



(12) 发明专利

(10) 授权公告号 CN 111832880 B

(45) 授权公告日 2024.04.16

(21) 申请号 202010304391.1

(22) 申请日 2020.04.17

(65) 同一申请的已公布的文献号
申请公布号 CN 111832880 A

(43) 申请公布日 2020.10.27

(30) 优先权数据
19169995.8 2019.04.17 EP

(73) 专利权人 ABB瑞士股份有限公司
地址 瑞士巴登

(72) 发明人 B.施密特 M.霍伦德 F.伦德斯

(74) 专利代理机构 中国专利代理(香港)有限公司 72001
专利代理师 李啸 刘春元

(51) Int. Cl.
G06Q 10/0639 (2023.01)

(56) 对比文件
CN 103842924 A, 2014.06.04
JP 2005135010 A, 2005.05.26
KR 101507103 B1, 2015.04.01
US 2014214374 A1, 2014.07.31

US 2015112498 A1, 2015.04.23

US 2018260560 A1, 2018.09.13

US 7249049 B1, 2007.07.24

CN 102298319 A, 2011.12.28

CN 109416534 A, 2019.03.01

CN 106845397 A, 2017.06.13

CN 107045283 A, 2017.08.15

CN 1379887 A, 2002.11.13

US 2016170406 A1, 2016.06.16

US 2017301093 A1, 2017.10.19

US 2017097635 A1, 2017.04.06

邢永康.一种基于Markov链模型的动态聚类方法.《计算机研究与发展》.2003,(第2期),全文.

张鹏伟.面向小批量生产过程的多变量贝叶斯控制图.《系统工程学报》.2013,(第3期),全文.

周昊飞.基于深度置信网络的大数据制造过程实时智能监.《中国机械工程》.2018,(第10期),全文. (续)

审查员 蒲韵霜

权利要求书1页 说明书25页 附图16页

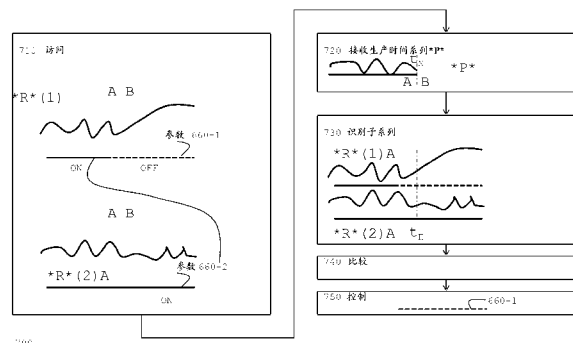
(54) 发明名称

对正在进行的生产批运行的质量指示符的计算机实现的确定

(57) 摘要

为了控制执行生产过程(200)的生产批运行(220)的技术设备(110),控制器模块采用如下方式执行方法(700):它访问具有来自先前执行的批运行的数据的参考时间系列(*R*(1)、*R*(2))。参考时间系统(*R*、*R*(1)、*R*(2))与技术设备(110)的参数(660-1、660-2)有关。在技术设备执行生产批运行(220)时,模块接收(720)具有数据的生产时间系列(*P*)、识别(730)参考时间系列(*R*、*R*(1)、*R*(2))的子系列(*R*(1) A、*R*(2) A),并且比较(740)所接收的时间系列(*P*)和参考时间系列的子系列(*R*(1) A、*R*(2) A)。

(2)A)。这得出相似性或非相似性的指示。在相似性情况下,在生产批运行的持续期间通过使用参数(660-1)作为控制参数和/或通过指示来自参考时间系列的参数作为给操作者的推荐,模块控制(750)技术设备。



CN 111832880 B

[接上页]

(56) 对比文件

VO THI THANH VAN.时间序列分析技术的研究.《中国博士学位论文全文数据库 信息科技辑》.2014,I138-32.

Chu, Fei.Researches on statistical performance monitoring and control of batch processes based on bdkpca method.《Northeastern University (People's Republic of China)》.2008,全文.

Yu, Chunyang.Group-based production

scheduling for make-to-order production.《JOURNAL OF INTELLIGENT MANUFACTURING》.2015,第26卷(第3期),585-600.

Aminata Kane.“Multivariate Time Series Representation and Similarity Search Using PCA”.《ADVANCES IN DATA MINING: APPLICATIONS AND THEORETICAL ASPECTS,ICDM 2017》.2017,第第10357卷卷第122-136页.

1. 一种计算机实现的方法 (700), 用于控制执行生产过程 (200) 的生产批运行 (220) 的技术设备 (110), 其中所述技术设备 (110) 以来自数据源 (120-v) 的集合的时间系列的形式提供数据, 其中所述数据源 (120-v) 与所述技术设备 (110) 有关, 所述方法 (700) 包括:

访问 (710) 具有来自所述生产过程 (200) 的先前执行的批运行 (210) 的数据的参考时间系列 (*R*、*R*(1)、*R*(2)), 其中所述参考时间系列 (*R*、*R*(1)、*R*(2)) 与所述技术设备 (110) 的参数 (660-1、660-2) 有关;

在所述技术设备 (110) 执行所述生产批运行 (220) 时:

接收 (720) 具有数据的生产时间系列 (*P*),

识别 (730) 所述参考时间系列 (*R*、*R*(1)、*R*(2)) 的子系列 (*R*(1)_A、*R*(2)_A),

比较 (740) 所接收的时间系列 (*P*) 和所述参考时间系列的所述子系列 (*R*(1)_A、*R*(2)_A), 从而得到相似性或非相似性的指示, 在相似性情况下, 在所述生产批运行 (210-B) 的持续期间通过使用所述参数 (660-1、660-2) 作为控制参数来控制 (750) 所述技术设备 (110),

其中对所述生产过程 (200) 的多个先前执行的批运行以及对所述生产批运行执行对参考时间系列 (*R*) 的步骤访问 (710), 并且其中通过以下来执行步骤比较 (740): 对属于多变量时间系列的单变量时间系列的连续数据值之间的预定义改变的出现进行计数、将计数器存储到矩阵内、并且确定所述矩阵之间的相似性度量。

2. 如权利要求1所述的方法 (700), 其中所述参考时间系列的所述子系列 (*R*(1)_A、*R*(2)_A) 通过所述参考时间系列与所述生产时间系列 (*P*) 之间的时间规整对准来识别。

3. 如权利要求1所述的方法 (700), 其中对带有所分配的质量指示符的参考时间系列执行对所述参考时间系列 (*R*、*R*(1)、*R*(2)) 的访问 (710)。

4. 如前述权利要求中的任一项所述的方法 (700), 其中对带有所分配的质量指示符的参考时间系列执行对所述参考时间系列 (*R*) 的访问 (710), 所分配的质量指示符与所述控制参数 (660-1) 具有预定义关系。

5. 如权利要求1所述的方法 (700), 其中识别 (730) 子系列并且比较 (740) 的步骤用经转换的参考时间系列 {R} # 和生产时间系列 {P} # 来执行, 其中所述转换是从多变量时间系列到单变量时间系列的转换。

6. 如权利要求1所述的方法 (700), 其中确定所述相似性度量包括确定从以下选择的任一个: 所述矩阵的元素上的差的绝对值的总和、曼哈顿距离、余弦相似性、以及莱文斯坦距离。

7. 一种适于执行如权利要求1-6中的任一项所述的方法 (700) 的计算机系统 (600)。

8. 一种计算机可读介质, 其上存储有计算机程序, 所述计算机程序在被加载到计算机的存储器内并且被所述计算机的至少一个处理器执行时, 执行如权利要求1-6中的任一项所述的计算机实现的方法的步骤。

对正在进行的生产批运行的质量指示符的计算机实现的确定

技术领域

[0001] 本公开一般涉及生产过程,并且更特定地,本公开涉及用于确定生产过程的批运行(batch-run)的质量指示符的计算机系统、方法和计算机程序产品。

背景技术

[0002] 在工业上,技术系统执行生产过程。要期望,生产过程和所得的产品两者都符合(或“遵从”)预定义的规范。然而,情况并不总是如此。因此,可以对生产过程的特定执行(或“批运行”)分配质量类别并且可以对特定产品分配质量类别。

[0003] 简而言之,质量指示符可以至少区分符合生产(conforming production)和不符合生产(non-conforming production)。符合通常与指示符“成功”相关联并且不符合通常与指示符“失败”相关联。质量指示符也可以区分得到符合产品的生产和得到不符合产品的生产。还可以使用另外的质量类别(例如,“第一选择”、“第二选择”)。质量指示符表示执行生产过程的技术系统的内部状态。

[0004] 在生产期间收集数据是信息源,其在被正确评估时可以导致(在进程的过程中的)改进。数据可以由测量信号(例如,生产设备的温度)、由与生产事件有关的控制指令(例如,用于打开或用于关闭特定阀、用于添加物料)或由状态指示符而产生。

[0005] 随着在诸如化工之类的特定工业中广泛采用批处理,可以收集单个批的数据。关于批控制的规约(convention)诸如在ANSI/ISA-88及等同标准(例如,IEC 61512-1:1997、IEC 61512-2:2001、IEC 61512-3:2008、IEC61512-4:2009)中被标准化。

[0006] 对于批处理,数据可以用作时间系列,即对后续时间点按时间顺序指数化的数据值系列。时间系列与特定批有关和/或与所得的产品有关。

[0007] 评估数据可以包括检测来自不同批运行的时间系列之间的相似性。如果相似性指数超过特定阈值,则可以认为相似性。或者,在另外的示例中,如果数据处理识别模式,诸如随时间而变的特性数据值,则可以认为相似性。

