象处理中的应用。借助小波变换, 图象信号可以被分解为许

多具有不同空间分辨率、频率特性和方向特性的子图象信号, 实现了低频长时特征和高频短时特征的同时处理, 有效地克服了付立叶分析在处理非平稳图象信号时的局限性, 同时更适合于人类的视觉特性, 为人们更有效地从图象的各种数据特性和视觉特性角度去压缩图象数据开辟了新的思路。

0.1 问题的提出

首先给出本人采用Daubechies2小波基 $h_2(n) = \{0.483, 0.837, 0.224, -0.129\}$ 对标准图象Lena(512×512×8bit)进行两级小波变换的结果, 原图和小波变换图如图1所示。表1给出了两级小波变换后各子图象系数的统计特性, 表2给出了两级小波变换后子图象能量分布关系情况。

从图1、表1和表2可以看出, 小波变换实现了能量的再分配, 绝大部分能量集中在二级低频子带 LL_2 中, 对它压缩的程度和信息保持情况将直接影响到图象的总体压缩比和恢复图象的保真度, 而由于低频子带系数动态变化非常大, 如果直接对低频子带系数进行较细致的均匀量化, 那么每个系数所需的比特数将是十分大的, 这样一来, 非但不能压缩数据量, 反而会比原图同样大小子块所需的数据量要大; 如果降低对低频子带系数的量化级数, 这样无

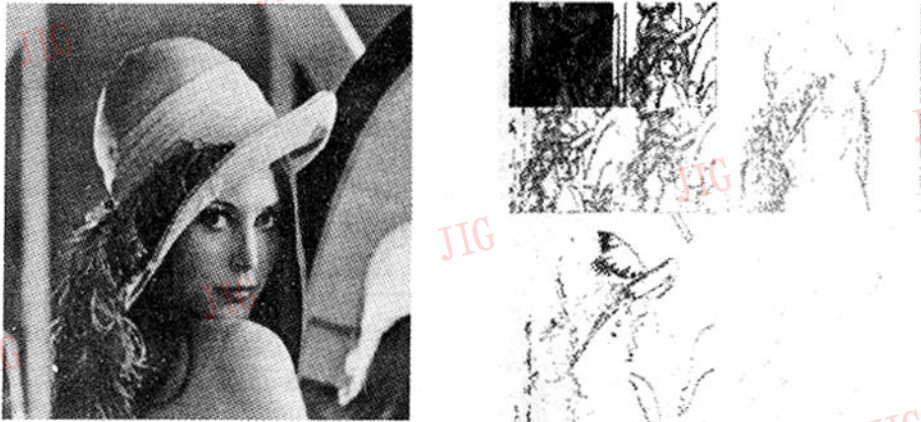


图1 Lena(512×512×8bit)及其二级小波变换

表1 图象 Lena(512×512×8bit)两级小波分解子图象数据统计特性

子图象	LL ₂	LH ₂	HL ₂	HH ₂	LH ₁	HL ₁	HH ₁
最大值	972.184	203.104	190.851	128.759	112.722	51.746	44.623
最小值	80.316	-187.257	-203.833	-96.334	-96.595	-60.118	-36.148
均值	496.264	-0.263	-0.132	-0.107	-0.115	-0.033	-0.007
均方差	35 109.99	692.445	254.474	133.627	71.027	29.558	11.629

表2 Lena(512×512×8bit)两级小波分解子图象能量分布关系情况

子带图象	子带平均能量	所占的百分比(%)	E_L/E_H (DB)
LL ₂	281328.5	99.48	60.87
HL ₂	254.5	0.09	52.2
LH ₂	692.5	0.25	66.47
HH ₂	133.6	0.047	85.4
HL ₁	15.13	0.021	71.95
LH ₁	71.04	0.101	92.86
HH ₁	6.4	0.009	

异于牺牲图象质量,结果只能获取很小的压缩比,可谓得不偿失,所以这部分往往是很难压缩的。

1 利用BP网络对二级低频子带进行压缩编码

实际上,对于代表低频分量的二级子块 LL₂,由

于它是去除原图象中高频成份而得到的,所以它的系数无论在行的方向还是列的方向都具有很强的相关性,图2给出了图象 Lena(512×512×8bit)二级低频子带 LL₂ 任选相邻两行的系数值分布情况。

