

# 基于仿射不变性特征的视点空间划分

孙洁, 马惠敏, 李凤亭

(清华大学 电子工程系, 北京 100084)

**摘要:** 为了简化三维物体的识别过程, 提高三维物体识别的识别率, 该文利用Multi-scale autoconvolution、Trace 变换、Zernike 矩3种仿射不变性特征, 对飞机、汽车、人等三维物体进行视点空间划分, 用尽可能少的不等间隔的三维物体的二维投影图像来表达三维物体, 并以此为依据进行三维物体识别。在此基础上提出一种针对不同类型物体的仿射不变性特征提取策略, 并建立一个实现三维物体任意姿态识别的软件系统平台, 应用Princeton 形状标准库中的部分模型对该平台进行测试。结果表明, 该方法能够取得较好的识别效果, 识别率在90%以上。

**关键词:** 模式识别; 三维识别; 视点空间划分; 仿射不变性特征

中图分类号: TP 391.4

文献标识码: A

文章编号: 1000-0054(2009)01-0053-04

## Viewpoint space partitioning based on affine invariant features

SUN Jie, MA Huimin, LI Fengting

(Department of Electronic Engineering, Tsinghua University, Beijing 100084, China)

**Abstract:** The viewpoint space of 3-D objects, such as aircraft, automobiles, and humans is partitioned using three affine invariant features based on a multi-scale autoconvolution, a trace transform, and the Zernike moment to express the 3-D objects within minimum two-dimensional projective images. The images are then nonuniformly projected for object recognition to simplify the 3-D recognition process and increase the recognition rate of 3-D objects. The paper presents an affine invariant feature extraction strategy for different 3-D objects that can recognize arbitrary postures of the 3-D object. Tests with models from the Princeton shape benchmark show that the good recognition effect can be achieved by the proposed method and the recognition rate is great than 90%.

**Key words:** pattern recognition; 3-D object recognition; viewpoint space partitioning; affine invariant feature

的很好的方法。但现有的三维识别算法都很复杂, 不能达到实时的要求。如何简化三维识别算法是目前的研究热点。经过美国Princeton 大学三维物体库的测试发现现有的三维识别算法没有一种通用的算法对各种物体的分类效果都很好, 也没有一种物体能使用任何分类算法都很好分类。已有15种算法在Princeton 平台上进行了测试, 基于二维投影的算法明显优于基于三维物体的算法<sup>[1]</sup>。基于二维投影的算法是对三维物体进行投影, 由二维投影图像表达三维物体进行处理识别。投影方法主要是均匀投影和非均匀投影。均匀投影的精度取决于投影的步长。步长过长, 一些本来存在差别的视图可能无法区分; 步长过短, 计算的复杂度会大大增加。非均匀投影可以改善这个问题, 但它的关键问题在于如何合并类似投影的。Zernike 矩<sup>[2]</sup>具有平移、尺度和旋转不变性, 是一种比较简单的形状描述子, 在已有的基于二维投影的三维识别中有所应用, 它可以比较好地描述物体的形状; 多尺度自卷积(multi-scale autoconvolution, MSA)特征是Rahtu等<sup>[3]</sup>提出的一种仿射不变性特征, 它在理论上可以抵抗仿射变形。文[4]中, 使用二维字母图像和Coil-100图像库测试了MSA特征, 得到MSA确实可以有效地抵抗仿射变换。基于Trace变换的特征是Petrou和Kadyrov<sup>[5]</sup>提出的一种仿射不变性特征, 在二维飞机识别和路标识别上表现良好。

本文将上述3种有效的抵抗仿射变性的特征应用于三维识别, 实现任意姿态下的三维物体识别。通过对3种特征的分析, 本文提出一种针对不同类型目标的仿射不变性特征提取策略。

收稿日期: 2007-09-27

基金项目: 国家自然科学基金资助项目(60502013);

