

矩不变量在目标形状识别中的应用研究^{*}

颜孙震 孙即祥 王晓华 钟山

(国防科技大学电子技术系 湖南长沙 410073)

摘要 本文提出了图像分割的一种新的算法, 简单实用。实验表明, 它优于经典的 max 方差准则门限法和 max 熵分割法。本文利用中心矩构成仿射不变量, 作为目标的特征。本文提出了利用上述各特征的一种数据融合算法。实验表明, 利用该算法进行识别时正确率高, 且所用时间短, 效果好。本文所提供的方法在对 18 类目标 138 幅图像进行识别时总识别正确率高于 85%。

关键词 识别模式, 仿射变换, 矩不变量 (moment invariants), 图像分割, 信息融合

分类号 TP391 41

The Research and Application of Moment Invariants in Object Shape Recognition

Yan Sunzhen Sun Jixiang Wang Xiaohua Zhong Shan

(Department of Electronic Technology, NUDT, Changsha 410073)

Abstract This paper first presented a novel and simple method of image segmentation. Experiments illustrate that this method is more effective than some traditional methods such as maximum-square-error threshold method and maximum-entropy threshold method. In the algorithm of object recognition, we construct object's characters based on four affine moment invariants. Furthermore, we proposed a data-fusion algorithm based on the above characters. The performance of the proposed technique in the recognition experiments of 138 pieces of image which gained from 18 kinds of objects, is also involved. The experiment result is satisfactory, which the correct-ratio of object recognition is above 85%.

Key words pattern recognition, affine transformations, moment invariants, image segmentation, information fuse

一个目标经过某种变化 (如观测视点的变化) 仍能被观察者视为同一模式, 这说明存在某些特征经历这种变化后仍保持不变, 而这些特征反映了不同模式之间的本质差异, 观察者也正是根据它们来作出判断。人们希望提取的特征应与目标所处的环境、条件无关, 这样的特征称为不变量。不变量的研究是模式识别领域的热点之一。

目标的主要特征大致地分为如下几类: (1) 形状特征; (2) 运动特征; (3) 光谱特征; (4) 灰度分布 (纹理) 特征。本文主要依据形状不变量进行识别。提取目标的形状不变量有许多方法。本文利用了一组由不高于三阶的中心矩组成的仿射矩不变量, 来提取目标的形状特征, 实验表明其抗干扰能力强, 识别可靠性高, 可适用于较复杂目标的识别。对于图像分割, 本文提出了一种基于类间类内距离比最大准则的灰度门限法, 并给出了该算法与最大方差准则门限法和最大熵分割法相比较的实验结果。我们提出了基于信息融合的思想的一种目标识别算法, 该算法与神经网络识别算法相比较, 也显示了良好的性能和较高的识别正确率。

1 仿射矩不变量

^{*} 1998年3月3日收稿

第一作者: 颜孙震, 男, 1974年, 本科生

1 1 概述

在图像识别中，不变量是指目标的特性在经历了如下的一个或几个变换后仍然保持不变的目标的特征量：(1) 目标尺度的改变；(2) 目标图像的平移；(3) 目标图像的旋转；(4) 仿射变换；(5) 透视变换。观察二维面目标的视点的不同，便等效于上述的一个或几个变换。变换 (1)、(2)、(3) 是 (4) 仿射变换的特例。实际成像系统是一个透视变换，透视变换是非线性的。但在满足一定条件的情况下，可以用仿射变换很好地近似透视变换

1 2 矩的定义和仿射矩不变量的构造

设图像函数 $f(x, y) \in L(R^2)$ ， D 为目标区域，则目标的 $p+q$ 阶原点矩定义为：

$$m_{pq} = \iint_D x^p y^q f(x, y) dx dy \tag{1}$$

其中 $p, q = 0, 1, 2, \dots$ ，目标的 $p+q$ 阶中心矩定义为：

$$\mu_{pq} = \iint_D (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy \tag{2}$$

