

Journal of Image
and Graphics

中国图象图形学报



ISSN1006-8961
CN11-3758/TB

2012 **6**
Vol.17 No.

中国科学院遥感应用研究所
中国图象图形学学会主办
北京应用物理与计算数学研究所

中国图象图形学报

Zhongguo Tuxiang Tuxing Xuebao

2012年6月 第17卷 第6期(总第194期)

目次

综述

人脸图像的年龄估计技术研究 王先梅, 梁玲燕, 王志良, 胡四泉(603)

图像统计模型参数估计中的期望最大值算法 李旭超(619)

图像处理和编码

鲁棒的梯度驱动图像修复算法 叶学义, 王靖, 赵知劲, 陈华华(630)

快速去块效应的线性规划方法 金剑秋, 刘春晓, 王勋, 章志勇(636)

DWT-SVD 域全盲自嵌入鲁棒量化水印算法 叶天语(644)

Mean-shift 滤波和直方图增强的图像弱边缘提取 冀峰, 高新波, 谢松云(651)

图像分析和识别

梯度点对约束的结构化车道检测 王永忠, 王晓云, 文成林(657)

多水平外区抑制的轮廓检测 闫超, 张建州(664)

基于局部投影与块 LBP 特征的图像检索 邹彬, 潘志斌, 胡森(671)

图像理解和计算机视觉

视觉陌生度驱动的增量自主式视觉学习算法 瞿心昱, 姚明海, 顾勤龙(678)

霍夫空间中多足球机器人协作目标定位算法 许家铭, 解仑, 王志良, 倪善超(687)

计算机图形学

虚拟 CCD 线阵星载光学传感器内视场拼接 张过, 刘斌, 江万寿(696)

“第八届数字电视与无线多媒体通信国际论坛”会议专栏

追踪盗版者的抗共谋数字指纹 李晓强, 张煌, 赵洋洋, 汪晶晶(702)

LBP 与鉴别模式结合的热红外人脸识别 谢志华, 伍世虔, 方志军(707)

电影中复杂事件的检索与识别 杜吉祥, 郭一兰, 翟传敏(712)

方向特征和网格特征融合的离线签名鉴别 杨丹凤, 吕岳(717)

视觉关注度的立体图像质量评价 张艳, 安平, 张秋闻, 张兆杨(722)

低开销的异构数据交换 赵凯, 赵正德(726)

灵活的多视点视频编码预测结构 张艳, 蔡灿辉(730)

跨层反馈在无线视频自适应编码中的应用 王垚中, 郑世宝, 张重阳, 刘勃(736)

中国图象图形学报

刊名题字: 宋 健

月刊(1996 年创刊)

第 17 卷 第 6 期

2012 年 6 月 16 日出版

主管单位 中国科学院
主 办 中国科学院遥感应用研究所
中国图象图形学学会
北京应用物理与计算数学研究所
主 编 李小文
编辑出版 《中国图象图形学报》编辑出版委员会
北京 9718 信箱 邮编 100101
电子信箱:jig@irsa. ac. cn
电话:010-64807995 010-82614429
网 址:www. cijig. cn
印刷装订 北京北林印刷厂
广告经营许可证 京朝工商广字第 0346 号
总 发 行 北京报刊发行局
订 购 全国各地邮局
国外发行 中国国际图书贸易总公司
(中国国际书店)
(北京 399 信箱 邮编 100044)

Superintended by Chinese Academy of Sciences
Sponsored by Institute of Remote Sensing Application,
CAS China Society of Image and Graphics
Institute of Applied Physics and Computational
Mathematics
Chief editor LI Xiaowen
Editor, Publisher Editorial and Publishing Board
of Journal of Image and Graphics
(P. O. Box 9718, Beijing 100101, China)
E-mail:jig@irsa. ac. cn
Distributed by Beijing Bureau for Distribution of Newspapers
and Journals
Domestic All Local Post Offices in China
Foreign China International Book Trading Corporation
(P. O. Box 399, Beijing 100044, China)
Printed by Beijing Beilin Printing House

ISSN 1006-8961 CN11-3758/TB CODE ZTTFXZ 国内邮发代号: 82-831 国外发行代号: M1406 国内定价: 45.00 元

Journal of Image and Graphics

(Monthly, Started in 1996)

Vol. 17 No. 6 June 2012

Contents

Review

- Age estimation by facial image: a survey Wang Xianmei, Liang Lingyan, Wang Zhiliang, Hu Siqian(603)
Expectation maximization method for parameter estimation of image statistical model Li Xuchao(619)

Image Processing and Coding

- Robust gradient driving image inpainting method Ye Xueyi, Wang Jing, Zhao Zhijing, Chen Huahua(630)
Fast image de-blocking by linear programming Jin Jianqiu, Liu Chunxiao, Wang Xun, Zhang Zhiyong(636)
Perfectly blind self-embedding robust quantization-based watermarking scheme in DWT-SVD domain Ye Tianyu(644)
Weak edge detection using Mean-shift filtering and histogram enhancement Ji Feng, Gao Xinbo, Xie Songyun(651)

Image Analysis and Recognition

- Gradient-pair constraint for structure lane detection Wang Yongzhong, Wang Xiaoyun, Wen Chenglin(657)
Contour detection based on multilevel inhibition Yan Chao, Zhang Jianzhou(664)
Image retrieval method based on local projection and block LBP feature Zou Bin, Pan Zhibin, Hu Sen(671)

Image Understanding and Computer Vision

- Visual novelty driven incremental and autonomous visual learning algorithm Qu Xinyu, Yao Minghai, Gu Qinlong(678)
Target localization algorithm for cooperative multi-soccer robots based on Hough space
..... Xu Jiaming, Xie Lun, Wang Zhiliang, Ni Shanchao(687)

