

## 神经网络在图象处理中的应用\*

马建波\*\*

(清华大学电子工程系)

**摘要** 图象处理是神经网络的一个重要应用领域。本文首先介绍了神经网络的模型、学习规则及其处理能力,在此基础上综述神经网络在图象处理中的应用,重点放在神经网络分类器和 Hopfield 网络解决图象处理中的最优化问题。

**关键词** 图象处理, 最优化, 神经网络, 分类器

**分类号** TN 941.1

神经网络是近年来学术界的一个热门课题,许多不同学科的科技人员正以极大的兴趣和热情投入神经网络理论、结构和应用的研究。

神经网络是源于大脑的一种计算结构。传统 AI 的方法本质上是把人脑看成是“黑箱”来研究的,而神经网络方法则试图走“白箱”的道路。现代计算机和算法计算擅长解决定义明确、解法明确的问题,如工程计算,而对于那些定义明确、解决不明确或定义和解法均不明确的问题,解决能力有限或十分困难。这类问题中的一些,如模式识别正是人脑所易于解决的。这就导致了研究人员研究神经系统,试图设计出具有类似大脑功能的智能计算系统,同时现代集成电路技术提供了实现大规模并行网络处理的可能,神经网络使我们能利用 VLSI 的发展来推动 AI 领域的研究。

神经网络经历了30多年的曲折发展过程<sup>[1]</sup>。我们认为,神经网络之所以成为一个热门课题是由于三个方面的原因:

(1) AI 研究所面临的困难。比如在计算机视觉、图象处理等领域,信息的描述问题一直是一个难于克服的难点,神经网络方法给逾越这个问题带来了希望。

(2) 神经网络本身的发展。这包括学习算法的改进<sup>[4]</sup>,理论基础的完善<sup>[2]</sup>,仿真研究的计算机系统的发展<sup>[3]</sup>,实现技术的进步。

(3) 神经网络模型提供了解决问题的新方法,有许多成功的应用范例。

目前,神经网络的研究侧重于三个方面<sup>[5]</sup>:①作为知识表示的工具;②作为学习系统;③作为计算的数学模型。有一种说法认为,逻辑思维是串行的,而形象思维则是并行的。神经网络作为高度并行处理的计算结构,无疑适于解决类似感知等非适应 (ill-posed) 问题。无论是用于自适应学习还是最优化计算,神经网络在计算复杂度和性能上的潜力远大于传统的方法,神经网络的本质十分适合解决图象处理和计算机视觉中的一些复杂而计算量大的问题,如模式识别、纹理分析、三维图象重建和识别、图象恢复、图象压缩和图象分割等。

\* 1991年9月3日收稿

\*\* 博士研究生

本文在描述神经网络的模型、学习规则之后,分析神经网络的信息处理能力,然后讨论神经网络在图象处理中的应用。

## 1 神经网络及其信息处理能力

### 1.1 神经网络的数学描述

神经网络机理的研究是一个十分重要的课题。S. Amari 描述了神经计算的数学基础<sup>[2]</sup>,从数学的观点考虑了多种神经计算结构的能力和局限,提出了层次网络变换、统计神经动力学、联想记忆及神经学习的一般理论。T. Poggio<sup>[3]</sup>以逼近理论的观点,将神经网络的学习问题看成是广义样条插值,通过成熟的逼近理论分析神经网络的学习机制。

目前,研究和应用中的神经网络模型大约有30多种,常用和比较著名的有14种<sup>[16]</sup>。笼统地分为前馈网络和反馈网络。前馈网络已成功用于声纳信号处理、语言识别、股市预测、图象压缩、自适应过程控制等;反馈网络则用于自组织学习、检索存贮模式和计算大量的最优化问题。

从数学观点来看,神经网络是具有如下特性的有向图:

- (1) 每个节点  $i$  与一个状态变量  $n_i$  相关联;
- (2) 每个两节点间的支路  $(i, k)$  与一个实数值的权  $\omega_{ik}$  相关联;
- (3) 每个节点  $i$  有一个实数值的偏置  $\theta_i$  与之关联;
- (4) 对每个节点  $i$ , 定义一个传递函数  $f_i[n_i, \omega_{ik}, \theta_i, (k \neq i)]$ , 确定节点的状态为其偏置、输入支路权值及相连节点状态的函数。

