

基于内在动机的智能机器人自主发育算法

任红格*, 向迎帆, 李福进

(华北理工大学 电气工程学院, 河北 唐山 063009)

(* 通信作者电子邮箱 414963045@qq.com)

摘要:针对两轮自平衡机器人在学习过程中主动性差的问题,受心理学内在动机理论启发,提出一种基于内在动机的智能机器人自主发育算法。该算法在强化学习的理论框架中,引入模拟人类好奇心的内在动机理论作为内部驱动力,与外部奖赏信号一起作用于整个学习过程。采用双层内部回归神经网络存储知识的学习与积累,使机器人逐步学会自主平衡技能。最后针对测量噪声污染对机器人平衡控制中两轮角速度的影响,进一步采用卡尔曼滤波方法进行补偿,以提高算法收敛速度,降低系统误差。仿真实验表明,该算法能够使两轮机器人通过与环境的交互获得认知,成功地学会运动平衡控制技能。

关键词:内在动机;自主发育;卡尔曼滤波;平衡控制;两轮机器人

中图分类号: TP183 **文献标志码:** A

Autonomous developmental algorithm for intelligent robot based on intrinsic motivation

REN Hongge*, XIANG Yingfan, LI Fujin

(College of Electrical Engineering, North China University of Science and Technology, Tangshan Hebei 063009, China)

Abstract: The initiative of two-wheeled self-balancing robot in the process of learning is poor. Inspired by intrinsic motivation theory of psychology, an autonomous development algorithm for intelligent robot based on intrinsic motivation was put forward. In the frame work of the reinforcement learning theory, the algorithm introduced human curiosity of intrinsic motivation theory as the internal driving force, and external reward signal into entire learning progress, and adopted double internal regression neural network for storage of knowledge learning and accumulation, which made robot gradually learn autonomous balance skill. Finally, aiming at the effects of measurement noise pollution on two-wheeled angular velocity of robot, further by adopting the method of Kalman filter to compensate, to speed up the algorithm convergence, and reduce the system error. Simulation experiments show that this algorithm can make the two-wheeled robot obtain cognition through interaction with the environment, therefore successfully learn balance control skill.

Key words: intrinsic motivation; independent development; Kalman filtering; balance control; two-wheeled robot

0 引言

在 20 世纪 80 年代末有学者提出了两轮自平衡机器人的概念,它属于智能发育机器人领域中一个全新的研究方向。于化龙等^[1]认为发育机器人与传统机器人相比,前者心智发育是建立在一个类似大脑的自然系统或人工嵌入式系统之上的,这种系统在其内在发育程序的控制下通过使用自己的传感器和执行器与环境进行自主实时的交流来发展心智能力。针对两轮自平衡机器人的平衡控制,人们提出了很多种控制方法:文献[2]结合 Q 学习和前馈(Back Propagation, BP)神经网络实现了对一级倒立摆的平衡控制;文献[3]通过采用经典线性二次型调节器(Linear Quadratic Regulator, LQR)算法对二级倒立摆进行平衡控制;文献[4]采用了经典强化学习(Reinforcement Learning, RL)与模糊神经网络相结合,实现了两轮机器人的平衡控制;文献[5]通过分层模糊控制来完成平衡控制目的。以上控制算法都是结合神经网络容错力强的特点,但是学习时间较长,自学习能力较差;并且都未考虑

测量噪声对两轮角速度的影响,不利于系统的稳定。

心理学家常常把强化信号区别为内在激励和外在激励。外在激励是从环境中得到的奖赏,而内在激励的产生仅仅是因为智能体自我的喜好而与外界环境因素无关。Oudeyer 等^[6]认为在智能机器人研究中,内在动机作为一种内部驱动力,同外界环境所发出的外部信号一同作用于机器人整个学习过程。2006 年 Pfeifer 等^[7]强调机器人物理本体与未知环境的交互来实现机器人对未知环境的学习和认知。2013 年 Cederborg 等^[8]从生物的自我意识下的探索出发,提出了内在动机(Intrinsic Motivation, IM)系统思想,并提出系统状态转移误差学习机,实现基于内在动机模型的机器人对未知环境的主动探索学习。心理学家 White^[9]曾指出内在激励行为是机体能力提高的重要过程,这种能力的获得一定会对以后可能遇到的各种问题有所帮助,并会为解决新的问题打下基础。

