

文章编号: 1002-0446(2002)03-0201-08

# 自然语言训练的机器人基本行为控制器\*

聂仙丽 蒋平 陈辉堂

(同济大学信息与控制工程系 上海 200092)

**摘要:** 本文探索了一种直接利用自然语言进行机器人运动技能训练的控制方法,提出了利用模糊神经网络结构作为基本行为控制单元,通过教师的自然语言指令完成针对某一特定行为的运动经验获取和控制器训练,这是一种更加自然的控制器构造方式,以基本运动单元为基础,可以进一步实现机器人复杂任务的语言编程与控制.所提控制方法最终在一个轮式移动机器人系统上进行了语言训练实验.

**关键词:** 模糊神经网络; 语音训练; 自然语言编程

**中图分类号:** TP24 **文献标识码:** B

## ROBOT MOTION BEHAVIOR TRAINING BY NATURAL LANGUAGE INSTRUCTIONS

NIE Xian-li JIANG Ping CHEN Huir-tang

(Dept. of Information and Control Engineering, Tongji University, Shanghai, 200092)

**Abstract** A control method that trains robot to learn motion skills by natural language is explored in this paper, where fuzzy neural networks are used as a general structure of controller for different motion primitives. For any specified motion primitives, the controller acquires knowledge and is trained by instructions of natural language from a supervisor. This might be a more natural way to construct a motion controller. Based on the trained motion primitives, more complex motion tasks can be implemented. Experiments on natural language training for motion primitives have been done on a wheeled robot.

**Keywords:** fuzzy neural networks, language instruction based training, natural language programming

### 1 引言(Introduction)

语言是人类文明进步的标志,也是人们最方便的交流方式.通过自然的语言交流,实现人机交互式的对话形式,使机器能理解我们的语言,并按照我们的指示去工作,这是人类最大的梦想.随着语音识别技术的发展,人类梦想的实现变成了可能.

目前机器人编程技术仍依赖于手工的硬编码方式,但作为应用于人们日常生活的服务机器人,不可能要求使用者在掌握某种机器人语言后方可操作使用,一种更加易于接受的方法是采用人自身采用的自然语言交流方式进行指令理解、控制器构建.因此将语音识别技术、编程技术、控制技术结合将对机器人真正走进人类生活起极大的促进作用,语言交互为一般使用者提供了友好、直观的接口方式,人们不

必了解专门知识就可以自由地训练机器人,让机器人更好地为他们服务.为了达到目的,我们希望设计一种借助于自然语言的机器人编程方案,完成任务编程、运动编程、行为编程与控制.通常机器人运动行为可以分解为一组下层基本运动单元序列<sup>[1]</sup>,因此基本运动单元训练成为本研究的基础,本文将把注意力集中在这一基本运动单元控制器的自然语言构造方法上,通过教师口头指令,自动生成针对不同基本行为的运动控制器,如跟踪目标、避障、直线行走等.

但是,由于自然语言具有很大的模糊性,特别是当用语言指令来控制机器设备时,操作员所给出的语言指令也具有不确定性,同样的条件有可能给出不同的指令等,这些不确定性大部分表现在知识与

\* 基金项目: 国家自然科学基金资助(60175028), 国家教委优秀青年教师资助计划.

收稿日期: 2001-07-30

信息的模糊性上. 模糊逻辑由于其本身的言语化描述特点, 在处理自然语言时具有较大的优势, 故而得到广泛的研究. Sugeno 和 Park<sup>[2]</sup>提出了 FULLINS (Fuzzy-Learning based on Linguistic Instruction). FULLINS 具备利用先验知识从语言指令中获取新知识的能力, 它已被应用在拖车倒车控制和直升飞机飞行控制中<sup>[3]</sup>. Miyamoto 等<sup>[4]</sup>提出了一种基于定性语音模型的语言指令学习方法, 并且将这一方法用在化学反应过程的模糊控制当中. 模糊分类系统 Fuzzy Classifier System (FCS) 是一种非常有效的知识获取方法<sup>[5-7]</sup>, Fujii 及 Hasegawa 等提出了基于 FCS 的学习系统<sup>[8,9]</sup>, 这一系统能理解人类的语音指令并根据对语言指令的解释改变自己的行为. 这些系统都能通过自然语言和人类交流, 但它们的前提是要获得关于系统的先验知识, 组成基本规则库, 语言指令用于对规则库进行调整、更新. 本文提出了一种无需任何先验知识的控制器构造方法, 机器人能自动从教师语言指令中获取知识并组织规则库, 经过语言指令训练后机器人可以无需教师参与自主完成任务.

