

从AlphaGo到BetaGo ——基于任务可完成性分析的定性人工智能的定量实现

苏剑波[†], 陈叶飞, 马哲, 黄瑶, 向国菲, 陈若冰

(上海交通大学 电子信息与电气工程学院, 上海 200240)

摘要: AlphaGo程序利用深度学习算法和蒙特卡洛树搜索算法在围棋领域取得了突破性进展, 用定量分析技术实现了围棋这一传统定性式的任务. 此次突破, 实现了对定性人工智能的定量研究, 这对人工智能体的研究具有重要借鉴意义. 对于人工智能体而言, 执行任务前需要考虑任务的可完成性, 对下棋任务而言, 任务目标是取得胜利, 因此, 本文先从任务可完成性角度出发, 分别从特征完备性、表征空间构建及基于表征空间的搜索角度分析AlphaGo程序. 其次, 人工智能体在任务完成过程中, 不可避免地受到各种扰动的影响, 对AlphaGo而言, 本质是对人下棋过程的建模, 因此, 本文从抗干扰的角度出发, 分析了AlphaGo的缺陷. 再者, 人工智能的研究是人类用科学技术的方式模拟大脑活动的过程. AlphaGo所体现出的围棋思想与人类棋手的围棋美学之间的差异, 也是定量分析与定性描述之间的差异. 因此, 本文从美感评价角度对AlphaGo进行了分析和展望. 通过上述三个角度, 本文诠释了AlphaGo程序所包含的原理以及对量化分析定性人工智能体研究的借鉴意义. 本文认为, AlphaGo虽然取得了里程碑式的进展, 但在定性描述(如: 美学, 艺术)以及系统未知扰动方面仍存在大量问题值得研究. 人工智能的跨越式发展, 即从Alpha级别提升至Beta级别, 应该包含对事物定性分析的能力. 最后, 希望人工智能算法的研究工作者通过本文能更关注于挖掘定性描述与定量分析之间的关联, 并进一步将人工智能算法提升至BetaGo乃至更高.

关键词: AlphaGo; 特征完备性; 表征空间; 任务可完成性; 抗干扰; 美感评价; 人工智能

中图分类号: TP273 文献标识码: A

From AlphaGo to BetaGo — Quantitative realization of qualitative artificial intelligence based on task realizability analysis

SU Jian-bo[†], CHEN Ye-fei, MA Zhe, HUANG Yao, XIANG Guo-fei, CHEN Ruo-bing

(School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: AlphaGo, using deep learning and Monte Carlo Tree searching algorithms, has achieved great progress in the game of Go, which realized the qualitative task by quantitative analysis. Such progress has realized the quantitative research of qualitative AI, which has a significant reference value for researchers. For AI, task realizability should be taken into consideration before executing the task. The task goal of the game of Go is winning of the game. Therefore, AlphaGo is firstly analyzed in the aspect of task realizability, including feature completeness, establishment of representation space and searching method based on representation space. Secondly, during the process of task executing, AI will confront various disturbance inevitably. The essence of AlphaGo is modeling the process of playing chess of human. Hence, the drawbacks of AlphaGo is analyzed in the aspect of disturbance rejection. Thirdly, the research of AI is a simulation of human brain activity by scientific technology. The beauty evaluation difference between AlphaGo and human player reflects the difference between quantitative analysis and qualitative description. Therefore, AlphaGo is analyzed and prospected in the aspect of beauty evaluation. In this paper, the principle of AlphaGo and the significant reference value of quantitative realization of qualitative AI are annotated through the three aspects mentioned above. Though AlphaGo has achieved remarkable progress, we consider that there are still plenty of problems remain to be studied in aspects of qualitative description (e.g. beauty evaluation, art) and unknown disturbance of system. The significant progress of AI should contains the ability of qualitative analysis, which is an leap from Alpha level to Beta level. At last, the analysis in this paper is supposed to make AI researchers pay more attention to dig the relationship between qualitative description and quantitative analysis and enhance the AI to BetaGo and higher level.

Key words: AlphaGo; feature completeness; representation space; task realizability; disturbance rejection; beauty evaluation; artificial intelligence

收稿日期: 2016-06-24; 录用日期: 2016-12-09.

[†]通信作者. E-mail: jbsu@sjtu.edu.cn; Tel.: +86 21-34204276.

本文责任编辑: 胡跃明.

国家自然科学基金项目(61533012, 61521063)资助.

Supported by National Natural Science Foundation of China(61533012, 61521063)

1 引言(Introduction)

不久前, AlphaGo与李世石的人机围棋对战以比分4:1结束, 此次胜利体现了以深度学习为代表的人工智能算法的强大, 同时彰显了以定量化手段分析并研究定性行为的可行性. 随后, 在非正式线上比赛中, 谷歌深智团队(DeepMind)以Master为名替AlphaGo建立账号, 获得了60胜0负1平的战绩. 一直以来, 围棋被视为人类所有棋类游戏中最高难度的体现, 是人类定性分析智慧的体现. 这主要因为围棋拥有较大的棋盘, 采用传统方式遍历所有可行解的方法并不适用, 且其局势分析复杂, 很难设计一套人工特征(handcraft feature), 并据此确立出一个简单的估值函数. AlphaGo的算法^[1]充分利用了计算机程序定量分析的特点和能力, 采用深度卷积网络和蒙特卡洛树两大核心算法, 综合考虑落子概率即监督学习策略网络(supervised learning policy network, SLPN)、当前形势下的胜率即估值网络(value network, VN)和传统快速走棋策略(rollout policy, RP)形成估值函数, 从对弈的结果来看, 这套方案成功将定性人工智能的任务转化为计算机定量求解的过程.

近两年, 谷歌公司的深智团队发布了一系列令人瞩目的研究成果, 从基于Atari视频游戏的深度强化学习算法^[2]到围棋人工智能(AlphaGo). 这些人工智能算法的共性是都采用深度学习(deep learning)与强化学习(reinforcement learning)^[3-5]相结合的训练方式, 从而获得在该领域胜于人类智能的表现结果. 一般而言, 深度学习算法需要大量的带标签的训练样本来构建网络, 且深度卷积神经网络(deep convolution neural network, DCNN)本身不具备时序性, 对下棋这类有时序性的问题缺乏有效的处理手段. 近年来已有学者将深度卷积神经网络(DCNN)与具有时序性质的循环神经网络(recurrent neural network, RNN)相结合, 在无标图像描述自动生成中取得了应用^[6]. 强化学习算法是以结果为导向, 通过最大化累积奖励的方式来学习得到最优策略. 强化学习的引入, 使得深度学习算法以胜率为导向, 进一步优化网络层参数. 同时, 采用不以胜负为导向的落子概率初始化深度学习网络, 使其具备基本的特征抽取能力, 以深度卷积神经网络初始化一个基本的特征集. 初始得到的深度学习网络不以胜负结果为目的, 仅仅完成状态的感知, 即状态到动作的映射. 最终, 通过强化学习算法将这个特征集合不断完善, 从而形成较为完备的特征集合. 从某种意义上说, 深度学习与强化学习算法解决了抽象特征抽取的工作, 将视觉信息通过深度卷积网络不断学习并得到较完备的特征集合. 该特征集合的完备性是针对特定任务而言, 即是否能够足以支持一个任务的完成, 就围棋而言即达到胜利的目标. 而蒙特卡洛搜索过程可视为在该完备特征集所构成的高维空间中搜索估

计代价的过程.

人工智能体是人类用机电方式的自我克隆, 机电方式的自我克隆不仅仅体现在肢体上面, 更体现在智能方面. 人工智能是西方定量科学发展的产物, 而围棋则是东方定性化的一个典型体现. 这次人类冠军棋手与人工智能对弈, 不仅仅是西方现代科技与寻求定性分析的东方智慧之间的博弈, 更是采用定量分析算法演绎传统定性分析智能的一次尝试. 除了围棋, 东方定性分析中还有很多暂时没有突破的研究方向, 例如对书法美学的研究. 其中已有学者对书法中的美学进行了量化分析的尝试^[7]. 虽然R. Salakhutdinov通过贝叶斯学习的方式让计算机自主生成人类书写的笔画^[8], 但东方的软笔书法不同于西方硬笔, 硬笔笔画的宽度深浅缺少变化, 且字母个数有限, 不论是花体还是印刷体通过简单训练即可书写漂亮, 而东方文字具有象形意义和结构美感, 不同的书法体具有不同的艺术效果. 而且东方文字数量庞大, 笔画写法众多, 间架结构繁杂. 在软笔书法上, 力道的掌握、下笔行笔收笔的技巧, 都是人类经过大量刻苦练习和老师的指导缓慢形成的肌肉记忆, 难以通过知识库的方式形成规律教给机器人. 其中的训练与指导和深度学习与强化学习具有内在的联系性. 与此同时, C. Li等^[9]和M. Bethge等^[10]试图借助深度学习算法来研究人创作艺术绘画的方式. 通过将绘画分成内容(content)与艺术样式(style), 成功将两幅不同类型的绘画进行合成, 实现了抽象图形的定量化分析. 人工智能除了实现定量运算外, 更应包含对事物定性分析的能力, 本文将拥有这一功能的算法称之为Beta级. 所以, 此次AlphaGo程序对人工智能算法的提升具有很强的借鉴意义.

