

一种图像理解的知识基系统 V 语言

姚庆栋 刘济林 徐胜荣 华中

(浙江大学信息与智能系统研究所, 浙江, 杭州, 310027)

摘要 介绍作者研究的一种用于图像理解的知识基系统 V 语言的特点、性能及其应用情况。

关键词 图像理解, 知识基系统。

引言

把从红外辐射、可见光或雷达传感器得到的图像进行分析, 找出所需目标, 做出场景的解释称为图像理解. 应用于不同的领域又有不同的名称, 如机器视觉在线控制、机器人视觉、自动目标识别(ATR)、视觉导引等等. Weems^[1]把图像理解处理分成3个层次是具有代表性的, 即: 低层处理(是在像素上的处理, 可改善图像质量, 得到轮廓图形及二值化处理等等), 中层处理(以团块分割得到点、线、面的特征描述, 进行匹配, 提出可能的目标的假说), 高层处理(根据特征及邻域关系做出识别和场景解释). 理解系统在不同层次的工作需要不同的软件, 低层处理需要各种算法, 高层处理需要知识处理^[2,3]. 许多报道认为中层处理有必要引入知识处理, 并且, 中、高层处理结果反馈到低层处理的控制, 可以选择或改进低层处理的算法和参数. 这样, 知识处理也就进入低层处理. 因此, 知识基系统成为图像理解系统的关键支撑软件.

早期如 Levine^[4]用于图像分割和 Mckeown 等^[5]用于解释航空图片, 都采用规则基产生式系统. 规则作为知识表达方式的特点是颗粒度较小, 且容易掌握编写, 但应用起来有一定困难, 例如, 一条规则的前提是道路上有一目标, 且可能是运动的车辆, 但在图像中识别道路并非易事. 单纯从底层数据向上的正向推理是很困难的. 规则数多的系统, 如果用穷举搜索的策略, 盲目性很大, 耗费在匹配的计算量也很大. 理解系统要反映出视觉判定目标的一些特点, 即对一个目标识别要反复处理提高置信度后才能作出判断. 许多情况下, 按需要识别的目标进行从顶层往下的推理比较容易, 或者从底层数据驱动到一定程度提出假设进行逆向推理验证也容易得多.

从文献[1~5]可看出图像理解的知识基系统发展趋势为: (1) 黑板多知识源协同工作系统, 把问题求解分割成较小的不同的领域问题专家知识源, 协同合作解决复杂的理解问题. (2) 正向数据驱动和逆向推理相结合, 能把理解目标的模型以决策树或语义网络形式存放在

系统中,便于逆向推理或按事例推理,这种方式又称为模型驱动。(3)知识表达采用框架式结构,容易把事实的描述和过程性描述结合在一起,框架之间及其槽的关系是一种描述性表达。而在槽中安排守护可引发过程检索的决策树、语义网络和规则等等。每个框架是一客体,又与软件面向客体发展趋势符合。

我所在 1989 年研制成低层处理的实时图像处理 ZRIP—1 系统^[6],随即研究建立了一个图像理解系统平台,本文介绍为这个系统研制的支撑软件知识基系统 V 语言。第 1 节简介第一版本的规则基 V.1 语言,第 2 节介绍改进成为面向客体式的 V.2 版本,第 3 节简单介绍应用情况和结论。

1 规则知识基 V.1 简介

1992 年底,我们研究成的第一个知识基系统是产生式规则基系统,取名为 V(VISION)语言,现作为第一个版本 V.1。

V.1 的设计思想是要与图像理解问题求解模型相配,并且是在并行处理器环境和并行 C 环境下的高级图像理解系统语言,它本身是一个协同式黑板专家系统。所谓协同式是指把知识系统划分成许多领域知识源,每一个知识源是一个推理功能模块,用多个知识源模块协同工作完成复杂的图像理解任务。这种做法和解题模型匹配,可以降低计算工作量。图 1 是图像理解问题模型。图像序列经过低层处理分割,标记成为团块数据结构,由特征提取知识源处理得到表示点、线、面的特征。由分裂和融合知识源把团块改组成为与目标相关的区域。由邻域关系知识源正向或逆向推理提出目标假设,经过景物分析知识源做出识别和场景解释,验证知识源负责推理的过程。黑板是工作存储器,即数据结构。每一知识源是一个推理模块,都有自己的模式匹配、冲突清除、问题规划等功能。与并行处理器系统相配的黑板层次和知识源装备在一个子结点上,硬件上一个处理器和本地存储器组合的子节点如图 2 所示。推理机、产生式规则系统与 OPS-83 方法类似。每个子节点也可以划成几个黑板层次,或包含几个知识源,视整个系统在工作时任务均衡分配而配置。