本文提出了以模糊神经网络作为一般的控制器形式, 针对具体行为, 通过教师语言指令自动产生与修正模糊规则, 这是一个自然语言理解并自动构造以传感器信号为反馈的控制器学习过程. 模糊逻辑和人工神经网络都可以表达和处理不确定性信息, 但各自具有优缺点. 模糊规则长于表示用语言描述的经验知识、定性知识, 但通常不具有学习能力, 模糊控制以模糊集合论、模糊语言变量及模糊逻辑推理为基础, 模拟人的近似推理和决策过程, 通常假设设计者已获得了专家知识和经验并选择了合适的隶属度函数和模糊规则. 目前, 在工业上投入运行的模糊控制器, 大都由一组模糊规则组成, 形成模糊控制规则表, 通过一定的模糊推理机制确定控制作用<sup>[10]</sup>. 这些模糊规则是人们对受控过程认识的归纳和控制经验的总结, 本文所提方法正是这一表格化经验的口授获得方法, 教师通过语言与机器人交互, 机器人在与教师的对话中获得经验知识, 自动形成规则库. 人工神经网络可以通过样本学习的方法, 将网络的输入输出关系以权值方式存储下来, 但网络内部的知识表达方式不清楚, 以致在学习时只能从任意初始条件开始, 不能利用已有的初始经验和知识, 学习时间长, 容易陷入局部最小, 且学习的结果完全取决于训练样本, 因而面临着如何生成可靠的训练样本的难题. 模糊神经网络将以上两者结合起

来, 优势互补, 一方面可以用语言描述的规则构造网络, 使网络的权值具有明显的意义; 另一方面引入学习机制, 改善原有规则, 提高了知识表示能力<sup>[11]</sup>. 我们所提出的语言指导下的运动控制器采用模糊神经网络结构, 将模糊控制和神经网络控制两者结合起来, 获得了一种基本运动控制器的统一结构, 使用者可以根据需要训练机器人的各种行为, 不同行为对应于一个网络实现, 因此这是一个使机器人由无知到有知的学习与积累过程. 为了加快训练, 提高学习效率, 本文借鉴增强式学习思想, 通过教师语言告诉机器人其控制性能的总体评估, 这一评估过程有效地加快了训练进程.

最后我们采用 Nuance 语音识别系统软件, 实现了一个移动机器人基本行为学习系统, 并针对一仿真机器人进行了各种运动行为的语言训练与控制. 实验结果表明机器人能有效地获取人的经验知识, 经过训练后, 机器人在不同的环境下均能很好地完成任务, 无需人的进一步参与.

## 2 系统结构 (System structure)

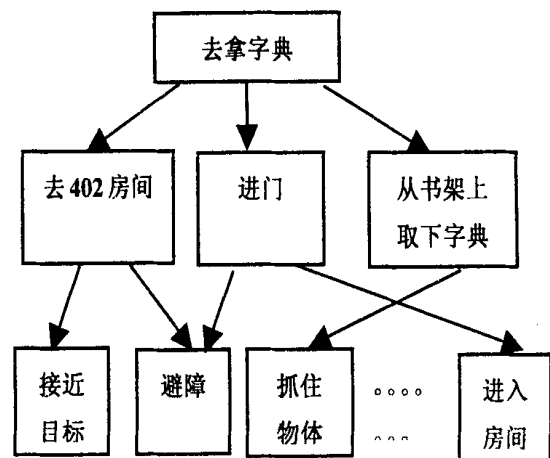


图1 自然语言编程的结构框图

Fig. 1 Structure of natural language programming

机器人运动任务通常包含三个层面, 抽象化的运动目标, 具体化的运动目标序列, 基本的运动行为单元, 如: 抽象化的运动目标: “将字典拿来”可分解为“到书房门口, 进门, 到书架, 拿书?”等串行动作序列, 同时应满足一些并行行为, 如: “若有障碍物则躲避”, 这些运动序列由基本行为单元构成, 如: “接近目标”, “进门”, “避障”, “抓取”等. 作为构成这一三层运动控制结构的基础? 机器人基本行为学习系统, 主要完成机器人基本行为的训练, 在机器人掌握了

基本技能后, 我们可通过与机器人的交互, 以自然语言的方式教机器人完成更高一级的任务, 这是本研究最终目标“自然语言编程”的基本思路. 本文将把研究注意力放在基本运动行为的语言训练方法上, 它是实现这一设想的基础, 机器人将通过教师口授自动构造不同行为的控制器, 学会各种基本运动行为.

机器人基本运动行为的语言训练系统是一个基于模糊神经网络的控制器, 教师观察机器人运动情况, 并给出相应的语言指令, 而语言指令将用于更新或调整模糊神经网络. 具体结构如图 2 所示, 下面对其进行分别阐述.

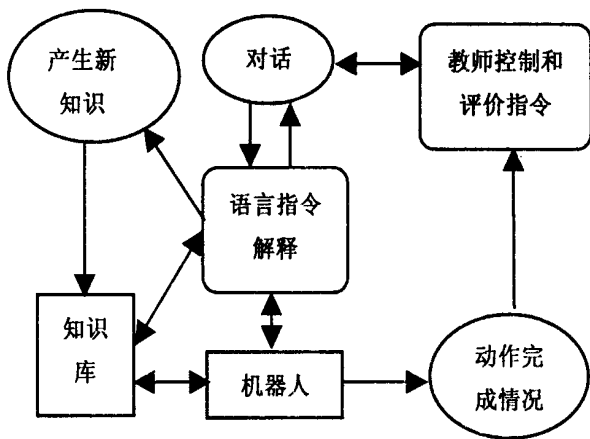


图 2 系统结构

Fig. 2 System structure

### 2.1 语言指令处理模块

语言指令处理模块是指图 2 中所有人与机器人的交互部分, 包括教师与机器人之间的对话、教师语言指令的理解也即指令解释, 这个模块组成了人机语言交互界面.