从李世石角度而言, 第4盘的第78手直接导致AlphaGo程序判断失误, 从某种意义上说, 正是其系统模型对外生干扰不稳定的体现. 人工智能的本质是人类试图以科技为手段, 建造一种可以模仿人类思想和行为的算法机制. 而人类思想存在一定随机性即所谓的灵感, 此类随机性以控制理论的角度来看, 应属于系统扰动一类, 其是否归属于可建模或不可建模的扰动还值得探讨. 对于人工智能系统而言, 能否有效处理此类扰动对模拟人类思想具有十分重要的意义.

从AlphaGo角度而言, 尽管自我对弈过程能产生大量棋谱从而对深度卷积网络的训练起到推进作用, 但并没有真正意义上实现自我学习. 首先, 围棋是一个博弈的过程, 这一特性使得AlphaGo算法能够实现自我迭代更新. 其次, 围棋任务中存在明确量化的目标即提高胜率, 从而使得深度卷积网络能够实现从图形图像中抽取对任务完备的特征并与胜率或落子概率建立起联系. 再者, 在输入图像方面, AlphaGo算法不仅仅使用棋盘落子状态图, 还加入了人对棋盘的先验分析结果, 从而生成49通道的输入图像(在训练

策略网络时使用48个通道,在训练估值网络时使用49个通道,增加的通道用于表示目前玩家执子的颜色).这也就意味着,当前的人工智能依然离不开人对任务的判断和假设,不能自主地从任务中获取这些信息,对AlphaGo算法而言,依然需要人从底层数据中根据经验进行初步的特征抽取.然而实际问题中,能有效符合上述3点的情况并不常见.另外,AlphaGo算法在整个过程中有两处近似估计.第1,在使用蒙特卡洛模拟胜率最高路径的过程中,对实际胜率进行了近似估计.即假设蒙特卡洛搜索结果近似等于实际穷举得出的胜率.本文从表征空间的角度认为:这一近似导致代价估计函数在完备特征集所构成的高维空间中错误地给出了搜索代价.第2,使用深度卷积网络对落子概率进行估计,由于是博弈问题,所以包含了对对手落子的估计.即假设AlphaGo算法对对手的落子概率估计是正确的,而实际系统训练建模过程中,李世石的历史棋谱在整个训练样本中屈指可数.据此,本文从抗干扰的角度分析系统建模和系统的抗干扰能力.同时,AlphaGo算法假设深度卷积网络对当前局势的胜率估计是正确的.即从特征完备性的角度而言,深度卷积网络所提取的特征对胜率估计这个任务而言是完备的.综上所述,可以将第4盘的失误归因于上述假设及估计的失效.正因如此,本文将从系统建模角度出发,对系统抗干扰能力提出了一些见解;从对任务可完成性的角度(即特征完备性和表征空间)来分析这一搜索过程.

对于深度学习、强化学习方法的研究和理解,有助于推动各个领域的发展.本文主要将从以下角度来诠释AlphaGo算法所体现的人工智能:1)任务可完成性角度,即特征完备性以及表征空间下AlphaGo算法剖析.2)抗干扰角度,即从抗干扰能力的角度分析AlphaGo算法以及相关人工智能系统.3)美感评价角度,即从人类美学角度分析AlphaGo算法及其意义.

2 任务可完成性下的AlphaGo原理分析(Analysis on AlphaGo under view of task realizability)

对于人工智能体而言,由于受到各种内外部约束,完成任务的能力有限.因此,在执行给定的任务之前,都应考虑当前系统的状态是否能够完成给定的任务,只有当任务可完成时,进行任务规划才有意义.此外,针对基于视觉信息的任务,如:机器人视觉伺服控制^[11-15],如何构建完备的特征集来提供完成任务所需的有效信息也是至关重要的.其中,完备特征集是构建表征空间(representation space, R-Space)^[16-17]的基础,也是分析任务可完成性的必要条件,对人工智能系统的任务规划和实现至关重要.而目前缺乏基于上述问题的研究.AlphaGo程序作为人工智能体的一部分,将围棋图像的感知与决策相映射,在视觉交互方

面具有较高的研究价值,本文将从特征完备性及任务可完成性的角度剖析AlphaGo算法的原理及意义.

2.1 特征完备性(Feature completeness)

人工智能体完成任务的过程就是信息特征变化的过程^[18].人工智能体的每一个操作过程,必须能通过特征表达出任务完成状态和执行情况,即一定存在至少一个信息特征集合,能定义任务完成的不同阶段和状况^[19-20],这个特征集合,就可以成为面向特定任务的完备特征集合.一般来说,完备特征集合不是唯一的^[21].技术路线上,需要设计视觉交互过程和优化控制算法^[22],从已有的不完备的特征集合出发,寻找邻近的完备特征集合,确定与已有特征集合互补成完备的特征集,并给出有效的寻找策略和评价手段,即如何根据任务来确定一个完备特征集,以及特征集合不完备时,如何向完备特征集合过渡^[23-24].

在视觉伺服控制领域,文[25]以机器人视觉伺服任务为研究对象,从特征完备性角度对视觉伺服系统进行分类,并分析指出:从不同结构的系统所采集到的图像中提取到的特征具有不同的完备性.在模式识别领域,文[26]以构建大量冗余稀疏字典的方式重构信号,并提出了过完备字典的概念(over complete dictionary).此后,涌现出大量基于稀疏的字典学习方法,此类方法均试图通过构建一个完备的特征字典即完备特征集以实现不同的任务:信号重构^[27]、图像复原^[28-30]、人脸识别^[31-36]、图像分类^[37]、去噪声^[38]、特征选择^[39-40]等.在早期的基于手工特征(handcraft feature)的研究中,大量特征算子被提出,并用统计的方式模拟人对物体的匹配和分类,如:LBP^[41-44]、LPQ^[45-46]、Gabor^[47-51]、HOG^[52-53]、SIFT^[54]、VLAD^[55]等.其中,S.Lazebnik通过spatial pyramid matching(SPM)^[56]的方法对特征进行多层次的抽取,取得了不错的进展,并激发了最大池化(max-pooling)的思想及相关研究^[57].随后,在浅层特征的基础上提出了关于中间层特征(mid-level feature)学习^[58]的方法.但手工特征需要大量人工设计和介入,难以形成完备的特征集合^[59].与此同时,在对神经网络的研究中,G.Hinton等在2006年首次提出了深度学习的概念^[60].文[61]提出了经典的深度学习方式:采用非监督的方式对神经网络逐层(layer-wise)贪婪训练,之后采用有监督的方式再次训练(fine-tune).其后,深度学习在语音及视觉领域均有不错的应用和发展^[62-69].最具代表性之一的研究工作是Krizhevsky在ImageNet上对深度卷积神经网络(deep convolution neural network)的应用^[70].对深度卷积置信网络的研究表明^[63-64],在第1层网络中,主要提取的特征是一些基本的纹理信息^[71],如:点、线等;在第2层网络中,主要提取到的是些物体局部信息,该信息是对前一层纹理信息的组合;在第3层网络中,主要提取到的是更进

一步的整体信息, 即更接近物体外貌的信息^[72-74].

上述方法从本质上, 都试图通过一种学习算法, 构建一个完备的特征集合, 从而实现特定的视觉交互任务. 深度卷积网络算法在训练上采用随机梯度下降(stochastic gradient descend, SGD)的算法实现误差灵敏度的反向传递(back propagation, BP), 从而实现网络层的学习^[75], 即特征集合的学习和建立. 其网络中大量的参数在一定程度上将样本进行记忆, 并通过参数的形式体现在网络中. 无监督的深度学习过程, 是为了构建一个初始化的特征集, 该特征集对特定任务而言不一定是完备的. 随后, 在此基础上, 通过有监督的深度学习过程, 逐步地从不完备特征集过渡到任务完备的特征集. 因此, AlphaGo程序中所采用的深度学习以及以胜负结果为导向的训练策略作为一种有效的特征学习手段值得我们研究和关注.

在围棋特征方面, AlphaGo首先对大小为 19×19 的棋盘进行人为的特征提取, 提取出包含原始状态图在内的共48个通道的图像用于后续建模. 其次, 采用深度卷积网络对大小为 19×19 的48通道图像进行特征抽取^[1], 先后建立了监督学习策略网络(SLPN)、强化学习策略网络(reinforcement learning policy network, RLPN)以及估值网络(VN)(如图1所示).