低频子带 LL₂ 系数的这种内在相关性实际上是一种信息冗余。试想如果最大限度地去除这种冗余度,系数的动态范围势必减小,相应编码所需的比特位也将减少,由此便可在最小损失图象质量的情况下提高数据的压缩比。这种思想正是预测编码的出发点。经典的线性预测编码(DPCM)方法根据相邻象素的已知值作线性预测,输出预测误差图象;一般来说预测误差图象的方差较之原始图象数据的方差要小,也即误差图象信息间的相关性比起原始图象的相关性要弱,从而达到去冗余、压缩数据的目的。

对于线性预测编码,我们可以将它看成一个“时不变的”数字处理系统,但是图象信息是高度非线性的;如果能把“时不变的”DPCM系统改进为“时变的”DPCM系统,也即自适应DPCM系统,使它更加匹配于图象的局部内容,则它将具有更大的灵活性,并将获得进一步的压缩效果,提高恢复图象的主观

质量。

从理论上讲,预测编码问题可以归纳为映射和优化的问题,而人工神经网络,如基于 BP 算法的前向神经网络(forward neural network based Back-Propagation,简称 BP 网络)实现的恰好是一种由输入空间到输出空间的非线性、非正交的数学映射。由此可见,基于 BP 算法的前馈神经网络与非线性预测图象编码在数学本质上具有同构性。

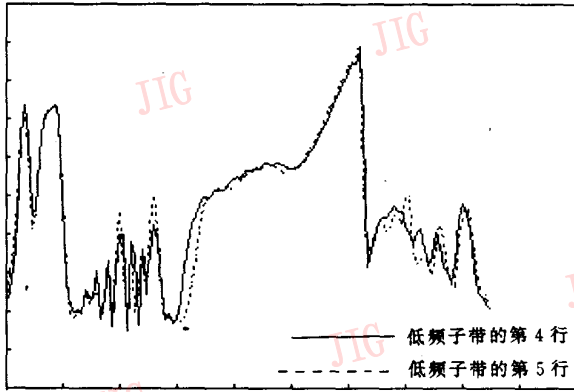


图 2 Lena(512×512)

Funashi 和 Hecht-Nielsen 分别证明了 3 层前馈人工神经网络所实现的非线性映射可以一致逼近紧集上的连续函数或按 L_2 范数逼近紧集上平方可积的函数。所以,我们在这里采用 3 层 BP 网络。具体算法如下:

设输入层为 N 维矢量 $X = \{x_1, x_2, \dots, x_N\}$; 隐含层有 L 个神经元,其输出矢量 $Y_h = \{y_{h1}, y_{h2}, \dots, y_{hL}\}$ 即为输出层的输入矢量;输出层含有 P 个神经元,输出矢量为 $Y = \{y_1, y_2, \dots, y_p\}$;输入层对隐含层作用的权系数构成 $N \times L$ 维的权重矩阵 $\bar{W}^{(1)}$,隐含层对输出层作用的权系数构成 $L \times P$ 维权重矩阵 $\bar{W}^{(2)}$;则有以下关系:

$$\text{隐含层输出 } y_{hj} = F\left[\sum_{i=1}^N x_i \omega_{ij}^{(1)} - \theta_{hj}\right]$$

$$\text{网络最终输出 } y_j = F\left[\sum_{i=1}^L y_{hi} \omega_{ij}^{(2)} - \theta_j\right]$$