语言指令处理模块基于自然语言对话机制, 教师通过麦克风输入语音指令, 如“前进”、“向左转”、“向右转”、“后退”、“角度大一点”、“角度小一点”等, 来控制移动机器人的动作, 使其达到预定的目标. 同时教师可对机器人动作执行情况给予评价, 如果认为机器人走得好, 可以说“很好”, 系统会增加相应动作的评价值(可信任度), 若认为机器人任务完成很差, 教师评价为“差”, 系统会降低相应动作的评价值, 我们将在后面的算法一节中进一步说明. 教师与机器人的一个对话示例如下:

机器人: 请输入指令  
 教师: 向前(不清楚)  
 机器人: 我没听懂, 请重试  
 教师: 向前  
 教师: 向右  
 .....  
 教师: 很好

### 2.2 知识库模块

该模块由一个三层网络组成(如图 3 所示), 其主要功能是存储由传感器输入到控制输出的模糊控制规则, 并以操作人员的语音指令为输入, 完成网络模糊规则的动态生成和实时调整.

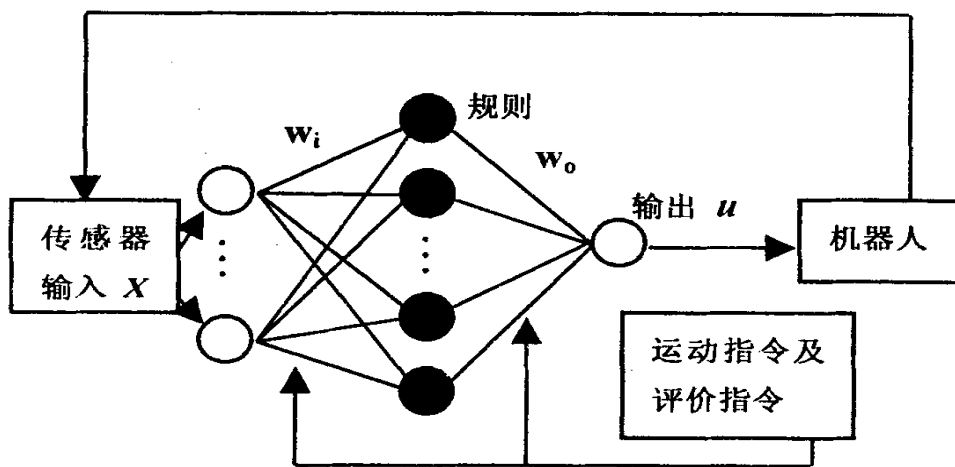


图 3 模糊神经网络的结构

Fig. 3 Structure of fuzzy neural network

网络的第一层为输入层, 每个节点表示一个传感器输入变量, 将传感器输入量规一化后, 作为第一层的输入. 网络的第二层为规则层, 每个节点表示一条规则. 这些规则是根据语言指令并由传感器输入状态, 采用自组织竞争方式自动获得输入空间聚集点. 每条规则的可信任度由教师对控制性能的评价决定. 该层与第一层的连接完成模糊规则前提条件的匹配, 实现“模糊逻辑与”运算.

网络的第三层为输出层, 采用有教师学习方式根据教师指令调整网络权值, 其输出为数值化的控制量, 如针对移动机器人输出层应包括两个神经元, 分别对应机器人运动速度及转动角度, 当指令为“向左”或“向右”时, 转动角度值相应大于零或小于零, 当指令为“向前”、“向后”时, 转动角度为零, 而速度值应分别大于零或小于零.

### 3 模糊神经网络的训练算法 (Training algorithm of the fuzzy neural networks)

为实现一个针对不同基本行为的通用的控制器, 我们采用模糊神经网络作为基本结构. 对任意给定的不同的基本行为, 都相应产生一个新的相似的神经网络结构. 机器人以这一神经网络作为控制器从任意的初始状态出发, 教师观察其动作执行情况, 给出语言指令以使机器人完成给定的任务, 神经网络根据教师给出的指令进行学习或添加新的神经元 (模糊规则). 具体训练算法如下:

(1) 初始规则数  $k=1$ , 相应地, 规则层的神经元个数为 1. 随机选择输入输出层的权值  $w_i, w_o$ .

(2) 采样机器人传感器的反馈信号向量  $X \in R^n$ , 这一反馈信号的具体形式依赖于机器人所采用的传感器类型.