[0008] 来自过去的特定批运行(“历史批运行”或“参考批运行”)的时间系列可以充当参考时间系列。可以对参考分配一个或多个质量类别。为了保持上文提到的简化类别,可以按符合生产、不符合生产、符合产品或不符合产品对参考时间系列进行分类。另外的简化只使用成功和失败类别。在生产和产品两方面(以及潜在地另外的更详细的规范)都符合的特定批运行可以被视为“黄金批(golden batch)”。

[0009] 如本文使用的,来自特定批运行(其正在进行)时间系列是生产时间系列。生产时间系列与参考时间系列之间的相似性可以指示类别,诸如符合生产、不符合生产(符合产品或不符合产品),从而期望达到“黄金批”。

[0010] 然而,检测时间系列之间的相似性并不像比较数字与阈值或类似物那么容易。存在至少两个约束:

[0011] 着眼于第一约束,作为时间系列基础的时间区间不是都一样的,即使生产过程相同。批处理并不意味用相同时间长度(或持续时间)执行每个特定生产批运行。生产批运行

的持续时间通常因批而异。对于不同的持续时间存在许多原因。例如,化学反应由于环境状况变化而造成具有可变持续时间。操作者潜在地由于后勤原因(例如,油箱已满、下一个设备被占用)而造成使过程暂停并且在后来的时间点恢复它们。操作者动作也具有可变持续时间。

[0012] 着眼于第二约束,数据不包括随时间而变的一个数据值,诸如提到的温度,但数据通常源于多得多的源。对于另外的物理现象可以存在诸如压力之类的测量值,可以存在诸如电机旋转速度、阀的打开或关闭状态之类的参数。另外的参数指提到的控制指令(例如,用于打开/关闭阀、添加物料)和/或用于状态指示符。换言之,数据是多变量数据。

[0013] 动态时间规整(DTW)是用于彼此比较时间系列并且使时间系列彼此对准的算法的总称。

[0014] 对DTW和对DTE软件的概述例如在[1]Toni Giorgino: "Computing and Visualizing DynamicTime Warping Alignments in R: the dtw Package(计算且可视化R中的动态时间规整对准:dtw包)"(Journal of Statistical Software第31卷(2009),第7版)中可获得。更简而言之,DTW允许比较时间系列,即使时间基础不同也如此。技术可用于考虑时间差异。例如,引用[1]的图1图示,可以对准两个时间系列以及可以通过调查对准距离来计算相似性。引用[1]还解释了适应多变量数据的方法。

[0015] 然而,存在另外的约束:DTW算法中的一些可能忽略时间系列的一些特性,使得例如无法识别指示(批运行或产品的)特定失败的特性模式。这样的非期望效应称为“过度激进(over-aggressive)规整”。换言之,特性模式可以指示在生产期间发送的问题,如果数据处理不够灵敏,则特性模式可能消失。

[0016] 再进一步地,使批运行与质量类别相关可能太迟,尤其当类别指示失败的时候也如此。提到的约束可能造成延迟。

[0017] 存在对找到考虑这些约束的方法的需要。

发明内容

[0018] 技术设备在参考和生产批运行中执行生产过程。该技术设备与多个数据源有关并且以多变量时间系列的形式提供数据。数据源与技术设备有关:数据源可以是设备的部分,或数据源对应于流向设备或从设备流出的数据。多变量时间系列包括源特定单变量时间系列。

[0019] 为了解决对生产批运行分类中的潜在迟滞(lateness),计算机单独对生产过程的后阶段识别质量类别。以二阶段(two-phase)方法开始,后续阶段是第一阶段和第二阶段(或初始阶段和最终阶段)。

[0020] 在具有第一阶段和第二阶段的类别的二阶段方法中,处理参考数据,从而用阶段特定类别来区分多个参考批运行。简而言之,类别可以是二进制类别。可选地,可以应用语义,诸如失败/成功、失败/失败、成功/成功以及成功/失败。参考数据还包括已应用于参考批运行的技术参数的集合。这些是参考条件。

[0021] 根据该参考数据,计算机获得类别转变条件,对于第一阶段的特定质量类别到第二阶段的特定质量类别的转变。换言之,两个阶段的质量类别经由通常在参数向量中可用的至少一个质量改变参数而彼此有关。

[0022] 参考条件可以分化为y到x、y到y、x到x、x到y的条件。(在语义情况下:失败到成功条件、失败到失败条件、成功到成功条件和成功到失败条件)。

[0023] 尽管在生产批运行的第一阶段执行特定生产过程,但计算机确定一个或多个参考批运行的第一阶段的相似性(或非相似性)。然后通过参考的第一阶段相似而识别质量类别,计算机导出生产批运行的第一阶段的质量类别。

[0024] 计算机然后应用参考条件与根据生产数据可用的条件(生产条件)相比较。要注意,在该时间点,第二阶段的生产条件尚不应用于生产过程。

[0025] 计算机然后将生产批运行的第一阶段和参考条件(中的至少一些)的质量类别传递给技术设备的操作者。这是技术设备作为技术系统的状态的指示。

[0026] 传递质量类别可以通过识别在第一阶段具有相同质量类别并且在第二阶段已经导致成功质量的参考批运行而完成。

[0027] 所识别的参考批运行可以经由用户界面来传递,可视化地例如通过显示(所识别的参考批运行的)时间系列的轨迹。

[0028] 高亮转变条件是可能的。高亮失败条件(作为警报或类似物)和高亮成功条件(作为推荐或类似物)之间存在不同。

[0029] 如果第一阶段已导致失败,则存在失败到成功条件和失败到失败条件。操作者然后可以选择合适的条件(即,参数)来控制技术设备,通常是导致失败到成功转变的条件。如果应用失败到成功条件不可能,则操作者可以取消生产批运行。

[0030] 如果第一阶段已导致成功,则存在成功到成功条件和成功到失败条件。操作者然后可以选择合适的条件(即,参数)来控制技术设备,通常是导致成功到成功转变的条件。操作者将避免设备受成功到失败条件的参数所控制。

[0031] 从操作者的角度来看,指示第一阶段的质量类别是监测的结果,并且指示参考数据作为转变条件允许操作者整体预测生产阶段的成果。

[0032] 换言之,计算机执行计算机实现的方法来识别初始阶段结束时或后续阶段开始时的质量类别。知道初始阶段的质量类别和知道历史参考(对于初始和后续阶段)的(技术设备的)操作者可以在后续阶段开始时干预生产过程。

[0033] 这是有利的,因为操作者可以引入校正动作。

[0034] 将生产过程分成两个(或以上)后续阶段便于解释。然而不要求在特定时间点划分过程。(这甚至将会很困难,因为不同的批运行具有不同的持续时间或批运行时间。)

[0035] 用来自生产批运行的第一阶段的数据来动态划分阶段,这是可能的,所述数据包括被收集直到计算机开始与参考数据比较的时间点的数据。

[0036] 计算机可以重复比较来更新质量类别。

[0037] 用于重复比较的前提条件之一是

[0038] ●选择相对节省计算的算法,其在条件仍然可以作为参数而应用于第二阶段(即,实时要求)时的时间点,第一次提供质量类别,该算法诸如例如包括比较单变量时间系列而代替比较多变量时间系列的算法。

[0039] ●识别时间系列内的特性部分,来分离阶段,使得可以以阶段特定方式获得参考数据。

[0040] 确定相似性可以使用某一方法,通过该方法将多变量时间系列(来自生产过程的

参考执行或执行的部分)转换为单变量的参考时间系列。相似的转换可以应用于生产执行的多变量时间系列,使得可以通过比较单变量时间系列来确定相似性。

附图说明

- [0041] 图1图示了用于两个后续批运行中的生产过程的执行的概述;
- [0042] 图2图示了用于两个后续批运行中的生产过程的过程的概述,并且还图示了计算机实现的方法;
- [0043] 图3图示了可应用于后续批运行的写规约(writing convention)的概述的表;
- [0044] 图4图示了计算机的框图;
- [0045] 图5图示了执行生产过程的技术设备,其中多变量时间系列中的数据是结果之一;
- [0046] 图6图示了在使用源特定因子的情况下将多变量时间系列转换为单变量时间系列的计算机执行方法步骤;
- [0047] 图7图示了由参考批运行和生产批运行产生的经转换的时间系列的评估,其中经转换的时间系列是单变量时间系列;
- [0048] 图8图示了计算机实现的方法的流程图;
- [0049] 图9-11通过示例图示了用于解释机器学习来获得源特定因子的时间系列的轨迹;
- [0050] 图12图示了还未完成但在控制中的生产批运行的用例场景;
- [0051] 图13图示了用于确定相似性指数(或其他指示符)以用于比较多变量时间系列的方法;以及
- [0052] 图14图示了可以与这里描述的技术一起使用的通用计算机装置和通用移动计算机装置的示例。

具体实施方式

- [0053] 写规约
- [0054] 人的参与由他们的职能来定义。如本文使用的,“操作者”是与技术设备交互的人。“用户”是与计算机交互的人。操作者在各种场合可以变成用户,例如当他或她使用来自计算机的信息(诸如状态信息、质量类别等)来改变与技术设备的交互时。关于(技术设备的)内部状态的信息使操作者能够采取校正动作。在一些情形中,用户可以是专家用户。专家用户可以从操作者那里获得专门的知识。
- [0055] 时间系列可以具有诸如由“单”花括号{}符号化的“单变量”或由“双”花括号{{}}符号化的“多变量”等性质。通过转换已获得“单变量”性质的时间系列可以由#标记。对于不重要的多种变量的时间系列由* *给出。
- [0056] 时间点由带指数k的“ t_k ”符号化。时间区间可以作为由方括号的闭区间给出,如在 $[t_1, t_k]$ 中,其中极限点 t_1, t_k 属于该区间。除非另外陈述,否则连续时间点之间的持续时间(“时隙”)相等: $\Delta t = t_{k+1} - t_k$ 。
- [0057] 1概述
- [0058] 图1图示了两个后续批运行210和220中的生产过程200的执行的概述。生产过程200由技术设备(参阅图4的110)执行。首先出现的是参考批运行210,并且其次出现的是生产批运行220。通过执行生产过程200,技术设备提供来自源的集合的数据。数据源与技术设