式中 $(\bar{x}, \bar{y}) = (m_{10}/m_{00}, m_{01}/m_{00})$ 为目标区域灰度质心

矩是用来描述目标区域的灰度分布特征的，在 $f(x, y)$ 分片连续且在有限区域内有非零值时，则其各阶矩皆存在，且 $f(x, y)$ 唯一地确定一个矩序列 $\{m_{pq}\}$ (或 $\{\mu_{pq}\}$)，反之，矩序列 $\{m_{pq}\}$ (或 $\{\mu_{pq}\}$) 也唯一地确定 $f(x, y)$ 。

下面我们不加证明地给出构造矩不变量的有关定理：

(1) 配极多项式 (A polar polynomial)：

$$\sum_{i=0}^k (-1)^i \binom{k}{i} x^{k-i} y^i \tag{3}$$

当 k 为偶数时，经归一化后为矩不变量

(2) Hankel行列式：

$$\begin{vmatrix} \mu_{0k} & \mu_{1k-1} & \dots & \mu_{q-1k-q+1} \\ \mu_{1k-1} & \mu_{2k-2} & \dots & \mu_{qk-q} \\ \dots & \dots & \dots & \dots \\ \mu_{u-1k-u+1} & \mu_{uk-u} & \dots & \mu_{u+q-2k-u-q+2} \\ \mu_{0k'} & \mu_{1k'-1} & \dots & \mu_{q-1k'-q+1} \\ \mu_{1k'-1} & \mu_{2k'-2} & \dots & \mu_{qk'-q} \\ \dots & \dots & \dots & \dots \\ \mu_{v-1k'-v+1} & \mu_{vk'-v} & \dots & \mu_{v+q-2k'-v-q+2} \end{vmatrix} \tag{4}$$

经归一化后为矩不变量。其中： $k = q + u - 2$ ， $k' = q + v - 2$

(3) 关于 x 的多项式

$$f(x) = \sum_{i=0}^k \binom{k}{i} x^{k-i} \tag{5}$$

的判决式 (Determination) 除以 μ_{0k} ，并归一化后为矩不变量。

现根据 (1)、(2)、(3) 构造出仅使用 3阶矩的 4个矩不变量：

$$I_1 = (\mu_{20} \mu_{02} - \mu_{11}^2) / \mu_{00}^4 \tag{6}$$

$$I_2 = (\mu_{30} \mu_{03} - 6 \mu_{30} \mu_{03} \mu_{21} \mu_{12} + 4 \mu_{30}^3 \mu_{12} + 4 \mu_{12}^3 \mu_{03} - 3 \mu_{21}^2 \mu_{12}^2) / \mu_{00}^{10} \tag{7}$$

$$I_3 = (\mu_{20}(\mu_{21} \mu_{03} - \mu_{12}^2) - \mu_{11}(\mu_{30} \mu_{03} - \mu_{21} \mu_{12}) + \mu_{02}(\mu_{30} \mu_{11} - \mu_{21}^2)) / \mu_{00}^7 \tag{8}$$

$$I_4 = (\mu_{30} \mu_{03} - 6 \mu_{20} \mu_{11} \mu_{12} \mu_{03} + 9 \mu_{20}^2 \mu_{02} \mu_{12}^2 + 12 \mu_{20} \mu_{11} \mu_{21} \mu_{03} + 6 \mu_{20} \mu_{11} \mu_{30} \mu_{03} \mu_{02} - 18 \mu_{20} \mu_{11} \mu_{02} \mu_{21} \mu_{12} - 8 \mu_{11}^3 \mu_{30} \mu_{03} - 6 \mu_{20} \mu_{02} \mu_{30} \mu_{12} + 9 \mu_{20} \mu_{02} \mu_{21}^2 + 12 \mu_{11}^2 \mu_{02} \mu_{30} \mu_{12} - 6 \mu_{11} \mu_{02}^2 \mu_{30} \mu_{21} - 6 \mu_{20} \mu_{02} \mu_{03} \mu_{21} + \mu_{02}^3 \mu_{30}) / \mu_{00}^{11} \tag{9}$$

