

# 一种层次化的目标表示方法

戴亚非

(哈尔滨工业大学计算中心, 哈尔滨 150001)

**摘要** 提出一种支持自底向上的目标识别策略的线画图象的表示方法。该表示建立在一种基于线→角→面的层次上, 它的低层表示可以使预识别过程能够以简单的统计方法实现, 高层表示支持在模型知识约束下的匹配。

**关键词** 目标识别, 目标表示

## 1 引言

目标识别是建立在目标表示之上的, 事实证明单纯地使用自底向上或自上而下的匹配均不能获得好的识别效果<sup>[1]</sup>。数据驱动的预识别结合模型驱动的匹配兼顾了两者的优点, 不失为一种可行的方法。为此, 本文提出一种层次化表示目标拓扑结构的方法, 目的是使每一层次的表示都能支持一种相应的匹配策略。

## 2 线画图的 EJAG 表示

一个三维多面体目标的图象经分割后得到一个边缘集合, 这些边缘之间存在的拓扑特征表现为线之间的方向性, 角之间的相交性和面之间的邻接性。根据 L. G. Shapiro 等人提出的自底向上的关系金字塔描述方法<sup>[2]</sup>, 本文提出了建立在上述 3 种关系之上的层次表示 EJAG (Edge-Junction-Area Graph)。

EJAG 是一个表示视图等价类<sup>[3]</sup>的层次图, 最底层由独立的边缘组成, 在此之上是由直线相交组成的独立的角, 然后是由交角构成的拓扑结构, 最高层是由边角围出的多边形。EJAG 可以表示成一个三元组  $\langle F, T, D \rangle$ , 其中:

$F = \{f_1, f_2, \dots, f_n\}$  是一个有限集, 表示所有图象特征;

$T = \{t_1, t_2, \dots, t_m\}$  是一个有限集, 表示所有特征类型;

$D = \{D_0, D_1, \dots, D_s\}$  是一个有序序列, 其中每一项表示一种关系结构,  $s$  表示层次结构中的层数。

$D_0$  是由  $n_0$  个关系组成的有序序列  $\langle R_0^1, R_0^2, \dots, R_0^{n_0} \rangle$ , 每个  $R_0^i$  对应着一个特征类型  $t_i, t_i \in T$

$R_0^i = \{ \langle f, A_f \rangle \mid \text{if } f \in F \wedge \text{type}(f) = t_i \}$   
式中:  $\text{type}(f)$  表示  $f$  的类型,  $A_f$  是  $f$  的属性向量。

$D_1$  是由  $n_1$  个关系组成的有序序列  $\langle R_1^1, R_1^2, \dots, R_1^{n_1} \rangle$ , 每个  $R_1^i$  定义了一种关系, 它所涉及的对象是  $D_0$  中的特征。  $R_1^i$  的元素  $r_1^i$  是一个  $q$  元组,

$$r_1^i = \langle (n_1, f_1), (n_2, f_2), \dots, (n_q, f_h) \rangle$$

这里  $n_i$  对应着  $D_0$  中某一关系  $R_0^i$  的名称,  $f_j$  则对应着  $D_0$  中的一个特征, 它的类型与  $R_0^i$  中定义的类型相同。在 EJAG 中, 线→角→面分别为第 0 层, 第 1 层, 第 2 层和第 3 层, 如图 1 所示。

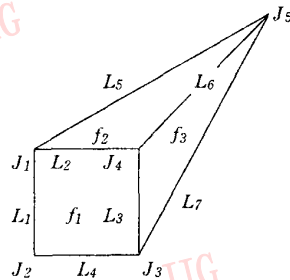


图 1 线画目标的 EJAG 表示

Fig. 1 EJAG representation of line-drawing

Level 3:

$$\{ \langle (4\text{-edges } f_1), (3\text{-edges } f_2), L_2 \rangle \}$$

- {(4-edges  $f_1$ ), (3-edges  $f_3$ ),  $L_3$ }
- {(3-edges  $f_2$ ), (3-edges  $f_3$ ),  $L_6$ }

Level2:

- {(3-lines  $J_1$ ), (3-lines  $J_4$ ),  $L_2$ }
- {(3-lines  $J_1$ ), (2-lines  $J_2$ ),  $L_1$ }
- {(3-lines  $J_3$ ), (2-lines  $J_2$ ),  $L_4$ }
- {(3-lines  $J_1$ ), (3-lines  $J_5$ ),  $L_5$ }
- {(3-lines  $J_4$ ), (3-lines  $J_5$ ),  $L_6$ }
- {(3-lines  $J_4$ ), (3-lines  $J_3$ ),  $L_3$ }

Level 1:

- { $J_5$ : ( $L_5, L_6, L_7$ )}
- { $J_1$ : ( $L_1, L_2, L_5$ )}
- { $J_3$ : ( $L_3, L_4, L_7$ )}
- { $J_4$ : ( $L_2, L_3, L_6$ )}
- { $J_2$ : ( $L_1, L_4$ )}

Level 0:

- { $L_1, L_2, L_3, L_4, L_5, L_6, L_7$ }

### 3 层次表示与匹配

#### 3.1 低层表示对预识别的支持

预识别的目的是在对一个目标作出精确识别之前,通过相对简单的匹配估计目标所对应的模型,因此要求算法简单快捷。利用 EJAG 中最低两层的边缘和交角,根据投影不变特性组成群体特征,这些特征具有与模型无关、相互独立、易于抽取等特点,以支持预识别所要求的统计特性。这些群体特征是<sup>[4]</sup>:平行性、凸凹性、方向性、多面性和复杂性。

