

## ◎ 博士论坛 ◎

# 基于评价选择策略的 Internet 多机器人协作控制研究

李响<sup>1,2</sup>, 王越超<sup>1</sup>LI Xiang<sup>1,2</sup>, WANG Yue-chao<sup>1</sup>

1. 中国科学院 沈阳自动化研究所 机器人学重点实验室, 沈阳 100016

2. 中国科学院 研究生院, 北京 100049

1. Robotics Laboratory of Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 100016, China

2. Graduate School of Chinese Academy of Sciences, Beijing 100049, China

E-mail: xiangli@sina.cn

LI Xiang, WANG Yue-chao. Coordination control of Internet-based multi-robot based on evaluation selective strategy. *Computer Engineering and Applications*, 2007, 43(36): 1-3.

**Abstract:** In this paper, we construct a neural group network controller for resolving the problems that exist in the Internet-based multi-robot system, such as: delayed instruction, low efficiency and poor collaboration ability, etc. During the learning process, multi-type neuron coming from the different function regions forms the dynamic neural group through connection. This structure may describe the complex mapping relationship among the external conditions, internal states and behaviors of robot. Through the evaluation to the internal connection weights, it can achieve the best behavior coordination of the multi-robot. We adopt the Internet-based soccer robot system as the experimental platform, and describe the learning algorithm. Simulation result shows that it successfully realizes the shooting mission by one team robots. The proposed model and method improve the coordination ability of multi-robot, and fulfill the stability and real-time requirement of system.

**Key words:** Internet; multi-robot; coordination; neural group network; evaluation

**摘要:** 针对 Internet 多机器人系统中存在的操作指令延迟、工作效率低、协作能力差等问题, 提出了多机器人神经元群网络控制模型。在学习过程中, 来自不同功能区域的多类型神经元连接形成动态神经元群集, 来描述各机器人的运动行为与外部条件、内部状态之间复杂的映射关系, 通过对内部权值连接的评价选择, 以实现最佳的多机器人运动行为协调。以互联网足球机器人系统为实验平台, 给出了学习算法描述。仿真结果表明, 己方机器人成功实现了配合射门的任务要求, 所提模型和方法提高了多机器人的协作能力, 并满足系统稳定性和实时性要求。

**关键词:** Internet; 多机器人; 协作; 神经元群网络; 评价

**文章编号:** 1002-8331(2007)36-0001-03 **文献标识码:** A **中图分类号:** TP18

近年来, 多机器人系统的控制和规划已得到了广泛的关注<sup>[1]</sup>, 有着十分广阔的应用前景。随着网络技术的成熟与发展, 使得通过 Internet 实现多机器人控制成为可能<sup>[2]</sup>。由于 Internet 的传输特性所带来的随机时延的存在, 令机器人无法实时获得准确的操作命令, 操作者也不能得到及时的信息反馈, 造成了系统的工作效率低, 多机器人之间的协作能力差, 甚至会影响到系统的稳定性和机器人自身的安全。因此, 需要建立一个合适的控制模型来解决这些问题。

本文借助神经元群网络建模思想<sup>[3,4]</sup>, 认为来自各功能区域的神经元之间的时空连接结构, 能够建立起外界环境信息与机器人的运动行为之间的复杂对应关系; 通过评价规则来实时选择最适合当前状态的机器人运动行为, 从而实现 Internet 环境下多机器人运动行为的协调与合作。利用该方法设计的多机器

人控制器, 能够保证 Internet 多机器人遥操作系统不受网络随机时延的影响, 灵活、安全地实现协调作业。

## 1 神经元群网络

神经元群网络从模拟人脑的功能和结构角度出发, 整个网络按照需求划分为多个功能区域, 每个功能区域包含若干数量具有相同结构和功能参数的神经元单位, 形成神经元群。各功能区域间存在着从生物学角度定义的连接方式, 区域内部的神经元单位之间也可能存在着横向连接, 如图 1 所示。通过模拟脑神经系统依靠不同功能神经元群集之间的兴奋—协调与抑制—竞争的作用来进行信息处理的动力学原理, 指导网络的学习和工作<sup>[5]</sup>。

