



(12) 发明专利申请

(10) 申请公布号 CN 105082132 A

(43) 申请公布日 2015. 11. 25

(21) 申请号 201410335997. 6

(22) 申请日 2014. 07. 15

(30) 优先权数据

14/285, 867 2014. 05. 23 US

(71) 申请人 通用汽车环球科技运作有限责任公司

地址 美国密歇根州

(72) 发明人 J. W. 威尔斯 D. W. 佩顿  
R. M. 尤伦布洛克 L. 库

(74) 专利代理机构 北京市柳沈律师事务所  
11105

代理人 葛青

(51) Int. Cl.

B25J 9/16(2006. 01)

B25J 19/00(2006. 01)

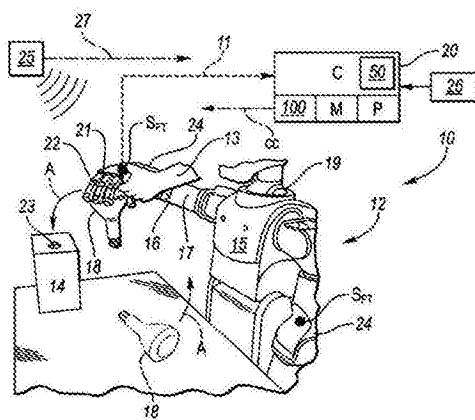
权利要求书2页 说明书10页 附图5页

(54) 发明名称

力 - 扭矩任务的快速机器人模仿学习

(57) 摘要

一种训练机器人以自主执行机器人任务的方法，包括让末端执行器运动通过预定机器人任务的多个状态，以在一组 n 个训练示范中将任务示范给机器人。方法包括测量训练数据，包括在末端执行器运动通过多个状态时经由力 - 扭矩传感器测量至少直线力和扭矩。从训练数据提取关键特征，其被分段为控制基元的时间序列。识别相邻时间序列段之间的转变。在同一任务的自主执行期间，控制器检测转变且在控制模式之间自动地切换。机器人系统包括机器人、力 - 扭矩传感器、编程为执行方法的控制器。



1. 一种用于训练机器人以自主执行机器人任务的方法,上述机器人任务需要通过机器人的末端执行器对物体施加直线力和扭矩,方法包括:

让末端执行器运动通过预定机器人任务的多个状态,以由此在 n 个训练示范的组中将预定机器人任务示范给机器人;

测量一组训练数据,包括在末端执行器运动通过预定机器人任务的多个状态时经由力 - 扭矩传感器测量至少直线力和扭矩;

经由控制器从测量的一组训练数据提取关键特征,包括将所述测量的一组训练数据分段为控制基元的时间序列以及识别时间序列的相邻段之间的转变;

在通过机器人进行的示范任务的随后自主执行期间,经由控制器检测转变;和  
响应于所检测的转变,在多个不同控制模式之间自动地切换。

2. 如权利要求 1 所述的方法,其中让末端执行器运动包括反驱动末端执行器。

3. 如权利要求 1 所述的方法,其中让末端执行器运动包括经由用户输入装置命令末端执行器。

4. 如权利要求 1 所述的方法,其中检测转变包括经由控制器执行评分函数,以计算联机 / 实时数据中的数据点代表其中一个转变的可能性。

5. 一种机器人系统,包括:

机器人,具有末端执行器;

至少一个力 - 扭矩传感器,相对于机器人定位且可操作以测量通过末端执行器施加到物体的直线和旋转力;和

控制器,与机器人和所述至少一个力 - 扭矩传感器通信,且具有处理器和在其上记录了指令的存储器,所述指令用于训练机器人以自主执行机器人任务,所述机器人任务需要通过末端执行器将直线力和扭矩施加到物体,其中控制器配置为执行所述指令以使得控制器进行:

在一组 n 个机器人任务示范期间,记录末端执行器通过多个机器人任务状态的运动;

测量一组训练数据,包括在末端执行器运动通过多个状态时经由至少一个力 - 扭矩传感器测量通过机器人施加到物体的直线力和扭矩;

从测量的一组训练数据提取关键特征,包括将所述测量的一组训练数据分段为分开的控制基元的时间序列以及识别时间序列的段之间的转变;

在通过机器人进行的示范任务的随后自主执行期间检测转变;和

响应于检测的转变在多个不同控制模式之间切换。

6. 如权利要求 5 所述的机器人系统,其中控制器配置为通过记录通过末端执行器的反驱动赋予的运动而记录末端执行器的运动。

7. 如权利要求 5 所述的机器人系统,进一步包括用户输入装置,其编程为在所述一组 n 个任务示范期间命令末端执行器运动。

8. 如权利要求 5 所述的机器人系统,其中机器人包括腕部,且其中力 - 扭矩传感器嵌入腕部中。

9. 如权利要求 5 所述的机器人系统,其中控制基元包括位置控制基元、混合力控制基元、和目标力控制基元,且其中控制器被编程为检测每一个控制基元之间相应的转变。

10. 如权利要求 5 所述的机器人系统,其中控制器通过执行评分函数而检测转变,以由

此计算联机 / 实时数据中的数据点代表其中一个转变的可能性。

## 力 – 扭矩任务的快速机器人模仿学习

### 技术领域

[0001] 本发明涉及力 – 扭矩任务的快速机器人模仿学习。

### 背景技术

[0002] 机器人是电机械装置,其能使用一系列机器人联动件 (link) 来操作物体。机器人联动件通过关节互连,每一个关节可以独立地或互相依赖地被一个或多个伺服驱动器驱动。每一个机器人关节代表独立的控制变量或自由度。末端执行器是直接执行任务 (例如抓持工作工具或叠放部件) 的末端联动件。通常,经由闭环的基于力、阻抗或位置的控制规则将机器人控制到期望目标值。

[0003] 在制造中,存在对更灵活的制造和加工的需要,其能以最小量的停机时间生产新的或更具变化的产品。为了完成该目标,机器人平台应该能快速使其自身适应新的任务,而不需要重新编程和编译的时间。任务示范 (也称为模仿学习) 是用于实现这种性能灵活性的进化方法。然而,现有的基于任务示范的训练方法在所需的训练示范的数量和总的静态计算工作量方面并不理想。

### 发明内容

[0004] 本文公开一种机器人系统和相关的控制方法。系统和方法目的是改善现有的人辅助的任务示范方法,其中人操作者实体地给机器人示范任务,例如经由手动或机器人通过具体轨迹、停止位置和末端执行器动作的自动命令的运动。本方法使得机器人系统中的机器人从给定任务的一小组示范中获得具有任务知识的基元组 (primitive set),其中示范任务具体需要施加直线和旋转力 / 扭矩,以便实现期望的最终结果。具体说,预设方法有助于 (基于不多于两个或三个的总任务示范) 实现机器人数据分析,以便确定有用的不变任务特征组。不变任务特征随后用于通过同一任务的自主执行而引导机器人通过许多变量条件的未来动作。