本文受到心理学内在动机理论启发,针对以上问题提出一种基于内在动机的自主发育算法(Reinforcement Learning based on Intrinsic Motivation, IM-RL)。该算法以强化学习为

收稿日期:2015-04-20;修回日期:2015-05-25。

基金项目:国家自然科学基金资助项目(61203343);河北省自然科学基金资助项目(E2014209106)。

作者简介:任红格(1979-),女,河北石家庄人,副教授,博士,主要研究方向:人工智能;向迎帆(1990-),男,河北唐山山人,硕士研究生,主要研究方向:检测技术及智能装置;李福进(1956-),男,河北唐山山人,教授,博士,主要研究方向:检测技术及智能装置。

框架,将内在动机信号作为内部奖励,驱动机器人的学习进程,并采用双层内部回归神经网络作为知识的学习与积累空间。最后根据文献[10]所提出的方法采用卡尔曼滤波方法进行补偿,可以有效地降低噪声影响,提高算法的收敛速度,降低系统误差。

1 基于内在动机的自主发育算法

1.1 自主发育模型结构

本文以强化学习为框架,结合内在动机驱动生物体自主学习的提点,设计了一种基于内在动机的自主发育算法,算法的控制结构如图1所示。

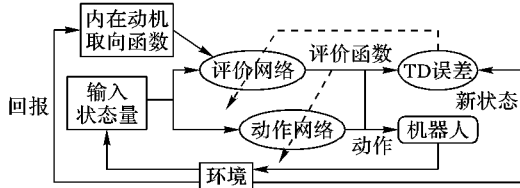


图1 基于内在动机的自主发育算法结构

本文设计的学习算法由评价网络和动作网络两部分组成,并且它们都使用了内部回归神经网络。评价网络采用了时间差分(Temporal Difference, TD)预测误差方法对评价函数进行逼近,并且根据内在动机机制计算出内部奖励,与外部输入信号一起作用于评价网络,为动作网络提供一个更加精准的评价函数来决定当前动作的优劣。动作网络利用评价网络的输出来优化动作选择,从而提高机器人最优动作选择概率,体现了智能体的自主学习能力。

1.2 自主发育算法设计

基于内在动机的自主发育算法可以表示为一个8元组计算模型,记为:

$$IM-RL = \{S, A, f, j, r, p(s_i, a_i), H, V(t)\}$$

其中各个元素的含义如下。

1) S : $IM-RL$ 的内部状态集合。 $S = \{s_i | i = 1, 2, \dots, n\}$, S 为系统所有可能的状态集合, s_i 表示第 i 个离散状态, n 为所有可能产生状态的个数。

2) A : $IM-RL$ 的动作集合, $A = \{a_i | i = 1, 2, \dots, m\}$, a_i 表示第 i 个动作行为, m 为所有动作行为的个数。

3) f : $IM-RL$ 的转移状态函数,一般由环境或者系统模型来决定。

4) J : $IM-RL$ 中由内在动机决定的取向函数。

5) r : $IM-RL$ 中的奖赏信号函数。由于内在动机机制的引入,使得奖赏信号分为两部分:一部分是由内在动机作为内部奖励函数所产生的内在强化信号 r_{im} ;另一部分为环境所产生的外在强化信号 r_{ex} 。所以总的奖赏信号函数定义为 $r = \varphi r_{im} + \varphi r_{ex}$, 其中 φ 和 φ 为参数。

6) $p(s_i, a_i)$: $IM-RL$ 中的动作选择概率,定义为在状态 s_i 下选择动作 a_i 的概率值。

7) H : $IM-RL$ 的信息熵,它作用于整个学习过程,主要是用来衡量 $IM-RL$ 的自主认知程度的。其中 H_i 定义为在状态 s_i 下的信息熵:

$$H = H(S, A) = \sum_{i=0}^n p_i H_i = \sum_{i=0}^n \sum_{j=1}^m p(s_i, a_j) H_i(s_i, a_j) = \sum_{i=0}^n \sum_{j=1}^m p(s_i, a_j) \lg p(s_i, a_j) \quad (1)$$

