

时频域混合的融合多特征协方差矩阵 目标识别算法

王大伟^{1,2}, 朱娟^{1,2}, 孙宏海¹, 王延杰¹

(1 中国科学院 长春光学精密机械与物理研究所, 吉林 长春 130033; 2 中国科学院 研究生院, 北京 100039)

摘 要: 基于特征融合的目标识别中, 为了有效选择融合特征, 提高目标识别率, 如何权衡所选择的特征间的相关性、选择有效的特征融合策略是其中的关键问题. 为了解决军事目标在畸变情况下的快速正确识别, 提出了多特征协方差矩阵目标识别方法, 与经常采用的单一特征目标识别方法相比, 该方法具有数据维数更低, 特征间相关性更小的优点. 该方法分为如下几个步骤: 首先, 针对同一目标分别提取几组特征并建立各自的特征向量; 然后, 利用特征向量构建协方差矩阵; 最后, 利用提出的归一化的 fisher 判别方法进行模式分类, 并将其用于多姿态目标识别. 在哥伦比亚大学的 coil-20 图像库上进行试验, 获得了 99.54% 的平均识别率, 试验结果同时证明了算法具有很强的鲁棒性.

关键词: 目标识别; 协方差矩阵; 特征级融合; 线性判别分析

中图分类号: TP31

文献标识码: A

文章编号: 1000-7180(2010)03-0029-04

Object Recognition Algorithm of Time-Frequency Domain Feature Integration Based on Covariant-matrix

WANG Da-wei^{1,2}, ZHU Juan^{1,2}, SUN Hong-hai¹, WANG Yan-jie¹

(1 Changchun Institute of Optics, Fine Mechanics and Physics, CAS, Changchun 130033, China;

2 Graduate School of the Chinese Academy of Sciences, Beijing 100039, China)

Abstract: In object recognition based on feature fusion, the key issues are the correlation of selected features and the strategy of feature fusion. For military target recognition with rapidness and high accuracy, we motivate the object recognition algorithm of time-frequency domain feature integration based on covariant-matrix, the new method with lower dimensions and more de-correlated among features compared to traditional method with single feature. The proposed method comprises three steps: firstly, extracting several sets of features with the same pattern and establishing each feature vectors; then, constructing co-variable matrix with these vectors; and finally, doing pattern recognition using improved normalized Fisher Linear Discriminant Analysis. At last, we demonstrated our improved algorithm on coil-20 image library, Colombia University, and obtained 99.54% average recognition accuracy. Our experimental results validated the wonderful robustness of our method.

Key words: object recognition; covariant matrix; feature-level fusion; linear discriminant analysis

1 引言

在目标的观测识别中, 目标通常以不同的姿态出现在视场中, 且存在旋转、缩放、平移等问题, 还常常受到噪声污染和光照变化的影响. 在这种情况下,

进行有效的特征选择, 并把这些特征以一定的方式融合起来进而实现正确的目标识别就显得尤为重要. 最初的方法是采用边缘和角点特征进行目标识别, 而这类特征很容易受到噪声的影响, 识别率不高; 后来又发展出了基于矩特征, 变换域特征和统计

指示特征的目标识别方法. 上述的各种方法在计算复杂度, 正确识别率和特征表达能力等方面各有优缺点. 而在实际工程应用中, 在获得较高识别率的前提下兼顾算法复杂度变得尤为重要.

特征级图像融合方法是近年来提出的一种图像中间层的处理方法, 它不但适用于多传感器图像融合, 也可以对同一图像的不同特征进行融合处理, 能够获得比决策级融合更好的性能^[1-2]. Oncel Tuzel 首先提出了利用协方差矩阵描述目标的方法^[3]; Zhang 等人提出基于多特征协方差矩阵的融合方法进行目标的匹配跟踪^[4]. 基于文中的应用, 我们改进了算法, 在特征融合上, 更好地选择有效特征进行特征融合, 并把该算法用于多目标识别中; 在目标分类上, 针对噪声和光照的干扰, 造成目标不能正确分类, 提出了归一化的 fisher 判别方法. 试验结果表明, 文中算法对旋转、缩放、平移、噪声和光照变化等极端条件都有很好的抵抗能力.