Computer Graphics

- Inner FOV stitching algorithm of spaceborne optical sensor based on the virtual CCD line
..... Zhang Guo, Liu Bin, Jiang Wangshou(696)

Special Issue of the IFIC' 2011

- Anti-collusion fingerprinting scheme capable of tracing pirate Li Xiaoqiang, Zhang Huang, Zhao Yangyang, Wang Jingjing(702)
Infrared face recognition using LBP and discrimination patterns Xie Zhihua, Wu Shiqian, Fang Zhijun(707)
Recognize and retrieval complex events in real movies Du Jixiang, Guo Yilan, Zhai Chuanmin(712)
Off-line signature verification based on combination of direction feature and grid feature Yang Danfeng, Lv Yue(717)
Stereo image quality assessment based on visual attention Zhang Yan, An Ping, Zhang Qiuwen, Zhang Zhaoyang(722)
Low overhead of heterogeneous data exchange Zhao Kai, Zhao Zhengde(726)
Flexible prediction structure for multi-view video coding Zhang Yan, Cai Canhui(730)
Cross-layer feedback based adaptive coding for wireless video transmission
..... Wang Yaozhong, Zheng Shibao, Zhang Chongyang, Liu Bo(736)

中图法分类号: TP18 文献标识码: A 文章编号: 1006-8961(2012)06-0678-09

论文引用格式: 瞿心昱, 姚明海, 顾勤龙. 视觉陌生度驱动的增量自主式视觉学习算法[J]. 中国图象图形学报, 2012, 17(6): 678-686.

视觉陌生度驱动的增量自主式视觉学习算法

瞿心昱, 姚明海, 顾勤龙

浙江工业大学信息工程学院, 杭州 310023

摘要: 针对传统机器学习框架下设计智能机器人造成的视觉任务执行时学习主动性差、对不确定情况适应性差、知识与能力扩展性差等问题,立足近年来新提出的认知发育思想,提出一种由视觉陌生度驱动的增量自主式视觉学习算法。算法根据在线主成分分析(PCA)计算视觉陌生度,作为Q学习内部动机,以PCA子空间的更新作为知识的主动学习与积累,并由以视觉陌生度为内部动机的Q学习引导,使得机器人能根据所学知识对所“见”场景的陌生程度来决策下一步如何学习。实验结果表明,该算法具有自主探索与学习性能、主动引导机器人学习新知识的能力,以及在线、增量地获取积累知识并发育其智能的能力。

关键词: 认知发育;内部动机;视觉陌生度;在线主成分分析;Q学习

Visual novelty driven incremental and autonomous visual learning algorithm

Qu Xinyu, Yao Minghai, Gu Qinlong

College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China

Abstract: In intelligent robot design, the traditional machine learning paradigm is commonly used. However, the traditional methods cause problems in visual tasks such as low learning initiative, lack of adaptability with uncertainty and bad expansibility of knowledge and ability. According to the new research direction called cognitive development learning, a visual novelty driven incremental and autonomous visual learning algorithm is proposed, in which the internal motivation is defined as visual novelty which is calculated by online PCA. The autonomous learning and accumulation of knowledge is implemented in the form of updating PCA subspace, which is guided by internally motivated Q-learning using visual novelty. Equipped with the proposed algorithm, a robot makes the next learning decision by judging the novelty between learned knowledge and what is seen now. Experimental results show that the algorithm has the ability of autonomous exploring and learning, actively guiding the robot to learn new knowledge, acquire knowledge and develop intelligence online and in incremental manner.

Key words: cognitive development; internal motivation; visual novelty; online principal component analysis; Q-learning

0 引言

使智能机器人具备像人一样的智能一直是人工智能领域的目标之一。而目前的智能机器人大多采用传统的机器学习框架,针对视觉学习与识别任务,

其主要局限性有以下3方面^[1-3]:1)传统机器学习框架采用离线形式,学习和执行两个阶段相对独立,导致机器人无法像人一样在主动探索学习过程中同时执行任务,即在线实时地主动感知、处理、执行。一旦遇到新视觉样本就需要重新离线学习。同时,离线学习需要大量测试数据,人工而非机器人主动

收稿日期:2011-07-10;修回日期:2011-11-16

基金项目:国家自然科学基金项目(61070113);浙江省大学生科技创新活动计划(新苗人才计划)(2010R403071)

第一作者简介:瞿心昱(1985—),男,浙江工业大学控制理论与控制工程专业博士研究生,主要研究方向机器学习、模式识别、机器视觉。E-mail: qxy2010@163.com

地大量收集这些数据很麻烦且困难。2)传统学习方法都是基于人类编程者的知识表达,而不是智能体自我感知形成的知识,这样的知识无法预知不确定情况,无法从根本上改善智能体面对不确定环境与任务时的适应性。3)传统学习方法学习的视觉内容局限性很大,一旦任务稍有改变,需要重新学习,只能局限于较为简单或单一的事先制定好的视觉任务,无法在此基础上解决更为复杂的视觉任务,无法达到像人一样的智能发育。

