

# 基于距离信息扩展原理的物体识别方法

陈向荣 林学闯 朱志刚 石定机

(清华大学计算机科学与技术系,北京 100084)

**摘要** 利用稀疏距离图象与灰度图象相结合,通过射影变换可使灰度图象特征有条件地获取距离信息,这称为基于射影变换的距离信息扩展原理。利用这些距离信息可计算物体表面平面图形的欧氏空间不变量并对景物中的物体产生假设。由于初始假设建立在可靠的基础上,并使用了带反馈的混合控制策略,使这种基于模型的物体识别定位技术具有实用的前景。此文讨论了整个方法的原理及其实施方案,并实验验证了原理的正确性与有效性。

**关键词** 基于模型的视觉,物体识别,数据融合,射影变换,距离信息扩展

## 1 引言

基于模型的物体识别方法可按使用二维灰度数据和三维深度数据划分为两大类型<sup>[1,2]</sup>。这两种类型的方法各有其优缺点<sup>[3]</sup>。二维灰度图象的获取和底层处理较为容易,但是所获得的二维数据与三维模型之间的匹配则非常复杂和困难;相反,完好的三维数据与三维模型之间的匹配相对容易,但困难在于如何快速获取完整可靠的三维特征。另外,单独使用深度数据或灰度数据都不能完整地描述所有的景物特征。深度数据具有描述光滑表面的能力,而灰度数据能够提供表面反射和阴影轮廓信息。这种互补性使得二者的结合成为一个有希望的研究方向。在深度数据与灰度数据结合的研究中,早期的工作使用简单的启发式方法,例如:Duda等<sup>[4]</sup>先用Hough变换在深度图象中检测平面块,然后通过分析灰度图象的直方图来获得更多的表面。Gil等<sup>[5]</sup>用“与”操作把深度图和灰度图中找到的边缘结合起来。Chu等<sup>[6]</sup>用表面拟合的方法分割深度图象,用统计的方法分割灰度图象,然后组合二者得到边缘图象。目前研究较多的方法是利用深度数据与灰度数据对图象进行区域分割与分类,利用分类后的区域信息完成对物体的识别。Zhang等<sup>[3]</sup>利用标定的深度图象与灰度图象对边缘进行语义上的标记并对深度数据进

行重建。Asada<sup>[7]</sup>利用投影到物体表面上的条纹光计算表面法向,并与从相应灰度图象中提取的边缘信息一起形成对景物的描述。Wang等<sup>[8]</sup>通过网格光计算表面法向,然后利用灰度图象与结构光图象的融合重建三维物体形状。

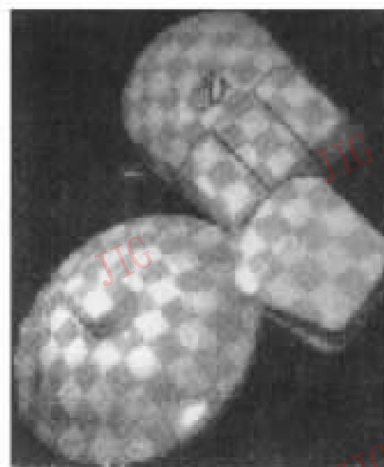


图1 稀疏网格光照下的物体图象  
Fig.1 Sample image.

物体识别定位方法又可按所适用的物体类型划分,分为简单形体的物体和复杂形体的物体。所谓简单形体的物体是指其表面能够用简单表面(平面、二次曲面)近似的物体。目前,在简单形体的物体识别的方法中,大都以检测简单表面为

主要技术。这种方法一般要求距离图象较为致密,以便有足够的信息来获取和计算景物的二维边缘信息和三维表面形状信息,如法向、曲率等,进而完成对物体表面的分割和拟合,计算量较大。另一方面,相当数量的几何形状简单的物体其表面无法采用简单

