



(12)发明专利申请

(10)申请公布号 CN 111077769 A

(43)申请公布日 2020.04.28

(21)申请号 201910994492.3

(22)申请日 2019.10.18

(30)优先权数据

102018217966.2 2018.10.19 DE

(71)申请人 罗伯特·博世有限公司

地址 德国斯图加特

(72)发明人 C.丹尼尔 M.福尔普

(74)专利代理机构 中国专利代理(香港)有限公司

司 72001

代理人 方莉 李雪莹

(51)Int.Cl.

G05B 13/04(2006.01)

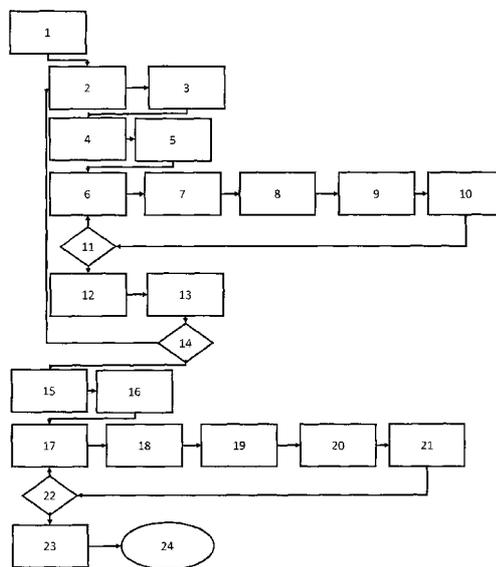
权利要求书1页 说明书5页 附图1页

(54)发明名称

用于控制或调节技术系统的方法

(57)摘要

介绍了一种用于控制或调节技术系统的方法,其中通过贝叶斯优化来确定控制或调节的参数并且在贝叶斯优化中通过采集函数来确定下一个测试点。在此,通过学习方法来自动地确定贝叶斯优化中的采集函数,借助于所确定的采集函数在贝叶斯优化中获知控制或调节的参数并且根据所获知的参数来操控或调节技术系统。



1. 用于控制或调节技术系统的方法,其中通过贝叶斯优化来确定控制或调节的参数,其中在贝叶斯优化中通过采集函数来确定下一个测试点,其特征在于,通过学习方法来自动地确定贝叶斯优化中的采集函数,借助于所确定的采集函数在贝叶斯优化中获知控制或调节的参数并且根据所获知的参数来操控或调节技术系统。

2. 根据前述权利要求中任一项所述的方法,其特征在于,在对于所述技术系统的控制或调节的模拟中学习所述采集函数。

3. 根据权利要求2所述的方法,其特征在于,在所述模拟中,对于控制或调节的优化通过贝叶斯优化作为学习方法的训练任务被选择。

4. 根据权利要求2或3中任一项所述的方法,其特征在于,在所述模拟中随机选择用于控制或调节技术系统的条件。

5. 根据权利要求2至4中任一项所述的方法,其特征在于,在所述模拟中用初始的采集函数来确定用于以贝叶斯优化来对技术系统进行优化的控制或调节的过程品质,并且根据所确定的过程品质来对所述初始的采集函数进行适配。

6. 根据权利要求2至5中任一项所述的方法,其特征在于,在所述模拟中在确定所述采集函数之后在实际的技术系统上进行过程优化。

7. 根据前述权利要求中任一项所述的方法,其特征在于,迭代地实施所述学习方法直至中止标准、尤其是达到预先确定的过程品质、所述方法的预先确定的持续时间或者预先确定的迭代次数。

8. 根据前述权利要求中任一项所述的方法,其特征在于,通过强化学习或模仿学习来确定所述采集函数。

9. 根据前述权利要求中任一项所述的方法,其特征在于,所述技术系统是工具、机床、机器人或车辆。

10. 用于优化机器学习系统的方法,尤其是用于优化神经网络,其中通过贝叶斯优化来确定学习系统的参数,其中在贝叶斯优化中通过采集函数来确定下一个测试点,其特征在于,通过学习方法来自动地确定贝叶斯优化中的采集函数,借助于所确定的采集函数在贝叶斯优化中获知参数,并且根据所获知的参数来训练所述机器学习系统。