由于采用传统的机器学习框架设计的机器人智能系统已越来越不能满足机器人对复杂环境与复杂高级智能的要求,近年来研究人员通过学科交叉,对智能产生的机理深入研究后,提出认知发育的新概念^[3-4]。利用认知发育学习方法来解决上述问题目前还处于初始阶段,存在大量有待探索的理论和实践问题^[1,5-6]。Brooks 等人^[7]以及 Asada 等人^[8],认为传统机器人的特点是:功能狭窄,任务特定,在理想环境下进行离线学习,通过结合神经科学的研究进展提出模拟脑皮层中额叶、感知、驱动等区域的计算模型来处理复杂环境中高级智能问题,强调机器人与教师、环境进行交互学习逐渐发展其智能。Kawamura 等人^[2]提出认知控制的概念,并借鉴人脑处理问题的方式,提出基于 Actor-Critic 网络人脑前额叶皮层工作记忆计算模型^[9]和基于情感驱动的决策制定方法^[10]。Weng^[11-12]提出直观无协方差计算增量主成分分析(CCIPCA)和增量分级回归树(IHDR)方法建立机器人感知映射和认知映射,使机器人知识和行为能力在回归树上不断发育。Pfeifer 等人^[13]强调机器人肢体和环境的作用来发育机器人智能,即具身认知思想。Oudeyer 等人^[14]从生物个体自主探索角度提出建立认知发育的内部动机系统思想,构造系统状态转移的预测误差学习机,计算预测误差作为内部动机来驱动系统智能发育。国内认知发育领域的研究还很少,文献[15]对 Weng 的自主心智发育进行跟踪研究,文献[16-17]仅对认知发育机器人进行了综述。目前众多研究方法都还只是很初步的探索,存在的问题主要有面对海量感知信息特别是高维视觉信息有维数灾难问题;方法中用到的如神经网络等传统方法,这些传统方法中存在如过拟合、在线单次训练时间依然很长等问题;如何在海量知识与行为空间高效搜索最优策略也是关键问题之一。因此,认知发育

学习领域还没有形成成熟统一的研究思路,理论还很不完善,是现在的热点研究方向。

本文研究目的在于对认知发育学习领域进行初步探索,在认知发育思想下提出适合于智能体移动机器人视觉学习与识别任务有效的新方法,在一定程度上解决采用传统机器学习方法设计机器人智能系统遇到的问题。借鉴生物学和心理学中内部动机的概念^[18-19],提出一种基于视觉陌生度内部动机的 Q 学习算法,引导机器人主动探索与增量在线学习视觉知识。该算法结合作者前期工作中提出的自适应子空间在线 PCA 算法^[20],根据在线 PCA 计算视觉陌生度,作为 Q 学习的奖励函数,同时以 PCA 子空间的更新作为知识的主动学习与积累,由基于内部动机的 Q 学习引导,使得机器人能够在线地自主探索与学习视觉知识,并根据其自身对环境的视觉陌生度来主动调整学习策略。最后通过移动机器人视觉场景学习实验来验证本文算法的自主探索与学习性能,主动引导机器人学习新知识的方式与能力,以及在线地、增量地获取积累知识并发育其智能的能力。

1 本文方法架构

基于视觉陌生度的增量自主式视觉学习算法在学习架构上与传统机器学习架构有所不同,主要特点在于本文算法是增量地、在线地学习,且借鉴生物学和心理学的相关概念,将视觉陌生度作为 Q 学习内部动机奖励函数设计,使得 Q 学习的引导更具有主动性。

当智能体(机器人)运行本文算法进行场景探索与学习时,根据当前读入场景图像与已学习场景知识来计算视觉陌生度,作为 Q 学习的内部动机奖励。这样智能体通过自身学习的知识与经验,主动决定下一步的探索与学习,一旦任务或环境出现变化,智能体可根据内部动机主动选择和改进下一步行为与策略,而不像传统的 Q 学习中人为定义奖励函数,导致任务或环境稍有变动,智能体就很难甚至无法适应,从而出现重大错误,必须通过人为修改奖励函数来实现调整。同时,知识积累与更新采用自适应子空间在线 PCA 的子空间更新来实现,这一增量地、在线地知识积累方式也是传统机器学习框架不具备的。图 1 所示为本文算法的架构示意图。

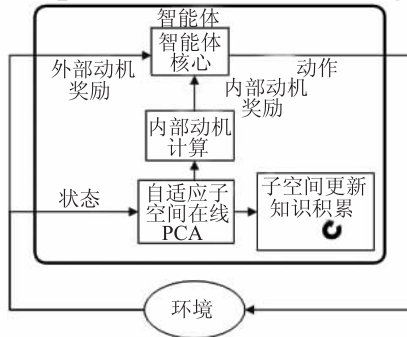


图 1 本文学习方法架构

Fig. 1 The architecture of the proposed learning method

2 基于视觉陌生度的内部动机

生物学范畴的食欲、求生等本能,以及心理学中描述的陌生度、好奇心、厌倦度等都是内部动机,生物系统通过它来维持生存和发展其智能。受其启发,建立这些内部动机的数学描述,就能将其应用于内部动机强化学习的计算模型中,引导智能体自主探索与学习,发育其智能。采用视觉陌生度来定义内部动机,将其用于 Q 学习的奖励函数设计。

设智能体感知输入为视觉输入,即可表示为定义在时间步长上的图像序列 $S_v = (x_1, x_2, \dots, x_t, x_{t+1}, \dots)$, 其中 $x_t = [x_{t1}, x_{t2}, \dots, x_{tm}]^T$ 为第 t 时间步长的 m 维图像输入的向量表示。视觉陌生度 N_v 定义为当前时间步长 t 的图像输入 x_t 与智能体对感知输入的经验知识预测值 \hat{x}_E 的差,即

$$N_v = \|x_t - \hat{x}_E\| \quad (1)$$

式中, \hat{x}_E 指智能体通过当前具备的知识或经验预测的当前感知输入。将此 N_v 作为内部动机函数,用于 Q 学习奖励函数设计。

但是,在智能机器人的应用中,通常输入图像的维数是很大的,从而造成计算量巨大,难以满足实时性的要求。PCA 方法在模式识别领域有着广泛的应用,其算法计算简单,降维特性使高维图像样本最大程度地转化为低维特征,大大减少计算量,因此采用 PCA 方法作为本文图像的降维是可行的。同时,本文对图像降维应用在图像序列的强化学习中可知,智能体的感知学习方式是增量的和在线的。而传统的 PCA 方法对数据的处理是批量的、离线的,因此需要增量地、在线地进行 PCA。采用自适应子空间在线 PCA 算法来实现图像增量地、在线地降维^[20]。最后

