

基于深度强化学习的智能暖气温度控制系统

李涛^{1,2}, 魏庆来^{1,2}

(1. 中国科学院自动化研究所复杂系统管理与控制国家重点实验室, 北京 100190;

2. 中国科学院大学, 北京 100049)

摘要: 研究如何通过暖气设备自适应地调节室温, 提升室内环境的舒适度, 具有非常重要的意义。因此, 提出了基于双深度 Q 网络方法的智能暖气温度控制系统, 根据人的表情信息控制暖气设备的阀门开度, 实时调整室温。首先, 介绍针对原始输入状态的预处理算法。然后, 设计通过双深度 Q 网络方法学习控制暖气设备阀门开度的最佳策略。最后, 通过仿真结果验证提出的方法的有效性。

关键词: 深度强化学习; 暖气设备; 温度控制; 疲劳检测; 图像处理

中图分类号: TP301.6

文献标识码: A

doi: 10.11959/j.issn.2096-6652.202037

Intelligent heating temperature control system based on deep reinforcement learning

LI Tao^{1,2}, WEI Qinglai^{1,2}

1. The State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Science, Beijing 100190, China

2. University of Chinese Academy of Sciences, Beijing 100049, China

Abstract: It is of great significance to study how to adjust the room temperature adaptively through heating equipment to improve the comfort of the indoor environment. Therefore, a double deep Q network method was developed to control the valve opening of heating equipment to adjust the indoor temperature in real time via human expressions. Firstly, the preprocessing algorithm for the original input state was introduced. Secondly, a double deep Q network method was designed to learn the optimal control policy of the valve opening of heating equipment. Finally, simulation results were given to illustrate the effectiveness of the method proposed.

Key words: deep reinforcement learning, heating equipment, temperature control, fatigue detection, image processing

1 引言

随着我国经济、科技水平的不断提高, 越来越多的人开始追求更加舒适的居住、工作环境。在影响环境舒适度的众多因素中, 温度是非常重要的因素之一。相关研究^[1]表明, 过高或过低的温度都会使人感到不舒适, 降低人的工作效率。暖气设备是常见的室温调节装置。研究如何根据室内人员的需

求通过暖气设备自动地调节室温, 对于提升室内环境的舒适度具有非常重要的意义。

传统的通过暖气设备控制室温的方法包括比例-积分-微分控制^[2]、动态矩阵预测控制^[3]以及模糊控制^[4]等。在传统的控制方法中, 室温的设定值通常是预先设定的固定值, 控制器需要调节暖气设备, 使得室温尽可能稳定在设定值附近。然而, 相关研究^[5]表明, 稳定的室温无法有效提升人的工作

收稿日期: 2020-11-18; 修回日期: 2020-11-30

通信作者: 李涛, litao2019@ia.ac.cn

基金项目: 国家自然科学基金资助项目 (No.61722312, No.62073321)

Foundation Items: The National Natural Science Foundation of China (No. 61722312, No. 62073321)

效率,甚至会降低人的工作效率。因此,本文抛弃了常规的固定室温设定值的控制思路,直接以消除人的困倦为控制目标,根据人的状态实时调整室温。本文设计了一种将人的表情作为输入的自学习控制器。这种自学习控制器能够根据人的表情自主学习到合适的控制策略,通过调整暖气设备的阀门开度控制室温,从而尽可能地消除室内人员的疲劳,提升室内人员的工作效率。

户外天气情况的变化和室内人员的自由活动使得房屋的热模型具有很强的不确定性。这导致在实际应用时基于模型的控制方法往往无法取得良好的效果。因此,许多研究者使用无模型的学习算法,尤其是强化学习方法^[6]来研究室温控制问题。在参考文献[7]中,Barrett E 使用贝叶斯学习方法预测室内人员的活动情况,使用 Q 学习方法学习控制策略。Barrett E 使用 Bang-Bang 控制方法控制加热设备,控制器的输入状态包括室温、时间和天气等。显然,由于传统强化学习算法的限制,控制器的输入状态只能是人们已经提取好的环境信息,无法直接从高维信息中自动学习到控制需要的特征。