(3) 计算输入向量  $X$  和输入层聚类中心  $w_i$  的距离, 其中  $w_i \in R^n$  对应于规则  $j$  的前提条件:

$$D_i = \|X - w_i\| \quad 1 \leq i \leq k \quad (1)$$

然后定义输入向量  $X$  与规则  $j$  前提的符合度

$$R_i = \begin{cases} (M - D_i)/M & \text{if } D_i \leq M \\ 0 & \text{if } D_i > M \end{cases} \quad (2)$$

其中  $M$  是一个给定的阈值, 表示和聚类中心接受域的最大距离,  $R_i \in [0, 1]$ . 这对应于模糊控制中利用三角形隶属度函数将输入量模糊化.

(4) 根据式(2)找出针对当前采样  $X$  的优胜规则  $c$ .

$$c = \arg \max_i (R_i)$$

$$1 \leq i \leq k \quad (3)$$

如果  $c=0$ , 则表明没有优胜规则.

定义一个宽度为  $N$  的优胜规则窗口  $Y$

$$Y = [c_1 \dots c_N] \quad (4)$$

该窗口记录最近  $N$  条优胜规则, 其中  $c_N$  是最近的优胜规则. 把  $Y$  左移一位, 并令  $c_N = c$ .

(5) 如果没有语言指令, 转到第 7 步, 输出本次采样控制量.

(6) 如果有语言指令, 根据所给指令修改优胜窗口的所有规则的输入、输出权值.

(1) 运动指令:

优胜规则窗口  $Y$  中输入层权值  $w_{icj}$  的 Hebbian 学习:

$$w_{icj}(k) = w_{icj}(k-1) + \alpha^* \eta^{N-j} (X - w_{icj}(k-1)) \quad j \in [1, N] \text{ and } c_j \neq 0 \quad (5)$$

其中  $\alpha$  是学习步长, 且  $0 < \alpha \leq 1$ .  $\eta$  是遗忘因子,  $0 < \eta \leq 1$ , 反映当前语言指令对历史规则的作用程度.

优胜规则窗中的输出层权值  $w_{ocj}$ ,  $j \in [1, N]$ , 的有教师学习:

$$w_{ocj}(k) = w_{ocj}(k-1) + \tau^* (\mu^* L) \quad (8)$$

其中  $\tau$  是一个可变的步长, 而  $\mu$  是副词变量  $\mu \in [-1, 1]$ , 如“大一点”、“小一点”等,  $L$  为语言指令动作向量, 即教师指令中“向左转”、“向右转”等对应的基本运动向量, 设  $L = [V \ \Delta\theta]^T$ , 如果指令是“向左转”, 则转角  $\Delta\theta$  大于零, “向右转”则小于零, “向前”和“向后”时转角均为零, 而  $V$  分别对应于大于零和小于零的值. 因此教师给出的运动指令包含了动作向量  $L$  及修饰此向量的副词. 副词是可选的, 没有时认为  $\mu$  等于 1. 式(6)中的  $\mu^* L$  给出了语言指令的控制解释.

由于语言指令的模糊性, 可变学习步长  $\tau$  由下式决定:

$$\tau = \delta \left\| \sum_T \mu L \right\| \quad (7)$$

其中  $\left\| \sum_T \mu L \right\|$  是在时间间隔  $T$  内教师所给出的运动指令的总体作用大小. 选取这一增益形式的理由是因为如果教师连续多次说同方向运动指令, 如“向左”“向左”“大一点”..., 显然教师认为目前控制修正  $\mu^* L$  不够, 需要一个更大的增益来使输出权值较快地达到期望值. 相似地, 如果教师在给定的时间间隔内给出一系列运动方向相反的指令, 如给出指令“向左转”后又给出“向右转”等, 这时对应于控制振荡的发生, 应该减少增益才能使输出较为稳定. 而式(7)

中  $\sum_T \mu L$  正是对发生这些情况的一种估计, 对指令累加就能达到增加或减小学习增益的目的. 式中的系数  $\delta$  是与教师评价指令有关的函数, 即由相应规则的可信度决定, 下面将进行具体阐述.

#### (2) 教师评价指令

教师可根据机器人动作执行情况给予评价, 如“好”或“差”, 我们定义了奖赏函数如下

$$r = \begin{cases} 0.1 & \text{“好”} \\ -0.1 & \text{“差”} \\ 0 & \text{其它} \end{cases} \quad (8)$$

由于教师通常给出的是一段时间内动作的总体评价, 所以我们对优胜规则窗口中所有规则的评价值进行修正:

$$\begin{aligned} q_{cj} &= q_{cj} + r \\ j &\in [1, N] \\ cj &\neq 0 \end{aligned} \quad (9)$$

这就是所谓增强学习中的一步 Q 学习<sup>[12]</sup>, 动作执行单元接收传感器的输入, 然后确定下一个行为, Q 函数  $Q(x, a)$  与每一个状态-行为联系起来, 是在状态为  $x$  的前提下, 执行动作  $a$  后所期望的未来的奖赏<sup>[13, 14]</sup>. 通常情况下, 规则  $cj$  在以往历史过程中的控制效果越好, 其评价值  $q_{cj}$  越高. 因此我们定义如下学习增益  $\delta$  与评价值  $q_{cj}$  的关系式为一 Sigmoid 函数

$$\delta = \frac{1}{1 + e^{q_{cj}}} \quad (10)$$

这意味着教师评价较高的规则, 其修改的可能性及修改的幅度值要小于评价低的规则, 即尽可能维持优秀规则, 而调整不良规则.