备(参阅图4)有关:数据源可以是设备的部分,或数据源对应于流向设备或从设备流出的数据。

[0059] 由于批运行序列,来自参考批运行210的数据是与生产批运行220的数据有关的历史数据。

[0060] 两个批运行提供多个变量中的数据。在图1中,数据以图表形式由轨迹图示,并且数据还图示为多变量时间系列 $\{R\}$ (来自参考批运行210的参考数据)和 $\{P\}$ (来自生产批运行220的生产数据)。在示例中,“多变量”简化为“双变量”,轨迹因此具有指数1和2。

[0061] 如由时间系列的形状不同的轨迹图示的那样,生产批运行220可以得到生产数据,其与参考数据相似或不相似。

[0062] 使用来自引用[1]的计算技术或使用其他技术,计算机实现的比较431(水平箭头)在 $\{P\}$ 与 $\{R\}$ 之间是可能的。该“多到多”比较得到可能与质量类别有关的陈述(例如“相似”或“不相似”)。例如,确定 $\{P\}$ 与 $\{R\}$ 之间的相似性可以指示生产批运行220具有与参考批运行210相同的质量类别(不管类别是“成功”还是“失败”)。

[0063] 将连同图13解释比较的备选方法。

[0064] 如将解释的那样,计算机实现的转换410和420(垂直箭头,细节在图4-5中)可以将多变量时间系列 $\{R\}$ 和 $\{P\}$ 转换成单变量时间系列 $\{R\}^{\#}$ 和 $\{P\}^{\#}$ 。换言之,这是“多到单”转换。后续比较430也导致结果陈述。

[0065] 因为多到单转化410/420(这里图示为“双到单”转换)本质上去除了信息,因此转换410/420必须保留比较430所需要的尽可能多的信息。描述通过时间系列的“特性部分”来描述该信息或将该信息描述为轨迹的“特性形状”。特性部分或形状的序列可以被视为“签名”(因为轨迹看起来像人的签名)。

[0066] 因为批运行210和220通常具有不同的持续时间:“批运行区间”,因此比较时间系列(转换410/420之后的多变量比较431,或单变量比较430)可以包括对准(具有细节的引用[1])。为了使解释简单化,要假设生产批运行已完成,但本文的教导可以应用于部分批运行,即持续的运行。用于处理部分批运行(实时)的特定示例连同图12解释。

[0067] 转换410/420随着转换因子向量 610^* 而变,这里图示为向量(对于 $v=1$ 和 $v=2$, α_1 、 α_2),其具有对应于两个轨迹(参阅指数1和2)的两个因子。原则上,存在两个可选的备选方案来获得转换因子向量 610^* :

[0068] ●通过手动评估参考数据 $\{R\}$,其中连同图6来解释,或通过

[0069] ●如贯穿该描述的主要部分而解释的机器学习(ML),其中从多个候选因子向量选择转换因子向量 610^* ,其中向量 610^* 是训练的结果(受监管,或不受监管)。

[0070] 可选地,转换因子向量 610^* 可以作为计算机模块——因子模块的输出而获得,该因子模块执行机器学习(ML、虚线框、图4和图8中具有在图9-10中的示例的概述)。

[0071] 作为输入,因子模块使用来自多个另外的参考批运行210-R'和210-Q'的数据。此外,因子模块确定这些参考批运行的特性。可选地,特性的确定可以通过与提到的专家用户交互来执行。

[0072] 因子模块执行步骤序列,转换和对准步骤在这些步骤之中(在环中)。因子模块重复或并行执行序列,并且如果检测到遵从预定义的精度条件,则序列的执行结束。

[0073] 2带有更多细节的概述

[0074] 图2带有更多细节图示了两个后续批运行210和220(即,“批运行”)中的生产过程200的执行的概述。生产过程200由技术设备(参阅图4)执行。如图示的,时间从左向右推移。如提到的,首先出现的是参考批运行210,并且其次出现的是生产批运行210。

[0075] 从操作者的角度来看,生产批运行220是已经完成的“当前”批运行(在区间 $[t_1, t_N]$ 末端的时间点 t_N 处)。操作者有兴趣知道(生产批运行220的)质量类别,并且如果可能的话,还有兴趣知道正在生产的产品的质量类别。

[0076] 在实施例(参阅图12)中,在有机会修改生产批运行(正在进行的)的情况下,操作者可以确定在较早的时间点的质量类别。换言之,在时间点 t_N ,操作者需要获得作为执行批运行的技术系统的设备的状态指示符。潜在地在操作者设置的不同条件下,在 t_N 后设备可以继续执行该批运行(根据状态指示符)。

[0077] 批运行210、220得到可以标记为矩阵的数据。简而言之,可以根据源(指数 v)并且根据离散时间点来区分单个数据值,使得矩阵是多变量时间系列 $\{\{\}\}$ 。

[0078] 参考批运行210得到以已经提到的多变量时间系列 $\{\{R\}\}$ 形式的参考数据,如由方框210右侧的箭头所符号化的那样。质量类别 Q_R 也可以分配给参考批运行210。 Q_R 可以具有不同的属性,例如二进制类别具有属性“成功”和“失败”。

[0079] 取决于优化目标(对于生产过程而言),也可以使用其他质量类别。例如,使用提到的批运行区间作为准则,批运行可以归类为长批运行或短批运行(另外的类别是可能的)。参考数据可以用于估计能耗(例如,电能、热、压缩空气),使得具有“低”、“中”或“高”消耗的批运行。涉及事件(诸如设备组件失败)的参考数据可以潜在地帮助识别使设备尽可能长地保持可操作的并且避免设备的应力、磨损等的批运行(和设备参数)。

[0080] 要注意,在参考批运行210结束后可能不会立即知道 Q_R ,但 Q_R 在以后的时间点变得可获得。

[0081] 相似地,生产批运行220得到以多变量时间系列 $\{\{P\}\}$ 的形式但还不在质量类别 Q_p 中的生产数据。要期望,在生产执行220(参阅图11)期间或之后不久(参阅图1)尽可能快地识别质量类别 Q_p 。

[0082] 如通过参考431(参阅图1)所图示的那样,计算 $\{\{R\}\} \odot \{\{P\}\}$ 可以得到相似性指数 S (或分数)。运算符 \odot 使已知方法符号化,参阅上文提到的引用[1]。

[0083] 如果确定了相似性(例如,通过比较指数和阈值,或以其他方式),则生产批运行220可以与具有相同属性的 Q_p (如参考批运行210)相关联。例如,对于 Q_R =“成功”和 $\{\{R\}\}$ 与 $\{\{P\}\}$ 之间的相似性,生产批运行220也与 Q_p =“成功”相关联。

[0084] 假设被简化。在二进制指示符的情况下,与成功-参考的相似性可以导致与成功相关联,而与成功-参考的非相似性并未自动导致与失败相关联。

[0085] 但比较受上文提到的约束的影响。

[0086] 图2还至少部分图示了计算机实现的方法400。方法400包括转换步骤410/420的序列并且包括比较步骤430。另外的步骤在图8中被图示。比较步骤430使单变量时间系列 $\{R\}$ 与 $\{P\}$ 之间(而不是多变量时间系列之间)的计算机实现的比较符号化。比较步骤430可以包括相似性指数 S' 的计算。

[0087] 运算符 $\odot\#$ 使对于比较步骤430的已知方法的使用符号化,而且使另外的方法(其可应用于单变量时间系列)的备选使用符号化。使时间系列对准可以是比较的部分。

[0088] 可以使用指数 S' 来代替指数 S ,用于确定相似性(或非相似性)。取决于参考的质量类别 Q_R ,也可以识别生产批运行的质量类别 Q_P 。

[0089] 如将在下文(图4-5)详细解释的,转换41/420包括使数据值与源特定因子 α_v 相乘并且根据离散时间点对相乘的值求和。在图2中,相乘由“*”符号化,并且求和由“ Σ_v ”(v=1至V)符号化。

[0090] 转换410得到经转换的参考时间系列{R}#并且转换420得到经转换的生产时间系列{P}#,两个都是单变量时间系列。步骤430代表这些单变量时间系列{}#之间而不是多变量系列{()}(如在431中)之间的计算机实现的比较。

[0091] 3写规约

[0092] 图3图示了可应用于批运行210、220(参阅图1)的另外的写规约的概述的表。一般来说,数据D可以分化为参考数据R(涉及参考批运行210,参阅图1)和生产数据P(涉及生产批运行200,参阅图1)。一般化为D是方便的,因为在一些步骤(诸如转换410/420)中,数据处理是相似的。

[0093] 因此对数据值使用小写字母“d”、“r”和“p”。变量指数v通过类型(从v=1到v=V)来识别数据源并且对于R和对于P都是共同的。时间指数一般由“k”(从k=1到k=K)表示,或对于参考数据R由“m”(m=1到m=M)表示并且对于生产数据P由“n”(n=1到n=N)表示。由于潜在地不同的区间,很方便地引入了分化,如在上文解释的那样。批运行区间由 $[t_1, t_k]$ (一般)或 $[t_1, t_m]$ 和 $[t_1, t_n]$ (特别对于R和对于P)表示。

[0094] 有时,为了描述机器学习,描述使用 R' 和 Q' 而代替R,但数据点将被称作r和q(没有撇号)。

[0095] 4计算机和计算机模块

[0096] 图4图示了计算机600(或计算机系统)的框图。如本文使用的,术语“计算机”以单数给出,并且术语“计算机”一般代表适于执行计算机实现的方法的步骤的实体。该规约被简化。