近年来,深度学习取得了巨大进展^[8],深度强化学习也得到了越来越多的关注^[9]。在参考文献[10]中,Wei T S 采用深度 Q 网络方法学习控制策略,控制器的输入状态包括时间、室温以及室外环境的干扰。尽管参考文献[10]使用了深度神经网络,但是仍然使用人们已经提取好的环境特征作为控制器的输入,没有发挥深度神经网络自动提取特征的能力。在参考文献[11]中,Wei Q L 等人使用深度 Q 网络方法,将人的表情和空调温度设定值作为控制器输入,通过控制空调设备调节室温。然而,参考文献[11]中的空调设备不同于本文所要研究的暖气设备。第一,空调制热会造成室内空气干燥,降低人的舒适度;而暖气设备由下向上加热室内空气,使热空气分布更均匀,人感觉较舒适,因此暖气设备比空调更有利于提升人的舒适度。第二,空调通常具有自动调温功能,只需给定温度设定值,空调即可自动地将房间温度调整到设定值附近,完成闭环控制;而大多暖气设备并不具有自动调温功能,只能通过控制暖气设备的阀门开度控制暖气设备向房间中输入的热量,间接调节室温,进行开环控制。基于这两方面的不同,有必要研究基于深度强化学习的智能暖气控制方案。

本文以消除人的疲劳、提升人的工作效率为直

接目标,将人的表情作为控制器的输入状态,使用双深度 Q 网络方法,通过暖气设备控制室温。接下来,本文将依次介绍总体方案设计、仿真环境与实验结果。

2 总体方案设计

本节将系统阐述以人的表情为输入状态,通过暖气设备自适应调整室温的总体方案。

本文的控制目标是根据人的状态实时调整室温,因此本文采用能够直接反映人的困意的人脸表情图像作为控制器的输入状态 $\mathbf{x}$ 。同时,将当前暖气设备的阀门的开度也作为输入状态的一部分。

控制器通过控制暖气设备的阀门开度调节室温。在本文中,设定暖气设备的阀门开度有 9 个档位,分别是全闭、1/8 开度、1/4 开度、3/8 开度、半开、5/8 开度、3/4 开度、7/8 开度和全开。控制器的可执行动作共有 3 种,分别是将阀门开大一个档位、将阀门关小一个档位和保持阀门开度不变。

为了减少控制器的计算负担,需要对控制器的输入状态 $\mathbf{x}$ 进行预处理。状态预处理过程共分为两步:第一步是执行人脸检测步骤,从图像中获得人的表情;第二步是从人脸表情图像中截取人眼图像。暖气设备的阀门开度将直接作为深度神经网络最末层的一个输入特征,这一点将在叙述双深度 Q 网络算法的神经网络结构时进一步说明。

控制器的原始输入状态是一幅图像,它在计算机中的存储形式是 $Q \times P \times 3$ 维的矩阵,参数 Q 和 P 取决于图像的分辨率。在人脸检测步骤中,首先将 $Q \times P \times 3$ 维的原始输入图像 $\mathbf{x}$ 灰度化,得到 $Q \times P \times 1$ 维的灰度图像 $\mathbf{x}_g$ 。然后,使用参考文献[12]中的人脸检测器检测灰度图像 $\mathbf{x}_g$ 中的人脸位置,得到人脸所在矩形区域的左上角坐标 (x_l, y_l) 和右下角坐标 (x_r, y_r) 。在截取人眼图像的步骤中,以人脸检测步骤中得到的 (x_l, y_l) 和 (x_r, y_r) 为参考文献[13]中的人脸关键点检测器的输入状态,从灰度图像 $\mathbf{x}_g$ 中估计 68 个人脸关键点的坐标。在这 68 个人脸关键点坐标中,第 37 个坐标 (x_{re}, y_{re}) 和第 46 个坐标 (x_{le}, y_{le}) 分别代表右眼的最右端和左眼的最左端。在 (x_{re}, y_{re}) 上减去一个常数 M 得到 (x'_{re}, y'_{re}) , 在 (x_{le}, y_{le}) 上加上一个常数 M 得到 (x'_{le}, y'_{le}) 。 (x'_{re}, y'_{re}) 和 (x'_{le}, y'_{le}) 分别代表了眼部图像的左上角和右下角坐标。人眼周围坐标点的相对位置如图 1 所示。接

