

智能型目标识别系统*

沈定刚 戚飞虎

(上海交通大学光纤技术研究所, 上海 200052)

李春茂

(香港科技大学计算机系, 香港)

摘要 介绍了图像目标识别技术中的图像分割、不变性参数提取和目标分类. 利用图像目标的均匀性和相应知识自适应地分割和提取图像目标, 被提取的每个图像目标的不变性参数由归一化过程和 Zernike 矩提取, 并利用 MPNN 模型将图像目标分类. 实验结果表明该识别系统能识别光照不均匀或复杂背景下的图像目标.

关键词 目标识别, 图像分割, 不变性参数提取.

引言

图像目标识别是图像处理和模式识别中的一个重要课题^[1], 就整个识别过程而言, 可将此过程分解成目标信号的预处理、目标信号的特征提取与选择、目标特征的自动分类. 当然, 由于使用方法不同, 识别过程所经历的阶段也不同, 例如, 有些人工神经网络方法^[2,3], 将特征选择和目标分类合并在一起, 由神经网络统一完成.

在预处理阶段首先对接受到的图像目标进行噪声滤除和图像分割, 以便获得用于目标特征选择与计算的二值图像. 噪声滤除通常使用中值滤波器完成, 主要因为它能保护边缘, 使其不受滤波影响. 图像分割是图像处理中的关键, 并影响后续处理, 现有的图像分割方法可分为灰度阈值分割法、迭代像素法、基于表面的分割法和基于模糊集合理论的分割法^[4], 但各种分割方法均存在着局限性. 本文采用目标知识来分割非均匀光照或复杂条件下的图像目标.

特性提取是预处理与分类器之间的桥梁, 选择与抽取能充分反映目标基本形状与属性的特征, 是图像目标识别的关键问题之一, 其有效性对后续处理至关重要. 特征提取的要求是: 同一类目标的特征相似性大, 不同目标的特征差异性大. 由此可见, 研究不变性特征是重要的, 本文利用归一化过程和 Zernike 矩提取目标的不变性特征.

就分类方法而言, 可分成传统的模式识别方法、人工智能方法和神经网络方法. 神经网

*国家攀登计划认知科学(神经网络)重大关键项目基金资助项目
本文 1994 年 7 月 11 日收到, 修改稿 1994 年 9 月 20 日收到

络因其具有并行处理、自学习和高容错性,使其对不确定模式识别具有独到之处.本文利用 MPNN (Modified Probabilistic Neural Network)^[5]来识别图像目标.

1 基于目标均匀性和有关知识的图像分割

图像的阈值分割是图像分割中研究和应用最多的方法,30 余年来,人们提出了多种阈值选取方法.但是,对于非均匀光照图像,仅仅以灰度值为标准来分割图像是完全不够的.同时,对于一些背景与目标间灰度较接近的图像,阈值的小量变化(即使是 1)也将导致图像分割质量相差甚大.因而,独立考虑图像灰度的分割方法是不适用于复杂图像分割的.

本文利用图像目标的均匀性和有关目标的知识作图像分割的依据.首先,在图像平面中以水平扫描顺序检测每个未被搜索像素的区域生长情况,由此而生长成的区域具有均匀性,即该区域中的每一个点至少与其 $(2r+1) \times (2r+1)$ (例如 $r=1$)邻域内的一个像素间的灰度差小于阈值 TH_1 (例如 $TH_1=2$).已被生长的像素点,其灰度值应赋为-1,使得下一次区域生长时不会再去搜索这些点.其次,利用目标的有关知识(大小范围、周长面积比范围等)判别被生长区域形成的点集是否为潜在目标,如果是潜在目标,则赋该区域中的像素灰度值为-2(-2 代表目标),否则,区域中的像素保持为-1.不断重复以上过程(灰度值小于 0 的点不成为搜索对象),直至水平扫描完成.这样,潜在目标为图像中灰度值为-2 的像素,将这种点赋为 255;同时将灰度值为-1 的像素赋灰度为 0,由此完成了图像分割任务.整个图像分割过程的算法描述如下.

