

## 基于径向矢量的不变性形状分类方法\*

叶芩芸 胡栋梁 戚飞虎

(上海交通大学计算机科学与工程系, 上海, 200030)

TP391.41  
TP18

**A** **摘要** 介绍一种利用径向矢量提取形状特征的方法, 着重于分析二维图形的不变性, 提出一种带有方向因子的径向矢量描述. 该矢量以图形边界弧长为自变量, 完整地刻画了图形特点, 克服了非凸图形识别中存在的多义性. 以归一化的特征矢量为输入, 采用三层神经网络为分类器, 在字符及军事目标不变性识别中取得了令人满意的结果.

**关键词** 形状分类, 径向矢量, 不变性识别, 神经网络.

模式识别

## 引言

形状的不变识别一直是模式识别和计算机视觉领域的重要部分. 对二维图形, 形状识别的不变性主要表现为位移、方向、尺度不变性. 大部分形状不变性识别方法都以图形的边界表示为基础.

Dubois<sup>[1]</sup>等人首先提出使用一个以形心为基点、等旋转角度间隔采样的矢量序列近似描述图形的边界, 该矢量的模由形心各边界点的直线距离决定. 但该方法仅限于凸性图形, 因为非凸图形有可能在某些矢量方向上有多个边界点, 造成一个(个角度)对多(个矢量)的情况, 从而无法用上述的简单序列完全描述图形. 为克服这个缺陷, Dubois 提出了展开径向矢量的方法, 即将出现一对多的径向矢量序列按边界跟踪方向依次展开, 以获得一一对应的序列. 尽管如此, 由于跟踪中没有记录相邻径向矢量的相角变化方向, 对不同图形有可能得到相同的展开曲线描述, 造成图形描述的歧义. 此外, 该模型不适用于形心落在图形边界外或非闭合图形的情况, 并且难于处理边界的毛刺. Gupta<sup>[2]</sup>采用类似的径向投影序列作为多层感知网络的输入分类特征, 将等分角度间隔改成沿边界各像素跟踪的方法, 较 Dubois 更为精确, 但仍未解决歧义性问题. K. Tsinkolias<sup>[3]</sup>采用了相同的特征描述方法用于统计模式识别. N. R. Pal<sup>[4]</sup>在前人的基础上提出用多边形逼近方法, 利用边界点间角度变化特性作为多层神经网络输入分类特征. 上述各种描述方法都采用了径向矢量模型, 但都无法避免该方法的固有缺陷, 在应用上受到很大限制. 本文针对这一问题提出了有效的解决方法.

## 1 原始特征提取

在前人工作的基础上, 本文对径向矢量(RV)定义一个方向因子, 使径向矢量序列的各分量除了静态的模长和本身方向以外, 还拥有动态的相角变化方向, 见图1. 称之为动态, 是因为它以跟踪序列号为自变量, 也可认为与跟踪时间有关. 这个方向因子由边界的跟

\* 国家自然科学基金资助项目

本文1995年6月5日收到, 修改稿1995年10月31日收到

踪方向决定,它确保了边界描述的唯一性,避免了特殊图形造成的歧义.同时,本文提出的方法在形成径向投影序列时,自变量是跟踪的真实弧长,避免了等角度投影方法描述非凸图形带来的一对多的情况,使图形特征保证唯一.

一般来说,大部分图形可以通过其周边轮廓加以区分,事实上,许多算法正是利用图形的边界信息作为分类特征.本文亦采用径向矢量序列来描述图形边界.

设二维图形的边界可用一闭合曲线(如图2)表示为.

$$L(r)=0,$$

这里  $r$  为二维矢量,表示二维空间中的点.设  $C$  为该曲线的形心,用  $r_c$  表示.定义矢量集合

$$RV_L = \{V | L(V + r_c) = 0\}, \quad (1)$$

则曲线上任一点  $R$  可以表示  $R = V + r_c$ . 当图产生平移  $r_T$ , 即  $L_T(r) = L(r - r_T)$  时,新的形心  $r_{CT} = r_c + r_T$ , 对应的矢量集合变为

$$RV_{LT} = \{V | L_T(r + r_{CT}) = 0\} = \{V | L(V + r_c) = 0\} = RV_L. \quad (2)$$

可见,对于集合  $RV_L$ , 平移不变性是显然的.  $RV_L$  即所谓的径向矢量集,如对其沿图形边界方向离散化,可得到一个径向投影序列  $(RV) \{RV_i\}_{i=0}^N$ , 且满足  $L(RV_i + r_c) = 0, RV_0 = RV_N, |RV_{i+1} - RV_i| = l_i, i = 0, 1, \dots, N-1$ .



