

基于小波描述子和神经网络的形状识别

夏 敏, 刘宏申

(安徽工业大学 计算机学院, 安徽 马鞍山 243002)

摘 要:提出一种把小波描述子和神经网络相结合的形状识别方法。通过小波描述子提取待识别物体的形状特征, 所提取的特征值不受物体位移、缩放和旋转的影响, 接着用多层分类器网络对物体的形状进行识别分类, 并采用 BP 算法对神经网络进行学习和训练。最后得出令人满意的实验结果。

关键词:形状识别; 小波描述子; 多层感知器; BP 算法

中图分类号: TN911.73

文献标识码: A

文章编号: 1673-629X(2007)03-0106-03

The Shape Recognition System Based on Wavelet Descriptors and Neural Networks

XIA Min, LIU Hong-shen

(School of Computer Science, Anhui University of Technology, Ma'anshan 243002, China)

Abstract: The application of combining wavelet and neural networks to the 2-D pattern recognition on the basis of its shape is presented. First, use wavelet descriptors to extract the shape features. The wavelet descriptors provide the appropriate invariances to the scale, rotation and translation. Then use neural networks to recognize and classify the shape of the 2-D patterns. The type of neural networks is the multilayer perceptron. The shape recognitions systems constructed neural networks carry out studying and training by using the BP algorithm. Finally, a very good experiment result is obtained.

Key words: shape recognition; wavelet descriptors; multilayer perceptron; BP algorithm

0 引言

形状识别^[1]技术是计算机视觉技术的一种, 其目的是让计算机自动识别出图像中物体的形状, 该技术在生物特征识别、医学研究、机器人视觉等方面都有很广泛的应用。小波变换提供了多分辨率表示, 在粗尺度给出形状的全局信息, 在细尺度上给出局部信息, 所以小波轮廓描述子^[1]受轮廓的局部畸变影响较小, 轮廓描述精度高, 并且具有平移、旋转和尺度不变性。神经网络具有良好的容错性能与泛化能力, 近年来大量用于模式识别当中, 它使识别系统对噪音不敏感。文中介绍一种把小波描述子和神经网络相结合的形状识别方法。先用小波变换对形状进行预处理得到形状特征的小波描述子, 再把所得小波描述子作为多层分类器网络的形状特征的输入信号进行分类识别, 并采用 BP 算法对系统进行学习、训练。

1 特征提取——小波描述子

小波变换属于多分辨率变换, 它不同尺度上对信号分解。应用小波进行轮廓表示时, 需要选择小波系数的限制级数来描述轮廓, 并且要对小波系数进行归一化, 从而达到平移、缩放、旋转不变性的要求。称这些小波系数为小波描述子^[2]。

1.1 小波描述子的定义

小波函数族 $W_{m,n}$ 可如下定义^[3]:

$$W_{m,n}(t) = \frac{1}{\sqrt{2^m}} W\left(\frac{t - n2^m}{2^m}\right) = 2^{-m/2} W(2^{-m}t - n) \quad (1)$$

对于给定的形状函数 $f(t)$, 小波变换系数为:

$$D_{m,n} = \langle f, W_{m,n} \rangle = \int f(t) \bar{W}_{m,n}(t) dt \quad (2)$$

而由小波变换系数重建 $f(t)$ 的公式为:

$$f(t) = \sum_{m=m_0+1}^{\infty} \sum_{n=-\infty}^{\infty} \langle f, W_{m,n} \rangle W_{m,n}(t) + \sum_{m=-\infty}^{m_0} \sum_{n=-\infty}^{\infty} \langle f, W_{m,n} \rangle W_{m,n}(t) \quad (3)$$

其中 m_0 与截断系数时所需的精度有关。如果设尺度函数 $S_{m,n}(t)$ 为 $S_{m,n}(t) = 2^{-m/2} S_{m,n}(2^{-m}t - n)$, 并

收稿日期: 2006-05-20

基金项目: 安徽省教育厅自然科学基金项目(2006KJ018A)

作者简介: 夏 敏(1977-), 女, 安徽马鞍山人, 硕士研究生, 研究方向为计算机图像处理; 刘宏申, 教授, 研究方向为计算机图像处理、模式识别。

将式(3)右边第一项用 $S_{m,n}(t)$ 的线性组合来代替,则可将式(3)写为:

$$f(t) = \sum_{n=-\infty}^{\infty} \langle f, S_{m,n} \rangle S_{m,n}(t) + \sum_{m=-\infty}^{m_0} \sum_{n=-\infty}^{\infty} \langle f, W_{m,n} \rangle W_{m,n}(t) \quad (4)$$

