

基于距离图象的定性识别¹⁾

庄青 吴立德 袁梅

(上海复旦大学计算机科学系 200433)

摘 要

本文提出的定性描述图表示方法能在无模型指导的情况下有效地实现遮挡物体的分离,并将识别过程转化为图或子图同构的问题,还提出了用“提纯”与“搜索”相结合的递归过程来判定图或子图同构。实验证明了该方法的有效性和稳定性。

关键词: 定性描述图,定性识别,图或子图同构。

一、引 言

三维物体识别研究如何利用灰度图象或距离图象完成对场景中物体的识别,其中确定三维物体的表示是问题的关键,许多研究者对此进行了大量的探索,提出了各种各样的表示方法,大致可分为基于体积元^[1-5]和基于表面片^[6]的两类,尽管这些表示方法对物体的定量程度各不相同,绝大多数的研究者都试图用尽可能精确的形式表示物体,这样会带来识别过程中巨大的运算量。

本文提出了定性描述图 (Qualitative Description Graph) 的物体表示方法,这种方法简单、直观,具有良好的稳定性,而且能在无模型指导的情况下有效地实现遮挡物体的分离,从而简化了复杂场景的处理。在获得了物体的定性描述图以后,识别过程转化为图或子图同构的问题。

二、定性描述图及其性质

2.1. 定性描述图的引入

点、边、面是刻画三维物体表面的重要特征,这些特征及其关系是把握物体主要结构所不可缺少的。本文提出的定性描述图的概念,利用点、边、面的定性特征及其相互关系,因而对物体的描述比较简洁。

2.2. 定性描述图的形式表示

记定性描述图为 $G_Q = (V_Q, R_Q)$, 其中 V_Q 是特征元素的集合, R_Q 是特征元素之间关系的集合。 V_Q 可分为三个子集 V_J, V_E 和 V_S , 分别表示点、边、面的集合, R_Q 也可分为三个子集 R_{JE}, R_{ES} 和 R_{SS} , 分别是点-边、边-面、面-面关系的集合。下面讨论各个子集中元素的性质。

本文于 1991 年 4 月 24 日收到。

1) 高等学校博士学科点专项科研基金资助项目。

1) 如果 $v \in V_E$, 则 v 具有以下性质:

i) 起始点和终结点的三维坐标;

ii) 属性

(a) 二维几何特性: 直线或曲线;

(b) 三维几何特性: 阶跃边、凸折边或凹折边。其中阶跃边是由于深度不连续而形成的, 折边是由于法向不连续而形成的;

2) 如果 $v \in V_J$, 则 v 具有以下性质:

i) 度数: 与之相关的边的个数;

ii) 类型: 每个 v 属于且只属于以下一类:

(a) O型交点: 某条边的孤立端点;

(b) L型交点: 由两条边形成的交点且此两边中至少有一条阶跃边;

(c) Y型交点: 由三条或更多的折边形成的交点;

(d) A型交点: 至少有三条边且其中至少有一条阶跃边形成的交点;

(e) T型交点: 其中“—”表示被分成两段的阶跃边, “|”表示一条阶跃边或折边;

3) 如果 $v \in V_S$, 则 v 具有以下性质:

i) 度数: 与之相关的边的个数;

ii) 面的法向: 该面上各点法向的平均;

iii) 可见面的 3D 面积;

iv) 质心: 该面上各点三维坐标的平均;

v) 面的类型: v 属于以下一种:

(a) 平面: 该面各点的高斯曲率 K 和平均曲率 H 满足 $K = 0, H = 0$;

(b) 凸脊面: 该面各点 $K = 0, H < 0$;

(c) 凹脊面: 该面各点 $K = 0, H > 0$;

(d) 凸峰面: 该面各点 $K > 0, H < 0$;

(e) 凹峰面: 该面各点 $K > 0, H > 0$;