[0097] 更详细地,计算机600可以具有专门针对特定步骤序列(或特定方法,诸如方法300和400,参阅图8)的执行的模块603和604。相似性模块604由两个子模块604-i和604-ii图示。还构思,不同的模块(和/或子模块601-i、601-ii)可以属于不同的物理计算机。

[0098] 为了方便起见,根据模块的主要功能来标记所述模块:数据存储库(模块)650、因子模块603和相似性模块604。

[0099] 将在下文解释数据流(具有到模块的输入和输出数据)。为了简单起见,图4只图示了一些数据流:预处理模块630和存储库650向因子模块603提供时间系列{{R'}}和{{Q'}},并且向相似性模块604提供{{R}}和{{P}}。因子模块603向相似性模块604提供转换因子向量 $610*$,并且相似性模块604提供质量类别 Q_P 。

[0100] 5技术设备

[0101] 图5图示了执行生产过程200(即,“批运行”)的技术设备110,其中多变量时间系列中的数据是结果之一。因为图5可应用于参考批运行210和可应用于生产批运行220(同样可应用于图1中的210-R'和210-Q'),说明由一般首字母缩写D来指代数据。图5还图示了在执行生产过程200期间从数据源120收集数据。

[0102] 图和描述被简化了。如本文使用的,技术设备110(执行参考批运行210)和技术设

备110(执行生产批运行220)在物理上可以是相同设备。这是方便的,但并不要求。还构思,对于参考和对于生产批运行的技术设备可以不同。同样的原理还应用于参考批运行R'和Q',参阅图1。

[0103] 换而言之,提供被用作参考数据的数据的技术设备可以被称作“参考设备”;并且提供要与参考比较的数据的技术设备被称作“生产设备”。因为分化为“参考”和“生产”的术语是相对分化(而不是绝对分化),因此生产批运行可以在将来用作参考。

[0104] 由模块630预处理数据是可选的并且包括归一化。数据在数据存储库650中可用作数据值 d_{vk} 。计算机600(即,因子模块603和相似性模块604)具有对数据存储库650(例如,计算机600的部分中的存储库)的访问。数据存储库650由本领域内已知的计算机存储器和/或数据库实现。在模块603和604接收数据之前执行收集数据(参阅图8)。

[0105] 通过示例,设备110被图示为包括具有搅拌液体115的电机/混合器112、具有加热液体115的加热器113和具有允许添加(或去除)液体115的阀114/116的罐111。

[0106] 设备110还图示有多个V数据源120-1、120-2...120-v...120-V(统称为“数据源120”)。数量V对应于“多变量”。对于是工业设备的技术设备110,数据从不同类型的源变得可获得,在它们之中:

[0107] ●提供测量值(例如,电机的旋转速度、液体的温度、液体量、来自实验室的数据等等)的源,

[0108] ●提供控制指令(例如,用于打开阀来添加液体、用于关闭阀等)的源,或

[0109] ●提供状态指示符(例如,特定阀被打开或被关闭)的源。

[0110] 数据源120可以不同地被实现。例如,测量值和状态指示符确实通常来自传感器。或者,控制指令可以来自控制器计算机(未图示),其控制设备110的操作。不需要数据源120物理连接到设备110。

[0111] 在批运行期间,数据源120提供数据值 d_{vk} (其中指数v识别特定数据源120-v并且指数k识别离散时间点 t_k)。

[0112] 批运行具有包括时间点的区间 $[t_1, t_k]$ 的时间长度(即持续时间),其中

[0113] ●时间点 t_1 ,其中在批运行开始时指数 $k=1$,以及

[0114] ●时间点 t_k ,其中在批运行结束时时间 $k=K$ 。

[0115] 假设在 t_1 之前还未收集数据,并且在 t_k 之后不再收集数据,批运行持续时间可以被计算为 $t_k - t_1$ 。

[0116] 不同的数据源120可以在不同时间点提供数据。例如,传感器可以使用不同的采样率(例如,每分钟感测温度对每秒感测旋转)。或者,当出现特定事件(其改变状态,例如阀从状态“关闭”到状态“打开”)时,状态指示符可以变得可用。

[0117] 本领域内技术人员可以使用内插/外推技术使时间归一化为公共 Δt (因此,还归一化结果公共K、公共横坐标)。内插/外推可以由可选的预处理模块630(单独对于 $v=1$ 到 $V=V$)执行。

[0118] 本领域内技术人员也可以使值归一化(公共纵坐标)。在示例中,数据值 d_{vk} 已归一化为在0(作为最小值)与1(作为最大值)之间的数值。归一化去除了测量单位和与数据相关联的其他信息。归一化可以使用最小/最大值(具有单位)。例如,对于具有每分钟60个周期的最大旋转速度的电机,归一化极值对于最大旋转是 $d_{vk}=1$ 并且对于静止是 $d_{vk}=0$ 。状态数据

可以被归一化,例如对于“阀关闭”是 $d_{vk}=0$ 并且对于“阀打开”是 $d_{vk}=1$ 。归一化可以由可选的预处理模块630执行,或以其他方式执行。

[0119] 数据值 d_{vk} 也可以具有负符号($d_{vk}<0$),但在该示例中这并未图示。

[0120] 着眼于存储数据值 d_{vk} 的数据存储库650,图通过属于曲线图(或“轨迹”)的点来图示它们,所述曲线图(或“轨迹”)在纵坐标具有 $[0,1]$ 值和横坐标具有时间 $[t_1, t_k]$ 。图每时间系列只图示了3个点。轨迹是连接点的线。轨迹是方便的标记,但再次注意数据值 d_{vk} 可用于离散时间点 t_k 。

[0121] 如在下面,具有来自源120-v的数据的单变量时间系列将由 $\{D_v\}$ 表示,所述 $\{D_v\}$ 对于区间 $[t_1, t_k]$ 中的连续时间点具有(归一化)数据值 $\{d_{v1} \cdots d_{vm} \cdots d_{vM}\}$ 。 $\{D_v\}$ 也可以写为具有V列和K行(或V行和K列)的矩阵。 $\{D_v\}$ 是单变量时间系列。

[0122] 在技术设备110的粒度级,多变量时间系列 $\{\{D\}\}$ 指属于生产过程200的特定批运行(210或220)的单变量时间系列 $\{D_v\}$ 的集合。集合可以标记为 $\{\{D\}\} = \{\{D_1\} \cdots \{D_v\} \cdots \{D_V\}\}$ 。

[0123] 在示例中,多变量时间系列 $\{\{D\}\}$ 作为在代表性时间点 t_1, t_k 和 t_k 分别具有数据值 d_{v1}, d_{vk} 和 d_{vk} 的单变量时间系列 $\{D_1\}, \{D_2\}, \{D_v\}$ 和 $\{D_V\}$ 的集合而给出。其他数据值由连接线符号化。单变量时间系列具有特性,诸如在下列示例中:

[0124] $\{D_1\}$ 具有代表性数据值 $d_{11}=0, d_{1k}=1$ 和 $d_{1K}=0.1$,其代表上升到最大值且下降到端值的测量值。 $\{D_2\}$ 具有代表性数据值 $d_{21}=0.2, d_{2k}=0.5$ 并且 $d_{2K}=0.8$,其代表大致线性上升值。要注意,不必达到最大值1。 $\{D_v\}$ 具有代表性数据值 $d_{v1}=0.1, d_{vk}=1$ 和 $d_{vK}=0.1$,其代表具有近似在 t_k 出现的峰值的事件。 $\{D_V\}$ 具有代表性数据值 $d_{V1}=1, d_{Vk}=1$ 和 $d_{VK}=1$,其代表在过程期间未改变的状态。

[0125] 在 $\{\{D\}\}$ 是参考 $\{\{R\}\}$ 的情况下,与 $\{\{P\}\}$ 的相似性必须通过考虑这些特性来检测。例如, $\{R_1\}$ 的上升将在 $\{P_1\}$ 出现, $\{R_n\}$ 中的峰值将在 $\{P_n\}$ 出现等等。然而,特性模式(诸如上升或峰值)并未在相同时间出现(参阅第一约束)。此外,操作多变量时间系列 $\{\{D\}\}$ 意味第二约束,如上文解释的那样。

[0126] 6多到单转换

[0127] 图6图示了在使用源特定因子 α_v 的情况下执行将多变量时间系列 $\{\{D\}\}$ 转换成单变量时间系列 $\{D\}\#$ 的方法步骤410/420的计算机600。一般解释该转换,其中转换 $\{\{D\}\}$ 到 $\{D\}\#$ 代表转换 $\{\{R\}\}$ 到 $\{R\}\#$ 和 $\{\{P\}\}$ 到 $\{P\}\#$,视情况而定。(要注意,转换还可应用于步骤310,图8。一般进行参考600,因为转换可以由模块603和604执行)。

[0128] 转换410/412包括子步骤相乘512和求和514。

[0129] 图示出针对输入(即 $\{\{D\}\}$)的时间图(简而言之,取出纵坐标和横坐标)、针对相乘步骤512的右侧箭头、针对中间结果(即,诸如具有经乘以的数据值的经乘以的时间系列)的时间图、针对求和步骤514指向下的箭头和针对结果(即, $\{D\}\#$)的时间图。

[0130] 在第一子步骤512中,计算机600使值 d_{vk} 与预定义源特定因子 α_v 相乘,从而有经乘以的值 $d_{\sim vk}$ (即,经乘以的值),其是: $d_{\sim vk} = d_{vk} * \alpha_v$ 。相乘可应用于 $k=1$ 到 $k=K$,同样地,不存在 k 上的 α_v 的改变。