下来,根据 (x'_{re}, y'_{re}) 和 (x'_{le}, y'_{le}) 从灰度图像 x_g 中截取人眼灰度图像,并将其大小调整为 $q \times p \times 1$, q 和 p 是预设的参数。人眼灰度图像就是预处理后的状态,设为 $\tilde{x}$,其维度为 $q \times p \times 1$ 。

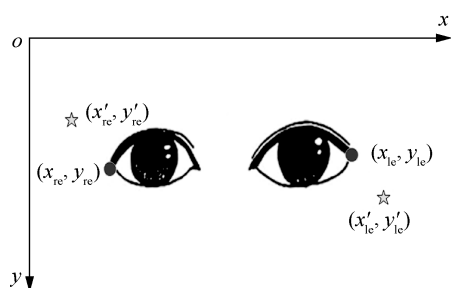


图1 人眼周围坐标点的相对位置

本文最主要的控制目的是消除人的困意,因此对于困意的检测是至关重要的。研究表明,人的困意与人眼的平稳张开程度有着密切的关系,因此本文使用非眨眼状态下的眼睛张开程度来衡量人的困意。困意检测分为两步:第一步是检测人是否处于眨眼状态,第二步是检测非眨眼状态下眼睛的张开程度。这两个步骤都需要实时检测眼睛的张开程度。

在状态预处理过程的截取人眼图像步骤中,已经检测出了人脸的68个关键点。在这68个关键点中,左右眼各占6个。参考文献[14]使用眼睛纵横比(eye aspect ratio, EAR)的平均值衡量人眼的张开程度。受参考文献[14]启发,本文也使用眼睛纵横比来衡量眼睛的张开程度。但是考虑到眨眼检测的实时性要求以及双眼动作的同步性,本文仅使用右眼的纵横比来衡量眼睛的张开程度。右眼纵横比 EAR^r 如式(1)所示:

$$EAR^r = \frac{\|d_2 - d_6\| + \|d_3 - d_5\|}{2\|d_1 - d_4\|} \quad (1)$$

其中, d_2 、 d_3 、 d_5 、 d_6 是分布在右眼上下边缘的关键点坐标, d_1 、 d_4 是分布在右眼左右边缘的点坐标。从式(1)可以看出, EAR^r 等于右眼的纵向距离除以横向距离。当眼睛处于张开状态时, EAR^r 基本保持不变;当眼睛闭合时, EAR^r 接近0;当眨眼时, EAR^r 迅速变化。因此,本文实时检测人眼的 EAR^r 值,通过判断 EAR^r 是否快速变化判断人是否处于眨眼状态,将 EAR^r 平稳时的值作为衡量人困意的依据。

奖励函数的作用是使控制器在与外部环境的

交互中逐渐学习到正确的控制策略,告诉控制器哪些动作应当被鼓励。考虑到本文设计的控制器的主要目标是消除人的困意,设计奖励函数如式(2)所示:

$$r_t = \begin{cases} -1, & EAR_t^r \leq EAR_{t-1}^r \\ 10, & EAR_t^r = EAR_{\max}^r \\ 0, & \text{其他} \end{cases} \quad (2)$$

其中, EAR_t^r 和 EAR_{t-1}^r 分别表示状态 x_t 和状态 x_{t-1} 的 EAR^r 值。在 $t-1$ 时刻,控制器的输入状态为 x_{t-1} ,控制器执行动作 u_{t-1} 后,得到下一个时刻的状态 x_t ,此时即可由 x_t 和 x_{t-1} 计算奖励 r_t 。奖励函数表达了这样的理念:当控制器的动作使人困意增大时,即 $EAR_t^r \leq EAR_{t-1}^r$,这个动作是不被鼓励的;当控制器的动作使人兴奋时,即 $EAR_t^r = EAR_{\max}^r$,这个动作是被极度鼓励的。