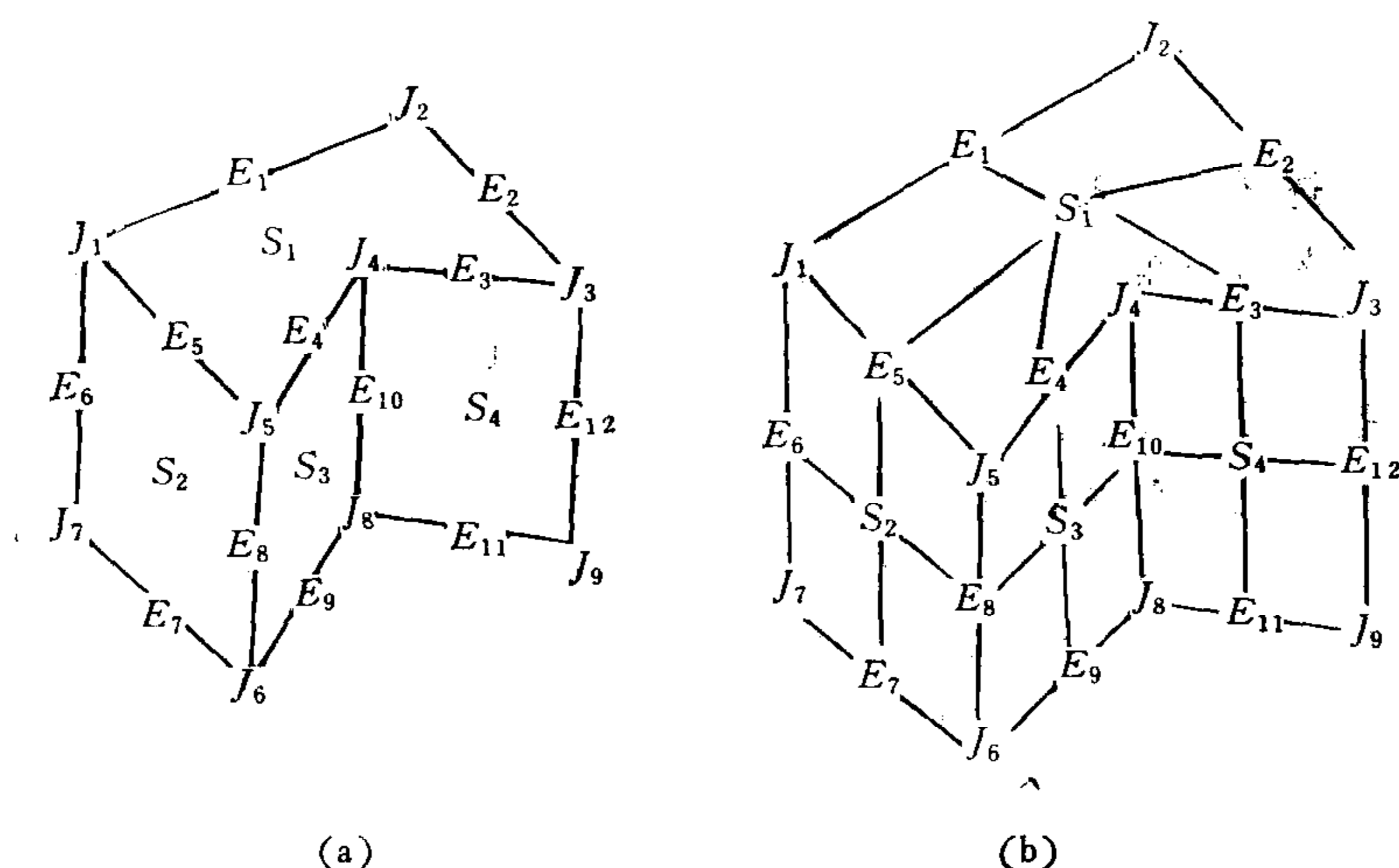


图 1

(f) 复杂曲面：该面各点具有其它的 KH 符号或该面至少有两个区域具有不同的 KH 符号；

4) 如果 $r = (v, v') \in R_{JE}$, 它表示 $v \in V_J$, $v' \in V_E$, 而且 v 是 v' 的一个端点；

5) 如果 $r = (v, v') \in R_{ES}$, 它表示 $v \in V_J$, $v' \in V_E$, 而且具有下列某个定性描述：

i) v 是 v' 的边界；

ii) v 是 v' 的一条遮挡边；

6) 如果 $r = (v, v') \in R_{SS}$, 它表示 $v, v' \in V_S$, 且 v, v' 以折边相连, 它有属性 θ , 是 v 和 v' 法向的夹角。

