

# 基于神经网络的二维不变性目标识别方法

贾财潮

(上海交通大学计算机科学与工程系, 上海 200030)

于 洵 张季涛

(西安应用光学研究所, 西安 710065)

**摘 要** 提出了一种基于多层前馈神经网络的二维不变性目标识别方法。利用傅里叶描述器提取具有旋转、平移及尺度不变性的目标形状特征。由于所识别的工业工具具有一个自由度, 它们的形状有一定的动态变化范围, 导致同一目标的形状特征矢量的不唯一性。文中采用含有两个隐层的多层前馈网络学习及识别这些特征矢量。在实验中, 对四类机械工具进行测试, 并将所提出方法与最近邻分类器进行比较。结果表明, 具有反向传播(BP)学习算法的多层前馈网络对噪音和形状特征变化具有鲁棒性, 且它还能判断未训练样本。

**关键词** 神经网络, 目标识别, 傅里叶描述器, 形状特征。

## 1 引 言

在视觉模式识别中, 由于成像距离、方向以及位置等因素的变化, 使得图像发生旋转、平移及尺度变化, 因而提取旋转、平移及尺度不变性的特征, 对目标识别是非常必要的。文献[1, 2]证明了傅里叶描述器是一种基于目标轮廓数据的全局性信息提取方法, 它的一系列系数具有不变性性质, 并应用于手写体及空中目标识别等领域<sup>[3, 4]</sup>。在现代工业环境如工厂自动化装配线上, 机器人视觉系统的一个主要任务就是要识别产品零件和机械工具。大部分机械工具具有一个自由度, 如钳子、扳子等它们的形状随着其爪角的变化而变化, 因而表征该类目标的形状特征矢量就有一定的动态变化范围。由于这些特征矢量在特征空间中占居着一个非凸区域, 要识别该类目标是一项困难的任務。统计模式识别方法如最近邻分类器等, 由于固有的缺点不能正确分类非凸集数据。一个有大于等于 2 隐层的前馈神经网络能够对任何可分非凸集合进行正确分类<sup>[5]</sup>。本文通过傅里叶描述器, 提取 4 类具有一个自由度机械工具的形状特征, 用多层前馈网络训练并识别这些特征矢量。比较了神经网络方法与广泛应用的最近邻分类器的识别性能, 并讨论了噪音对识别性能的影响。

## 2 傅里叶描述器和目标形状特征提取

为了有效地识别目标, 需要获得表示目标的特征矢量。如何从物体的图像数据中提取其

特征, 一直是目标识别的一个重要问题。由于目标的边界轮廓包含了其主要信息, 因此选择形状特征作为目标的特征矢量, 并用傅里叶描述器表达该特征。傅里叶描述器的优点是对目标具有平移、旋转及比例不变性, 且能通过去除高频分量达到显著地压缩数据的目的。

### 2.1 图像数据获取

用 CCD 摄像头采集目标以得到其灰度图像, 图像的分辨率为  $256 \times 256$ , 每幅图有 256 级灰度, 分别存储在  $256 \times 256$  的矩阵中。应用 Sobel 算子对灰度图像进行边缘提取, 进而获得其外形轮廓数据  $(x[m], y[m]) \quad 1 \leq m \leq l$ 。图 1 给出了一个钳子图像轮廓的获取过程, 图 2 表明了其随着爪角变化的轮廓图像。

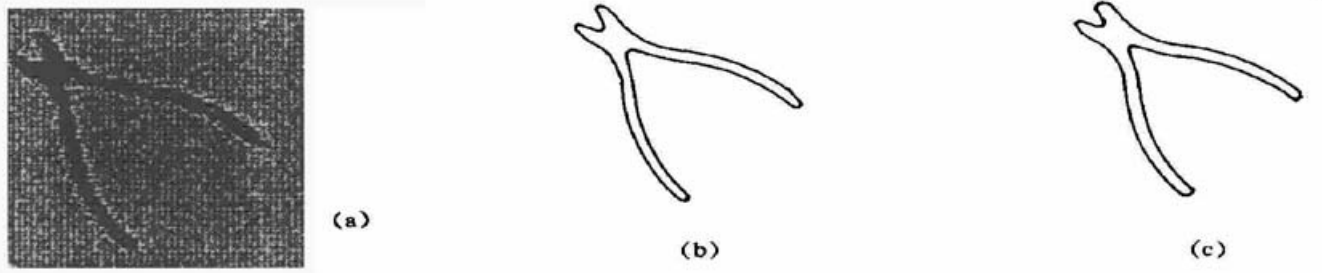


Fig. 1 Procedure for obtaining boundary image. (a) Digitized image of pliers. (b) Edge image extracted using Sobel operator. (c) Boundary image