图1 径向矢量表示的边界  
Fig. 1 Shape representation  
by radius vectors

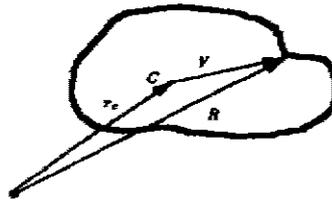


图2 闭合曲线的RV表示  
Fig. 2 RV representation  
for closed shapes



图3 非闭合曲线表示方法  
Fig. 3 Method describing  
open shapes

RV 序列可以成功地描述一些凸或非凸图形,然而在描述某些形心处于边界以外的图形时却出现了问题(对于非闭合曲线,可将其视为宽度很小的闭合曲线,如图3). 下面的例子将说明 RV 序列存在的问题:图4和5是两个不同图形,但得到的 RV 序列是相同的,这将造成判断的歧义,它是由图形的对称性造成的.可以证明,当非闭合曲线  $L$  关于形心中心对称时,其衍生的闭合曲线所对应的 RV 序列可能与另一闭合曲线  $L'$  相同,只要  $L'$  是由  $L$  及其关于某过形心的直线轴对称图形组合而成.这里对称轴过  $L$  的两个端点,且图形组合时不出现连续重叠的片断.当然,这两种图形的 RV 跟踪路径必然是不同的,这就需要一个方向因子对它们进行区分.

如图6所示,我们定义方向因子为

$$\alpha_i = \text{sgn}[(RV_i \times P_i) \cdot k], \quad (3)$$

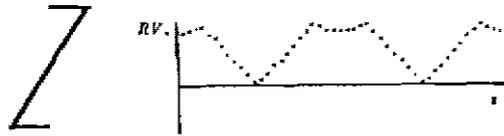


图4 图形Z的RV表达  
Fig. 4 RV representation for shape Z

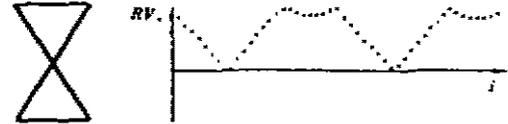


图5 图形X的RV表达  
Fig. 5 RV representation for shape X

其中  $P_i = RV_{i+1} - RV_i$ ,  $i, j, k$  分别为笛卡尔坐标系中沿  $x, y$  和  $z$  轴的单位矢量. 从以上的定义可知  $\alpha_i$  的取值范围是  $-1, 0, 1$ .

当将 RV 的模值乘上因子  $\alpha_i$  后, 考虑前面的例子, 可以看出加权的 RV 序列将是不同的. 这样就保证了原始特征与图形间关系的唯一性. 下面我们简单证明: 不同的加权 RV 序列对应不同的闭合曲线, 相同的加权 RV 序列对应平移的闭合曲线. 这就是说, 可由加权 RV 序列唯一地恢复一个闭合曲线. 如图 7 所示, 由于跟踪提取 RV 序列时, 相邻 RV 矢量

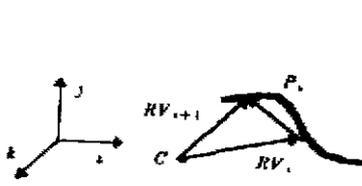


图6 方向因子  $\alpha_i$  的定义  
Fig. 6 Definition of the directional factor  $\alpha_i$

