

基于多层神经网络的交叉线自动识别

曹爱军 史杏荣 孙贞寿 梅建华

(中国科学技术大学信息处理中心, 合肥 230027)

摘要 神经网络是一个非线性动力学系统,在许多方面都表现出强大的生命力,特别为信息科学界所瞩目.将图形信息输入计算机是CAD、GIS等应用系统中不可避免而又十分繁重的工作,人们一直在追求一种自动化的方法来解决这一问题.该文在深入分析Hopfield模型及多层神经网络的理论基础、学习算法的基础上,尝试使用神经网络的方法进行交叉线的自动识别,提出了一种全新的思路,并且用BP学习算法实现了这一思路.在实验模拟中,该算法显示了优异的性能.

关键词 地图自动识别 多层神经网络 交叉线识别 BP算法

中图分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2000)02-0149-04

Intersecting Lines Recognition Based on Multilayer Network

CAO Ai-jun, SHI Xing-rong, SUN Zhen-shou, Mei Jian-hua

(Information Processing Center, University of Science and Technology of China, Hefei 230027)

Abstract Neural Network is a non-linear dynamics kinetics system, which can do a lot of complex jobs, especially in information science. Inputting image information into computers is an inevitable and troublesome job in many practical systems such as CAD and GIS, etc. An automatic method is needed badly. After analysing fundamental theories of Hopfield model and multi-layered network, a new method to solve the problem realized by a BP algorithm is proposed. During experimental simulations, this algorithm really showed fine performance.

Keywords Map recognition, Multilayer network, Intersecting lines recognition, BP Algorithm

0 引言

孙贞寿等^[1]成功地提出了一种用Hopfield神经网络识别交叉线的模型和算法.该算法将检测到的含有交叉线的图象映射到一个Hopfield神经网络之上,每一个神经元对应于一个像素,神经元间的连接权重表示为像素间的关系函数.在这种权连接下,神经网络的状态可以用一个能量函数来表示,将交叉线识别归结为最优化问题.通过能量函数最小化,可以将交叉线段分离开来.

该交叉线识别算法的中心思想是将交叉线作为一种模式,交叉的每一条线均作为该模式的一个状态分量;Hopfield网络的主要任务是将要识别的交

叉线与它所能识别的模式对应起来,得到模式的各个状态分量,从而确定交叉线的走向.实现此中心思想的核心是所给出的计算权重的经验性公式;这个经验性公式具备怎样的存储能力无法从理论上求知,因此它只能识别出常见的交叉方式.另外,模型的尺寸定为 16×16 ,Hopfield网络的存储能力有一个理论上的公式 $N / (2 \times \ln(N))$,其中 N 为神经元的个数.由此我们可以计算出模型的最大存储能力为 $256 / (2 \times \ln 256) = 23$.这显然远小于实际存在于一个 16×16 图象中的交叉方式.

为提高模型的识别能力,可以修正权重公式或增加模型的尺寸.但是我们不能指望获得一个能识别所有交叉模式的权重计算公式,因此该方法无法从根本上使模型的性能得到改善;如果增加模型的

尺寸,那么模型所能识别的模式数与模式空间的大小也将同时增加,且模式空间要大得多,因此也不能通过这种方式来提高模型的识别能力.

本文从认知学的角度出发提出了一种解决交叉线识别的新思路.

1 解决交叉线识别问题的新思路

根据人眼的视觉特性,如图1所示.假设眼睛从一条线的起点沿该线向前看,在没有遇到交叉点时很容易地找出下一点;当遇到交叉点时,人眼具备穿越交叉点的能力.因此对此过程可以作下述理解:从起点开始毫不费力地找出下一点,是一个获取样本信号的过程,是一个学习过程;遇到交叉点时则为一个判断过程,用已获得的样本判断下一点.这个描述完全符合实际的人眼的视觉特性.这样,将自动识别交叉线的问题描述为:已有样本信号 $\{x_1, x_2, \dots, x_N\}$,用适当的模型进行训练后对待判点 $\{y_1, y_2, \dots, y_M\}$ 进行判别分类.