下面给出一个物体定性描述图的例子。图 1 中(a)是一个楔型物体, (b)是它的定性描述图, 其中包括点、边、面及其关系的属性。

三、定性识别

在获得了场景的定性描述图后, 先要对场景中的各个物体进行分离, 然后再对分离后的各个物体进行定性识别。

3.1. 物体的分离

基于定性描述图的物体分离主要是利用阶跃边和由遮挡引起的 T 型结点的特性。其基本思想是：先去掉定性描述图中的阶跃边, 然后根据图的连通性, 形成若干子图, 其中每个子图代表一个物体, 再把相应的阶跃边及其与面、点的关系加入子图, 并重新处理 T 型结点经分离后所形成的 L 型结点与 O 型结点, 最后就形成了各个物体的定性描述图。具体算法如下：

step 1. 根据 G_Q 构建一个新的子图, 记为 $G^* = (V^*, R^*)$, 其中

$V^* := V_S \cup V_E - \{v \in V_E \mid v \text{ 是阶跃边}\}$;

$R^* := R_{ES} - \{(v, v') \in R_{ES} \mid v \notin V^*\}$ 。

step 2. 检测 G^* 的连通性, 并产生 $k (\geq 1)$ 个连通块, 然后将 V_S 划分成

$V_S^i := \{v \in V_S \mid v \text{ 在第 } i \text{ 个连通块中}\}, i = 1, \dots, k$ 。

step 3. 对每一个 V_S^i 构造相应的 $G_Q^i = (V_Q^i, R_Q^i)$ ：

$V_Q^i := V_S^i \cup V_J^i \cup V_E^i$, $R_Q^i := R_{JE}^i \cup R_{ES}^i$, 其中

$V_E^i := \{v \in V_E \mid v \text{ 是 } V_S^i \text{ 中面的边界}\}$, $V_J^i := \{v \in V_J \mid v \text{ 是 } V_E^i \text{ 中边的端点}\}$,

$R_{ES}^i := \{(v, v') \in R_{ES} \mid v \in V_E^i \text{ 且 } v' \in V_S^i\}$, $R_{JE}^i := \{(v, v') \in R_{JE} \mid v \in V_J^i \text{ 且 } v' \in V_E^i\}$ 。

step 4. 重复下列四个运算式直到 V_J^i 中不存在 L 型结点 v , 它在 V_J 中的对应点是 T 型结点：

$$V_E^i := V_E^i \cup \{\bar{v}\} - \{v_1, v_2\}, \quad V_J^i := V_J^i - \{v\},$$

$$R_{ES}^i := R_{ES}^i \cup \{(\bar{v}, v_5)\} - \{(v_1, v_5), (v_2, v_5)\},$$

$$R_{JE}^i := R_{JE}^i \cup \{(v_3, \bar{v}), (v_4, \bar{v})\} - \{(v, v_1), (v_1, v_2), (v_3, v_1), (v_4, v_2)\},$$

其中：i) $v_1, v_2 \in V_E^i$, $v_3, v_4 \in V_J^i$, $v_5 \in V_S^i$, 并且有 $(v_1, v_5), (v_2, v_5) \in R_{ES}^i$ 和 $(v, v_1), (v_1, v_2), (v_3, v_1), (v_4, v_2) \in R_{JE}^i$ ；

ii) \bar{v} 是一条新的阶跃边(将阶跃边 v_1, v_2 合并而成), 它的二维几何特性由 v_1, v_2 的二

维几何特性决定;

step 5. 重复下列两个运算式直到 V_j^i 中不存在 O 型结点 v , 它在 V_j 中的对应点是 T 型结点:

$$V_j^i := V_j^i - \{v\}, R_{jE}^i := R_{jE}^i - \{(v, v')\}.$$

3.2. 定性识别

在完成了对场景中各个物体的分离以后, 由于模型也是用定性描述图表示的, 所以物体的定性识别就对应于判断图或子图的同构.

1) 候选模型 给定场景中某个物体的定性描述图, 记为 G_0^s , 将候选模型的定性描述图记为 G_0^m . 首先选用一些简单的条件来排除大量不可能的模型.

(a) 当 G_0^s 为某个未被遮挡的物体时, 利用图同构的如下必要条件选取 G_0^m :

i) $|V_s^m| = |V_s^s|$, ii) $|V_E^m| = |V_E^s|$, iii) $|V_j^m| = |V_j^s|$;

(b) 当 G_0^s 为某个被部分遮挡的物体时, 利用如下子图同构的必要条件来选取 G_0^m :

i) $|V_s^m| > |V_s^s - \bar{V}_s^s|$, ii) $|V_E^m| > |V_E^s - \bar{V}_E^s|$, iii) $|V_j^m| \geq |V_j^s|$,

其中, \bar{V}_s^s, \bar{V}_E^s 分别是被遮挡的面和边的集合.