[0005] 通过本发明的控制器提供的控制架构的实施例包括两个主任务模块,即训练阶段模块 (TPM) 和执行阶段模块 (EPM),其有效地将总的处理分为分开的训练和执行 (后训练) 阶段。本文使用的术语“模块”在计算机领域意味着所有必需的硬件 (例如处理器、数据总线、存储器装置、输入 / 输出 (IO) 装置) 和所有其他必要的相关硬件以及任何下层软件编程指令,其执行最终使得控制器和机器人以下文所述的方式发挥功能。

[0006] 在公开的控制架构中,TPM 允许用户提供小量示例性任务示范,例如正好 1-3 个任务示范。机器人的直线力 – 扭矩传感器 (其可以连接到或嵌入在机器人的腕部或其他合适结构中,且是位于机器人工作环境中的可选环境传感器) 在示范任务 (一个或多个) 期间收集机器人性能数据。从任务示范,控制器确定机器人的末端执行器经过末端执行器相对于任务示范环境中物体的运动范围的位置和取向。控制器还处理来自力 – 扭矩传感器的数据。组合的收集数据组在本文被称为训练数据组且可被暂时地记录在存储器中。

[0007] 使用收集的训练数据组,控制器接下来执行时间分段处理,其中控制器的处理器

将任务示范的时间序列分为不同任务段。即时间序列经由控制器的逻辑模块而被分析,以识别某些段之间的转变事件。该功能可以在控制器的事件描述符训练器模块(EDTM)中实现。可以分析任务段以识别特征事件描述符,如下所述其随后被控制器使用以检测通过机器人进行的任务的实际自主执行中相同或相似的转变。

[0008] EPM(其是本发明控制器的第二主控制模块,用于任务执行)施加控制基元的训练序列,即基本控制动作,例如在任务示范/模仿学习之后让末端执行器朝向或沿目标表面运动。EPM还在任务执行期间施加学习参数和将运行时间数据中的特征匹配所学习的特征,以便精确地确定在不同控制策略之间何时转变,“控制策略”是指用于在相应的任务段中获得给定动作的具体必要控制命令。如上所述的各种概念和术语将参考各种附图在下文详细描述。

[0009] 所公开的发明的益处是能从很少的简单任务示范中快速学习涉及直线和旋转力目标的机器人任务,下文称为力和扭矩目标。在基于训练的以前的方法中,力和扭矩目标可以仅通过消耗时间的统计过程来知晓,统计过程要求数百训练例子以生产统计相关数据。从工厂应用的方面看,本发明可以提供优于这样的统计学方法的优点。

[0010] 本发明也可以提供优于其他模仿学习方法(基于末端执行器的获得位置目标)的优点。通过包括直线力和扭矩目标,本方法提供一种机器人能力,其中给定工作任务的成功完成更取决于检测实际的感知线索,随后获得具体的实体末端执行器位置或取向。示例性工作任务(其中该优势可以是尤其有益的)例如包括抓持和插入电灯泡或其他部件到套筒,或固定使用卡扣连接的部件。在本文使用电灯泡例子是为了说明简单和一致。

[0011] 根据优选实施例,提供一种用于训练机器人的方法,以自主执行机器人任务,上述机器人任务需要通过机器人的末端执行器对物体施加直线力和扭矩,方法包括:末端执行器运动通过预定机器人任务的多个状态,由此在n个训练示范的组中将预定机器人任务示范给机器人;测量一组训练数据,包括在末端执行器运动通过预定机器人任务的多个状态时经由力-扭矩传感器测量至少直线力和扭矩;经由控制器从测量的一组训练数据提取关键特征,包括将测量的一组训练数据分段为控制基元的时间序列和识别时间序列的相邻段之间的转变;在通过机器人进行的示范任务随后的自主执行期间经由控制器检测转变;和响应于所检测的转变在多个不同控制模式之间自动地切换。

[0012] 根据一个优选方面,其中n≤5。

[0013] 根据更优选的方面,其中n≤2。

[0014] 根据一个优选方面,其中检测转变包括检测位置控制基元、混合力控制基元、和目标力控制基元每一个之间的转变,且其中将测量的一组训练数据分段为控制基元的时间序列包括将测量的一组训练数据分段为位置控制基元、混合力控制基元、和目标力控制基元的组合。

[0015] 根据优选实施例,提供一种用于训练机器人的方法,以自主执行机器人任务,上述机器人任务需要通过机器人的末端执行器对物体施加直线力和扭矩,方法包括:反驱动末端执行器通过预定机器人任务的多个状态,由此在一组n个训练示范中将预定机器人任务示范给机器人,其中n≤3;测量一组训练数据,包括在反驱动末端执行器通过预定机器人任务的多个状态时经由嵌入机器人腕部中的力-扭矩传感器测量至少直线力和扭矩;经由控制器从测量的一组训练数据提取关键特征,包括使用测量的一组训练数据的滚动平均

将测量的一组训练数据分段为控制基元的时间序列和识别时间序列的相邻段之间的转变；在随后通过机器人进行的示范任务自主执行期间经由力-扭矩传感器测量至少直线力和扭矩作为联机/实时数据；在通过机器人进行的示范任务随后的自主执行期间检测联机/实时数据中的转变包括检测位置控制基元、混合力控制基元、和目标力控制基元每一个之间的转变；和响应于检测的转变在多个不同控制模式之间自动地切换。

[0016] 优选地，其中检测转变包括执行评分函数，所述评分函数计算联机/实时数据中的数据点代表其中一个转变的可能性。

[0017] 在下文结合附图进行的对实施本发明的较佳模式做出的详尽描述中能容易地理解上述的本发明的特征和优点以及其他特征和优点。

## 附图说明

[0018] 图1是机器人系统的示意图，其具有机器人和用于经由有限数量的力和扭矩任务示范训练机器人的控制器。

[0019] 图2是具有学习控制基元的示例性力-扭矩任务的示意图。

[0020] 图3是示例性训练数据的时间曲线图，其形式是图1所示的机器人的示例性力测量，垂直轴线显示力且水平轴线显示时间，且训练数据被分段和分类为控制基元的序列。

[0021] 图4A-C共同描述了根据三个可能情况的图1所示机器人的末端执行器的各示意性轨迹。

[0022] 图5是用于图1所示控制器的系统架构的示意图。

[0023] 图6是可作为本文公开方法一部分使用的坐标变换步骤的示意图。

## 具体实施方式

[0024] 参考附图，其中几幅图中相同的附图标记指示相同或相似的部件，在图1中显示例性机器人系统10。机器人系统10包括机器人12和控制器(C)20，所述控制器被编程为执行方法100，且由此经由人辅助的任务示范训练机器人12来执行力-扭矩任务。机器人12可以包括躯干15、具有末端执行器21的臂16，且可能地在示例性拟人实施例中包括头部19。末端执行器21可以配置为任何合适的装置，以用于执行示范任务，例如抓持器或具有指状物22的拟人的手，所述指状物22附接到腕部24或其他臂段17，如所示的。这样的实施例中指状物22可以是马达驱动的指部、延伸部或其他抓持器。