2 图像分割算法研究

景物分割的目的在于准确地找出三维物体的边缘, 一般来说, 实际得到的二维图像的边界和三维物体的边缘之间通常满足某种对应关系。

从图像中提取目标, 即图像分割技术, 可以从两个不同途径实现: 一是提取目标的边缘, 一是提取目标的区域。考虑到边缘特征受噪声影响较大, 而区域特征受噪声影响相对来说要小得多, 使用灰度门限法进行区域分割又比较简单, 我们提出了一种新的区域分割算法, 并取得了较好的效果。

为了说明方便, 在这里我们认为图像只有前景和背景两个区域。实际上该算法稍加扩展即可用于多目标分割。

一般地, 在统计判决门限法中经常运用的是日本学者大津展之提出的最大方差准则最佳门限法, 即使类间距离 (方差) 取最大而得到的最佳门限^[6]。还有一种常用的方法是最大熵准则灰度门限法, 其基本思想是选择最佳分割灰度门限, 使图像的熵达到最大。实验表明, 上述两种方法在实际应用中分割效果并不总是很好, 尤其是对质量较差的图像而言。

在本文中, 我们用类间距离 (方差) 与类内距离 (方差) 之比作为最佳分割的目标函数。其推导过程如下:

设图像的灰度级为 L , 图像中灰度值是 i 的像素个数为 n_i , 则总的像素数 $N = \sum_{i=1}^L n_i$, 各灰度值出现的频率为 $p = \frac{n_i}{N}$ ($i = 1, \dots, L$)。又设以灰度 k 为门限将图像分割成两个区域, 灰度为 $1 \sim k$ 的像素和灰度为 $(k+1) \sim L$ 的像素分别构成区域 A 和 B , 区域 A 和 B 的频率分别为:

$$p_A = \sum_{i=1}^k p_i \equiv p(k) \tag{10}$$

$$p_B = \sum_{i=k+1}^L p_i = 1 - p(k) \tag{11}$$

区域 A 和 B 的平均灰度分别为:

$$\bar{u}_A = \frac{1}{p_A} \sum_{i=1}^k i p_i \equiv \frac{u(k)}{p(k)} \tag{12}$$

$$\bar{u}_B = \frac{1}{p_B} \sum_{i=k+1}^L i p_i \equiv \frac{u - u(k)}{1 - p(k)} \tag{13}$$

区域 A 中的所有灰度 (级) 到区域 B 的平均灰度的类间距离为 $\sum_{i=1}^k (i - \bar{u}_B)^2 p_i$ 。

区域 B 中的所有灰度 (级) 到区域 A 的平均灰度的类间距离为 $\sum_{i=k+1}^L (i - \bar{u}_A)^2 p_i$ 。

区域 A 和区域 B 之间的距离平方和 (类间距离) 为

$$\mathcal{C}_1^2(k) = \sum_{i=1}^k (i - \bar{u}_B)^2 p_i + \sum_{i=k+1}^L (i - \bar{u}_A)^2 p_i \tag{14}$$

区域 A 和区域 B 内部的聚合程度 (类内距离) 为

$$\mathcal{C}_2^2(k) = \sum_{i=1}^k (i - \bar{u}_A)^2 p_i + \sum_{i=k+1}^L (i - \bar{u}_B)^2 p_i \tag{15}$$

根据以上两式我们构造统计量: $d^2(k) = \frac{\mathcal{C}_1^2(k)}{\mathcal{C}_2^2(k)}$ (16).