表面拟合技术。(图 1 中杯盖,插头和鼠标器),表面形状相对简单,但用简单表面去拟合却十分困难。事实上,很多实际物体的表面上存在平面图形,例如平面的边缘,柱体或锥体的母线组成的轮廓线,球的大圆,图 1 中杯盖的外缘,插头的边缘,鼠标器表面间的交线等等。尤以杯盖为例,其表面是光滑过渡的曲面,使用拟合方法的困难在于如何准确地对表面进行分割和分类。但是,尽管其表面没有一块平面,它的外缘却是平面图形,可以很容易地利用网格线的间断点检测出来。这些平面图形对其所属的物体具有很强的描述能力,R. C. Bolles<sup>[9]</sup>对此有较为详细的论述。针对具有平面图形的物体,本文提出一种新的基于模型的三维物体识别方法。首先在图象中检测平面图形,通过相关距离信息建立射影变换。然后通过射影变换由平面图形的图象坐标获取其欧氏空间坐标,进而计算平面图形的欧氏空间不变量,并以这些不变量作为匹配与识别的依据。该方法的优点在于:(1) 采用灰度图象与稀疏距离图象结合的方法,对于距离图象的致密程度及灰度图象的质量要求都较低;(2) 不必拟合物体表面,而只需通过距离图象和灰度图象的结合检测平面图形,计算量较小;(3) 通过计算物体表面平面图形的欧氏空间坐标及不变量,得到较为可靠的三维特征,以此做为与模型库进行匹配的依据,从而提高了匹配和识别的效率。

## 2 原理描述

### 2.1 稀疏距离图象与灰度图象的获取

就目前的研究状况而言,致密距离图象的实时获取和处理有待进一步的研究,而稀疏距离图象的获取则相对简单,不少学者已提出了快速准确的获取方法。本文所采用的方法其原理与参考文献<sup>[10]</sup>中所述相似。摄取图象的装置由一台摄象机及一台投影器组成,景物中的物体由投影器光源及环境光源共同照射。因此灰度图象与投影器投射的网格图象可显现在同一图象上(图 1)。

利用标定网格图象获取距离信息关键在于对图象中的网格结点能够有效地进行辨认。利用网格结点在图象中只能沿各自相应的外极线轨迹移动的特点,合理选取网格方向与外极线方向,适当限制空间视场范围,选择合适的网格间距,可使对应于网格每个结点的外极线之间有足够分辨的距离,从而可从结点在图象中的坐标唯一地确定结点所在物体表面点的欧氏坐标,并可由此确定每条网格线上各点的

欧氏坐标,从网格图象获得了稀疏距离图象。

### 2.2 基于射影变换的距离信息扩展原理

采用上述方案获取的稀疏距离图象不仅为其所在图象点提供欧氏空间的坐标,其线条在图象中所呈现的形状还包含了物体相应表面的形状信息,因而可用来对物体表面平面图形进行检测。通过以下所述的基于射影变换的距离信息扩展原理,可以为这些平面图形赋予欧氏坐标值。

针孔模型的摄象机的图象点和空间点之间可以建立 2D-3D 对应关系。已知图象点  $U$  的齐次坐标  $(u, v, 1)^T$ , 空间点  $P$  的齐次坐标  $(x, y, z, 1)^T$ , 该对应关系可以表示为:

$$U \cong M_{3 \times 4} P, \quad m_{34} = 1$$

其中符号表示带比例因子的等价关系。这是一个一对多的对应关系,已知空间中一点就可以确定其在图象中的位置;而已知图象中一点则对应空间中通过摄象机光心的一条直线。若想从图象点唯一确定空间中一点,则需要附加约束条件。当图象中某个区域为空间中平面图形的成像时,上述关系可以得到进一步的简化。我们知道,在透视投影下对应的两个平面之间可以建立射影变换,而图象平面正是空间平面在透视投影下的像平面。已知空间平面方程后,不妨规定平面上点  $P$  的三维空间欧氏坐标  $(x, y, z)^T$  为其二维齐次坐标。空间处于同一平面的所有点的  $P$  与  $U$  之间可用一平面射影变换  $A$  表示,即

$$\forall P \in \pi \quad P \cong A_{3 \times 3} U \quad (1)$$

矩阵  $A$  可以利用距离图象中共平面的 4 个点的双重坐标  $(P, U)$  计算确定。式(1)的一个重要用途在于,如果空间一点的确属于该平面,则在  $A$  确定之后,可由其图象坐标  $U$ , 按式(1)计算齐次坐标  $P$ , 再利用平面方程的约束获取其欧氏空间坐标。这就是基于射影变换的距离信息扩展原理,也是本文提出的物体识别方法的数学基础。