#### 7 增加新规则(添加新的神经元)

如果教师给出了一条运动指令, 而当前传感器状态通过式(3)选择没有优胜规则存在, 则添加一个新的神经元且  $k = k + 1$ . 假定输入的传感器向量为  $X_s$ , 则设定该神经元输入层的权值为  $X_s$ , 输出层的权值为  $\mu * L$ , 并令评价值  $q$  为 0. 则等价于在知识库中增加了一条新的模糊规则如下

$$\text{if } X \text{ is } X_s \text{, then result is } \mu * L \text{ credit is } q \quad (11)$$

#### 8 输出本采样周期的控制量

$$u = \sum_i R_i w_{oi} \quad (12)$$

#### 9 返回第二步.

第一步到第九步给出了基于语言指令训练的基本行为控制器的控制方法. 我们在此算法中综合考虑了自然语言的特征, 模糊规则学习等因素. 该算法

已在一台 Pentium III 的计算机上实现. 在下一节中, 我们要将此算法应用在轮式移动机器人的基本行为训练上.

## 4 语言训练的实验研究 (Experiments of training by language instructions)

我们针对一个仿真移动机器人进行了语言训练的实验研究. 该机器人仿真环境是一个  $600\text{cm} \times 500\text{cm}$  的“房间”, “房间”中包括了静态目标和动态目标. 静态目标是一个方形物体, 其位置可以随意设定, 机器人可以从不同的位置以不同的方位出发来接近目标. 动态目标是一个移动的球, 它碰到墙后会反弹回来. 我们对机器人做了接近静态物体、跟踪动态目标及避障等基本行为的训练. 我们采用了面向对象的编程技术<sup>[15]</sup>来设计基本行为学习系统, 用 java 语言编制而成<sup>[16]</sup>. 其中语言指令处理模块采用 Nuance 语音识别系统<sup>[17]</sup>, 实现了人与机器人的自然语言对话. 我们首先介绍 Nuance 语音识别系统, 然后对机器人的运动学模型进行了说明, 最后展示了实验结果.

### 4.1 Nuance 语音识别系统

Nuance 语音识别系统是由美国 Nuance 公司开发的具有世界领先水平的语音识别软件, 它提供先进的语音识别、话者身份确认及语言理解功能. 它主要由三个部分组成 RecClient、RecServer 及 Resource Manager. 一个 Nuance 系统的应用程序可以由多个 RecServer 和多个 RecClient 组成, 而且这些 RecClient 和 RecServer 可以分布在一个局域网中的不同机器上. 其分工如下: RecClient 进程负责语音输入、输出及输入的预处理, 应用程序的界面等. RecServer 进程负责语音识别, 话者身份验证及自然语言理解.

Resource Manager 管理在多个 RecServer 之间负责识别任务的分配.

Nuance 系统的语音识别流程如下: 语音信号输入, 通过 RecClient 的预处理后, 转换成声音波形, 交给 RecServer, 进行所谓的前端处理 (Front-end processing), 也即对信号进行频谱分析, 提取声音特征, 然后综合考虑声音模型 (Acoustic Models)、字典文件 (Dictionaries) 及语法文件 (Grammars), 从中找出与此声音特征匹配的词句, 经过解释后, 输出此词句的含义. 其中声音模型由 Nuance 系统提供, 主要用于组成词句的音素识别, 如英语单词 pat 和 bat, 其中的 p 和 b 就是两个不同的音素. 字典文件包含

对音素发音的描述,负责把音素序列组合成单词,用户可以建立自己的词典.识别语法(Grammars)则定义了可被识别的词或句子及这些词句的含义,只有被语法包含的词才可被识别.语法有专门的语言.GSL(Grammar Specification Language).我们根据控制要求,编写了如下汉语语法文件:

```
. Choices
[
  [ (xiang4 zuo3 zhuan3) ] { speaker_said "a" }
  [ (xiang4 you4 zhuan3) ] { speaker_said "b" }
  [ (xiang4 qian2) ] { speaker_said "c" }
  [ (xiang4 hou4) ] { speaker_said "d" }
  [ (da4 yil dian3) ] { speaker_said "i" }
  [ (da4 yil dian3 dian3) ] { speaker_said "j" }
  [ (xiao3 yil dian3) ] { speaker_said "h" }
  [ (xiao3 yil dian3 dian3) ] { speaker_said "g" }
  [ (hao3) ] { speaker_said "9" }
  [ (cha4) ] { speaker_said "6" }
]
```

Nuance 采用汉语拼音来识别汉字,用数字 1, 2, 3, 4, 5 分别代表汉语声调的第一声、第二声、第三声、第四声和轻声.如[(zuo3) (xiang4 zuo3 zhuan3)] {speaker\_said "a"}, zuo3 就是汉语的“左”, xiang4 zuo3 就是汉语的“向左”,后面的 {speaker\_said "a"} 代表 Nuance 系统在识别出“左”或“向左转”这两个词后所返回的结果为符号“a”.语法文件编译后,就生成一个包含了特定声音模型、字典文件和语法文件的包(package),以后的语音识别都在这个 package 上进行. Nuance 的语法有动态语法和静态语法之分,动态语法可以动态修改语法文件,并自动编译执行.静态语法预先编译好语法文件,可以加快识别速度,提高识别精度,两者各有优缺点,在此不详述.