前文主要介绍了状态预处理算法和奖励函数的设计。有了上面的铺垫后,接下来介绍用于控制暖气设备阀门开度的双深度Q网络算法。

首先,介绍使用近似Q函数的深度Q网络。为了衡量在状态 x 下执行控制动作 u 所能获得的平均报酬,引入动作值函数,也称Q函数。由于本文的状态是图像,因此使用深度神经网络参数化动作值函数 $Q(\tilde{x}, u; \theta)$, $Q(\tilde{x}, u; \theta)$ 被称为Q网络,其中 θ 是Q网络的参数。Q网络的输入是预处理后的状态 $\tilde{x}$ 和控制动作 u ,输出是状态 $\tilde{x}$ 下执行动作 u 获得的平均报酬。在本文中,将Q网络的结构设计为卷积神经网络。Q网络的输入是预处理后的状态 $\tilde{x}$ 。经过多层卷积层, $\tilde{x}$ 的空间维度减小,通道维度增加。最后一层卷积层的输出将被展开成一个一维向量 v ,这个向量就是由卷积神经网络从图像中自动提取出的特征。暖气设备当前的阀门开度也是控制器输入状态的一部分,在这里,将阀门开度与 v 合并,输入卷积层之后的全连接层。

其次,介绍双深度Q网络方法的经验回放机制和目标网络机制。在双深度Q网络方法中,笔者使用了经验回放机制打破样本之间的关联性,从而减小神经网络更新时的方差。经验回放机制为:在每一个时间步,控制器接收状态 x_t ,做出动作 u_t ,得到奖励 r_t ,获得下一个状态 x_{t+1} ,控制器将得到的经验 (x_t, u_t, r_t, x_{t+1}) 存储到缓冲区 D 中。当需要训练Q网络时,就从 D 中抽取小批量的经验来训练Q网络。双深度Q网络方法引入目标网络机制来降低训

练 Q 网络时输入值和目标值之间的相关性,增强算法的收敛性。目标网络机制为:建立一个与 Q 网络 $Q(\tilde{x}, u; \theta)$ 结构相同的目标网络 $\hat{Q}(\tilde{x}, u; \theta^-)$ 。 $\hat{Q}(\tilde{x}, u; \theta^-)$ 的参数 θ^- 周期性地更新,在两次更新的间隔, θ^- 保持不变。

接下来,介绍双深度 Q 网络方法中 Q 网络目标值的计算方法。在深度 Q 网络方法中,计算 Q 网络的目标值如式 (3) 所示:

$$Y_t^{\text{DQN}} \equiv r_t + \gamma \max_{u_{t+1}} Q(\tilde{x}_{t+1}, u_{t+1}; \theta^-) \quad (3)$$

其中, γ 是折扣因子。但是,式 (3) 中的 $\max$ 操作会造成 Q 函数的过估计,为了避免这种情况,双深度 Q 网络方法使用式 (4) 计算 Q 网络的目标值:

$$Y_t^{\text{DoubleDQN}} \equiv r_t + \gamma Q\left(\tilde{x}_{t+1}, \arg \max_u Q(\tilde{x}_{t+1}, u; \theta); \theta^-\right) \quad (4)$$

最后,介绍双深度 Q 网络方法的损失函数。在第 i 个训练回合, Q 网络中的参数为 θ_i , 目标网络中的参数为 θ_i^- 。第 i 个训练回合的训练目标是使 Q 网络尽可能接近目标网络,在第 i 个训练回合的损失函数如式 (5) 所示:

$$L_i(\theta_i) = \frac{1}{2} [r_t + \gamma Q(\tilde{x}_{t+1}, \arg \max_u Q(\tilde{x}_{t+1}, u; \theta_i); \theta_i^-) - Q(\tilde{x}_t, u_t; \theta_i)]^2 \quad (5)$$

本文提出的以人的表情为输入状态,通过暖气设备自适应调整室温的方案与参考文献[11]中的方案主要有以下几点不同。

第一,被控对象不同。本文使用暖气设备调节室温,而参考文献[11]使用空调调节室温。使用空调制热会造成室内空气干燥,使人不舒服,从提升人的舒适度的角度来看,使用暖气制热比使用空调更合适。