用标准术语,节点称为神经元,支路称为突触,偏置叫做开启阈值。传递函数一般具有  $f(\sum_k \omega_{ik} n_k - \theta_i)$  的形式,其中  $f(\cdot)$  或者是阶跃函数,或者是分段线性函数,或者是平滑递增的 S 形函数。没有输入支路的节点称为输入神经元,前馈网络是没有环路的有向图。

### 1.2 神经网络的学习规则

神经网络的学习规则是传递函数的关键所在,用来修正连接权值的各种学习规则,都是作为经验函数而提出来的。目前大约有六类学习规则:

- (1) Grossberg: 输入加权平均的竞争学习规则;
- (2) Hebb: 相互重合输入的相关学习;
- (3) Kohonen: 形成一个与一个特定的概率密度函数相一致的向量集,是一种无监督的学习方法。输出可以看成是输入样本特征的一种表达<sup>[6]</sup>;
- (4) Kosko/Klopf: 形成时间顺序事件序列的表示<sup>[17]</sup>;
- (5) Rosenblatt: 一种有监督的学习规则,调整感知器线性分类器;
- (6) Widrow: 最小化均方误差代价函数<sup>[1]</sup>。

神经网络可采用一个或多个学习规则,少数神经网络没有学习规则,这些网络常用于解决固定的问题。当动用学习规则进行学习时,有三种训练策略:

- (1) 有监督训练: 给定网络输入数据和所期望的输出;
- (2) 分级训练: 只给定输入数据,不提供期望输出,给出一个性能得分;
- (3) 自组织: 只提供输入数据,期望网络自己形成一些有用的组态。

### 1.3 神经网络的信息处理能力

神经网络从一开始就被应用到信息处理中去,神经网络能够完成如下信息处理运算<sup>[16]</sup>:

- (1) 数学映射逼近: 根据样本集  $(x_1, y_1), (x_2, y_2), \dots, (x_L, y_L)$  自调整,形成函数的近似  $f: ACR^n \rightarrow B \subset R^m$ , 其中  $y_i = f(x_i)$  或  $y_i = f(x_i) + n$ ,  $n$  是平稳噪声过程。
- (2) 概率密度函数估计: 通过响应  $R^n$  中根据固定概率密度函数  $\rho$  选择的样本向量  $x_1, x_2, \dots$  自组

织形成一个等概率的稳定点集。

(3) 从二值数据库抽取相关知识：响应数据库记录输入，自调整形成数据库的记录域间有关统计公共关系的共同模型。

(4) 形成拓朴连续及统计保留映射：基于根据固定概率密度函数选择的输入数据，调整、自组织这种映射，最终的映射常常反映数据空间不同项之间的相似性。

(5) 最近邻模式识别：通过与预先分类和存储的样本模式集比较，进行模式分类；可用于空间和时空模式，而且可以运用层次存贮的模式来压缩存贮量。

(6) 数据归类：根据数据自组织形成可选择粒度的类别，在粒度的峰值点处可有限次改变类别；对任何不符合已有类别的新目标，可以形成新的类别。

## 2 神经网络与图象处理

图象处理是大数据量、大运算量的多维信息处理。广义上它不仅包括图象质量的改善、变换、图象的数据压缩，还包括特征提取、模式识别、图象理解等力图实现动物视觉功能的计算机视觉课题。图象的理解是一个非适应问题，研究中存在着很多困难<sup>[19]</sup>：

(1) 由于图象是三维景物的二维投影，一幅图象本身不具备复现三维景物的全部几何信息。这是个维数丢失问题。

(2) 图象在形成过程中受到多种因素的影响，但具体分清每个因素所产生的具体影响是很困难的。

(3) 图象本身不具有精确理解三维景物的全部信息，这就需要知识的导引，但表示和运用所需知识并非易事。

(4) 实现图象理解，在工程上面临着可观的数据量和运算量。

Marr<sup>[18]</sup>奠定视觉计算理论以来，这个领域的发展没有预想的那么快，进入90年代之后可以说是停滞不前。神经网络作为大规模并行计算结构，作为一种全新的方法，无疑给图象处理的研究注入了新的活力，带来了新的希望。

