

基于 Zernike 矩的三维目标多视点特性视图建模*

彭 辉, 黄士科, 陶 琳, 张天序

(华中科技大学 图像识别与人工智能研究所, 图像信息处理与智能控制教育部重点实验室,
湖北 武汉 430074)

摘要: 目标的模型化是三维目标识别的一个重要环节。由于从多视点特征法建立的模型能够避免从二维图像中恢复目标的三维结构, 在三维目标识别中应用比较广泛。使用镶嵌式多面体分割视觉空间以得到三维目标的多视点投影图像, 采用较小的镶嵌单元尺寸以完备地描述目标各种姿态的信息。Zernike 矩具有旋转不变性, 并能够方便地由各阶矩重建图像。计算多面体各个单元的投影图像的 Zernike 矩, 所需的矩的最高阶次由重建图像与原图像的差别确定, 通过聚类方式合并那些图像的矩变化很小的投影对应的几个小单元, 每个新单元可以提取一幅特性视图, 实现较小数目的多视点特性视图的建模。实验证明了这种方法的可行性。

关键词: 三维目标识别; 特性视图; 视觉空间; Zernike 矩

中图分类号: TP391 **文献标识码:** A **文章编号:** 1007-2276(2005)03-0292-05

Multi-view modeling of 3-D target based on Zernike moments*

PENG Hui, HUANG Shi-ke, TAO Lin, ZHANG Tian-xu

(Key Laboratory of Ministry of Education for Image Processing and Intelligent Control, Institute for Pattern Recognition and Artificial Intelligence, Huazhong University of Science and Technology, Wuhan 430074, China)

Abstract: Modeling is an important step in 3-D target recognition. Avoiding the recovery of 3-D object, multi-view modeling is used widely in 3-D object modeling. Multiple 2-D views are obtained by projecting a tessellated regular polyhedron onto the viewing sphere. In order to describe all shape and pose information, a fine partition of the viewing sphere is necessary, which leads to computation complexity. Zernike moments have a rotation invariant property. A method is used to decrease the number of 2-D projections view by extracting the Zernike moments of each 2-D view, whose max order is decided by the difference between the original image and the reconstruction one. Those partitions with Zernike moments of corresponding views are similar are merged and a characteristic view is obtained for a new partition. Experimental results show the validity of the method.

Key words: 3-D target recognition; Characteristic view; Viewing space; Zernike moments

收稿日期: 2004-06-08; 修订日期: 2004-07-15

* 基金项目: 国家自然科学基金重点项目(60135020); 国防科技预研重点项目(413010701)

作者简介: 彭辉(1980-), 男, 湖北安陆人, 硕士, 主要从事目标识别、图像信息处理方面的研究。

0 引言

三维目标识别是计算机视觉的一个重要研究内容,是通过目标在二维平面上的投影所形成的图像进行识别的,由于三维目标在不同视点呈现不同的姿态,使得三维目标的识别成为一个比较复杂的课题。多视点特征法在三维目标识别中得到了广泛应用。多视点特征是用不同视点的多个二维投影描述三维目标,目标可以与二维图像中提取出的特征直接匹配,避免了从二维图像中恢复目标的三维结构问题;另一方面,近年来在二维图像的目标识别研究中,图像区域不变矩集合能够较好地反映形体特征,由矩构成的模式特征被广泛应用于工业、医学以及军事上,取得了不错的成果。设想将具有良好不变特性的矩从二维拓展到三维,用于多视点三维模型,这是一个很有吸引力的课题。本文尝试在计算 Zernike 矩的基础上进行三维目标的模型化。通过将尺寸较小的镶嵌式规则多面体投影到视图球上得到能够充分描述三维目标的诸多投影图像,计算每幅图像的 Zernike 矩,使用聚类的方式减少视图数目,提取目标的特性视图,将三维目标的立体形状信息缩减至若干个特性视图,识别只需在这些特性视图上进行。