11. 计算机程序,所述计算机程序被设置用于执行根据前述权利要求中任一项所述的方法。

12. 计算单元,具有根据权利要求11所述的计算机程序。

用于控制或调节技术系统的方法

技术领域

[0001] 本发明涉及用于控制或调节技术系统的方法以及一种为此而设置的计算机程序。

背景技术

[0002] 贝叶斯优化 (Bayesian Optimization) 能够用作黑箱过程的全局优化方法、尤其是用于优化与基础技术过程相应的函数的参数, 其中所述技术过程由其输入和输出来定义。为此, 能够在贝叶斯优化中迭代地选择测试点, 其目的经常是, 用尽可能少的测试来找到实现函数最优值的输入参数。

[0003] 为此目的, 贝叶斯优化利用所谓的采集函数 (Acquisition Functions) 来选择下一个测试点。这些采集函数典型地手动定义, 参见 E. Brochu、V.M. Cora 和 Nando de Freitas 的“A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning”, CoRR 2010。

发明内容

[0004] 提出了一种用于控制或调节技术系统、尤其是工具、机床、机器人或车辆的方法。在此, 通过贝叶斯优化来确定控制或调节的至少一个参数, 其中在贝叶斯优化中通过采集函数来确定迭代的优化过程的下一个测试点。现在提出, 可以通过学习过程、优选通过强化学习或模仿学习来自动地确定贝叶斯优化中的采集函数, 其中而后借助于所确定的采集函数在贝叶斯优化中获知控制或调节的参数并且根据所获知的参数来操控或调节技术系统。

[0005] 在此, 采集函数 (Acquisition Function) 尤其是用于选择 (潜在或实际的) 未来的测试点的比较的评价函数。

[0006] 贝叶斯优化 (Bayesian Optimization) 尤其是一种无梯度的方法, 用于通过对于测试点的迭代的评估来对未知函数进行全局优化。

[0007] 强化学习尤其是用于对顺序的决策问题进行优化的机器学习框架。在此, 所谓的策略尤其是应该通过强化学习来优化的决策函数。

[0008] 优选在对技术系统的控制或调节的模拟中学习采集函数, 其中在模拟中, 对于控制或调节的优化通过贝叶斯优化作为学习方法的训练任务被选择。

[0009] 在模拟中, 优选随机选择用于控制或调节技术系统的条件或环境参数。用初始的采集函数, 在模拟中优选确定用于以贝叶斯优化对技术系统进行优化的控制或调节的过程品质, 并且根据所确定的过程品质对初始的采集函数进行适配或优化, 并且因此学习所改进的采集函数。在使用所学习的采集函数的情况下, 优选在确定采集函数之后在模拟中在实际的技术系统上进行过程优化。

[0010] 优选迭代地实施用于采集函数的学习方法, 直至达到中止标准、尤其是达到预先确定的过程品质、方法的预先确定的持续时间或预先确定的迭代次数。

[0011] 本发明能够自动地、自学习地建立采集函数。这些所学习的采集函数在总体上功

能更强大并且可以更快地评估。此外,能够针对特定的问题类别或要求优化地学习这些采集函数,从而能够实现另外的效率优点、例如较少数量的测试点。所述采集函数能够进行更有效率的并且更加有效的搜索,也就是说,用较少的尝试就找到最优值,并且所找到的最优值具有较小的后悔程度(Regret,与最优函数值的差别)。对于所学习的采集函数的评估通常所需要的计算时间明显少于诸如“熵搜索(Entropy Search)”的更加复杂的常规使用的采集函数中的一些采集函数(相对于几个小时<1s)。

附图说明

[0012] 下面参照附图并且借助于实施例对本发明进行详细解释。在此,图1示意性地示出了用于对技术系统进行控制或调节的方法的示例性的流程。

具体实施方式

[0013] 贝叶斯优化(Bayesian Optimization)能够用作黑箱过程(Black-Box-Prozesse)的全局优化方法。为此,能够迭代地选择测试点,其目的通常是,用尽可能少的测试来找到实现函数最优值的输入参数。为此目的,贝叶斯优化利用所谓的采集函数(Acquisition Functions)来选择下一个测试点。

[0014] 已知的采集函数包括:

- “改进的可能性(Probability of Improvement)”:在实践中,这种采集函数可能导致无效的本地搜索;