[0025] 非限制性的示例性力-扭矩工作任务（出于简单和一致的目的在下文使用）是对电灯泡形式的物体18的抓持。然而，可在本发明的范围内使用任何数量的其他力-扭矩任务。人操作者13（出于简单的目的，在图1中仅显示其手臂）可简单地通过示范任务对机器人12示范新的力-扭矩工作任务。例如，在图1所示的例子中，操作者13通过将电灯泡抓持且插入到套筒形式的固定结构14中而手动地引导机器人12，如通过箭头A的轨迹示出的。

[0026] 经由操作者13通过机器人12施加的直线力和扭矩（即旋转力）在任务示范期间通过一个或多个力-扭矩传感器(S<sub>FP</sub>)测量，所述传感器定位在机器人的末端执行器21处或附近，例如嵌入在腕部24中。通过以这种方式反驱动机器人12而进行的任务示范允许操作者13在任务的示范过程中感觉和施加适当的力和扭矩，例如相对于固定结构14抓持、

旋转和定位物体 18，例如将物体 18 插入固定结构 14 的套筒 23 中。即操作者 13 抓持给定的力 - 扭矩传感器 ( $S_{FT}$ ) 后方的腕部 24，从而力 - 扭矩传感器 ( $S_{FT}$ ) 可检测与操作者 13 相同的力。还有，操作者 13 可以伸缩机器人 12，在这种情况下操作者 13 观察而不是感觉力，例如根据各种显示器或软件提供的读数。

[0027] 图 1 的控制器 20 包括逻辑模块 50，其包括两个主控制模块：训练阶段模块 (TPM) 和执行阶段模块 (EPM)。参考图 5 在下文详细描述 TPM 和 EPM。通常，逻辑模块 50 进行两个不同阶段的操作，即任务训练阶段（阶段 I）和任务执行阶段（阶段 II）。在任务训练阶段期间，操作者 13 提供期望任务的相对小量的示例性任务示范，术语“小”在本文的意义是在一个实施例中总共不多于五个任务示范，在大多数情况下不多于两个或三个的示范就足够。

[0028] 从该相对有限数量的训练例子，控制器 20 从力和扭矩传感器  $S_{FT}$  获得力 - 扭矩信号（箭头 11）以及末端执行器 21 的位置和取向的有关细节，例如经由性能数据（箭头 27）。性能数据（箭头 27）被一个或多个额外传感器 25 收集且从其输出，例如关节角度传感器、视觉系统传感器、点云摄像头、和 / 或类似物。性能数据（箭头 27）可以包括末端执行器 21 相对于物体 18 运动的数据追踪。任务示范期间收集的数据的组合在本文中被称为训练数据组（箭头 11T）。

[0029] 简要地参见图 5，力 - 扭矩逻辑模块 50 包括训练阶段（阶段 I）和执行阶段（阶段 II），如上所述。图 1 的控制器 20 接收训练数据组 11T 且经由基元分段模块 (PSM) 52 对接收的训练数据组 11T 执行分段操作。这种操作可以包括将被示范任务的时间序列分为不同的活动段，其中每一个作为结果的段对应于图 1 所示的机器人 12 的单个控制模式或“任务基元”。

[0030] 简要地参见图 3，提供这种分段的例子。给定任务示范的量 ( $F_M$ ) 以其三维形式绘制，如示例性测量的力分量  $F_x$ 、 $F_y$  和  $F_z$ ，同时针对时间 ( $t$ ) 绘制测量的力分量，以产生示范任务的时间序列 48。其他示范数据可以包括机器人 12 和 / 或其末端执行器 21 的位置、取向、扭矩和关节角度。出于简单的目的，在图 3 的例子中仅显示力。

[0031] 分段过程可以形成任务段 S1、S2、S3 和 S4。在测量的力中，阈值改变，例如阈值可被控制器使用，以区别在段 S1 和 S2、S2 和 S3、S3 和 S4 之间的过渡，例如整个示范任务跨过持续时间  $t_0 - t_4$  的情况下在  $t_1$ 、 $t_2$  和  $t_3$  发生的那样。线  $L_1$ 、 $L_2$ 、 $L_3$ 、 $L_4$  指示了每一个段 S1-S4 的相对力水平，例如 X、Y、Z 分量的滚动平均值。尽管在图 3 的例子中显示了力，但是图 1 的控制器 20 可以对例如位置、取向和扭矩这样的值执行相似的操作。

[0032] 再次参见图 5，在已经确定如图 3 中的活动段（其共同限定任务基元序列 (TPS) 58）之后，图 1 的控制器 20 接下来分析这些段，以识别段之间的过渡事件，例如  $t_1$ 、 $t_2$ 、 $t_3$  处在图 3 中的垂直线。这可以在事件描述符训练器 (EDT) 60 中实现，例如处理这些段的计算机模块，以获得一组特征事件描述符 (CED) 68。特性事件描述符 68 随后被控制器 20 使用，以在任务执行阶段（阶段 II）期间检测之前示范任务的实际执行中的控制模式或段之间的转变。

[0033] 图 5 的逻辑模块 50 的其余结构和操作细节描述如下。这种描述以任务训练阶段（阶段 I）中的训练数据组 11T 的处理的描述为开始，且随后描述构建特征事件描述符 68 过程中得到的任务基元序列 58 的使用，以便允许图 1 的机器人 12 执行之前示范的力 - 扭矩

任务。

[0034] 再次参见图 1, 控制器 20 可以包括任何所需的硬件和处理指令, 其适用于执行本方法 100, 且按照需要用于输出控制信号 (箭头 CC) 到机器人 12, 例如执行例如抓持和插入物体 18 到套筒 23 中的自主任务的命令, 如之前通过操作者 13 示范的。控制器 20 可以实施为一个或多个数字计算机或主机, 其每一个具有一个或多个处理器 (P) 和存储器 (M), 即有形的、非瞬时存储器, 例如光学或磁只读存储器 (ROM), 以及随机存取存储器 (RAM), 电可编程只读存储器 (EPROM) 等。

[0035] 控制器 20 也可以包括高速时钟、模拟数字 (A/D) 电路、数字模拟 D/A 电路、和任何所需的输入 / 输出 (I/O) 电路、I/O 装置、通信界面、信号调节和缓冲器电子器件等。输入装置 26 可以与控制器 20 分开或集成在一起。输入装置 26 可以是鼠标、控制杆或其他控制装置, 适用于在图 5 的任务训练阶段 (阶段 I) 通过人示范的任务来遥控机器人 12。