1 Zernike 矩特性

自从 Hu 提出并构造矩的七个不变量函数以来,大量学者研究了矩的快速计算、几何矩的推广,构造能够在更多几何变换下保持不变的矩。但是,几何矩是定义在非正交的基集 $\{x^p y^q\}$ 上的,由该基集构造尽可能多的不变矩不仅计算复杂,而且得到的矩信息冗余,增加了识别环节的计算复杂度,另外,使用这些矩目前无法方便地恢复重建图像^[1]。几何矩先天的不足无疑制约了其更大范围的应用。由 Alireza 引入的 Zernike 矩是定义在圆盘上的正交矩,具有简单的旋转变换性质,很容易构造旋转不变量——旋转目标图像不会改变图像 Zernike 矩的大小,可用作旋转不变量特征来描述图像。另外,与几何矩不同,这些矩特征很容易创建到任意高阶,而且这些矩便于重建图像。因此,Zernike 矩的出现引起了人们极大的兴趣。Al-

Rawi Mohammed 研究了 Zernike 矩的快速算法^[2],Simon 等研究了 Zernike 的精度问题^[3],参考文献^[4]推广了 Zernike 的尺度不变性。这些研究大多关注于 Zernike 矩在二维平面的应用,如何将其用于三维空间,扩展其使用范围,仍然值得探索。

1.1 Zernike 的定义

Zernike 引入了一个复数多项式集,这个复数多项式集在单位圆内,即 $x^2+y^2=1$ 内形成一个完备正交集,用 $\{V_{nm}(x,y)\}$ 来表示这个多项式集,这些多项式的形式为:

$$V_{nm}(x,y)=V_{nm}(\rho,\theta)=R_{nm}(\rho)\exp(jm\theta)$$

式中

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

为多项式,其中 n 为正整数或者零; m 为小于 n 的正整数或负整数; $n-|m|$ 为偶数, $|m| \leq n$; ρ 为从原点到点 (x,y) 的向量的长度; θ 为 ρ 与 x 轴在逆时针方向的夹角。

Zernike 矩是图像函数在这些正交基函数上的投影。由于 Zernike 矩定义在单位圆内,圆外的点不参与计算,在计算给定图像的 Zernike 矩时,需要把图像的中心作为原点,图像中像素点的坐标被映射到单位圆内,即 $x^2+y^2 \leq 1$ 。在单位圆外趋近于零的连续图像函数的 m 重 n 阶 Zernike 矩为:

$$A_{nm}=\frac{n+1}{\pi} \int \int_{x^2+y^2 \leq 1} f(x,y) V_{nm}^*(\rho,\theta) dx dy$$

式中 $V_{nm}^*(\rho,\theta)$ 为 $V_{nm}(\rho,\theta)$ 的共轭函数。对于离散的数字图像,上式中的积分被求和代替,变为:

$$A_{nm}=\frac{n+1}{\pi} \sum_x \sum_y f(x,y) V_{nm}^*(\rho,\theta), x^2+y^2 \leq 1$$

1.2 Zernike 的不变性质

Zernike 矩具有简单的旋转不变性质,旋转时每个 Zernike 矩只是相位发生了变化,模值不变。若以 A_{nm} 、 A'_{nm} 分别代表旋转前后图像的 n 阶 m 重 Zernike 矩,可以证明^[5]:一幅旋转了角度 α 的图像,其每阶 Zernike 矩与图像旋转前对应的 Zernike 矩之间有如下

关系, $A'_{nm} = A_{nm} \exp(-jm\alpha)$ 。从这个简单的性质可知, 一个图像函数的 Zernike 矩的模值大小在图像旋转前后保持不变。这样, Zernike 矩的大小 $|A_{nm}|$ 可以作为图像函数的旋转不变量特征。

Zernike 矩只具有旋转不变性。为了方便地通过特征向量识别目标, 需要构造二维平面上的旋转、平移、缩放不变量。为获得尺度和平移不变量, 在提取旋转不变量之前, 先要对图像进行尺度和平移标准化处理(由于在将图像映射到单位圆这一坐标转换时即完成了平移标准化过程, 实际上只需进行尺度标准化), 然后从标准化后的图像中提取旋转不变的 Zernike 特征。使用 B2 飞机图像进行了实验, 表 1 为 B2 飞机旋转前后 Zernike 矩的部分实验结果, 表中 μ 为样本均值, σ 为样本标准差, $\sigma/\mu\%$ 表示 $|A_{nm}|$ 的值偏离相应均值的百分数, 实验数据表明, 旋转前后, Zernike 矩的不变特性保持良好。

