

一种吸引分岔知识网模型及其应用实例*

周昌乐

(杭州大学计算机科学系 杭州 310028)
(北京大学视觉与听觉国家重点实验室 北京 100871)

摘要 本文基于非线性动力学,特别是托姆的形态发生学思想,针对视觉学习,给出了一种吸引分岔知识网模型,用于解决知识表示和获取问题.通过引入皮亚杰发生认识论中的概念,模型拥有的学习功能包括强化、同化、顺应、聚合、分裂和遗忘;这样就给出了一个学习视觉知识的完整方法.3个应用系统的结果表明,该模型及其学习方法,对于解决实际问题,是有效和适用的.

关键词 知识表示,知识获取,学习方法,视觉计算模型,吸引分岔网.

在知识表示和获取方面,可以说任何一种模型及方法的提出,都是出于对解决实际问题的需要.比如语义网、脚本、产生式、框架以及近年来较新的面向对象、基于过程、概率逻辑、模糊知识和信念网等表示方法,都是在人工智能的发展中,为解决各种具体问题而应运而生的.但由于最近几年神经网络方法的重新兴起,在知识学习研究方面,偏重于学习算法研究而忽略了知识表示模型的发展性研究.针对这种情况,我们在视觉计算研究中,因为需要,借助于皮亚杰发生认识论原理^[1]和托姆形态发生学思想^[2],提出了一种知识表示模型,我们称之为吸引分岔知识网,并相应地给出了基于其上的知识动态积累方法.

顾名思义,吸引分岔知识网是一种基于形态发生相空间中吸引域和分岔点及轨迹等思想的知识表示方法.其目的是要表示在解决视觉问题过程中,状态相空间的整体刻画及其动态发展,以便通过自学习,使其解决问题的计算更为有效.

在本文中,我们将就这种知识表示模型及其动态积累方法作系统的介绍.内容包括模型的形式定义、知识的动态积累原理、动态模型的性能刻画及调节准则以及3个应用实例的简单说明.

1 吸引分岔知识网模型

为了能有效地刻画视觉问题解决过程中各状态及其相互关联情况,有助于视觉问题解决的计算实现,就需要有一种能反映问题解决过程中行为规律的知识表示模型,形式上我们有如下定义.

* 本文研究得到浙江省自然科学基金资助.作者周昌乐,1959年生,副教授,主要研究领域为计算视觉,模式识别,计算语言学.

本文通讯联系人:周昌乐,杭州 310028,杭州大学计算机科学系

本文 1995-06-23 收到修改稿

定义 1.1. 设视觉理解过程的状态特征全集为

$$Z = \{z^{(i)} | i=1, 2, 3, \dots, n\}$$

每个特征 $z^{(i)}$ 的取值范围为 $D^{(i)}$, 并记

$$Z(D) = \prod_{i=1}^n D^{(i)} = D^{(1)} \times D^{(2)} \times \dots \times D^{(n)}$$

则对于任意给定一个特征子集

$$Z_s = \{z_j^{(j)} | j=1, 2, \dots, n_s\} \subseteq Z$$

及各特征 $z_i^{(j)}$ 的取值范围 $D_i^{(j)} \subseteq (\exists i) D^{(j)}$, 我们称

$$Z_s(D_s) = \prod_{i=1}^{n_s} X^{(i)}$$

为 $Z(D)$ 的一个可变更域, 其中

$$X^{(i)} = \begin{cases} D^{(i)} & Z_s \text{ 中不存在 } z^{(i)} \\ D^{(i)} & Z_s \text{ 中有 } z^{(i)} \end{cases}$$

并称 Z_s 为该域的实质性特征集.

定义 1.2. 设 $Z(D)$ 定义如上, $Z_s(D_s)$ 为其上的一个可变更域. $y \in Z_s(D_s)$ 为任意给定可变更域 $Z_s(D_s)$ 中的一个特征向量值, 则我们称

$$u_s = (Z_s(D_s), p_s(y))$$

为 $Z(D)$ 上的一个吸引分岔知识单元, 其中记 $Y(u_s) = Z_s(D_s)$ 为 u_s 的吸引值域, $p_s(\cdot)$ 为作用在 $Z_s(D_s)$ 和 y 上的分岔算式.