2 特征描述

2.1 灰度和彩色直方图特征

灰度或者亮度是图像的基本信息, 它反映了目标的空间分布. 考察一幅 $m \times n$ 大小的图像^[5], 它实际上可以看做维空间的一个点. 用形如 $f(x, y)$ 的 2 维函数表示一幅图像. 在空间坐标 (x, y) 上 f 的值或者幅度是一个正的标量值, 其物理意义由源图像决定. 实际上严格说来, 图像的灰度并不等同于亮度, 亮度是指光源在每立体角度上发出的光通量, 是与光源相关联的值. 尽管如此, 目前在灰度图像处理时, 亮度与灰度已经通用. 颜色特征也是目标的基本特征, 在军事领域中, 虽然颜色特征很容易被伪装, 但是在工业检测, 或交通识别等复杂场景系统中, 目标颜色比较鲜明, 利用彩色特征往往可以极大的提高目标检测率^[6]. 颜色特征除了可用 RGB 颜色空间, 还可以采用 HSI 颜色空间和 Lab 颜色空间等. 由于灰度或者彩色直方图特征是基于统计特性的, 所以它具有优秀的旋转、缩放不变性.

2.2 空间距离

目标的边缘和角点特征虽然具有光照不敏感的优点, 但是其在协方差矩阵的融合中, 需要把二维矩阵的转化为一维向量, 这改变了边缘或角点信息的二维位置关系. 而对于确定目标, 其空间相对位置关系不随旋转、缩放而变化. 这里首先利用 Canny 算子提取目标的边缘信息, 然后计算目标的重心, 通过式 (1) 计算边缘像素与目标区域重心的距离作为特

征, 为不致在协方差矩阵中空间几何距离特征的数量级过小, 图像其他像素到目标重心的距离我们用一个适当的值代替. 为获得与目标尺寸无关的特征, 需要利用目标的尺寸对其进行归一化.

$$d(x, y) = [(x - x_G)^2 + (y - y_G)^2]^{\frac{1}{2}} \quad (1)$$

2.3 梯度特征

梯度特征体现了目标的边缘信息, 反映了图像局部灰度的变化情况, 是图像的基本特征之一, 而且对光照变化具有很好的克服能力. 单从字面上来说, 梯度算子就是求取图像的一阶导数, 对于二维图像而言, 导数的求取就具有方向性, 简便起见, 一般可以使用 x 方向的梯度特征或者 y 方向的梯度特征. 常用的梯度算子还有: Sobel 算子、Prewitt 算子、Roberts 算子, 文中使用了 Sobel 算子提取梯度特征.

2.4 傅里叶相位谱

傅里叶变换是一个重要的图像处理工具, 它可以把图像分解成其正弦和余弦分量. 在傅里叶域图像中, 每一点都代表了一个在空间域图像中的频率. 在数字图像处理中, 通常采用二维傅里叶变换 (2-D DFT). 对于一幅 $N \times N$ 大小的图像, 其二维傅里叶变换如式 (2) 所示:

$$F(k, l) = \frac{1}{N^2} \sum_{a=0}^{N-1} \sum_{b=0}^{N-1} f(a, b) e^{-i2\pi(\frac{ka}{N} + \frac{lb}{N})} \quad (2)$$

式中, f 代表图像的灰度值.

很多研究者已经指出, 在描述图像特征方面, 傅里叶相位谱比傅里叶幅度谱重要很多, 不同的目标图像通常有着不同的傅里叶相位谱^[7]. 为提高傅里叶相位谱的图像对比度, 取得合理的灰度数量级, 需要用对数变换对傅里叶变换后的图像幅度作如式 (3) 所示的处理:

$$f_{\text{new}} = \log(1 + \text{abs}(f)) \quad (3)$$

式中, f 表示傅里叶变换后的幅度, f_{new} 是对数变换后的幅度. 在融合多特征协方差矩阵中, 傅里叶相位谱具有旋转、平移、缩放不变性, 但是需要指出的是傅里叶相位谱对噪声过于敏感, 在极端条件下往往不能使用.

3 特征融合策略和协方差矩阵构造

虽然特征级融合在理论上存在如何寻找紧密的特征表达的问题, 但是它与逐个像素的处理方法相比, 明显地降低了运算量, 并且取得了较好的结果. 良好的识别特征应该具有如下的特点:

(1) 可区分性: 目标和背景, 以及目标之间应该具有良好的区分性.

(2) 可靠性:描述同类目标的特征在特征空间中应该具有良好聚类特性。

(3) 独立性:同一目标的各个特征应该互不相关。

(4) 维数低:系统的计算复杂度是随着特征量的增加而迅速增长的.通常采用映射或者变换的方法把高维空间中的样本或特征引入低维处理。

已经有学者论述,协方差矩阵在可区分性、可靠性和特征维数方面是融合多维特征描述目标的有效方法^[1,5].对于独立性的判别,引入了如式(4)所示的公式。

$$J(f_i, f_j) = \frac{|C_{i,j}|}{\sqrt{|C_{i,i}|} \times \sqrt{|C_{j,j}|}} \quad (4)$$