表 1 B2 飞机图像在不同旋转角度下 Zernike 矩的部分示例

Tab.1 Sample of normalized B2-images rotated different angles

Rotation angle	$ A_{20} $	$ A_{22} $	$ A_{31} $	$ A_{33} $
30°	109.073	7.833	10.092	7.504
45°	108.267	7.513	9.386	7.157
60	107.762	7.539	9.664	7.147
76°	107.725	7.646	9.874	7.314
90°	107.575	7.586	9.740	7.193
120°	107.274	7.506	9.629	7.114
135°	108.115	7.741	9.987	7.344
150°	107.911	7.652	9.879	7.284
μ	107.963	7.627	9.781	7.257
σ	0.509	0.108	0.2098	0.122
$(\sigma/\mu)\%$	0.471%	1.418%	2.145%	1.684%

2 多视点投影模型^[6-8]

2.1 多视点模型简介

解决三维目标识别的一个途径是使用多个二维视图描述三维目标, 将三维目标识别的问题转化为二

维图像中目标识别问题。因此首先要计算出目标的一组有限的二维视图, 这组二维视图能够充分代表三维目标在空间的各种姿态形状, 实际获取目标的二维图像后, 使用这幅图像与上面得到的那组视图进行匹配识别。问题的关键是这一组视图集合要尽可能小并能代表无限数目的可能视图。

三维目标不同姿态的投影成像即是三维目标在不同成像平面处的正交投影。为方便研究建模, 不考虑成像平面与目标的距离, 即认为成像平面与目标等距离。三维目标向各平面正交投影方向形成的空间就是从各平面投影点处观察目标所形成的视觉方向空间, 用一个单位球表示。这个单位球上的每一点对应相同坐标下的方向向量。假设一个被缩小的无穷小目标置于球体中心, 沿视觉方向 P 的正交投影对应于从球上 P 点看目标, 我们把这个球称作视觉球, 球上的点被称作正交投影的视点, 如图 1 所示。

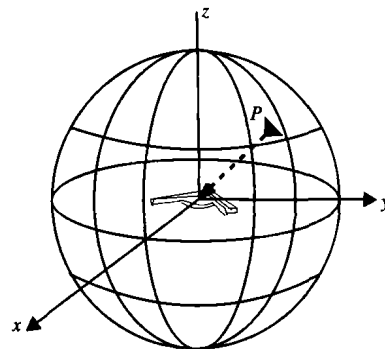


图 1 视觉空间的单位球表示

Fig.1 Viewing space represented by a unit sphere

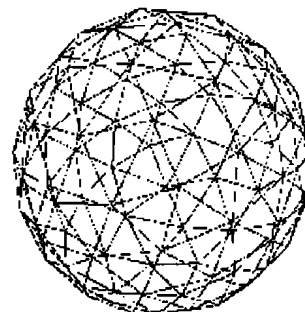


图 2 80 面体对视觉空间的均匀划分

Fig.2 Viewing space partitioned uniformly by projecting a tessellated regular polyhedron with 80 faces

2.2 视觉空间的划分

因为想要得到描述不同姿态的三维目标的投影,

必须把视觉球划分为一个个小区域,每个区域描述目标的一种姿态。又因为,某些不同区域描述的目标姿态很接近,从这些区域得到的诸多二维投影的某些特征保持不变或变化很小,可以认为这些小区域属于同一类型,可将其合并,区域内的视点称为一般视点。而从另外一些区域得到的三维目标的诸多二维投影差别很大,那些特征发生了较大的变化,认为在这些区域边缘发生了视觉事件(如目标的一个面被遮挡或者几条不相交的棱边投影至一点),不能将这些区域合并,这些发生视觉事件的视点称为视觉视点。在合并后的每个区域中,从该区域的诸多视点得到的正交投影中选择一个代表性视图,来描述从该区域观察目标所得的目标形状姿态信息,这样每个区域的代表视图即为一个特性视图。在匹配识别阶段,识别只需要在未知图像与这些代表性视图之间进行。划分视觉空间的一个常用方法是均匀划分方法,即把一个镶嵌式规则多面体投影到球面上,认为每个小格就是一个一般视点的集合——一个分区。该方法分区容易,计算直接,但镶嵌式规则多面体的尺寸难以确定。为了得到物体比较全面的特性视图特性,必须精细地划分视觉空间,选择小尺寸的镶嵌单元,数量庞大的单元会导致巨大的计算量;另一种方法就是方位图或形态图方法,其基本思想是根据投影图像的视图结构不变划分视觉空间。根据各种视图原则,确定使图像的视图结构发生变化的各视觉视点,方位并没有统一的定义,根据应用各有不同,它强调的是视图的结构关系。方位图方法的不足在于可操作性不强,不能通过程序计算得到目标。所以此方面的工作以理论研究为主。结合均匀划分和方位图方法的思想,将视觉空间小尺寸均匀划分后,在每个分区均得到三维目标的一个特性视图,使用 Zernike 矩描述这些小分区的特性视图,通过聚类将其中矩特征变化很小的分区合并,减少特性视图的数目,实现多视点建模。