8) $V(t)$: $IM-RL$ 中的评价函数。

在该算法中,内在奖励函数由内在动机理论代替,假定智能体在一个陌生环境中运行,其输出量为 x_t ,并将期望输出定义为 \tilde{x}_t ,那么这两者的差值 $r_{im} = x_t - \tilde{x}_t$ 定义为系统的内部奖励函数。当系统在 t 时刻选择动作 a 时,状态会从 s_t 转移到 s_{t+1} ,如果 $r_{im}(t+1) - r_{im}(t) < 0$,即系统产生的误差比前一时刻的误差小,说明 $t+1$ 时刻选取的动作使系统达到期望目标状态的效果比 t 时刻选取动作的效果更好,同时也说明系统在 $t+1$ 时刻的取向性大;反之,如果 $r_{im}(t+1) - r_{im}(t) > 0$,则说明系统在 $t+1$ 时刻的取向性小。

在心理学中描述的陌生度、好奇心、厌倦度等都可以被看作是一种驱动力,它们可以使人类或其他生物对未知环境或任务进行自主的探索与学习。以上驱动因素可以归结为生物体去探索和学习过程中由内在动机引发的取向机制。本文将引入学习算法中,使机器人能够模拟人类心理工作机理,使其具有自主学习能力,提高学习主动性和效率性。

而在基于内在动机的自主发育算法中,本方法认为在内在动机机理的驱动下,使得内部动作评价函数 $V(t)$ 逐渐趋近于0,从而能使两轮机器人能保持最合适的平衡姿态。下面来定义评价函数:

$$V(t) = r(t+1) + \gamma r(t+2) + \gamma^2 r(t+3) + \dots \quad (2)$$

其中, $\gamma \in [0, 1]$ 为折扣因子。

在 $t-1$ 时刻的评价函数为:

$$V(t-1) = r(t) + \gamma r(t+1) + \gamma^2 r(t+2) + \dots = r(t) + \gamma [r(t+1) + \gamma r(t+2) + \dots] = r(t) + \gamma V(t) \quad (3)$$

这就证明了在 $t-1$ 时刻的评价函数可以由 t 时刻的评价函数表示。将 $V(t)$ 作为一个预测器来预测 $V(t-1)$ 。之后可以建立 TD 误差差分公式,即

$$\phi = r(t) + \gamma V(t) - V(t-1) \quad (4)$$

在评价网络中将两轮机器人的状态量和控制量作为输入,当 TD 误差趋近于0时,评价网络可以近似逼近式(2)中的评价函数。然后系统将 $V(t-1)$ 的值存储于网络中,这样就避免了再计算 $V(t+1)$ 的值。

最后系统的内在动机取向函数 $J(t)$ 为:

$$J(t) = \frac{1 - e^{-\lambda V(t)}}{1 + e^{-\lambda V(t)}} \quad (5)$$

其中 λ 为取向函数的参数。当 $J(t)$ 的值越小,这就表明对应的动作奖励越少,系统中内在动机的取向性越小;相反,当 $J(t)$ 越大,对应的动作奖励就会越大,表明系统中内在动机的取向性越大。

1.3 内部回归神经网络

该算法中使用文献[11]中所提出的双层内部回归神经网络(Internal Regression Neutral Network, IRN)进行训练。它具有学习速度快、拓扑结构简单、逼近精度高等特点,结构如图2所示。

在 IRN 中,隐含层各个节点的延时信号作为输入信号与输入层的信号一起传送到隐含层,这样每一次都有新的信号加入进来,从而增加了数据训练的强度,同时还提高了系统对动态信息的处理能力,这种网络更加适合复杂系统的稳定控制。