式中, $C_{i,j}$ 表示特征间的协方差, $C_{i,i}$ 和 $C_{j,j}$ 表示特征的方差.通过上式可以判定特征间的独立性,优化特征选择。

对于 $M \times N$ 大小的图像区域 R ,其具有 d 维特征的协方差矩阵可以用如式(5)的表达式构造:

$$C_{\text{image}} = \frac{1}{M \times N} \sum_{k=1}^{MN} (f_k - \mu_R) \times (f_k - \mu_R)^T \quad (5)$$

式中, μ_R 是图像 R 中各个像素对应特征的均值.协方差矩阵是一个对称矩阵,其对角线上的元素代表每个特征的方差,而非对角线上的元素代表了各个特征之间的相关性。

采用协方差矩阵融合多维目标具有如下优点:

(1) 协方差矩阵融合多维特征,每增加一种特征,在协方差矩阵中数据仅仅增加一维,而不会像直方图那样随着特征维数的增加其计算量呈指数级增长;且协方差矩阵是对称矩阵,对 N 种特征构成的协方差矩阵,在分类处理中只要计算 $(N^2 + N)/2$ 个值,大大降低了计算复杂度。

(2) 协方差矩阵的维数仅与所选择的特征个数有关,与目标尺寸和旋转无关.而且,只要选择合适的特征,协方差矩阵对于目标图像的旋转、缩放基本保持不变.文中选择了基于统计特性的特征,用这类特征构造协方差矩阵对于目标的旋转、平移和缩放都有良好的抵抗能力。

(3) 融合多维特征协方差矩阵的方法相当的灵活,可以根据不同的目标和环境特点选择不同的特征进行协方差矩阵的构建.比如,当处理灰度图像时可以舍弃 RGB 颜色特征。

为实现目标匹配,文献[5]采用了如式(6)的协方差矩阵距离度量公式:

$$\rho(C_i, C_j) = \sqrt{\sum_{k=1}^d \ln^2 \lambda_k(C_i, C_j)} \quad (6)$$

式中, $\lambda_k(C_i, C_j)$ 是协方差矩阵 C_i 与 C_j 的广义特征值,可由式(7)计算:

$$|\lambda C_i - C_j| = 0 \quad (7)$$

然而,式(7)存在一个代数上的问题,即不能保证其总有非奇异的 λ 解.实际的试验结果也并不理想,见第4节的试验结果比较.为解决这一问题,文中提出了归一化的 Fisher 线性判别方法^[8]。

4 归一化的 Fisher 线性判别分析

一般说来,模式识别问题有基于距离度量和基于判别函数两大类方法.协方差矩阵的距离计算一般采用式(6)和式(7),但是不能总保证式(7)有实数解,从而使得分类的正确率大大地下降.我们转而寻求基于判别函数的方法,其中 Fisher LDA 是一种计算量小而有效的分类方法.而在军事目标的观测识别中,目标经常受到噪声、光照等条件的干扰,使得传统的 Fisher LDA 方法找不到正确的待分类目标的正确分类阈值,针对这种情况,我们提出一种归一化的 Fisher LDA 方法。

Fisher LDA 考虑的是把 d 维空间的样本投影到一条直线上,把数据维数压缩到一维.关键的问题是如何根据实际情况找到这条最好的、最易于分类的投影线.通常为达到上述目的,需要最大化如下的目标函数:

$$J_F(w) = \frac{w^T S_b w}{w^T S_w w} \quad (8)$$

式中, S_b 称为样本类间离散度矩阵,由于 S_b 实际上是两个向量的外积,因此其秩至多为 1,并且是对称和半正定的; S_w 称为样本类内离散度矩阵,它与全部样本的样本协方差矩阵成正比,也是对称半正定的.样本类间离散度矩阵和样本类内离散度矩阵的定义如式(9)、(10)所示:

$$S_b = (m_1 - m_2)(m_1 - m_2)^T \quad (9)$$

$$S_w = S_1 + S_2 \quad (10)$$

式中, $m_i = \frac{1}{N} \sum_{x \in X_i} x$, $i = 1, 2$, $S_i = \sum (x - m_i)(x - m_i)^T$, $i = 1, 2$. 式(8)中 $J_F(w)$ 是广义 Rayleigh 商,利用 Lagrange 乘数法可以得出使 $J_F(w)$ 取得极大值的 w^* 如式(11)所示:

$$w^* = S_w^{-1}(m_1 - m_2) \quad (11)$$

这个 w^* 就是使得类间散布和类内散布的比值达到最大的线性函数.接下来只要确定出阈值点

y_0 , 并将分类样本在 w^* 上的投影点 y_n 与 y_0 相比较, 便可作出分类决策。

从理论上讲, 在有很多训练样本的情况下并不能使最小误差概率减小。而且在工程实践中, 往往会出现诸如噪声污染、光照变化、局部遮挡等极端情况, 目标在 w^* 上的投影点就会发生平移或波动, 有时就不能获得好的分类效果甚至完全不可分。这种情况下, 我们总是愿意为了得到操作的方便性和实用性, 相应的牺牲掉理论上的分类效果。所以, 如何求取阈值 y_0 , 并如何使得两类的投影点在 y_0 两侧合理分开以进行有效识别就称为一个主要的问题。这里我们提出了归一化 Fisher 线性判别分析, 该算法分为如下三个步骤:

(1) 在由训练集计算分类阈值时, 把投影点 $y = w^{*T}x$ 得到的训练值归一化。首先把训练值平移到大于 0 的区域上, 然后求得平移后的均值 $\bar{y}$, 把平移后的训练值中大于 $\bar{y}$ 的数和小于 $\bar{y}$ 的数分别除以和乘以一个系数 φ , 这样可以把训练值归一化并把两类数据分的更开一些(文中 $\varphi = \max(y)$), 最后通过归一化后的数值计算得新的归一化分类阈值 y_0 。

(2) 测试集计算出特征协方差矩阵后, 由 $y = w^{*T}x$ 同样可以得到测试集的投影点, 然后把这些投影点按照(1)中类似的方法进行归一化处理, 得到归一化后的投影数值, 然后把这些数据与步骤 1 中得到的阈值 y_0 进行比较, 规矩决策规则:

$$\begin{aligned} y > y_0 &\rightarrow x \in \omega_1 \\ y < y_0 &\rightarrow x \in \omega_2 \end{aligned} \quad (12)$$

就可以判断 x 属于什么类别。

(3) 对于 K 类的分类问题, 采用两两分类的, 得到 $K(K-1)/2$ 个判别结果, 再用投票的方法得到最终输出, 使 x 属于投票数最多的类。当投票数为平局时, 在实验中取类别下标较小的类。

5 试验结果

实验结果证明, 文中提出的融合特征方法与依靠单一特征的目标识别方法相比, 在识别率有显著的提高; 与依靠协方差矩阵距离度量的分类方法相比, 文中提出的归一化 fisher 判别分析方法也极大的提高了分类识别率, 而且不用反复求解一元高次方程, 避免出现一元高次方程出现奇异解的情况, 计算也更为简便。

为了验证文中算法对噪声污染和光照变化的克服能力, 选择灰度(亮度)特征、梯度特征和空间距离

特征组成特征协方差矩阵, 在不同大小的高斯噪声条件下、不同强度的光照条件下进行了分类试验。由分类结果可以看出, 目标即使在受到很严重的噪声污染和强烈的光照变化情况下仍然能够获得较高的分类正确率。

6 结束语

针对军事目标的多角度、多姿态识别问题, 文中利用协方差矩阵融合多维特征特有的优点, 合理地选择了融合特征, 使得所提出的识别算法具有很强的旋转、缩放、平移不变性, 与基于单一特征的目标识别方法相比, 识别率有显著的提高; 与依靠协方差矩阵距离度量的分类方法相比, 文中提出的归一化 fisher 判别分析方法避免求解有可能奇异的一元高次的协方差矩阵广义特征值方程, 极大地提高了分类识别率, 使得计算也更为简便。

参考文献:

- [1] Perry Lindberg. An optics approach to sensor fusion for target recognition[J]. SPIE, 1992(1699): 152-157.
- [2] Kuhnert K D. Towards creating abstract features of complex objects - the fusion of contour points in significant contour sections for object recognition[J]. Pattern Recognition Letters, 2002(23): 713-718.
- [3] Förstner W, Moonen B. A metric for covariance matrices [R]. Dept. of Geodesy and Geoinformatics, Stuttgart University, 1999: 113-128.
- [4] Tuzel O, Porikli F, Meer P. Region covariance: a fast descriptor for detection and classification[C]// proc. 9th European Conf. on Computer Vision. Graz, Austria, 2006: 589-600.
- [5] Xuguang Zhang. Integrated intensity, orientation code and spatial information for robust tracking[C]// ICIEA. China: Changchun, 2007: 1842-1846.
- [6] Jianzhong Fang, Guoping Qiu. A colour histogram based approach to human face detection[C]// Proc. Conf. on Visual Information Engineering. UK, Nottingham University, 2003: 133-136.
- [7] Shioyama T, Hamanaka J. Recognition algorithm for hand-printed Chinese characters by 2D-FFT[J]. Pattern Recognition, 1996(3): 225-229.
- [8] Yang Wankou. Fuzzy inverse fisher discriminant analysis for face recognition[J]. Wavelet Analysis and Pattern Recognition, 2007, 1(2/4): 107-111.

作者简介:

王大伟 男, (1983-), 博士研究生。研究方向为图像融合、目标识别。