国家“八六三”高技术项目(2006AA 01Z115)

作者简介: 孙洁(1983—), 女(蒙古), 内蒙古, 硕士研究生。

通讯联系人: 马惠敏, 副教授, E-mail: mhmpub@tsinghua.edu.cn

实现任意姿态下的物体识别是自动驾驶、自动驾驶技术实现的难点。三维识别加入了更多的关于目标的信息, 更符合人的感知系统, 是解决这个问题

### 1 视点空间采样

本文实现的三维识别系统有2大部分组成,训练部分和识别部分,如图1所示。训练部分又分为投影图获取、特征提取、视点空间划分、训练分类器几部分。首先,对所输入的三维物体投影,得到各角度投影的海量二维投影图像;其次,对各投影图像提取仿射不变性特征,聚类分析视点空间划分,将一个三维物体由其不同视点下的二维图像组合表达;最后,用这些信息训练分类器。识别部分使用由训练得到的分类器对测试图像分类识别。

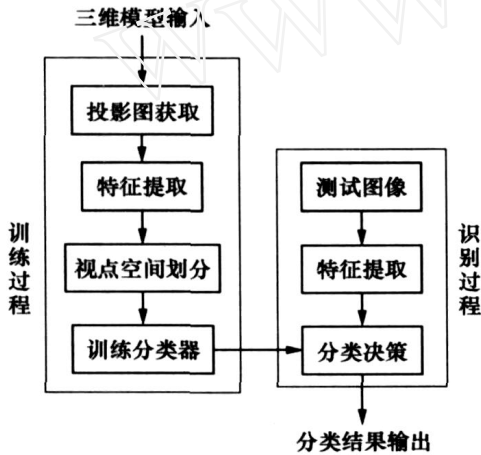


图1 三维识别系统框图

三维物体的投影图的获取如图2所示,在整个球面上对三维物体进行密集的等步长采样,得到整个空间内的投影图。

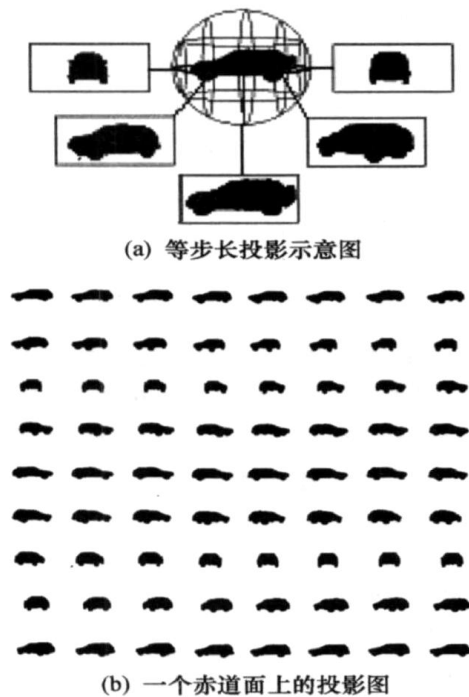


图2 投影示意图

本文的研究对象主要是飞机、汽车、人等三维物体。在现实世界的三维物体中,飞机、汽车、人是比较有代表性的三维物体。对于飞机模型,它在形状上外轮廓多棱角,轮廓复杂,汽车模型则比较典型的光滑凸面体,人的模型是轮廓比较光滑的复杂形状。因此,本文针对这3类模型进行研究测试,有其代表意义。在Princeton库<sup>[1]</sup>中选取各类的几个模型进行训练、测试,其中有4种飞机,2种汽车和人的模型,如图3所示。对每个模型在赤道面上每隔5°一次投影,投影一周后得到72幅图像,用于下面的特征提取视点空间划分。