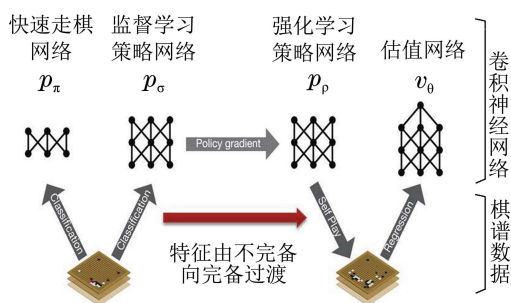


图 1 AlphaGo算法策略实现了特征由不完备向完备的过渡

Fig. 1 AlphaGo realized the transformation from incompleteness feature to completeness feature

第1阶段中, 监督学习策略网络是通过棋圣堂围棋服务器(kiseido go server, KGS)上的3000万个人类专业棋手对弈棋谱数据训练得到. 该网络根据当前的棋盘状态, 给出下一步落子的概率. 监督学习策略网络(SLPN)是一种典型基于深度卷积神经网络的分类器, 其输出层采用Softmax进行分类. 虽然称之为监督学习网络, 但从任务可完成性角度而言(即胜负), 该网络是一种无监督的网络, 仅仅实现了一个初始化的特征集合, 对于下棋任务而言未必是完备的. 采用全部输入特征的策略网络落子预测准确率仅为57%, 虽远高于其他方法, 但仍不能达到完成任务即胜利的要求.

第2阶段中, 强化学习策略网络被建立, 该网络仍是一种分类模型, 输出下一步落子的概率. 该网络在第1阶段的基础上, 使用强化学习进一步对策略网络

进行优化. 训练时, 强化学习策略网络的参数由第1阶段产生的网络参数进行初始化. 通过自弈的方式, 不断产生出大量的新的数据样本, 并在随后的训练过程中按式(1)即预期胜率最大化更新参数

$$\Delta \rho \propto \frac{\partial \log p_{\rho}(a_t | s_t)}{\partial \rho} z_t, \quad (1)$$

其中: z_t 只有在终止时刻 T 时, 根据胜负给出+1和-1的奖赏, 其余状态为0; a_t 是 t 时刻不同落子位置的描述; s_t 是 t 时刻棋盘状态; ρ 是深度卷积网络参数. $p_{\rho}(a_t | s_t)$ 反映当前棋盘状态 s_t 时, 由参数 ρ 生成的深度卷积网络所给出的不同位置的落子概率. 这种策略, 以任务胜负结果为导向, 实现特征集合从不完备向完备特征集合的过渡.

第3阶段中, 使用第2阶段的强化学习策略网络(RLPN)进一步自我对弈, 根据最终胜负结果, 建立棋盘状态与胜负概率的回归模型. 在结构上, 该网络的输出层由Softmax层被替换为回归层. 此外, 该网络与前面二者的区别在于, 前者是分类模型, 后者是回归模型; 前者是对落子的经验或先验判断(即先验概率^[1]), 后者是对落子后胜负概率的估计(即某种概率密度函数). 在训练过程中, 该网络所提取的特征集合进一步地向完备特征集合过渡, 即通过当前状态来判断任务完成的概率(即胜负).

从特征完备性的角度而言, AlphaGo程序通过监督学习策略网络(SLPN)完成初步特征集合的建立, 实现对任务完成的不同阶段和状况的描述, 即棋盘下棋状态特征的抽取. 其后逐步通过强化学习策略网络(RLPN)以任务完成性(即胜负率)为导向优化特征集合, 实现从已有的不完备的特征集合出发, 寻找近邻的对任务相对完备特征集合. 基于梯度下降算法的反向传播算法从某种意义上而言, 提供了一种有效的特征寻找和评价策略, 即根据任务来确定一个完备的特征集合. 另一方面, 由于batch-training策略, 深度学习算法得到的解是局部极小点, 难以获得全局最优解^[75]. 虽然在深度学习的解的曲面上存在大量局部极小点, 但这些点的极小值都是接近的, 且最终效果也接近. 这一点上也印证了完备特征集合不是唯一的.

2.2 表征空间(Representation space)

人工智能体在执行给定的任务之前, 都应考虑当前系统的状态是否能够完成给定的任务, 只有当任务可完成时, 进行任务规划才有意义. 在机器人运动规划领域中, 文[76]和文[16]提出了基于状态空间的机器人运动规划与多机器人协调的方法, 初步探讨了机器人系统的任务可完成性及其表达(如式(2)所示), 针对任务完成性修改机器人系统配置及任务约束, 将任务转化为可完成, 但并未形成统一的理论.

$$\begin{cases} (\zeta^0 \in A) \cap (\zeta^g \in A), \\ \exists P \subset A, P = \{\zeta^0, \dots, \zeta^g\}, \end{cases} \quad (2)$$

其中 ζ^0, ζ^g 分别表示系统的起始状态与任务的终止状态,且均位于可达空间 A 中,且可达区域内存在 ζ^0 与 ζ^g 间的联通路径 P .文[16]分析了传统运动规划任务的建模方法存在的优势和不足.文[17]在此基础上,提出了基于表征空间的任務可完成性评价框架.在该框架下,机器人执行任务的过程被视为表征空间中的轨迹变迁过程.这样,人工智能体的任务规划问题就简单地转化为表征空间中的路径搜索问题.

虽然基于任务空间(task space)和配置空间(configuration space)^[77]的建模方法均在机器人运动规划领域得到了非常广泛的应用,但在用于解决复杂规划任务时则会显示出其固有的局限性^[16].在进行运动规划任务的求解中,往往会受到多种不同的约束,比如目标约束,环境障碍约束,以及机器人本身硬件配置带来的物理约束等.随着任务复杂度的增加,在多种约束的作用下很可能无法在任务空间或配置空间求解得到一条可行的路径.因此,为了能够完成给定的运动规划任务,需要扩展传统的基于任务空间和配置空间的建模方法,以使不可行的任务转化为可行的任务.Hsieh等建立了“约束松弛”的概念^[78],OlfatiSabe在这一概念的基础上,首要考虑任务规划过程中权重高的约束,得到初始的可行路径,而后再逐渐根据其余的约束对得到的路径进行调整^[79].Haue则通过移除最少数目的约束来使运动规划任务可完成^[80].虽然上述理论和方法均能在各自的场景下解决不可行任务的转化问题,但难以推广到一般的运动规划任务中.Su和Xie首次针对一般的机器人运动规划任务提出了基于表征空间(representation space, R-Space)的建模方法^[16],其核心思想在于扩展了传统的任务空间和配置空间的维度,把影响任务完成的主要因素加入其中,以此得到一个能够完备描述运动规划任务的空间^[81].文[82]以二自由度机械臂的运动规划任务为研究对象,对表征空间中任务的可完成性以及不可完成任务的转化问题进行了研究(如图2所示).文[83]利用表征空间模型对多机器人的编队运动规划问题进行了案例应用研究.给出了机器人学研究的新方向.

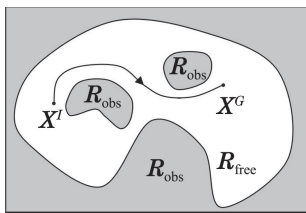


图2 二维表征空间示意图

Fig. 2 Example of a 2-dimensional representation space

对上述空间搜索所采用的路径搜索方法,是指通过初始条件和节点扩展规则最终构造出一棵连接初始节点和目标节点的树的过程.该过程可用两种不同

方法实现,分别为广度优先和深度优先搜索.然而上述两种传统方法存在重大缺陷,即以上二者均通过穷举方式进行搜索,当搜索空间较大时效率低.文[84]首次提出了一种高效的启发式搜索算法,其基本思想是在搜索的过程中对每一个搜索的结点进行评价,得出最好的节点,然后从所得到的最好的结点继续搜索直至到达目标结点.通过这种方式减少了大量的无谓搜索,极大地提高了效率.其中,启发式算法以其较低的计算复杂度、不易陷入极小值、效率高等特点被广泛采纳和研究.基于概率的算法是为了改善经典方法的效率,包括概率路线图(probabilistic roadmap, PRM)^[85-86]和快速扩展随机树(rapidly exploring random tree, RRT)^[87]等.与传统方法相比,这类方法的主要优点是高效快速.另外,在运动规划领域还存在着其他的方法,如水平集(level set)^[88]和语言几何(linguistic geometry, LG)^[89].要解决局部极小问题,很多启发式和基于启发式的算法被广泛使用.例如,模拟退火(simulated annealing, SA)技术与势场法的组合.其他启发式方法还包括人工神经网络(artificial neural network, ANN)、遗传算法(genetic algorithms, GA)、粒子群优化(particle swarm optimization, PSO)、蚁群算法(ant colony, ACO)、Stigmergy、小波理论、模糊逻辑(fuzzy logic, FL)和禁忌搜索(tabu search, TS).启发式算法不能保证找到一个最优解决方案,甚至可能无法找到可行解,但如果能够找到,其计算速度一般远高于传统方法.

