

文章编号: 1002-0446(2000)05-0384-06

小波神经网络在两足步行机器人爬斜坡中的应用*

张 克 傅佩琛 强文义

(哈尔滨工业大学机器人研究所 105 室 150001)

摘 要: 针对传统的神经网络中神经元模型在结构和信息存储能力上存在的不足, 本文提出了一种基于广义小波基函数网络的神经元集聚模型. 这种小波神经网络不仅收敛速度快, 非线性逼近能力更好, 而且具有内部结构变尺度、自适应调整和广义信息存储等智能化特点, 更符合生物原型的实际情况. 静态学习和准动态学习仿真实验证明这种神经网络结构的有效性.

关键词: 两足步行机器人; 小波神经网络; 无导师强化学习

中图分类号: TP24 文献标识码: B

1 引言

近年来, 很多学者就两足步行机器人在不同的环境中行走进行了研究. 大部分工作都是利用传感器来探测环境的变化, 从而修正步态规划, 处理过程比较繁琐. 本文根据对神经元机空间分布结构和状态转换的研究, 基于网络嵌套思想, 利用小波良好的局域化和变尺度的特性, 提出了一种以广义小波基函数网络为特征作用函数的神经元集聚模型, 从而设计了具有自适应调整和广义信息存储等智能化特征的小波神经网络结构. 当两足步行机器人在斜坡行走时, 通过对安装在脚底的力传感器反馈力的测量, 神经网络机构学习后逐渐并自动的产生新的步态. 由于力传感器作为两足步行机器人稳定与否的指示器, 所以不需要计算坡度和重新规划步态. 采用无导师强化学习算法训练神经网络, 这种学习方法的优越性是: 算法不需要外在的反馈信号、在线学习和算法简单. 另外, 传感器的噪声特性在学习过程中也被考虑进去^[1].

2 HIT-III 型两足步行机器人简介

本文的研究对象是 HIT-III 型两足步行机器人. 该机器人是高 1.1m、重 38kg、电气驱动、10 个自由度仿人型两足步行机. 其坐标定义如图 1 所示, 自由度分配如图 2. 由于脚的质量比小腿的质量小的多, 且其运动的速度较低, 在运动学分析中将它与小腿作为一体来考虑. 因为两个腰关节的分布分别在一个机座上(见图 2), 只有它们同时转动并且转动相同的数量值时, 机器人才能在腿部位置不改变的情况下实现躯体转动, 所以躯体的运动和腿部运动可以解耦. 这样两足步行机器人就简化为一个 5 连杆机构. 模型及杆件参数如图 3 所示.

3 小波神经网络控制器

3.1 控制器的组成和功能

如果步行机器人在没有姿态调整的情况下, 从平面走向斜坡或其它非平地行走环境时, 重心轨迹将会与平面行走时的重心曲线相偏离, 这意味着步行机器人将出现行走姿态不稳定. 这

* 收稿日期: 1999-08-01

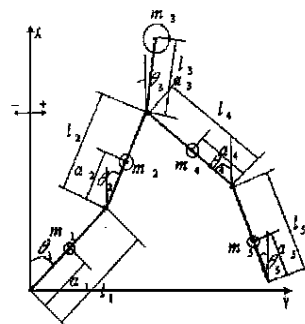
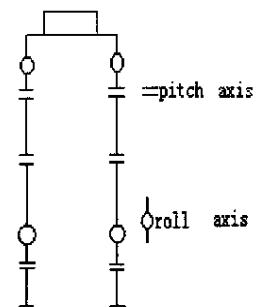
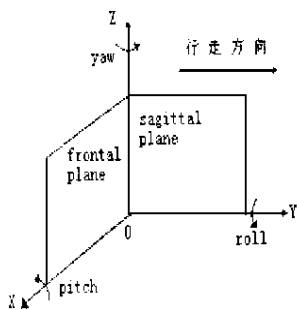


图 1 机器人的运动坐标

图 2 HIT-III 的自由度分配

图 3 两足步行机器人简化模型

时就需要利用神经网络控制器, 通过足底的压力传感器的输入认识到行走环境的这一变化, 并开始调节行走步态, 最终得到一族新的稳定行走步态, 从而机器人实现非平地环境的稳定行走. 本文设计的神经网络控制器如图 4 所示.