**基金项目:** 国家自然科学基金(the National Natural Science Foundation of China under Grant No.60334010)。

**作者简介:** 李响(1980-), 女, 博士生, 主要研究方向为网络多机器人系统, 智能控制方法; 王越超(1960-), 男, 博士生导师, 研究员, 主要研究方向为机器人控制。

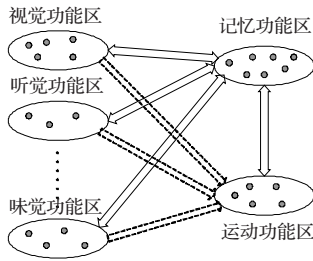


图1 神经元网络示意图

## 2 神经元网络控制器的设计

### 2.1 控制器结构

如图2所示为神经元网络控制器结构示意图。采用正—逆向联合建模法,  $d$  为遥操作机器人的期望输出(延迟的操作命令);  $u$  为遥操作机器人的控制输入;  $y$  为遥操作机器人的实际输出;  $y^*$  是神经元网络给出的一个近似期望输出; 误差  $e$  由两部分组成: 即期望输出  $d$  与实际输出  $y$  之间的误差  $e_1$ , 期望输出  $d$  与近似输出  $y^*$  之间的误差  $e_2$ 。神经元网络控制器的训练是由期望输出  $d$  和误差  $e_1, e_2$  联合来完成的。这样, 神经元网络即可学习到受控对象的正逆向动力学特性。随着学习的进行, 误差  $e_2$  将逐渐减小, 或保持在合适的范围内。当由于网络随机时延的影响, 无法实时接收到操作者的控制命令时, 神经元网络控制器将取代外部控制, 成为主控制器, 对遥操作机器人进行控制, 从而提供了处理随机时延的控制方法。提高了遥操作系统的操作性能, 保证了系统的稳定性和安全性。

延迟的操作者控

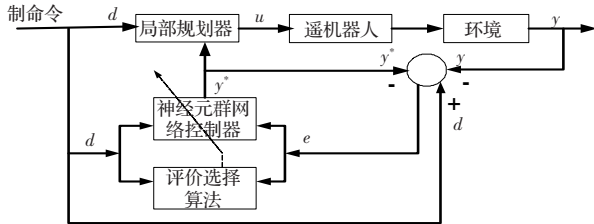


图2 神经元网络控制器结构示意图

### 2.2 学习规则

由  $r$  个输入和  $s$  个输出构成的两个神经元之间的连接模型如图3所示。基本思想是, 在计算机仿真的每一周期(迭代步)  $t$  内, 通过调节网络权值近似于输入或输出来训练神经元节点响应特定的输入或输出, 实现学习和联想记忆功能。

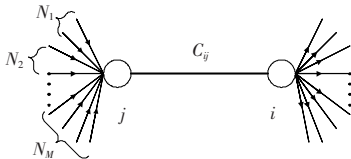


图3 两神经元连接模型

对于所有的神经元群单位, 单位  $i$  的总输入由公式(1)给出; 单位  $i$  的活性值由公式(2)给出; 权值变化规律定义为公式(3); 公式(4)是一个分段线性函数。  $M$  是生物学定义的不同连接类型;  $N_i$  是投射到单位  $i$  的每一类型  $M$  的连接数;  $\omega$  确定单位活性的余留时间;  $\eta$  是学习率;  $\sigma_i$  是特定激活阈值;  $g_i$  是比例因子;  $\varepsilon$  是衰减常数  $<1$ ;  $k_1, k_2$  是两个斜率;  $0 < \theta_1 < \theta_2 < 1$  是两个阈值;  $\overline{v}(t)$  是评价区域的瞬时平均活性;  $c_{ij}(0)$  是  $t=0$  时的初始连接权重。以上参数的初始值不是随机选取的, 是由生物学特性明确规定的<sup>[6]</sup>。  $\Delta c_{ij}(t)$  是逐渐减小的, 最终达到  $c_{ij}(t) = s_j(t)$  或

$c_{ij}(t) = s_j(t)$ , 通过自学习, 达到了应用该模型来响应输入或响应输出的目的。

从公式(2)、(3)中可以看出, 由神经元  $j$  到神经元  $i$  的连接权值不仅取决于神经元  $j$  的输入, 还与神经元  $i, j$  自身的活性值有关; 并且神经元单位活性值也能反映出连接权值的变化情况。这是因为在神经元群网络理论中, 认为来自不同功能区域的神经元具有不同的特性参数值, 从真正生物模拟的角度考虑不应将其忽略。

$$A_i(t) = \sum_{j=1}^M \sum_{j=1}^{N_j} c_{ij}(t) s_j(t) \tag{1}$$

$$s_i(t+1) = \varphi(\tanh(g[A_i(t) + \omega s_i(t)]))$$

$$\varphi(s) = \begin{cases} 0 & s < \sigma_i \\ s & \text{or} \end{cases} \tag{2}$$

$$\Delta c_{ij}(t) = \varepsilon(c_{ij}(0) - c_{ij}(t)) + \eta s_j(t) F(s_i(t)) \overline{v}(t)$$

$$F(x) = \begin{cases} 0 & x < \theta_1 \\ k_1(\theta_1 - x) & \theta_1 \leq x < (\theta_1 + \theta_2)/2 \\ k_1(x - \theta_2) & (\theta_1 + \theta_2)/2 \leq x < \theta_2 \\ \frac{k_2 \tanh(\rho(x - \theta_2))}{\rho} & \text{or} \end{cases} \tag{4}$$