2) 结点对应矩阵 结点对应矩阵不仅被用来判定 G_0^s 和 G_0^m 间的同构, 而且被用来表示搜索路径中的各个状态, 记 $M = [m_{ij}]$ 是一个 $|V^s|$ 行 \times $|V^m|$ 列的矩阵, 它的每一个元素取 0 或 1, 如 $m_{ij} = 1$, 那么表示 G_0^s 中的第 i 个结点 ($v_i \in V^s$) 对应于 G_0^m 中的第 j 个结点 ($v_j \in V^m$). 为了方便起见, 假定 V^s 中前 $|V^s - \bar{V}_s^s \cup \bar{V}_E^s|$ 个结点表示未被遮挡的结点.

(a) 初始矩阵 $m^0 = [m_{ij}^0]$: 即初始状态;

$$m_{ij}^0 = \begin{cases} 1, \text{ 如果 i) } v_i \in V_j^s, v_j \in V_j^m, \text{ 且这两个交点的类型和度数相同} \\ \text{或 ii) } v_i \in V_E^s, v_j \in V_E^m, \text{ 这两条边的二维和三维几何特性相} \\ \text{同, 且它们关联的面结点个数为 } (1 \text{ 或 } 2) \\ \text{或 iii) } v_i \in V_s^s, v_j \in V_s^m, \text{ 这两个面的类型相同, 且当 } v_i \in \bar{V}_s^s \text{ 时,} \\ |A(v_i) - \bar{V}_E^s| < |A(v_j)|, \text{ 而当 } v_i \in V_s^s \text{ 时, } |A(v_i)| = |A(v_j)|, \\ \text{这里 } A(v) \text{ 表示与面 } v \text{ 相关的边的集合;} \\ 0, \text{ 其它.} \end{cases}$$

(b) 匹配矩阵 $M' = [m'_{ij}]$: 即终结状态;

判定图同构或子图同构也就是得到一个匹配矩阵 M' , 使得, 如果 $m'_{ij} = 1$ ($i \leq |V^s - \bar{V}_s^s \cup \bar{V}_E^s|$), 那么 G_0^s 中的第 i 个结点在这个图或子图同构下与 G_0^m 的第 j 个结点对应, 可见 M' 必须满足

$$\text{i) } \sum_{j=1}^{|V^m|} m'_{ij} = 1, (i \leq |V^s - \bar{V}_s^s \cup \bar{V}_E^s|);$$

$$\text{ii) } \sum_{i=1}^{|V^s - \bar{V}_s^s \cup \bar{V}_E^s|} m'_{ij} \leq 1, (\text{当 } |\bar{V}_s^s \cup \bar{V}_E^s| \neq 0 \text{ 时, 才允许出现小于号});$$

$$\text{iii) } ((v_{i1}, v_{i2}) \in R^s \iff (v_{j1}, v_{j2}) \in R^m \left| \begin{array}{l} m'_{i1j1} = 1 \\ m'_{i2j2} = 1 \end{array} \right., (\bar{V}_s^s \cup \bar{V}_E^s = 0));$$

$$((v_{i1}, v_{i2}) \in R' \iff (v_{j1}, v_{j2}) \in R^m \mid \begin{matrix} m'_{i1j1}=1 \\ m'_{i2j2}=1 \end{matrix}, (\bar{V}'_S \cup \bar{V}'_E \neq 0));$$

3) “提纯”与“搜索” 根据 M' 应满足的条件和 M' 的定义可知: $(m'_{ij} = 1) \Rightarrow (m_i^i = 1)$, 因此, 求 M' 的过程可看作从初始状态 M^0 出发, 反复搜索下一个状态, 直至达到状态 M' 的过程. 在此采用“提纯”与“搜索”相结合的递归过程, 它充分利用了定性描述图的特有属性, 使得搜索过程变得简洁而有效.

(a) “提纯”

虽然 M^0 已限制了许多不可能的对应, 但通常搜索空间仍然很大, 因此在搜索过程中引入一个“提纯”过程, 它利用当前状态 M 提供的信息将 M 中某些元素的值从 1 变为 0, 即删除树搜索的一些后继状态.