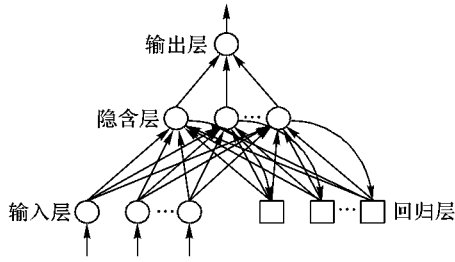


图2 内部回归神经网络结构

IRN 输入输出数学关系描述如下:

$$O_{ij}(t) = \sum_{i=1}^{netI} w_{cij}(t) X_{ci}(t) + \sum_{i=1}^{netH} w_{rij}(t) y_{ci}(t-1) \quad (6)$$

$$y_{ij}(t) = \frac{1 - e^{-O_{ij}(t)}}{1 + e^{-O_{ij}(t)}} \quad (7)$$

$$V(t) = \sum_{j=1}^{netH} w_{oj}(t) y_{oj}(t) \quad (8)$$

其中: $netH$ 和 $netI$ 分别为 IRN 中隐含节点和输入节点的个数, $X_{ci}(t)$ 为 t 时刻的第 i 个输入信号, $O_{ij}(t)$ 和 $y_{ij}(t)$ 分别为第 j 个隐含层节点的输入与输出, $V(t)$ 为网络的输出; $w_{cij}(t)$ 为 IRN 输入层到隐含层的权值, $w_{rij}(t)$ 为 IRN 回归层到隐含层的权值, $w_{oj}(t)$ 为 IRN 隐含层到输出层的权值。

2 卡尔曼滤波方法

文献[12]指出在系统实际的运行过程中,输入输出信号都会或多或少地受到外部噪声的干扰,主要表现在由不可控因素所造成的系统内部产生噪声和在观测信号中会掺杂观测噪声,这些都会影响系统的精确性。所以本文采用卡尔曼滤波方法进行补偿,从而得到两轮机器人的最优姿态。本文采用卡尔曼滤波方法对系统的状态值进行估计,并且建立状态方程和观测方程对信号进行处理,最终满足最小均方误差的估计。将内部回归网络所训练出的机器人两轮角速度数据作为观测信号,利用卡尔曼滤波器估计过程状态并从噪声的测量获得反馈,通过反馈控制为求解最小方差提供了一种有效的递归解法。

系统的离散线性模型描述为:

$$x(k) = \theta x(k-1) + Bu(k) + \omega(k) \quad (9)$$

$$z(k) = Hx(k) + v(k) \quad (10)$$

$$\text{其中 } \theta = \begin{bmatrix} 0.1891 & 0.2329 & 0.1125 & 0 \\ 0.2329 & 0.1891 & 0.1125 & 0 \\ 0.1125 & 0.1125 & 0.1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0.0328 & 0 \\ 0 & 0.0328 \\ -0.0328 & -0.0328 \\ 0 & 0 \end{bmatrix}, \quad H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ 分别为方程参}$$

数。 $\omega(k)$ 和 $v(k)$ 分别为过程噪声和测量噪声。在实际运行过程中它们的值变化较快,所以把它们看成均值为 0 的高斯白噪声,且它们的协方差用 Q 和 R 表示。

根据文献[13]中的证明过程,下面直接给出卡尔曼滤波算法公式:

$$x(k|k-1) = \theta x(k-1|k-1) + Bu(k) \quad (11)$$

$$p(k|k-1) = \theta p(k-1|k-1) \theta' + Q \quad (12)$$

$$x(k|k) = x(k|k-1) + kc(k)(z(k) - Hx(k|k-1)) \quad (13)$$

$$kc(k) = \frac{p(k|k-1)H'}{Hp(k|k-1)H' + R} \quad (14)$$

$$p(k|k) = (I - kc(k)H)p(k|k-1) \quad (15)$$

其中: $p(k|k-1)$ 为 $x(k|k-1)$ 的协方差, $kc(k)$ 为卡尔曼增益系数。再对系统的运行状态进行估计时,需要给定 $x(0|0)$ 和 $p(0|0)$ 的值,因为随着滤波器的工作 x 和 p 会逐渐收敛,所以给定它们为一个随机矩阵。

3 实验结果与仿真

针对非完整式两轮自平衡机器人而言,它是一个本征不稳定的系统,在实现各种运动之前,首先要保证机器人能够保持自身平衡,所以两轮机器人的姿态平衡是进行运动控制的首要条件。为了验证本文设计算法的有效性和可行性,对机器人的运动平衡性能进行了仿真实验和结果分析。