#### [0036] 任务示范

[0037] 参见图 2, 示例性基元序列 30 示出了基本训练例子。在图 1 的操作者 13 与机器人 12 关联地执行任务 (通过手动地反驱动机器人 12、通过经由来自输入装置 26 的输入信号遥控操作机器人 12 或两种情况都有) 时, 任务示范数据通过控制器 20 而被记录。图 1 的机器人 12 的状态 (包括末端执行器 21 的位置、其取向、抓持器 / 工具状态和力 / 扭矩传感器信号 (箭头 11)) 被周期性地记录在存储器 M 中。

[0038] 例如, 在图 2 中, 图 1 的机器人 12 可以在第一点 P<sub>1</sub> 开始, 且随后经由接近轨迹 33 直线地运动到第二点 P<sub>2</sub>, 直到图 1 的末端执行器 21 在第二点 P<sub>2</sub> (例如墙壁上的点) 处接触第一表面 32。在第二点 P<sub>2</sub> 处, 末端执行器 21 可以以力 F<sub>1</sub> 压靠第一表面 32, 且可能地以扭矩 T<sub>1</sub> 绕第二点 P<sub>2</sub> 旋转, 所述扭矩 T<sub>1</sub> 以与力 F<sub>1</sub> 同样的方式被收集。

[0039] 接下来, 图 1 的操作者 13 可以将末端执行器 21 从第二点 P<sub>2</sub> 沿第一表面 32 通过运动轨迹 35 滑动到第三点 P<sub>3</sub>。末端执行器 21 可以施加力 F<sub>2</sub> 直到形成与第二表面 34 (例如地面) 的接触。图 1 的末端执行器 21 可能这样运动: 通过第二扭矩 T<sub>2</sub> 而绕第三点 P<sub>3</sub> 旋转, 且其后经由离开轨迹 37 运动到第四点 P<sub>4</sub> 以由此完成示例性任务基元序列 30。图 2 的示范的任务基元序列 30 仅仅是说明性的。本领域技术人员应理解力 - 扭矩任务的其他组合可以以相似的方式示范。

#### [0040] 基元库

[0041] 再次参见图 5, 在本文使用的术语“行为基元”描述相对简单的机器人动作。行为基元依赖于可被图 1 的机器人 12 执行的一些控制参数。可存在多于一种类型的行为基元, 其可以从可用任务基元的基元库 (L<sub>p</sub>) 56 中选择。每一个任务基元类型可以具有下列性能: (1) 相关的机器人控制器, 例如图 1 的控制器 20, 其能执行任务基元, (2) 用于在任务示范数据的给定时间段将控制参数调试为“最佳拟合”的逻辑, 和 (3) 用于评估性能 (2) 最佳拟合控制参数和示范数据段之间相对成本或误差的逻辑, 图 3 显示了示例性任务段 S1-S4。

[0042] 在具体实施例中, 图 1 的控制器 20 可以使用三个不同基元类型的基元库 56: 自由运动、约束力运动、和目标力运动。将依次描述这些示例性基元类型每一个。

#### [0043] 自由运动

[0044] 术语“自由运动”是指基于位置的运动, 其不取决于力 - 扭矩传感, 例如图 2 中从第一点 P<sub>1</sub> 到第二点 P<sub>2</sub> 的运动。自由运动包括图 1 的末端执行器 21 从开始位置和取向到目

标位置和取向的控制,例如经由位置的线性内插和取向的四元数插值。

[0045] 本发明的控制器 20 从给定时间段的第一和最后时步 (time step) 提取位置和取向参数,形成三维 (3D) 笛卡尔空间中的线  $l_p$  和 3D 翻滚 - 俯仰 - 摆动空间 (roll-pitch-yaw space) 中的  $l_r$ 。n 时步的段上的成本可以如下计算:

$$[0046] \quad cost(prim) = \sum_{i=1}^n \alpha \|P_{rlp\perp}(pos_i)\| + \beta \|P_{rlp\perp}(rot_i)\| + \gamma \|force_i\|$$

[0047] 其是收集的数据和线性近似值之间总几何误差加上对任何观察到的力的校准罚值 (calibrated penalty)。

#### [0048] 约束力运动

[0049] 术语“约束力运动”是指混合力 - 位置值,其具有运动所沿的一个方向和维持恒定约束力的另一方向。这对模拟沿表面滑动或拖曳行为来说是适当的基元,例如沿表面 32 发生在图 2 的第二点  $P_2$  和第三点  $P_3$  之间的运动。针对该基元以与自由运动参数中发生同样的方式提取位置和取向参数。还提取针对约束力向量  $f_c$  的约束值,作为所述段期间观察的力的方向和平均大小。由此在 n 时步的段上确定成本,例如如下:

[0050]

$$cost(prim) = \sum_{i=1}^n \alpha \|P_{rlp\perp}(pos_i)\| + \beta \|P_{rlp\perp}(rot_i)\| + \gamma \|force_i - f_c\| + \delta I(\|force_i\| \leq f_{thresh})$$

[0051] 其是位置和取向与其线性近似之间的总几何误差加上来自约束力的偏差,加上力大小低于阈值 ( $f_{thresh}$ ) 时的罚值。

#### [0052] 目标力运动

[0053] 术语“目标力运动”在本文用于另一基元,其是描述直到目标力被保持的运动的混合力 - 位置值。该基元适用于模拟插入或接触行为。图 1 的控制器 20 以与如上所述的自由运动参数相同的方式提取位置和取向参数,且还提取目标力作为运动结束期间观察的平均力,具体说是在图 1 的末端执行器 21 被定位在结束位置的距离阈值中时。n 时步的段上运动的成本可以如下计算:

$$[0054] \quad \sum_{i=1}^n \alpha \|P_{rlp\perp}(pos_i)\| + \beta \|P_{rlp\perp}(rot_i)\| + \gamma \|force_i - f_c\| \cdot I(\|pos_i - pos_n\| \leq pos_{thresh}) \\ + \delta (\|force_i\| \cdot I(\|pos_i - pos_n\| > pos_{thresh}))$$

[0055] 其是位置和取向与其线性近似之间的总几何误差加上来自运动结束期间目标力的偏差,加上运动结束之前用于任何力的罚值。

#### [0056] 优化

[0057] 在图 5 中,在训练阶段 (阶段 I),控制器 20 使用优化模块 (OB) 54。在该逻辑阶段,控制器 20 分析给定任务的示范数据和产生任务基元序列 (TPS) 58,即将用于模仿和进一步分析示范任务的行为基元。控制器 20 经由 TPS58 找出对任务示范数据最佳拟合的基元序列,且可以针对全部基元、其类型、基元何时开始和结束以及其每一个控制参数来优化该基元序列。

[0058] 用于优化模块 54 的一个可行方案是迭代爬山优化的一种变体,其是本领域已知的过程。在该方法中每一次爬山从随机产生的基元序列和转变时间开始。在每一次迭代中,