Fig. 2 The variation of the shape of a pair of pliers along the jaw angle

### 2.2 傅里叶描述器

对边界轮廓数据  $(x[m], y[m]) \quad 1 \leq m \leq l$ , 应用离散傅里叶变换有:

$$\left. \begin{aligned} a[k] &= \frac{1}{l} \sum_{m=1}^l x[m] \exp[-jk(2\pi/l)m] \\ b[k] &= \frac{1}{l} \sum_{m=1}^l y[m] \exp[-jk(2\pi/l)m], \end{aligned} \right\} \quad 0 \leq k \leq l-1 \quad (1)$$

因为  $a[0] = \frac{1}{l} \sum_{m=1}^l x[m]$ ,  $b[0] = \frac{1}{l} \sum_{m=1}^l y[m]$ , 它们仅表示图像的中心位置信息。为了获得图像的不变性表示, 首先剔除  $a[0]$  和  $b[0]$ , 然后令:

$$\left. \begin{aligned} r[k] &= \sqrt{|a[k]|^2 + |b[k]|^2} \\ s[k] &= r[k]/r[1], \end{aligned} \right\} \quad 1 \leq k \leq l-1 \quad (2)$$

上二式中,  $|a[k]|$  和  $|b[k]|$  是复数  $a[x]$  和  $b[k]$  的模值。由傅里叶变换的幅-相关系知,  $|a[k]|$  和  $|b[k]|$  分别对  $x[m]$  和  $y[m]$  具有平移及旋转不变性。因而  $r[k]$  就刻划了二维图像的平移、旋转不变性特性, 文献[1] 进一步发展了比例不变性描述器  $s[k]$ 。

由于傅里叶系数的低频部分主要决定于图像的整体轮廓, 其高频部分仅刻划了外形的细节, 且高频部分容易退化为噪音, 因此高频部分对目标识别的作用不大, 通常可以不考虑高频分量。这样还能达到显著地压缩数据的目的。本次实验中我们取前 16 个系数作为每个目标

的形状特征矢量, 表 1 给出了图 2 中(a)、(b)、(c)、(d)四幅图像的傅里叶系数。计算表明,  $\sum_{k=1}^{16} s[k]$  已构成目标轮廓总长度  $\sum_{k=1}^l s[k]$  的 99.9%。从表中也可看出, 随着  $k$  的增大  $s[k]$  快速地衰减。

Table 1. The variation of the Fourier descriptors of pliers in fig. 2 along with the jaw angle

$k$	(a)	(b)	(c)	(d)	$k$	(a)	(b)	(c)	(d)
1	1.00000	1.00000	1.00000	1.00000	9	0.02236	0.01376	0.02324	0.01855
2	0.88149	1.00080	1.04270	0.86790	10	0.04249	0.04244	0.03523	0.04454
3	0.21087	0.28314	0.24978	0.25863	11	0.02692	0.00891	0.02458	0.01104
4	0.02507	0.04999	0.03017	0.04337	12	0.01794	0.02751	0.02668	0.02762
5	0.10293	0.08078	0.11820	0.09058	13	0.02461	0.01833	0.02519	0.02067
6	0.03276	0.05252	0.04607	0.04194	14	0.02347	0.02278	0.02218	0.02282
7	0.08743	0.06151	0.09310	0.07043	15	0.07543	0.01758	0.01071	0.01716
8	0.03408	0.04306	0.04124	0.04656	16	0.01292	0.01828	0.01861	0.01754

### 3 神经网络结构及网络训练

神经网络分类器的主要优点是它结构的鲁棒性及具有并行分布计算能力, 从而能容许目标特征有一定的动态范围并执行快速分类。本文使用了具有两个隐层的多层前馈网络结构, 用文献[6]中修改的反向传播算法进行学习。反向传播网络是应用最广泛的网络模型之一, 它能满足一般的模式识别的需要<sup>[7]</sup>。正如前面所述, 本文的特征矢量有 16 个元素 ( $s[1] \cdots s[16]$ ), 自然地网络输入层有 16 个神经元, 两个隐层神经元数分别选择为 20 和 14, 输出层有 4 个神经元, 表示要分类的 4 类目标。神经元函数采用  $s$  型函数, 且每个神经元有一阈值  $\theta_i$ , 权值调整按下式:

$$\Delta w_{ij}(k+1) = \alpha \delta_{pi}(k+1) O_{pj} + \eta \Delta w_{ij}(k)$$

式中  $\alpha$  为学习率,  $\eta$  为动量系数, 在该实验中分别取为 0.3 和 0.5,  $O_{pj}$  是输入样本为  $p$  时第  $i$  个神经元的输出,  $\delta_{pi}$  为第  $i$  个神经元的误差。