蒙特卡洛树^[90-91]是一种以概率统计理论为指导的启发式算法.AlphaGo程序采用蒙特卡洛树搜索的方式(Monte Carlo tree, MCT)将监督学习策略网络(SLPN)、估值网络(VN)以及传统围棋走棋策略(RP)相结合.该程序从当前围棋棋盘状态出发,以上述3种策略网络对之后的走棋过程进行模拟(simulation)、结果评估(evaluation)以及估值备份(backup),并最终模拟出一条期望胜率最高的路径.该路径是一条期望胜率较高的路径,是下一步落子即新增节点中估计代价最优的节点.其对棋盘空间搜索的过程中,通过监督学习策略网络和估值网络建立完备的特征集合,在该特征集合所构成的空间中以先验落子概率与期望胜率加权最大化为目标进行探索.最后,根据当前状态下模拟得到的最优路径,进行下一步走棋,并不断重复上述过程.

上述方法搜索的特征空间即非配置空间又非任务空间,对其进行启发式搜索过程可以用表征空间的理论加以诠释.对表征空间而言,这里的胜率最大化目标即所要找寻的目标节点,当前棋盘的状态即当前节点,走棋策略中所抽取的特征集合对应了构成表征空间的完备的围棋特征集.启发式算法在对其空间进行搜索时的难点在于如何建立一个代价估计函数(如

式(3)所示), 用于判断当前探测节点到目标节点之间的距离, 从而选取下一步的新增节点.

$$f(n) = g(n) + h(n), \quad (3)$$

其中: $f(n)$ 是节点 n 的估值函数, $g(n)$ 表示从初始节点到当前节点 n 的实际代价, $h(n)$ 表示从节点 n 到目标节点的代价估计函数.

深智团队研发的AlphaGo程序给出了一种可行的方案. 其中的模拟和评估过程对应了启发式算法中的代价估计函数, 即对从当前节点到目标节点的代价(胜率和落子概率)进行估计. 监督学习策略网络(SLPN)、估值网络(VN)以及传统围棋走棋策略(RP)起到了启发式搜索算法中代价估计的作用. 每一次的落子对应了启发式算法中新增节点的过程. 最终, 不断迭代增加节点, 直到达到目标节点(即胜利)或空间搜索完毕(即棋盘下满). 在失利的第4盘中, AlphaGo程序对比赛输赢走势一度判断错误, 最终导致失败. 上述过程从表征空间的角度而言, 是由于代价估计函数设计(即卷积网络的训练)上存在一定缺陷, 即该函数认为人类棋手的那几手落子在表征空间中的代价函数中属于代价较高的部分, 所以程序并没有遍历或深度遍历此节点及以下分支. 我们认为AlphaGo程序所构建的表征空间是一种高维度的空间, 隐含在了深度卷积网络中. MCT所实现的是在此空间中不断使用代价函数进行估计(如式(4)所示), 并寻找一条路径, 连接至终点即“胜利”.

$$a_t = \arg \max(Q(s_t, a) + u(s_t, a)), \quad (4)$$

其中 $u(s, a)$ 反映了先验落子概率, 由监督学习策略网络(SLPN)提供. $Q(s, a)$ 反映了当前形势下的期望胜率, 该节点处胜率是对期望胜率的近似估计, 由所有后继节点的胜率平均获得, 叶节点的胜率 $V(s_L)$ 由估计网络(VN)胜率 $v_\theta(s_L)$ 与快速走棋网络(RN)胜率 z_L 线性加权得到:

$$u(s, a) \propto \frac{P(s, a)}{1 + N(s, a)}, \quad (5)$$

$$N(s, a) = \sum_{i=1}^n 1(s, a, i), \quad (6)$$

$$Q(s, a) = \frac{1}{N(s, a)} \sum_{i=1}^n 1(s, a, i) V(s_L^i), \quad (7)$$

$$V(s_L) = (1 - \lambda)v_\theta(s_L) + \lambda z_L, \quad (8)$$

其中: $N(s, a)$ 反映蒙特卡洛搜索树算法中节点访问次数, $P(s, a) = p_\rho(a|s)$ 是由深度卷积网络给出的先验落子概率. 遍历进行到 L 时, 节点记为 s_L . 在训练过

程即构建代价函数的过程中, AlphaGo采用的是自弈的方式, 虽然扩充大量的训练样本对深度学习而言是有利的, 但这种演化方式无法引入不同的对弈策略, 故缺陷必然存在.

综上所述, AlphaGo程序通过深度卷积网络与蒙特卡洛树的结合, 建立了一个对下棋任务而言完备的特征集合, 并在由此构成的表征空间中设计了一种代价估计函数, 从而选取下一步的新增节点. 整个过程中, 特征集合逐步从不完备向完备过渡. 高维度的表征空间隐含在了深度卷积网络中, 并以胜负概率、落子概率相体现, 从而给出了从当前节点到目标节点的代价估计方法. 另一方面, 实现了由抽象图形到定量结果的转化. 其思想对人工智能领域的发展具有重要的意义和作用.

3 抗干扰下AlphaGo分析(Analysis on Alpha Go under view of robustness)

人工智能体在任务完成过程中, 不可避免地会受到各种未知内/外部扰动的影响, 如: 未建模动态、模型不确定性及未知扰动等. 这些扰动会恶化系统性能, 甚至使系统不稳定, 导致智能体在表征空间中找不到一条路径达到最终目标. 因此, 在智能体任务执行过程中, 充分考虑系统可能受到的各种扰动, 对于智能体最终任务的完成有着极其重要的意义. 近年来, 抗干扰控制策略的研究及应用受到了诸多领域学者的广泛关注, 许多期刊和会议都出现了关于抗干扰控制的专刊和讨论组, 如: American Control Conference (June, 2013)¹, ISA Transactions(July, 2014)², IEEE Transactions on Industrial Electronics(Sept., 2015)³, 控制理论与应用(Dec., 2013)⁴等. 其中, 基于自抗扰控制(active disturbance rejection control, ADRC)^[92]和干扰观测器(disturbance observer, DOB)^[93]的控制策略以其灵活的结构和对扰动的有效估计能力, 受到学术界和工程界的青睐. 这类控制策略中, 均采用一个具有观测器作用的结构对未知外部干扰和未知系统动态进行在线估计, 进而对控制量进行补偿, 使得系统抗干扰性能及鲁棒性得到提升.

值得注意的是, ADRC的内环观测器, 即扩张状态观测器(extended state observer, ESO), 通常基于积分链式模型在时域设计. 因此, ESO不仅能估计系统的总扰动, 也能观测系统的状态, 进而ADRC可设计为状态反馈的形式. 而DOB通常基于先验模型信息在频域设计, 主要补偿系统外扰和内部不确定性, 外环控制器通常设计为输出反馈形式. 总之, 基于观测器的抗干扰控制策略不仅能充分利用已知的系统信息, 更

¹<http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=6573242>

²<http://www.sciencedirect.com/science/journal/00190578/53/4>

³<http://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=41>

⁴<http://jcta.alljournals.ac.cn/ctacn/ch/index>

能在不牺牲系统标称性能的前提下有效处理未知干扰和未建模动态的影响. 因此, 这类方法在机器人无标定视觉伺服^[11-15]、机器人运动控制以及多机器人协调^[94]、无人机飞行控制^[95-96]等领域得到广泛的应用.

在无标定视觉伺服领域, Su团队应用ADRC理论发展了从具体应用到机器人视觉伺服一般理论的一套体系^[13-14, 20, 97-99], 取得了一些成果: 文[98]对ADRC在机器人无标定手眼协调问题的控制器设计进行了完整的讨论, 展现了ADRC用来解决这个问题的潜力和良好性能. 通过图像雅克比矩阵建立视觉反馈与机器人执行器之间的非线性映射关系

$$\dot{P} = J(W)U, \quad (9)$$

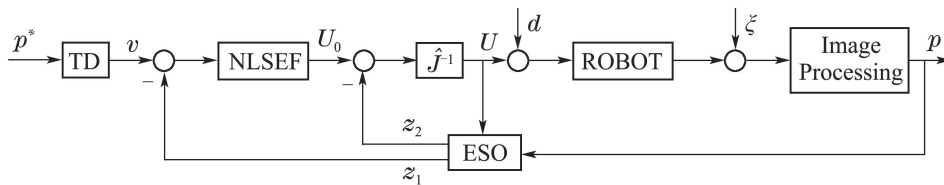


图3 基于ADRC的无标定视觉伺服结构图

Fig. 3 ADRC-based control structure of uncalibrated visual servoing

与此同时, Yin和Su等着重研究了能直接应用于DOB设计过程的优化评价函数, 并考虑系统相对阶次和扰动的内模阶次^[100-106]. 文[103]采用DOB实现柔性机械臂的关节解耦控制(如图4所示).