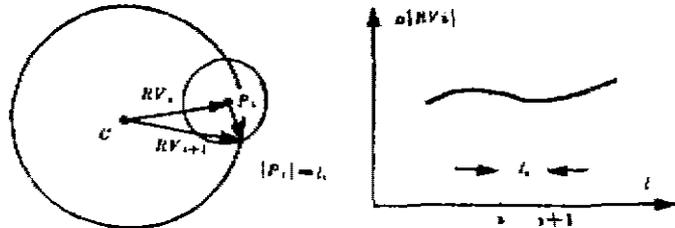


图7 加权 RV 序列及其唯一对应闭合曲线  
Fig. 7 Weighted RV sequence and its uniquely corresponding closed shape

的间距  $L_i$  是固定的, 所以确定了  $RV_i$  以后, 矢量  $RV_{i+1}$  的方位将限于  $RV_i$  端点为圆心、 $L_i$  为半径的圆上, 另一方面,  $RV_{i+1}$  又必然处于以形心  $C$  为圆心,  $|RV_{i+1}|$  为半径的圆上. 如两圆相交, 可利用  $RV_{i+1}$  的加权因子在两个交点中唯一确定  $RV_{i+1}$  对应的真实边界点; 如两圆相切,  $RV_{i+1}$  的选择也是唯一的; 考虑到由 RV 序列恢复一闭合曲线的过程其实是从该闭合曲线得到 RV 序列的逆过程, 两圆相离的情况是不存在的. 由以上分析可知, 一个加权的 RV 序列对应唯一的闭合曲线.

至此, 我们已通过加权的 RV 序列获得了图形的原始特征, 即  $\{RV_i' = |\alpha_i RV_i|\}_{i=0}^N$ . 该序列克服了特殊图形带来的歧义性, 适用于任意闭合及非闭合曲线.

## 2 分类

利用加权 RV 序列可以实现对各种图形的分类. 从前面的分析可以看出, 图形模式的平移不会引起 RV 序列的变化, 图形的旋转将反映在 RV 序列的循环移位上, 而图形的尺度变化将仅仅影响 RV 序列的幅度变化. K. Tsinkolias 等人曾利用图形的主轴来确定 RV 序列产生的起点, 以消除图形旋转的影响. 但这种方法不适用对称图形, 而且仍然存在上述的歧

义问题. H. Kwon<sup>[5]</sup>建议用高阶神经网络克服图形的旋转,但训练费时,计算量大. 本文通过简单的自相关运算和尺度归一化可以得到位移、方向、尺度不变的特征序列,这为后面的分类提供了可靠的依据.

我们定义  $R_k = \sum_{i=0}^{N-1} RV'_i \cdot RV'_{i+k \bmod N}$  ( $k = 0, 1, \dots, N$ ), 并进行幅度的归一化, 得到

$$R'_k = (R_k - \bar{R}) / \sigma, k = 0, 1, \dots, N; \tag{4}$$

其中

$$\bar{R} = \frac{1}{N} \sum_{i=0}^{N-1} R_i, \sigma = \frac{1}{N-1} \sqrt{\sum_{i=0}^{N-1} (R_i - \bar{R})^2}.$$

由此得到具有位移、方向、尺度不变性质的特征表达.

我们采用在模式分类方面有成功应用的 NIN-NHID-NOUT 结构三层 BP 网络进行模式分类. 其中 NIN 是输入神经元个数, 取为 RV 自相序列的归一化长度; NOUT 是输出神经元个数, 我们取为分类类别数; NHID 是隐层神经元个数. 这里网络训练时网络的权值和阈值仅对误分类模式按如下方式进行调整:

$$\left. \begin{aligned} \Delta W_{ij}(t+1) &= -\eta \frac{\partial E}{\partial W_{ij}}(t) + \alpha W_{ij}(t), \\ \Delta \theta_j(t+1) &= -\eta \frac{\partial E}{\partial \theta_j}(t) + \alpha \theta_j(t). \end{aligned} \right\} \tag{5}$$

进行训练的样本集包括标准样本和少量噪声样本. 为提高训练的速度和质量, 训练时在完成理想模式(不含噪声)的基础上逐渐增加噪声样本. 在训练中我们采用变速率学习方案, 即学习速率根据输出误差的变化率作相应调整:

$$\eta = C \cdot \ln \left| \frac{E(t)}{E(t+1)} \right|, \tag{6}$$

其中 C 是介于 0 和 1 之间的常数. 实验中我们选用  $\alpha=0.8, C=1.0$ . 图 8 示出实验误差下降曲线, 可见变速率的学习方案使网络快速平稳地收敛.