[0131] 因子可以在上文提到的转换因子向量610*中总结(参阅图1和图4)并且还可以标记为 $(\alpha_1, \alpha_2 \cdots \alpha_v \cdots \alpha_V)$ 。

[0132] 在示例中, $\{D_1\}$ 乘以因子 $\alpha_1=0.5$ 导致经乘以的时间系列 $\{D_1\}\sim, \{D_2\}$ 乘以因子 $\alpha_2=2$ 导

致经乘以的时间系列 $\{D_2\} \sim$; $\{D_v\}$ 乘以因子 $\alpha_v=0.5$ 导致经乘以的时间系列 $\{D_v\} \sim$; 并且 $\{D_v\}$ 乘以因子 $\alpha_v=0.1$ 导致经乘以的时间系列 $\{D_v\} \sim$ 。

[0133] 在第二子步骤514中, 计算机600通过根据离散时间点对那经乘以的值 $d_{\sim vk}$ 求和而对经乘以的时间系列求和, 即 $d\#_k = \Sigma d_{\sim vk}$ 。

[0134] 求和符号 Σ 被理解为从 $v=1$ 到 $v=V$ 的总和。求和得到经转换的时间系列:

[0135] $\{D\}\# = \{\Sigma d_{\sim v1} \dots \Sigma d_{\sim v2} \dots \Sigma d_{\sim vk} \dots \Sigma d_{\sim vK}\} = \{d\#_1 \dots d\#_2 \dots d\#_k \dots d\#_K\}$ 。

[0136] 经转换的时间系列 $\{D\}\#$ 在下文在图的右侧图示。 $\{D\}\#$ 是单变量。 $\{D\}\#$ 可以被视为来自添加的多个系列 $\{D_v\}$ 的覆盖系列。

[0137] 着眼于 $\{D\}\#$ 的曲线图, 它具有特性曲线, 或打比方来说, 是特性“签名”。如将在下文解释的, “签名”可以被视为特性形状的序列。

[0138] 该签名保留原始多变量时间系列 $\{D\}$ 的大部分特性。签名未保留全部特性, 而保留足够的特性来执行比较430。

[0139] 在示例中, $\{D\}\#$ 已经继承了上升到最大值 (从 $\{D_1\}$ 到 $\{D_v\}$), 并且已经继承了从 $\{D_2\}$ 逐渐增加。它还已经继承了 $\{D_v\}$ 的状态, 尽管对签名的贡献较低。

[0140] 不久返回上文的图1中引入的转换因子向量 610^* , 从而手动评估诸如参考数据 $\{R\}$ 之类的的数据是可能的。简而言之, 因子放大或减少单个时间系列的贡献。在示例中, 本领域内技术人员可以评估具有大致恒定值的时间系列 $\{D_v\}$ 。对 $\{D\}\#$ 的签名的贡献可以被忽视, 因此因子 α_v 可以被设置成0。

[0141] 但在解释ML备选方案中的因子 α 之前, 描述不久查看上文关于图1-2的论述。转换步骤410可以应用于来自生产过程200的多个批运行210、220的数据 $\{D\}$, 从而分别得到 $\{R\}$ 和 $\{P\}$ 。相似性 S' 可以如解释的那样 (在图2中) 被检测。换而言之, 如果来自两个执行的数据的经转换的时间系列相似, 则质量类别可以假设为也是相同的。

[0142] 例如, 批运行210 (参考) 具有质量类别“成功”并且得到 $\{R\}\#$ (与图5中的 $\{D\}\#$ 一样)。批运行220 (生产) 得到 $\{P\}\#$, 其看上去像平行线。相似性因此将未被检测到。质量类别“成功”将未分配 (给特定生产执行)。

[0143] 在备选方案中, 转换因子向量 610^* 可以从手动评估参考数据 $\{R\}$ 获得。

[0144] 在解释用于获得转换因子向量 610^* 的因子 α 的细节之前, 描述论述了用例。

[0145] 7比较

[0146] 图7图示了由参考批运行和由生产批运行产生的经转换的时间系列的比较, 其中经转换的时间系列是单变量时间系列。

[0147] 在第一情况下, 参考批运行210具有质量类别“成功”, 并且具有特性签名 (参阅图5中的 $\{D\}\#$)。生产批运行220具有相似的特性签名。签名显得“更长”, 因为生产批运行220花更长的时间来完成 (比参考批运行210)。由于相似性 ($S' = \text{YES}$, 参阅图2), 相同的质量类别 (“成功”) 也分配给生产批运行220。本领域内技术人员可以比较这样的签名, 即使它们的长度不同也如此。比较可以包括计算相似性指数或类似物。适当的工具是可用的 (参阅引用 [1])。

[0148] 在第二情况下, 参考批运行210与第一情况相同, 但生产批运行220具有不同的特性签名。图通过不同的曲线图来图示非相似性。非相似性 ($S' = \text{NO}$) 结合二进制类别得到“失败”。

[0149] 图7(和其他图)使用轨迹中的签名来增强理解,但本领域内技术人员不必画出轨迹。

[0150] 在图5的示例中,使用因子 α_v 使得{D}#中的签名的特性保留。为了继续图5的示例,因子 $\alpha_v=0.5$ 使经乘以的时间系列 $\{D_v\}$ ~中的 $\{D_v\}$ 的增加对于经转换系列{D}#中的“可见性”而言保持足够大。换言之,选择因子使得转换是稳健的,从而确定指示特定质量类别的相似性。下面解释如何可以获得因子。

[0151] 专家用户的参与是可选的,并且参与不要求专家来识别因子。

[0152] 8方法

[0153] 图8图示对于计算机实现的方法300和400的流程图。

[0154] 方法300是用于通过机器学习来获得因子的一个方法(即,因子方法),并且方法400是用于确定生产过程200的特定批运行的相似性(或非相似性)的一个方法。

[0155] 两个方法都描述了可以被不同的计算机模块(分别是因子模块603和相似性模块604,参阅图4)执行的步骤序列。

[0156] 以方法400开始是方便的,因为已经解释了步骤中的一些。

[0157] 同样如由虚线矩形图示的那样,方法400可以分化为计算机实现的方法401(是参考方法)和方法402(是生产方法)。

[0158] 为了执行方法400(具有方法401和402),相似性模块604需要转换因子向量 610^* ,参阅图4。如提到的,可以通过ML从因子模块603获得转换因子向量 610^* ,从而执行方法300。在备选方案中,转换因子向量 610^* 可以从手动评估参考数据{R}#而获得。

[0159] 通过方法401(具有步骤序列405和410),相似性模块604的参考子模块对生产过程200的参考批运行210(或对于由参考批运行210产生的产品)提供特定质量类别的参考。参考具有经转换的时间系列{R}#的形式。

[0160] 在方法402(具有步骤序列415、420和430)中,相似性模块604的生产子模块确定对于生产过程200的生产批运行220(或对于由生产批运行220产生的产品)的质量类别。

[0161] 方法400分化为方法401和402对于存储参考{R}#达相对长时间的情形和对于相同参考{R}#充当用于比较不同生产数据的参考的情形是方便的,所述生产数据潜在地来自其他生产现场的生产设备所接收。

[0162] 在步骤接收405中,相似性模块604例如从数据存储库650(参阅图5-6)接收数据作为{R}#。

[0163] 在步骤转换410中,使用转换因子向量 610^* ,相似性模块604将多变量参考时间系列{R}#转换成经转换的参考时间系列{R}#,如用子步骤512和514(参阅图6)所解释的那样。相似性模块604可以将{R}#存储在数据存储库650中,或存储在别处。

[0164] 在可选的步骤读取412中,相似性模块604从数据存储库650或从别处读取(即,接收)经转换的参考时间系列{R}#。要注意,这里引入读取412只是为了方便解释。在相同物理计算机上实现子模块604-i(针对方法401)和604-ii(针对方法402)的情况下,读取只包括访问存储器中的数据。

[0165] 在步骤接收415中,相似性模块604例如从数据存储库650(参阅图4-5)或(没有中间存储)从设备110(也参阅图5)接收多变量生产时间系列{P}#。

[0166] 在步骤转换410中,使用转换因子向量 610^* (即使用相同的因子向量),相似性模块

604将多变量生产时间系列 $\{P\}$ 转换成经转换的生产时间系列 $\{P\}$ #,如利用子步骤512和512(参考图5)所解释的那样。

[0167] 在步骤比较430中,相似性模块604比较 $\{R\}$ #与 $\{P\}$ #(即,比较经转换的参考时间系列和经转换的生产时间系列)。由此,相似性模块可以使用时间规整(即,时间规整操作)。本领域内技术人员可以应用时间规整并且可以针对另外的细节回顾引用[1]。还构思使用其他方法,诸如可选地根据图13的方法。

[0168] 如上文提到的,在转换期间(在步骤410和420中)保留特性是比较步骤430的条件。

[0169] 描述现在给出由因子模块603执行的方法300的概述。鉴于图1,描述改变了首字母缩写。

[0170] 在步骤305中,因子模块403从至少两个参考批运行210-R'和210-Q'接收参考数据,参阅图1。(要注意,Q'与Q不相同)。该步骤可以像步骤405或415(例如,R'代替R,Q'代替P)那样执行。

[0171] 在下面,参考数据将被标记为R'和Q',诸如对于第一多变量时间系列 $\{R'\}$ 和第二多变量时间系列 $\{Q'\}$ 。第一多变量时间系列 $\{R'\}$ 至少包括具有来自第一源120-1(参阅图5)的数据的第一(单变量)时间系列 $\{R'_1\}$,和具有来自第二源120-2的数据的第二(单变量)时间系列 $\{R'_2\}$ 。第二多变量时间系列 $\{Q'\}$ 至少包括具有来自第一源120-1的数据的第一(单变量)时间系列 $\{Q'_1\}$,和具有来自第二源120-2的数据的第二多变量时间系列 $\{Q'_2\}$ 。更一般地, $\{R'\}$ 和 $\{Q'\}$ 两者都包括来自V个源的V个单变量时间系列。