Nuance 系统为编写基于对话的语音识别应用程序提供了极大的方便,用户可以使用 Microsoft Visual C++ 编程语言直接调用 Nuance C++ recognition API 函数,或用 Microsoft Visual Basic 开发基于 Nuance ActiveX Speech Channel API 的应用程序.但是,最为方便的是采用 Nuance 提供的 Java 类库 SpeechObjects<sup>[18]</sup>,用 JDK1.2 以上的版本编写应用程序.笔者就是采用 SpeechObjects 实现了语音控制机器人的对话系统.

#### 4.2 机器人及控制目标

目前使用的移动机器人有两种方向控制方式:

方向盘或差动控制.前者的方位由一个类似于舵的装置控制,机器人的运动由速度  $V$  和角速度  $\omega$  确定.而后者的左右轮可以独立控制,利用两轮速度的差异达到控制方向的目的,控制量为左右轮速度  $V_r$  和  $V_l$ . ( $V, \omega$ ) 和 ( $V_r, V_l$ ) 的关系如下

$$\begin{bmatrix} V \\ \omega \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 \\ 1/w & 1/w \end{bmatrix} \begin{bmatrix} V_r \\ V_l \end{bmatrix} \quad (13)$$

其中,  $w$  为两后轮轮间距.  $V_l$  和  $V_r$  分别为左轮和右轮的前进速度.假定此移动机器人在水平面上运动,车轮纯滚动无滑动,即轮子与地面接触点之间相对速度为零.在反馈控制中都以轮基中点  $q = (x, y, \theta)$  为基点进行,其中  $x, y$  分别为移动机器人的位置坐标,  $\theta$  为移动机器人相对于外部坐标系  $X$  轴的姿态角,以逆时针方向为正,则可得到平面运动的轮式移动机器人运动学方程为

$$\dot{q} = \begin{bmatrix} \cos\theta & 0 \\ \sin\theta & 0 \\ 0 & 1 \end{bmatrix} u(X) \quad (14)$$

控制示意图如图 4 所示.为了使机器人跟踪运动目标,机器人分别具有两种传感器输入量  $X = [D1, \varphi]$ , 其中  $D1$  为两者的相对距离,  $\varphi$  为机器人和目标的相对角度,它们可方便地由摄像机或激光扫描仪获得.同时机器人的控制量为  $u(X) = [V, \Delta\theta]$ , 由图 3 模糊神经网络输出,其中  $V$  为机器人运动速度,  $\Delta\theta$  为角速度的离散实现,即  $\tilde{\omega} = \Delta\theta/\Delta t$ .

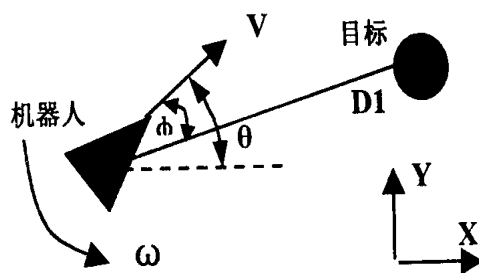


图 4 机器人控制参数示意图

Fig. 4 Denotations of robot control parameters

#### 4.3 实验结果

我们用所提方法对机器人进行了一系列基本行为的训练,如接近目标和避障等.当我们处理目标跟踪的问题时,目标如图 4 所示,训练机器人避障时我们把移动的球作为障碍物.所有的程序都在一台 Pentium III 的计算机上实现.下面仅给出机器人跟踪静态目标的仿真结果.我们让机器人从不同的初始状态出发接近静态目标,在训练了一定次数后,机

器人能在无教师指导的情况下自动接近目标, 并可以自动跟踪动态目标, 达到了预期目的. 设定输入层学习参数  $a = 0.1$ ,  $T = N = 5$ ,  $\eta = 0.9$ , 副词  $\mu$  的取值分别为 0.4、- 0.4、0.2 和 - 0.2, 对应于指令中的“大一点”、“小一点”、“大一点点”和“小一点点”. 动作向量  $L = [V \ \Delta\theta]$ , 当指令为“向右转”和“向左转”

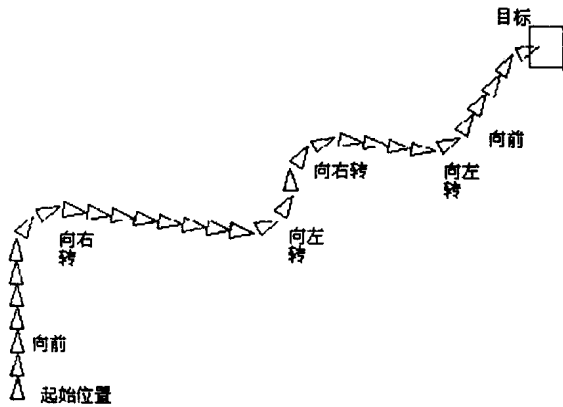


图 5 语言指令插入点说明  
Fig. 5 Examples of instruction training

时,  $V$  为 0, 而  $\Delta\theta$  分别为  $-30^\circ$  和  $30^\circ$ , 指令为“向前”和“向后”时,  $V$  分别为 9 和 - 9, 而  $\Delta\theta$  为 0.