- “预期的改进(Expected Improvement)”:这种采集函数试图补偿“改善的可能性”函数的缺点,但是因此搜索范围很广并且可能找不到最优值,而是可能在参数范围之间来回跳转;

- “置信上限(Upper Confidence Bound)”:这种采集函数可能非常有效,但是对于高维度的问题来说经常会提供相对较差的结果;

- “熵搜索(Entropy Search)”:在许多情况下这种采集函数提供很好的结果,但是经常要求很长的计算时间(经常数小时至数天)。

[0015] 通常应该坚持的是,对于已知的采集函数来说会发现一般的优点和缺点,但是这些优点和缺点只能用于选择对所确定的技术过程进行优化的依据,而不能用作确定的决策点。也就是说,在实践中,典型地会测试多个采集函数,以用于针对问题来选择相应最佳的采集函数。

[0016] 现在提出一种方法,利用该方法能够借助于学习方法来自动地确定要使用的采集函数。在此能够学习采集函数,方法尤其是:通过贝叶斯优化作为学习方法的训练中的任务来提供未知函数的优化。

[0017] 为此,在第一步骤中,优选获知或提供对技术系统或技术过程的相关函数类别的代表。这能够通过对于通用函数类别的预定、像例如多项式来进行或者在不同条件下通过对于有待控制或调节的技术系统的特殊的模拟来进行。

[0018] 如果使用相应的方法用于控制或调节技术系统、例如用于优化磨削过程,则能够在相应的模拟中描绘系统属性、例如输入材料、具体的磨削方法或磨削过程的环境条件。

[0019] 在第二步骤中,模拟地学习用于有待优化的过程的最优的采集函数。为此,优选重

复地实施以下步骤：

- a. 首先能够随机地选择诸如环境条件的条件，以用于固定模拟环境；
- b. 借助贝叶斯优化，在目前的策略下（采集函数）寻找函数最优值；
- c. 根据所查明的数据能够学习所改进的策略。

[0020] 现在能够将在第二步骤中所学习到的策略用作用于实际的系统的采集函数，以用于优化技术过程。

[0021] 在第二步骤的学习方法中，在模拟过程中，优选为学习方法的每次迭代 k 来执行贝叶斯优化的用 i 指示的步骤：

- GP $_k, i$: 在时刻 i 所学习到的过程模型，
- θ_k, i : 所选择的测试点，
- q_k, i : 所产生的过程品质。

[0022] 借助于这些数据，能够要么通过强化学习要么通过模仿学习来学习策略。

[0023] 自由学习（强化学习）：

在自由学习模式下未预定目标-采集函数；换言之应该学习全新的采集函数。在这种情况下，使用所谓的无模型的强化学习方法，以用于训练代表着最优的采集函数的策略。这样的方法例如是策略梯度方法、Q-学习方法、行动者-批评者方法（Actor-Critic-Methoden）。对于这些方法中的每一种方法来说，大量的算法可用，考虑将所述算法用于所描述的学习过程。

[0024] 自由学习的优点是通常最终更好的解决方案，该解决方案也能够专门针对所期望的应用类别，以用于在那里实现效率优点。

[0025] 目标指向的学习（模仿学习）：

就目标指向的学习而言，将所存在的采集函数预定为目标。优点是，可以快得多地对采集函数的所学习的代表进行评估。为此，借助于目标-采集函数来产生数据，然后将所述数据用在监督学习方法（Supervised-Learning-Verfahren）中，以用于训练所述策略至再现所观察到的行为。

[0026] 优点是更加简单的学习问题，因为只需要模仿已知的行为并且根本无需学习新的行为。因此也可能在总体上减少了学习时间。

[0027] 如果使用强化学习，则能够在每次迭代 k 之后通过强化-方法来适配策略。如果使用模仿学习，则在建立所有数据点之后才对策略进行适配。

[0028] 所提出的方法例如能够用于优化各种技术控制或调节过程、尤其是用于控制或调节技术系统、尤其是机器人、车辆、工具或机床，以用于根据所采集到的输入参量来确定并且使用用于控制或调节的输出参量。作为所述调节或控制的输入参量，尤其考虑传感器数据或者根据传感器数据所获知的参量。传感器数据能够源自技术系统的传感器或者能够由技术系统从外部来接收。根据输出参量或者控制或调节参量，通过计算单元用操控信号来操控或者用调节信号调节技术系统的至少一个执行器。例如，就这样能够控制机器人或车辆的运动或者对车辆的驱动单元或驾驶员协助系统进行控制。更加具体的用途例如包括磨削过程、超声波焊接过程的优化、接合过程的优化、钻孔过程（例如钻深孔过程）的优化、PID控制器的调整、车辆或机器人的参数化期望轨迹的优化等等。