用 BNF(Backus Norm Form)表示的 V.1 语言的语法如下。

```

<V 语言文件> ::= <知识源说明表>
<知识源说明表> ::= <知识源说明> | <知识源说明> <知识源说明>
<知识源说明> ::= <知识源名说明> <规则事实集说明> <处理器配置说明>
<知识源名说明> ::= Knowledge-Source <字符串>
<规则事实集说明表> ::= <规则说明> <事实说明> <规则事实集说明表>
<规则说明> ::= (<规则名> <规则前提表> <演绎符> <规则行为>)
<演绎符> ::=  $\diamond$  | <正向演绎符> | <逆向演绎符>
<正向演绎符> ::=  $\diamond$  | <规则强度>
<逆向演绎符> ::=  $\diamond$  | <<规则强度>> =
<规则名> ::= <字符串>
<规则前提表> ::= <规则前提> | <规则前提> <规则前提表>

```

<规则前提> ::= <模式>
 <模式> ::= (<谓词><参数表>)|~(<谓词><参数表>)
 <参数表> ::= <原子>|<原子><参数表>
 <原子> ::= <符号>|<变量>|<整型数>|<浮点数>|<字符串>
 |<一维数组>|<二维数组>|NULL
 <规则强度> ::= <<浮点数>>
 <规则行为> ::= <C 语言程序体>
 <事实说明> ::= <规则强度><模式>|<模式>
 <处理器配置说明> ::= Configuration{<place 语句组><connect 语句组>}
 <place 语句组> ::= <place 语句>|<place 语句><place 语句组>
 <connect 语句组> ::= <connect 语句>|<connect 语句><connect 语句组>

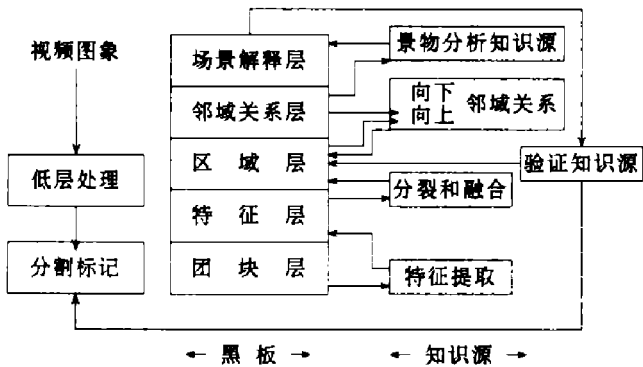


图 1 图像理解问题求解模型

Fig. 1 The modle for evaluating the image understanding

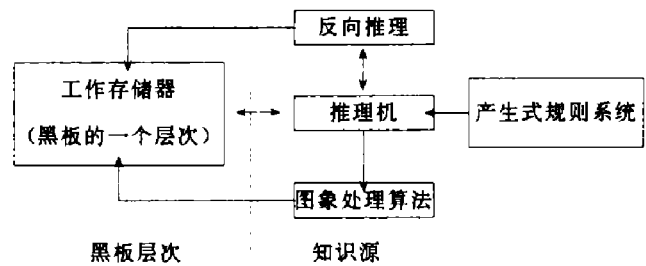


图 2 一个子节点的结构

Fig. 2 The structure of a subjoint-point

这个语法与一般产生式规则系统类似. 在我们的图像理解系统中知识源就是用这个语法写成的,我们以识别航空图中桥梁知识源中的两条规则为例. 其中一条是

```
[connect_river]
  (river_piece River_1) (river_piece River_2)
  (Sigma_for_Space Sigma)
  =<1.0>>{
max_distance=1.5 * Sigma * WIDTH_OF_BRIDGE/STD_SIGMA;
if (distance_of_river_piece (River1, River2,
  max_distance, line) <
  max_distance) {
  pattern:=(connect River_1 River_2 line);
  Store_pattern (pattern)
}
}
```