$M$  分别取 200 和 300, 我们对  $M$  取两个不同值的情况做了比较, 图 5 给出了一次语言训练实例, 图中“运动指令”标记了在训练中教师语言指令的插入点.

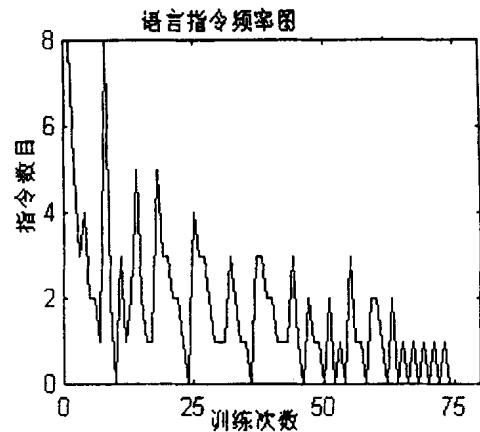


图 6 语言指令频率图( $M = 200$ )  
Fig. 6 Numbers of the movement instructions( $M = 200$ )

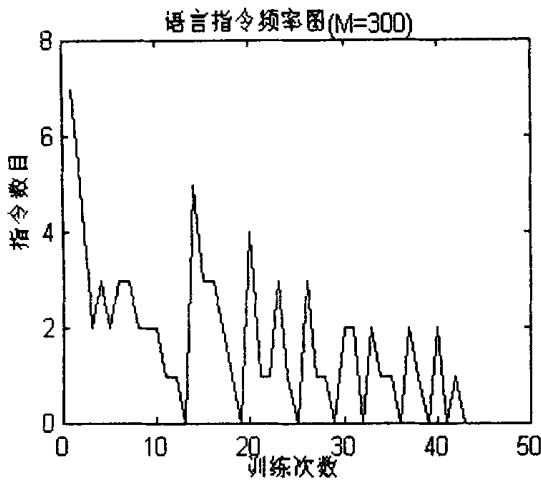


图 7 语言指令频率图( $M = 300$ )  
Fig. 7 Numbers of the movement instructions( $M = 300$ )

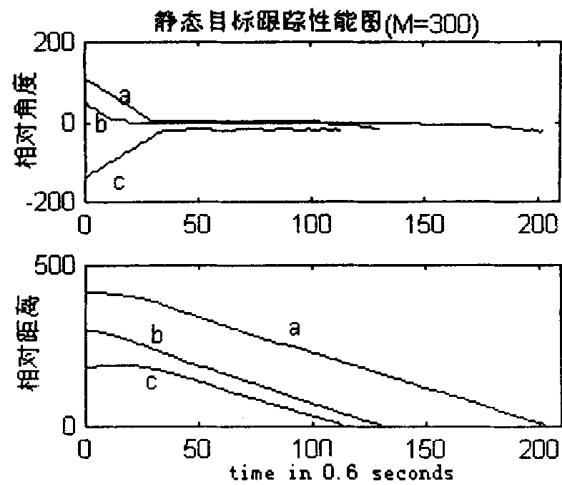


图 8 静态目标跟踪性能图  
Fig. 8 Performance of static target tracing

图 6、图 7 分别统计了在反复训练过程中教师语言指令数, 它反映了教师的参与情况和机器人掌握技能情况. 从实验结果可以看出语言指令数目在整体趋势是随着训练次数的增加而减少, 表明机器人掌握的知识在增长, 而越来越不需要教师的参与. 图中的各个峰值表示了机器人由不同的初始状态出发, 而当前规则库无法完全包括这一新的状态, 因此

需要教师重新给予语言指令、增加指令数目. 我们注意到  $M = 200$  时训练次数要大于  $M = 300$  的情况, 且语言指令的数目也要多于  $M = 300$  的情况. 原因如下: 由于控制器的输出是由多条模糊规则共同作用的结果, 较小的  $M$  对应于较精细的输入空间划分, 因此较小的  $M$  适合于学习有较高精度要求和较复杂的映射关系, 但需要更多的训练样本以产生更多

的控制规则;而较大的  $M$  具有较大的输入接收域,通常它可以更加充分地利用模糊推理的泛化能力,用较少的规则数拟合输入输出关系,但通常更加适合于一些简单映射关系的拟合. 在我们所提方法中,由于针对机器人的基本运动行为学习,所以这种关系本身较为简单,与  $M = 200$  相比,采用  $M = 300$  将可以以更少的指令数、更少的规则数达到我们的基本行为要求.