图3 模型库代表图像

### 2 特征提取及视点空间划分

#### 2.1 特征提取

投影可以分成中心投影和平行投影,实际中的成像过程一般都是中心投影过程。但是,由于本文针对的飞机、汽车等人造交通工具是属于刚体,其透视变换可以近似为仿射变换;另外在成像设备与物体相隔足够远后,中心投影也可以近似成平行投影。因此,提取仿射不变性特征来消除小视点变化对投影的影响,划分视点空间。

提取的仿射不变性特征有 Zernike 矩<sup>[2]</sup>、MSA<sup>[3]</sup>和基于T race 变换的特征<sup>[5]</sup>。

Zernike 矩本质上是单位圆上的一组正交基,是径向上的 legendre 多项式和切向上的三角多项式的乘积,对于径向参数为  $n$ , 切向参数为  $m$  的 Zernike 矩的表达式为

$$Z_{nm} = \frac{n+1}{\pi} \int_0^{2\pi} \int_0^1 f(x_i, y_j) V_{nm}^*(x_i, y_j) \rho^{n-2s} \rho^s d\rho d\theta \quad (1)$$

其中:

$$V_{nm}(x_i, y_j) = V_{nm}(\rho \cos \theta, \rho \sin \theta) = R_{nm}(\rho) \exp(jm\theta), \quad (2)$$

$$R_{nm}(\rho) = \sum_{s=0}^{(n-|m|)/2} (-1)^s \rho^{n-2s}$$

$$s! \left[ \frac{n+|m|}{2} - s \right]! \left[ \frac{n-|m|}{2} - s \right]! \quad (3)$$

$f(x_i, y_j)$  表示图像的灰度函数。

理论上, Zernike 矩类似于把信号展开成一系列的系数,选取系数越多,恢复图像形状时精度越

高。它可以达到任意高的维数要求。作为识别分类, 前36个矩已经可以满足需要; 因此, 本文选取前36维。Zernike 矩它具有平移、放缩和旋转不变性。

M SA 是从概率论的角度出发, 构造一个仿射不变量。对于图像  $f(x)$ , 定义概率密度为

$$p(x) = \frac{f(x)}{\int f(x)}$$

选取3个独立的随机变量, 组成新的变量

$$u = \alpha(x_1 - x_0) + \beta(x_2 - x_0) + x_0 \quad (4)$$

经过仿射变换  $x = Tx + t$ , 式(4)变为

$$u = \alpha(x_1 - x_0) + \beta(x_2 - x_0) + x_0 \quad (5)$$

由  $u = Tu + t$ , 得

$$f(u) = f(T^{-1}u - T^{-1}t) =$$

$$f(T^{-1}(Tu + t) - T^{-1}t) = f(u) \quad (6)$$

得到的图像概率在仿射变换下是相同的, 故其  $k$  阶矩相同, M SA 为其一阶矩。因此, 它在理论上完全可以抵抗可矩阵化的仿射变换。同时它将概率运算中出现的复杂的卷积都转化为快速 Fourier 变换后的乘积运算, 大大降低了其运算复杂度。最终的表达式为

$$F(\alpha, \beta) = E[f(u_{\alpha\beta})] =$$

$$\frac{1}{f(0)^3} \int \int \int \hat{f}(-\xi) \hat{f}(-\alpha\xi) \hat{f}(-\beta\xi) \hat{f}(-\gamma\xi) d\xi \quad (7)$$

其中:  $\hat{f}$  为  $f(x)$  的 Fourier 变换,  $\gamma = 1 - \alpha - \beta$

因此, M SA 特征根据  $(\alpha, \beta)$  的不同选取, 有无穷多的维数。但是由于计算的复杂性, 即使使用了快速 Fourier 变换, 其运算量在点数很高的时候仍不可小视, 因此在考虑到计算量, 仅在  $(0, 1)$  范围内取值, 取19维的M SA 特征。