[

这条规则名称为 Connect_river(可相连的河),从航空图处理得到团块(river_piece River_1)和(river_piece River_2),并有一空间距离 Sigma 值.触发了这条规则后,计算 max_distance 值,并调用函数 distance_of_river_piece 检查两块间距离.如果小于 max_distance 值,则以置信度为 1.0 作为可相连的河模式,存储到黑板的特征以供进一步推理使用.另一条规则为

```
[get_bridge
  (bridge position)
  <<1.0>={
  image_pattern;=(image Aerial_image);
  retrieve (image_pattern);
  connect_pattern;=(connect River_1 River_2 Line);
  retrieve (connect_pattern);
  while (connect_pattern != Nil){
    position=get_bridge (Aerial_image, Line);
    pattern;=(bridge position);
    Store_pattern(pattern);
    connect_pattern=next_pattern(connect_pattern)
  }
}
```

这是一条逆推理规则,检索航空图片及可连接河流模式找到桥的位置.

在黑板上记录的数据都用二元组矢量表达,例如特征层中(line (linenum @ LMI) (r @ R1) (alpha @ ALPHA1) (Supporter @ SP!))表示线,其特征描述为线的标号,从 Hough 变换得到的到原点法线距离 r 、角度 α 及支持点数.二元组的第二项@标识符表示变量值(例如 r 的值是变量@R1),在以后推理运算中例化后就用一个具体值置换.各知识源都配有与一般产生式系统类似的推理机程序,并配有 3 类库函数(支持黑板的库函数,如上例规则中的 Store(pattern)功能是把模式写到黑板上去;支持逆向推理的库函数,如上例中 Retrieve(pattern)功能为检索已在黑板上的模式;及支持图像处理的库函数).每个知识源可自成系统工作,装备在一个硬件处理器——存储器节点上.各知识源集成一个系统协同并行操作,所以 V 语言是一个并行图像理解的语言^[7].

王伟明^[8]的实验说明,产生式规则系统在执行时模式匹配占用了大量计算时间,大的规则基系统采用 Rete 网络来提高效率^[9].在 V 语言中,分割成的各知识源较小,对规则数少于 40 条的知识源,采用完全状态保存的线性匹配算法效率较高,不必采用 Rete 网络方法.

另一个实验是对 2 个处理器——存储节点系统(设计了典型参数规则)测试黑板系统的性能,数据说明知识源的推理运算占 60.8%,用于黑板读写转播及通信等待开销占 7.8%,知识源内进程管理、通信信道管理及同步管理等管理开销占 31.4%.这在实时处理执行速度非常重要时是需要考虑的.

因此,我们在硬件上布置了一条数据总线,用双端口 RAM 做工作存储器(即黑板),读

写黑板就可以同时进行,以减少这方面的开销.

另一方面把知识基系统改编成面向客体框架式结构 V. 2 版本,减少管理开销,并且改进规则匹配效率,如下节所述.

2 V. 2 语言

93 年后,我们开始 V. 2 版本的研制工作. 其出发点是:(1)减小各种开销,改善规则匹配效率;(2)系统中能方便地安置决策树、语义网络以改进系统推理功能和效率;(3)在界面上要更方便于用户和系统的交互. V. 2 版本以面向客体框架式风格编写,达到了这些改进目标.