在语言训练后,不再需要教师的参与机器人就可以利用所学的规则库完成指定基本行为,图 8 描述了利用训练好的神经网络( $M = 300$ )控制机器人从不同初始位置出发到达目标的过程,在图中  $a, b, c$  分别表示机器人从三个不同初始位置出发的运动曲线. 可以看出,所提方案使  $a, b, c$  情况下的相对距离都接近于 0,即到达了指定目标,而且无振荡现象. 在  $c$  初始状态下,可以看出相对距离是单调减少的,但相对角度的减少却略呈阶梯形,这与人类判断具有模糊性有关,人们只有在认为超过某一容忍程度后才发出指令,不同的人其容忍度不同,而作为具有自组织能力的模糊神经网络,所生成的模糊规则继承了教师在训练过程中所表现出的某些习惯. 图 8 也反映了机器人已经很好地掌握了所要求的基本行为“接近目标”.

## 5 结论(Conclusion)

本文提出了一种基于自然语言训练的机器人基本行为的模糊神经网络控制方法,并利用此方法完成了机器人基本行为语言训练系统,为实现更高级的机器人自然语言编程提供了基础. 该模糊神经网络能从机器人与教师的对话中获取知识,并形成知识库,语言训练实验证明这种方法具有普遍性,对于不同的基本行为都可用同样的神经网络结构组成控制器. 在未来的工作中,我们将以本文赋予机器人的基本行为能力为基础,研究复杂运动的自然语言编程,机器人通过语言方式与人交互,自主地完成指定任务.

## 参考文献 (References)

- Schaal S. Is imitation learning the route to humanoid robots?. *Trends in Cognitive Sciences*, 1999, 3(6): 233- 242
- Sugeno M, Park G. Learning Based on Linguistic Instructions Using Fuzzy Theory. *the J of Japan Society for Fuzzy Theory and Systems*, 1992, 4(6): 1164- 1181
- Sugeno M, Park G. An Approach to Linguistic Instruction Based Learning and its Application to Helicopter Flight Control. *Proc of fifth Int'l Fuzzy Systems Assoc. World Congress'93*, 1993, (IFSA'93): 1082- 1085
- Miyamoto A, Goto K, Itoh O, Migita H, Sugeno M. Method for Adjustment Control Rules by Linguistic Instructions. *Japanese Journal of Fuzzy Theory and Systems*, Allerton Press, Inc, 1996, 8(5): 847- 861
- Mokato FUJII, Takeshi FURUHASHI. A Proposal of Human-Machine Interactive System through Linguistic Instructions Using Fuzzy Classifier System. Nagoya University, Japan
- Nakaoka K, Furuhashi T, Uchikawa Y. A Study on Apportionment of Credits of Fuzzy Classifier System for Knowledge Acquisition of Large Scale Systems, *Proc of third IEEE Int'l Conf on Fuzzy Systems*, 1994, (FUZZ- IEEE'94): 1797- 1800
- Furuhashi T, Nakaoka K, Uchikawa Y. A Study on Fuzzy Classifier System for Finding Control Knowledge of Multi- Input System. *Journal of Studies in Fuzziness and Soft Computing*, 1996, 8: 489- 502
- Hasegawa T, Kamei D, Furuhashi T, Uchikawa Y. A Study on Autonomous Understanding of Linguistic Instructions and Behavior Improvement Using Fuzzy Classifier System. *12th Fuzzy System Symposium in Japan*, 1996. 459- 462
- Fujii M, Hasegawa T, Furuhashi T, Uchikawa Y. A Study on Understanding of Linguistic Instructions and Behavior Improvement. *6th Intelligent System Symposium in Japan*, 1996. 185- 188
- 周志坚, 毛宗源. 智能控制中的模糊控制. 神经网络控制计遗传算法. *广东自动化与信息工程*, 1998, 19(1)
- Rober Fuller. *Neural Fuzzy Systems*. Abo Akademi University
- Watkins C. Learning from delayed rewards, PhD Thesis, University of Cambridge, England
- Richard S Sutton, Andrew G Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, 1998
- Sutton R S, Barto A, William s R. Reinforcement learning is direct adaptive optimal control. *Proc of ACC*, Boston, June 1991
- 印旻, Java. 面向对象程序设计教程. 高等教育出版社, 1999
- Cay S. Horstmann and Gary Cornell. *Java 2 核心技术* 北京工作室译, 机械工业出版社, 2000
- Nuance Communications, Inc. *Introduction to Nuance System* 1999
- Nuance Communications, Inc. *SpeechObjects Developer's Guide*, 1999

## 作者简介:

聂仙丽 (1977-), 女, 硕士研究生. 研究领域: 机器人控制与智能控制.

蒋平 (1963-), 男, 教授, 博士生导师. 研究领域: 机器人控制与智能控制.

陈辉堂 (1933-), 男, 教授, 博士生导师. 研究领域: 机器人控制与智能控制.