在每一周期(迭代步)  $t$  内, 网络内部总有一些连接权值会表现为增强趋势(反之则为抑制), 并抑制其它权值的响应程度, 那么这一时刻由这些权值所连接在一起的神经元群体就表示对该时刻  $t$  的响应结果。因此, 在记忆评价区域始终存储着当前  $t$  时刻每一神经元单位的活性值, 在与  $t+1$  时刻活性值进行比较, 获得变化趋势后, 则更新为  $t+1$  时刻活性值, 评价过程正是依据变化趋势和瞬时活性平均值  $\overline{v}(t)$  进行的。

## 3 基于 Internet 遥操作多机器人系统建模

### 3.1 多机器人任务描述

多机器人协作行为是指多机器人系统在已知或未知环境中, 通过各机器人运动行为的协调与综合, 实现期望的任务目标。本文研究的任务描述如下: 通过己方 3 个足球机器人的协作, 实现射门任务。通过悬挂在场地上方的摄像机获得机器人和目标(球)的分布情况, 机器人采用什么样的行为策略是未知的。总体目标是最快速度地选择最有效的运动行为策略, 将球射入球门区域。

### 3.2 构造模型

图4为用于实现多机器人控制的神经元网络模型, 采用共用一个视觉区域和记忆评价区域的局部并行(虚线框部分)网络结构。为适合多机器人的应用, 网络的训练和信息处理是并行的<sup>[5]</sup>。如果整个输入样本可划分为  $n$  个样本集, 则定义  $n$  个并行网络, 由  $n$  个样本集轮流对其训练, 训练结果经过评价记忆区域进行选择, 得到最适合当前系统状态的输出。

本文中, 视觉部分将识别到的带有不同色标的 3 个足球机器人小车作为 3 个输入样本, 因此局部并行网络结构如图5所示。其定义的输入矢量为  $[R_x, R_y, d, \theta]$ : 分别为机器人的位置信息  $[R_x, R_y]$ ; 与球的距离  $d$ ; 小车正方向与球—车连线的夹角  $\theta$ , 用来实时计算机器人的线速度  $v$  和角速度  $\omega$ 。因此, 视觉区域包含 3 个神经元单位, 用于实现 3 个机器人样本类的划分; 运动规划区域包含  $3 \times 4$  个神经元单位, 用于处理输入矢量; 运动控制区域包含 3 个神经元单位, 用于实现转动、定点和定向 3 个基本动作; 记忆评价区域包含 18 个神经元单位, 用于实现活性值的存储和评价分析, 各区域神经元之间实现局部互连。

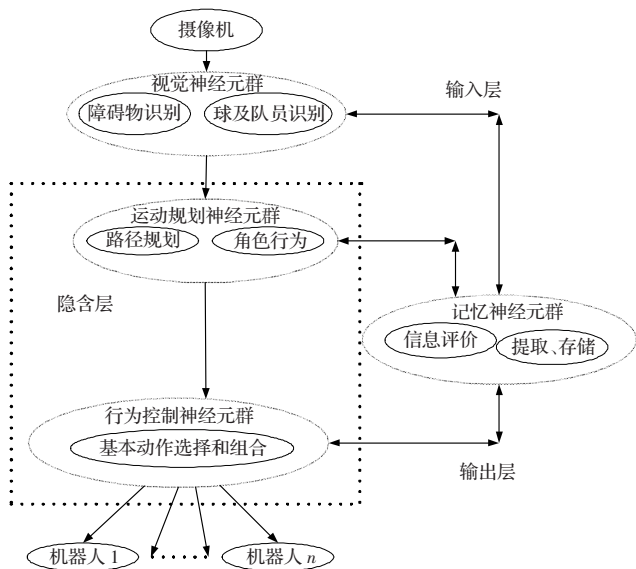


图4 带有共享区的多机器人神经网络结构示意图

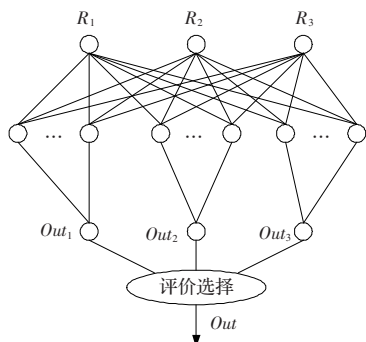


图5 局部并行网络结构

### 3.3 学习算法

每个机器人可以通过视觉区域获得自身和球的位姿信息,并通过共享的记忆区域获得自身和其它机器人的评价信息,以进行选择。具体的学习算法如下:

**步骤1** 周期  $t$  时刻,初始化  $\omega, \sigma_i, g_i$  和  $c_{ij}(0)$  的参数值,计算初始活性值和连接初值。

**步骤2** 根据视觉区域识别的机器人数量对输入样本集进行划分;

**步骤3** 任选一组输入样本提供给运动规划区域(即:局部并行网络),计算  $t$  时刻的  $A_i(t), S_i(t), \Delta c_{ij}(t)$  值。

**步骤4** 根据评价信息,获得由具有增强趋势的连接权值所

确定的一组来自不同区域的神经元单位,产生选择性运动行为输出。

**步骤5** 选取下一组输入样本返回步骤3,直到所有输入样本全部提供给网络。

**步骤6** 令  $t=t+1$ ,返回步骤2,直至结束。

### 4 仿真实验及分析

图6给出了仿真实验的初始界面图,3个机器人从初始位置开始运动直到将球送入球门区域,各功能区域的特性参数和初始连接值、学习率等由表1和表2给出。

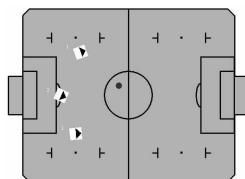


图6 仿真实验初始界面

表1 不同区域的功能参数

功能区域	$g$	$\sigma$	$\omega$
视觉	1.356	0.282	0.236
运动规划	2.000	0.100	0.300
行为控制	2.000	0.050	0.150
记忆评价	1.920	0.112	0.206

表2 不同区域间的连接初值及参数值

连接(区域)	$c_{ij}(0)$	$\eta$	$\varepsilon$	$\theta_1$	$\theta_2$	$k_1$	$k_2$
视觉→运动	0.120 0	0.020 0	0.002 0	0.01	0.16	0.10	0.16
运动→行为	0.022 0	0.007 6	0.000 6	0.14	0.16	0.10	0.16
视觉↔记忆	0.377 5	0.035 0	0.000 5	0.14	0.31	0.45	0.10
运动↔记忆	0.001 6	0.050 0	0.000 5	0.01	0.29	0.45	0.15
行为↔记忆	0.070 0	0.004 0	0.000 4	0.10	0.50	0.45	0.15

图7是整个执行过程的3次截图,3个机器人通过实时获得的  $d$  和  $\theta$  信息计算各自的  $v, \omega$  值,通过评价选择,(a)时刻1号机器人移动到球处,将球踢出,2、3号机器人进行协作;(b)时刻1、2号机器人配合3号机器人完成踢球动作;(c)时刻1号机器人将球送入球门区域,3个机器人通过协调合作最终成功地完成了射门任务。图8是对应时刻的神经元单位活性值分布情况。

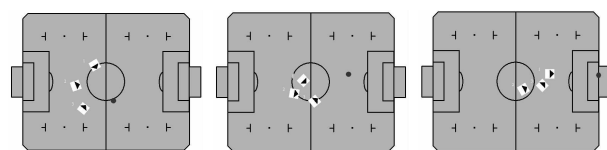


图7 3个机器人合作射门的仿真结果

仿真实验中,神经元群网络的平均更新周期在300ms左右,完成整个射门任务耗时109.4s,能够满足真实赛场的实时

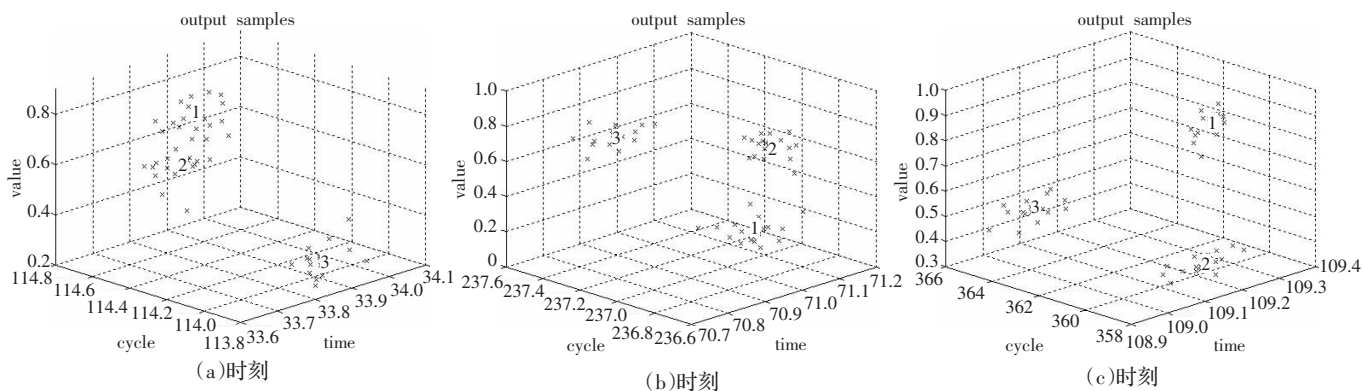


图8 神经元活性值分布情况