

# 基于模型的三维物体识别<sup>1)</sup>

林应强 吴立德

(复旦大学计算机科学系 上海 200433)

**摘 要** 实现了一个完整的基于模型的三维物体识别系统, 它可识别灰度图象中包含的物体, 如对遮挡加以限制, 还可识别被遮挡的物体. 该系统能实现物体的自动建模, 也可先定性识别某一物体的立体图对以获取高层知识, 然后在高层知识的指导下准确地匹配立体图对中相对应的特征. 此外, 还提出了利用最能表示物体特征的表面(特征面)来识别物体的方法, 以提高系统抗噪声的能力. 大量实验证明, 该系统具有相当的稳健性.

**关键词** 三维物体识别, 关系属性图, 特征面, 特征面匹配, 立体对应.

## 1 引言

三维物体识别是计算机视觉的重要组成部分, 它的任务是识别出图象中有什么类型的物体, 并给出物体在图象所反映的场景中的位置和方向. 如果在识别的过程中要利用有关物体的先验知识(model), 则称为基于模型的三维物体识别. 它的基本思想是选取物体的某些特征作为基元, 在确定基元之间的相互关系后, 将物体表示成一个关系属性图. 物体的模型就是识别系统在识别物体前所获得的物体的表示——关系属性图. 系统通过从图象中抽取出的物体的关系属性图与物体的模型进行匹配来识别图象中包含的物体.

目前, 已有很多关于三维物体识别的研究, 一般介绍见文献[1, 4], 较为具体的介绍见文献[2, 3]. 本文实现的三维物体识别系统以灰度图象作为输入、以 2 维表面为基元建立关系属性图. 由于噪声的存在, 难以从真实图象中得到物体的完整的关系属性图. 为此, 本文提出了特征匹配的方法, 即通过识别最能表示物体特征的面(特征面)来识别物体. 这样, 即使得不到完整的关系属性图, 但只要能得到物体的一个或多个特征面, 系统仍有可能正确地识别物体. 另外, 若提供一对立体图对, 系统还可利用通过定性识别获取的高层知识来准确地匹配立体图对中相对应的特征. 限于文章篇幅, 这部分内容从略.

## 2 模型的建立及场景中允许出现的物体

### 2.1 模型的建立

本文实现的识别系统的建模方法与文献[2]相同. 只是本系统在建模时对每一模型得到 26 张侧面图.

### 2.2 场景中允许出现的物体

1) 国家自然科学基金资助项目.

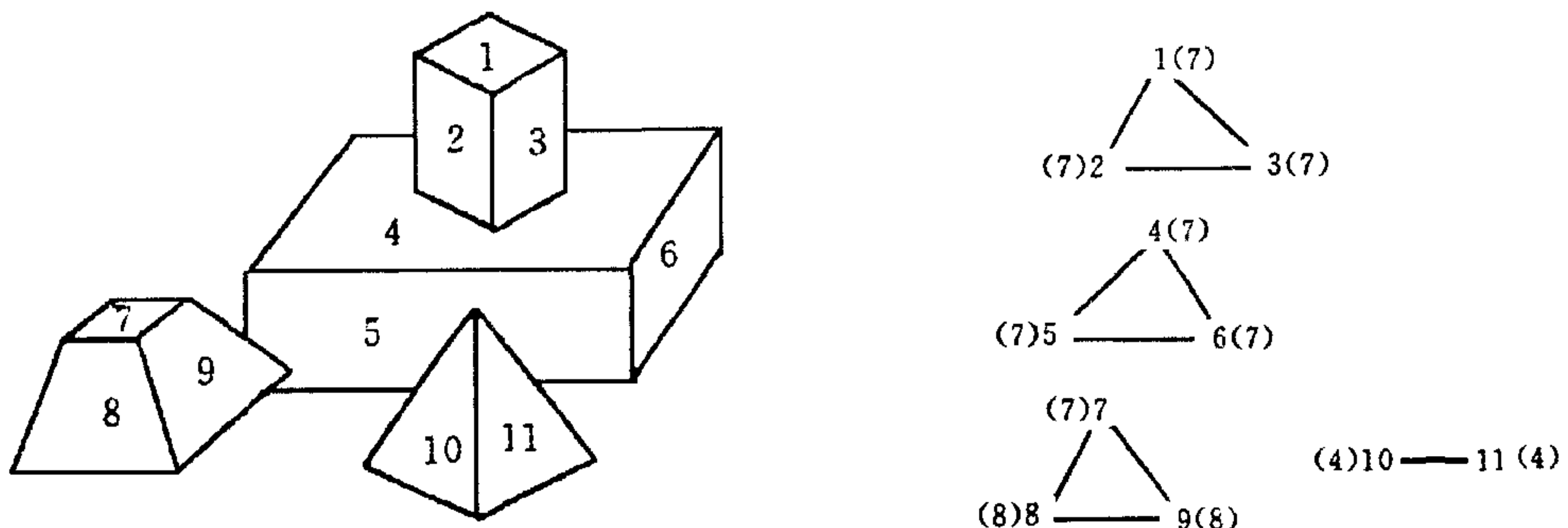
本系统对场景中允许出现的物体有所限制,模型库中只存放了允许出现的物体的关系属性图,系统只能识别存放在模型库中的物体.但模型库中存放的物体可视具体应用的要求而改变,因而本文介绍的方法是通用的.