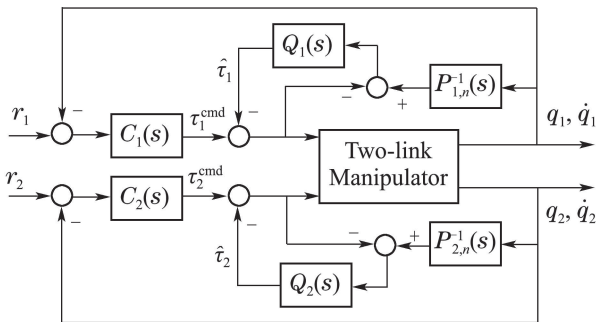


图4 基于DOB的二自由度柔性机械臂控制框图

Fig. 4 DOB-based control structure for 2-DOF robotic manipulator

从上述角度再看李世石与AlphaGo的对弈, 特别是李第4盘的第78手, 导致AlphaGo程序“崩溃”乱走棋, 当前深度学习的某些本质不足便一览无余. 人工智能的本质是人类试图以计算的方式, 建立一套可以模仿人类思想和行为的算法机制. 可以说, 无论AlphaGo采用多少层深度网络、定义何种评价函数、增加多少硬件设备以及完成多少次的自我对弈, 均只是对人类思维过程的逼近, 不可能得到人类思想的模型, 也就是说建模误差必然存在. 另一方面, 人类思想存在一定的随机性即所谓灵感, 以及中日围棋界中存

其中 $W = (x_w, y_w, z_w)$ 表示目标点在世界坐标系下的位置. 图像雅克比矩阵 $J(W)$ 是描述系统参数和外部扰动等未知参数的时变矩阵. 根据ADRC的基本思想, 粗略估计被控对象模型 $\hat{J}(W)$ 作为系统标称模型, 将真实模型与初始模型之间的偏差以及所有外扰作为总扰动.

$$\dot{P} = \hat{J}(W)U + a(t), \quad (10)$$

其中: $a(t) = (J - \hat{J})U + \xi(t) + d(t)$, $\xi(t)$ 和 $d(t)$ 分别表示图像噪声以及未知外部扰动. 设计ESO对总扰动进行估计, 并结合微分跟踪器 (tracking differentiator, TD) 和非线性状态误差反馈控制 (nonlinear state error feedback, NLSEF), 得到一个耦合的基于ADRC的系统模型. 如图3所示.

在的求道派的对弈策略等, 这一类现象不可预见性以及不可建模性对于AlphaGo已有模型来讲便是外部未知扰动, 而在人类棋手对弈过程中, 这些现象是普遍存在的. 因此, 当系统建模误差以及外部扰动存在时, 特别是李在第4盘第78手之后, 由于在AlphaGo的评价体系中认为这一步是小概率事件, 为了缩减算法计算开支, 系统直接忽略这一事件, 从而并没有采取一定的应对措施, 最终导致AlphaGo程序“崩溃”乱走棋. 可以想象, 若李对AlphaGo的运行机理完全明白 (当然这是不可能的), 李完全可以故意落在AlphaGo的小概率空间中, 那么AlphaGo或者机器学习程序是根本不可能击败人类的.

综上所述, 从系统抗干扰的角度来讲, AlphaGo采用深度学习网络对人类思维过程进行建模不可避免地存在误差, 同时, 在模拟对弈过程时采用了自我对弈, 而非与李世石的模型对弈, 这也是整个系统模型不精确的地方. 另一方面, 蒙特卡洛搜索算法虽然能减少搜索广度和时间, 但必然导致系统存在大量的未知干扰, 而AlphaGo本身并没有建立一套有效地处理扰动的机制, 最终导致了系统失稳. 例如, 当对手出现模型中小概率的落子时, 应当对此进行一定的衡量和评估. 所以, 若能借鉴抗干扰控制理论的一些思想, 建立如干扰观测器一类的策略实现对系统建模误差和未知干扰进行在线估计, 并将估计结果纳入下一步对弈策略的考虑, 或许能从本质上改善AlphaGo的运行效率和运行结果.

4 美感评价下的 AlphaGo 分析(Analysis on AlphaGo under view of beauty evaluation)

通过研究人工智能, 人类想要用科学技术的方式模拟自己的大脑活动. 计算是人类大脑的一项重要活动, 更是计算机擅长的事情. 除了计算以外, 大脑的其他活动, 如一些感性模糊的感受是更复杂多变而微妙的部分, 也是人类区别于机器和其他动物的高级特性. 美感便是其中一种. 研究美感既能对人类的思维有更深层次的认识, 也有益于创造让人感到舒适愉悦的服务和环境, 提升人机交互的效果和体验, 是人工智能和人机交互的重要组成部分.

人工智能中的美感, 即人工智能功能在人机交互的过程中给人带来的艺术舒适度. 美感是一种很主观的感受, 但是在大多数人群中又有一些共识. 技术路线上, 需要把人们模糊主观的感受剖析清楚, 确定美感的影响因素和决定方式, 根据其复杂程度借助图论、计算机视觉、学习方法等工具对其进行建模, 进而提出设计手段实现美感度高的人工智能和人机交互功能.

在艺术层面上, C. Li等^[9]和M. Bethge等^[10]试图借助深度学习算法来研究并模拟人创作艺术绘画的方式. 通过将绘画分成内容 (content) 与艺术样式 (style), 成功将两幅不同类型的绘画进行合成, 实现了抽象图形的定量化分析. 如图5所示, 左上角大图是原始内容, 艺术样式由各个名画(即3幅小图)提供, 其余3幅图分别是原始内容与名画艺术样式合成后的结果.



图 5 内容与艺术样式的创作合成^[10]

Fig. 5 The creation combining content with art style^[10]

同时, 另有一些工作^[107-108]对照片美感等进行研究. 文^[107]分别在颜色、纹理和特性通道上提取照片的图像块, 在此基础上建立起从有标记的照片训练样本到样本图像块, 到测试照片的图像块, 到测试图像的美感评价概率模型, 获取各个图像块的矩阵, 并利

用流形转换成 d 维特征向量, 根据概率模型计算出测试照片的美感评分

$$\begin{aligned} \gamma(I^*) &= p(I^*|I^1, \dots, I^H) = \\ &= p(I^*|G^*) * p(G^*|G) * p(G|I^1, \dots, I^H), \end{aligned} \tag{11}$$

其中: I^j 是有标记的照片训练样本, I^* 是照片测试样本, G 是训练样本的图像块, G^* 是测试样本的图像块. 表达式中的各个概率的计算如下:

$$\begin{aligned} p(I^*|G^*) &= p(I^*|G_1^*, \dots, G_T^*) = \\ &= \frac{p(G_1^*, \dots, G_T^*|I^*)p(I^*)}{p(G_1^*, \dots, G_T^*)} = \\ &= \prod_{t=1}^T \prod_{j=1}^{N_t^*} p(G_t^*(j)|I^*), \end{aligned} \tag{12}$$

$$\begin{aligned} p(G^*|G) &= p(G_1^*, \dots, G_T^*|G_1, \dots, G_T) = \\ &= \prod_{t=1}^T \prod_{j=1}^{N_t^*} p(G_t^*(j)|G_1, \dots, G_T), \end{aligned} \tag{13}$$

$$\begin{aligned} p(G|I^1, \dots, I^H) &= p(G_1, \dots, G_T|I^1, \dots, I^H) = \\ &= \prod_{t=1}^T \prod_{j=1}^{N_t} p(G_t(j)|I^1, \dots, I^H). \end{aligned} \tag{14}$$

但是很少有工作从科学技术的角度研究中国文化中的美感. Su等人在已有的机器人书法研究^[7]的基础上展开了对中国书法美学的探讨和研究, 尝试用西方定量的科技描述方式来解读东方的哲学和文化. 从笔画的尺寸、倾斜角、位置、宽度变化等因素入手, 分析书法作品的协调性、平衡性、构图分布等美感要素, 给出综合的美感评价, 并在此基础上参照字帖进行机器人书法创作, 在线规划汉字笔画, 写出符合人的美感评价的书法作品(如图6所示). 很明显, 如果不采用基于美感评价的在线规划, 书写效果不美观, (b)中的横撇、撇和捺有明显的过度的倾斜, 各个笔画的相对尺寸不协调, 粗细变化不合理; 相对而言, 采用基于美感评价的在线笔画规划方法, 书写效果更佳, (c)中书法整体平衡性更好, 整个字很协调, 构图疏密得当, 更符合我们的审美.



图 6 机器人书法创作

Fig. 6 Robotic calligraphy performance

围棋被视为人类所有棋类游戏中难度最高的一个, 不仅是人类的智慧壁垒, 而且是东方文化的美学和道韵的体现. 围棋与书法的美感方面具有一定的相关性.

书法的笔画、结构都有一定规则需要遵循,其美感通过白纸黑字的形式呈现出来,而围棋则同样是在一定规则下,以黑白子的形式加以体现. 书法作品的美感是通过执行者用笔体现在纸上,再透过视觉信息传递出来;围棋的道韵也是通过下棋双方借由黑白棋子体现在棋盘上,再透过视觉信息传递出来. 与书法不同之处在于,下棋双方通过博弈落子实现各自布局,从而体现出各自的棋风和道韵. 而书法仅是由单一对象完成的. AlphaGo算法充分利用计算机程序的特点和能力,采用深度卷积网络和蒙特卡洛树两大核心算法,综合考虑了落子概率(SLPN)、当前形势下的胜率(VN)和传统快速走棋策略(RP),形成估值函数,计算出每一步的走法. 整个方法总体上以结果为导向,以达到胜利为目的,却忽视了其中的美感和韵味. 自古高手下围棋,虽然步步为营,但也不斤斤计较,下棋的战术战略中无不透露着人生哲理与气度. 围棋是孤独中的修炼,讲究道、天、地、将、法,综合考虑基本道理、时机、布局、对手的风格意图、规则. “奇”和“正”也是贯穿整个过程的重要因素. 指的是布防构置要堂堂正正,同时不能泥古不化,要机动灵活,出其不意. 奇正相生,衍生出万千变幻. 此外,围棋作为一种策略性棋类,还属于“琴棋书画”四艺之一. 棋艺是要在已有一定水准的棋手之间,在对局的同时,感受对方的气质和性格,从而进行双方内在沟通,而不止于争输赢. 在双方较量的过程中,体会出现的各种定式、棋型、手筋,感受其中的奇妙和美. 这些美感和道韵恰是围棋的魅力所在.