导出在线 PCA 意义下的视觉陌生度表达式。

设当前时间步长为 t , 已经历 $t-1$ 步 $t-1$ 个图像的计算。定义变量:当前子空间 U_t , 当前图像均值 \bar{x}_t , 当前图像投影系数矩阵 A_t , 当前时间步长 t 的图像输入 x_t , x_t 在当前子空间下的重建 \hat{x}_t 和投影 a_t , 更新后子空间 U_{t+1} , 更新后图像均值 \bar{x}_{t+1} , 更新后图像投影系数矩阵 A_{t+1} 。

当第 t 时间步长图像输入 x_t 到来时,通过计算与图像重建等价的低维投影系数向量集的 PCA 来获得新子空间,将 x_t 投影到当前子空间 U_t , 得到投影

$$a_t = U_t^T(x_t - \bar{x}_t) \quad (2)$$

重建图像

$$\hat{x}_t = U_t a_t + \bar{x}_t \quad (3)$$

计算重建误差向量

$$r_t = x_t - \hat{x}_t \quad (4)$$

计算 r_t 的欧几里德范数 $\|r_t\|$, 由于 r_t 正交于 U_t , 则由 r_t 计算新的基

$$\check{U}_t = \left(U_t \quad \frac{r_t}{\|r_t\|} \right) \quad (5)$$

在新基下构建新的系数向量矩阵

$$\check{A}_t = \begin{pmatrix} A_t & a_t \\ 0 & \|r_t\| \end{pmatrix} \quad (6)$$

式中, A_t 为前 $t-1$ 个图像输入的系数向量矩阵。对 \check{A}_t 计算 PCA, 得关于系数向量矩阵 \check{A}_t 的均值 $\check{\mu}_t$, 特征向量 \check{U}'_t 。此时可计算出在加入 x_t 后新的子空间下的图像投影系数向量矩阵

$$A_{t+1} = \check{U}'_t{}^T(\check{A}_t - \check{\mu}_t \mathbf{1}) \quad (7)$$

新图像均值

$$\bar{x}_{t+1} = \bar{x}_t + \check{U}_t \check{\mu}_t \quad (8)$$

以及新子空间

$$U_{t+1} = \check{U}_t \check{U}'_t \quad (9)$$

对每一个读入样本,将其分为 3 种情况,即新类、与已学习样本相似程度较小的已知类、与已学习样本相似程度较大的已知类。其中后两类均属于已知类,根据与已学习样本相似程度决定是否需要学习,并更新子空间。设定两个不同的阈值来对这 3 种情况作出区分: θ_{class} , 类间重建误差阈值; $\theta_{distance}$, 类内距离阈值,根据 θ_{class} 和 $\theta_{distance}$ 阈值判断,将子空间更新分 3 种方式。新子空间根据以上阈值判断 $\|r_t\|$ 大小,分别采用增维 $dim(U_{t+1}) = dim(U_t) + 1$, 保持维数 $dim(U_{t+1}) = dim(U_t)$ 和维持子空间不变 $U_{t+1} = U_t$ 的自适应更新方式来获得。

求出的 \mathbf{A}_{t+1} , $\bar{\mathbf{x}}_{t+1}$ 和 \mathbf{U}_{t+1} 作为下一图像输入 \mathbf{x}_{t+1} 进行增量的、在线的 PCA 降维的初始参数,如此迭代计算进行子空间更新,智能体对原始高维图像序列的在线感知转化为低维向量的在线感知,计算量大为减小。

在线 PCA 算法下,将 $\|\mathbf{r}_t\|$ 作为视觉陌生度,表示为 n_v ,则视觉陌生度的表达式为

$$n_v = \|\mathbf{r}_t\| = \|\mathbf{x}_t - \hat{\mathbf{x}}_t\| = \|\mathbf{x}_t - (\mathbf{U}_t \mathbf{a}_t + \bar{\mathbf{x}}_t)\| \quad (10)$$

可知每个时间步长 t 时的视觉陌生度 n_v 可以通过当前子空间 \mathbf{U}_t 计算,并以增量在线的形式计算。

3 内部动机 Q 学习增量自主学习算法

强化学习^[21]是一种无监督的在线学习,它采用“试错法”学习机制,通过感知环境状态和从环境中获得不确定奖赏值来学习动态系统的最优行为策略。因此,强化学习被广泛应用于智能系统的控制中^[22-23],被证明是一种通过经验逐步提高智能系统自主学习能力的有效计算方法^[24],其中 Q 学习^[25]是一种应用最广的强化学习算法。

经典的 Q 学习方法的奖励函数是根据特定任务人为设计的,如导航任务的奖励函数是针对导航任务设计的等,一旦任务改变,就需要人类编程者针对新任务重新设计奖励函数。这样就造成经典的算法无法满足认知发育智能的非特定任务要求,无法自主地增量地发育其智能、学习新经验。因此,若将内部动机加入经典的 Q 学习算法,扩展其奖励函数的表示,则 Q 学习奖励函数可以表示为内部动机奖励和外部动机奖励(即经典的奖励,如任务成功后奖励信号为 +1 等)的复合,即

$$R = \xi r_{in} + \eta r_{ex} \quad (11)$$

$$\xi, \eta \in [0, 1], \xi + \eta = 1$$

式中, r_{in}, r_{ex} 分别为内部动机函数和外部动机函数, ξ, η 分别为 r_{in}, r_{ex} 的权重。可以看到,当 $\xi = 0$ 且 $\eta = 1$ 时, $R = r_{ex}$ 即简化为经典 Q 学习中的奖励函数。若将第 2 节推导的视觉陌生度 n_v 作为内部动机函数 r_{in} , 即 $r_{in} = n_v$, 则基于内部动机 Q 学习的奖励函数为

$$R = \xi n_v + \eta r_{ex} = \xi \|\mathbf{r}_t\| + \eta r_{ex} = \xi \|\mathbf{x}_t - \hat{\mathbf{x}}_t\| + \eta r_{ex} \quad (12)$$