通过平面图形的欧氏坐标,可以计算其各种欧氏不变量,并利用这些不变量与模型库中的物体模型建立匹配关系,从而产生物体识别的初始假设。

### 2.3 基于射影不变量的假设检验方法

对待识别物体产生了初始假设后,该物体的所有特征在欧氏空间的坐标可以按模型提供的数据计算与预测,这就为假设的检验创造了条件。对预测特征的检验仍可以采用射影几何原理来实现,与利用光心的假设检验方法相比,其优点在于不必对摄象

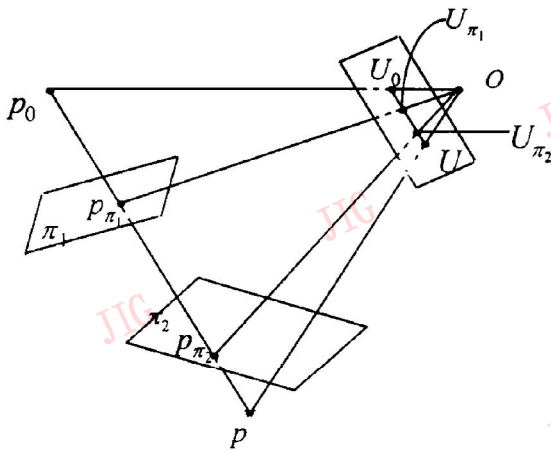


图 2 利用射影变换确定 P 与 U 的关系

Fig. 2 Determine the relation between P and U using projective transformation.

机的内外参数进行标定。此时所要解决的问题是,已知某空间点的欧氏坐标 P 预测其相应的图象坐标 U。

如果空间一点不在平面  $\pi$  上,它的 P 与 U 之间的关系不再能用式(1)表示。仿照 shashua 在<sup>[11]</sup>中的做法,P 与 U 之间的关系可借助两个参考面及一参考点进行计算(图 2)。设空间有两个参考平面  $\pi_1$  与  $\pi_2$ ,它们与图象之间的射影变换关系可分别表示为:

$$\begin{cases} P \cong A_1 U & \text{或 } U \cong A_1^{-1} P & \text{如果 } P \in \pi_1 \\ P \cong A_2 U & \text{或 } U \cong A_2^{-1} P & \text{如果 } P \in \pi_2 \end{cases} \quad (2)$$

假设 P 与另一空间点  $P_0$  的连线与该两平面  $\pi_1, \pi_2$  分别交于  $P_{\pi_1}, P_{\pi_2}$  两点,则应有  $U_{\pi_1} \cong A_1^{-1} P_{\pi_1}$  及  $U_{\pi_1} \cong A_2^{-1} P_{\pi_2}$ ,其中  $U_{\pi_1}, U_{\pi_2}$  分别表示  $P_{\pi_1}, P_{\pi_2}$  相应的图象坐标。如果  $P_0$  点的图象坐标为  $U_0$ ,按交比不变性应有

$$(UU_0, U_{\pi_1} U_{\pi_2}) = (PP_0, P_{\pi_1} P_{\pi_2}) \quad (3)$$

其中  $(\times \times, \times \times)$  表示共线 4 点的交比。因此在  $\pi_1, \pi_2$  及  $P_0$  确定的条件下,空间点 P 的图象坐标 U 可按式(3)计算其预测值,由预测的准确与否检验相应假设的正确性。

### 3 系统框图

利用上述原理进行物体识别定位的系统框图如图 3 所示。输入图象为经过标定的稀疏网格光照下的景物图象,可以由网格线得到稀疏的距离图象。首先根据距离图象中网格线的断点和转折点检测平面