[0029] 例如，在给定磨削过程（不够准确的）模拟情况下对实际的磨削过程进行优化。能

够通过不同的调节参量在其行为方面改变模拟。所述调节参量例如能够是工具材料、工件材料、环境参量、过程噪声等等。借助于所描述的方法,能够根据给定的模拟来学习用于一般磨削过程的最优的或得到优化的采集函数,方法是:在学习过程中有规律地改变模拟参数。这种采集函数引起用于对实际的磨削过程进行参数化的、尤其有效的贝叶斯优化过程,因为基础的采集函数专门用于磨削过程(或确定类别的磨削过程)。

[0030] 图1示意性地示出了所提出的方法的示例性的流程,所述方法用来学习用于贝叶斯优化的采集函数。

[0031] 在第一步骤1中使策略(采集函数)初始化。在接下来的步骤2中,确定或者随机选择用于模拟的环境条件、例如材料参数。对于这些环境条件来说,在接下来的步骤3中开始模拟技术过程、这里是磨削过程。在这种情况下,首先在步骤4中将第一个计数器设置为零。在接下来的步骤5中,根据第一计数器的计数器值零使过程模型初始化。在接下来的步骤6中,根据第一计数器的当前的计数器值来提供过程模型。当前的策略在步骤7中根据所提供的过程模型来选择过程参数。在接下来的步骤8中,在模拟中鉴定或者根据所模拟的过程品质来评价过程参数。相应地在接下来的步骤9中,根据鉴定或评价情况对过程模型进行更新或适配。在接下来的步骤10中,将第一个计数器提高一个步骤。在接下来的步骤11中,根据中止标准(例如过程品质或迭代的次数)来结束步骤6至10的回环,并且在步骤12中进行分支或者分支到步骤6,并且重新连续执行步骤6至10。在步骤12中对奖励函数进行评估,该奖励函数根据所确定的用于所选择的参数的过程品质来分配奖励值(Reward)。在接下来的步骤13中,通过学习方法(例如强化学习)来更新策略(采集函数)。随后,在接下来的步骤14中确定,是否在另一个回环中通过在步骤2中的分支还再次执行模拟,或者是否(例如以学习方法的足够的收敛)结束模拟并且在步骤15中进行分支。

[0032] 在先前的模拟之后,用步骤15开始在实际的系统上进行过程优化。在步骤15中,使第二计数器初始化或者将其设置为零。在接下来的步骤16中,使与第二计数器的计数器水平零相应的过程模型初始化。在接下来的步骤17中,根据第二计数器的当前的计数器值来提供过程模型。在步骤18中,与第二计数器的当前的计数器值相应的当前策略根据所提供的过程模型来选择过程参数。在接下来的步骤19中,在实际的系统中鉴定或者根据实际的系统上的过程品质来评价过程参数。相应地在接下来的步骤20中,根据鉴定或评价情况来更新或适配过程模型。在接下来的步骤21中,将第二计数器提高一个步骤。在接下来的步骤22中,根据中止标准(例如过程品质或迭代的次数)来结束步骤17到22的回环,并且在步骤23中进行分支或者分支到步骤17,并且重新连续执行步骤17到22。

[0033] 在步骤23中,提取用于优化过程的最优的或者得到优化的参数,并且能够用最优的参数来执行技术过程、这里是磨削过程,也就是说,在使用确定的参数的情况下操控或调节技术系统。

[0034] 所描述的方法也能够用于优化机器学习系统、尤其是神经网络。在此,能够通过贝叶斯优化来确定学习系统的参数、尤其是超参数(Hyperparameter)。在贝叶斯优化中,通过采集函数来确定下一个测试点,通过学习方法来自动地确定贝叶斯优化中的采集函数,并且借助于所确定的采集函数在贝叶斯优化中获知所述参数。然后根据所获知的参数来训练机器学习系统(das maschinelle Lernsystem)。机器学习系统、例如神经网络在此尤其是通过计算机来实现、也就是说不仅用于确定采集函数的计算还有用于训练过程的计算都通