图 1 的控制器 20 对解决方案施加离散数量的修改，评估每一个经修改的解决方案的成本，且用最低成本的修改替换解决方案。一些修改可以依赖于比例因数，其可在检测到局部最小值时增加，以发现更好的序列。在最大比例因数下发现局部最小值时结束爬山。从不同的随机初始该状态执行少次数爬山，以找到总的最小值。

[0059] 在 N 段上具有 i 任务基元的序列 S 的成本可以如下计算：

$$[0060] \quad cost(S) = \sum_{i=1}^N cost(prim_i, seg_i) + \alpha N$$

[0061] 其中相对于具有许多步骤的更复杂序列， $\alpha$  用于支持具有更少步骤的序列。可在优化期间在序列上执行的修改是将两个基元之间的转变时间向前或向后移动 x 个时步，其中 x 为比例因数，将两个相邻基元合并为覆盖这两段的一个基元，将一段分为两个小段，或将一个基元类型换为另一个。

[0062] 简要地参见图 4A-C，示出了用于给定的简化二维任务示范的几个可能的段，X 和 Y 维度绘制在相应的水平和垂直轴线上。如上所述的优化效果显示在图 4C 中。线 42A 显示了图 1 的示例性的电灯泡插入任务期间自由空间中末端执行器 21 的位置的轨迹。线 41A 显示了基元序列，任务基元之间的转变通过点 43 标记。图 4A 显示了随机初始化基元序列，其对示范数据是相对差 (poor) 的匹配。图 4B 显示了另一任意序列，其更好地匹配任务示范，但是其使用许多不必要的基元。图 4C 显示了通过图 5 的优化模块 54 产生的优化段，其更接近地匹配训练数据且使用少量基元。

[0063] 事件描述符训练器

[0064] 再次参见图 5，事件描述符训练器 (EDT) 60 代表在事件已经发生时的力 - 功信息，且用于检测应该导致机器人 12 动作改变的有意义的力 - 扭矩事件。因为如上所述，示范的任务在不同任务阶段被分段为不同运动基元，所以重要的发现是正好在这些转变必须发生之前发生的特征感知事件。这些事件随后在任务被执行时变成可触发运动基元之间转变的信号。示例性实施例中的事件描述符训练器 60 可以包括三个部分：(1) 凸点检测器 (SPD) 62，(2) 力 - 扭矩描述符产生器 (FTDG) 64，和 (3) 描述符选择器 (DS) 66。

[0065] 凸点检测器 62 在训练数据组 11T 中找出时间点，在该点处事件更可能发生，且将该点标记为凸点。力 - 扭矩描述符产生器 64 随后接收该信息 63 且产生用于每一个凸点的力 - 扭矩描述符 65。力 - 扭矩描述符产生器 64 可以通过将凸点的力 - 扭矩数据映射到不同坐标系（其更好地区分不同事件）来操作。

[0066] 描述符选择器 66 随后在所有训练例子中将相似的力 - 扭矩描述符 65 分组且将凸点力 - 扭矩描述符 67 的组输出。每一个组可以被赋予评分 (score)。在该事件下最可能由凸点构成的组随后被控制器 20 选择。特征事件描述符 (CED) 68 随后从凸点力 - 扭矩描述符 67 产生。将依次描述这些要素每一个的细节。

[0067] 凸点检测器

[0068] 图 5 的阶段 I 中的凸点检测器 62 运行为在训练数据组 11T 中找出事件更可能发生的时间点。这些时间点在控制器 20 的逻辑中被标记为凸点。被标记为凸点的时间点越多则所需训练时间越长。例如，图 1 的机器人系统 10 可以使用以下等式，以将从给定传感器 (i) 收集的原始力 - 扭矩数据  $D_i$  中的凸点进行标记：

$$[0069] \quad Salient(t) = \begin{cases} true, & \text{if } \exists i: \Delta D_i(t+1) \cdot \Delta D_i(t) \leq 0 \\ false, & \text{otherwise} \end{cases}$$

[0070] 上述等式在时间点 t 处标记凸点, 在该时间点处任一维度下的力或扭矩值形成顶点或底点。图 5 的凸点检测器 62 实质上用作一类第一阶段过滤器, 其将“不可能具有”的事件的时间点过滤掉。

[0071] 力 - 扭矩描述符产生器

[0072] 力 - 扭矩描述符产生器 64 产生用于通过凸点检测器 62 检测的每一个凸点的力 - 扭矩描述符 65。为了产生力 - 扭矩描述符 65, 图 1 中收集的原始力 - 扭矩数据的部分使用固定时间窗被首先提取, 上述时间窗具有以凸点 s 为中心的长度 L。分段的原始数据 W(s) 为  $L \times K$  矩阵, 其中 K 为使用的传感器的数量。F-T 描述符 65 随后通过将用于每一个凸点的分段的原始数据 W(s) 进行转变而形成, 所使用的步骤如下:(1) 维度增加, (2) 解除偏置, (3) 坐标变换, 将描述这些步骤每一个。

[0073] 维度增加 : 在该可选步骤中, 来自图 1 的机器人 12 的示例性的六维(六轴线)原始数据的维度被增加, 包括额外信息, 例如力 - 扭矩的衍生或末端执行器 21 的运动方向。包含该步骤可以增加公开本文的方法的总体鲁棒性。

[0074] 解除偏置 : 不同力 - 扭矩传感器具有不同偏开值, 如本领域已知的。解除偏置数据 U(s) 可以如下获得 :

$$[0075] \quad U_{lk}(s) = W_{lk}(s) - \frac{\sum_{l=1}^L W_{lk}(s)}{L}, \forall l \forall k$$

[0076] 其减去了分段数据中用于每一个传感器维度的分段的原始数据  $W_{lk}(s)$  的平均值。

[0077] 坐标变换 : 简要地参见图 6 所示的示意性变换 90, 该步骤通过将解除偏置的原始数据 U(s) 92(例如收集的用于力分量  $F_x, F_y, F_z$  的原始力数据) 变换到过滤器坐标 96 而产生力 - 扭矩描述符 65。过滤器坐标 96 使用校准的一组过滤器 94 作为输入基础。过滤器坐标 96 中每一个基础  $f_t$  具有与原始数据的时间窗相同的长度。Haar 小波集、Daubechies 小波集和经设计的非正交小波集对该步骤来说都是有效的替换。变换等式可以是如下的 :

$$[0078] \quad D(s) = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix} \cdot U(s)$$

[0079] 其中 n 为使用的过滤器的数量。变化到过滤器坐标的数据为产生的力 - 扭矩描述符 65。

[0080] 描述符选择器 : 再次参见图 5, 描述符选择器 (DS) 66 将通过力 - 扭矩描述符产生器 64 产生的相似的力 - 扭矩描述符 65 分组成多个组  $G_j, j \in J$ , 其中 J 为组的总数量。代表被训练事件的组被图 1 的控制器 20 基于评分功能和现有的知识来选择。事件描述符 68 随后从包含在选择的组中的力 - 扭矩描述符 65 产生。将依次描述这两个步骤每一个(即描述符的分组和描述符的选择)。