尽管对每一物体可得 26 张侧面图,但由于物体的对称性,从某些侧面图得到的关系属性图是同构的,建模时系统模型库中只保留物体的不同构的关系属性图(见文献[2, 3]).另外,有些侧面图不具有代表性,没有很好地反映物体的特征.因此,识别物体时,若系统发现从某幅图象无法确定物体类型,则应要求用户再提供该物体的另一幅图象,或从模型库中选取一个可能性最大的物体.

从所有物体的侧面图中,可以总结出场景中允许出现的表面的类型,类型用数字来表示,例如 4,7,8 分别表示三角形、平行四边形、梯形;对边数大于等于 5 的表面,其类型编号规定为曲边数目 $\times 100 + (\text{边的数目} - 5) + 30$ .

### 3 从图象中提取关系属性图

本系统输入的图象为灰度图象,尺寸  $256 \times 256$  象素,灰度级为 256 级.图象输入识别系统后,先进行边缘求取和断边连接等预处理(若输入人工生成的图象,不必进行预处理).经过预处理后,图象中物体的每一个面变成由象素组成的圈——围线.因此,要利用围线追踪算法从图象中提取物体的每一个面,作为建立关系属性图的基元.但传统的基于象素的区域围线追踪算法在追踪形状较为复杂的区域的围线时常出现错误,如文献[5]中的围线追踪算法(见文献[6]).文献[7]曾提出基于边过程的区域围线追踪算法,彻底解决了二值图象中的区域围线追踪问题.追踪出每个面的围线后,就可以建立面的连接图以找出每个面的所有相邻面.当两个面的围线有一部分公共象素时这两个面互为相邻面.面的连接图表示了图象中面的连接关系,但它不是关系属性图,因为图中每一节点(表面)的类型尚未确定,在多物体情形,被遮挡物体还未与遮挡它的物体相分离.为确定面的类型,必须对围线进行分割、拟合等处理,得到围线的结构.得到每一个面的围线结构后,可利用“T”节点来确定被遮挡面并分离被遮挡面与遮挡它的面,同时尽可能恢复被遮挡面的形状.分离物体后,便可确定每个面的类型,得到物体的关系属性图.对于形状不能完全恢复的被遮挡面,要利用模型库中有关物体的知识确定其类型.图 1 所示为一场景及其关系属性图.



(a)场景(图中的数字是表面的编号)

(b)关系属性图(括号中的数字是表面的类型)

图 1 场景及其关系属性图

## 4 三维物体识别

### 4.1 通过匹配关系属性图实现物体识别

本系统识别物体的方法是将从图象中抽取出的物体的关系属性图与存放在系统模型库中的关系属性图进行匹配. 若匹配成功, 则系统识别了图象中的物体; 反之, 系统将表示不认识该物体. 关系属性图的匹配实际上是图论中的图同构问题. 有关图同构的算法是十分常用的, 这里从略.

### 4.2 利用特征面匹配实现三维物体识别

在真实图象中, 由于噪声的存在, 系统难以得到物体的完整的关系属性图, 为此, 我们提出了利用特征面识别物体的方法. 特征面有两类, 若某一表面  $S$  在同一物体的一张(或多张)关系属性图中出现, 且在至少某一张关系属性图中没有与  $S$  同类型的其它表面, 则  $S$  为该物体的第一类特征面(KF1); 若从该物体的所有侧面图中得到的关系属性图只包含一种类型的表面(该类面可能在某一侧面图中重复出现, 但该侧面图中不含别的类型的面), 则该面为该物体的第二类特征面(KF2).