在基于内在动机的自主发育算法中,动作网络为 $net^3[4, 8, 1]$, 输入为两轮机器人的左右轮角速度,位移和机身自身倾角四个状态量,输出则为系统的控制量。评价网络为 $net^3[5, 8, 1]$, 输入为机器人四个状态量和控制量,输出是系统的评价函数 $V(t)$ 。在每次实验中,当机器人的尝试次数超过 1000 次或者一次尝试的平衡步数超过 20000 步时,则停止机器人的学习并重新开始另一次实验。如果机器人在其中一次尝试中经历 20000 步后还能保持平衡,则认为机器人已经学会平衡控制技能了。每次实验失败后,将初始状态及各个权值重新复位为一定范围内的随机值,再重新学习。平均 50 次的实验结果,实验平均 74 次学习失败后就能够实现平衡控制,且每次实验均能完成目标,成功概率为 100%。

机器人在没有干扰的未知环境下,采用基于内在动机的智能机器人自主发育算法,经过不断的学习,大约需要经历 350 步左右,即 3.5 s 左右就完成了自平衡控制,表现了其较快的自主学习能力,仿真结果中前 3000 步的各状态量响应曲线和评价函数与误差曲线如图 3 所示。

图 4 为本文所提出的基于内在动机的自主发育算法(IM-RL)与经典强化学习(RL)^[4] 仿真结果对比。可以看出 IM-RL 算法收敛速度优于后者,很快就能使机器人学会平衡控制,并且较经典 RL 算法鲁棒性明显提高。图 5 为两种算法的学习曲线,通过对比图 5(a) 和 (b) 可以发现:经典 RL 经过 108 次错误试探后才能实现机器人的自主平衡控制;而本文提出的 IM-RL 算法经过 84 次错误试探后就能实现机器人的自主平衡控制。因此,由以上仿真实验可以看出,加入内在动机机制的 IM-RL 算法能够使机器人获得更好的动态性能和较快的学习速度,同时,也体现了机器人较强的自主学习能力和平衡控制技能。

为了克服测量噪声对系统的影响,将卡尔曼滤波器用于两轮机器人平衡实验中。在实现平衡控制中,机器人的左右两轮角速度受到噪声干扰最为强烈,因此本文着重消除测量噪声对左右两轮角速度的影响。为了模拟真实系统,在输出中加入均值 $\mu = 0$ 、方差 $\sigma^2 = 0.1$ 的高斯分布噪声。图 6 分别为左右两轮角速度对比,可以看出通过运用卡尔曼滤波方法有效地消除了噪声的干扰。