经典 Q 学习迭代公式

$$Q(s, a) \leftarrow Q(s, a) + \alpha [r + \gamma \max_{a'} Q(s', a') - Q(s, a)] \quad (13)$$

修改为基于内部动机 Q 学习的迭代公式

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [R + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)] \Leftrightarrow Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [(\xi \|\mathbf{x}_t - \mathbf{x}_t\| + \eta r_{ex}) + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)] \quad (14)$$

式中, s_t 是强化学习智能体的当前时间步长感知状态,可以表示为

$$s_t = \mathbf{x}_t \quad (15)$$

a_t 是强化学习智能体的当前时间步长动作,也就是智能体在当前感知下采取何种行为去和环境交互,即采用何种行为去主动感知其视觉图片输入; α 为学习率,表示学习收敛速度, γ 为折扣因子,越大表明后步骤奖励对本步骤影响越大。

基于此视觉陌生度内部动机 Q 学习的增量自主式视觉学习算法的具体步骤如下:

1) 初始化 Q 学习的值函数 $Q(s_t, a_t)$, 并选定学习率 α , 折扣因子 γ , 以及内部动机和外部动机奖励函数的权重 ξ, η 。

2) 读入 t 时间步长图片,通过式(2)—(4)计算其投影,重建和残差向量 \mathbf{r}_t 。

3) 计算视觉陌生度内部动机 $n_v = \|\mathbf{r}_t\| = \|\mathbf{x}_t - \hat{\mathbf{x}}_t\|$ 。

4) 如果 $\|\mathbf{r}_t\| > \theta_{class}$, 则通过式(5)—(9)的在线 PCA 算法计算新子空间 \mathbf{U}_{t+1} , 且 $dim(\mathbf{U}_{t+1}) = dim(\mathbf{U}_t) + 1$, 接着执行步骤 8), 如果 $\|\mathbf{r}_t\| < \theta_{class}$, 则执行步骤 5)。

5) 计算新图片投影和已有所有图片投影的欧氏距离 $d_{Euclidean}$ 。

6) 如果 $d_{Euclidean_min} > \theta_{distance}$, 则通过式(5)—(9)在线 PCA 算法计算新子空间,舍去其最后一个分量,使 $dim(\mathbf{U}_{t+1}) = dim(\mathbf{U}_t)$, 且执行步骤 7); 如果 $d_{Euclidean_min} < \theta_{distance}$, 直接执行步骤 7)。

7) 将 $d_{Euclidean_min}$ 对应的已学习图片类别作为识别输出。

8) 探索并选择智能体的动作, $\max_{a_{t+1}} Q(s_{t+1}, a_{t+1})$, 由此引导智能体主动感知下一时间步长 $t+1$ 时的视觉感知信号,即图片输入。

9)更新智能体行为 Q 值函数,即

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [(\xi \|x_t - x_t\| + \eta r_{ex}) + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)]$$

10)由此不断更新智能体主动感知视觉图片输入的行为方式和能力。

11)返回步骤 2),进入下一时间步长,智能体继续感知视觉图片输入。

算法流程如图 2 所示。



图 3 实验移动机器人 AS-R

Fig. 3 Mobile robot AS-R used in the experiment

实验中,机器人作为本文算法运行智能体的具体对象,由 Sony 二自由度云台摄像头及 Osprey 视频采集卡采集的视觉图像作为算法的图像输入,其原始大小为 320×240 像素。为了减小计算量,提高实时性,将其缩小至 60×45 像素,即每一时间步长机器人智能体在本文算法下的图像输入为 2 700 维,且将彩色图像转换成灰度图像。

实验在实验室内进行,让机器人学习与识别在同一个位置的 3 个不同角度下所输入的场景图像,验证本文基于视觉陌生度的增量自主式视觉学习算法的自主性、增量性。实验平面示意图如图 4 所示,黑色圆代表机器人,初始状态为面向过道,实验设计机器人行为状态:不动,左转 90° ,右转 90° ,其中转动行为需要恢复初始状态(面向过道),以便下一定义行为能顺利执行。机器人分别读取图像输入场景为过道、门廊、办公室。实验验证基于视觉陌生度的增量自主式视觉学习算法引导机器人选择 3 种行为,主动探索与学习这 3 个场景,同时又不不断更新主动学习行为,积累新知识。图 5 为机器人实际获取的 3 个场景的图像样本。针对学习率 α ,由于本文提出方法的学习是在线实时的,因此在参考经验数值的基础上,选取较大学习率,保证一定精确性的同时使 Q 学习快速收敛; γ 参考经验值,取较大的 γ 值。实验中 Q 学习参数 $\alpha = 0.6, \gamma = 0.8$ 。