在网络的学习过程中,首先训练样本矢量由输入层经由隐含层到达输出层,求出目标矢量与实际输出的误差。通常衡量误差 E 的方法是建立误差函数,如采用熟知的均方误差函数:

$$E = \frac{1}{2} \sum_j (y_i - t_j)^2$$

其中 t_j 为期望输出。如果没有达到精度要求,则利用该误差从后向前修正各层之间的权重,使输出误差减小。具体修正权重系数时,采用梯度下降法求

得优化的权值,即:

$$\Delta \omega_{ij} = -\eta \frac{\partial E}{\partial \omega} = \eta \delta_i y_j$$

$$\omega_{ij}^{new} = \omega_{ij}^{old} + \Delta \omega_{ij}$$

其中 δ_i 是反传误差。当节点 j 为输出层的节点时:

$$\delta_i = y_i(1 - y_i)(t_i - y_i)$$

当节点 j 为隐含层的节点时:

$$\delta_i = y_i(1 - y_i) \sum_k \delta_k \omega_{ki}$$

其中 k 取 j 节点所在层之上一层的节点。照此方法,直至输出的均方误差达到要求的精度。

将上述 BP 网络应用于构造非线性自适应预测器,在学习训练阶段,将上一行低频子带数据送到输入层作为输入信号,而相邻下一行低频子带数据送到输出层作为目标信号,当网络训练达到所需的精度时,整个网络就可以实现了对低频子带的非线性预测功能。

2 实验结果

以 Lena(512×512×8bit)为例,基于 BP 算法的前馈神经网络在经过 1 000 步左右的学习训练之后,网络的训练精度达到 0.012 左右;图 3 给出了低频子带任取一行数据与由相邻的上一行数据通过训练好的人工神经网络进行非线性预测输出的拟合情况,以及两者之间的误差分布;

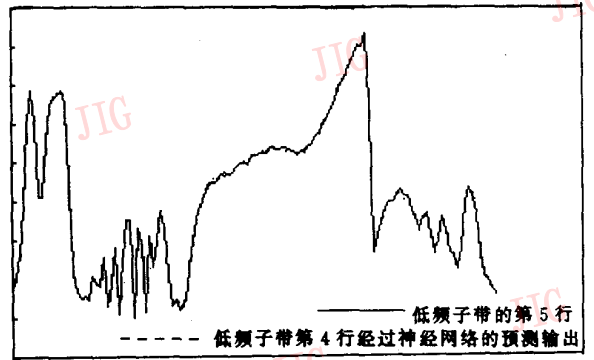


图 3

图 4 给出了低频子带任取相邻两行直接相减的误差分布情况和由前一行通过神经网络预测下一行的误差分布情况。

图 5 给出了相邻行直接相减的误差统计分布和相邻行通过神经网络进行非线性预测输出的误差统计分布比较情况。

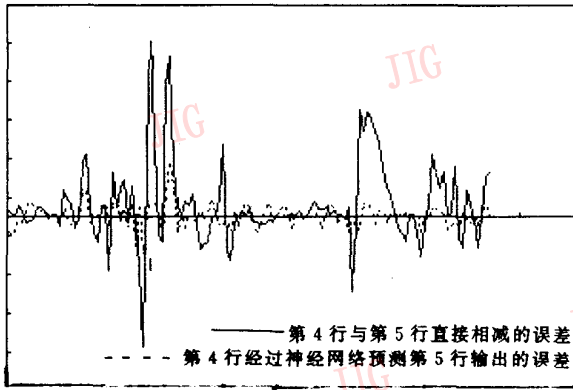


图4

表3给出了 Lena(512×512×8bit)的原始低频子带系数 LL₂、相邻行直接相减输出的误差系数以及经由神经网络的非线性预测输出的误差系数之间最大值、最小值、均值和均方差指标比较。

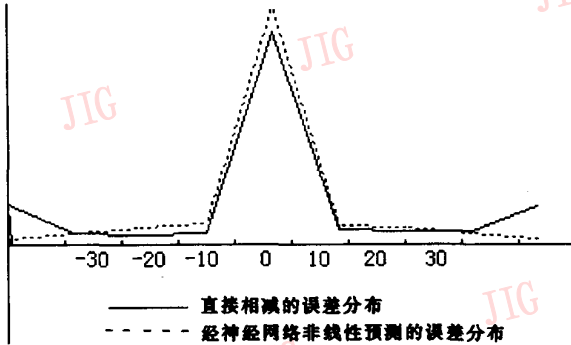


图5

表3 Lena(512×512×8bit) LL₂系数、直接误差系数和非线性预测误差系数

	原始低频子带系数	相邻行直接相减的误差系数	经由神经网络非线性预测的误差系数
最大值	972.2	596.24	126.67
最小值	80.32	-555.5	-101.91
均值	496.3	0.513	-0.056
均方差	35 109.99	3 908.47	252.18