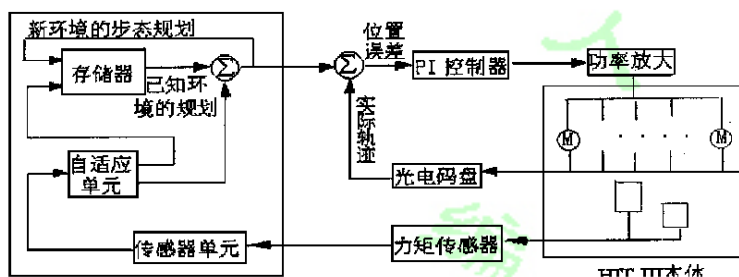


图 4 HIT-III 型两足步行机器人控制系统结构

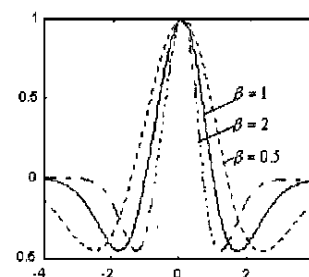


图 5 小波函数 $\psi(x)$ 的图像 ($\beta=0.5, 1, 2$)

3.2 基于广义小波基函数的神经元集聚模型

神经网络控制器的主要组成部分是自适应单元, 它负责修正关节的运动轨迹. 自适应单元是由很多具有刺激和抑制突触连接的神经元组成. 传统的神经网络模型大多是建立在 Hebb 所提出的神经网络信息存储方式的基础上, 而现代神经生理学的研究^[2]表明: 1. 人脑的神经元, 在处理某一具体问题, 有相当一部分处于休眠状态, 而且休眠和工作状态可以发生转移. 同时, 处于空间不同区域的神经元分工应有所不同, 表现为神经元的机能空间非均匀分布结构; 2. 神经元对外界刺激的反应同时也具有宏观意义上的集聚特性, 即同一机能区域内处于工作状态的部分会通过空间整合, 从整体的层面上对外界刺激做出某种反应, 这种分布结构和机能都表现出集聚特性的神经元全体, 称之为一个神经元的集聚 (Neurons Assemblage). 我们采用一广义的小波基函数网络来构造神经元的集聚, 其母波函数选择墨西哥帽状小波

$$\psi(x) = (1 - \beta x^2)e^{-\frac{\beta x^2}{2}}, \quad \beta \text{ 为形状参数} \quad (1)$$

从图 5 中看, 该小波具有良好的局域特性. 从而以一族这样的小波函数为基础构造神经元的集聚来逼近非线性映射 $y = f(x)$ 时, 对每一对样本对 (x_k, y_k) , 只需调整集聚内相应位置的小波函数而对其它函数没有影响. 这与生物神经元的集聚在神经元生理活动中总是部分神经元工

作, 而另一部分则处于休眠状态的情形相类似. 采用这样的结构不仅大大简化网络学习的复杂程度, 而且使集聚内的权值的调整变成一个简单的线性调整问题.

该神经元集聚的数学模型为

$$f(x) = \sum_{k=0}^L c_{j,k} \psi_k(x) + U[2^j(x+1) - b_0(L+1)] + c_{L+1} \quad (2)$$

式中小波基函数为式(1)的墨西哥帽状小波.