3 实验结果

由于目标相对摄像机运动,在摄像机成像平面上得到的投影图像呈现不同的姿态,这种情况相当于目标静止、摄像机反向运动。假设将目标固定置于视觉球的中心,摄像机沿着球体的各个镶嵌面对目标进行

观察,得到目标多视点的二维投影。视觉球的中心目标为 B2 飞机模型,如图 1 所示。

3.1 视觉空间的初步划分

在对视觉空间的初步划分实验中,将 80 个面的镶嵌式规则多面体投影到球体表面,这样球体表面被均匀划分为 80 个区域,每个区域对应于一个单独的视点,每个视点均可得到目标在该视点的正交投影,一共得到目标在镶嵌多面体上的 80 个正交投影。80 个面的镶嵌式规则多面体如图 2 所示。

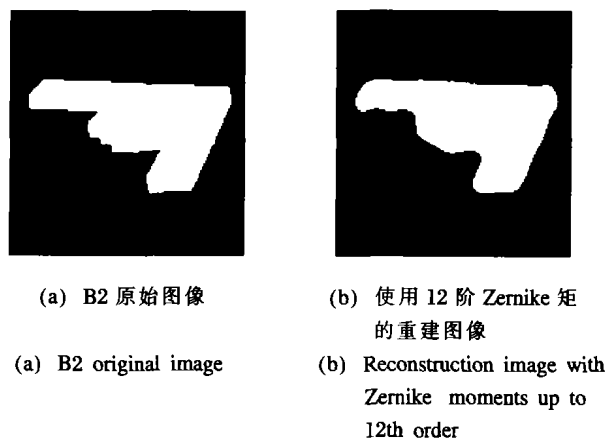


图 3 使用 Zernike 矩进行图像重建示例
Fig.3 Illustration of image reconstruction by Zernike moments

3.2 Zernike 矩特征提取

Zernike 矩的低阶矩描述了目标图像的大尺度信息,高阶矩描述了目标图像的细节信息,显然阶次越高的矩越能充分描述目标图像,体现图像特征;另一方面,尽管 Zernike 矩能够很容易地创建到任意高阶,但是阶次的增加对应于不变特征向量的维数的增加,无疑使后期的识别计算变得更加复杂。因此,权衡选择正确的最高阶 Zernike 矩十分重要。根据计算所得的 Zernike 矩可以方便地重建出原始图像,用 Hamming 距离来衡量 Zernike 矩的描述能力。Hamming 距离定义为重建图像与原始图像的误差,即不相同像素的个数。实验使用了最高阶次为 12 的 Zernike 矩,这 12 阶矩重建所得图像与原始图像的相似度为 97.4%,Hamming 距离相对原始图像为 2.6%,这种误差是可以接受的,即 12 阶 Zernike 足以描述图像特征而参与识别。图 3 为重建示例。