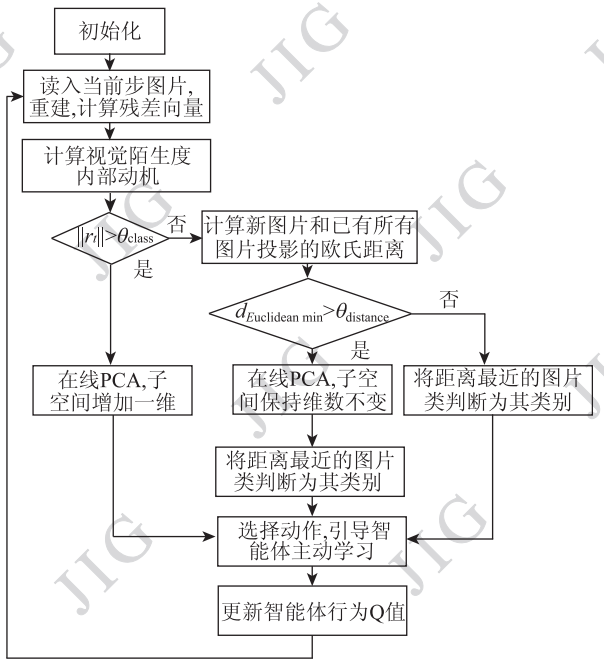


图 2 本文算法流程图

Fig. 2 Flow chart of the proposed algorithm

4 实验与讨论

4.1 实验设计

针对本文提出的基于视觉陌生度的增量自主式视觉学习算法,设计实验将其应用在智能移动机器人视觉场景学习与识别上,验证算法的自主探索与学习性能,主动引导机器人学习新知识的方式与能力,以及在线地、增量地获取积累知识并发育其智能的能力。实验硬件平台为上海广茂达智能移动机器人研究版 AS-R, Sony 二自由度云台摄像头及 Osprey 视频采集卡,如图 3 所示。学习算法程序的软件平台采用 Visual C++ 6.0 及 OpenCV 机器视觉开源库。

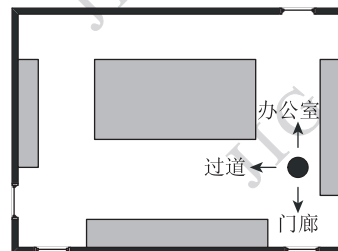


图 4 机器人实验环境

Fig. 4 The robot experiment environment



图 5 机器人学习与识别的 3 个场景样本

Fig. 5 Three scene samples for learning and recognition

4.2 实验结果与分析

实验 1 将机器人置于图 4 所在位置,让其在本文算法的引导下自己探索学习 3 个场景,并记录机器人对各场景探索情况的变化趋势,分析其随学习步数的变化规律以及内部动机在引导机器人自主探索中的作用,并且与传统的强化学习方法进行比较。实验使机器人探索学习 500 个时间步长,并将其分为 5 个时间段进行分析,每个时间阶段含 100 个时间步长的学习过程。分析机器人从初始状态开始,通过 500 个时间步长的主动探索来学习 3 个场景的情况。如图 6 所示,采用本文算法,机器人根据视觉陌生度作为主动探索选择场景学习的驱动信号。初始状态时,机器人读入的是“过道”场景,并且由于一开始机器人对“过道”场景陌生度较高,因此在初始几个时间阶段中,机器人选择此场景较多,同时随着“过道”陌生度降低,其他两个场景也越来越多的得到探索与学习,最终 3 个场景的陌生度趋于一致,即都已学习足够,算法对这 3 个场景的选择也趋于均等。

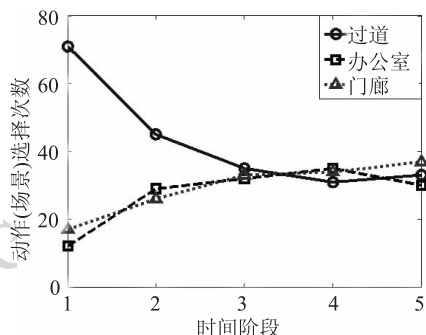


图 6 本文算法探索与学习场景过程

Fig. 6 Exploring and learning process of the proposed algorithm

将本文算法与传统的 Q 学习算法进行同一实验比较,其中传统 Q 学习采用人为给予奖励信号,即外部动机奖励信号。实验中给“过道”场景以最大正奖励,记录机器人场景选择学习过程如图 7 所示。可见随着学习步数的增加,机器人选择“过道”次数显著增加。但这种选择学习机制是依靠外部的、人为的奖励,无法做到依靠智能体自身判断去执行,一旦任务的稍许改变,最大正奖励可能就不能给“过道”场景了,而调整奖励函数就必须重新人为设计,无法满足智能体主动探索、调整和学习以及发育其智能的要求。

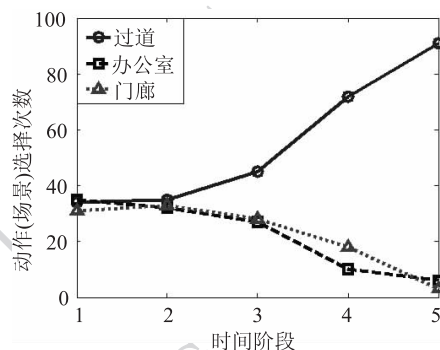


图 7 传统方法探索与学习场景过程

Fig. 7 Exploring and learning process of the traditional algorithm

实验 2 为了进一步验证本文算法主动探索和引导机器人学习新知识的能力,重复以上实验,并在机器人对 3 个场景已学习充分后,在“过道”场景中放入一个新物体,记录机器人接来自主探索与学习的过程。本文选择在第 301 个时间步长时在过道中放入此新物体,如图 8 所示。记录放入物体后的学习过程如图 9 所示,可见,当“过道”场景放入物

体后,此场景发生变化,和机器人已学习的 3 个场景都不相同,不再为机器人所“熟悉”,即“过道”场景的陌生度增加。算法由此增加的陌生度再次引导机器人选择“过道”场景学习,第 4 个时间阶段选择“过道”的动作有了显著增加,把“过道”学习充分后,3 个场景的选择再次趋于均等。