定义 1.3. 我们称有向超图 $G = (U, E, W)$ 为定义在 $Z(D)$ 上的一个吸引分岔知识网. 其中节点集 U 中的元素为 $Z(D)$ 上的吸引分岔知识单元; 有向联边集 E 中的元素为满足诸种语义谓词的节点有序对; 而 W 为定义在 U 和 E 上的权值函数, 反映了每个节点或联边的重要程度. 进一步地, 如果语义谓词仅取

- (1) $Q_c(u_i, u_j) = Y(u_j) \subset Y(u_i)$
- (2) $Q_p(u_i, u_j) = \dim(Y(u_j)) < \dim(Y(u_i))$

且满足

$$(3) E = E_c \cup E_p \text{ 及 } E_c \cap E_p = \emptyset$$

则称 G 为 $Z(D)$ 上的一个狭义吸引分岔知识网, 其中

$$E_c = \{(u_i, u_j) | Q_c(u_i, u_j)\} \quad E_p = \{(u_i, u_j) | Q_p(u_i, u_j)\}$$

分别称为 G 的类聚关系集和构成关系集, 并称

$$U_p = \{u_j | (u_i, u_j) \in E_p\}$$

$$U_c = \{u_i | (u_i, u_j) \in E_c\}$$

$$U_s = \{u_j | (u_i, u_j) \in E_c\} \cup (U - U_c - U_p)$$

$$U_b = (U \cup U_p) - \{u_i | (u_i, u_j) \in E_p\}$$

$$U_o = U - \{u_j | (u_i, u_j) \in E\}$$

分别为组元节点集、类聚节点集、单元节点集、基元节点集和起始节点集, 称 $\bigcup_{u_0 \in U_o} Y(u_0)$ 为 G 的吸引值域, 记为 $Y(G)$.

注意, 可以规定每个单元节点均对应于视觉问题解的一种结果, 那么也就可以用单元节

点及其吸引值域的分布来反映一个知识网的功能.

定理 1.1. 我们称一个狭义吸引分岔知识网 G 是一致的, 当且仅当其满足

$$(\forall u \in U_c)(Y(u) = \bigcup_{u' \in U_p(u)} Y(u'))$$

这里 $U_p(u) = \{u' | (u, u') \in E_c\}$. 现设 G_1, G_2 为 2 个一致的狭义吸引分岔知识网, 如果有 $U_{s1} = U_{s2}$, 则存在一组子域 $Y_k, k=1, 2, \dots, l$, 使得

$$Y(G_1) = Y(G_2) = \bigcup_{k=1}^l Y_k = Y_1 \cup Y_2 \cup \dots \cup Y_l$$

成立.

有了定理 1.1 的保证, 对于任意给定的一个狭义吸引分岔知识网, 只要保证其一致性, 我们总可以通过对其进行不破坏单元节点的改造来获得一个功能不变, 具有更好结构的网络. 比如, 可以通过设置一个总类聚节点, 用类聚关系将所有起始节点联接起来, 就可以形成一个仅有唯一起始节点 u_0 和功能不变的吸引分岔知识网, 且有

$$Y(u_0) = Y(G) \text{ 及 } U_s = \{u_j | (u_i, u_j) \in E_c\}$$

以后我们总假定, 所给出的吸引分岔知识网, 都是仅含唯一起始节点的, 一致的狭义吸引分岔知识网.

2 知识网的动态积累原理

上面给出的知识网模型, 只是讨论了其静态结构, 对于学习系统而言, 更重要的是知识的动态积累问题. 本节我们将通过阐述知识网的吸收机制和调节机制, 来给出知识网动态积累发展的一般原理.

知识的吸收是指将理解过程中获得的新情况, 作为新知识, 吸收到知识网中. 视情况不同, 知识的吸收可分为知识的强化、同化和顺应.