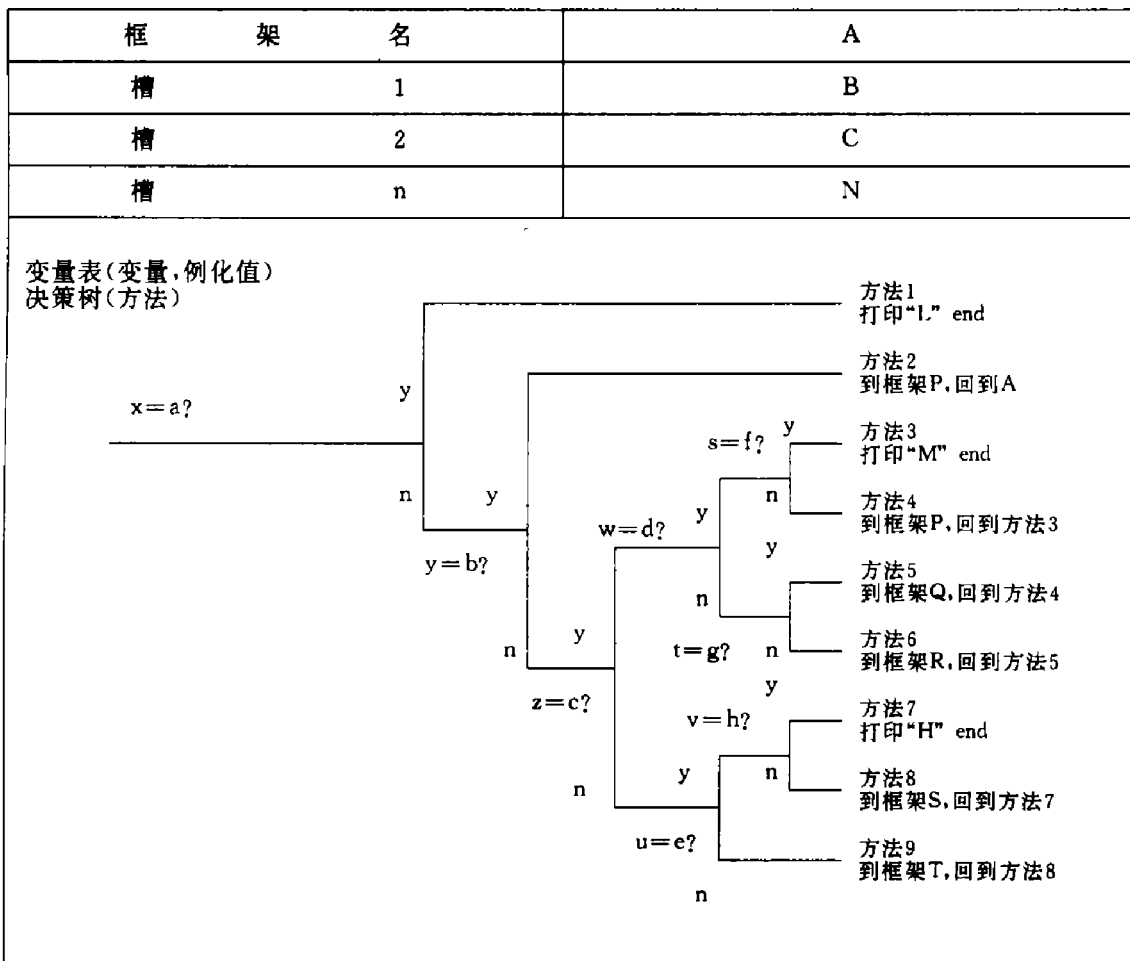


图 3 一个框架结构举例

Fig. 3 An example of a frame structure

在 V. 1 版本中事实描述为二元组矢量表达,而在 V. 2 中一个框架就是一个二元组矢量描述,但框架之间构成类和客体、双亲和子女的关系,构成了语义关系,对事实的描述就丰富而有层次.

另一方面为了减少规则模式的匹配,把有关连的规则化成语义网络或决策树,文献[10]说明这些表示是等价的. 例如图 3 所示框架.

框架之间构成推理网络如图 4,这个网络反映了图 3 框架部分的决策树形成的框架网络关系,框架之间的交互采用面向客体风格的消息传递方法.在 V.2 语言版本中仍然如 V.1 中一样包含支持推理及图像处理的库函数.在执行方面,如上所述构成了框架网络.尽量避免环境语言 C 函数的调用,而用指针传递.采用了网络和决策树结构,减少了很多匹配运算.以测试程序“猴子摘香蕉”在同一处理器执行比较.执行用 V.1 编写的程序速度为 190 条规则/s,而执行用 V.2 编写的程序等效于 3000 条规则/s,由此可见,在执行速度上 V.2 版本的改进是很可观的.