[0172] 在步骤310中,因子模块603通过以下来确定特性

[0173] ●确定311单变量时间系列的特性部分,以及

[0174] ●确定312特性部分之间的关系。

[0175] 在下列图中,(时间系列的)特性部分将图示为特性部分(标记为①②③),并且关系将通过虚线箭头(在特性形状之间)来图示。特性形状的序列是(特定批运行的)签名。有时步骤参考311和312被添加到图。

[0176] 有关的特性部分(步骤311/312)的数字 Ω 小于K。该数字 Ω 对于来自不同源的时间系列而言可以是不同的。在下文的示例中,对于第一示例,数字是 $\Omega=1$,数字是 $\Omega=2$ (对于第二示例,部分①②)并且是 $\Omega=3$ (对于第三示例①②③)。

[0177] 从不同角度看来,在步骤310中,因子模块603识别如何涉及时间系列的目标。因子模块603对不同的候选因子向量执行下列步骤320-350。

[0178] 应用候选因子向量可以通过重复步骤320-350(即,在对于F个向量的环中)、通过并行执行步骤320-350(即通过并行操作子模块)或通过其组合来实现。当在步骤360中评估示出了特性保持而不管多到单变量转换时,应用候选因子向量停止。图8的描述假设重复实现。

[0179] 在步骤320中,因子模块603在具有至少两个因子的候选因子向量 (α_1, α_2) 中选择因子:对于(单变量)时间系列 $\{R'_1\}$ 和 $\{R'_2\}$ 是第一因子 α_1 以及对于(单变量)时间系列 $\{R'_2\}$ 和 $\{Q'_2\}$ 是第二因子 α_2 ,其分别具有来自第一源120-1和第二源120-2的数据。

[0180] 候选因子向量的初始选择可以是随机选择。

[0181] 在步骤330中,因子模块603将第一多变量时间系列 $\{R'\}$ 和第二多变量时间系列 $\{Q'\}$ 转换成经转换的第一时间系列 $\{R'\}$ #和经转换的第二时间系列 $\{Q'\}$ #。由此,因子模

块603相应地应用子步骤512(使用候选向量因子相乘)和514(求和)。

[0182] 在步骤340中,因子模块603使经转换的第一时间系列 $\{R'\}_\#$ 与经转换的第二时间系列 $\{Q'\}_\#$ 对准。因子模块603由此使用DTW(参阅引用[1])。由此,经转换的第二时间系列 $\{Q'\}_\#$ 的数据值与经转换的第一时间系列 $\{R'\}_\#$ 的数据值对准。这可以通过对 $\{Q'\}_\#$ 的数据值中的一些分配新的时间点指数而实现。可以考虑对准包括确定(单变量时间系列的)特性部分和确定关系。更详细地,由于对准步骤340,数据值中的至少一些分配给与批运行(即,210-Q')的时间区间不同的时间区间。

[0183] 在步骤350中,因子模块603测量部分特定对准位移 ΔT 。如本文使用的,位移是 $\{R'_v\}$ (在步骤310/312中确定)的特性部分的原始时间点与已经有关的 $\{Q'_v\}$ (在步骤310/314中确定)的特性部分的对准时间点之间的距离——在时间点指数中测得。

[0184] 对大致所有特性部分和它们的关系(数字 Ω)执行测量 ΔT 。

[0185] 在步骤360中,因子模块603通过对部分特定位移 ΔT 求和($\sum \Delta T_\omega$ ($\omega=1$ 至 Ω))来评估它们。转换因子向量610*是总和最小的向量。

[0186] 可选地,总和被计算作为绝对值的总和 $\sum |\Delta T_\omega|$ 。

[0187] 尽管图8图示因子模块603重复执行步骤320、330、340和350(针对被修改的因子向量),因子模块603也可以并行执行步骤(图8还图示具有多个方框的步骤)。

[0188] 要注意,重复不一定使位移(ΔT)和/或它们的总和减小。因此存储与候选因子向量有关的位移(ΔT)是有利的,至少是只要还未识别转换因子向量。

[0189] 图8还图示一个或多个计算机程序或计算机程序产品。计算机程序产品在被加载到计算机的存储器内并且被计算机的至少一个处理器执行时执行计算机实现的方法的步骤。因此换而言之,图8中的框图示了方法可以在程序的控制下通过模块实现。本领域内技术人员可以对计算机模块选择合适的分配方法,如在示例中,对因子模块603选择方法300、对相似性模块604选择方法400(通过子模块的步骤序列)。

[0190] 9-11用于获得转换因子向量的示例

[0191] 图9-11通过示例图示了用于解释机器学习(ML)来获得源特定因子 α_v 的时间系列的轨迹。对于说明性示例,技术设备110在两个(或更多个)参考批运行210-R'和210-Q'(即批运行,参阅图1),已经执行生产过程200两次。说明更加被简化了。事实上,存在超过两个批运行。

[0192] 从参考批运行210-R'和210-Q'来看,至少两个多变量时间系列是可用的,在示例中是多变量时间系列 $\{\{R'\}\}$ 和 $\{\{Q'\}\}$ 。图从左到右图示了轨迹。

[0193] 要注意,源(参阅图5中的源120)不一定根据它们的类型被分化。换而言之,源提供测量值、控制指令还是状态指示符并不重要。

[0194] 多变量时间系列通过它们的轨迹图示为 $\{\{R'\}\}$ (通常在上文)和 $\{\{Q'\}\}$ (下文)。

[0195] 参考批运行210-R'和210-Q'共享相同的质量类别,诸如例如 Q_R ="成功"。参考批运行210-R'和210-Q'得到相同的数据是不一定的。数据通常不同。例如在持续时间和/或在数据值中存在差异。

[0196] 9第一示例

[0197] 图9图示了第一示例。为了方便起见,线被编号。为了简单起见,对于该说明性示例,假设以下:

[0198] ●参考批运行210-R'和210-Q'得到只可以是1或0的数据值。

[0199] ●在第一、第二和第三连续时间点 $[t_{k-1}, t_k, t_{k+1}]$ 的数据值(0, 1, 0)的序列是“峰值”。

[0200] ●与在线0中一样,对于两个批运行,从 $[t_1, t_k]=[1, 10]$ 存在 $K=10$ 个时间点。

[0201] ●如果来自第一源120-1的数据在两个批运行中示出至少一个峰值,无论什么时候,则批运行相似(并且具有相同的质量)(相似性准则)。

[0202] 如提到的,两个参考批运行210-R'和210-Q'相似,并且方法300提供可以用于(在方法400中)为后续批运行(参阅批运行210(参考)和220(生产))确定相似性(或非相似性)的转换因子向量610*。但再次假设210-R'与210-Q'之间的相似性。

[0203] 线2和4示出双变量时间系列 $\{\{R'\}\}$ 。如在线2中, $\{R'_1\}$ 具有特定数据值作为 $r_{12}=1$,其中其他数据值应是零(由点图示)。换言之,来自第一源120-1的数据具有集中在 t_2 的峰值。如在线4中, $\{R'_2\}$ 具有特定数据值 $r_{28}=1$,并且否则是零。换言之,来自第二源120-2的数据具有集中在 t_8 的峰值。

[0204] 如在线6和8中,第二参考批运行210-Q'在 $\{Q'_1\}$ 中得到峰值 $q_{15}=1$ 并且在 $\{Q'_2\}$ 中得到两个峰值 $q_{23}=1$ 和 $q_{27}=1$ 。

[0205] 如通过下划线图示的,因子模块603确定特性部分。峰值中的一些也是特性部分,但不是全部。在示例中,用于确定特性部分和用于使特性部分相关的预定义的规则根据相似性准则导出。

[0206] 在该简化示例中,因此预定义的规则用于识别 $\{R'_1\}$ 和 $\{Q'_1\}$ 中的每个中的至少一个峰值,并且用于使第一出现的峰值彼此相关、使第二出现的峰值彼此相关等等。图9图示了只具有一个峰值的示例,但这并不改变规则。

[0207] 根据预定义的规则,集中在时间点 $k=2$ 的 $\{R'_1\}$ 中的峰值(来自第一批运行R'的数据)与集中在 $k=5$ 的 $\{Q'_1\}$ 中的峰值有关。来自第二源120-2的数据被忽略,因为预定义的规则与该源无关。

[0208] 特性部分(即,峰值)之间的关系通过虚线箭头来图示。要注意,确定特性部分还对其他部分“解除分类”。在示例中,来自源120-2的数据被“解除分类”。

[0209] 描述忽略了几个词的特性并且解释下一个步骤:如在线10和12中,因子模块603选择候选因子向量(1, 0)(向量610-1,步骤320)。在示例中,选择是随机选择。以候选因子向量(1, 0)开始的原因只是为了方便使该描述保持简短。

[0210] 因子模块603将 $\{\{R'\}\}$ 转换成 $\{R'\}\#$ 并且将 $\{\{Q'\}\}$ 转换成 $\{Q'\}\#$ (步骤330,参阅子步骤512、514)。因子 $\alpha_1=1$ 保持 $\{R'_1\}$ 中和 $\{Q'_1\}$ 中的峰值并且因子 $\alpha_2=0$ 从 $\{R'_2\}$ 和 $\{Q'_2\}$ 中去除峰值。求和导致在时间点 $k=2$ 具有值1的经转换 $\{R'\}\#$ 和在时间点5具有值1的经转换 $\{Q'\}\#$ 。到目前为止,没有什么在时间上移位