从 $1 \sim L$ 改变 k , 使 $d^2(k) = \max$ 的 k 值便是区域分割门限, 这时前景和背景的灰度差别尽可能地大, 而前景和背景各自的灰度分布差别又尽可能地小, 从而达到在上述意义下的最佳分割图像的目的。

这种方法的特点, 一是可分性更大, 二是对于噪声有较好的适应性, 它也可向多门限选择推广, 以用于多目标分割。

3 融合识别算法

3 1 集中式融合算法的原理

统计判决法以其思想严谨、方法成熟和较好的效果在模式识别中起着重要的作用。而信息融合技术通过对多源信息的有机有效的“组合”，充分利用各种信息使信息互补。本文中我们将运用经典的统计理论和新兴的融合思想对不同算法计算的结果进行“融合”，从而提高识别的正确率。

利用多种信息对目标进行识别，可以看作是在多维特征矢量 $x = (x_1, x_2, \dots, x_D)^T$ 下判断 M 个事件中某一事件 k_j 发生的可能性，即要获得这样一个判别概率：

$$P(k_j | x) = \frac{p(x | k_j)P(k_j)}{p(x)} \quad (j = 1, \dots, M) \quad (17)$$

假设矢量 x 各分量是相互独立的，则有：

$$P(k_j | x) = \frac{\prod_{i=1}^D p(x_i | k_j) P(k_j)}{p(x)} \quad (18)$$

又根据贝叶斯公式可得：

$$P(k_j | x) = \frac{\prod_{i=1}^D P(k_j | x_i) p(x_i) P(k_j)}{[P(k_j)]^D \cdot p(x)} \quad (19)$$

又由矢量 x 各分量的相互独立性有：

$$P(k_j | x) = \frac{\prod_{i=1}^D P(k_j | x_i)}{[P(k_j)]^{D-1}} \quad (20)$$

上式中， $P(k_j | x_i)$ 为在矢量 x 的第 i 个特征分量 x_i 下对事件 k_j 发生的判别概率，而 $P(k_j | x)$ 为条件联合判别概率，这样就把条件联合判别概率和各个分量的判别概率联系在一起了。具体到我们的研究中， $P(k_j | x_i)$ 即为用某一矩不变量单独进行判断时的判别概率（后验概率），而 $P(k_j | x)$ 则是依据几个矩不变量作为后验条件的判别概率。

3 2 后验概率的确定

以上已经得到了总的判别概率和各个分量的判别概率之间的关系，现在要解决如何确定用某一矩不变量单独进行判断时的判别概率 $P(k_j | x_i)$ 的问题。

根据贝叶斯公式，后验概率 $P(k_j | x_i)$ 可由类别 k_j 的先验概率 $P(k_j)$ 和 x_i 的类概率密度 $p(x_i | k_j)$ 确定，即：

$$P(k_j | x_i) = \frac{p(x_i | k_j)P(k_j)}{p(x_i)} = \frac{p(x_i | k_j)P(k_j)}{\sum_{k=1}^M p(x_i | k_k)P(k_k)} \quad (21)$$

由于所统计的图像数量繁多、数据庞大，多为加权和运算，据中心极限定理，我们有理由认为这些矩不变量对于某一类模式是近似服从正态分布的。大量实验表明，上述观点是正确的。由此我们就可以得到 x_i 的条件概率密度函数为：

$$p(x_i | k_j) = \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left\{-\frac{(x_i - \mu_j)^2}{2\sigma_j^2}\right\} \quad (22)$$

式中 μ_j 和 σ_j^2 分别代表其均值和方差，可由大量统计算得。

根据以上所述的方法和公式，就可以很容易地获得某一矩不变量下的后验概率，进而得到总的判别概率。

4 实验结果

4 1 图像分割算法实验结果

我们用 18 类目标共 138 幅图像对本文提出的分割算法的有效性进行验证，其中每一类的图像的大小、位置、放置方向都有所不同，且有不同的信噪比。对比算法为最大方差准则最佳门限法和最大熵准则最佳门限法。图 1 图 2 图 3 为部分实验结果。