设图像大小为 $M \times N$, $f(i, j)$ ($0 \leq i < M, 0 \leq j < N$)表示像素 (i, j) 的灰度值, P_1 为指向路径堆栈 $\{Road-X(P_1), Road-Y(P_1)\}$, P_2 为目标堆栈 $\{Obj-X(P_2), Obj-Y(P_2)\}$ 的指针, P_3 为指向边界堆栈 $\{Edge-X(P_3), Edge-Y(P_3)\}$ 的指针.图像分割算法包括:

(1)平滑待分割图像.

(2)赋 $i=0, j=-1; P_1=-1, P_2=-1, P_3=-1$.

(3) $j++$;若 $j=M$,则 $i++, j=0$;若 $i=M$,转向(9).

(4)若 $f(i, j) < 0$,则转向(3);否则, $ii=i, jj=j$.

(5)检测 $f(ii, jj)$ 与其邻域 $(2r+1) \times (2r+1)$ 内 $f(ii+k, jj+l)$ ($0 \leq ii+k < M, 0 \leq jj+l < N, -r \leq k, l \leq +r$)间的灰度差,如果灰度差小于 TH_1 (例如 $TH_1=2$),则 $\{f(ii+K, jj+L)=-1$ (表明已搜索过), P_1++ , $Road-X(P_1)=ii+K$, $Road-Y(P_1)=jj+L$ (记录搜索路径); P_2++ , $Obj-X(P_2)=ii+K$, $Obj-Y(P_2)=jj+L$ (记录可能组成目标的位置)};如果灰度差大于 TH_2 (例如 $TH_2=3$),则可能存在边界,记录其位置并赋负的灰度值, $\{f(ii+K, jj+L)=-1, P_3++$, $Edge-X(P_3)=ii+K$, $Edge-Y(P_3)=JJ+L$ };

(6)如果路径堆栈未空,则从路径堆栈中释放出一个地址,即 $\{ii=Road-X(P_1), jj=Road-Y(P_1), P_1--\}$,并转向(5);否则,转向(7).

(7)利用有关目标的知识,判断由潜在目标堆栈形成的点集是否可能为目标,如果可能,则该栈指向的像素灰度值赋为-2;否则,保持为-1.

(8)恢复边界堆栈 $\{Edge-X(P_3), Edge-Y(P_3)\}$ 所指向像素为其原来灰度值,并转向(3).

(9)将灰度为-2的像素置灰度值为255,其余为0.

完成以上算法操作后,潜在图像目标在图像平面内已转化成二值图像.如果图像目标本身比较均匀,那么利用以上算法就可找到整体目标,并由目标堆栈确定,即可将目标从图像平面中提取出来.如果目标本身非均匀(可能因光照原因使得目标由几种灰度的团块组成),则利用上述算法不能提取整体目标,但对二值化图像再作以上算法的操作,因目标已由灰度为255的像素组成,整体性的团块目标就可被提取出来.

利用上述算法分割图像的结果如图1和2所示,图中(a)是原始图像、(c)是二值化图像、(d)是从图像中提取出来的目标.如果利用阈值分割法分割图1和2的图像,其结果会很差.我们选择0~255中的任一个灰度值作为阈值,其分割结果仍让人无法接受,图1和2中的(b)是我们认为效果最好的阈值分割结果.

2 不变性参数提取

使用图像分割法从图像平面中提取出来的目标,其位置、大小和方向 PSR (Position Scale Rotation) 往往与训练目标不同,因而进入网络分类之前,提取目标的 PSRI (PSR Invariant) 参数是完全必要的.提取 PSRI 参数的过程为:先进行目标位置和大小归一化,再用 Zernike 矩提取旋转不变性参数.

设 $g(x)$ 是从图像平面中提取的目标,其大小为 $R \times L$. 其实,因目标的大小不一致,使得 R 和 L 随每个目标而变化.目标 $g(x, y)$ 的中心坐标点可由它的几何矩来表示,即

$$X_0 = M_{10} / M_{00},$$

$$Y_0 = M_{01} / M_{00},$$

这儿 $M_{pq} = \sum_{x=0}^{R-1} \sum_{y=0}^{L-1} x^p y^q g(x, y)$ 是几何矩.