下面介绍基于特征面的物体识别方法, 要说明的是, 用特征匹配识别物体时, 有可能出现识别错误. 先介绍索引库的建立. 库中每一节点是一特征面  $S$ , 所有包含该特征面的物体(即  $S$  是该物体的特征面)组成该节点指向的链表, 链表中每一单元包含了物体的类型以及  $S$  在该物体中的权重. 权重计算如下: 对第一类特征面, 用特征面在该物体的所有侧面图中出现的次数除以该特征面在模型库的所有物体的侧面图中出现的次数, 它是位于 0 和 1 之间的一个数. 由于边数多的面集中出现在某一物体中, 它在该物体中的权重接近 1, 而边数少的面在多个物体中出现, 在每一物体中的权重较小, 这就给边数多的面以较高的优先权; 对第二类特征面, 令其权重为 1. 最后将每一特征面指向的链表按权重从大到小的顺序排序, 识别时首先找到权重大的物体.

建立索引库后, 便可进行基于特征面的物体识别. 一般而言, 对识别所赋予的识别值越大, 则识别的相对可靠性越高.

#### 识别算法.

输入: 从图象中提取的物体的不完整的关系属性图. 输出: 可能的物体种类.

符号:  $result$  为识别值, 初值为 0;  $PARTI$  为图象中的物体的关系属性图;  $PARTM$  为模型库中的物体;  $PARTM-GR$  是  $PARTM$  的关系属性图.

1) 从不完整的关系属性图中找出特征面, 并决定是 KF1 还是 KF2. 其方法与建索引库时找特征面的方法一样. 若是 KF2, 则该物体只有一个特征面.

2) 从索引库中找出每一特征面指向的物体, 并将每一特征面指向的物体组成一个集合, 有几个特征面就有几个集合.

3) 若特征面不止一个, 则作所有集合的交集以找出包含所有特征面的物体.

4) if 特征面不止一个且交集不空, then 对其中每一物体( $PARTM$ )do 5)–14)步.

5) 令  $result1 = 0$ , 找出  $PARTM$  的所有关系属性图, 对每一张关系属性图  $PARTM-GR$  do 6)–12)步.

6) 若  $PARTM-GR$  包含的面不少于  $PARTI$  包含的面,  $result1 = 11$ ; 否则, 取下一

PARTM-GR 并转第 6)步;若 PARTM-GR 已取完,转第 13)步.

7)–11)若 PARTI 中部分特征面与其在 PARTM-GR 中的对应面有相同个数的相邻面,则  $result1=12$ ;若所有特征面均满足上述条件,则  $result1=13$ ;若 PARTI 中部分特征面与其在 PARTM-GR 中的对应面有相同个数的相邻面且在其相邻面之间能建立一一对应,则  $result1=14$ ;若所有特征面均满足上述条件,则  $result1=15$ ;若能从 PARTI 与 PARTM-GR 中提取特征子图且特征子图能匹配,则  $result1=16$ .

12)若  $result1>result$ ,则  $result=result1$ ;令  $result1=0$ ,并取下一 PARTM-GR 转第 6)步;若 PARTM-GR 已取完,转第 13)步.

13)若  $result>0$ ,则将 PARTM 作为 PARTI 的匹配物体,并保存  $result$  的值.

14)令  $result=0$ ,取下一 PARTM 并转第 5)步,若交集集中的 PARTM 已取完,则转第 15)步.

15)若已找到匹配,转第 42)步;否则  $result=0$ ,转第 16)步. endif

16)if 特征面不止一个 且  $result=0$  then

17)找出边数最多且指向链表不空的特征面,若其边数 $<5$ ,则作所有集合的并集;若边数 $\geq 5$ ,则选取该面指向的链表作为并集;若所有特征面指向的链表均为空,则返回不识别.

18)对并集中的每一物体(PARTM) do 19)–27)步.

19)令  $result1=0$ ,找出 PARTM 的所有关系属性图,对每一张关系属性图 PARTM-GR do 20)–24)步.

20)若 PARTM-GR 包含的面不少于 PARTI 包含的面, $result1=1$ ;否则,取下一 PARTM-GR 并转第 20)步;若 PARTM-GR 已取完,转第 25)步.