[0211] 如在线14中,因子模块603使 $\{R'\}\#$ 和 $\{Q'\}\#$ 已经对准(步骤340)。在示例中,对准保持对于 $\{R'\}\#$ 的区间 $[1, 10]$ 中的时隙并且将 $\{Q'\}\#$ 的至少一些数据值重新分配到其他时间点,在示例中,通过使值1从时间点5移位到时间点1。这得到移位的经转换的时间系列 $\{Q'\}\#\sim$,其已与经转换的时间系列 $\{R'\}\#$ 对准。(~符号指示时间尺度改变了)。

[0212] 该对准没有忽略历史:在图中,移位值1(在 $\{Q'\}\#\sim$ 中)仍然通过下划线来图示。换言之,因子模块603跟踪到该值早些已经被识别了(在步骤310中)。因子模块603还跟踪关

系(参阅箭头)。

[0213] 在步骤350中,因子模块603现在测量关系的部分特定位移 ΔT (或移位),是 R' 与 Q' 之间的时间距离。位移是部分特定的,因为它涉及($\{R'_1\}$ 中的第一特性部分到 $\{Q'_1\}$ 中的第一特性部分的)第一关系。在示例中,只存在一个特性部分和一个关系($\Omega=1$)。

[0214] 在示例中, ΔT 是零,因为所识别的值(在 $\{R'_1\}$ 中在时间点2的值1并且值1来自 $\{R'_2\}$ ~中的原始时间点5)现在在相同的时间点。

[0215] 在步骤360中,因子模块603评估位移。如上文解释的,因子模块603计算位移的总和。在简化 $\Omega=1$ 中,可以跳过求和。在示例中,零位移代表由于简化而造成的理想情形。零位移意味因子向量(1,0)(参考610-1)变成转换因子向量610*(参阅图4)。因子模块603继续提供步骤370(并且提供转换因子向量(1,0)供在方法400中使用)。

[0216] 为了更加精确,峰值(作为(0,1,0)的序列)还要求在(0,1,...)之前和在(...,1,0)之后的数据值的零位移,1处于中心,但这也符合。

[0217] 为了总结示例,峰值的出现仍然在经转换的时间系列中,使得经转换的时间系列可以被因子(1,0)转换。

[0218] 由于因子的初始值并不重要,因此因子模块603在步骤320中可以具有随机选择的候选因子向量(1,1)。转换(步骤330)将导致经转换的时间系列 $\{R'\}$ #和 $\{Q'\}$ #,如在线16和18中的那样。在对准步骤中,因子模块603将使 $\{Q'\}$ #向左移位一个位置。然而,鉴于历史, $k=3$ 处的1与所识别的值(带下划线的值)不相同。部分特定位移 ΔT 将更大。因子模块603将必须重复步骤,参阅图8。

[0219] 10第二示例

[0220] 图10图示第二示例,再次鉴于图8中的方法300的流程图。图10具有3个部分:图10A、图10B和图10C。与(图9的)第一示例相比,第二示例更加复杂(sophisticated),例如在下面:

[0221] ●批运行 $\{R'\}$ 和 $\{Q'\}$ 得到数据值,所述数据值是在0与1之间的正实数。为了简单起见,纵坐标排除在外。

[0222] ●在由圆符号①和②识别的图中存在特性部分。部分①应集中在区间 $[t_{k-\lambda}, t_{k+\lambda}]$ (指数 k 和 λ)中的时间点 t_k 周围。

[0223] ●批运行 $\{R'\}$ 得到来自区间 $[t_1, t_{100}]$ 的数据值,并且批运行 $\{Q'\}$ 在区间 $[t_1, t_{150}]$ 中花费更多时间。

[0224] 如在图10A中,多变量时间系列 $\{R'\}$ 包括单变量时间系列 $\{R'_1\}$ 和 $\{R'_2\}$ 并且多变量时间系列 $\{Q'\}$ 包括单变量时间系列 $\{Q'_1\}$ 和 $\{Q'_2\}$ 。着眼于 $\{R'_1\}$,值(来自源120-1)略微增加直到近似 t_{30} 为止,急剧增加直到 $t_k=t_{40}$ 为止,保持大致恒定直到 t_{65} ($[t_{30-10}, t_{30+10}]$ 中的①)为止、急剧减小直到 t_{70} 为止、保持大致恒定直到 t_{80} 为止、再次下降(集中在 t_{85} 的②)并且保持恒定。着眼于 $\{Q'_1\}$,值示出相似的增加/上升/减小等模式,但在时间上延伸更多:①集中于 t_{65} 并且②集中于 t_{125} 。

[0225] 仍然在图10A中,因子模块603执行步骤310并且由此通过确定311单变量时间系列 $\{R'_1\}$ 、 $\{Q'_1\}$ 的特性部分①和②并且通过确定312特性部分(参阅虚线箭头)之间的关系来确定特性。

[0226] 因子模块603可以根据预定义规则将时间系列划分成部分。通过研究轨迹对时间

系列的性质分类也是公知的(例如,称为“曲线描绘”)。在部分①的示例中,因子模块603已通过确定拐点(即,其中轨迹的二阶导数改变其符号)来确定特性形状和与特性形状有关。

[0227] 可以进一步假设在备选方案中,因子模块603可以通过与专家用户交互已经执行确定。专家用户已可视化地检查了时间系列并且已注释了它们。用于示出轨迹的图形工具和用于从用户获得输入的工具在本领域内可用。如提到的那样,在具有轨迹的图示中,特性部分是特性形状。注释得到对应形状的确定的。在示例中,注释由虚线箭头(其连接轨迹)符号化。用户可以在对应形状之间划线,并且因子模块603可以识别对应的时间点。

[0228] 换而言之,在步骤310中,因子模块603确定是时间系列的特性的部分,但在不同批运行(批运行)中,自动通过应用预定义规则或通过专家用户交互(即,具有受监管训练的ML)。

[0229] 然而,除与专家用户的交互(可选)外,因子模块603不再与用户交互。不是选择合适的因子(α_1, α_2)的人类用户,而是因子模块603。由此,因子模块603执行与相似性模块604后来将执行的步骤大致相似的步骤(参阅方法400)。图10B和图10C给出另外的细节。

[0230] 图10B和图10C图示对F=3个不同的候选因子向量610执行的步骤,其中轨迹在图的左侧、在中心和在右侧。该方法还指示可以并行执行步骤。

[0231] 图10B图示因子模块603执行步骤320(选择因子向量610)和330(转换)。对F=3个因子向量,图示经转换的时间系列 $\{R'\} \#$ (即,与 α 因子(α_1, α_2)相乘,并且求和):

[0232] ● (1, 0) (候选因子向量610-1)

[0233] ● (0.5, 0.5) (候选因子向量610-2) 以及

[0234] ● (0, 1) (候选因子向量610-3)。

[0235] 经转换的时间系列 $\{R'\} \#(1, 0)$ 将看起来像 $\{R'_1\}$ 并且对于 $\{R'\} \#(1, 0)$ 的数据轨迹将是特性签名。经转换的时间系列 $\{R'\} \#(0.5, 0.5)$ 将保持来自 $\{R'_1\}$ 的一些特性,而 $\{R'\} \#(0, 0)$ 将失去它们。

[0236] 要注意,原始时间点的图示部分①和②:在 t_{30} 和在 t_{85} 的形状①,以及在 t_{85} 和 t_{125} 的形状②。然而,经转换的时间系列 $\{R'\} \#(0, 1)$ 失去特性,甚至更糟:①和②和被置于无关于($\alpha_1=0$)原始(在图10A中)的形状处。

[0237] 图10C图示在对于3个因子集的变量中因子模块603执行步骤340和350。

[0238] 对于步骤340,因子模块603使用具有DTW的对准工具(例如,引用[1])。不管先前确定的形状,因子模块603使 $\{R'\} \#$ 和 $\{Q'\} \#$ 对准到公共时间尺度(在对应于 $\{\{R'\}\}$ 的区间 $[t_1, t_{100}]$ 的区间 $[\sim t_1, \sim t_{100}]$ 中,参阅图10A)。经转换的时间系列 $\{Q'\} \#$ 中的数据点与经转换的时间系列 $\{R'\} \#$ 中的数据点对准。要注意, $[t_1, t_{150}]$ 到 $[\sim t_1, \sim t_{100}]$ 的映射不一定是线性的(即,不是2/3因子)。

[0239] 忽视的原因如下:在执行方法400中,相似性模块604未执行这样的确定步骤。

[0240] 因此,原始已经“定位”在 $\{\{Q'\}\}$ 中的特定时间点的数据值(例如, t_{65} 处的①和具有到 t_{125} 的②)现在“定位”在别处。换而言之,原始①和②被重新定位。

[0241] 对于步骤350,知道时间尺度上的原始“位置”的因子模块603单独对每个部分计算重新定位或位移 ΔT 。

[0242] 对于因子向量(1, 0),来自第二源 $\{R'_2\}$ 、 $\{Q'_2\}$ 的数据的影响是零(与零相乘),使得位移 ΔT 对于①和②两者都是零。

[0243] 对于因子向量 $(0.5, 0.5)$, 来自第二源 $\{R'_2\}$ 、 $\{Q'_2\}$ 的数据的影响仍然是限制的, 并且位移是 $\Delta T \textcircled{1}=10$ 和 $\Delta T \textcircled{2}=20$ 。总和 ($\Omega=2$) 可以被计算为 $\Sigma \Delta T=30$ 。