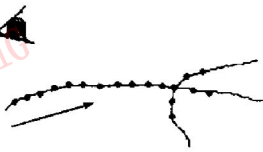


图1 人眼观察过程示意图

其次从图象处理方法本身来看,把一幅图象当成一个数量场的描述也是可行的.在图象编码中曾经有把一幅图象当成是一阶或二阶马尔科夫场的例子(其实际意义是某一象素可能只与左边或上边的象素相关)^[2].为方便起见,可以把一幅图象看作是由若干彼此无关的独立的数量场所迭加而成的一个合成的数量场.对于交叉线识别任务来说只要研究其中一个相关的数量场即可.把图象中的某条直线看作是与之相关的数量场的一条等值线,这样交叉线识别的任务就变成寻找数量场等值线上的点.即:对于数量场 $F_i = F_i(x, y)$,其中 F_i 具体表达式未知,现已知其等值曲线 $F_i(x, y) = C$ (C 为给定常数)上的若干点 $\{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$,要求判别 $\{y_1, y_2, \dots, y_M\}$ 中属于此等值线的点.这样,交叉线的识别问题完全等价于一个函数逼近一条曲线的问题.

那么神经网络能否完成函数逼近这个任务呢?显然,神经网络模型中的前馈网络具备映射作用能力.从系统观点来看,这是一个高度非线性的映射,它的信息处理能力也来自于简单的非线性函数的多

次复合. Kolmogorov 于 1957 年提出的连续函数表示定理^[3,4]:

给定任一连续函数 $\Phi: E^m \rightarrow R^n$, $\Phi(x) = Y$, 这里 E 是闭单位区间 $[0, 1]$ (因而 E^m 是 m 系统单位立方体), Φ 可以精确地由一个三层神经网络实现,此神经网络的第一层有 m 个单元,中间层有 $(2m+1)$ 个处理单元,第三层有 n 个处理单元.

虽然该定理未提供如何构造这样的3层网络的方法,但从理论上保证了其存在性与可行性.

2 用于交叉线识别问题的多层神经网络

接下来的问题是针对交叉线识别的函数逼近问题确定多层神经网络的具体结构和神经元的特性.由BP构造定理,网络可分为3层或5层,如取为3层.样本点坐标为 (x_i, y_i) ,所以输入层只要2个神经元,分别接收 x 和 y 的坐标值.输出层由输入到输出映射的状态数确定,因为数量场的输出只要一个实数值,所以输出层只要一个神经元,其期望输出对一条曲线的任何样本点 (x_i, y_i) 均为一个常数 C . C 即为等势线 $F(x, y) = C$ 中的常量 C .在彩色或具有灰度值的曲线识别时, C 可取其彩色码或灰度值.在实验时对黑白二值的细化光栅图象中的各种线对象采用自动跟踪,所以 C 取为0.隐层结点数可调,从5个到10个不等,视收敛情况而定.输入与输出层的神经元的特性为线性,隐层神经元的特性则是Sigmoid曲线,连续可微且限幅;其最小值和最大值分别取0.1和0.9.网络结构如图2所示.

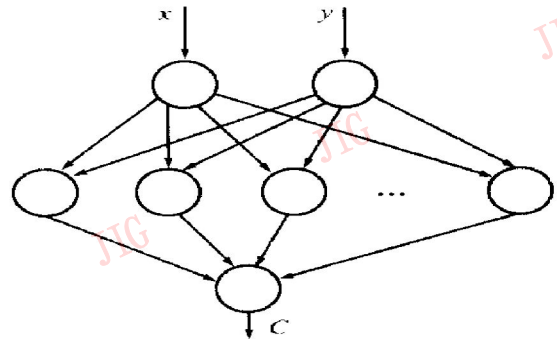


图2 应用于交叉线识别的多层神经网络示意图

经过细化、除刺等预处理后的线划图,从起始点出发,只要未遇到交叉点就可以自动控制获得线的样本集,即将下一个点的坐标 (x_i, y_i) 加入到线的样本点集,此过程无须人工干预即可自动进行.在遇到交叉点后,则用已经获取的样本集对网络采用BP

