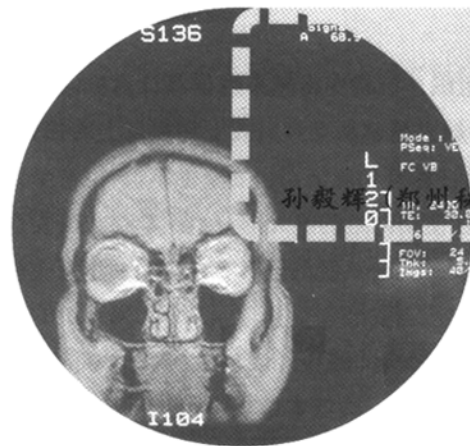


# 基于神经网络的图像识别系统



**摘要:** 本文介绍了一种以神经网络为基础并结合专家系统的方法进行图像识别的系统。着重介绍了在设计中较为困难的部分—抽象概念的表示和推理机的设计, 这些是系统的核心部分决定着系统的整体性能。

**关键词:** 图像识别 神经网络 神经元

目前计算机图像识别技术已越来越广泛应用于各个领域, 但是根据一些现有的软件来看识别效果并不理想。本文介绍了一种基于神经网络并与专家系统相结合的方法进行图像识别的系统。

神经元只接收一个像素, 再将输入层中的像素送至隐含层, 同样一个神经元只接收一个像素。因为隐含层中每个神经元都对应一个坐标, 所以每个像素也相应有一个坐标值。隐含层就是利用这些坐标训练出像素之间的位置关系。例如: 只向系统输入一些坐标为(1、2),(1、5),(1、11),(1、13)的一些像素, 虽然每两个像素之间的坐标不一样, 但是它们之间的位置关系都一样, 最终系统将收敛于一种状态, 归纳出“上下关系”这一概念。用同样的方法可以训练其他一些位置关系。如图2中A、B两神经元的坐标为(1、4),(4、1), 通过坐标差可以反映出A、B两点间的位置关系, 坐标差为(1-4、4-1), 得(-3、3), x为-3表示B在A的下方, Y为3表示B在A的右方, 所以B在A的右下方。注意这里并不是将坐标差的正负与“上”“下”简单地对应起来, 而是利用训练时归纳出的概念进行推导得出的结果。

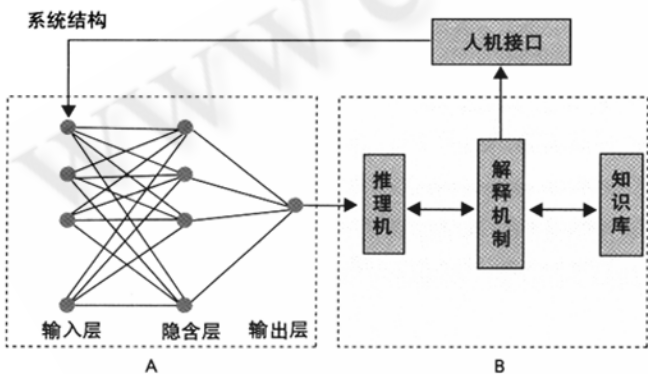


图 1

如图1所示, 该系统有两部分组成, A为神经网络, B为专家系统。图像由输入层进入系统经A预处理后送至B, 由专家系统推导出图像的基本特征。系统在识别不同类型的图像时其训练方法与识别过程有所不同, 以下分别加以介绍。

## 1 单色图像的识别

图像由像素组成, 像素的不同排列形式组成了具有不同特征的图像, 所以只要检测出每个像素之间的位置关系, 就可以确定图像的基本特征。像素主要是在隐含层中确定其位置关系。隐含层的结构如图2所示, 神经元间的权值为1, 为了便于操作可以为每个神经元确定一个坐标值。

首先将图像文件输入至输入层, 在输入层中一个神

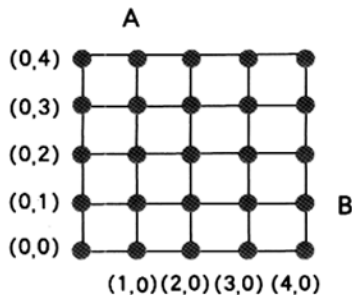


图 2

有了上面的基础后就可以对系统进行一些简单图形的训练。向系统分别输入一些不同方向的直线如图3所示, 系统就以像素A为基准分别计算出A与其余各个像素之间的坐标差, 这就得出了一组坐标差, 这组坐标差反映了像素

A与其他各象素之间的位置关系。将象素间的位置关系综合在一起分析就能得出所有象素的分布规律,最后得到图形的特征,在这里就是直线的特征。这一过程由推理机完成。对于不同方向的直线来说它们都属于同一类的图形,所以最终推导出的特征是一致的。各类图形的特征都存放在知识库中,系统就是利用不同的特征来识别不同的图形。注意图形的特征并不是用图像中象素的坐标或坐标差直接表示,而是由象素的坐标差推导出的一类抽象概念。关于抽象概念的表达比较复杂,而且还涉及到前期专家系统学习的一些问题,由于篇幅所限这里不再详细讨论。