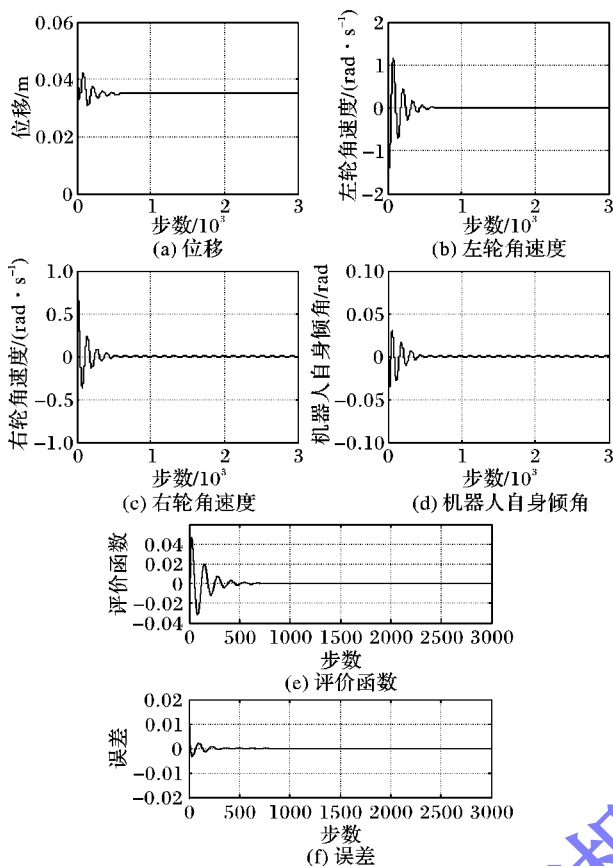


图3 实验仿真结果

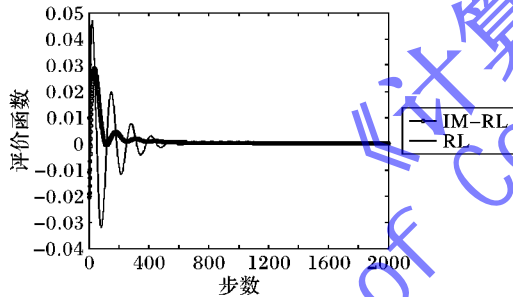
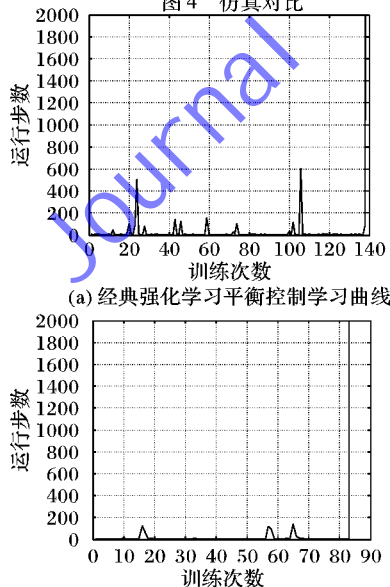


图4 仿真对比



(b) 基于内在动机的自主发育算法平衡控制学习曲线

图5 两种算法的平衡控制学习曲线

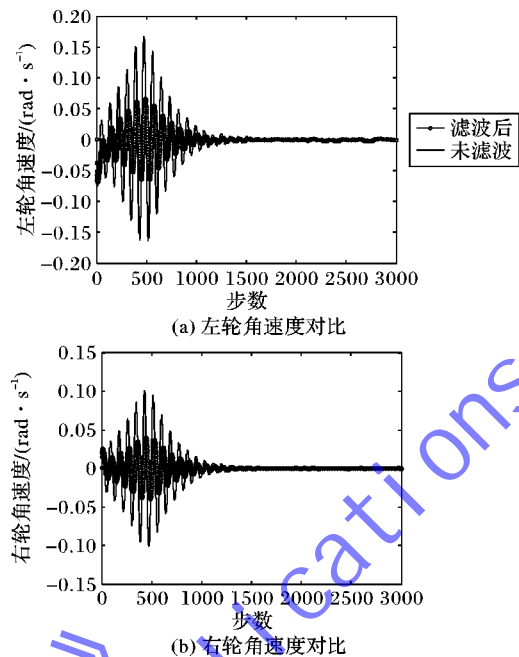


图6 左右两轮角速度对比

4 结语

本文在有关发育心理学、神经科学和发育机器人学的研究基础上,提出了一种基于内在动机的智能机器人自主发育算法。本文将这种算法应用在两轮机器人上,使机器人渐进的学会了行为与奖赏信号或惩罚信号之间的对应关系,并且能够根据内在动机取向机制确定动作的评价值,从而使机器人完成了自平衡控制。最后进行仿真对比实验,结果表明,机器人在较短时间内完成平衡控制,并且加入卡尔曼滤波算法后,解决测量噪声对左右轮角速度的影响,有利于机器人最优平衡姿态的学习。

参考文献:

- [1] YU H, ZHU C, LIU H. A survey on developmental robotics [J]. Transactions on Intelligent Systems, 2007, 2(4): 34-39. (于化龙, 朱长明, 刘海波. 发育机器人研究综述[J]. 智能系统学报, 2007, 2(4): 34-39.)
- [2] JIANG G. The control of inverted pendulum based on Q learning algorithm and BP neural network [J]. Acta Automatica Sinica, 1998, 24(5): 662-667. (蒋国飞. 基于Q学习算法和BP神经网络的倒立摆控制[J]. 自动化学报, 1998, 24(5): 662-667.)
- [3] PATHAK K, FRANCH J, AGRAWAL S K. Velocity and position control of a wheeled inverted pendulum by partial feedback linearization [J]. IEEE Transactions on Robotics, 2005, 21(3): 505-513.
- [4] RUAN X, CAI J, CHEN J. Learning to control two-wheeled self-balancing robot using reinforcement learning rules [J]. Computer Measurement and Control, 2009, 17(2): 321-323. (阮晓钢, 蔡建炎, 陈静. 基于强化学习规则的两轮机器人自平衡控制[J]. 计算机测量与控制, 2009, 17(2): 321-323.)
- [5] LI M, JIAO L, QIAO J. Hierarchical fuzzy control for self-balancing two-wheel robots [J]. Control Engineering of China, 2009, 16(1): 80-82. (李明爱, 焦利芳, 乔俊飞. 自平衡两轮机器人的分层模糊控制[J]. 控制工程, 2009, 16(1): 80-82.)
- [6] OUDEYER P Y, KAPLAN F. What is intrinsic motivation? A typology of computational approaches [J]. Frontiers in Neurobotics, 2007, 1: Article 6.