根据小波变换的特点,上式右边第一项可以看作 $f(t)$ 在 2^{-m_0} 尺度下的模糊草图,第二项则是对 $f(t)$ 的细节补充。如果将 $\langle f, S_{m,n} \rangle$ 称为第 m 层尺度系数, $\langle f, W_{m,n} \rangle$ 称为第 m 层小波系数,则所有系数组成与形状 $f(t)$ 对应的小波轮廓描述符。

由式(4)知,如果选择所有系数都作为描述子,则轮廓的信息量完全没有损失。但在实际应用中有时并不需要那么精确,可以用某前 j 层小波特征来识别 $f(t)$ 。由实验得到,当轮廓数据点在 256 以上时,分级计算应该不少于 2 级,当然也不能太大,否则丢失的细节过多,一般可取 $j \leq 5$ 。

并且这些系数还要经过归一化处理,使它们不受尺度、位置和旋转的影响。归一化处理后得到了一组具有缩放、平移和旋转不变性的小波描述子。

1.2 变换系数的计算

设尺度系数为 $S_{m,n}$,小波系数为 $W_{m,n}$,把小波变换看成一个卷积、亚采样的过程,则可得如下迭代计算系数的公式:

$$S_{m+1,n} = \sum h_{m,l-2n} S_{m,l} \quad (5)$$

$$W_{m+1,n} = \sum g_{m,l-2n} W_{m,l} \quad (6)$$

其中 $h_{m,l-2n}$ 和 $g_{m,l-2n}$ 为与所选小波基有关的滤波器,卷积的范围 l 与当前尺度 m 有关,并满足 $l \in \{0,1,2, \dots, 2^{m+1}-1\}$ 。

2 神经网络识别

在此选用多层感知器^[4,5]网络来进行识别分类,并运用 BP 算法对系统进行训练。在神经网络中,多层感知器是最有力的工具,在用神经网络讨论模式识别问题时,多层感知器占有重要位置。

2.1 多层感知器的结构

多层感知器是一个前馈网络结构,其中只有相邻两层的神经元互相连接。图 1 给出一粗略的多层感知网络结构。

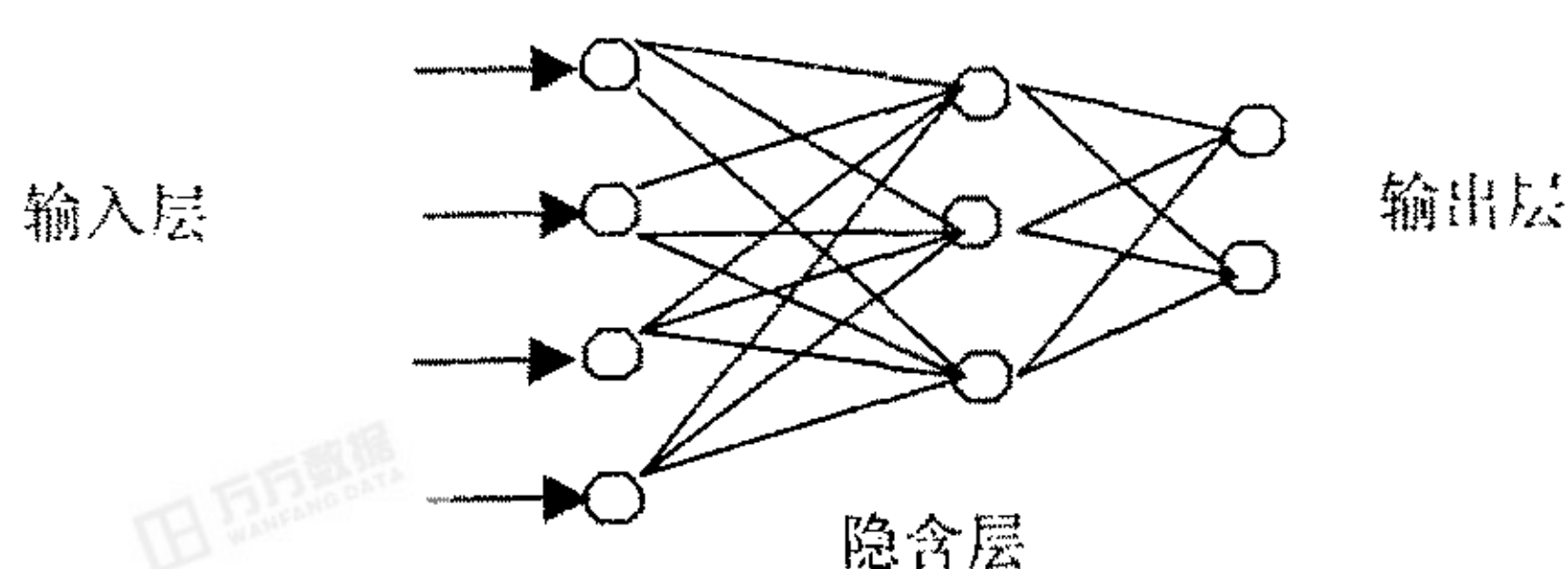


图1 前馈多层感知器结构

将代表待识别形状的输入矢量输入至输入层,并传至后面的隐含层,最后通过连接权输出到输出层。该网络中每个神经元通过求输入权值和及经非线性兴奋函数传递结果来工作,其数学描述如下:

$$out_i = f(net_i) = f\left(\sum_j w_{ij} out_j + \theta_i\right) \quad (7)$$

这里 out_i 是所考虑层中第 i 个神经元的输出; out_j 是前一层第 j 个神经元的输出。对非线性兴奋函数 f 经常采用的是 sigmoid 函数:

$$f(net_i) = \frac{1}{1 + e^{-\frac{net_i}{Q_0}}} \quad (8)$$

2.2 网络训练——BP 算法

多层感知器网络的学习任务是使网络转向网络误差最小化,网络误差是对所有训练的输入输出对的计算。定义网络误差如下^[6]:

$$E_p = \frac{1}{2} \sum_{j=1}^N (t_{pj} - o_{pj})^2 \quad (9)$$

式中 E_p 是第 p 个表征矢量的误差, t_{pj} 是第 j 个输出神经元的期望值, o_{pj} 是第 j 个输出神经元的实际值。在网络的学习阶段采用梯度下降训练算法中比较简单的反向传播算法来调节权值。中心思想是调节权值使网络总误差最小。反向传播训练算法所需的微分方程是:

$$\Delta w_{ji} = \eta \delta_{pj} o_{pi} \quad (10)$$

其中, η 为学习速率, δ_{pj} 为 L 层神经元 j 的误差信号, o_{pi} 为 $L-1$ 层神经元 i 的输出。误差信号 δ_{pj} 可表示为:

$$\text{对输出神经元 } \delta_{pj} = (t_{pj} - o_{pj}) o_{pj} (1 - o_{pj}) \quad (11)$$

$$\text{对隐含层神经元 } \delta_{pj} = o_{pj} (1 - o_{pj}) \sum_k \delta_{pk} w_{kj} \quad (12)$$

o_{pj} 为 L 层神经元 j 的输出, o_{pi} 为 $L-1$ 层神经元 i 的输出, δ_{pk} 为 $L+1$ 层神经元 k 的误差信号。

多层感知网络的形状特征的输入信号由小波处理产生,输出信号指明了输入矢量 x 属于哪个特定的模式类别。为了达到最好的训练网络性能,必须用最小限度的多层感知结构以使能够缩小误差量达到一个低的、可以接受的水平。此外,训练样本的数量应该足够大,以保证 Vapnik - Chervonenkis 需要^[7]。

3 实验结果

实验对象是图 2 中的 4 个钥匙形状。先对形状进行相关处理得到形状函数,此形状函数是关于弧长的周期函数^[8,9],接着选择 Daubechies 周期化小波 D4 作为小波基进行小波变换得到小波描述子。提取前 3-