### 2.1 图象处理中的神经网络雏形

神经网络的思想可以说很早就能从图象处理中找到痕迹。Marr-Hildreth<sup>[18]</sup>于1979年所提出的边缘检测理论就是一例。

Marr-Hildreth的边缘检测方法，是从动物视觉理论导出的，对思想并不复杂，首先用高斯函数对图象进行滤波，然后进行Laplacian运算，最后检测过零点作为边缘，故又称 $\nabla^2 G$ 算子，其中

$$G(x, y) = \frac{1}{2\pi\sigma} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right); \nabla^2 G(x, y) = k \left[ 2 - \frac{x^2 + y^2}{\sigma^2} \right] \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)$$

$\sigma$ 是高斯函数的尺度参数。

这个算子的结果与动物视觉特性十分吻合，可以说是对视神经元的模拟。 $\nabla^2 G$ 算子可表示成一个单一的perceptron单元，如图1。

$X$ 是一定邻域内(2D)图象灰度， $W$ 是各输入的权值，对应于 $\nabla^2 G$ 算子。

这只是个很简单的雏形。事实上神经网络的含意及能力要比这些深远得多，神经网络可以用来完成很多图象处理任务。下面的几个小节中，我们把神经网络在图象处理中的应用分为模式识别和最优化计算两大类进行讨论。

### 2.2 神经网络用于模式识别

神经网络用于模式识别是一个活跃且重要的课题，已有大量的理论模型和应用实例。Lippmann给

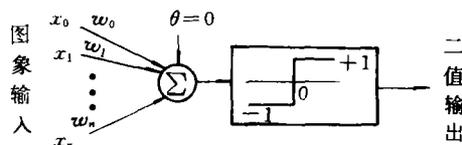


图1  $\nabla^2 G$ 算子的神经元表示

出了神经网络分类器的类别图，如图2。

这些模型或算法均可用于图象识别。以单个或多层 Perceptron 为例：单层 Perceptron 结构的判定域是以超面为界的半空间，只能解决线性可分的问题；两层 Perceptron 结构的判定域为凸开/闭区域；三层及多层 Perceptron 结构的判定域则是任意形状的，理论上能胜任所有的模式识别任务。