简单地说, 强化是对已有知识确定性的强调; 同化是对知识在量上的积累, 属于吸引值域一般化过程; 而顺应则是对知识在质上的积累, 通过形成新的知识单元来完成, 所以是吸引值域的特殊化过程.

图 1 给出了吸收机制的实现原理, 其中遍历采用最左优先策略, 并假定理解所获得的结果是以一种反映景物构成关系的树结构形式给出的, 树中的节点值反映了该节点所代表景物成份的特征向量值. 这样, 对树中每个节点进行遍历, 按图 1 流程, 就可以完成一次知识的吸收处理. 其中各部分功能实现的判别准则是根据当前理解树节点的向量值与知识网中最吻合知识单元 u 之间的关系作出的, 分为如下 3 种情况:

(1) 强化: 如果 $y \in Y(u)$, 则通过加强知识单元 u 及其构成关系的权值来完成吸收处理;

(2) 同化: 如果表征 y 值的特征子集 Z_y 与表征 $Y(u)$ 的特征子集 Z_u 一致, 但 $y \notin Y(u)$, 则令 $Y(u) = Y(u) \cup \{y\}$, 实际上是变更了 $Y(u)$ 的吸引值域. 当然, 如果这一变更导致 $Y(u)$ 超限(事先规定的某个阈限)则转入顺应处理;

(3) 顺应: 不满足(1)和(2), 则进行顺应处理, 依据 y 和 Z_y , 创立一个新知识单元 u_y , 令 $Y(u_y) = Z_y(\{y\})$, 并在继承性构造了(利用 u 的分岔算式)相适应的分岔算式后, 将 u_y 挂入 u 的上层类聚节点上, 建立类聚关系, 赋初始权值.

这里, 所谓继承性构造分岔算式, 指的是根据基本指令集, 将 $Y(u_y)$ 与 $Y(u)$ 进行实质性

特征比较后,参照 u 的分岔算式,来构造适用于 u_y 节点情况下的新分岔算式,具有遗传的特点.

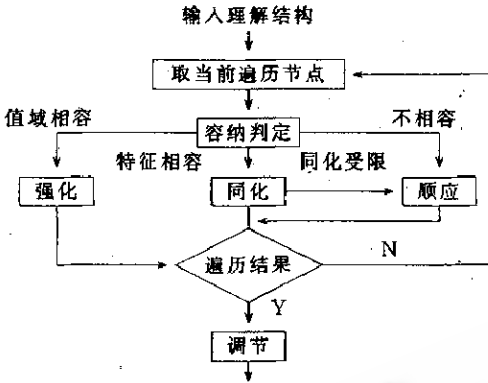


图1 吸收机制原理图

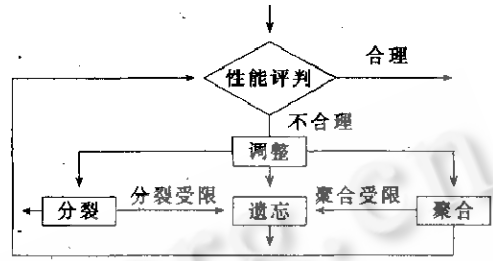


图2 调节机制原理图

只要在吸收处理的过程中,同时兼顾诸成份节点构成关系的及时修正,通过整个理解树的遍历处理,我们就可以完成整个理解结果的知识吸收.

不过,通过吸收,虽然解决了知识的增长问题,扩大了可理解景物的范围,但也会因知识的杂乱堆积,使知识网的性能下降,甚至出现理解混乱的现象.所以除了吸收机制外,我们还需要一种调节机制,使得知识网的性能/容量比保持最佳,真正提高理解能力.

所谓调节,就是对知识网进行重组和改造,并作为定理 2.1 保证功能不变的更一般情况,要求调节前后的知识网在功能上应是有增无减的.调节操作主要包括调整、遗忘、分裂和聚合,如图 2 所示.具体各部分功能简述如下:

(1)性能评判:按照一定的性能标准,对知识网进行性能综合评判,作为进行调节的依据.