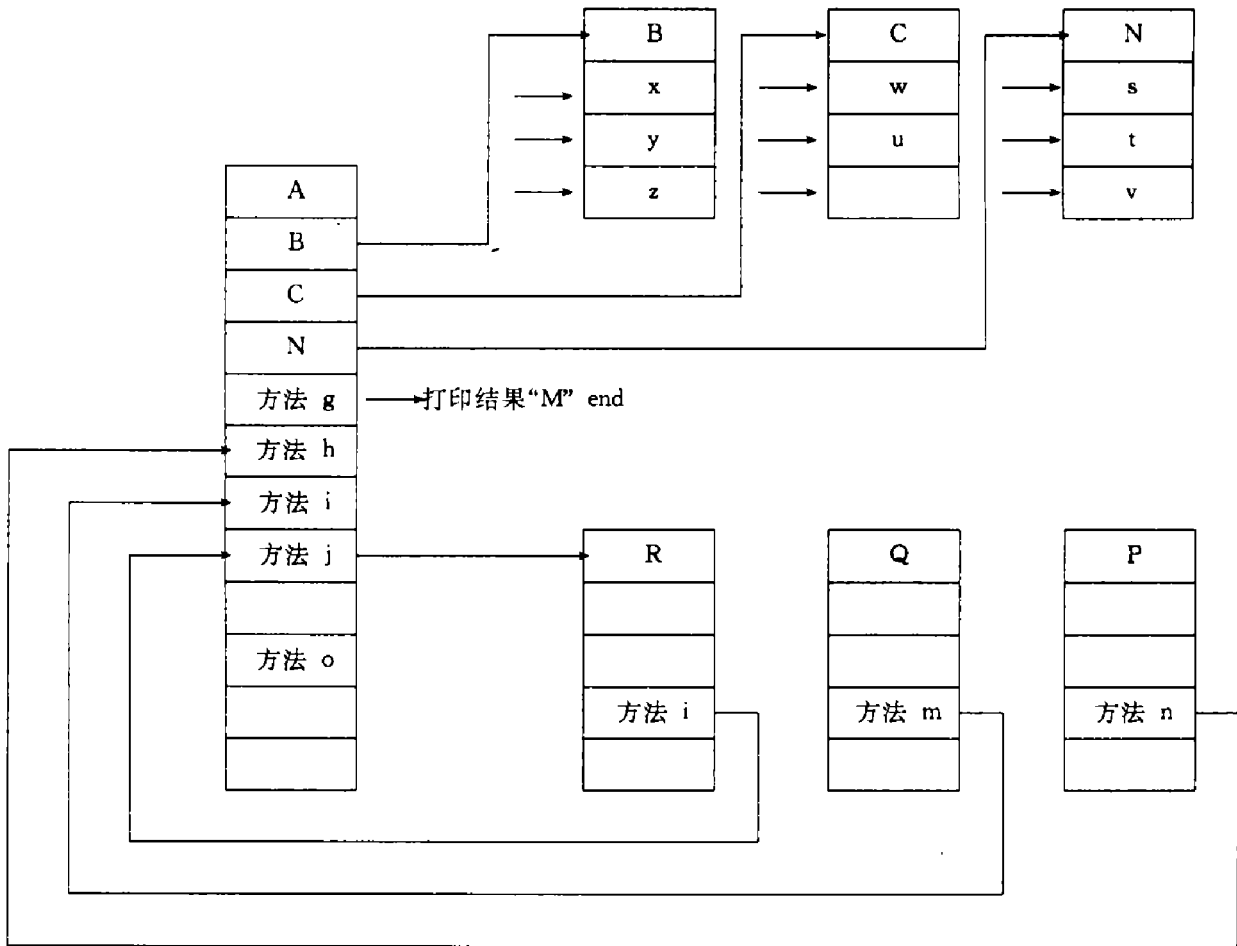


图 4 框架构成推理网络

Fig. 4 The inference-network formed by the frames

在用户界面上,V.2 版本采用窗口方式,汉字交互.用户编好框架就可显示.与规则基相比较,框架表达更接近于人的习惯表达方式.由于框架是面向客体风格的,并且是封装的,仅通过消息的交互发生关系,其工作软件便于由许多研究者各自单独编写,然后集成在一起,比规则基要方便得多,这对于图像理解解题方面尤其重要.

3 简单应用和结论

由于硬件系统正在制造中,V 语言在 386PC 机上用 Turbo C 执行并在 Transputer 的 2 个节点开发系统中做过并行工作的试验.图 5 是有桥梁的一幅航空图片和一幅遥感图片,采

用 V 语言编程执行后能从图片中找出桥梁目标.