衡量目标 $g(x, y)$ 大小的参数为目标零阶几何矩 M_{00} , 假设大小规格化后的图像,其零阶几何矩均为 M_{00}' , 则目标 $g(x, y)$ 相对标准样本的大小因子为 $d = \sqrt{M_{00} / M_{00}'}$. 设目标位置、大小归一化后的图像函数为 $g^R(x, y)$, 其大小统一成 64×64 , 则目标的位置和大小归一化过程可归结为以下函数的映射式(具体过程可参考文献[7]):

$$g^R(x, y) = g(ax + X_0, ay + Y_0).$$

其中 $0 \leq x, y < 64, 0 \leq ax + X_0 < R, 0 \leq ay + Y_0 < L$.

若能计算 $g^R(x, y)$ 的旋转不变性特征,就可完成对目标 $g(x, y)$ 的 PSRI 参数提取工作,本文使用 Zernike 矩来提取旋转不变性参数. Zernike 矩是一种正交矩,因其具有旋转不变性和容易构造高阶矩等优点已被广泛应用.图像 $g^R(x, y)$ 的 Zernike 矩是该图像在一组正交多项式上的投影,而这组多项式 $\{V_{nm}(x, y)\}$ 在单位圆内 ($x^2 + y^2 \leq 1$) 是正交的,即

$$\iint [V_{nm}(x, y)] \cdot V_{pq}(x, y) dx dy = \begin{cases} \pi / (n + 1), & n = p, m = q \\ 0, & \text{其它} \end{cases} \quad (1)$$

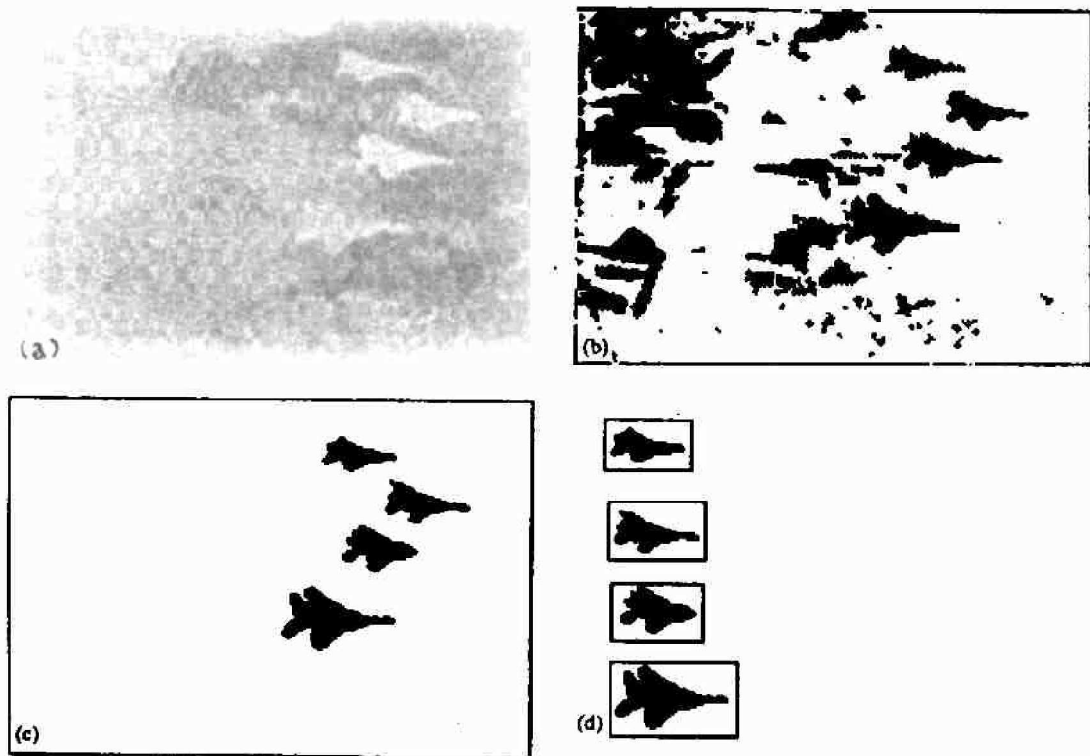


图 1 基于知识的图像分割结果 1

(a)原始图像,(b)阈值分割结果,(c)本方法的分割结果,(d)提取的目标

Fig. 1 Knowledge-based image segmentation result 1 (a)gray image, (b)segme-tation by threshold method, (c)segmentation by the proposed method, (d)objects extracted