“提纯”过程给出了保留当前状态 M 中 $m_{ij} = 1$ 的必要条件, 当该条件未被满足时, 将 $m_{ij} = 1$ 变为 $m_{ij} = 0$. 如果 $m_{ij} = 1$, 即 G'_0 中的 v_i 与 G''_0 中的 v_j 对应, 那么 v_i 的邻点与 v_j 的邻点也应存在对应关系. 定义一个二分图 $B_{ij} = (X_i \cup Y_j, \bar{M}_{ij})$ 来表示 G'_0 中 v_i 的邻点与 G''_0 中的邻点 v_j 之间的对应关系.

$$X_i = \begin{cases} \{v \in V'_E \mid (v_i, v) \in R'_{jE}\}, & \text{如果 } v_j \in V'_j; \\ \{v \in V'_j \cup V'_S \mid (v, v_i) \in R'_{jE} \text{ 或 } (v_i, v) \in R'_{ES}\}, & \text{如果 } v_j \in V'_E; \\ \{v \in V'_E - \bar{V}'_E \mid (v, v_i) \in R'_{ES}\} \cup \{v \in V'_S - \bar{V}'_S \mid (v, v_i) \in R'_{SS}\}, & \text{如果 } v_j \in V'_S; \end{cases}$$

$$Y_j = \begin{cases} \{v \in V''_E \mid (v_j, v) \in R''_{jE}\}, & \text{如果 } v_j \in V''_j; \\ \{v \in V''_j \cup V''_S \mid (v, v_j) \in R''_{jE} \text{ 或 } (v_j, v) \in R''_{ES}\}, & \text{如果 } v_j \in V''_E; \\ \{v \in V''_E \mid (v, v_j) \in R''_{ES}\} \cup \{v \in V''_S \mid (v, v_j) \in R''_{SS}\}, & \text{如果 } v_j \in V''_S; \end{cases}$$

\bar{M}_{ij} 是一个对应矩阵中与 X_i 对应的行和 Y_j 对应的列组成的 $|X_i| \times |Y_j|$ 子阵.

显然, 如果当前状态 M 中某个 $m_{ij} = 1$ 被保留, 它必须满足以下条件: 二分图 $B_{ij} = (X_i \cup Y_j, \bar{M}_{ij})$ 存在最大匹配, 使得 X_i 中的每个结点都有 Y_j 中的某个结点匹配. “提纯”过程的具体算法如下:

step 1. $i := 1$, change := false;

step 2. for $j := 1, \dots, |V''|$,

if $m_{ij} = 1$ 但二分图最大匹配条件不满足 then $m_{ij} = 0$, change := true;

step 3. if $\sum_{j=1}^{|V''|} m_{ij} = 0$ then goto FAIL 出口(失败);

step 4. $i := i + 1$;

if $i \leq |V''|$ then goto step 2;

step 5. if change = false then SUCCEED 出口(成功), goto step 1;

“提纯”算法中的第三步, 测试 M 中的某一行是否全为 0, 如果全为 0, 则表明 G'_0 中的某个结点不可能与任何 G''_0 中的结点对应, 即 M' 不可能在 M 的后继状态中, 所以从失败出口退出.

(b) “搜索”

基于“提纯”的搜索过程简洁而有效, 其基本思想是: 先产生初始矩阵, 对每一行进行“提纯”, 如果“提纯”失败, 那么搜索不到同构; 否则, 反复搜索下一个状态直至找到终结状

态 M' 。其具体算法如下:

- step 1. $M := M^0$; $j_c := 0$, $d := 0$; $d' := |V' - \bar{V}'_s \cup \bar{V}'_k|$; “提纯” M ; if FAIL 退出, then 不能搜索到同构, 算法结束; 将 M 压入栈中;
- step 2. $d := d + 1$; if $d > d'$ then 已搜索到同构, 算法结束;
if $\sum_{j=1}^{|V^m|} m_{jj} = 1$ then goto step 2;
- step 3. $j_c := j_c + 1$, if $j_c \leq |V^m|$ 且 $m_{dj_c} = 0$ then goto step 3;
if $j_c > |V^m|$ then goto step 6;
对任何 $i_c < d$, if $m_{i_c i_c} = 1$ then goto step 3;
- step 4. 对任何 $j \neq j_c$, $m_{dj} = 1$;
“提纯” M ;
if FAIL 退出 then 从栈顶复制 M 且 goto step 3;
- step 5. if $d = d'$ then 已搜索到同构, 算法结束;
将 d, j_c 和 M 压入栈中;
 $j_c := 0$, goto step 2;
- step 6. 将 M 从栈中弹出;
if 栈空 then 不能搜索到同构, 算法结束;
将 j_c 和 d 从栈中弹出并从栈顶复制 M , goto step 3;