(下转第2610页)

多分类的旋转机械故障诊断领域。

参考文献:

- [1] ZHANG J, FENG Z, CHU F. Rolling bearing fault feature with time - the wavelet energy spectrum extraction method [J]. Journal of Mechanical Engineering, 2011, 47(17): 33 - 49. (张进, 冯志鹏, 褚福磊. 滚动轴承故障特征的时间 - 小波能量谱提取方法[J]. 机械工程学报, 2011, 47(17): 44 - 49.)
- [2] YANG J, ZHOU P. Empirical mode decomposition and Laplace wavelet in the application of locomotive diesel engine gear fault diagnosis [J]. Journal of Mechanical Engineering, 2011, 47(7): 109 - 115. (杨江天, 周培钰. 经验模态分解和 Laplace 小波在机车柴油机齿轮系故障诊断中的应用[J]. 机械工程学报, 2011, 47(7): 109 - 115.)
- [3] HE T, LIN Y, HAO P, *et al.* Local mean decomposition in the application of gear fault diagnosis research [J]. Vibration and Shock, 2011, 30(6): 196 - 201. (何田, 林意洲, 郝普刚, 等. 局部均值分解在齿轮故障诊断中的应用研究[J]. 振动与冲击, 2011, 30(6): 196 - 201.)
- [4] SU W, WANG F, ZHU H, *et al.* Feature extraction of rolling element bearing fault using wavelet packet sample entropy [J]. Journal of Vibration, Measurement and Diagnosis, 2011, 31(2): 162 - 166. (苏文胜, 王奉涛, 朱泓, 等. 基于小波包样本熵的滚动轴承故障特征提取[J]. 振动、测试与诊断, 2011, 31(2): 162 - 166.)
- [5] RUAN Q, RUAN Y. Digital image processing [M]. Beijing: Publishing House of Electronics Industry, 2011: 138 - 140. (阮秋琦, 阮宇智. 数字图像处理[M]. 北京: 电子工业出版社, 2011: 138 - 140.)
- [6] AN J, TIAN B, SUN Y, *et al.* An algorithm for direct sequence spread spectrum signal detection based on intrinsic time-scale decomposition [J]. Journal of Electronics and Information Technology, 2010, 32(5): 1178 - 182. (安金坤, 田斌, 孙永军, 等. 一种基于 ITD 算法的直扩信号检测算法[J]. 电子与信息学报, 2010, 32(5): 1178 - 182.)
- [7] LIAO X, WAN Z, XIONG X. Based on ELM and LS-SVM rolling bearing fault diagnosis method [J]. Journal of Chemical Industry, 2013, 64(12): 4667 - 4673. (廖星智, 万舟, 熊新. 基于 ELM 与 LS-SVM 的滚动轴承故障诊断方法[J]. 化工学报, 2013, 64(12): 4667 - 4673.)
- [8] CHENG J, MA X, YANG Y. Rolling bearing fault diagnosis method based on permutation entropy and VPMCD [J]. Journal of Vibration and Shock, 2014, 33(11): 119 - 123. (程军圣, 马兴伟, 杨宇. 基于排列熵和 VPMCD 的滚动轴承故障诊断方法[J]. 振动与冲击, 2014, 33(11): 119 - 123.)
- [9] CHENG J, MA X, YANG Y. Gear fault diagnosis method based on VPMCD and EMD [J]. Journal of Vibration and Shock, 2013, 32(20): 9 - 13. (程军圣, 马兴伟, 杨宇. 基于 VPMCD 和 EMD 的齿轮故障诊断方法[J]. 振动与冲击, 2013, 32(20): 9 - 13.)
- [10] ZHANG K, CHENG J, YANG Y. Local mean decomposition method and its application in the rolling bearing fault diagnosis [J]. The Chinese Mechanical Engineering, 2009, 20(22): 2712 - 2717. (张亢, 程军圣, 杨宇. 局部均值分解方法及其在滚动轴承故障诊断中的应用[J]. 中国机械工程, 2009, 20(22): 2712 - 2717.)