[0081] 将描述符分组 : 通过图 5 的力 - 扭矩描述符产生器 64 产生的所有力 - 扭矩描述符 65 被自动地重新分组为多个组  $G_j$ 。每一个组含有来自每一个训练例子的一个力 - 扭矩描

述符 65。产生所有可能的组。每一个组随后被赋予类似的评分，取决于这些力 - 扭矩描述符 65 在欧几里得空间中彼此有多靠近。用于此的示例性等式如下：

$$[0082] \quad \text{similarity score}(G_j) = \frac{1}{\sum_{i=1}^M \sum_{k=1}^M \|D_i - D_k\|_2}, \quad D_i, D_k \in G_j$$

[0083] 其中 M 为被使用的训练例子的数量。

[0084] 描述符选择：在该步骤中代表被训练的力 - 扭矩事件的事件组从在先步骤中产生的所有组中选择。事件组选择基于事件评分功能，其是有关事件更可能发生的时间点的在先评分和相似性评分的函数。

[0085] event score = similarity score(G<sub>j</sub>) • prior score(G<sub>j</sub>)

$$[0086] \quad \text{prior score}(G_j) = \prod_{D \in G_j} \frac{1}{1 + e^{-\alpha \frac{t_D}{L_D}}}$$

[0087] 在先评分 G<sub>j</sub> 为 s 型函数的乘积，其在分段的训练数据端部具有更高可能性。变量 α 是常数，变量 t<sub>D</sub> 是描述符 D 的时间点，且变量 L<sub>D</sub> 是含有描述符 t<sub>D</sub> 的具体训练段的长度。在先函数可用提供额外信息的任何函数代替。事件描述符 E 通过计算具有最高评分的组中描述符的平均值而产生，即：

$$[0088] \quad E_{ij} = \frac{\sum_{D \in G} D_{ij}}{M}$$

[0089] 执行阶段

[0090] 参考图 5 的阶段 II，任务执行阶段，仅图 1 的机器人系统 10 的两个或三个训练例子就足以获得足够信息来执行示范任务。通过重复之前所学习的控制基元序列，且随后使用用于这些基元的所学习的力 - 扭矩约束和目标，从而运行执行阶段。为了精确地知道在学习的序列中何时从一个基元切换到下一个，机器人系统 10 还依赖于检测转变事件的能力。从任务训练阶段（阶段 I），事件描述符 69 已经被学习，其对每一个转变来说是具体的，且这些事件描述符 69 将允许事件检测器（ED）70、另一计算机或逻辑模块确定何时触发行为控制模块（BCM）86，例如控制器 20 的行为控制逻辑和硬件，且由此切换到新的控制策略。

[0091] 行为控制模块 86 被编程为或配置为管理图 1 的机器人 12 的动作的执行。行为控制模块 86 使用如前所述的三个控制策略，自由运动、约束力运动、目标力运动，以控制机器人 12 的动作。这些控制策略每一个具有其自身的具体反馈控制需求，其在学习的基元分段阶段被参数化。基元之间的切换通过来自事件检测器 70 的事件触发。

[0092] 事件检测器

[0093] 事件检测器 70 使用从阶段 I 的事件描述符训练器 68 产生的事件描述符 69，以检测实际的 / 联机数据中的具体事件。因为每一个事件与被训练的事件描述符 68 相关，以检测不同事件，所以需要不同的事件描述符 68。阶段 II 中的事件检测器 70 包括两个部分：另一力 - 扭矩描述符产生器（FTDG）72（与阶段 I 的 64 处所示的不同）和事件分类器（EC）74。力 - 扭矩描述符产生器 72 产生力用于联机 / 运行时间数据的扭矩描述符 73。事件分类器 74 随后将来自联机数据的力 - 扭矩描述符 73 是否是阶段 I 中通过事件描述符训练器 60 在先训练的具体事件进行分类。

[0094] 力 - 扭矩描述符产生器

[0095] 阶段 II 的力 - 扭矩描述符产生器 72 从联机数据 85 产生力 - 扭矩描述符 73, 例如从图 1 的传感器 S<sub>FT</sub> 和 25, 其在图 5 中被共同示出为传感器套件 82。联机数据 85 被设定为具有等于事件描述符训练器 60 中使用的时间窗长度的时间长度。力 - 扭矩描述符 72 通过在阶段 I 的事件描述符训练器 60 的力 - 扭矩描述符产生器 64 中应用相同的三个步骤而产生, 即维度增加、解除偏置、和坐标变换, 如上所述。

[0096] 事件分类器

[0097] 阶段 II 的事件分类器 74 将联机数据 85 是否为与事件描述符 E 相同的数据进行分类, 即通过特征事件描述符 68 从阶段 I 输出事件描述符 69。这可如下获得:(1) 评分函数可以用于计算联机数据代表事件的可能性。评分函数可以使用通过力 - 扭矩描述符产生器 64 产生的力 - 扭矩描述符 65(D<sub>T</sub>) 和事件描述符 E 之间的欧几里得距离:

$$[0098] \text{Scoring function} = \frac{1}{\|D_T - E\|_2}$$

[0099] (2) 如果评分函数的输出超过校准的分类阈值, 则联机数据被分类为事件, 其在图 5 的决定模块 75 发生。

[0100] 分类阈值可以使用事件描述符训练器 60 和事件检测器 70 两者经由训练数据的“弃一法交叉验证 (leave one out cross validating)”而通过控制器 20 学习。给予最小数量误报的值随后被选择为分类阈值。通过经由图 1 的控制信号 (箭头 CC) 从其促动器 84 而来的命令动作, 机器人 12 随后运动通过用于完成所命令之前学习的任务的必要轨迹。由此, 总的任务被表示为通过参数化基元 (具有如上所述的特征事件描述符) 限定的运动段序列。在执行时, 通过调用相关的控制基元执行初始运动段, 直到通过事件分类器 74 检测到事件。事件必须匹配与当前段的结束对应的事件描述符。在该事件被触发时, 控制器 20 自动地切换到下一段。

[0101] 使用上述方法, 图 1 的操作者 13 能在比少量示范更多地示范任务, 且让机器人 12 从有限数量的示范中学习。如此, 操作者 13 能为控制器 20 传授任务知识, 而不必是机器人编程或机器人控制方面的专家。图 1 的系统 10 和相关的方法 100(经由图 5 的逻辑模块 50 举例说明)可以在不能用常规的工业机器人技术执行的各种情况下实施。