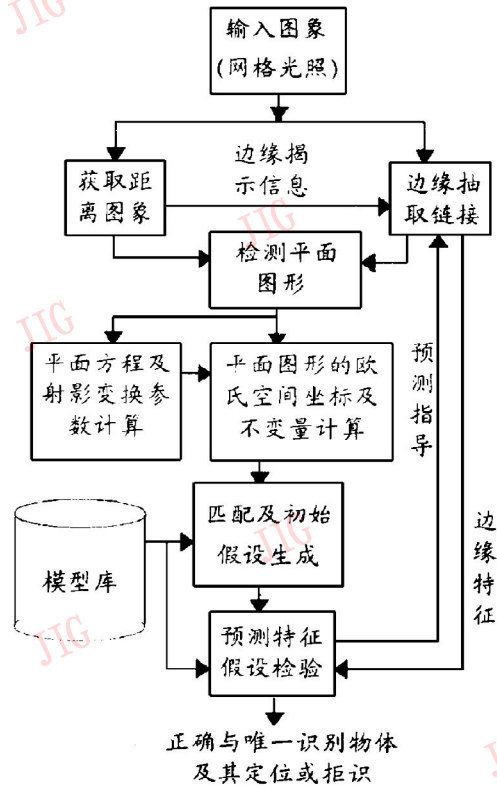


图 3 系统框图

Fig. 3 System diagram.

图形,并利用灰度图象的边缘图象对这些离散点进行链接,对于检测到的平面图形,可以通过相关距离信息和对应的灰度图象中的位置建立射影变换,然后通过射影变换由平面图形的图象坐标获取其欧氏空间坐标,从而可以进一步计算平面图形的欧氏空间不变量,利用这些不变量与模型库中的模型进行匹配,能够产生有效的初始假设。考虑到物体表面材料反射系数、光源条件、二次反射、阴影等多方面因素对抽取高质量的边缘图象的不利影响,系统采用带反馈的混合控制策略,即利用初始假设可以预测和指导灰度图象中低反差边缘特征的提取,并对初始假设进行验证。

假设检验的过程采用证据累积的方法,证据来自两个方面:由初始假设预测的灰度特征是否在灰度图象中出现;由初始假设预测的区域上的距离图象是否与模型相符合。当证据累积到一定程度时,认为假设成立,否则检验下一个假设。所有假设都不成立则对下一个平面图形进行处理。

物体的模型分为匹配模型和假设检验模型 2 个部分,分别用于对平面图形的匹配和对初始假设的检验。匹配模型,即平面图形的模型组织包含平面图形整体特征和局部特征。整体特征包括平面图形的

形状类型,面积,长度等。局部特征包括角度,曲率等。利用它们与景物中检测到的平面图形进行匹配,形成有关物体类型和方位的初始假设。根据假设检验的需要,物体的模型的另一部分以面和边的形式组织。得到了初始假设之后,可以利用可见边的信息预测灰度特征,利用可见面的信息预测它在距离图象中所处区域,利用这一区域中的距离图象信息检验是否与该面相符。

## 4 实验结果

使用本方法,我们对图 4 中一个杯盖进行识别

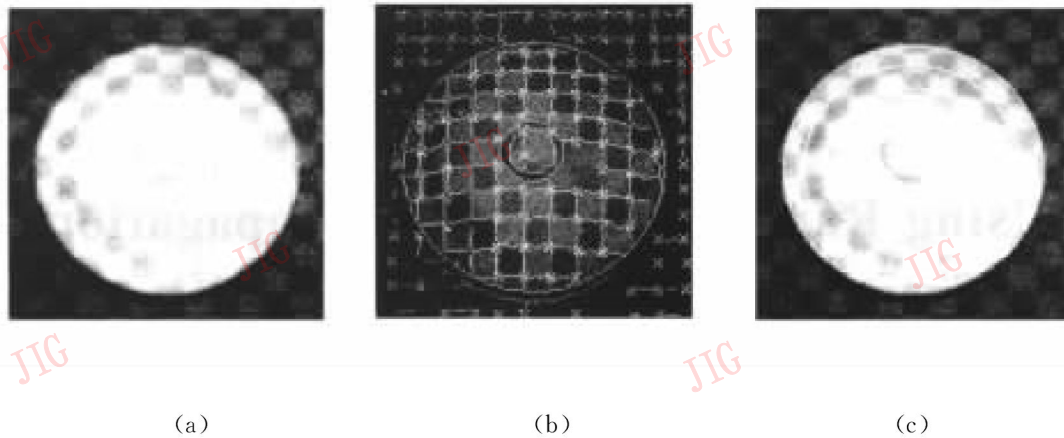


图 4 实验结果

Fig. 4 Experimental results.