四、实 验

1. 圆柱体

图 2(a) 是一圆柱体, (b) 是它的定性描述图, 其中 S_1 是平面, S_2 是凸脊面, E_1 和 E_5 是曲的阶跃边, E_2 是曲的折边, E_3 和 E_4 是直的阶跃边, J_1 和 J_2 是 A 型结点, J_3 和 J_4 是 L 型结点。

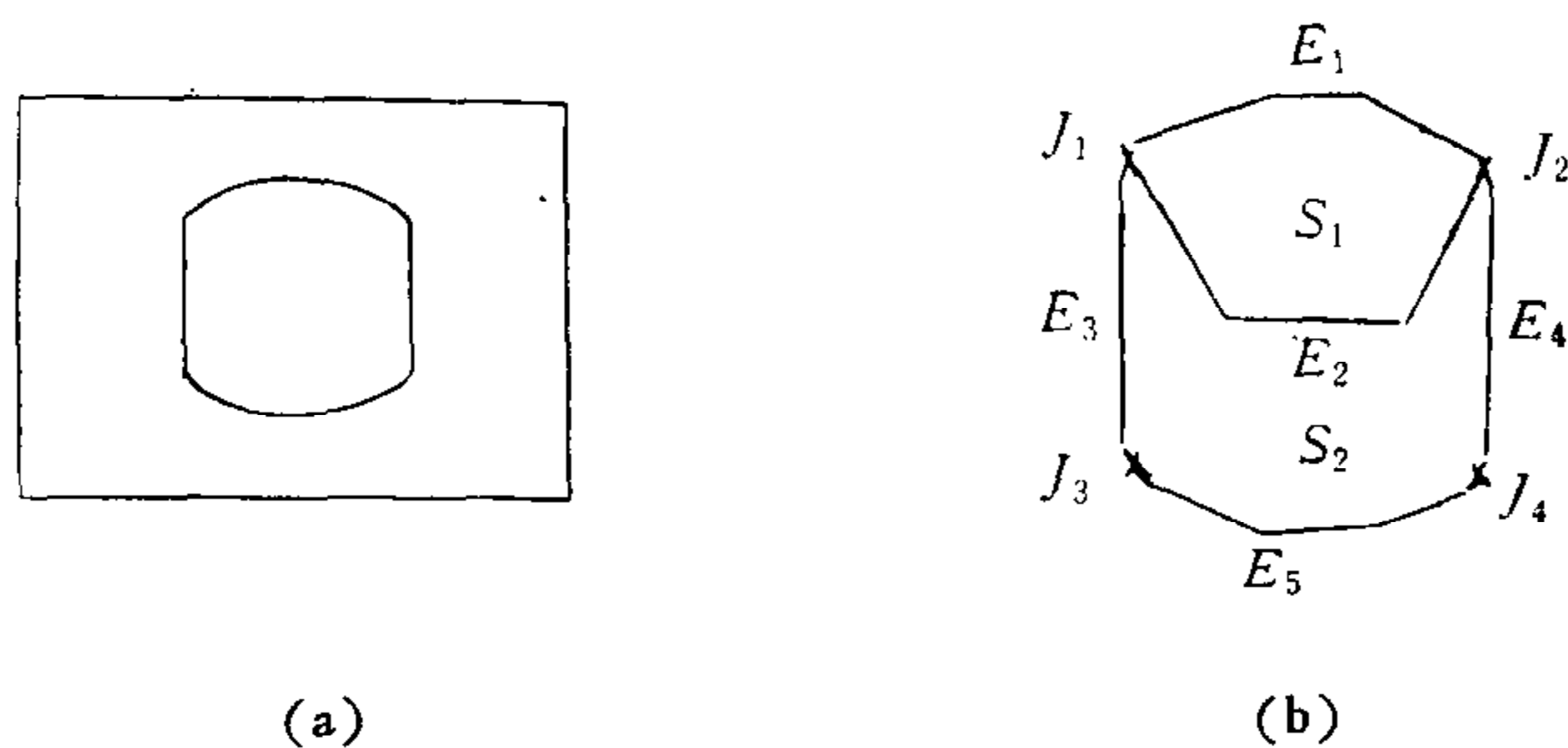


图 2

2. 复杂场景

在已存在控孔立方体(物体 I)、楔型物体(物体 II)、锥型(物体 III)等模型的情况下, 对此场景进行初始识别。图 3(a) 是场景, (b) 是它的定性描述图, 表 1 是识别结果。