[0102] 尽管已经对执行本发明的较佳模式进行了详尽的描述, 但是本领域技术人员可得知在所附的权利要求的范围内的用来实施本发明的许多替换设计和实施例。

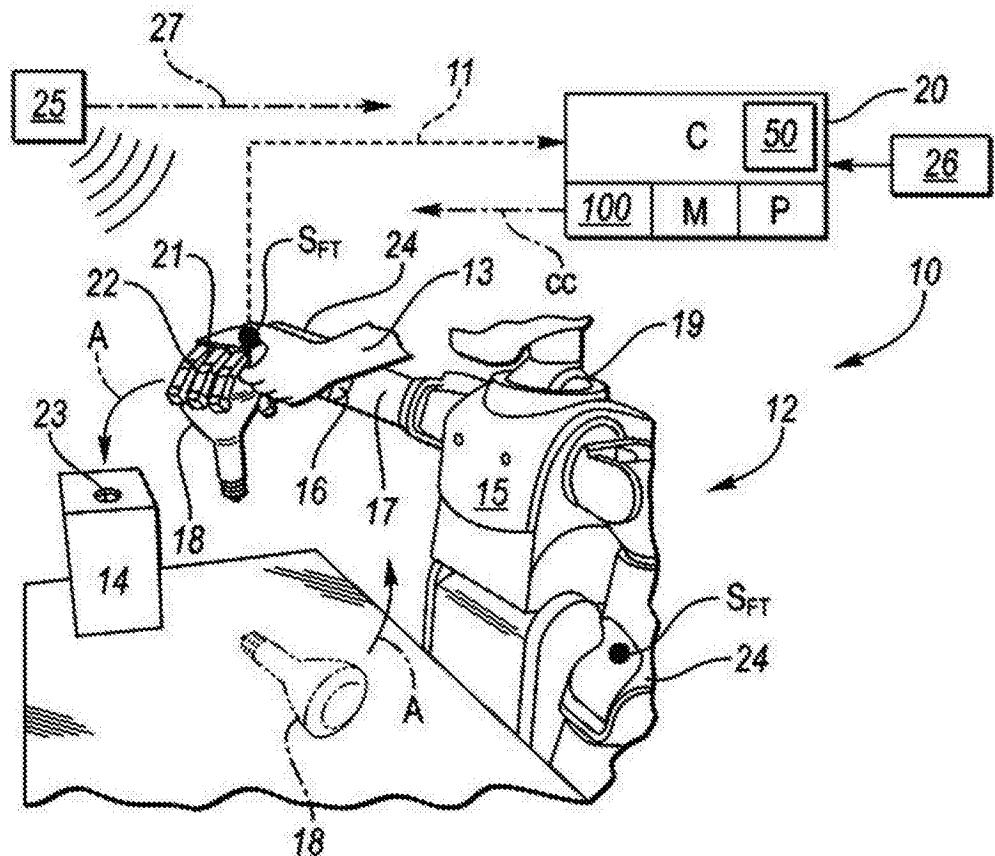


图 1

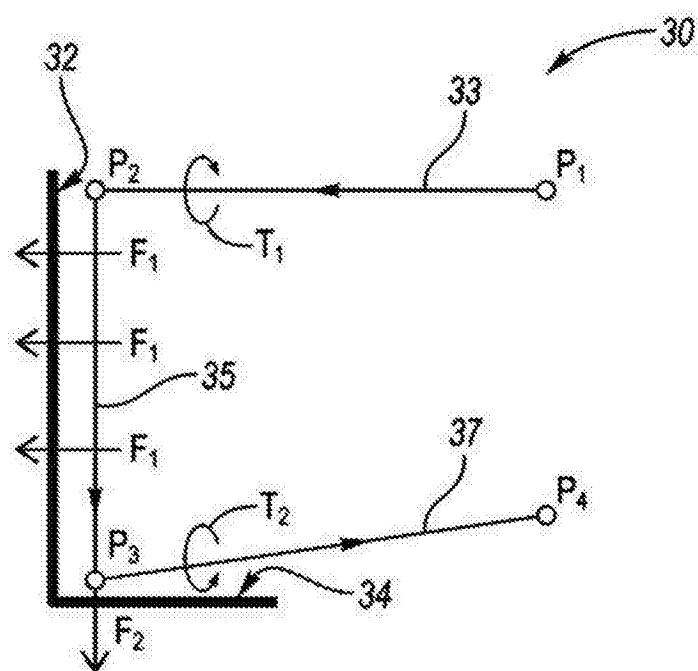


图 2

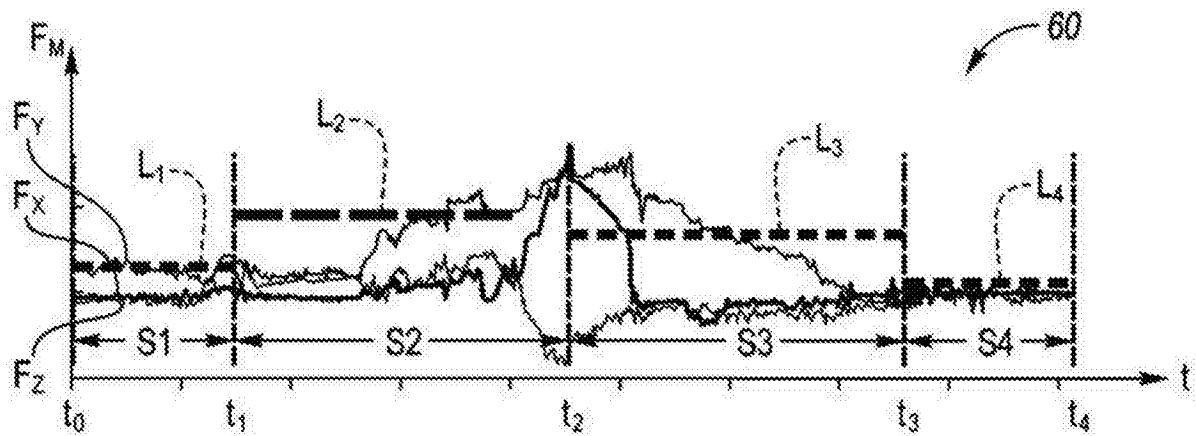


图 3