(2)调整:调整主要是进行一致性检查,包括重复节点的检查、矛盾节点的检查、分岔算式适应性检查等,并对检查结果进行相应的增删和修改操作,然后按照性能评判结果去进行分裂、聚合和遗忘处理.

(3)分裂:在需要进行分裂的情况下,对所选节点进行分裂操作,分成 2 个或多个节点,并建立诸种相关的有向联边.

(4)聚合:聚合为合并和聚类 2 种.聚类是将多个类聚节点合而为一,形成出度更大的类聚节点;合并则是针对非类聚节点,将多个同种节点合并为一个节点.

(5)遗忘:遗忘包括删除节点、联边或压缩节点的吸引值域、修改节点和联边的权值等,以便在某种情况下提高整个知识网的性能.

注意,在进行了分裂、聚合和遗忘后,自然应该补建或修改各种相关的有向联边,不破坏知识网的完整性和一致性.

有了上述吸收和调节机制的实现,那么就不难给出吸引分岔知识网的积累过程:

(1)开始,只有一个起始节点,其分岔算式含有先验的理解、吸收和调节诸功能.

(2)当有输入景物图象到达知识网时,起始节点自动激活并执行分岔算式,对图象景物进行理解、吸收和调节,得到新的知识网.

(3)在任何时候,一个激活节点完成分岔算式后,就自动激活其后继节点,除非当前节点

已是单元节点为止,此时给出理解结果,并进行吸收和调节处理.

(4)每隔一定的时间,时钟自动激活起始节点,进行知识网的例行调节处理.

于是,只要最初给定一个初始节点及其先验分岔算式,那么通过对输入景物图象的不断处理,知识网就会不断壮大和完善.这便是知识网动态积累的基本原理.

3 知识网性能分析及调节准则

知识网的动态调节是依赖于知识网性能评判的,所以实施知识网的调节机制,离不开对知识网性能的确定和分析.通常,性能的确定与具体视觉系统的目标有关.在这里,出于通用性考虑,我们仅给出如下与具体问题无关的性能定义,作为知识网性能分析和调节准则的一般性依据.

定义 3.1. 设 $G=(U, E, W)$ 为一个一致的狭义知识网, u_0 为其唯一的起始节点,则我们定义如下参数来作为知识网 G 的性能描述:

$$(1) \text{时间消耗率: } K_1 = |E_c| / |U_c|$$

$$(2) \text{空间消耗率: } K_2 = |U| / h_m(Y(G))$$

$$(3) \text{知识内聚度: } K_3 = |U_s| / \sum_{u_i \in U_s} h_c(Y(u_i))$$

$$(4) \text{理解确定率: } K_4 = |U_s| / \sum_{u_i \in U_s} h_m(Y(u_i))$$

$$(5) \text{理解失误率: } K_5 = \frac{h_m(\bigcup_{u_i \neq u_j \in U_s} Y(u_i) \cap Y(u_j))}{h_m(Y(G))}$$

$$(6) \text{理解分辨率: } K_6 = |U_s| / h_m(Y(G))$$

$$(7) \text{知识分离度: } K_7 = (\sum_{u_i \neq u_j \in U_s} h_d(Y(u_i), Y(u_j))) / |U_s|^2$$

其中 h_m, h_c 和 h_d 分别为测度函数、非连通域计数函数和区域中心之间欧氏距离函数.

知识分离度和内聚度,反映的是单元节点吸引值域在整个特征空间 $Z(D)$ 上的分布状况;理解分辨率反映的是对景物理解区分的精细程度;理解确定率反映的是对景物理解的肯定程度;而理解失误率则反映的是对景物理解的失误概率;最后,时间和空间消耗率反映的是知识网的结构分布,通常与知识网动态积累中的时空代价有关.

性质 3.1 对于定义 3.1 给出 G 的诸种性能参数,其均为非负实数,且满足如下性质:

(1) $K_6^{-1} \leq K_4$; (2) $0 \leq K_5 \leq 1$; (3) $0 \leq K_3 \leq 1$; (4) 在 $Y(G)$ 恒定的条件下,性能参数均随知识网结构变化而变化,也即知识网的性能参数是可调节的.