检验时在模型的母线上取 3 个点,分别对应杯盖表面上的 3 个圆,根据假设预测这 3 个圆在图象中的

位置,与图象中的网格线求交点,判断交点处距离图象的值是否与模型相符。实验数据如下表所示:

	统计点数	平均误差(mm)	最大误差(mm)	标准值(mm)	检测值(mm)
拟合	12	0.3	0.6	48.4(半径)	48.1(半径)
假设检验	63	0.9	1.6	—	—

物体模型的视觉系统中检验。

## 5 结束语

通过稀疏距离图象及灰度图象结合,利用物体表面的平面图形实现物体识别定位的新方法。该方法利用两种信息的互补作用,提高了灰度图象特征提取的效率。初始假设的准确与歧义性小,为使用带反馈的混合控制策略创造了条件,并有利于降低对低层次处理的难度。这对提高本方法的实用性是十分重要的。使用这种方法的代价在于识别物体的类型受到一定的限制,即物体必须具有至少一个可检测的平面图形。上述方法将在一个具有一定数量的

## 参考文献

- 1 Lowe O G Three-Dimensional Object Recognition from Single Two-Dimensional Images. *Artificial Intelligence*, 1987, 31: 355~395.
- 2 Stein F, Medioni G. Structural indexing: Efficient 3D object recognition. *IEEE Trans PAMI*, 1992, 14(2): 125~145.
- 3 Zhang G, Wallace A. Physical modeling and combination of range and intensity edge data, *CVGIP: Image Understanding* 58, 1993, 191~220.
- 4 Duda R O, Nitzan D, Barrett P. Use of range and reflectance data to find planar surface regions, *IEEE Trans PAMI*, 1979,

- (3):259~271.
- 5 Gil B , Mitiche A, Aggarwal J K. Experiments in combining intensity and range edge maps. CVGIP 1983,21: 395~411.
  - 6 Chu C ,Nandhakumar N,Aggarwal J K. Image segmentation using laser radar data. Pattern Recognit,1990 23,(6) : 569~581.
  - 7 Asada M ,Ichikawa H , Tsuji A. Determining surface orientation by projecting a stripe pattern, IEEE Trans PAMI,1988: 749~754.
  - 8 Wang Y F, Cheng D I. Three-dimensional shape construction and recognition by fusing intensity and structured lighting, Pattern Recognition 25, 1992,1411~1425.
  - 9 Bolles R C ,Horaud P,Hannah M H . 3DPO: A three-dimensional part orientation system, Proceedings IJCAI-83 karlsruhe. F R G 1983,1116~1120.
  - 10 Blake A, Mccowen D, Lo H R. Trinocular Active Range-Sensing. IEEE Trans. PAMI, 1993, 15(5):477~483.
  - 11 Sashua A. Projective structure from Uncalibrated images: structure from motion and recognition. IEEE Trans. PAMI, 1994, 16(8): 778~790.



陈向荣,1994 年获得清华大学计算机系学士学位,同年在清华大学计算机系攻读博士学位。目前的研究方向是计算机视觉和图象处理。

## Using Range Information Propagation Transform in Object Recognition

Chen Xiangrong, Lin Xueyin, Zhu Zhigang, Shi Dingji

(Department of Computer Science & Technology Tsinghua University, Beijing 100084)

**Abstracts** Using sparse range image and gray level image in combination, features extracted from the gray level image can be assigned conditionally with corresponding range values by means of a projective mapping. It is called range information propagation transformation. The range information propagated in this way can be used to calculate the Euclidean invariant of corresponding surface planar patterns and hence to generate initial hypothesis. Since initial hypotheses are extracted based on reliable information, and hybrid control strategy with feedback is used, the proposed method can be used in practice situations. In this paper the principle of the method and its implemental issues are addressed, and some preliminary experimental results are presented to demonstrate its effectiveness and efficiency.

**Keywords** Model-based vision, Object recognition, Data fusion, Projective transformation, Range information propagation