### 3 在字符与军事目标识别中的应用

字符识别是模式识别的重要分支, 字符的不变性识别是该领域中一个令人感兴趣的方向. 由于字符可由其边界的不同唯一地确定, 文中的边界描述方法也可以在解决字符不变性识别问题上得到应用. 此外, 我们还对军事目标作了识别实验.

图 9 为部分飞机和字符的训练样本和测试样本. 训练样本集包括标准位置、尺寸

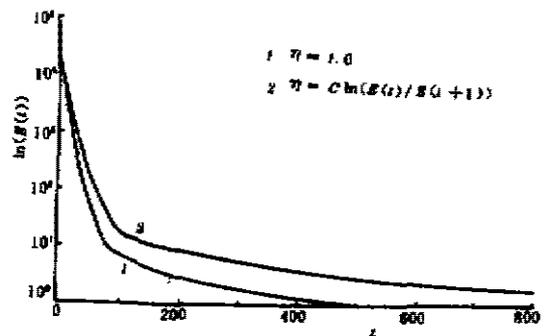


图 8 不同学习速率误差下降曲线比较  
Fig. 8 Comparison of error curves between different learning rates

和方向的标准样本和少量噪声样本(字符  $10 \times 3$  个,飞机  $5 \times 3$  个),测试集包括任意位置、尺寸和方向的样本(字符  $10 \times 20$  个,飞机  $5 \times 20$  个). 实验中飞机图像的大小为  $64 \times 64$ , 字符图像由 HP- I C 扫描仪以 200DPI 的分辨率扫入, 尺寸任意. 图 10 是使用归一化的加权 RV 自相关序列表达不同方向字符的例子.

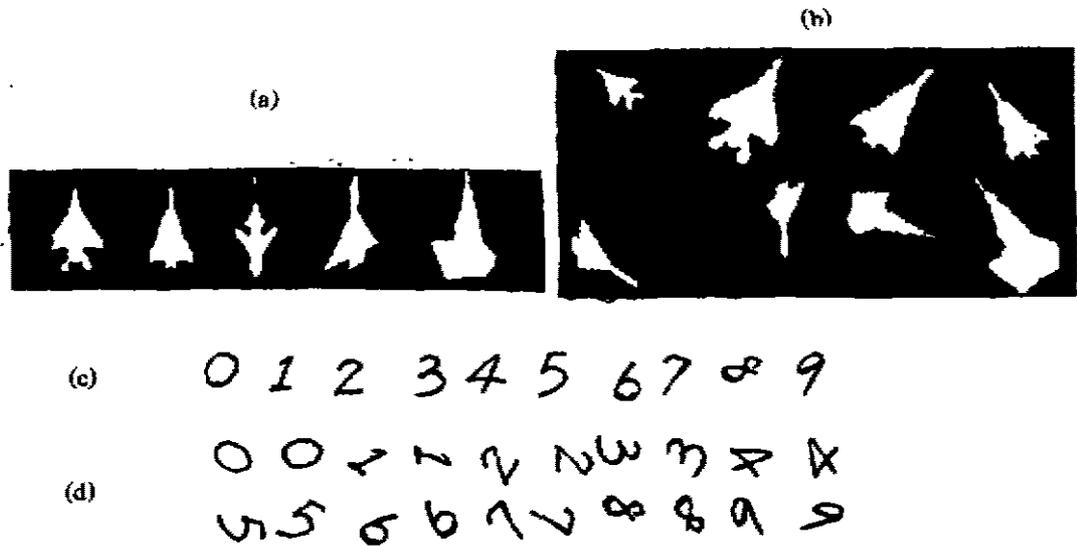


图 9 (a)部分飞机训练样本,(b)部分飞机测试样本,

(c)部分字符训练样本,(d)部分字符测试样本

Fig. 9 (a) Training set of military targets (part of airplanes), (b) testing set of military targets (part of airplanes), (c) training set of characters (part), (d) training of characters (part)

实验结果表明,对训练样本,识别率可达 100%,而对测试样本,字符识别率为 87%,飞机为 96%. 这一方面是由于图像预处理工作不够完善,另一方面是数字 6 和 9 在旋转不变性意义下的相似性造成的.