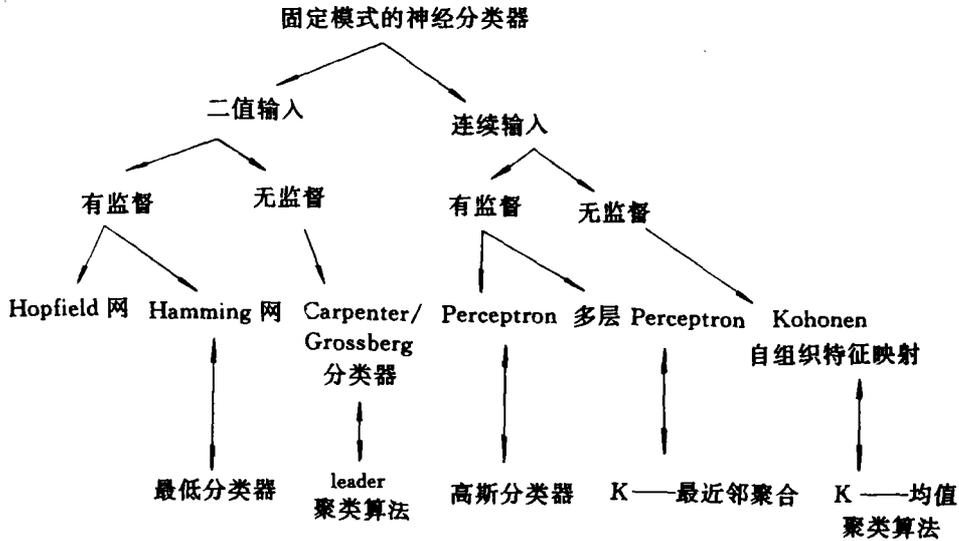


图 2 神经分类器的分类

目前，神经网络用于图象识别方面的研究多数集中在字符识别上，网络模型大多是多层 Perceptron，训练算法以 BP 为主流<sup>[5]</sup>，众多的研究都报道了较高的识别率。这方面的应用主要有两个思路：(1) 直接用图象模式训练神经网络分类器。Martin<sup>[14]</sup>用  $15 \times 24$  的灰度数组通过 BP 算法训练多层 Perceptron，在 5% 的拒识率上达到了 97% 的识别率。Burr<sup>[23]</sup>进行了识别手写大写字母的研究，在规整之后投影到模板中的 13 个线段上，用 BP 算法训练有 13 个输入、26 个输出的神经网络分类器，识别率是 90%~95%。显然，上述的两种方法尚不能保证图象的旋转、平移和尺度变化的情况，于是就有了解决问题的第二种思路。(2) 对输入模式进行预处理，抽取各种不变特征作为神经网络分类器的输入。Khotanzad<sup>[10]</sup>运用了 Zernike 矩对多层 Perceptron 进行训练，得出了较好的结果，并指出多层 Perceptron 的分类性能好于三种传统的统计分类器：贝叶斯分类器、最近邻分类器和最小平均距离分类器，特别是在存在噪声的情况下。

虽然神经网络字符识别有其成功的一面，但是有一个问题是不容回避的：训练时间和容量问题。选择大量有代表性的训练样本是获得高识别率的最关键因素，而样本集的增大必然要求加大容量，也增加了训练时间。这使得仿真计算更加令人难以忍受，即使用 VLSI 实现，也是不能忽视的。

神经网络分类器也可以用于图象两个最基本特征：边缘与纹理的识别上。神经网络在图象识别方面目前还未必比传统方法优越。我们认为，如果神经网络能避开十分困难的特征抽取和描述问题，并且能提高容量和速度的话，才能真正确立它在图象识别中的地位。

### 2.3 神经网络用于图象中的最优化计算

图象处理中的很多问题都可以转化为某个函数最小化的问题。从自 Hopfield 成功地用神经网络解决了具有 NP 完全性的组合优化问题，其思想被广泛应用到信息处理领域。也许正是因为 Hopfield 网络

才导致了当前的神经网络热潮。

### (1) Hopfield 网络与模拟退火

Hopfield 网络是一种对称连接的互连网络,对于这种网络 Hopfield 引入一个能量函数,或称 Lyapunov 函数(离散网格格):

$$E = -\frac{1}{2} \sum_i \sum_j T_{ij} V_i V_j - \sum_i V_i I_i$$

其中  $T_{ij}$  是单元  $i$  和  $j$  之间的连接权值,  $V_i$  是单元  $i$  的输出,  $I_i$  是单元  $i$  上的输入偏置。这种网络与 VLSI 有直接的对应,可以把每个单元称为神经元。单个神经元的运动方程是:

$$dU_i/dt = -U_i/RC + \sum_j T_{ij} V_j + I_i$$

其中  $V_i = g(U_i) = 1/2(1 + \tanh U_i/U_0)$  是  $S$  形函数的输出,  $RC$  是输入时间常数。

改变网络节点  $j$  的状态  $\Delta V_j$ , 能量的变化为  $-(\sum_i T_{ij} V_i + I_j) \Delta V_j$ , 令  $\Delta E_j = \sum_i T_{ij} V_i + I_j$ , 网络的运行规则是向能量单调减小的方向运行,即  $\Delta E_j \geq 0$  时,  $V_j$  取 +1,  $\Delta E_j < 0$  时,  $V_j$  取 -1。对网络节点的一个初始状态  $V(0)$ 。若网络  $k$  步运行后处于状态  $V(k)$ , 则

$$V_i(k+1) = f(\sum_j T_{ij} V_j(k) + I_i) \quad (1)$$

由于  $E$  是有界的,系统最后必然达到某个状态不再运行,或在某  $n$  个状态绕行。这些状态是非线性动力学系统的吸引子,或者说是能量函数的局部极小点。

Hopfield 网络解决 TSP 问题与人类解决此类问题的方法很相似,但是常收敛到局部极小点上。Hinton 将模拟退火应用到 Hopfield 能量模型中,导出了一种称为“Boltzman 机”的网络学习算法。模拟退火方法只是对 Hopfield 算法作了一点改动,把(1)式变为:无论节点  $j$  在  $k$  时刻的状态取何值,在  $k+1$  时刻其状态取 +1 的概率为

$$P(k+1) = \frac{1}{1 + e^{-\Delta E_j/T}}$$

$T$  是类似于物理系统的温度参数,当系统达到“热平衡”时,服从 Boltzmann 分布

$$P_\alpha/P_\beta = \exp(-(E_\alpha - E_\beta)/T)$$

这里,  $P_\alpha$  是网络处于  $\alpha$  状态的概率,  $E_\alpha$  是这个状态的能量,模拟物理系统退火如下:

初始温度足够高,使系统处于任一可能状态;然后缓慢降温,使系统在每一温度下保持平衡;当温度降为 0 时,能量函数达到全局最小点。

Hopfield 网络和模拟退火为解决最优化问题提供了一种重要的思想方法。目前大多数神经网络优化计算都是在这两者的基础上进行的<sup>[11]</sup>。

### (2) 用神经网络求解图象处理中的最优化问题

首先,我们来看一个浓淡点化灰度图象的神经网络方法<sup>[20]</sup>。

建立图象的 Hopfield 模型,稳态下每个神经元的输出为

$$y(m,n) = f(b_{m,n} + \sum_{(i,j) \in R'} w_{ij} y(m-i, n-j))$$

其中  $m, n$  为图象坐标,  $f$  是  $S$  型函数,  $(i, j)$  覆盖连接区域  $R'$ ,  $w_{0,0} = 0$ ,  $b_{m,n}$  是外部输入。取  $S$  型函数的极限情况——硬限幅函数,神经元输出的稳定状态是下列能量函数的局部极小:

$$E = -\frac{1}{2} \sum_{m=1}^N \sum_{n=1}^N y(m,n) \left[ \sum_{(i,j) \in R'} w_{ij} y(m-i, n-j) \right] - \sum_{m=1}^N \sum_{n=1}^N b_{m,n} y(m,n)$$

也就是说,每个神经元首先计算出一些近邻神经元输出的加权平均,把结果通过硬限幅函数来得到对应的二值象素值,在能量函数的局部极小处,所得到的浓淡点图象与原灰度图象的误差最小。

很多图象处理中的算法都可以表示成一个最优化问题。MIT 的研究人员已将许多视觉预处理问题表示成最小化能量函数和找出最佳贝叶斯估计的形式。这些都可以用神经网络实现,其中包括:(1) 表面插值;(2) 边缘检测;(3) 从阴影中确定物体形状;(4) 光流场估计;(5) 彩色分离;(6) 从运动图

象中确定其结构<sup>[7]</sup>。

用神经网络解决图象处理中最优化问题的报道很多。文献[24]用 Boltzmann 机进行属性图的匹配, 文献[25]提出把图象分割问题表示成最优化问题求解。Simic<sup>[15]</sup>提出了基于神经网络寻找两个同构图最优匹配的一般方法。Salem<sup>[9]</sup>用 Hopfield 网巧妙地解决了积木世界的线条标记问题。

我们认为用神经网络的方法解决图象处理中最优化问题的潜力很大, 关键在于如何把问题表述成 Hopfield 网络模型易于接受的形式。

其他类型的一些应用, 诸如图象序列压缩<sup>[22]</sup>, 特征选择<sup>[13]</sup>, 矢量编码<sup>[21]</sup>, 等等, 在此不一一列举。

### 3 结束语

在 Von Neumann 机结构中, 以神经网络硬件代替当前的硬件会提高运算的效率。神经网络技术在图象处理中的应用主要集中于两个方面: (1) 模式识别: 目前的神经网络分类器与传统分类器相比, 优点还不够突出, 人们期待着全新神经分类器的出现。(2) 最优化计算: 这类应用的面很广, 也很灵活, 潜力极大, 关键在于把图象处理算法表示成某个函数最小化的形式。

神经网络在图象处理中的应用尚处于起步和探索阶段, 理论和方法还不够成熟。神经网络能否给图象处理领域带来革命性的变革, 依赖于集成电路技术和神经网络理论的突破性进展, 以及对人类视觉和感知机理的进一步了解。

## Neural Network Application in Image Processing

Ma Jianbo

(Department of Electronic Technology, Tsinghua university)

### Abstract

Image processing is an important field of the application for neural network. In this paper, the models of neural network, learning rule and the capabilities of message processing are introduced, Then, the applications of neural network for image processing is described. The emphasis of this paper is put on the classifier of neural network and Hopfield network for solving optimum problems in image processing.

**Key words** neural network, image processing, optimization, classifier