3.3 特性视图的提取

视觉空间被均匀划分为 80 个分区后,每个分区对应一幅二维图像,这 80 幅图像之间存在巨大的数

据冗余,可以将矩特征变化很小甚至不变的视觉空间的分区合并,合并后的每个分区提取出对应的特性视图。实验使用了动态聚类方法中的 C -均值法进行聚类,选择每个聚合类中的代表作为特性视图。 C -均值的基础是误差平方和准则。设 N_i 是第 i 聚类 Γ_i 中的样本数目, m_i 是这些样本的均值,即 $m_i = \frac{1}{N_i} \sum_{y \in \Gamma_i} y$,把 Γ_i 中的各样本 y 与均值 m_i 间的误差平方和对所有类相加后为 $J_c = \sum_{i=1}^c \sum_{y \in \Gamma_i} \|y - m_i\|^2$, J_c 是误差平方和聚类准则,度量了用 c 个聚类中心 m_1, m_2, \dots, m_c 代表 c 个样本子集 $\Gamma_1, \Gamma_2, \dots, \Gamma_c$ 时所产生的总的误差平方。对不同的聚类, J_c 的值是不同的,使 J_c 极小的聚类即为误差平方和准则下的最优结果。实验中,由于 c 是未知的,假设类别逐步增加,对 $c=1, c=2, \dots$ 分别使用 C -均值算法。误差平方和随着 c 的增加而单调减小。当数据集中表现为 $c_{\text{最优}}$ 个很集中的聚类时, J_c 随着从一个聚类增加到 $c_{\text{最优}}$ 个聚类而迅速减小。当 c 再增加时,相当于将本来较密集的群再行分开,因此, J_c 虽有所减小,但减小的速度比较缓慢,直到 $c=N$ 时 $J_c=0$ 。根据 J_c - C 的曲线,可以得到曲线上拐点对应的类别数,这个类别数比较接近最优的聚类数 $c_{\text{最优}}$ 。如图 4 所示, A 点即为拐点,所以, 7 是比较合适的聚类类别数。

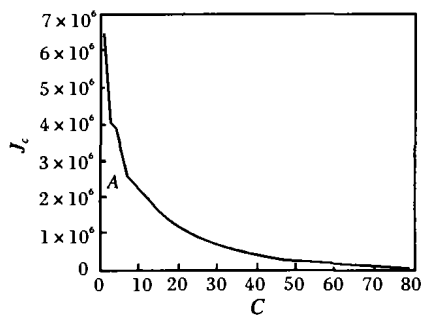


图 4 聚类 J_c - C 曲线图

Fig.4 Errors(J_c) under the influence of various numbers (C) of categories

图 5 为得到的 B2 的 7 个特性视图。因为 B2 的形状比较规则,得到的特性视图数目很小,同时也注意到由于 Zernike 矩的仿射不变性较弱,所得特性视图仍有继续缩减的空间,值得进一步研究。

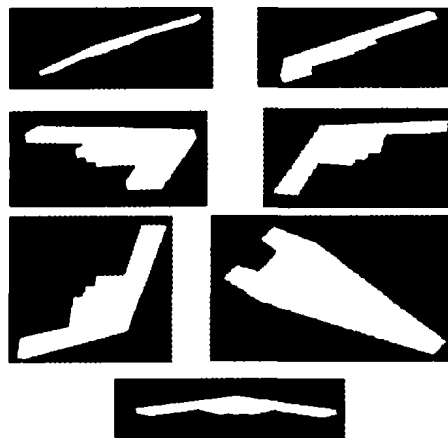


图 5 B2 的 7 个特性视图

Fig.5 7 characteristic views of B2

4 结 论

结合视觉空间划分中的均匀划分和方位图的思想,提出了一种新的视觉空间划分方法,这种方法利用了旋转不变 Zernike 矩构成的特征向量。通过对视觉空间的划分,能得到三维目标的很有限的若干个特性视图,这些特性视图代表在不同视点观察三维目标获得的形状姿态信息,从而完成对三维目标的建模。

参考文献:

- [1] Ziv Gigu, Jitendra Malik. Computing the aspect graph for line drawing of polyhedral objects[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996, 12(2): 113-122.
- [2] Al-Rawi Mohammed, Yang Jie. Practical fast computation of Zernike moments[J]. J Comput Sci & Technol, 2002, 17(2): 181-188.
- [3] Simon X Liao, Miroslaw Pawlak. On the accuracy of Zernike moments for image analysis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(12): 1358-1364.
- [4] 叶斌, 彭嘉雄. Zernike 矩不变性分析及其改进[J]. 红外与激光工程, 2003, 32(1): 37-41.
- [5] Alireza Khotanzad, Yaw Hua Hong. Invariant image recognition by Zernike Moments[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990, 12(5): 489-497.
- [6] 张海燕, 宋克欣, 王东木. 基于方面图技术的三维运动目标识别[J]. 中国图象图形学报, 2003, 8A(8): 866-871.
- [7] 席学强, 王润生. 一种用于遥感图像 3D 目标识别的基于形态的建模方法[J]. 计算机应用研究, 2000, 11: 6-9.
- [8] 席学强, 王润生. 计算机视觉模型中的知识及其表示[J]. 计算机工程与应用, 2000, 11: 45-47.