Trace 变换是广义的 Radon 变换, 它将整幅图像由一根根与原点的垂线的角度和离原点的距离确定的 Trace 线覆盖, 对每根线上的数据以函数  $T$  进行处理, 这样将原始图像转换成为一个由角度和距离为变量  $(p, \phi)$  的二维函数。在对这个函数在方向  $(p)$  和距离  $(b)$  坐标上分别以函数  $P$  和函数  $\phi$  处理, 由这3个函数本身的性质消除仿射变换的影响, 最后得到一维仿射不变的特征。其关键在于3种处理函数的选择, 对于一组处理函数, 特征提取过程如图4所示。选择4个  $T$  函数6个  $P$  函数14个  $\phi$  函数,  $4 \times 6 \times 14$  共336维特征, 再去除零特征后, 使用278维特征。

## 2.2 视点空间划分

三维物体的最终是由一系列不均匀采样的投影

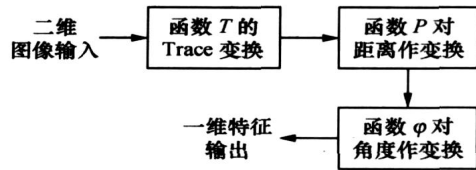


图4 使用基于 Trace 变换特征流程图

图表示的。以上得到了均匀采样后的投影图及其各自的特征, 采用最小距离的聚类方法, 将上面均匀采样后的投影图根据特征的 Euler 距离聚类, 划分视点空间。如图5所示, jeep3 模型在赤道面上均匀采样72幅图像后, 用 Trace 特征划分视点空间得到的15幅代表视点投影图像。

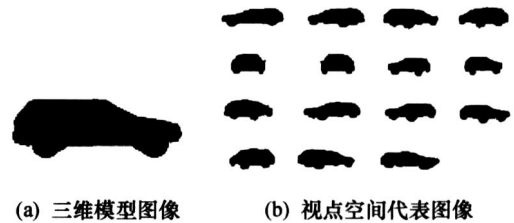


图5 视点划分后代表视点投影图像

## 3 实验结果

由于采样步长的限制得到的投影图是有限的, 经过视点划分后, 每个视点所包含的图像也是很少的, 所以在分类器上, 选择适于小样本分类的支持向量机 (support vector machine, SVM) 分类器<sup>[6]</sup>和原理, 算法简单的最近邻分类器。

训练库是 Princeton 库中选取的7个模型在赤道面上每隔  $5^\circ$  一次投影, 投影一周后得到72幅图像, 经过视点空间划分后, 每个模型划分为15个视点, 7个模型共504幅训练图像用于训练 SVM 过程。测试库图像为赤道面上每隔  $13^\circ$  的图像, 一周28幅, 7个模型共196幅。对其中一幅使用基于 Trace 变换特征的结果图如图6所示。

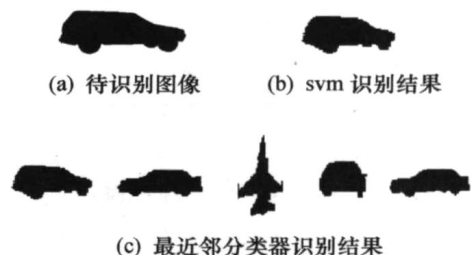


图6 用基于 Trace 变换特征的结果图

实验分成3组, 第1组是在整个测试库上测试

整体的识别效果;第2组是在飞机、汽车、人类间测试识别效果,最后一组在飞机、汽车类内测试识别效果,实验结果及分析如下。

在整个测试库上得到的结果见表1和表2。识别速度是在处理器P41.86G,512MB内存的计算机上得到的结果。从表1和表2可以看出,采用T trace和M SA特征,识别率明显好于Zernike矩,使用SVM分类器识别率都在90%以上,但其耗用时间也是成倍增长,T trace特征在时间和准确度有一个最好的折衷,SVM分类器的效果明显好于最近邻分类器,但在最近邻的前5个中基本可以找到结果。从此可以看到本文提出的方法是有效的,能得到较好的识别效果。