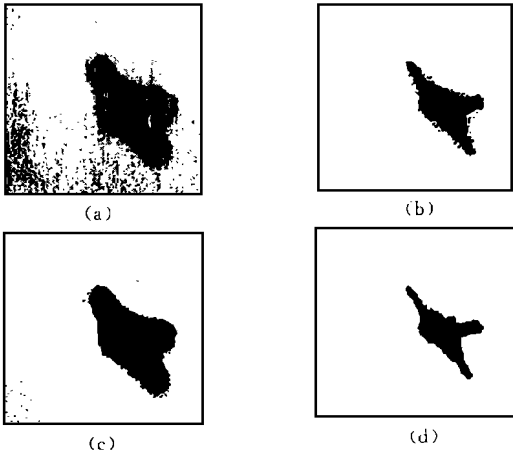


图 1 原始图像和分割结果

- (a) 原始图像
- (b) 用最大方差准则最佳门限法分割的结果
- (c) 用最大熵准则最佳门限法分割的结果
- (d) 用类间类内距离比最大准则最佳门限法分割的结果

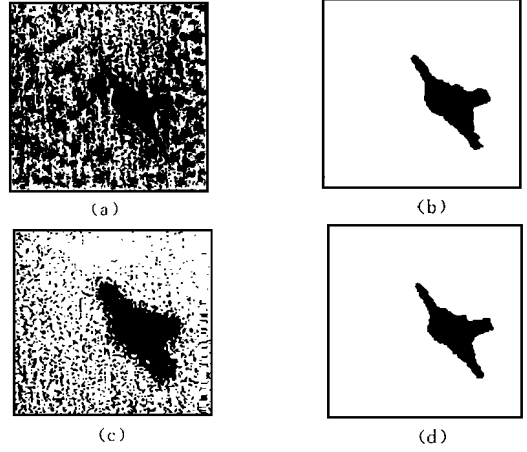


图 2 噪声污染图像和分割结果

- (a) 噪声污染图像 (信噪比为 6: 1)
- (b) 用最大方差准则最佳门限法分割的结果
- (c) 用最大熵准则最佳门限法分割的结果
- (d) 用类间类内距离比最大准则最佳门限法分割的结果

从图 1 图 2 图 3 可以看出, 用最大方差准则法和最大熵准则法分割的结果都不理想, 尤其是在复杂环境中和有噪声污染的情况下, 二者甚至不能从背景中分离出目标。而用类间类内距离比最大准则法分割得比较精确, 较好地合于目标。这充分说明, 此新的分割算法具有更优越的性能。

4.2 融合识别算法实验结果

我们把文中提出的融合识别算法和神经网络算法的识别结果作一比较。实验中采用了一种典型的 BP 多层神经网络。这种网络采用了误差反传 (error back propagation) 的训练算法。该网络采用了三层结构, 输入层结点数为 4, 中间层结点数为 6, 输出层结点数为 18。下表 (表 1) 给出了两种识别方法对 18 类目标 138 幅图像的识别正确率。

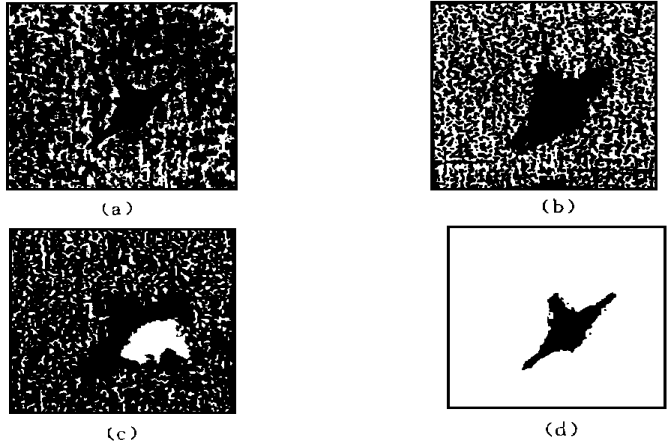


图 3 噪声污染图像和分割结果

- (a) 噪声污染图像 (信噪比为 7: 2); (b) 用最大方差准则最佳门限法分割的结果
- (c) 用最大熵准则最佳门限法分割的结果; (d) 用类间类内距离比最大准则最佳门限法分割的结果