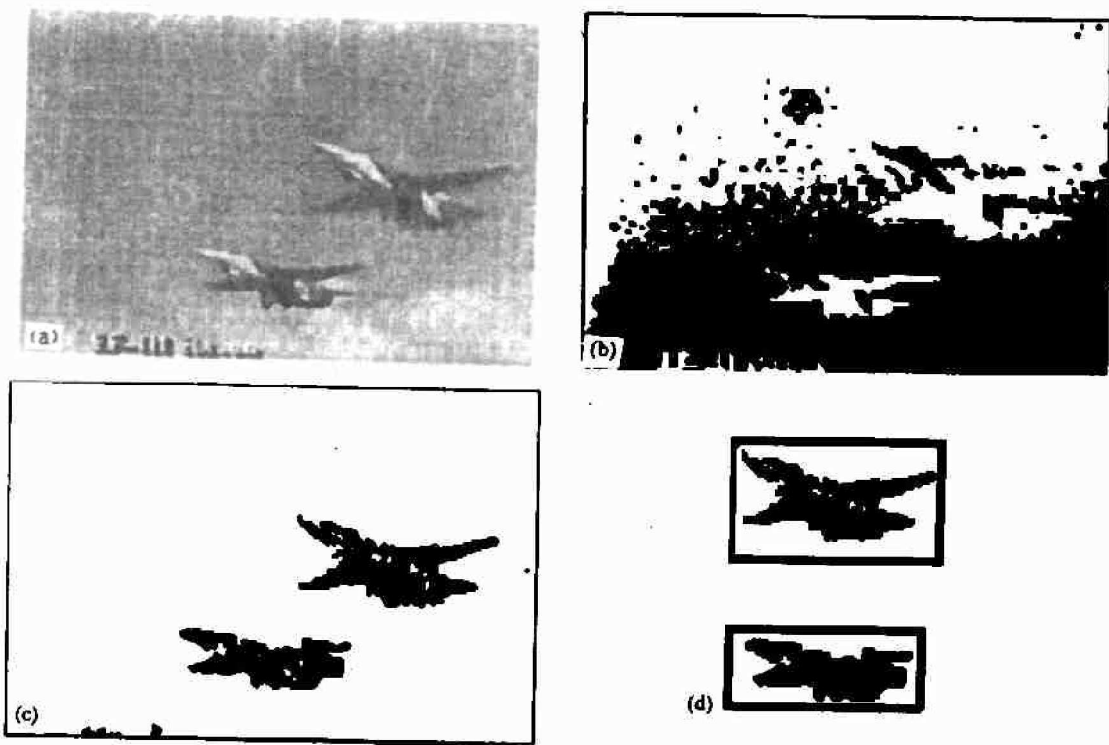


图 2 基于知识的图像分割结果 2

(a)原始图像,(b)阈值分割结果,(c)本方法的分割结果,(d)提取的目标

Fig. 2 Knowledge-Based image segmentation result 2. (a)gray image, (b)segme-tation by threshold method, (c)segmentation by the proposed method, (d)objects extracted

式(1)中, $V_{nm}(x, y)$ 的形式为 $V_{nm}(x, y) = R_{nm}(x, y) \cdot \exp(jm\theta)$, 这里 n 是正整数或 0, m 是负或正整数, 满足 $n - |m| = \text{偶数}$, $R_{nm}(x, y)$ 是径向多项式, 其表达式为

$$R_{nm}(x, y) = \sum_{s=0}^{(n-|m|)/2} (-1)^s \frac{(n-s)!}{s! \left(\frac{n+|m|}{2} - s\right)! \left(\frac{n-|m|}{2} - s\right)!} (x^2 + y^2)^{\binom{n-s}{2}} \quad (2)$$

重复率为 m 的 n 阶 Zernike 矩定义为

$$A_{nm} = \frac{n+1}{\pi} \iint g^R(x, y) [V_{nm}(x, y)]^* dx dy \quad (3)$$

Zernike 矩的不变性证明可参见文献[5].