过处理器来执行。学习系统的参数被存放在机器可读的存储器中。

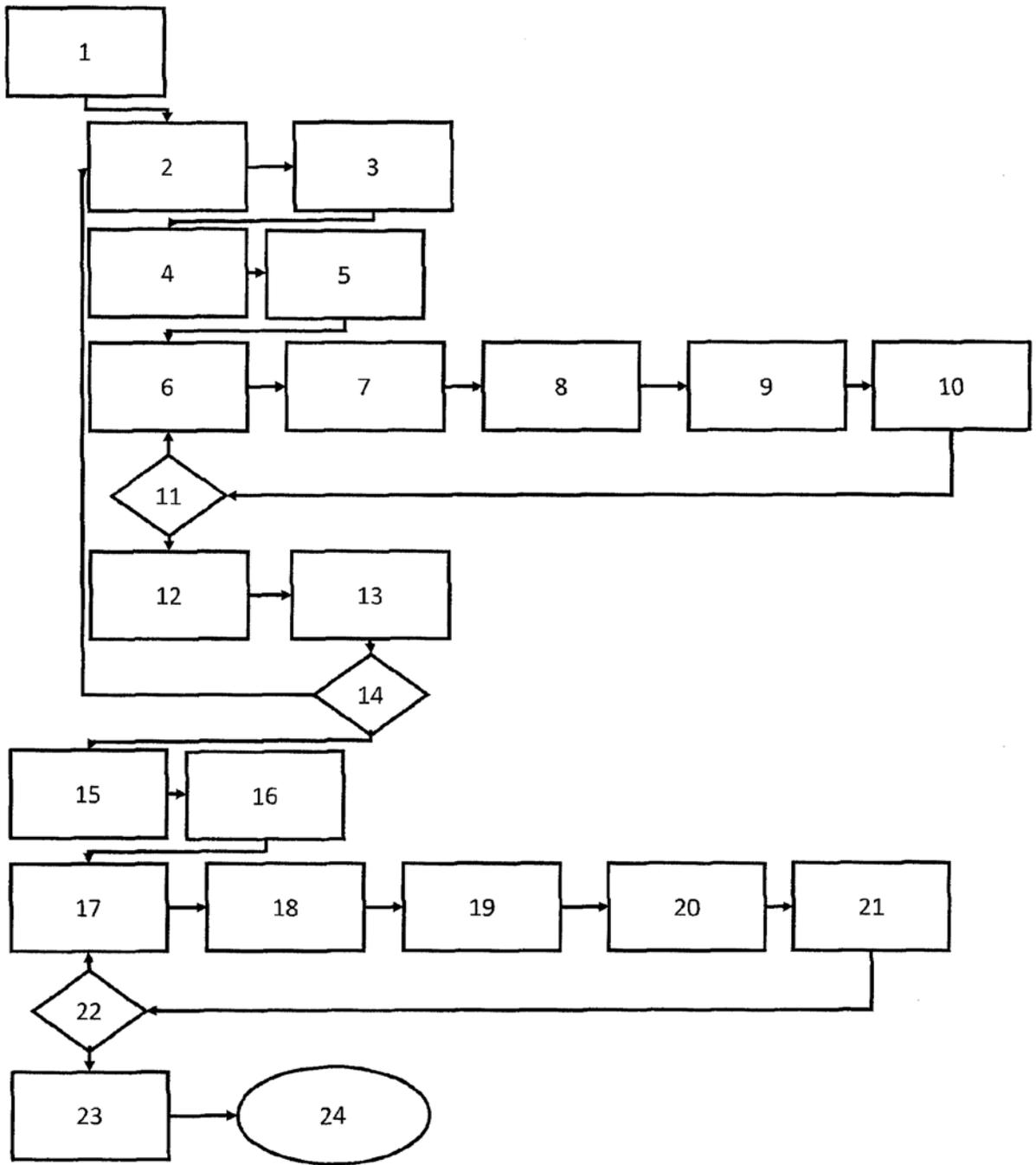


图 1