为了方便计算机实现, 采用倍频程规则离散化, 则相应二进小波基函数为

$$\psi_{jk} = 2^{j/2} \psi(2^j x - kb_0) \quad (3)$$

式中 b_0 为抽样速率参数; 函数 $U[2^j(x+1) - b_0(L+1)]$ 保证整个 $f(x)$ 是一个广义的 Sigmoidal 函数; c_{L+1} 为偏置项^[3]. 图 6 为这种神经元集聚模型(WNNAM)的结构示意图. 与传统的前向网络模型相比, 每个 WNNAM 在集聚层上多存储了 $L+2$ 个信息, 它与集聚间权值中的信息分布存储方式一道构成了新模型的广义信息存储特性. 此外, 利用小波所特有的变尺度特性, 还可以对 WNNAM 的规模进行自适应调整, 从而达到在给定逼近精度的前提下获得最优的网络结构^[4].

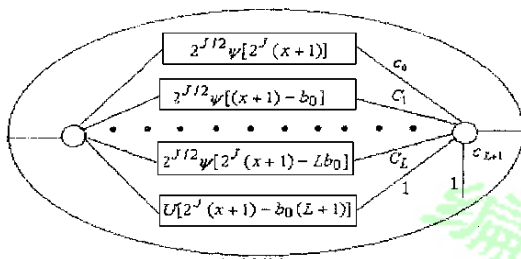


图 6 神经元集聚模型结构示意图

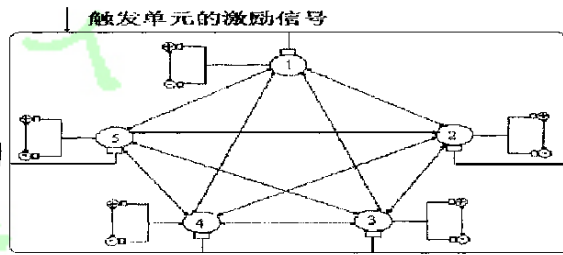


图 7 神经网络的刺激-抑制模型

3.3 神经网络结构模型

两足步行机器人在行走过程中, 前向平面的运动对侧向平面的运动可以解耦. 而且前向平面的关节运动在平面行走和斜坡行走是一样的, 因此神经网络的自适应机构主要调节侧向平面的关节运动. 由图 3 的简化模型, 本文设计图 7 所示的刺激-抑制相结合的神经网络模型.

图中的圆圈代表神经元, 白色的方块代表刺激性连接, 黑色的箭头代表抑制性连接. 标有数字号码的神经元叫关节神经元, 对每一个关节 j , 相应的关节神经元在起始点 pp 处都有一个神经元值 $w(j, pp)$ ($j = 1, 2, \dots, 5; pp = 0, 1, \dots, 7$). 每个关节神经元节点连着两对小神经元称为“方向神经元”(图中只画出爬坡的一对). 方向神经元在起始点 pp 处的神经元值 $v(ud, jn, h, pp)$ 为, 这里, $ud = 0, 1$, 代表机器人爬坡($ud = 1$)或下坡($ud = 0$); jn 代表关节数; $h = 0, 1$ 代表关节的调整方向, 0 表示负方向, 而 1 表示正方向, 分别相对应于图中的“-”、“+”号.

神经网络的工作原理如下: 假设两足步行机器人在爬一斜坡, 在每一个行走阶段末, 传感器单元读取来自脚底的力传感器的信号, 并计算重心的位置, 若结果表明此时的重心位置与期望的重心位置相比使机器人不稳定, 传感器单元送出激励信号使自适应单元中 5 个神经元中的一个被激活, 其它的处于抑制状态, 这种选择基于每个神经元的神经元值(neuron value). 这

里假设 3 号关节神经元首先被激活, 它将向连在它上面的两个方向神经元(爬坡)发出刺激信号. 假设带“+”的神经元首先被激活, 则带“-”的神经元被抑制, 结果关节 3 的角度增加. 这一过程一直进行, 直到没有刺激信号从传感器单元传来. 这表明机器人的行走重新达到稳定状态. 此时, 自适应单元不再活跃. 下一阶段关节的运动轨迹将从存储器中已经训练好的步态中获得.