AlphaGo算法充分利用计算机的快速大数据计算优势,达到很高胜算,是用西方计算技术对典型东方知识体系成功描述的展现. 但是忽视了围棋中双方风格脾性的沟通与棋局中的美学和道韵. 在下棋过程中记录、总结棋手的风格特点,有针对性地进行意图识别,能够提高算法的智能和自主性. 从图形图像或更抽象的角度评价各种定式、棋型和手筋的美感,甚至对道韵进行建模和模拟,能够更深层次地实现人工智能,为从西方科技的视角读懂东方文化做出更有效的尝试.

5 结论(Conclusions)

本文从任务可完成性角度出发,探讨了AlphaGo算法的原理. 在执行给定的任务之前,都应考虑当前系统的状态是否能够完成给定的任务,只有当任务可完成时,进行任务规划才有意义. 在此基础上,本文从任务可完成性的角度进行分析,并以特征完备性及表征空间为切入点,对AlphaGo算法的原理进行了阐释. 将深度卷积网络视为构建特征集的手段,将强化学习及自弈过程视为特征从不完备向完备过度的途径,将蒙特卡洛树搜索过程视为表征空间中代价函数估计的过程. 另一方面,在任务完成过程中,不可避免地会

受到各种未知内/外部扰动的影响,因此,本文从抗干扰的角度进行分析,将AlphaGo算法视为对人类棋手的建模,进而与控制论中抗干扰策略相结合进行讨论. 据此,将人类棋手的某些行为视为干扰,即当人类棋手行为不符合原先模型预测时,算法应当进行一定的调整和修正,从而进一步考察此行为对执行任务的影响. 最后,人类研究人工智能体的出发点是期望用科学技术的方式模拟自己的大脑活动. 围棋不仅仅包含了量化计算,更体现了东方文化中的艺术和美感,即定性的成分. 因此,本文从美感评价的角度出发,探讨了定性人工智能的定量研究过程中值得注意的问题. 本文从上述三个层面,对AlphaGo算法的原理和人工智能体的研究进行了探讨和分析,并指出了当前AlphaGo算法的局限性. AlphaGo虽然取得了里程碑式的进展,但在定性描述(如:美学、艺术)以及系统未知扰动方面仍存在诸多问题值得研究. 自我对弈的方式仅适用于博弈问题,并不能广泛适用于各个应用场景. 因此,AlphaGo的人工智能成分仍受到诸多制约. 人工智能的跨越式发展,应包含对事物定性分析的能力,即本文所认为的Beta级别. 而目前的AlphaGo算法仍然是一种解决定量问题的手段. 我们希望人工智能算法的研究工作者(尤其以研究AlphaGo为代表的深度学习研究工作者)通过本文能更关注于挖掘定性描述与定量分析之间的内在关联,并进一步将人工智能算法提升至BetaGo乃至更高.

参考文献(References):

- [1] SILVER D, HUANG A, MADDISON C J, et al. Mastering the game of Go with deep neural networks and tree search [J]. *Nature*, 2016, 529(7587): 484 – 489.
- [2] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Playing Atari With Deep Reinforcement Learning [G] // *NIPS Deep Learning Workshop*. 2013.
- [3] LITTMAN M L. Reinforcement learning improves behaviour from evaluative feedback [J]. *Nature*, 2015, 521(7553): 445 – 451.
- [4] LILLICRAP T P, HUNT J J, PRITZEL A, et al. Continuous control with deep reinforcement learning [J]. *Computer Science*, 2016, 8(6): A187.
- [5] ZHAO D, SHAO K, ZHU Y, et al. Review of deep reinforcement learning and discussions on the development of computer Go [J]. *Control Theory & Applications*, 2016, 33(6): 701 – 717.
- [6] KARPATHY A, FEI-FEI L. Deep visual-semantic alignments for generating image descriptions [C] // *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Boston, MA, USA: IEEE Computer Society, 2015: 3128 – 3137.
- [7] MA Z, XIANG Z, SU J. Robust visual servoing based Chinese calligraphy on a humanoid robot [J]. *High Technology Letters*, 2016, 22(1): 30 – 37.
- [8] LAKE B M, SALAKHUTDINOV R, TENENBAUM J B. Human-level concept learning through probabilistic program induction [J]. *Science*, 2015, 350(6266): 1332 – 1338.
- [9] LI C, WAND M. Combining Markov random fields and convolutional neural networks for image synthesis [J]. *CoRR*, 2016, abs/1601.04589.