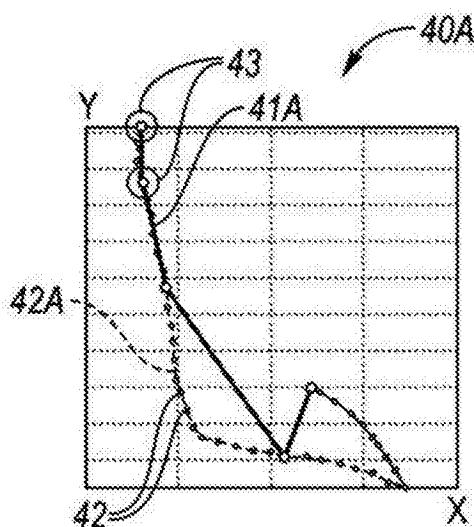


图 4A

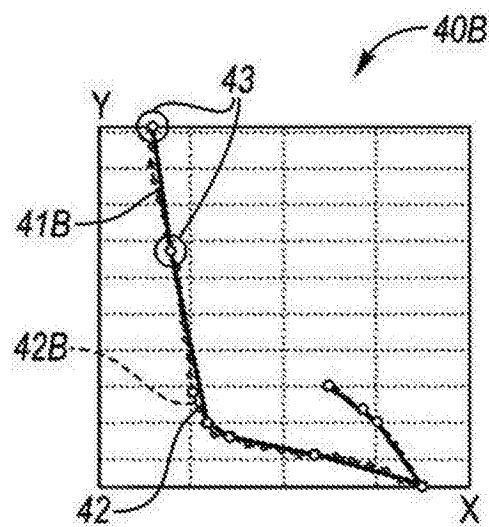


图 4B

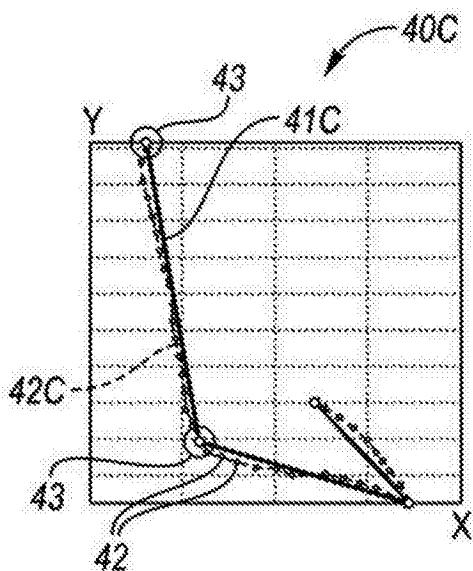


图 4C

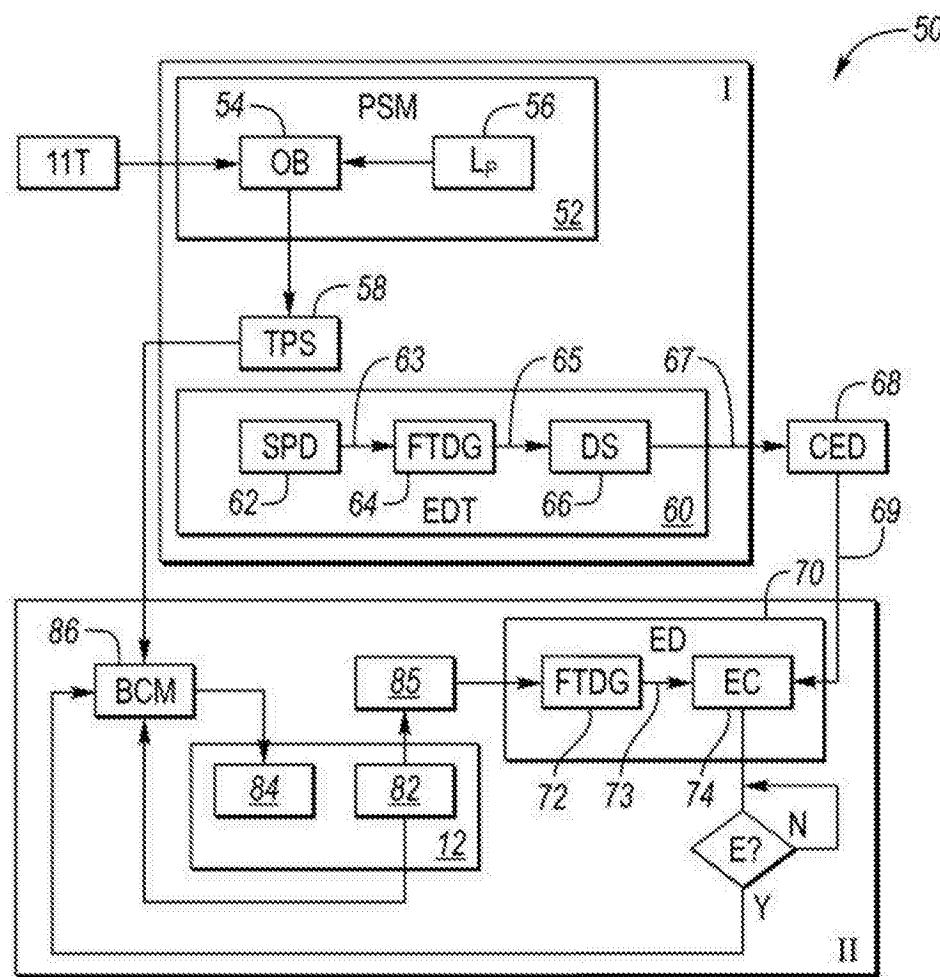


图 5

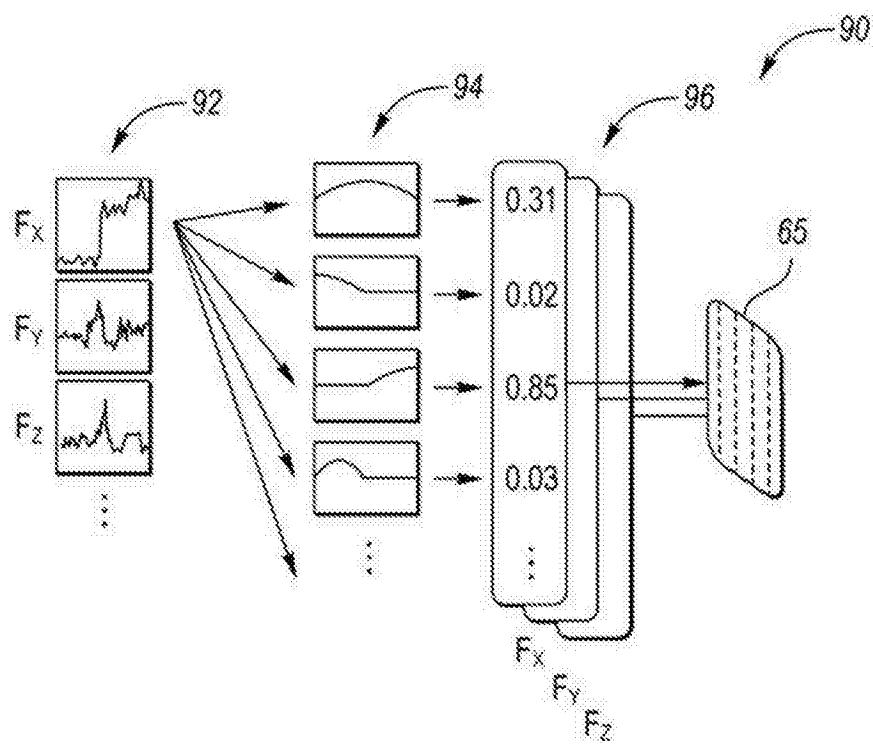


图 6