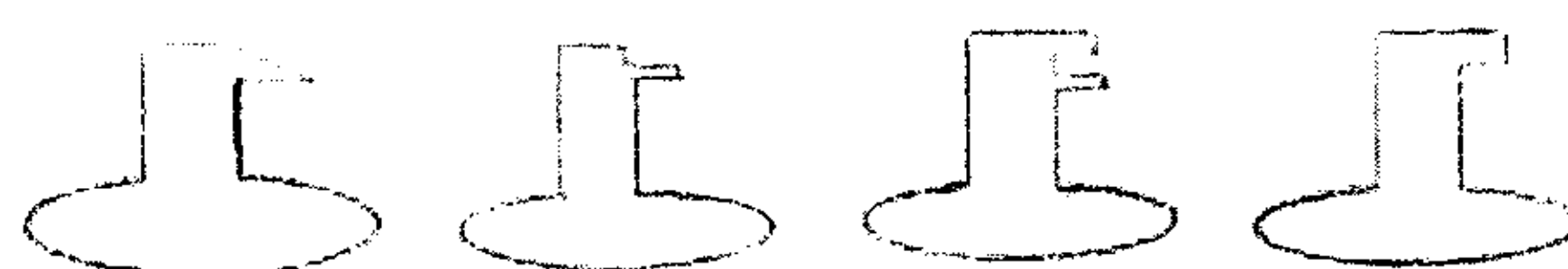


图2 实验对象

4 层小波特征来识别,用 20 个小波描述子系数表示形状,这些数据已证明能基本精确表达形状。

在进行了一些最初的试验后,多层感知网络的结构确定如下:20-10-4(20 个输入接点、10 个隐藏神经元和 4 个译码被分为 1 或 0 的输出神经元)。并且在训练阶段,向原始的样本进行旋转、平移和放大并且加入了噪音^[10]产生了新的学习模式,在实验中用了 300 个学习模式进行训练,达到了满意的效果。这样就增加了神经网络的一般化识别能力。在实验中信噪比值是从 40dB 到 3.6dB。40dB 的噪音形状几乎没有相应的影响,充满噪音的形状和有 3.6dB 噪音的形状意味着形状完全被噪音覆盖。训练后的系统通过对 100 个模式进行识别得到表 1 的数据,它表明了受干扰形状识别的结果。

表 1 加入噪音的识别数据

信噪比(dB)	识别率(%)
40	100
30	100
18	100
15	97.9
10	89.1
5	60.0

4 结 论

文中介绍了小波描述子和神经网络相结合的形状识别方法。通过实验表明该方法能够更加鲁棒地识别

和区分具有旋转、平移和尺度变化的物体。由于神经网络具有良好的容错性能与泛化能力,因此本方法对于噪声不敏感,对于信噪比大于 15dB 的形状有很好的识别能力。

参考文献:

[1] 丁险峰,吴 洪,张宏江,等.形状匹配综述[J].自动化学报,2001,27(5):678-694.

[2] Chuang Gene C H, Kuo C C J. Wavelet descriptor of planar curves: theory and application[J]. IEEE Trans Image Processing, 1996, 5(1):56-70.

[3] 杨翔英,章毓晋.小波轮廓描述符及在图像查询中的应用[J].计算机学报,1999,22(7):752-757.

[4] Pandya A S.神经网络模式识别及其实现[M].徐 勇,荆涛译.北京:电子工业出版社,1999.

[5] Navi R, Stantanu C, Subhashis B. Recognition of partially occluded objects using neural network based indexing[J]. Pattern Recognition, 1999, 32:1737-1749.

[6] 高 隽,张维勇,蒋建国,等.基于神经网络的形状识别系统及优化[J].系统工程与电子技术,1999,21(3):52-55.

[7] Vapnik V. The Nature of Statistical Learning [M]. New York: Springer, 1995.

[8] 刘宏申,秦 锋.确定轮廓形状匹配中形状描述函数的方法[J].华中科技大学学报,2005,33(4):13-16.

[9] 朱同林,彭嘉雄.轮廓形状匹配的函数小波特征方法[J].自动化学报,2001,27(6):855-859.

[10] Matsuoka K. Noise injection into inputs in backpropagation learning[J]. IEEE Trans Syst Man Cybern, 1992, 22: 436-440.

(上接第 25 页)

以便分页显示。在 Session 范围内缓存登陆用户个人信息。JSF 框架提供了方便的用户缓存实现方式,通过分别创建 ApplicationBean、SessionBean 并纳入 JSF 管理,在配置文件中设置 managed-bean-scope 属性为 application 或者 Session 实现数据缓存。

4 结束语

JSF 在基于 Java 的 Web 应用开发领域提供了一个可重用、可扩展、基于组件、工具友好的服务器端 UI (User Interface) 框架;Hibernate 完成对象和关系数据之间的映射;Spring 框架将业界广泛使用的技术集成进来,为解决 J2EE 平台技术间的使用隔阂起了重要作用。将该架构应用于网上选课系统的设计和开发,大大降低了表现层、控制层和数据访问层之间的耦合性,增强了系统的可扩展性和可维护性,同时利用缓存机制,减少了对数据库的直接访问次数,使系统具有良

好的性能。

参考文献:

[1] Hibernate org. Hibernate Reference Documentation[EB/OL]. 2005. <http://www.hibernate.org/hib-docs/v3/reference/en/html/>.

[2] Springframework. org. Spring Framework Documentation [EB/OL]. 2005. <http://www.springframework.org/documentation>.

[3] 朱庆生,葛 亮.新一代 Web 应用框架 JSF[J].计算机科学,2005(7):224-225.

[4] SHEN DEREK YANG. Put JSF to wrok[EB/OL]. 2004-06. <http://www.javaworld.com/javaworld/jw-07-2004/jw-0719-jsf-p.html>.

[5] 冯国仕,李志蜀.基于 Struts 与 Hibernate 集成架构的项目管理系统[J].计算机应用,2005(8):1884-1885.

[6] GEARY D, HORSTMANN C. JavaServer Faces 核心编程 [M].王 军,马振萍,等译.北京:电子工业出版社,2005.