第二,控制器施加给被控对象的动作不同。在参考文献[11]中,空调大多带有自动调温功能,因此,控制器可直接控制空调的温度设定值,空调与房间之间存在单独的控制回路,空调可以自动调节室温到设定值附近。而在本文中,大多暖气设备只有调节阀,控制器需要通过改变调节阀的开度调节室温,暖气设备与房间之间并不存在控制回路。

第三,状态预处理算法和 Q 网络的结构不同。在参考文献[11]中,空调温度设定值作为图像的一个通道与预处理后的图像合并,由深度神经网络自动提取特征。而在本文中,将阀门开度作为一个已

经提取好的特征直接输入最末层的全连接网络,大大减小了计算量。

第四,计算眼睛纵横比的方式不同。在参考文献[11]中,分别计算左右眼纵横比,再取平均值得到衡量双眼纵横比的值。而在本文中,考虑到双眼运动的一致性以及眨眼检测的实时性,使用右眼的纵横比代替双眼纵横比,减小了计算量,更便于算法实时运行。

第五,在参考文献[11]中,使用深度 Q 网络方法学习最优策略,而在本文中,为了避免深度 Q 网络方法造成的动作价值过估计现象,使用双深度 Q 网络学习控制策略。实验表明,本文中的双深度 Q 网络方法比参考文献[11]中的方法收敛更快,更平稳。

3 仿真环境与实验结果

本节将介绍仿真环境的搭建以及仿真实验的结果。智能体在仿真环境中的工作流程如下:控制器实时检测人眼的运动情况,在 t 时刻,控制器采集到人非眨眼状态下的表情,并将其作为 t 时刻的状态 x_t , 预处理 x_t 得到 $\tilde{x}_t$ 。接下来,控制器将 $\tilde{x}_t$ 输入 Q 网络,根据 ε -贪婪策略改变暖气设备的阀门开度,进而房间温度会随着时间改变。一段时间之后(假设此时房间温度已经达到稳定),控制器再次执行上述过程,直到能使人保持在最亢奋的状态。

为了模拟房间中的温度变化,本文建立了一个简化的房屋热模型。由热力学定律可解得室温变化的微分方程:

$$\tau_i(t) = \rho_i \sigma_i + \tau_o + (\tau_r(0) - \rho_i \sigma_i - \tau_o) e^{-\frac{1}{\rho_i \mu_i} t} \quad (6)$$

其中, τ_r 和 τ_o 分别代表室内温度和室外温度, μ_i 和 ρ_i 分别代表房间的热容和热阻, σ_i 代表房间的输入热量。由式 (6) 可以看出,稳态室温为 $\tau_\infty = \rho_i \sigma_i + \tau_o$ 。可以看到,房间稳态温度 τ_∞ 与暖气设备向房间输入的热量 σ_i 呈线性关系,因此,可以认为房间稳态温度 τ_∞ 与暖气设备的阀门开度近似为线性关系。

为了模拟人的表情随温度的变化,需要建立一个温度-表情数据库,在这个数据库中,人脸的图片与温度一一对应,输入房间温度,就能得到一张人脸表情的图片。设定房间温度的取值范围是 18 °C 到 34 °C 之间的 6 个值。各个温度下的眼睛闭合程度见表 1。从表 1 可以看出,当室内温度为 26 °C 时, EAR^r 最大。

表1 温度-EAR^r对照

序号	房间温度/℃	EAR ^r
1	18	0.095
2	20	0.212
3	22	0.243
4	24	0.329
5	26	0.381
6	28	0.324
7	30	0.271
8	32	0.118
9	34	0.105

在实验中, 每个回合开始时的初始温度将从 18℃、20℃、32℃和 34℃中随机选择。控制器将根据人的表情调整暖气设备的阀门开度。当房间温度被调整到最合适的温度时, 人的困意最少, 此时认为控制器已经达到目的, 本回合结束, 重新开始的回合, 房间温度将被重新初始化。