由图3至图5及表3可以看出,经过基于BP算法的前向神经网络进行非线性预测输出的误差图象具有数据动态范围小、去冗余性好的优点,且误差数据符合 Laplacian 分布,所有这些都利于对低频子带系数的压缩编码。

在非线性预测之后,我们可以对误差图象系数采用非均匀量化,并进行 Huffman 游程混合编码。表四给出了在不同置0门限条件下图象的二级低频子带压缩比,总体压缩比、总体信噪比和恢复图象的

主观质量;并同时给出了一幅恢复图象,如图6所示。

表4 不同置0门限下的压缩比、信噪比和主观质量

置0阈值	LENA(512×512×8)			
	二级低频子带压缩比	图象总体压缩比	信噪比	主观质量
$ T < 3$	2.76	21.87	33.7	很好
$ T < 6$	2.96	22.9	32.87	很好
$ T < 12$	3.66	25.2	31.75	良好
$ T < 16$	4.85	27.85	30.82	好
$ T < 20$	5.52	29.1	30.25	好

4 小结

综上所述,应用神经网络进行非线性预测,可以有效地去除低频子带 LL₂系数间的相关性;从而对低频子带 LL₂获得较为理想的压缩比和信噪比。此外,这种应用神经网络构造非线性系统的思想启发我们在处理一些非线性系统时,除了传统的通过分析建立数学模型方法以外,也可以将所处理的非线性系统看成一个黑匣子,充分利用神经网络优良的非线性映射能力,辅之先验认识条件的约束,使网络认识系统的特性,从而达到较好的匹配和拟合效果。



图6 Lena(512×512×8bit) (CR=29.1; PSNR=30.25)

参考文献

- 1 Rumelhart D E, Hinton G E, Williams R J. Learning Internal Representations by Error Propagation. Parallel Distributed Processing, 1986, 1.
- 2 Hopfield J, Tank D W. Neural Computation of Decisions in Optimization Problems. Biol. Cybern., 1985, 141~152.
- 3 Chua L O, Lin T. A Neural Network Approach to Transform Im-

- age Coding. Circuit Theory and Applications, 1988, 16(3):317~324.
- 4 程存学,朱家瑞. 图象压缩编码的新技术. 海军装技委通信分会第二次年会科学技术论文集, Nov. 1996.
- 5 李象霖. 数字图象处理. 北京: 中国科学技术大学出版社, 1992.
- 6 姚庆栋, 毕厚杰, 王兆华, 徐孟侠. 图象编码基础, 北京: 人民邮电出版社, 1984.



武 斌 1971年生, 硕士, 毕业于中国科学技术大学通信与信息专业, 现在海军电子工程学院从事通信与图象处理的研究工作。



李象霖 中国科学技术大学研究生院(北京)电子学部教授, 主要从事数字图象处理, 三维运动分析教学和科研工作。感兴趣的科研方向: 图象序列分析, 计算机视觉、数字图象处理与分析、遥感与医学图象处理。

Image Compression Method Research in Wavelet Transform Domain by Using BP Neural Network

Wu Bin

Li Xianglin

(Naval Electricmc Engineering College, Nanjing 210029) (School of Graduate USTC, Beijing 100039)

Abstract After discussing the second level low frequencies part of wavelet transform, which is rather difficult to be compressed, we can use the nonlinear mapping capacity of multi-layer free-forword neural network on BP algorithm. A scheme of image coding based on nonlinear predictor is put forward. Redundance among the wavelet coefficients can be rided effectively with the help of this method. On the condition of relatively low distortion degree, the rather satisfactory compress rate can be achieved.

Keywords Wavelet transforms, Nonlinear predict, BP network

敬告本刊邮发订户

请本刊邮发订户注意:

凡从邮局订阅1999年《中国图象图形学报》的订户, 请将订刊收据的复印件以及你们的联系地址、邮编、电话, 及时函告本刊编辑部, 以便向你们寄赠本刊B版:《图象图形世界》。

切切, 多谢各位的支持与合作。

本刊编辑部