图 3 一个方块代表一个象素(放大后的直线)

用上述方法可以对系统进行其他图形的训练,例如三角形训练,三角形有很多种,但训练时只需输入少数几种,系统就可以归纳出它们的基本特征,系统可以利用这一基本特征正确识别其他三角形。在识别简单图形的基础上系统可以识别复杂图形,因为复杂图形是由简单图形组成的。



图 4

从前面的例子看,只要计算出一组坐标差就可以推导出图形的基本特征,但是在实际应用中必须计算出图像中每一个象素与其他各个象素之间的坐标差,如果图像中有 $n$ 个象素,那么就会计算出 $n$ 组坐标差,然后根据这 $n$ 组坐标差共同分析得出图像的特征,这是因为当图像中有干扰时,只有共同分析 $n$ 组坐标差才能正确识别图像,这也为以后的模糊识别奠定了基础。

如图4所示,将A、B两图分别输入系统进行分析,最后推导出的特征是一致的,因为它们属于同一类图形。在系统识别的过程中,要计算出每一个象素与其他各象素之间的坐标差,所以A的运算量显然大于B,为了提高系统的运行速度,可以将被识别的图像先加以处理再进行分析。因为图像识别主要是检测反差即根据图像边缘信息就能判断图像特征,所以要先对图像进行边缘处理,只保留边缘信息。图4中A经边缘处理后得到B,然后系统对B进行分析,这样既提高了运行速度又不影响识别结果。检测边缘首先考虑一模板,模板定义了象素Z5的8邻域,如下:

Z1	Z2	Z3
Z4	Z5	Z6
Z7	Z8	Z9

边缘就是灰度值有明显区别的两区域的边界。如果象素Z5在边缘处,那么它与相邻的8个象素的灰度值就不完全一致,这种情况下保留Z5,如果Z5的灰度值与相邻的8个象素一致,就说明Z5不在边缘处,那么就将Z5删除。用这一方法将所有象素检测一边,不在边缘处的象素将被删除,剩下的就是边缘处的象素,如图4, A就转换为B。边缘检测是在输入层中完成。以上所讨论的是系统识别单色图像的过程,单色图像中每个象素用1位表示,只有两种象素,所以边缘检测与系统识别都比较容易实现。

## 2 灰度图像的识别

灰度图像中每个象素用8位表示,换句话说它是256级灰度图像,所以系统的分析过程复杂一些,前面讲的识别单色图像只是为识别灰度图像奠定基础。在灰度图像中如果按256级灰度对图像进行边缘检测,那么就会出现很多小的区域,过多的区域会导致识别过程的复杂化,甚至无法识别图像,所以在对图像进行边缘检测之前先将256级灰度图像转换为16级灰度图像,然后再进行边缘检测。这样虽然会丢失一些信息,但图像的整体特征一般不会受到影响。接下来在系统识别的过程中不但要计算出象素的位置关系,还要考虑到象素的灰度值,这就要求对系统进行全面的训练,例如:输入一些由不同灰度的象素组成的同一类图形。虽然训练的样本足够多,但是当系统投入使用的时候识别率并不高,因为在实际应用中遇到的图像比样本更多更复杂,所以必须在系统使用的同时不断的进行人机对话,这样才能随时纠正错误,逐步提高系统性能。

(下转第48页)

### 3 彩色图像的识别

彩色图像中每一个像素可以用8位表示,也可以用16位或24位表示。但在进行边缘检测前先将图像中的像素转换为8位表示即256色,如果实际情况要求不高也可用16色表示。经边缘检测后不同的像素出现不同的区域,识别过程与灰度图像类似,先训练一些样本然后通过人机对话的形式来提高系统性能。这时的系统只能识别二维图像,如果加以改进则可用于识别三维模型,方法是先由三维目标模型得到二维图像,即通过两个不同角度得到两幅同一物体的二维图像。将这两幅二维图像同时输入系统,经输入层后汇聚在隐含层(一个神经元存储两个像素)。由不同角度得到的两幅图像并不完全相同,所以在隐含层中确定两图像的不同之处,再经专家分析就可得出三维模型的基本特征。这部分在设计时涉及到:物体形状的分析,特体形状的分析,训练样本的选取等一系列的问题,尤其是程序设计较为复杂,这里只是做个简要介绍。

### 4 文字识别

可以将图像中的文字视为图形的一种。将图像文字输入系统,在隐含层中计算出组成文字的各个像素之间的位置关系,然后通过推理机推导出文字的特征。通过训练可以看出不同的文字特征是不一样的,但是不同字体的同一文字特征是相似的,所以只要把一种字体的特征记录下来,其他字体的文字都可以识别,并不需要一一训练,不过草体字还需进一步训练才能正确识别。■

#### 参考文献

- 1 《人工神经网络及应用》 袁曾任, 清华大学出版社
- 2 《神经网络与模糊控制》 张乃尧, 阎平凡, 清华大学出版社
- 3 《模式识别与图像处理并行计算机系统设计》 张大鹏, 哈尔滨工业大学出版社
- 4 《人工智能及其应用》 蔡自兴, 徐光 清华大学出版社