21)–23)若 PARTI 中部分特征面与其在 PARTM-GR 中的对应面有相同个数的相邻面,则  $result1=2$ ;若 PARTI 中部分特征面与其在 PARTM-GR 中的对应面有相同个数的相邻面且在其相邻面之间能建立一一对应,则  $result1=3$ .

24)若  $result1>result$ ,则  $result=result1$ ;令  $result1=0$ ,并取下一 PARTM-GR,转第 20)步;若 PARTM-GR 已取完,转第 25)步.

25)若  $result>0$  且边数最多并指向链表不空的特征面的边数 $\geq 5$ ,则令  $result=result+3$ .

26)若  $result>0$ ,则将 PARTM 作为 PARTI 的匹配物体,并保存  $result$  的值.

27)令  $result=0$ ,取下一 PARTM 并转第 19)步;若并集中的 PARTM 已取完,则转第 28)步.

28)若已找到匹配,转第 42)步. endif

29)if 特征面只有一个 then

30)选取该面指向的链表作为并集,若并集为空,则返回不识别.

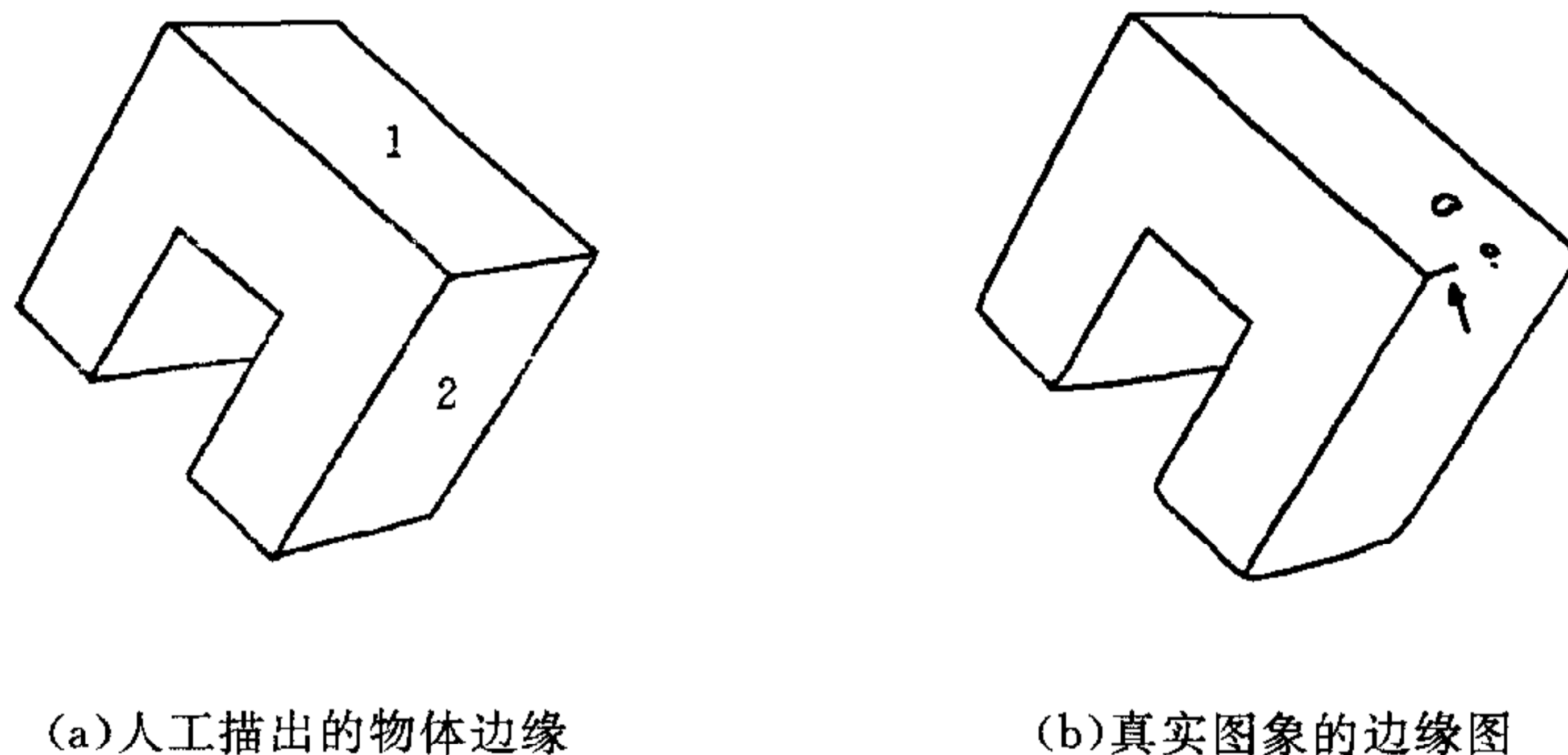
31)对并集中的每一物体(PARTM) do 32)–41)步;以下 32)–41)步与 18)–28)步的方法几乎一样,由于只有一个特征面,所以该面可能是物体的 KF2 面,若是 KF2 面,则将  $result$  的值加上 13. endif