- [10] GATYS L A, ECKER A S, BETHGE M. A neural algorithm of artistic style [J]. *CoRR*, 2015, abs/1508.06576.
- [11] SU J. Camera calibration based on receptive fields [J]. *Pattern Recognition*, 2007, 40(10): 2837 – 2845.
- [12] SU J, WANG J, XI Y. Incremental learning with balanced update on receptive fields for multi-sensor data fusion [J]. *IEEE Transactions on Systems Man and Cybernetics, Part B (Cybernetics)*, 2004, 34(1): 659 – 665.
- [13] SU J, QIU W, MA H, et al. Calibration-free robotic eye-hand coordination based on an auto disturbance-rejection controller [J]. *IEEE Transactions on Robotics*, 2004, 20(5): 899 – 907.
- [14] SU J, MA H, QIU W, et al. Task-independent robotic uncalibrated hand-eye coordination based on the extended state observer [J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 2004, 34(4): 1917 – 1922.
- [15] SU J, XI Y, HANEBECK U D, et al. Nonlinear visual mapping model for 3-D visual tracking with uncalibrated eye-in-hand robotic system [J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 2004, 34(1): 652 – 659.
- [16] SU J, XIE W. Motion planning and coordination for robot systems based on representation space [J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 2011, 41(1): 248 – 259.
- [17] WU B. *Evaluation of task realizability for robot systems* [D]. Shanghai: Shanghai Jiao Tong University, 2012.
- [18] BICCHI A, TONIETTI G. Fast and soft-arm tactics: dealing with the safety-performance trade-off in robot arms design and control [J]. *IEEE Robotics and Automation Magazine*, 2004, 11(2): 22 – 33.
- [19] TAHRI O, MEZOUAR Y, CHAUMETTE F, et al. Decoupled image-based visual servoing for cameras obeying the unified projection model [J]. *IEEE Transactions on Robotics*, 2010, 26(4): 684 – 697.
- [20] SU J. Robotic uncalibrated visual serving based on ADRC [J]. *Control and Decision*, 2015(1): 1 – 8.
- [21] OU M, LI S, WANG C. Finite-time tracking control for nonholonomic mobile robots based on visual servoing [J]. *Asian Journal of Control*, 2014, 16(3): 679 – 691.
- [22] ZHANG X, FANG Y, SUN N. Visual servoing of mobile robots for posture stabilization: from theory to experiments [J]. *International Journal of Robust & Nonlinear Control*, 2015, 25(1): 1 – 15.
- [23] FOMENA R T, CHAUMETTE F. Improvements on visual servoing from spherical targets using a spherical projection model [J]. *IEEE Transactions on Robotics*, 2009, 25(4): 874 – 886.
- [24] MALIS E, MEZOUAR Y, RIVES P. Robustness of image-based visual servoing with a calibrated camera in the presence of uncertainties in the three-dimensional structure [J]. *IEEE Transactions on Robotics*, 2010, 26(1): 112 – 120.
- [25] MA Z, SU J. Feature completeness in visual servoing [C] // *Chinese Control Conference*. Hefei: IEEE, 2012: 1 – 8.
- [26] WRIGHT J, MA Y, MAIRAL J, et al. Sparse representation for computer vision and pattern recognition [J]. *Proceedings of the IEEE*, 2010, 98(6): 1031 – 1044.
- [27] AHARON M, ELAD M, BRUCKSTEIN A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation [J]. *IEEE Transactions on Signal Processing*, 2006, 54(11): 4311 – 4322.
- [28] MAIRAL J, ELAD M, SAPIRO G. Sparse representation for color image restoration [J]. *IEEE Transactions on Image Processing*, 2008, 17(1): 53 – 69.
- [29] TAN W T, CHEUNG G, MA Y. Face recovery in conference video streaming using robust principal component analysis [C] // *The 18th IEEE International Conference on Image Processing (ICIP)*. Brussels: IEEE Signal Processing Society, 2011: 3225 – 3228.
- [30] LI Y, CAI C, QIU G, et al. Face hallucination based on sparse local-pixel structure [J]. *Pattern Recognition*, 2014, 47(3): 1261 – 1270.
- [31] WRIGHT J, GANESH A, RAO S, et al. Robust principal component analysis: exact recovery of corrupted low-rank matrices [J]. *Journal of the ACM*, 2009, 58(3): 1 – 73.
- [32] WRIGHT J, YANG A Y, GANESH A, et al. Robust face recognition via sparse representation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(2): 210 – 227.
- [33] ZHANG L, YANG M, FENG X. Sparse representation or collaborative representation: Which helps face recognition? [C] // *2011 International Conference on Computer Vision*. Barcelona: [s.n.], 2011: 471 – 478.
- [34] ARPIT D, SRIVASTAVA G, FU Y. Locality-constrained low rank coding for face recognition [C] // *The 21st International Conference on Pattern Recognition (ICPR)*. Stuttgart: IEEE, 2012: 1687 – 1690.
- [35] ZHAO Y, SU J. Sparse learning for salient facial feature description [C] // *IEEE International Conference on Robotics and Automation*. Hong Kong: IEEE, 2014: 5565 – 5570.
- [36] JIANG X, LAI J. Sparse and dense hybrid representation via dictionary decomposition for face recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(5): 1067 – 1079.
- [37] YANG M, ZHANG L, FENG X, et al. Sparse representation based fisher discrimination dictionary learning for image classification [J]. *International Journal of Computer Vision*, 2014, 109(3): 209 – 232.
- [38] ELAD M, AHARON M. Image denoising via sparse and redundant representations over learned dictionaries [J]. *IEEE Transactions on Image Processing*, 2006, 15(12): 3736 – 3745.
- [39] YANG A Y, WRIGHT J, MA Y, et al. Feature selection in face recognition: a sparse representation perspective [R]. USA: University of Illinois at Urbana-Champaign, 2007.
- [40] ZHAO Y, SU J. New sparse facial feature description model based on saliency evaluation of regions and features [J]. *International Journal of Pattern Recognition & Artificial Intelligence*, 2015, 29(5).
- [41] AHONEN T, HADID A, PIETIKÄINEN M. *Face Recognition with Local Binary Patterns* [M]. Berlin, Heidelberg: Springer, 2004: 469 – 481.
- [42] GONG Y, KUMAR S, ROWLEY H A, et al. Learning binary codes for high-dimensional data using bilinear projections [C] // *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Portland, Oregon, USA: IEEE Computer Society, 2013: 484 – 491.
- [43] CHEN D, CAO X, WEN F, et al. Blessing of dimensionality: high-dimensional feature and its efficient compression for face verification [C] // *2013 IEEE Conference on Computer Vision and Pattern Recognition*. Portland, Oregon, USA: IEEE Computer Society, 2013: 3025 – 3032.
- [44] LU J, LIONG V E, ZHOU X, et al. Learning compact binary face descriptor for face recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(10): 2041 – 2056.
- [45] CHAN C H, KITTLER J, POH N, et al. (Multiscale) Local phase quantisation histogram discriminant analysis with score normalisation for robust face recognition [C] // *IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops)*. Kyoto, Japan: IEEE Computer Society, 2009: 633 – 640.
- [46] CHAN C H, TAHIR M A, KITTLER J, et al. Multiscale local phase quantization for robust component-based face recognition using kernel fusion of multiple descriptors [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(5): 1164 – 1177.
- [47] LIU C, WECHSLER H. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition [J]. *IEEE Transactions on Image Processing*, 2002, 11(4): 467 – 476.
- [48] SHEN L, BAIL A. A review on Gabor wavelets for face recognition [J]. *Formal Pattern Analysis & Applications*, 2006, 9(2/3): 273 – 292.

- [49] ZHANG B, SHAN S, CHEN X, et al. Histogram of Gabor phase patterns (HGPP): a novel object representation approach for face recognition [J]. *IEEE Transactions on Image Processing*, 2007, 16(1): 57 – 68.
- [50] ZHANG W, SHAN S, QING L, et al. Are Gabor phases really useless for face recognition? [J]. *Formal Pattern Analysis & Applications*, 2009, 12(3): 301 – 307.
- [51] XIE S, SHAN S, CHEN X, et al. Fusing local patterns of Gabor magnitude and phase for face recognition [J]. *IEEE Transactions on Image Processing*, 2010, 19(5): 1349 – 1361.
- [52] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection [C] // *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Diego, CA, USA: IEEE Computer Society, 2005(1): 886 – 893.
- [53] DENIZ O, BUENO G, SALIDO J, et al. Face recognition using histograms of oriented gradients [J]. *Pattern Recognition Letters*, 2011, 32(12): 1598 – 1603.
- [54] LOWE D G. Distinctive image features from scale-invariant keypoints [J]. *International Journal of Computer Vision*, 2004, 60(2): 91 – 110.
- [55] ZISSERMAN A. All About VLAD [C] // *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Portland, Oregon, USA: IEEE Computer Society, 2013: 1578 – 1585.
- [56] LAZEBNIK S, SCHMID C, PONCE J. Beyond bags of features: spatial pyramid matching for recognizing natural scene categories [C] // *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. New York, USA: IEEE Computer Society, 2006(2): 2169 – 2178.
- [57] WANG Bo, GUO Jichang, ZHANG Yan. Learnable receptive fields scheme in deep networks for image categorization [J]. *Control Theory & Applications*, 2015, 32(8): 1114 – 1119.
(王博, 郭继昌, 张艳. 基于深度网络的可学习感受野算法在图像分类中的应用 [J]. *控制理论与应用*, 2015, 32(8): 1114 – 1119.)
- [58] BOUREAU Y L, BACH F, LECUN Y, et al. Learning mid-level features for recognition [C] // *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Francisco, CA, USA: IEEE Computer Society, 2010: 2559 – 2566.
- [59] CAO Z, YIN Q, TANG X, et al. Face recognition with learning-based descriptor [C] // *IEEE Conference on Computer Vision and Pattern Recognition*. Portland: IEEE, 2010: 2707 – 2714.
- [60] HINTON G E, OSINDERO S, TEH Y W. A fast learning algorithm for deep belief nets [J]. *Neural Computation*, 2006, 18(7): 1527 – 1554.
- [61] SCHOLKOPF B, PLATT J, HOFMANN T. Greedy layer-wise training of deep networks [M] // *Advances in Neural Information Processing Systems 19: Proceedings of the 2006 Conference*. [S.l.]: MIT Press, 2007: 153 – 160.
- [62] BENGIO Y. Learning deep architectures for AI [J]. *Foundations and Trends in Machine Learning*, 2009, 2(1): 1 – 127.
- [63] LEE H, GROSSE R, RANGANATH R, et al. Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations [C] // *Proceedings of the 26th Annual International Conference on Machine Learning*. Montreal, Quebec, Canada: ACM, 2009: 609 – 616.
- [64] LEE H, GROSSE R, RANGANATH R, et al. Unsupervised learning of hierarchical representations with convolutional deep belief networks [J]. *Communications of the ACM*, 2011, 54(10): 95 – 103.
- [65] HUANG G B, LEE H, LEARNED-MILLER E. Learning hierarchical representations for face verification with convolutional deep belief networks [C] // *2012 IEEE Conference on Computer Vision and Pattern Recognition*. Providence, RI, USA: IEEE Computer Society, 2012: 2518 – 2525.
- [66] LUO P, WANG X, TANG X. Hierarchical face parsing via deep learning [C] // *IEEE Conference on Computer Vision and Pattern Recognition*. Providence, RI, USA: IEEE Computer Society, 2012: 2480 – 2487.
- [67] SUN Y, WANG X, TANG X. Deep learning face representation from predicting 10,000 classes [C] // *2014 IEEE Conference on Computer Vision and Pattern Recognition*. Columbus, OH, USA: IEEE Computer Society, 2014: 1891 – 1898.
- [68] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions [C] // *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Boston, MA, USA: IEEE Computer Society, 2015: 1 – 9.
- [69] LIANG X, LIU S, SHEN X, et al. Deep human parsing with active template regression [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(12): 2402 – 2414.
- [70] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. *Advances in Neural Information Processing Systems*, 2012, 25(2): 1097 – 1105.
- [71] CHAN T H, JIA K, GAO S, et al. PCANet: A simple deep learning baseline for image classification? [J]. *IEEE Transactions on Image Processing*, 2014, 24(12): 5017 – 5032.
- [72] SIMONYAN K, VEDALDI A, ZISSERMAN A. Deep inside convolutional networks: visualising image classification models and saliency maps [J]. *CoRR*, 2013, abs/1312.6034.
- [73] ZEILER M D, FERGUS R. Visualizing and understanding convolutional networks [C] // *Proceeding of Computer Vision—ECCV 2014: 13th European Conference*. Zurich, Switzerland: Springer, 2014: 818 – 833.
- [74] MAHENDRAN A, VEDALDI A. Understanding deep image representations by inverting them [C] // *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Boston, MA, USA: IEEE Computer Society, 2015: 5188 – 5196.
- [75] LECUN Y, BENGIO Y, HINTON G. Deep learning [J]. *Nature*, 2015, 521(7553): 436 – 444.
- [76] XIE W. *The research on motion planning and coordination for robot systems* [D]. Shanghai: Shanghai Jiao Tong University, 2008.
- [77] LOZANOPEREZ T. Spatial planning: a configuration space approach [J]. *IEEE Transactions on Computers*, 1983, c-32(2): 108 – 120.
- [78] HSIEH M A, KUMAR V, CHAIMOWICZ L. Decentralized controllers for shape generation with robotic swarms [J]. *Robotica*, 2008, 26(5): 691 – 701.
- [79] OLFATI-SABER R. Flocking for multi-agent dynamic systems: algorithms and theory [J]. *IEEE Transactions on Automatic Control*, 2006, 51(3): 401 – 420.
- [80] HAUSER K. The minimum constraint removal problem with three robotics applications [J]. *International Journal of Robotics Research*, 2014, 33(1): 5 – 17.
- [81] XIANG Zhenzhen, SU Jianbo. Robotic hierarchical motion planning in its representation space [J]. *Control Theory & Applications*, 2015, 32(9): 1261 – 1268.
(项镇镇, 苏剑波. 表征空间中的机器人分层运动规划 [J]. *控制理论与应用*, 2015, 32(9): 1261 – 1268.)
- [82] WU B, SU J. Evaluation of task realizability for robot systems in representation space [C] // *Proceedings of the 10th World Congress on Intelligent Control and Automation*. Beijing, China: IEEE, 2012: 3816 – 3821.
- [83] ZHANG Y. Motion planning for multi-robot coordination on representation space [J]. *IFAC Proceedings Volumes*, 2014, 47(3): 3024 – 3029.
- [84] HART P E, NILSSON N J, RAPHAEL B. A formal basis for the heuristic determination of minimum cost paths [J]. *IEEE Transactions on Systems Science & Cybernetics*, 1968, 4(2): 100 – 107.