控制器训练过程中的回合总长度、回合总奖励、回合损失的变化分别如图2、图3、图4所示。

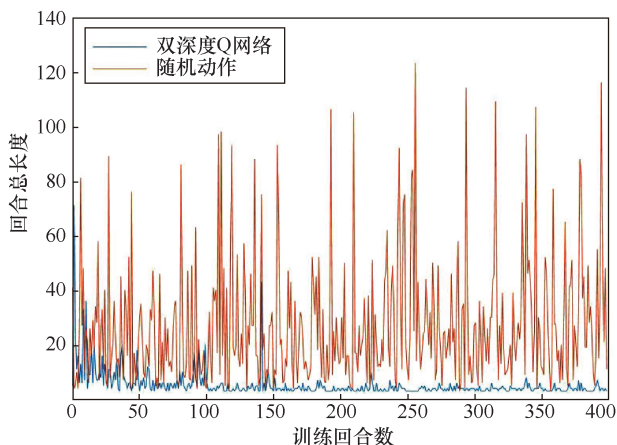


图2 回合总长度

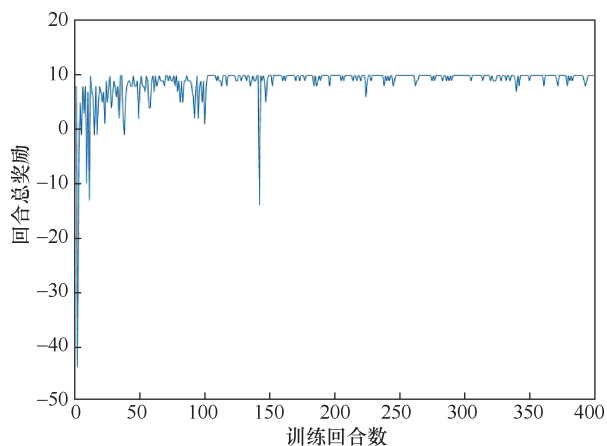


图3 回合总奖励

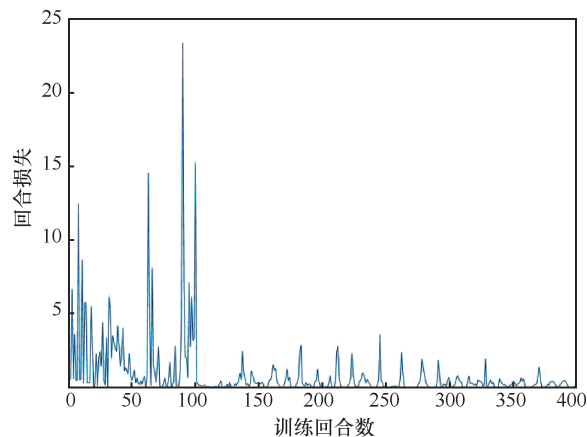


图4 回合损失

回合总长度代表调节暖气设备阀门开度的次数。从图2可以看出, 训练初期的回合总长度要明显长于训练末期。在训练末期, 回合总长度会稳定在3或4。在训练末期, 控制器仍对环境保持着一定的探索性, 因此训练末期的回合总长度会在3和4附近有微小的波动。为了展示所提方法的有效性, 将双深度Q网络方法与采取随机动作的方法进行对比, 通过对比可以看出, 双深度Q网络方法的回合总长度明显更小。回合总奖励表示智能体在这一回合中获得的奖励总数。随着训练的进行, 智能体能够在每个状态下都做出正确的决定。因此在训练末期, 回合总奖励稳定在10附近。同样地, 由于在训练末期, 控制器仍保持一定的探索性, 曲线在10附近有轻微的波动。回合损失指损失函数在这一回合结束时的值。从图4可以看出, 随着训练的进行, 回合损失逐渐接近于0。

4 结束语

本文提出了一种基于深度强化学习方法, 将人的表情作为输入状态, 以消除人的困意为直接目标的自适应暖气设备调温系统。通过使用双深度Q网络方法, 智能体能够直接从图像中学习到控制策略。仿真结果表明, 智能体在充分学习后能够控制暖气设备自适应地将室温调整到最合适的温度。

在未来的工作中, 笔者将使用预训练好的卷积神经网络作为Q网络的初值, 从而避免从头训练Q网络, 减小训练耗时。

参考文献:

- [1] HEDGE A, GAYGEN D E. Indoor environment conditions and computer work in an office[J]. HVAC&R Research, 2010, 16(2): 123-138.