具体方法为：(1)从图像中找出灰度变化比较平坦、长而蜿蜒的团块，把它假设为水域；(2)两块靠近的水域定为可连接水域；(3)检查两水域间团块是否窄长，成直线；(4)按文献 [11]介绍，检查两水域间团块两端是否和道路连接，如图 6 所示. 综上结果判断有否桥梁存在.

除图 5 所示由航空遥感图片检查桥梁以外，我们还做了从近景桥梁的图片分析判定桥梁及从航空图片中分割出建筑物. 这些工作证明了图像理解知识基系统 V 语言是成功的.



图 5 有桥梁的航空遥感图片

Fig. 5 Aeronautical and remote sensing pictures including bridges

在本工作中，我们得出以下几点结论：

(1)在硬件系统构成后，知识基系统 V 语言的装配要做到理解图像内容，还要做许多细致工作，包括 V 语言的定型和进一步改善.

(2)根据已做过的工作，图像理解方法的瓶颈在中层处理(如分割，特征提取)目前没有简单的或成熟的方法，并且耗费的计算量很大，如何应用知识基还需要进一步做工作. 这一点与许多研究者的看法一致.

(3)V 语言在并行硬件环境下工作、任务分割和调配及软件的优化还要做许多工作.

上述这些是我们今后工作的重点.

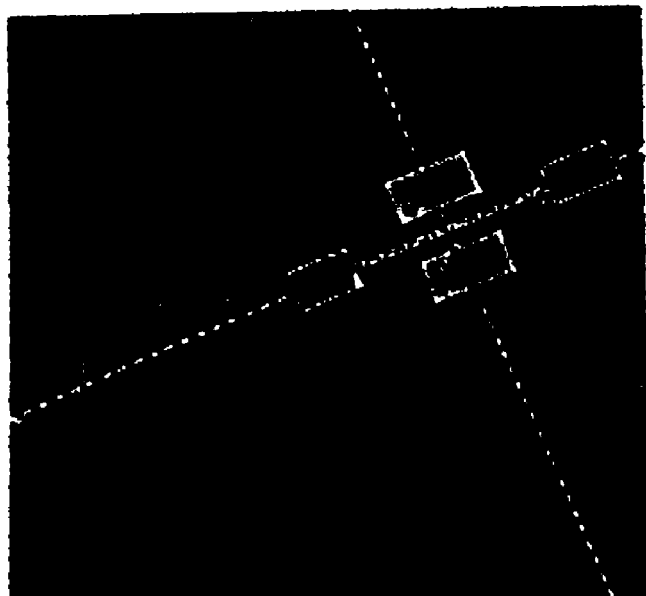


图 6 推理过程的图示

Fig. 6 The diagram showing the inference process

参考文献

- 1 Weems C C. *Proc. IEEE*, 1991, **79**(4):537~547
- 2 Weems C C, Riseman E E, Hanson A R. *Computer*, 1992, **25**(2):65~68
- 3 Gilmore J F. *Optical Engineering*, 1991, **30**(5):557~569
- 4 McKeown D M, Jr Harvey W A, Jr Mc Dermott J. *IEEE Trans. on PAMI*, 1985, **7**(5):570~585
- 5 唐国良, 荆仁杰, 姚庆栋, 等. 模式识别与人工智能, 1991, **4**(2):53~59
- 6 王伟明, 费旭东, 顾伟康, 等. 实时图像理解系统的专家系统语言, 浙江大学信息与智能研究所技术报告, 1991
- 7 王伟明. 用于实时视觉理解系统的并行知识基系统的结构与算法的研究, 浙江大学博士学位论文, 1992
- 8 Forgy C L. *Artificial Intelligence*, 1982, **19**(1):17~37
- 10 Nilson N J. *Principles of Artificial Intelligence*, 1980, Tioga Publishing CO.
- 11 Sauer L, Taskett J. *Proc. SPIE*, 1981, **292**:270~276

VISION LANGUAGE: AN IMAGE UNDERSTANDING KNOWLEDGE-BASED SYSTEM

Yao Qingdong Liu Jilin Xu Shengrong Hua Zhong

(Institute of Information and Intelligent System, Zhejiang University, Hangzhou, Zhejiang 310027, China)

Abstract The knowledge-based system Vision Language, its feature, performance and application are presented

Key words image understanding, knowledge-based system.