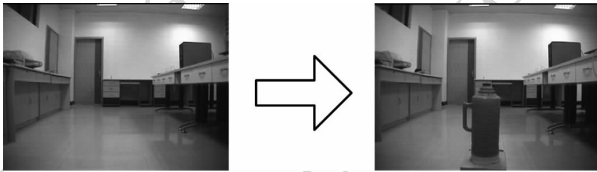


图 8 “过道”场景放入新物体

Fig. 8 New object put in the pathway scene

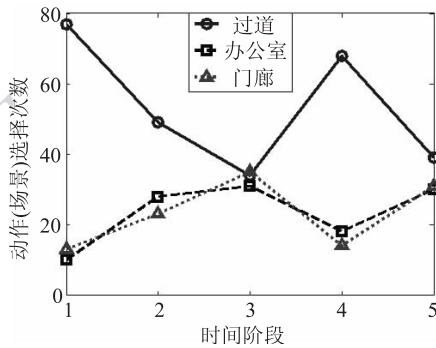


图 9 本文算法新物体主动探索与学习

Fig. 9 Autonomous exploring and learning of new object of the proposed algorithm

将外部动机奖励(即人为设定的奖励)考虑在内,分析内外部动机结合下新知识探索学习的情况。根据认知科学相关研究,外部动机奖励(或惩罚)对

于个体学习驱动影响要大于内部动机,即内外部动机权重关系可表示为 $\xi < \eta$, 实验中令 $\xi = 0.2, \eta = 0.8$, 重复以上新物体探索学习实验,并给予“办公室”场景最大外部正奖励。学习过程中各场景选择学习数变化如图 10 所示,可见,在内外部动机结合的情况下,场景更趋向于选择给予外部动机奖励。在第 4 个时间段时陌生度增加的“过道”场景选择数略有增加,说明算法依然能主动学习陌生新事物。因此,结合了内外部动机的学习算法在传统方法基础上可以更主动的学习新知识。

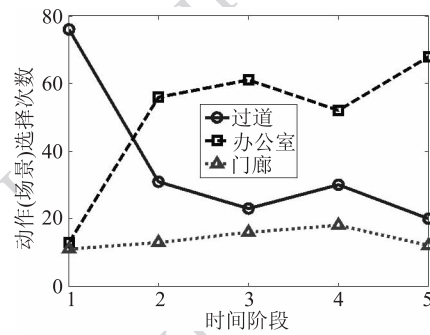


图 10 内外部动机结合新物体主动探索与学习

Fig. 10 Autonomous exploring and learning of new object with internal motivation and external motivation

实验 3 算法中的增量 PCA 部分使得机器人对场景的学习能实现最简单的发育能力,实验中以一个新场景的学习过程来研究算法对场景的增量学习性能,这里选择“门廊”场景,并通过增量 PCA 重建误差来描述发育学习情况,重建误差越小,则学习越充分。图 11 为“门廊”场景中的 5 个采样点时间步



图 11 “门廊”学习过程中 5 个采样点重建与重建误差

Fig. 11 Reconstruction and reconstruction error of the five sample points in hallway learning process

长上的原始读入图像、增量重建图像与重建误差的比较。如图12所示,选择10个采样点进行重建误差分析,在内部动机(即陌生度)的引导下,随着学习过程的推进,“门廊”场景不断被学习,重建误差不断减小,直至学习充分,这体现了本文算法的简单发育性能。

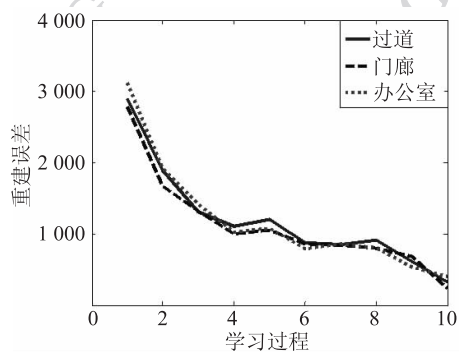


图12 3个场景学习的重建误差变化趋势

Fig. 12 Reconstruction error trend in the three scenes learning

以同一个场景学习过程重建误差的减小过程来描述本文算法的学习速率,并和传统的Q学习进行比较,如图13所示。可见,本文基于内部动机Q学习算法在最初时间步长时比传统Q学习重建误差下降快,即采用本文算法引导机器人自主探索与学习达到学习充分状态比传统Q学习要略快,这也验证了本文算法学习效率较高。

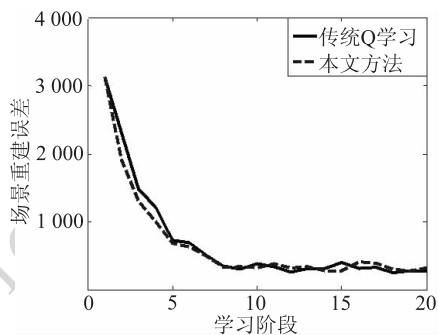


图13 本文算法学习效率比较

Fig. 13 Comparison of learning efficiency of the proposed algorithm with the traditional method

5 结论

针对传统机器学习方法与框架设计智能机器人存在的问题,结合近年来逐渐成为热点的认知发育学习思想,提出一种适合于智能移动机器人室内视

觉场景探索与主动学习的认知发育算法。算法通过结合基于视觉陌生度的内部动机Q学习和自适应子空间在线PCA算法,将增量PCA子空间的更新,由基于内部动机的Q学习引导,使得机器人能根据所学知识与所“见”场景的陌生程度决策下一步如何学习。通过实验验证了算法的自主探索与学习性能,主动引导机器人学习新知识的方式与能力,以及在线地、增量地获取积累知识并发育其智能的能力。本文算法在视觉陌生度的定义中带有一定的特殊性,如何更一般地设计内部动机是今后的研究重点之一,另外,随着知识地增多,仅用PCA子空间作为知识积累是不够的,探索更适合海量知识存储和搜索的发育算法也将是今后的研究重点。