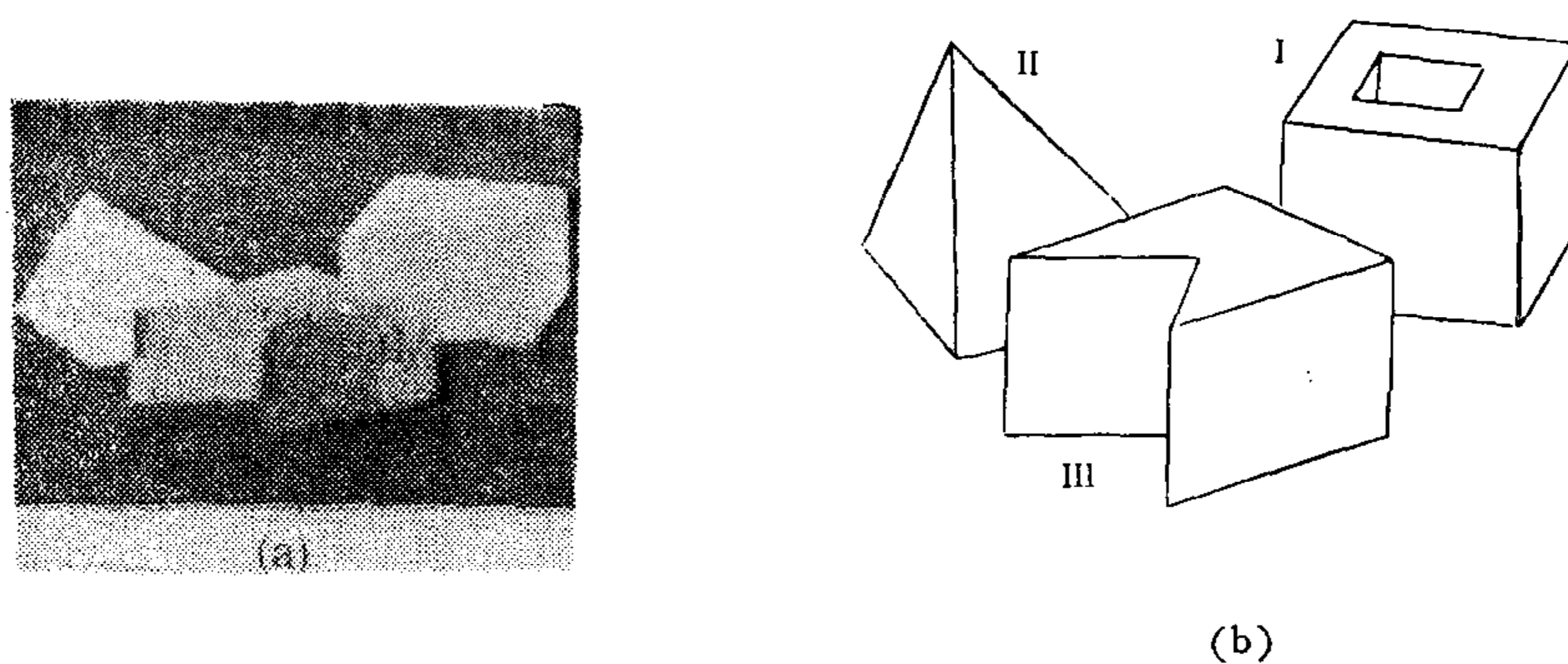


图 3

表 1

No.	识别结果
1	{物体 I}
2	{物体 III}
3	{物体 II}

参 考 文 献

- [1] Meagher, D. J., Geometric Modeling Using Octre Encoding, *Computer Vision, Graphics and Image Processing*, 19(1985), 129—147.
- [2] Brooks, R. A., Symbolic Reasoning Among 3D Models and 2D Images, *Artificial Intelligence*, 17(1981), 285—348.
- [3] Brooks, R. A., Model Based Three Dimensional Interpretation of Two Dimensional Images, *IEEE Trans of Pattern Analysis and Machine Intelligence*, 5(1983), 2, 140—150.
- [4] Horn, B. K. P., Extended Gaussian Images, *In Proceedings of the IEEE* 72(1984), 2, 1656—1678.
- [5] Bhanu, B. and Ho, C., CAD-based 3D Object Representation for Robot Vision, *Computer*, 20(1987), 8, 19—36.
- [6] Fan, T. J., Medioni, G. and Nevatia, R., Recog Evaluation and Merging Algorithms. *Proc. IEEE Trans of Pattern Analysis and Machine Intelligence*, 11(1989), 1140—1157.