- [2] 施志荣, 康玉文. 基于模糊自适应PID的分布式室温监控系统设计[J]. 齐鲁工业大学学报, 2017, 31(1): 78-80.
SHI Z R, KANG Y W. A control system design of distributed room temperature: based on fuzzy adaptive PID[J]. Journal of Qilu University of Technology, 2017, 31(1): 78-80.
- [3] 高秀娟. 集中供热系统室温控制策略研究[D]. 杭州: 浙江大学, 2014.
GAO X J. Research on room temperature control strategy for central heating system[D]. Hangzhou: Zhejiang University, 2014.
- [4] 李冰皓. 暖气热量智能控制系统设计及算法研究[D]. 石家庄: 河北科技大学, 2015.
LI B H. Intelligent control system design and algorithm study of heating calories[D]. Shijiazhuang: Hebei University of Science and Technology, 2015.
- [5] TSE W L, SO A T P. The importance of human productivity to air-conditioning control in office environments[J]. HVAC&R Research, 2007, 13(1): 3-21.
- [6] 王飞跃, 曹东璞, 魏庆来. 强化学习: 迈向知行合一的智能机制与算法[J]. 智能科学与技术学报, 2020, 2(2): 101-106.
WANG F Y, CAO D P, WEI Q L. Reinforcement learning: toward action-knowledge merged intelligent mechanisms and algorithms[J]. Chinese Journal of Intelligent Science and Technology, 2020, 2(2): 101-106.
- [7] BARRETT E, LINDER S. Autonomous HVAC control, a reinforcement learning approach[C]//Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Cham: Springer, 2015: 3-19.
- [8] 张钊. 人工智能进入后深度学习时代[J]. 智能科学与技术学报, 2019, 1(1): 4-6.
ZHANG B. Artificial intelligence is entering the post deep-learning era[J]. Chinese Journal of Intelligent Science and Technology, 2019, 1(1): 4-6.
- [9] 孙长银, 穆朝絮. 多智能体深度强化学习的若干关键科学问题[J]. 自动化学报, 2020, 46(7): 1301-1312.
SUN C Y, MU C X. Important scientific problems of multi-agent deep reinforcement learning[J]. IEEE/CAA Journal of Automatica Sinica, 2020, 46(7): 1301-1312.
- [10] WEI T S, WANG Y Z, ZHU Q. Deep reinforcement learning for building HVAC control[C]//Proceedings of the 54th Annual Design Automation Conference. [S.l.:s.n.], 2017: 1-6.
- [11] WEI Q L, LI T, LIU D R. Learning control for air conditioning systems via human expressions[J]. IEEE Transactions on Industrial Electronics, 2020(99): 1.
- [12] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]// 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2005, 1: 886-893.
- [13] KAZEMI V, SULLIVAN J. One millisecond face alignment with an ensemble of regression trees[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2014: 1867-1874.
- [14] SOUKUPOVA T, CECCH J. Eye blink detection using facial landmarks[C]//The 21st Computer Vision Winter Workshop, Rimske Toplice, Slovenia. [S.l.:s.n.], 2016.

[作者简介]



李涛(1996—), 男, 中国科学院自动化研究所复杂系统管理与控制国家重点实验室博士生, 主要研究方向为深度强化学习、自适应动态规划、最优控制。



魏庆来(1979—), 男, 博士, 中国科学院自动化研究所研究员, 复杂系统管理与控制国家重点实验室副主任, 中国科学院大学岗位教授, 青岛智能产业技术研究院智能技术创新中心主任, 主要研究方向为自主学习控制、平行控制、自适应动态规划、智能控制、最优控制及其工业应用。