选取每个目标不同位置、不同比例及不同爪角的 10 幅图像, 4 类目标共 40 幅图像组成学习样本。图 3 给出了四种目标部分训练样本的轮廓图像。设置权重系数和神经元阈值的初始值为区间  $[-0.5, 0.5]$  的随机数。经过 2000 次学习周期后, 误差减少到 0.03 以下。网络进行识别时, 定义输出矢量为:

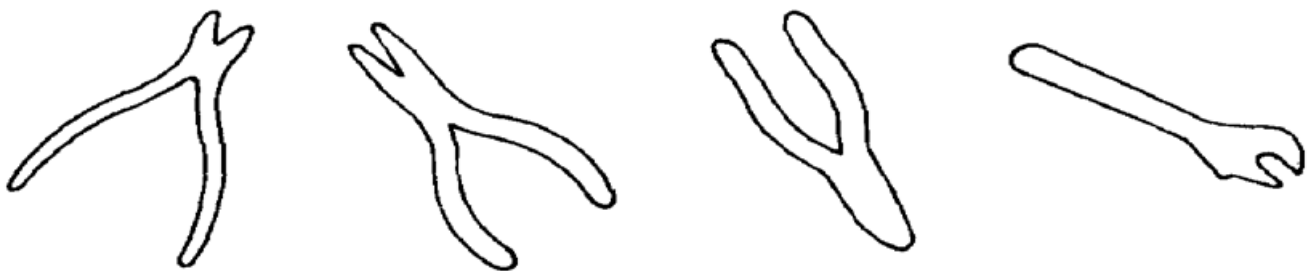


Fig. 3 Some of the boundary images of object 1, object 2, object 3 and object 4 used for neural network training

$$y_i = \begin{cases} 1 & y_i \geq 0.7 \\ 0 & y_i \leq 0.3 \\ \text{未知} & 0.3 < y_i < 0.7 \end{cases} \quad (3)$$

对测试样本若输出矢量为[ 1, 0, 0, 0], 则认为是目标 1, 若输出矢量为[ 0, 1, 0, 0] 则认为是目标 2, 其余依此类推。若  $0.3 \leq y_i \leq 0.7$ , 认为它不属于学习过的 4 类样本。

### 4 实验结果

在测试实验中, 制作了 4 类目标不同位置、不同比例及不同爪角的 80 幅图像( 每类 20 幅), 并将反向传播网络与最近邻域分类器进行了比较, 表 2 给出了测试结果。神经网络方法对 4 种目标都达到了 100% 的识别率, 而最近邻域分类器对目标 2 仅取得了 90% 的识别性能。对 20 幅原始图像添加高斯噪音( 均值为 0, 方差  $\sigma^2$  分别为 9、36 和 81), 再用 2 种方法进行测试, 结果如表 3。图 4 给出了图 1 目标的噪音图像。可以看出即使噪音有显著影响的目标图像, 神经网络方法仍达到 95% 的识别率, 最近邻分类器下降为 83.5%。进一步将 2 个未训练目标( 图 4) 的特征输入反向传播网络, 它的输出矢量分别为[ 0.04, 0.63, 0.0, 0.40] 和[ 0.12, 0.32, 0.02, 0.08], 按照(3) 式, 反向传播网络认为该类目标是未学习过的样本。而最近邻分类器无法判断未训练的目标。

Table 2. The accuracy of recognition under noise-free environment.

object	1	2	3	4
BP	100%	100%	100%	100%
nearest-neighbor	100%	90%	100%	95%

Table 3. The accuracy of recognition under noise environment.

variance ( $\sigma^2$ )	BP	nearest-neighbor
9	98.7%	93.0%
36	97.5%	93.5%
81	95.0%	83.5%