QUALITATIVE RECOGNITION OF 3D OBJECTS FROM RANGE IMAGE

ZHUANG QING WU LIDE YUAN MEI

(Computer Science Department, Fudan University, Shanghai 200433)

ABSTRACT

In this paper, a new 3D object qualitative representation called Qualitative Description Graph (QDG) is presented. This symbolic representation can separate effectively the objects with occlusion and simplify greatly the identification of objects in complex scenes. To improve the recognition efficiency, a recursive scheme integrating 'purifying' with 'searching' is proposed. Several experiments show that this approach is effective, robust and efficient.

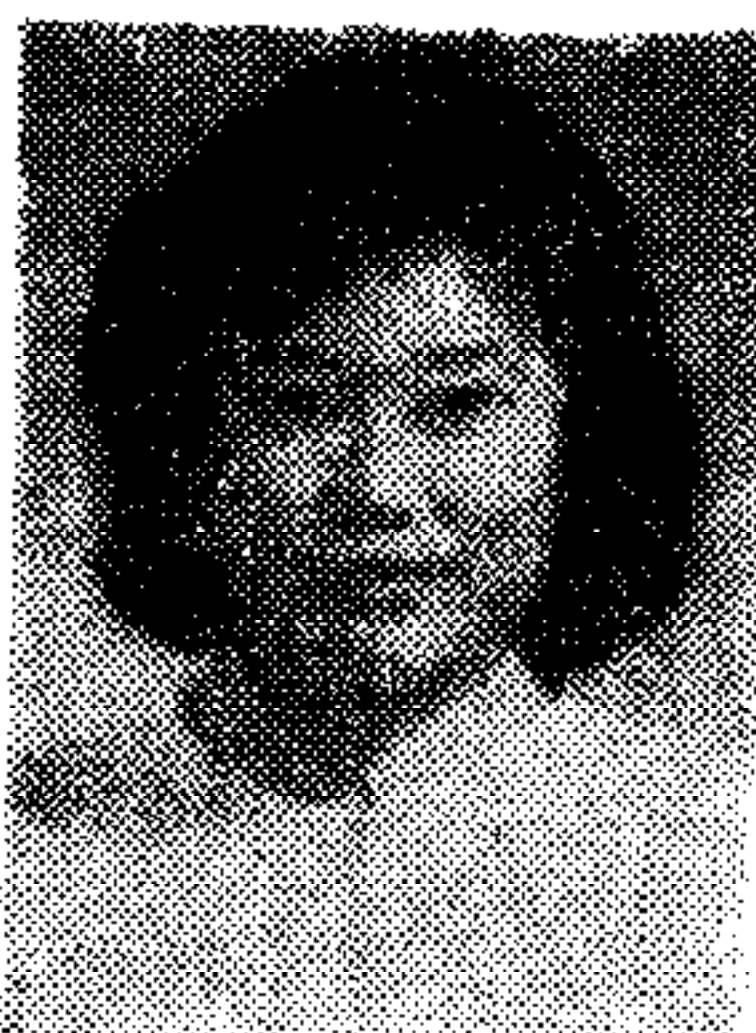
Key words: the representation of 3D objects; Qualitative Description Graph; qualitative identification; graph or subgraph isomorphism.



庄 青 1984年毕业于复旦大学计算机科学系,获学士学位,1984年被录取为该系硕士研究生,因成绩优秀于1987年2月直接攻读博士学位,并于1990年获博士学位。主要研究方向为计算机视觉,在“中国科学”和国内外学术会议上发表过多篇论文。



吴立德 毕业于复旦大学数学系,曾任该校计算机系主任(1982年—1985年),现为该系教授、博士生导师,兼任中国计算机软件专业技术资格和水平考试中心副主任。长期从事计算机软件和应用方面的研究,特别是计算机视觉和自然语言理解方面的研究,著有“概率论”、“计算机系统性能评价”等和论文70余篇。



袁梅 1989年毕业于复旦大学计算机科学系软件专业,获学士学位,现为该系计算机应用专业硕士研究生,主要从事计算机视觉、图象处理、模式识别和软件工程等方面的研究,已在国内外学术会议上发表论文若干篇。