表1 测试库识别结果

| 特征        | SVM 错误数 | 最近邻错误数 |      |      |
|-----------|---------|--------|------|------|
|           |         | 第1近邻   | 前2近邻 | 前5近邻 |
| Zernike 矩 | 41      | 49     | 33   | 4    |
| T trace   | 9       | 12     | 8    | 5    |
| M SA      | 9       | 21     | 10   | 1    |

表2 测试库识别时间结果

| 特征        | 特征提取时间/m s | 识别时间/m s |      |
|-----------|------------|----------|------|
|           |            | SVM      | 最近邻  |
| Zernike 矩 | 973        | 1.18     | 0.13 |
| T trace   | 1343       | 1.63     | 2.93 |
| M SA      | 12896      | 1.01     | 0.08 |

表3表示了测试库中得到的类间识别的识别结果。从表3可以看出M SA特征在类间的分类效果最好,在测试结果中,使用Zernike矩特征很难区分人和飞机模型这2种物体,T trace特征较之有很大提高,而M SA特征在这2种物体的分类有很好的结果。

表3 测试库类间识别测试结果

| 特征        | SVM 错误数 | 最近邻错误数 |
|-----------|---------|--------|
| Zernike 矩 | 32      | 35     |
| T trace   | 9       | 8      |
| M SA      | 4       | 9      |

表4显示的是测试库中对飞机类内和汽车类内识别的结果。从表4中看出,Zernike矩虽然在飞机类内的识别效果不是很好,但对于汽车这类模型的类内识别效果却表现很好,M SA在类间、总体及飞

机类内都表现不错,但在汽车的类内识别劣于另外2种特征。

表4 测试库类内识别测试结果

| 特征        | SVM 错误数 |    | 最近邻错误数 |    |
|-----------|---------|----|--------|----|
|           | 飞机      | 汽车 | 飞机     | 汽车 |
| Zernike 矩 | 15      | 0  | 13     | 1  |
| T trace   | 0       | 0  | 1      | 3  |
| M SA      | 0       | 5  | 8      | 8  |

由上述结果看出,基于T trace变换的特征综合效果最优。但是,仅在处理比较光滑的凸面体类似汽车模型时,Zernike矩在速度及准确度方面都有极大优势,而M SA特征对于目标的形状比较复杂的情况下分类效果较好。

## 4 结论

实验结果表明,没有一种特征适用于所有模型。在处理形状比较光滑的凸面体时,选用Zernike矩综合性能好,在处理形状比较复杂的模型时,使用M SA特征识别效果比较好。在处理各种形状的模型都比较多的情况时,使用基于T trace变换的特征综合性能比较好;应用本文的飞机、汽车和人模型测试,识别率在90%以上。

## 参考文献 (References)

- [1] Shilane P, Min P, Kazhdan N, et al. The princeton shape benchmark [C]//Proceeding of the IEEE Shape Modeling International 2004(SMI'04). Washington DC: IEEE Computer Society, 2004: 167-178
- [2] YE Bin, PENG Jiaxiang. Analysis and improvement of invariance of zernike moments [J]. *Infrared and Laser Engineering*, 2003, 32(1): 37-41.
- [3] Rahtu E, Sabo M, Heikkilä J. Affine invariant pattern recognition using multiscale autoconvolution [J]. *IEEE Trans on Pattern and Analysis and Machine Intelligence*, 2005, 27(6): 908-918
- [4] Rahtu E, Heikkilä J. Object classification with multi-scale autoconvolution [C]//Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04). Washington DC: IEEE Computer Society, 2004, 3: 37-40
- [5] Petrou M, Kadyrov A. Affine invariant features from the trace transform [J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2004, 26(1): 30-44
- [6] Hsu C W, Chang C C, Lin C J. A practical guide to support vector classification [EB/OL]. (2007-09-10). <http://www.csie.ntu.edu.tw/~cjlin>