算法进行训练, 待训练符合一定收敛条件后结束训练过程. 训练终止条件取所有样本点或最近的 20 个点实际输出与期望输出差的绝对值之和小于一个给定的门限. 待训练完成后用备选点的坐标作为网络输入, 计算其输出, 从这些输出点中选择最符合期望输出的点作为下一个点, 即误差最小的输入点作为下一个点, 从而越过了交叉点, 并继续进行线的自动跟踪. 其过程如图 3 所示.

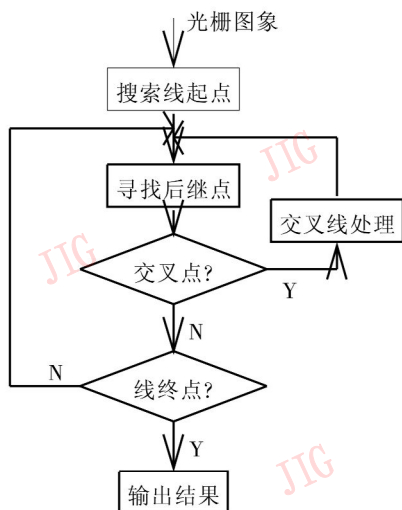


图 3 光栅图形矢量化线跟踪处理流程示意图

3 使用 BP 算法进行交叉线识别

许多应用任务很难用一系列顺序过程来完成. 例如进行较复杂的数据变换却没有预先定义好的映射; 或者输入数据中含有噪声, 这两种情况均不适合于用顺序过程来解决. BP 网络是一个适合于解决复杂模式及映射问题的人工神经网络. 其 BP 算法是一种供多层前馈网络所使用的监控式学习算法, 基本思想是 LMS 算法, 采用梯度搜索技术, 以期获得最小化网络的实际输出与期望输出的均方差. 网络的学习过程是一种误差边向后传播边修正权重的过程, 实际上包括了两个阶段: 前馈计算阶段和反向调整权重阶段. 当然并没有反馈存在, 实际运行仍然是单向的, 所以并不能将其视为一个非线性动力系统, 只能看作一个非线性映射系统^[5].

BP 算法具体步骤如下:

(1) 将输入矢量 $\mathbf{X}_p = (x_{p1}, x_{p2}, \dots, x_{pn})^T$ 应用于输入单元;

(2) 计算隐层单元的激励值 $net_{pj}^h = \sum_{i=1}^N W_{ji}^h x_{pi} \theta_i$;

(3) 计算隐层单元的输出值 $i_{pj} = f_j^h(net_{pj}^h)$;

(4) 计算输出单元的激励值

$$net_{pk}^o = \sum_{j=1}^L W_{kj}^o i_{pj} + \theta_k;$$

(5) 计算网络输出 $O_{pk} = f_k^o(net_{pk}^o)$;

(6) 计算输出层的误差

$$\delta_{pk}^o = (y_{pk} - O_{pk}) f_k^o'(net_{pk}^o);$$

(7) 计算隐层单元的误差

$$\delta_{pj}^h = f_j^h'(net_{pj}^h) \sum_k \delta_{pk}^o W_{kj}^o;$$

注意在权重更新前进行计算误差的值.

(8) 更新输出层权重;

(9) 更新隐层权重 $W_{ji}^h(t+1) = W_{ji}^h(t) + \eta \delta_{pj}^h x_{pi}$;

(10) 计算 $E_p = \frac{1}{2} \sum_{k=1}^M \delta_{pk}^2$.