显然,在性质 3.1 的基础上,通过进一步分析,我们就不难得到有助于调节处理的评判依据和准则,具体可归纳为如下 8 条规则.

规则 1: 如果 $K_3 \leq 1$, 则寻找 1 个含有 2 个以上分离区域的吸引值域之单元节点,对其进行分裂处理;

规则 2: 如果 $K_5 > 0$, 则寻找 1 个权值最低且其成员节点吸引值域交集最大的类聚节点,对其进行压缩处理,原则是减小权值较低的成员节点之吸引值域.如果成员节点的权值均高于某个阈值,则取消此次压缩操作.

规则 3: 如果 K_7, K_3 过低, 那么在知识网中, 选取某个权值最低且其成员的吸引值域在合并后能形成连通域的一类聚节点, 对其进行合并处理. 但此时, 如果合并后会减低 K_6 和 K_4 , 或被合并的成员节点存在不相容的组元节点, 则取消此次合并操作.

规则 4: 如果 K_5 过低且规则 1 和规则 2 均不满足, 则选择一个权值最低且其成员节点交集最大的类聚节点, 对其遗忘处理. 遗忘的原则是取一个权值最低、遗忘前与其它成员节点相交最大且对组元节点去留影响最小的成员节点进行.

规则 5: 如果知识网中出现重复节点, 则应对其进行简化处理, 取消重复节点并维持诸种关系.

规则 6: 经分裂、压缩、合并和遗忘后, 如果知识网诸节点之间产生不一致, 则应进行调整处理, 使其保持一致.

规则 7: 如果 K_2 过大, 在 K_1 容许的前提下, 则可酌情删除若干类聚节点, 以保证 K_2 维持在最佳范围之内.

规则 8: 在所有上述规则处理完毕之下, 只要空间资源允许, 如果 K_1 不合理, 则应适当对类聚节点进行聚类或分裂处理, 努力保证类聚节点的平均出度维持在 2~3 之间 (最佳为 e).

定理 3.1. 按照如上规则对知识网进行调节, 除空间消耗率外, 知识网的性能参数均保持有增无减 (对于失误率是有减无增) 且不改变知识网的一致性.

定理 3.1 告诉我们, 基于以上 8 条规则来进行知识网的动态调节, 对于给定的 $K_1 \sim K_7$ 7 种性能参数而言, 是可以改善知识网的内在理解能力的. 不过, 更一般地, 对具体视觉问题的理解计算, 往往还需要更多特定性能指标的综合考虑, 此时可以按照事先定义好的性能目标函数 $h(K_1, K_2, \dots, K_7)$, 并借助于数学规划方法, 来指导知识网的动态调节.

4 3 个应用实例系统的简要说明

一种好的知识表示和获取模型, 应该是能够经得起实际问题检验的. 自形成初步的吸引分岔知识网模型以来, 我们先后在简笔画脸谱识认自学习系统、歧义图形机器理解研究以及手写汉字识别系统中采用了该模型 (表 1 给出了 3 个实例中具体知识网模型的说明), 并取得了较好的效果, 我们分别简要说明如表 1.

1. 简笔画脸谱识认自学习系统: 该系统是要通过系统的自学习来完成对简笔画脸谱的自动分类识别. 实际系统是在 PC 机上用 C 语言实现的, 系统性能测试和分析表明, 就平均识认可信度、识认正确率而言, 在学习次数为 5 次时, 指标均达到 90% 以上; 而学习次数为 25 次时, 几乎达到 100% 的正确率.^[3]

2. 歧义图形的机器理解: 吸引分岔知识网的一个独有特点是通过突变类聚节点 (类聚节点的各成员节点吸引值域两两重迭), 它能很好地处理不满足分离性要求的模式, 而这一点在传统的模式识别方法上是不能想象的. 为了说明知识网的这一优势, 结合突变理论, 我们运用吸引分岔知识网及其积累原理, 具体实现了一个歧义图形的机器理解系统. 系统主要是针对兔鸭图形的区分识别展开的.^[4] 实验结果指出, 只要在原始图形中存在着微小差别, 通过实质性特征值的计算, 系统就能很好地区分出兔图和鸭图.