Zernike 矩中低阶代表图像目标的低频信息, 高阶表示目标的高频信息, 但是高阶矩所表示的图像目标之细节易于受噪声影响, 所以具体使用时取 Zernike 矩的前 D ($D=12$) 个低阶矩作为 PSI 目标 $g^R(x, y)$ 的旋转不变性参数. 由于 Zernike 矩采用正交基, 使得所提取的特征相关性小, 抗噪声能力强.

3 MPNN 模型用于特征分类

MPNN 模型的结构如图 3 所示, 设 \vec{S}_{K_i} 代表第 i 类模式的第 K 个训练样本, 样本维数为 D , 其中 $1 \leq i \leq \text{NUM}$, $1 \leq K \leq m_i$; 这儿 NUM 是模式类别数, m_i 代表第 i 模式的训练样本总数. 设 \vec{Z}_i 为第 i 类模式的均值, 即

$$\vec{Z}_i = \frac{1}{m_i} \sum_{K=1}^{m_i} \vec{S}_{K_i}, \quad 1 \leq i \leq \text{NUM} \quad (4)$$

O_i 表示网络的输出值, 采用如下作用函数表示神经网络的响应:

$$O_i = \exp\left[-A_i (\vec{X} - \vec{Z}_i)^T \begin{bmatrix} \sigma_{1i}^2 & & & \\ & \sigma_{2i}^2 & & \\ & & \dots & \\ & & & \sigma_{Di}^2 \end{bmatrix}^{-1} (\vec{X} - \vec{Z}_i)\right], \quad (5)$$

$$\epsilon(\vec{X}) = -\vec{B}^T \cdot \vec{O}. \quad (6)$$

这儿 $\vec{X} = [x_1 \dots x_D]^T$ 为未知目标的不变性参数, $\vec{Z}_i = [Z_{1i}, Z_{2i}, \dots, Z_{Di}]^T$ 为第 i 类样本中心, $\vec{B} = [B_1, B_2, \dots, B_{\text{NUM}}]^T$, $\vec{O} = [O_1, O_2, \dots, O_{\text{NUM}}]^T$, A_i 和 B_i 是调节参数, 如果各样本及各元素的重要性相同, 则 A_i 和 B_i 都是不随 i 变化的常数, σ_{ji}^2 是第 i 类样本的第 j 个元素的均方误差, 可用统计规则得到

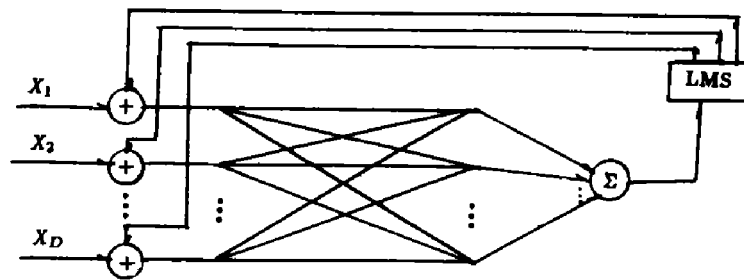


图 3 MPNN 模型的结构图
Fig. 3 The structure of MPNN model

$$\sigma_{ji}^2 = \frac{1}{m_i} \sum_{K=1}^{m_i} (S_{jKi} - Z_{ji})^2, \quad (7)$$

这儿 S_{jKi} 表示 \vec{S}_{Ki} 的第 i 个元素的值, 从式(6)式可知, ϵ 是一系列钟形函数的迭加, 均值样本矢量 \vec{Z}_i 不可能完全相同, 只要控制一定比例的 A_i 就可控制各钟形函数间的影响. 事实上, A_i 越大, 钟形函数越尖锐, 相互间的影响就越小.