表 1

原始图像 \ 识别算法	神经网络识别算法	融合识别算法
噪声小的原始图像	96.3%	100%
信噪比为 6: 1 的原始图像	92.2%	82.4%
信噪比为 7: 2 的原始图像	78.9%	69.5%

从表 1 可以看出, 本文提出的融合识别算法在已知的特征量较少的情况下, 其识别正确率略低于神经网络识别算法。下面是这两种算法在其它方面的一些比较:

1 判别时间的比较

两种算法对 18类 138幅图像的判别时间都不过是几秒钟,神经网络法所用时间稍长。

2 训练时间的比较

与融合识别算法相比,在训练时间上,神经网络法所用时间要多许多。由于在训练过程中神经网络法是随机抽取样本进行训练,所用时间不定,但都超过了 5分钟。

3 算法实现难易程度的比较

用神经网络法进行识别时,其训练过程是一个迭代的过程,它的收敛与否依赖于模式特性差别的大小。当算法收敛时,其收敛速度还和最小加权修正值有关,实现起来不如融合识别算法方便。

从上面的比较中我们可以看出各种算法的优劣,同时也表明了融合识别算法的有效性。

5 结论

本文较完整地讨论了包含图像分割、特征提取和模式识别等相关技术的一种目标识别方法。我们运用了一组四个在仿射变换下保持不变的矩不变量,实验表明,以此作为空间平面目标或薄体目标的模式特征,其抗噪能力强,识别率高,可用于低信噪比条件下复杂目标的识别,对于图像分割算法,我们提出了一种基于类间类内距离比最大准则确定灰度门限的方法,该算法与最大方差准则法及最熵准则法的实验结果比较,证明了该算法在所用环境下更为有效。在识别算法研究中,我们基于信息融合的思想导出了一种识别算法,通过与经典的 BP 多层神经网络识别算法的比较,该算法的优点也充分体现出来了。

摄像机的成像过程是从三维物体空间到二维图像平面的透视投影,是一个非线性的透视变换过程,因此文中提出的仿射距不变量在实际应用中有一定的局限性,然而正如前文所说的以及实验表明,在一定条件下,透视变换可以用仿射变换很好地近似。如何构造出在透视变换下保持不变的一般形体的透视不变量,并以此作为目标的模式特征对目标进行正确的识别,这是当前计算机视觉领域的一大难点和热点,也是我们后续工作的主要研究方向。

参考文献

- 1 M K Hu. Visual Pattern Recognition by Moment Invariants. IRE Trans., Inf. Theory 8, 1962, 179-187
- 2 Reforming The Theory Of Invariant Moments For Pattern Recognition. Pattern Recognition, 1992, 25 (7): pp. 723-730
- 3 Li Bing deng, Shen Jun. Fast Computation Of Moment Invariants. Pattern Recognition, 1991, 24 (8): 807-813
- 4 Chen Chaur din. Improved Moment Invariants For Shape discrimination. Pattern Recognition, 1993, 26 (5): 683-686
- 5 Jan Flusser, Tomas Suk. Pattern Recognition By Affine Moment Invariants. Pattern Recognition, 1993, 26 (5): 167-174
- 6 孙即祥. 数字图像处理. 石家庄: 河北教育出版社, 1992
- 7 周继成. 人工神经网络. 北京: 科学普及出版社, 1993
- 8 Taek Mu Kw on and Hui cheng. Contrast Enhancement For Backpropagation. IEEE Trans. On Neural Networks, 1996, 7 (22)