神经网络采用强化学习方法, 在学习过程中有效的选择不断被强化. 这种强化学习方法的原理已经被 *Barto* 等学者所证明, 并广泛应用^[5].

3.4 无导师强化学习算法

我们利用脚尖和脚跟的压力差作为强化信号, 也就是

$$\Delta f = f_{heel} - f_{toe} \quad (4)$$

这里 f_{heel} 代表足跟部分感受到的压力; f_{toe} 代表足尖部分感受到的压力. 训练算法如下:

(1) 确定当前起始点的数值 pp , 这里 $pp = 0, 1, \dots, 7$.

(2) 初始化: 如果不考虑预先学习的影响, 设所有的神经元值为零(第一遍学习).

(3) 如果 $|\Delta f - \Delta f_{des}| < \epsilon$, 则不需要进行学习, 反之, 学习过程将按如下步骤进行. 其中 Δf_{des} 是理想的压力差, 是一个小的正常数, 可以通过与传感器的噪声信号一致来确定.

(4) 设参数 ud :

$$\begin{cases} ud = 0 & \text{if } \Delta f > 0 \\ ud = 1 & \text{if } \Delta f \leq 0 \end{cases} \quad (5)$$

(5) 定义 $yw(j)$ 如下:

$$yw(j) = w(j, pp) + \text{noise}, \quad j = 1, 2, \dots, 5 \quad (6)$$

这里噪声是人为加入的随机干扰, 其取值范围在 $[-1, 1]$. 两个神经元值 $w(j, pp)$ 相近, 神经元值较小可能对维持机器人平衡更有利, 由于附加的随机噪声值可能被选中.

(6) 使具有最大 $yw(j)$ 的相应关节的关节数 $jn = j$.

(7) 定义 $yv(h)$ 如下:

$$yv(h) = v(ud, jn, h, pp) + \text{noise}, \quad h = 0, 1 \quad (7)$$

(8) 使具有最大 $yv(h)$ 的方向神经元的方向 $dir = h$.

(9) 根据上面的两个结果定义 $\Delta q(jn)$, 例如: 如果 $jn = 3, dir = 1$, 那么 $\Delta q(3) = +dq^\circ$, 其中 d_q 使一个小的正数.

(10) 计算新的压力差和强化信号 z :

$$\Delta f_{old} = \Delta f_{new} \quad (8)$$

$$\Delta f_{new} = \Delta f_{des} - \Delta f \quad (9)$$

$$z = \Delta f_{old} - \Delta f_{new} \quad (10)$$

(11) 用下列方程对神经元进行训练:

$$v(ud, jn, dir, pp) = v(ud, jn, dir, pp) + c_1 \text{ sigmoid}(s_1 \Delta f_{new}) z \quad (11)$$

$$w(jn, pp) = w(jn, pp) + c_2 \text{ sigmoid}(s_2 \Delta f_{new}) z \quad (12)$$

其中, c_1, c_2 是正的常数作为学习速率, s_1, s_2 是饱和度的标志, 如果响应的神经元的神经元值比预先定义的极限值 s_1, s_2 小, 这个标志为 1, 否则为 0.

(12) 返回第三步, 重复执行以上步骤.

上面的训练过程一直进行, 直到 $|\Delta f - \Delta f_{des}| < \epsilon$ 时自动停止. 整个训练过程我们称之为—

个训练周期. 此时 Δf 的值将比较接近期望值, 同时所有的神经元都已获得一个新值而不是原来的零值. 当受过训练的神经网络用于保持机器人在斜坡行走的稳定时(第二次学习), 就能比没有受过训练的神经网络用少得多的时间使机器人重新达到稳定状态.