参考文献(References)

- [1] Vernon D, Metta G., Sandini G. A survey of artificial cognitive systems: Implications for the autonomous development of mental capabilities in computational agents[J]. IEEE Transactions on Evolutionary Computation, 2007, 11(2): 151-180.
- [2] Kawamura K, Gordon S. From intelligent control to cognitive control[C]// Proceedings of the 11st International Symposium on Robotics and Applications, Budapest, Hungary: World Automation Congress, 2006:1-9.
- [3] Weng J, McClelland J, Pentland A, et al. Autonomous mental development by robots and animals [J]. Science, 2001, 291(5504): 599-600.
- [4] Asada M, Hosoda K, Kuniyoshi Y, et al. Cognitive developmental robotics: A survey [J]. IEEE Trans. on Autonomous Mental Dev., 2009, 1(1):12-34.
- [5] Steil J J, Wersing H. Recent trends in online learning for cognitive robot[C]// Proceedings of the European Symposium on Artificial Neural Networks. Bruges, Belgium: D-side Publisher, 2006:77-87.
- [6] Alami R, Chatila R, Clodic A, et al. Towards human-aware cognitive robots [C]//Proceedings of the Fifth International Cognitive Robotics Workshop (The AAAI-06 Workshop on Cognitive Robotics). Stanford, USA: AAAI Press, 2006.
- [7] Brooks R, Breazeal C, Irie R, et al. Alternative essences of intelligence[C]// Proceedings of the American Association of Artificial Intelligence. Madison, Wisconsin: MIT Press, 1998.
- [8] Asada M, MacDorman K, Ishiguro H, et al. Cognitive developmental robotics as a new paradigm for the design of humanoid robots[J]. Robotics and Autonomous Systems, 2001, 37(2-3): 185-193.
- [9] Kawamura K, Gordon S, Ratanaswasd P, et al. Implementation of cognitive control for a humanoid robot [J]. International Journal of Humanoid Robotics, 2008, 5: 547-586.

- [10] Gordon S M, Kawamura K, Wilkes D M. Neuromorphically inspired appraisal-based decision making in a cognitive robot[J]. IEEE Trans. on Autonomous Mental Dev., 2010, 2(1):17-39.
- [11] Weng J. Developmental robotics: Theory and experiments[J]. Int. J. Humanoid Robot, 2004, 1(2): 199-235.
- [12] Weng J. A theory for mentally developing robots [C]// Proceedings of the IEEE Second International Conference on Development and Learning, MIT. Cambridge, MA: IEEE Computer Society Press, 2002: 131-140.
- [13] Pfeifer R, Bongard J C. How the Body Shapes the Way We Think: A New View of Intelligence[M]. Cambridge, MA: MIT Press, 2006.
- [14] Oudeyer P Y, Kaplan F, Hafner V. Intrinsic motivation systems for autonomous mental development[J]. IEEE Transactions on Evolutionary Computation, 2007, 11(2): 265-286.
- [15] Gao Y, Chen D Y, Zhang L M. An exploration of autonomous developing robot with real time vision learning[J]. Journal of Fudan University, 2005, 44(6): 964-970. [高颖, 陈东岳, 张立明. 一种带有实时视觉特征学习的自主发育机器人探索[J]. 复旦大学学报(自然科学版), 2005, 44(6): 964-970.]
- [16] Qu X Y, Yao M H. Cognitive robotics: a survey [C]// Proceedings of Chinese Intelligence and Automation Conference. Nanjing, China: Southeast University, 2009: 718-725. [瞿心昱, 姚明海. 认知机器人研究综述[C]// 中国智能自动化会议论文集. 南京: 东南大学, 2009: 718-725]
- [17] Yu H L, Zhu C M, Liu H B, et al. A survey on developmental robotics[J]. Caai Transactions on Intelligent Systems, 2007, 2(4): 34-39. [于化龙, 朱长明, 刘海波, 等. 发育机器人研究综述[J]. 智能系统学报, 2007, 2(4): 34-39]
- [18] Deci E, Ryan R. Intrinsic Motivation and Self-determination in Human Behavior[M]. New York: Plenum, 1985.
- [19] Schultz W, Dayan P, Montague P. A neural substrate of prediction and reward[J], Science, 1997, 275:1593-1599.
- [20] Yao M H, Qu X Y. Hand gesture recognition based on online pca with adaptive subspace[J]. Pattern Recognition and Artificial Intelligence, 2011, 24(2): 299-304. [姚明海, 瞿心昱. 基于自适应子空间在线 PCA 的手势识别[J]. 模式识别与人工智能, 2011, 24(2): 299-304.]
- [21] Mitchell T M. Machine Learning[M]. Beijing: Machine Press, 2004: 263-280.
- [22] Xiao N F, Nahavandi S. A reinforcement learning approach for robot control in an unknown environment [C]// Proceedings of the IEEE International Conference on Industrial Technology. Bangkok, Thailand: IEEE Press, 2002:1096-1099.
- [23] Wang Y C, M. Usher J. Application of reinforcement learning for agent-based production scheduling[J]. Engineering Application of Artificial Intelligence, 2005, 18(1): 73-82.
- [24] Gao Y. Progress of Reinforcement Learning Research [M]// Machine Learning and its Application. Beijing: Tsinghua University Press, 2006: 116-134. [高阳. 强化学习研究进展[M]//机器学习及其应用. 北京: 清华大学出版社, 2006: 116-134]
- [25] Watkins C, Dayan P. Q-Learning [M]// Machine Learning, Boston: Kluwer Academic Publishers, 1992, 8: 279-292.