对所有 E_p 样本均小于一个可忍受的水平, 否则重复 (1) ~ (10) 步骤.

在软件模拟时, 隐层神经元的个数取 5, 如果收敛后减少其个数仍收敛, 减少隐层神经元数; 否则不变. 倘若在 5 个神经元迭代最大次数后仍未能收敛, 增加隐层神经元的个数, 并重复迭代过程, 直至收敛为止. 由 BP 构造定理, 样本空间在 $[0, 1]^n$ 中, 所以样本点坐标 (x, y) 按式 $((x - x_{min}) / (x_{max} - x_{min}), (y - y_{min}) / (y_{max} - y_{min}))$ 转换. 其中, (x_{min}, y_{min}) 为地图左下角的坐标, (x_{max}, y_{max}) 为地图右上角的坐标. 收敛门限取值太小, 则收敛时间长, 精度高; 反之, 收敛快, 精度低. 通常取 0.001 或 0.000 5.

图 4(a) 是从一张石油勘探图中截取扫描下来, 并经二值化处理后得到的 128×128 的黑白二值图象. 图中的 2 条线分别代表河流和地层缺失区的边界. 我们对河流进行了自动跟踪, 图 4(b) 是跟踪后的结果.

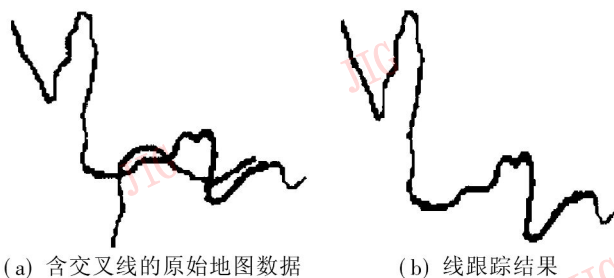


图 4 单条线的跟踪

图 5(a) 是某省部分公里原始地图经预处理后的 128×128 的黑白二值图象, 图 5(b) 是其中部分线的自动识别结果.

实验表明, BP 算法完全可以成功地训练图 2 所示的网络, 并进行交叉线的识别.



图5

4 结束语

无论是 Hopfield 模型还是应用 BP 算法或遗传算法的多层模型, 识别均依据于点集位置上天然的关联. 在地图中的各个要素客观地存在于地图上, 通过人们认知过程反映于人们的主观观念之中. 只有主观上的概念与客观存在相一致时, 识别结果才有意义; 否则必须通过人工补救的方式来修正. 可以看出, 在没有完全不同的机理或者说对人的认知本质没有突破性的进展的情况下, 地图的完全数字化还有很长的一段路. 现有的各种方法只是尽可能减少

人工干预的工作量. 本文提出的算法在大量的实验基础上为此作了大胆的尝试, 并取得了一定成果.

参考文献

- 1 焦李成. 神经网络理论. 西安: 西安电子科技大学出版社, 1990.
- 2 吴健康. 数字图象分析. 北京: 人民邮电出版社, 1989.
- 3 Lippman R P. An Introduction to Computing with Neural Nets. IEEE ASSP Magazine, 1987, 2.
- 4 孙贞寿等. 基于 Hopfield 模型的交叉线识别算法. 中国图象图形学报, 1998, 3(8): 684~ 687.
- 5 Richard J Mammone, Yehoshua Y Zeevi. Neural Networks: Theory and Application, Academic Press INC, 1991.



曹爱军 1974年生, 1996年获得中国科学技术大学电子工程专业学士学位, 同年9月攻读本校电子与通讯系统专业硕士研究生. 主要研究方向为智能信息系统、信号处理和多媒体技术. 已发表论文4篇.



史杏荣 1946年生, 教授, 1970年毕业于清华大学自动控制系. 主要研究方向为数据库、地理信息系统和多媒体信息系统等. 已在国内核心刊物上发表论文10篇, 出版教材或专著4本.