#### 4 结语

本文指出了传统径向矢量描述方法的固有缺陷,在其基础上提出了一种不变性模式识别方法,该方法可用于任意形状的曲线(包括非凸图形和形心落在图形边界以外的情况). 我们在传统的 RV 矢量序列中引入了一个方向因子,消除了 RV 序列描述某些特殊图形时产生的歧义. 后续的自相关处理和归一化使得到的特征序列具有不变性特性. 在分类过程中,我们采用了变速率学习的三层 BP 网. 字符和军事目标识别的实验结果表明,文中提出的特征提取方法可有效地解决不变性模式识别问题. 需指出的是:本文工作重点是指出并解决应用广泛的径向矢量方法的缺陷与不足,使之成为唯一的形状表达,虽然在实验中本文算法给出了令人满意的结果,但在构造实用系统时,作者建议采用多种特征识别算法集成的方案,以期达到良好的实用效果.

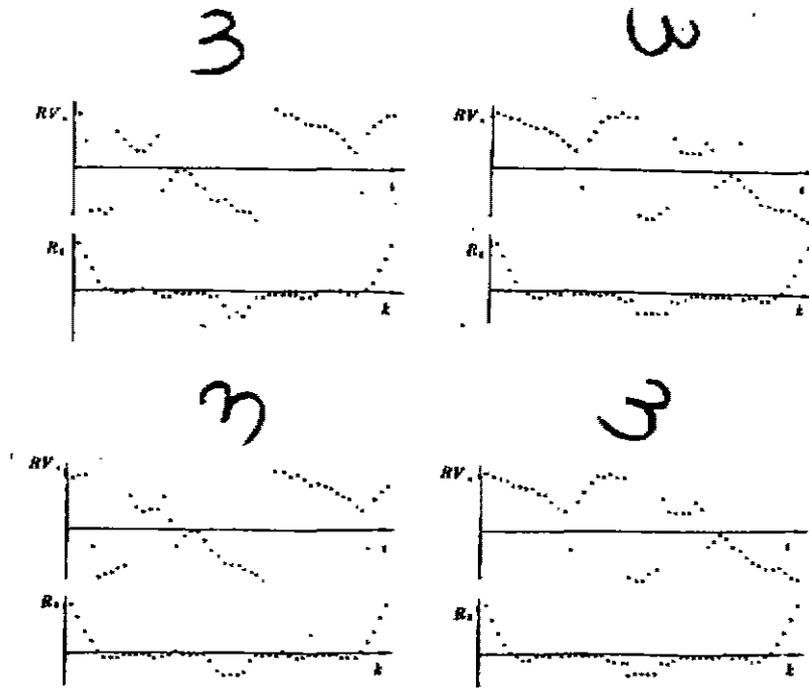


图 10 不变性特征  
Fig 10 Invariant features

#### 参考文献

- 1 Dubois S R, et al. *IEEE Trans. PAMI.*, 1986, 8(1): 55
- 2 Gupta L, et al. *Patt. Recog.*, 1990, 23: 563
- 3 Tsirikolias K, et al. *Patt. Recog.*, 1993, 26: 877
- 4 Pal N R, et al. *Patt. Recog.*, 1993, 26: 543
- 5 Kwon H, et al. *Proc. of IJCNN92*, 1992, 11: 213

## RADIUS VECTORS BASED INVARIANT SHAPE CLASSIFICATION\*

Ye Xiangyun Hu Dongliang Qi Feihu

*(Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200052, China)*

**Abstract** A feature extracting method based on radius vectors was introduced. By using a directional factor, the proposed feature gives a complete description of the shapes, and it eliminates the ambiguity that may occur in the traditional shape representation of concave shapes. The normalized features can be used for final classification. With the help of a three-layer BP neural network, satisfactory results in recognizing characters and military objects were carried out.

**Key words** shape classification, radius vector, invariant recognition, neural network.

---

\*The project supported by the National Natural Science Foundation of China