42)若已找到匹配,则返回物体类型及识别值.

## 5 实验

用大量人工生成的图象和一些真实图象对识别系统进行了测试,系统对人工生成的图象具有相当的稳健性,对真实图象也具有一定的稳健性.若能在边缘求取等视觉底层处理方面取得较好结果,将会提高系统对真实图象的稳健性.

目前的系统能正确分割类似图 1 的场景并识别其中的每一物体.图 2(b)是 Sobel 边缘算子作用到一真实图象并取阈值后获得的边缘图,图 2(a)是人工描出的该真实图象中物体的边缘,由于面 1 与面 2 之间灰度接近,加上噪声的干扰,它们之间的边缘未能求出.这样,图象中的物体有 3 个特征面,类型分别为 33(8 条边),31(6 条边),8(梯形).模型库中不存在同时包含这三个特征面的物体,系统选取最主要特征——类型为 33 的面,该特征在索引库中仅指向图 2(a)中的物体,类型为 33 的面是该物体的 KF1 面且权重为 1,即该类表面只出现在这一物体中.因此,尽管图象中物体的关系属性图不完整,系统还是正确识别了该物体.



(a)人工描出的物体边缘

(b)真实图象的边缘图

图 2

注.图中的小圈(噪声形成的假边)和箭头所指的断边在围线追踪时被去掉了.

## 参 考 文 献

- [1] Arman F, Aggarwal J K, Model-based object recognition in dense range images——A review. *ACM Comput. Surveys*, 1993, 3.
- [2] Dickinson S J, Pentland A P, Rosenfeld A. From volumes to views: An approach to 3-D object recognition. *CVGIP:IU* 1992, 55(2):130—154.
- [3] Dickinson S J, Pentland A P, Rosenfeld A. 3-D shape recovery using distributed aspect matching. *IEEE Trans. PAMI, special issue on Interpretation of 3-D Scenes, Part I*, 1992, 14(2):174—198.
- [4] 吴立德. 计算机视觉. 上海: 复旦大学出版社, 1993.
- [5] Pavlidis T. *Algorithms for Graphics and Image Processing*, Rockville: Computer Science Press, 1982.
- [6] 林应强, 吴立德. 区域围线追踪算法的改进. *模式识别与人工智能*, 1994, 7(3):215—226.
- [7] 吴立德, 林应强. 基于边过程的区域围线追踪和围线的树结构. *计算机学报*, 1996, 19(6):457—465.

## MODEL BASED 3-D OBJECT RECOGNITION

LIN YINGQIANG WU LIDE

(*Computer Science Department, Fudan Univ., Shanghai 200433*)

**Abstract** A model based 3-D object recognition system is realized. The system can extract the relational face-graph of the object in the original image, then by matching the relational face-graph with the one in the model base, the system can recognize what kind of object is in the image. If the stereo images of one scene are given, the system can accurately establish the correspondence between the two images using higher level knowledge about the scene. In addition, the characteristic-matching method has been proposed. The main idea of this method is to recognize the object by using the faces which can best represent the characteristics of the object. The robustness of the system has been proven by experiments.

**Key words** 3-D object recognition, relational face-graph, characteristic face, characteristic face matching, 3-D correspondence.

**林应强** 复旦大学计算机科学系助教、硕士. 主要从事模式识别、图像处理和计算机视觉方面的研究.

**吴立德** 复旦大学计算机科学系教授、博士生导师. 主要从事模式识别与计算机视觉、计算机语言学以及计算机应用等方面的研究. 曾获国家自然科学基金等多项奖励, 发表论文百余篇.