使反馈系统收敛的目的是让输入的待识别目标归入最相似的样本点, 从而使得 $\epsilon(\vec{X})$ 达到波谷处. 经分析可得 \vec{X} 的迭代式^[6]为

$$\vec{X}^{(u+1)} = (\vec{I} - 2u\vec{D}) \cdot \vec{X}^{(u)} + 2u\vec{E}. \quad (8)$$

这儿

$$\vec{E} = \left[\sum_{i=1}^{NUM} B_i O_i A_i \frac{1}{\sigma_{1i}^2} Z_{1i} \quad \sum_{i=1}^{NUM} B_i O_i A_i \frac{1}{\sigma_{2i}^2} Z_{2i} \cdots \sum_{i=1}^{NUM} B_i O_i A_i \frac{1}{\sigma_{Di}^2} Z_{Di} \right]^T,$$

$$\vec{D} = \begin{bmatrix} \sum_{i=1}^{NUM} B_i O_i A_i \frac{1}{\sigma_{1i}^2} & & & & \\ & \sum_{i=1}^{NUM} B_i O_i A_i \frac{1}{\sigma_{2i}^2} & & & \\ & & \cdots & & \\ & & & \cdots & \\ & & & & \sum_{i=1}^{NUM} B_i O_i A_i \frac{1}{\sigma_{Di}^2} \end{bmatrix}.$$

4 实验结果分析

实验中, 待识别的图像是从“沙漠风暴”录相带中截取的帧图像, 共有 14 幅, 每幅图像的大小为 512×352 , 且每幅图像中含有多目标. 如果用阈值分割方法或边缘提取等传统方法来分割这些图片, 其结果根本无法令人满意, 采用第 2 节中的智能型分割方法, 分割效果令人满意. 由于利用了有关目标的知识, 使得分割出来的二值化图像特别“干净”, 二值化图像平面中只有潜在目标. 关于分割方法的效果比较, 可参见图 1 和 2.

图 4 是从图片中提取出来的 5 类图像目标, 其不变性参数用于 MPNN 训练. 由于利用 Zernike 矩, 使得同阶矩相对同类目标不同姿态的变化率在 5% 以下. 同时, 根据 MPNN 利用概率分布理论, 使它能反映目标受噪声随机影响而造成的变化, 以致 MPNN 对 14 幅图片中 60 余个目标的识别成功率为 95%, 而利用 Hamming 网作为分类器时, 识别的成功率只为 84%. 主要原因是:



图 4 用于训练 MPNN 网络的 5 类图像目标
Fig. 4 Image targets used to train MPNN model

MPNN 模型由式(5)计算得到的等值曲面为超椭球曲面,而 Hamming 网反映的等值曲面为超球面,以致造成收敛区域的不同.

5 结语

本文利用知识型分割方法使得所提取的潜在目标完整,相对经典方法更体现了智能化;利用不变性参数法获得所提取潜在目标的 PSRI 特征;利用 MPNN 模型作为目标特征分类器,使得基于概率理论的 MPNN 模型的识别成功率高于 Hamming 网络.

参考文献

- 1 Michael W R. *IEEE Trans. Neural Networks*, 1990, **1**(1):28~43
- 2 Fukushima K. *Pattern Recognition*, 1982, **15**(6):455~469
- 3 Spirkovska L, Reid M B. *Pattern Recognition*, 1992, **25**(9):975~985
- 4 Pal N R, Pal S K. *Pattern Recognition*, 1993, **26**(9):1277~1294
- 5 Khotanzad A, Yong Y H. *Pattern Recognition*, 1990, **23**(10):1089~1101
- 6 沈定刚,戚飞虎. *模式识别与人工智能*, 1994,待发表.
- 7 沈定刚,戚飞虎. *上海交通大学学报*, **28**(3):64~71

A KNOWLEDGE-BASED TARGET RECOGNITION SYSTEM*

Shen Dinggang Qi Feihu

(*Research Institute of Optic Fiber Technology, Shanghai Jiaotong University, Shanghai 20052, China*)

John C. M. Lee

(*The Department of Computer Science, The Hongkong University of Science and Technology, Hongkong*)

Abstract The technology proposed in the image target recognition consists of image segmentation, invariant feature extraction and target classification. In this paper, a knowledge-based approach is proposed for adaptive image segmentation. Invariant features about image targets are extracted by Zernike moments. Object classification is completed by the MPNN model. The experimental results show that the proposed system can be used to recognize objects from the complicated background.

Key words target recognition, image segmentation, invariant feature extraction.

*The project supported by the Climbing Programme—National Key Project for Fundamental Research in China