表1 3个实例知识网的具体说明

实例名称 说明项目	简笔画脸谱识别	歧义图形理解	汉字识别字库
对象形态	具体脸谱	歧义图形	手写汉字
有效特征	拓扑属性、线条形状、位置分布	区分特征	拓扑指数、笔顺形态、组合图式
类聚节点	一类脸谱	歧义图形	满足一定性质的字集
单元节点	单种脸谱	单种理解	单类汉字
组元节点	构成要素图形	构成要素图形	汉字组成构件
基元节点	基本线条	基本图形	基本笔划
有向联边	类聚关系、构成关系和语义关系	形似关系、构成关系	类聚关系、构成关系、同构关系、形似关系
权值含义	节点为可信度 联边为关联强度	节点为可信度 联边为关联强度	节点为存在概率 联边为关联强度
系统目标	形成特征空间取值分布的最佳 聚类划分	形成歧义图形区分特征的最佳 突变点集	形成对给定汉字样本全集的最佳 分类
附加性能	保证 $K_5 = 0$ 强调 K_3 和 K_7	允许值域重迭限制实质性特征 个数	限制单元节点规模,放宽类聚 点规模,允许交叉类聚

3. 手写汉字识别字库:如果说以上2个系统仅是实验系统,那么把吸引分岔知识网模型应用到实用化手写汉字识别系统的构造上,便说明了其解决实际问题的适用性。通过吸引分岔网模型构造的手写汉字识别字库^[5],从目前已获得的结果看,其性能指标要比未用此模型时^[6,7]优越得多。

总之,从3个应用实例的实验中可以看出吸引分岔知识网,作为一种知识表示和获取模型,是有着一定的优越性和普遍适用性的。当然,作为一种尝试性研究工作,尽管已取得了一些明显效果,但尚有许多更深入的问题有待于进一步完善,特别是关于分岔算式的形式分析和动态构造问题尤为突出。

参考文献

- 1 皮亚杰. 发生认识论原理. 北京:商务印书馆,1986.
- 2 雷内·托姆. 结构稳定性与形态发生学. 成都:四川教育出版社,1992.
- 3 周昌乐,伍可,骆建华. 一种视觉适应性计算原理与实现. 见:杨忠祥等编,中国机器学习'93,第4届全国机器学习研讨会,北京:电子工业出版社,1993.
- 4 周昌乐,施项君. 歧义图形机器理解初探. 计算机应用与软件,1996,13(2):36~41.
- 5 周昌乐,郑春英. 一种手写汉字识别字库及其自动生成. 杭州大学学报,1995,22(2):158~163.
- 6 周昌乐,马希文. 基于互动计算的汉字楷书识别. 自动化学报,1992,18(4):502~504.
- 7 周昌乐,马希文. 一种汉字楷书特征提取方法. 电子学报,1993,21(2):84~85.

A MODEL OF THE ATTRACTING—BIFURCATING NETS AND ITS APPLICATION

Zhou Changle

(Department of Computer Science Hangzhou University Hangzhou 310028)
(Audio and Video National Laboratory Beijing University Beijing 100871)

Abstract Based on the idea of the nonlinear dynamics, especially the Thom's Morphogenesis, a knowledge model, called attracting—bifurcating net, is advanced to solve the problem of knowledge representation and acquisition for visual learning in this paper. By introducing the Piaget's concepts in genetic epistemology, the model possesses seven main functions which include consolidation, assimilation, reunion, accommodation, segmentation and forgetting, thus having given out a completed method for learning visual knowledge. The results three application systems show that the model and its learning method are effective and adaptable to solve visual problems.

Key words Knowledge representation, knowledge acquisition, learning method, visual computation model, attraction—bifurcating net.