4 仿真试验

4.1 静态学习

在静态学习中, 神经网络对一个单独的起始点进行训练, 直到对这个起始点机器人达到期望的稳定性. 当两足步行机器人行走在 10° 的斜坡上, 利用上节的算法对神经网络进行训练. 在实际应用中选择的母波为墨西哥帽函数

$$\psi(x) = \frac{2}{3} \pi^{\frac{1}{4}} (1 - x^2) e^{-\frac{x^2}{2}}$$

可以计算出 $w_0 = 0.752$, $x_0 = 0$, 取伸缩平移步长 $a = 2$, $b = 1$, $\epsilon_1 = 0.01$, 其它的训练参数为: $c_1 = 0.025$, $c_2 = 0.008$, $d_q = 0.015$, $\epsilon = 10$, (此处 ϵ 的单位是 0.1 牛顿), 饱和度标志的极限为: $S_1 = 10$, $S_2 = 10$.

仿真曲线如下: 图 8(a-e) 展示了修正后的各关节角轨迹, 其中虚线表示修正后的各关节角轨迹, 而实线表示了未修正的各关节角轨迹, 横坐标为时间轴, 纵坐标为角度; 图 9 展示了神经网络训练过程中机器人前向重心轨迹, 其中横坐标表示时间轴, 纵坐标表示相对于初始点前脚掌的距离, 单位为 m. 数字表示学习周期数. 可见经过 13 个周期的学习机器人前向重心轨迹上移直至与理想的重心轨迹重合, 此时强化学习停止.

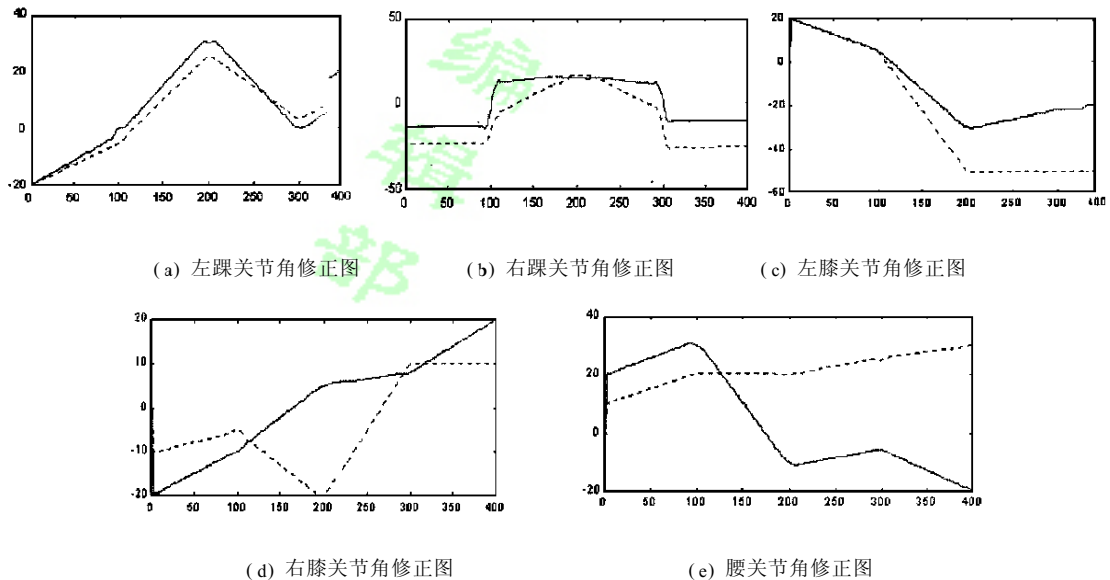


图 8 修正后的各关节角轨迹

4.2 准动态学习

在准动态学习中, 神经元训练在机器人行走过程中. 当神经网络选择一个新的神经元时, 由于这个神经元对于当前阶段没有经过训练, 因此可能导致关节运动的不连续. 为了避免这种不连续性, 我们使学习过程只在这一阶段的前半部分进行, 在训练中累积的运动的调整对当前的起始点执行. 而在这一阶段的后半部分, 通过神经网络机构使累积值线性的衰减到零, 从而

使机器人能平滑的到达下一个起始点. 准动态学习中, 我们选择的训练参数如下: $c_1 = 0.025$, $c_2 = 0.008$, $d_q = 0.02$, $S_1 = 10$, $S_2 = 10$, $\epsilon = 10$ (ϵ 的单位同上).