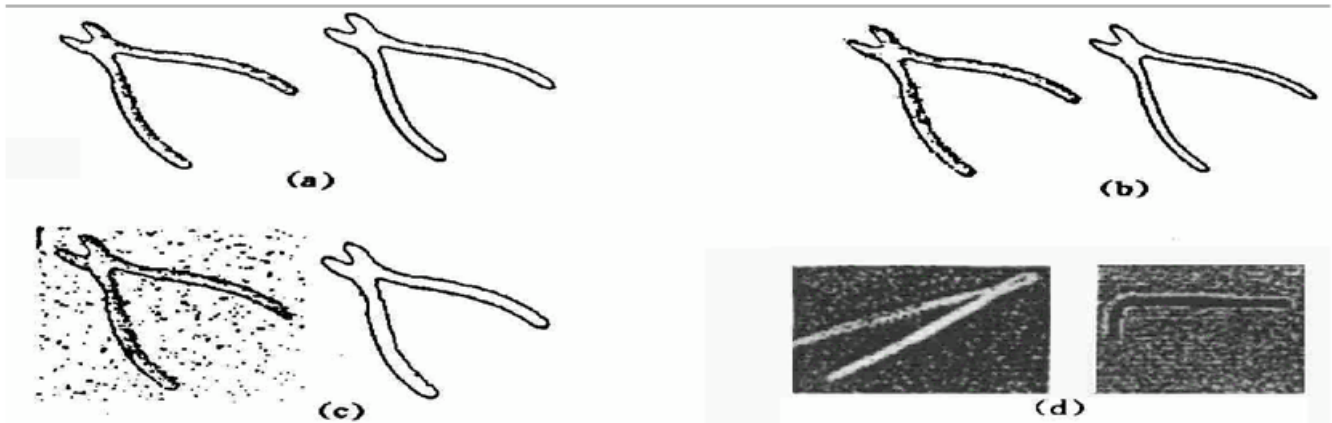


Fig. 4 The edge image and the boundary image in fig. 1, when Gaussian random noise is added.

(a)  $\sigma^2 = 9$ , (b)  $\sigma^2 = 81$ , (c)  $\sigma^2 = 36$ . The images of untrained objects shown in fig. (d)

将本文方法与 Casasent<sup>[8]</sup> 提出的神经网络目标识别方法进行对比。它采用光学相关技术处理感兴趣的图像区域, 进而提取其不变性特征, 然后用神经网络分类器对 3 类飞机达到了 98.7% 的识别率。这与本文识别方法达到了相当的识别精度。由于它使用自适应聚类神经网络分类器, 采用“胜者全赢”规则, 这样每当输入一个目标样本, 该网络总要找到激励最大的一个隐层神经元( 对应于某一分类类别), 并将其它隐层神经元输出抑制为零。因此该方法不能判别未训练目标。

**结 论** 本文提出了一种神经网络目标识别方法, 利用傅里叶描述器提取二维目标的不变性形状特征。对 4 种目标的图像样本测试表明, 该方法无论对变化的图像特征, 还是对噪音污染的目标都具有鲁棒性。它的分类结果优于传统的最近邻分类器, 且达到了较高的识别率。进一步该方法还可判断未训练目标。

### 参 考 文 献

- [1] Shridhar M, Badreldin A. High accuracy character recognition algorithm using Fourier and topological descriptors. *Pattern Recognition*, 1984, **17**(5) : 515~ 524
- [2] Persoon E, Fu K S. Shape discrimination using Fourier descriptors. *IEEE Trans. SMC*, 1977, **7**(3) : 171~ 179
- [3] Wallace T P, Wintz P. An efficient three-dimensional aircraft recognition algorithm using normalize Fourier descriptors. *Computer Graphics and Image Processing*, 1980, **3**(2) : 99~ 126
- [4] Venugopal K P. Invariant recognition of 2D objects using Alopex neural network. *Proc. SPIE*, 1992, **1709** : 182~ 190
- [5] 郑君里, 杨行峻. 人工神经网络. 北京: 高等教育出版社, 1992.
- [6] Vogl T P, Mangis J K, Riglev A K *et al.*. Accelerating the convergence of the back-propagation method. *Bio. Cyber.*, 1988, **59**(4) : 257~ 263
- [7] Gupta L. A neural network approach to robust shape classification. *Pattern Recognition*, 1990, **23**(6) : 563~ 568
- [8] Casasent D P. An optical correlator feature extractor neural net system. *Opt. Engng.*, 1990, **31**(5) : 971~ 978

## A Neural Network-Based Approach for Invariant Recognition of 2-D Objects

Jia Caichao

(Department of Computer Science and Engineering, Shanghai Jiaotong University, Shanghai 200030)

Yu Xun      Zhang Jitao

(Xi'an Institute of Applied Optics, Xi'an 710065)

(Received 12 January 1998; revised 15 July 1998)

**Abstract** A neural network-based approach is proposed for invariant recognition of two-dimensional objects. The industrial tools to be recognized have one degree of freedom, the dynamic range of their shapes leads the feature vector not uniquely defined even for a single object. The Fourier descriptors of objects boundary are taken as the features being invariant to translation, rotation and scale changes. A multilayer feedforward neural net with two hidden layer classifiers is utilized. The experimental studies involving four sorts of mechanical tools are carried out. The performance is compared to a nearest neighbor rule. It is shown that the approach is robust to not only noisy but also varying feature vector, and the neural network can recognize some untrained objects.

**Key words** neural networks, target recognition, the Fourier descriptors, shape features.