- [85] KAVRAKI L, LATOMBE J C. Randomized preprocessing of configuration for fast path planning [C] // *Proceedings of the 1994 IEEE International Conference on Robotics and Automation*. San Diego, CA, USA: IEEE Computer Society, 1994: 2138 – 2145 vol.3.
- [86] KAVRAKI L E, SVESTKA P, LATOMBE J C, et al. Probabilistic road maps for path planning in high-dimensional configuration spaces [J]. *IEEE Transactions on Robotics & Automation*, 1996, 12(4): 566 – 580.
- [87] LAVALLE S M. Rapidly-exploring random trees: a new tool for path planning [J]. *Algorithmic & Computational Robotics New Directions*, 1999: 293 – 308.
- [88] ANDREW A M. Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science [J]. *Robotica*, 2000, 18(1): 89 – 92.
- [89] STILMAN B. Network languages for complex systems [J]. *Computers & Mathematics with Applications*, 1993, 26(8): 51 – 79.
- [90] COULOM R. Efficient selectivity and backup operators in Monte-Carlo tree search [C] // *Computers and Games: 5th International Conference, CG 2006*. Berlin, Heidelberg: Springer, 2007: 72 – 83.
- [91] KOCSIS L, SZEPEVARI C. Bandit based monte-carlo planning [C] // *Proceedings of the 17th European conference on Machine Learning*. Heidelberg, Berlin: Springer, 2006: 282 – 293.
- [92] HAN J. From PID to active disturbance rejection control [J]. *IEEE Transactions on Industrial Electronics*, 2009, 56(3): 900 – 906.
- [93] CHEN W, BALLANCE D J, GAWTHROP P J, et al. A nonlinear disturbance observer for robotic manipulators [J]. *IEEE Transactions on Industrial Electronics*, 2000, 47(4): 932 – 938.
- [94] CHAI R. *Motion planning for multi-robot formation* [D]. Shanghai: Shanghai Jiao Tong University, 2013.
- [95] HUANG Tingguo, WANG Lu, SU Jianbo. Nonlinear disturbance rejection control of unmanned aerial vehicle attitude [J]. *Control Theory & Applications*, 2015, 32(4): 456 – 463.
(黄廷国, 王璐, 苏剑波. 无人驾驶飞行器姿态的非线性扰动抑制控制 [J]. 控制理论与应用, 2015, 32(4): 456 – 463.)
- [96] WANG Lu, SU Jianbo. Attitude tracking of aircraft based disturbance rejection control [J]. *Control Theory & Applications*, 2013, 30(12): 1609 – 1616.
(王璐, 苏剑波. 基于干扰抑制控制的飞行器姿态跟踪 [J]. 控制理论与应用, 2013, 30(12): 1609 – 1616.)
- [97] MA H, SU J. Uncalibrated robotic hand-eye coordination based on theory of coupled ADRC [J]. *Robot*, 2003, 25(1): 39 – 43.
- [98] SU J, QIU W. Robotic calibration-free handeye coordination based on auto disturbances rejection controller [J]. *Acta Automatica Sinica*, 2003, 29(2): 161 – 167.
- [99] MA Hongyu, SU Jianbo. Uncalibrated robotic 3D hand-eye coordination based on auto disturbance rejection controller [J]. *Acta Automatica Sinica*, 2004, 30(3): 400 – 406.
(马红雨, 苏剑波. 基于自抗扰控制器的机器人无标定三维手眼协调 [J]. 自动化学报, 2004, 30(3): 400 – 406.)
- [100] YIN Zhengnan, SU Jianbo, LIU Yantao. Design of disturbance observer with robust performance based on H_∞ norm optimization [J]. *Acta Automatica Sinica*, 2011, 37(3): 331 – 341.
(尹正男, 苏剑波, 刘艳涛. 基于 H_∞ 范数优化的干扰观测器的鲁棒设计 [J]. 自动化学报, 2011, 37(3): 331 – 341.)
- [101] YIN Zhengnan, SU Jianbo, GAO Xiuxing. Systematic design method of disturbance observer guaranteeing closed-loop system's robust stability [J]. *Acta Automatica Sinica*, 2012, 38(1): 12 – 22.
(尹正男, 苏剑波, 高秀行. 保证闭环系统鲁棒稳定性的干扰观测器系统性设计方法 [J]. 自动化学报, 2012, 38(1): 12 – 22.)
- [102] SU J, WANG L, YUN J. A design of disturbance observer in standard H_∞ control framework [J]. *International Journal of Robust and Nonlinear Control*, 2015, 25(16): 2894 – 2910.
- [103] YUN J, SU J. Design of a disturbance observer for a two-link manipulator with flexible joints [J]. *IEEE Transactions on Control Systems Technology*, 2014, 22(2): 809 – 815.
- [104] YUN J N, SU J, KIM Y I, et al. Robust disturbance observer for two-inertia system [J]. *IEEE Transactions on Industrial Electronics*, 2013, 60(7): 2700 – 2710.
- [105] MA Z, SU J. Robust uncalibrated visual servoing control based on disturbance observer [J]. *ISA transactions*, 2015, 59: 193 – 204.
- [106] WANG L, SU J. Robust disturbance rejection control for attitude tracking of an aircraft [J]. *IEEE Transactions on Control Systems Technology*, 2015, 23(6): 2361 – 2368.
- [107] ZHANG L, GAO Y, ZIMMERMANN R, et al. Fusion of multichannel local and global structural cues for photo aesthetics evaluation [J]. *IEEE Transactions on Image Processing*, 2014, 23(3): 1419 – 1429.
- [108] MEIJER H J A, STELLINGSMA K, MEIJNDERT L, et al. A new index for rating aesthetics of implant-supported single crowns and adjacent soft tissues—the Implant Crown Aesthetic Index [J]. *Clinical Oral Implants Research*, 2005, 16(6): 645 – 649.

作者简介:

苏剑波 (1969–), 男, 教授, 目前主要研究机器人无标定视觉伺服、意图识别、模式识别、人机交互等, E-mail: jbsu@sju.edu.cn;

陈叶飞 (1988–), 男, 博士研究生, 主要研究人脸识别及相关模式识别领域, E-mail: marschen@sju.edu.cn;

马哲 (1988–), 女, 博士研究生, 研究内容包括机器人视觉伺服控制和人机交互, E-mail: cheetahf@sju.edu.cn;

黄瑶 (1991–), 女, 博士研究生, 当前主要进行机器人视觉伺服方向的研究工作, E-mail: eden9121@sju.edu.cn;

向国菲 (1990–), 男, 博士研究生, 当前主要进行鲁棒控制、抗干扰控制策略及其应用等方面研究工作, E-mail: xiangguofei@sju.edu.cn;

陈若冰 (1994–), 男, 硕士研究生, 当前主要研究基于书法机器人的视觉伺服控制问题, E-mail: leonhardt@sju.edu.cn.