- [11] ZHANG K, CHENG J, YANG Y. Based on local mean decomposition of rational spline function method and its application [J]. Journal of Vibration Engineering, 2011, 24(1): 96 - 103. (张亢, 程军圣, 杨宇. 基于有理样条函数的局部均值分解方法及其应用[J]. 振动工程学报, 2011, 24(1): 96 - 103.)
- [12] SMITH J S. The local mean decomposition and its application to EEG perception data [J]. Journal of the Royal Society Interface, 2005, 2(5): 443 - 454.
- [13] CHENG J, LUO S, YANG B, *et al.* LMD energy torque and VPMCD in the application of the bearing fault intelligent diagnosis [J]. Journal of Vibration Engineering, 2013, 26(5): 753 - 754. (程军圣, 罗颂荣, 杨斌, 等. LMD 能量矩和变量预测模型模式识别在轴承故障智能诊断中的应用[J]. 振动工程学报, 2013, 26(5): 753 - 754.)
- [14] RAGHURAJ R, LAKSHMINARAYANAN S. Variable predictive models—a new multivariate classification approach for pattern recognition applications [J]. Pattern Recognition, 2009, 42(1): 6 - 16.
- [15] ZHONG B, HUANG R. Mechanical fault diagnosis [M]. Beijing: China Machine Press, 2006: 43 - 146. (钟秉林, 黄仁. 机械故障诊断学[M]. 北京: 机械工业出版社, 2006: 43 - 146.)
- [16] WEN X. Pattern recognition and condition monitoring [M]. Beijing: Science Press, 2007: 55 - 197. (温熙森. 模式识别与状态监控[M]. 北京: 科学出版社, 2007: 55 - 197.)
- [7] PFEIFER R, BONGARD J C. How the body shapes the way we think: a new view of intelligence [M]. Cambridge: MIT Press, 2006: 75 - 77.
- [8] CEDERBORG T, OUDEYER P Y. From language to motor Gavage: unified imitation learning of multiple linguistic and nonlinguistic sensorimotor skills [J]. IEEE Transactions on Autonomous Mental Development, 2013, 5(3): 222 - 239.
- [9] WHITE R W. Motivation reconsidered: the concept of competence [J]. Psychological Review, 1959, 6(6): 297 - 333.
- [10] DENG Z. The information fusion filtering theory and application [M]. Harbin: Harbin Institute of Technology Press, 2007: 120 - 124. (邓自立. 信息融合滤波理论及其应用[M]. 哈尔滨: 哈尔滨工业大学出版社, 2007: 120 - 124.)
- [11] WANG R, SUN L, RUAN X. Reinforcement learning based on internal regression neural network [J]. Control Engineering of China, 2005, 12(2): 138 - 140. (王瑞霞, 孙亮, 阮晓钢. 基于内部回归神经网络的强化学习[J]. 控制工程, 2005, 12(2): 138 - 140.)
- [12] QI Q, LI Z, TAN Z. Noise suppression of inverted pendulum system based on Kalman filter [J]. Control and Decision, 2010, 25(8): 1144 - 1148. (祁虔, 李祖枢, 谭智. 基于卡尔曼滤波的倒立摆控制系统噪声抑制[J]. 控制与决策, 2010, 25(8): 1144 - 1148.)
- [13] SONG W, ZHANG Y. Kalman filter [M]. Beijing: Science Press, 1991: 116 - 119. (宋文尧, 张牙. 卡尔曼滤波[M]. 北京: 科学出版社, 1991: 116 - 119.)

(上接第 2605 页)