仿真曲线如图 10: 其中曲线 1 是两足机器人第一遍行走在平地行走时的重心投影; 曲线 2 是步行机器人在第二遍行走时, 不经调整步态在 10° 斜面行走时的重心轨迹; 在第三遍行走时控制系统启动自适应单元, 经过训练的进行, 重心的轨迹越来越接近于期望的轨迹(曲线 1). 曲线 3 和 4 的分别代表经过 5 次和 10 次训练后的结果.

5 结论

试验结果表明, 这种小波神经网络控制器具有较快的收敛性、网络逼近能力强、自适应调整网络结构等特性. 并能让两足步行机器人在一不确定的斜坡上自适应地行走.

参 考 文 献

- 1 Salatian A W, Keon young yi Yuan F Zheng. Reinforcement Learning for a Biped Robot to Climb Sloping Surfaces. Journal of Robotic Systems, 1997, 14(4): 283- 296
- 2 Sejnowski T, Korch C, Churchland P. Computational Neuroscience. Science, 1998, 241: 1299- 1306
- 3 Qinghua Zhang. Using Wavelet Network in Nonparametric Estimation. IEEE Trans on Neural Networks. 1997, 8(2)
- 4 张山, 周凤岐. 一种基于广义子波基函数选择的智能化神经元集聚模型. 西北工业大学学报, 1998, 16(4): 495- 500
- 5 Gullapalli V, Franklin J A, Benbrahim H. Acquiring Robot Skills Via Reinforcement Learning. IEEE Control Syst Mag, 1994. 13- 24

AN APPLICATION OF WAVELET NEURAL NETWORKS IN THE BIPED ROBOT'S SLOPE CLIMBING

ZHANG Ke FU Peichen QIANG Wen-yi
(Robot Research Institute, Harbin Institute of Technology 150001)

Abstract We found that neuron model is inadequate owing to its defects such as those inherent in its structure and in its capability of information storage. So we propose an intelligent neurons assemblage model with generalized wavelet basis function network as its excited function. Not only the wavelet neural networks' convergence rate is much faster and its nonlinear approach capability is much better but also its intelligent characteristics, such as the variable-scale adaptive adjustment of structure and the generalized information storage, make it reflect much more faithfully the biological original. Static learning and pseudo dynamic learning are demonstrated to prove that the proposed mechanism is valid.

Keywords: Biped robot, wavelet neural networks, unsupervised reinforcement

作者简介:

张 克 (1974-), 男, 博士研究生. 研究领域: 智能控制, 计算机控制, 机器人控制.

傅佩琛 (1934-), 男, 教授. 研究领域: 两足步行机器人控制, 智能控制.

强文义 (1938-), 男, 教授, 博士生导师. 研究领域: 智能控制, 过程控制, 机器人控制.

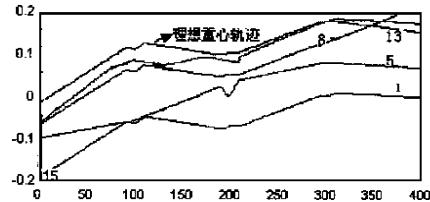


图 9 训练过程中重心轨迹调整

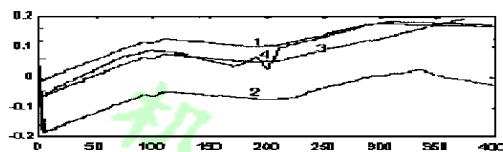


图 10 机器人在不同状态下的重心投影轨迹