从图象中抽取这些特征,构成一个特征向量:

$$V = \{v_1, v_2, v_3, v_4, v_5\}$$

其中每个分量是相互独立的,采用证据累积<sup>[5]</sup>的策略使匹配变得非常简单。首先通过量化公式将  $V$  中各分量以数值形式给出,根据不同类型的模型在上述 5 个特征上的不同取向,引入不同的加权值  $\lambda$ ,将  $V$  中各分量与不同的  $\lambda$  值相结合,可得到不同的数值,我们称其为可信度,使可信度最高的  $\lambda$  所对应的模型,则是要找的模型,这一过程用统计的方法可以达到快速预识别的目的。

#### 3.2 模型约束与目标的结构表示

预识别提供了目标的视图等价类的模型,于是基于模型约束的假设与检验的匹配方法得以实施<sup>[1]</sup>。EJAG 在这一级的表示提供了与模型有关的

结构信息,如各种交角相连接所形成的拓扑结构。其中可反映出以下两个主要特性:连通性和相邻性。

建立在本级表示上的识别目的是标注具有重要意义的特征(Landmark),因为它们能为下一步的匹配提供线索。例如多面体目标中由内部边缘构成的内角,通常是识别的重点。因此根据模型我们可以定义这样的规则:如果一个三线交角的所有相邻交角都是三线交角,则该角为内部交角。

#### 3.3 目标的高层表示与不完整图象的恢复

在 EJAG 的 Level 3 表示中,规定了构成视图等价类的多边形的个数、形状和相邻方式。表示对识别的支持反映在模型约束对丢失特征的预言。在本级表示上的识别目的是建立图象特征与模型之间的全面的对应关系,以便为计算目标的位置与朝向提供量度依据。如果根据模型预言到的特征抽取不到,则可预言丢失特征,并根据模型加以补充。可以说预识别是自底向上的匹配,而在这个过程中匹配是自上而下的,模型始终起引导和约束的作用,从而缩小了搜索空间,提高了匹配效率。

#### 3.4 表示层次与匹配的作用

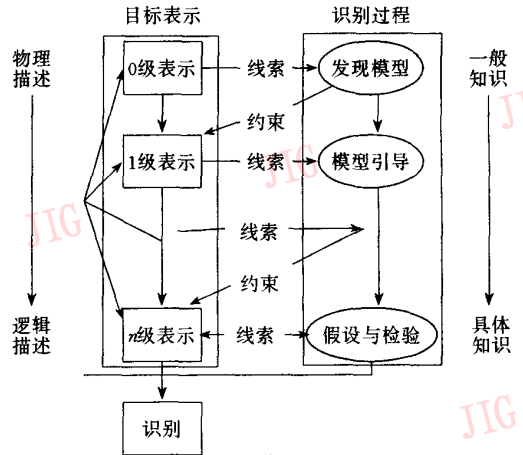


图 2 目标识别中表示与匹配的相互作用关系  
Fig. 2 The effecton between representation and matching in object recognition

从图 2 中可以看到,基于层次表示的匹配也是在不同级别上进行的。模型知识固然重要,但是当对所获取目标一无所知时,模型的约束是无法发挥作用的。因此应使初级表示与模型无关,通过简单特征

的匹配提供线索,发现模型。第2级表示提供的是与模型有关的知识,模型的约束对进一步的匹配起到限制搜索空间的作用,去指导那些未出现特征的抽取,只有识别进行到逻辑表示一级(最高级)时,才能发现图象处理中丢失的特征,才能对目标的具体特征赋予逻辑意义,可见识别也是在不同级别上渐次进行的。

#### 4 结 论

本文通过EJAG表示在目标识别中的应用,阐述了信息表示与信息处理之间相辅相成的作用,同时也表示了这样一种观点,自底向上的识别方法仍

然是一种行之有效的办法。

#### 参 考 文 献

- 1 Lowe D G. Three-Dimensional Object Recognition from Single Two-Dimensional Images. *AI*, 1987, 37 : 355~395.
- 2 Goodman A M, Haralic M, Shapiro L G. Knowledge Based Computer Vision. *Computer*, 1989, 22(12) : 43~54.
- 3 Besl A J, Jian R C. Three Dimensional Object Recognition. *Computing Surveys*, 1985, 17(1) : 95~145.
- 4 戴亚非. 多面体目标的参数化与结构化相结合的表达方法. *计算机工程与设计*, 1996, 17(2):10~15.
- 5 戴亚非. 用证据累积的方法识别不完整图象. *小型微型计算机系统*, 1994,15(4): 37~43.



戴亚非, 哈尔滨工业大学计算中心副教授, 1993年在哈工大计算机系获得博士学位, 主要研究方向: 目标识别, 计算机图形用户界面, CAI。

## A Hierarchical Representation for Object Recognition

Dai Yafei

(Computer Center of Haerbin Institute of Technology, Haerbin 150001)

**Abstract** A representation of line draw is put forward which supports the bottom-up strategy for object recognition. The proposed method is built upon the hierarchy of line-junction-area, whose lower level representation is used in pre-recognition for its statistical feature, and higher level supports object recognition based on the model constraints.

**Keywords** Object recognition, Object representation