[0244] 对于因子向量 $(0, 1)$, 来自第二源 $\{R'_2\}$ 、 $\{Q'_2\}$ 的数据的影响起主导作用, 并且位移是 $\Delta T \textcircled{1} = 30$ 和 $\Delta T \textcircled{2} = -20$ 。总和 ($\Omega=2$) 可以被计算为 $\Sigma \Delta T=10$ 。(要注意, 绝对值 $\|\cdot\|$ 的可选使用将导致总和是 0、20 和 50)。

[0245] 如在步骤 360 中, 因子模块 603 将候选因子向量 $(1, 0)$ 评估为导致 ΔT 的最小总和 (总和甚至是零)。因此, $(1, 0)$ 可应用于在方法 400 (参阅图 8) 中使用。在步骤 370 中, 因子模块 403 提供转换因子向量 610^* 作为 $(1, 0)$ 。使用 $(1, 0)$ 将是评估避免的过度激进规整的说明性示例。

[0246] 要注意, 在实现中, 因子模块 603 不只是评估两个批。通过多个重复 (或并行执行), 因子模块 603 识别另外的因子向量。在图 10 的示例中, 转换因子向量 $(\alpha_1, \alpha_2) = (1, 0)$ 将是相似性最高 (位移总和最低) 的结果。

[0247] 因此, 将方法 400 (图 8) 中的因子应用于参考批运行 (参阅图 7 的左侧) 和生产批运行 (参阅图 7 的右侧) 将允许确定批运行之间的相似性 (或非相似性)。

[0248] 11 第三示例

[0249] 图 11 图示第三示例。与 (图 10 的) 第一示例相比, 第三示例具有特性部分 $\textcircled{1}$ 、 $\textcircled{2}$ 和 $\textcircled{3}$ 。在步骤 311 中, 因子模块 603 确定来自第一源 (即, 在 $\{R'_1\}$ 和 $\{Q'_1\}$ 中) 的时间系列中的特性部分 $\textcircled{1}$ 、 $\textcircled{2}$, 以及来自第二源 (即, 在 $\{R'_2\}$ 和 $\{Q'_2\}$ 中) 的时间系列中的特性部分 $\textcircled{3}$ 。在步骤 312 中, 因子模块确定关系。来自第一源的时间系列之间存在两个关系, 并且来自第二源的时间系列之间存在一个关系。因子模块 603 执行如上文描述的其他步骤。因为来自两个源的时间系列具有特性, 所以忽略一个源 (诸如图 10 中的 $(1, 0)$ 和 $(0, 1)$) 的因子向量很可能不是转换因子向量 610^* 。

[0250] 对于因子向量 $(1, 0)$, 步骤 340 中的对准将使上面的 $\textcircled{1}$ 与下面的 $\textcircled{1}$ 对准, 并且将使上面的 $\textcircled{2}$ 与下面的 $\textcircled{2}$ 对准 (如果图示的话, 连接线将是垂直的, 如在图 9 中在线 10 与 14 之间那样)。然而, 将忽略特性部分 $\textcircled{3}$ (以及上面和下面的 $\textcircled{3}$ 之间的关系)。

[0251] 对于向量因子 $(1, 1)$, 步骤 340 中的对准将使上面的 $\textcircled{1}$ 与下面的 $\textcircled{1}$ 对准, 并且将使上面的 $\textcircled{3}$ 与下面的 $\textcircled{3}$ 对准, 并且将使上面的 $\textcircled{2}$ 与下面的 $\textcircled{2}$ 对准, 因为经转换的时间系列将保持特性部分 $\textcircled{1}$ 、 $\textcircled{2}$ 、 $\textcircled{3}$ 。 $\textcircled{1}$ 和 $\textcircled{2}$ 中的对准将导致大致零 ΔT , 但 $\{Q'_2\}$ 中的第二上升 (“第二”) 可以影响步骤 340 中的对准。因此, 对于 $\textcircled{2}$ 的位移 ΔT 可以变得更大 (与时间系列 $\{Q'_2\}$ 将没有这样的第二上升的情形相比)。

[0252] 在步骤 360 中, 将选择使 ΔT 的总和最小化的因子向量, 潜在地是因子向量 $(1, 0.8)$ 。因此, $\{Q'_2\}$ 中第二上升的影响被最小化。

[0253] 12 用例场景

[0254] 图 12 图示用于控制执行生产批运行 220 的技术设备 110 的计算机实现的方法 700 的流程图。执行批运行还未完成。方法 700 可以由控制器模块或由任何其他计算机执行。为了简单起见, 未图示控制器模块, 但要注意, 控制器模块可以实现计算机 600 的部分 (参阅图 4)。说明使用上文在图 3 中引入的规约。

[0255] 在上文解释的场景中, 生产批运行 220 的执行可以被视为完成 (参阅 t_N) 并且具有数据的时间系列可以被视为源于已完成的批运行。在上文解释的实施例 (图 1-11) 中, 相似

性S' 被假设为在完成生产批运行220之后例如通过比较430来确定。

[0256] 期望在修改异常批运行仍然可能的早期时间点识别质量指示符(或生产批运行220的质量类别)。要求时间来执行方法400的步骤420/430,但该运行时间可以短于整体批运行时间(即,持续时间 T_p)。确定生产后质量是有用的,但确定生产期间质量潜在地更加相关。

[0257] 在方法700的步骤710中,控制器模块访问具有来自生产过程200(参阅图1-2)的至少一个先前执行的批运行(210,参阅图1-2)的数据的至少一个参考时间系列。图12给出具有两个参考时间系列*R*(1)和*R*(2)的示例。带有()中的指数的标记代表两个先前执行的批运行。时间系列是单变量还是多变量并不重要,因此使用* *标记。在示例中,时间系列是双变量时间系列。控制器模块可以访问来自数据存储库650的时间系列,或以其他方式,参阅图4。

[0258] 参考时间系列中的至少一个具有充当生产批运行的目标的质量类别。在示例中,*R*(1)具有目标质量“成功”。*R*(2)具有非目标质量“失败”。分化成正目标和负目标(例如,成功/失败)充当示例,但还可以使用另外的类别(即,指示符的另外的粒度)。要期望,生产批运行得到目标质量。

[0259] 为了简单起见,步骤710用由对于双变量时间系列的轨迹图示的参考数据而可视化,其具有上面的线和下面的线(如在图1中)。

[0260] (至少一个)参考时间系列与技术设备110的参数610-1/610-2相关联。图12给出参数是双变量时间系列的部分的状态指示符(其由下面的线来符号化)的示例。状态可以是“OFF”(虚线,参数660-1)或“ON”(平线,参数660-2)。

[0261] 在步骤720中,当技术设备110执行生产批运行220时,控制器模块接收具有数据的生产时间系列*P*(这里也图示为双变量时间系列)。

[0262] 在步骤730中,控制器模块识别参考时间系列的子系列。分成子系列是从时间方面划分的。在示例中,利用两个参考时间系列*R*(1)和*R*(2),存在两个子系列*R*(1)_A和*R*(2)_A。指数“A”代表对应于已经过去的生产时间系列*P*的时间区间的阶段。

[0263] 语义(即,质量类别)可选地可以应用于阶段A和B,诸如失败/成功、失败/失败、成功/成功和成功/失败。阶段的类别是目标。

[0264] 如上文解释的,参考时间系列包括区间 $[t_1, t_M]$ 中的数据,参阅图3。为了使用该标记,子系列包括区间 $[t_1, t_{\Pi}]$ 中的数据(大写希腊字母 Π),其中 $\Pi < M$ 。末端指数 Π 对于每个子系列可以不同(因为参考批运行可以具有不同的持续时间)。

[0265] 同样,生产时间系列*P*包括区间 $[t_1, t_N]$ 中的数据。因为技术设备110在执行批运行,因此数据值的数量N在增加,但对于已经接收的生产时间系列*P*,N应是恒定的。对于阶段B的数据因此不再可用。换言之,识别(数据的)子系列在概念上将批运行的执行分成连续阶段A和B。虚垂线指示从阶段到阶段的转变。参考数据*R*可用于阶段A和可用于阶段B,作为参考数据*R*_A和*R*_B,但生产数据只可用于阶段A。

[0266] 为了简化解,可以假设对于图12的示例,数据值的数量相同: $N = \Pi(1) = \Pi(2)$ 。换言之,控制器模块确定生产运行的进程并且相应地选择子系列。

[0267] 要注意,识别可以通过上文已经解释的技术、通过与专家用户交互或通过机器学习(选择与上文解释的因子 α 相似的阶段)来执行。在下文将解释对该方法的修改。

[0268] 参数660-1和660-2分别可用于两个参考批运行。如提到的那样,图通过水平线(部分虚线、部分平线)来图示参数。在示例中,在参考时间系列 $*R*(1)$ 中,参数在 t_{Π} 前不久从ON切换到OFF,在参考时间系列 $*R*(2)$ 中,参数一直保持ON。

[0269] 仍然可以修改生产。对于阶段B的数据 $*P*$ 还不可用但这不是问题。

[0270] 在步骤740中,控制器模块比较所接收的生产时间系列和参考时间系列的子系列。

[0271] 在示例中,存在两个参考时间系列可用于比较。 $*R*(1)_A$ 和 $*R*(2)_A$ 。要注意,不是所有变量都需要被比较。在示例中,比较至少上文的轨迹。

[0272] 比较得到相似性或非相似性的指示。在示例中,子系列 $*R*(1)_A$ 和 $*R*(2)_A$ 两者都与 $*P*$ 相似。然而, $*R*(1)$ 和 $*R*(2)$ (在完整的持续时间A和B内,而不仅在初始A)先前得到不同的质量类别:参考批运行(1)得到“成功”(目标质量)并且参考批运行(2)得到“失败”。换言之,当控制器模块访问 $*P*$ 时,生产批运行在阶段A以“成功”类别运行,但潜在地作为“成功”或作为“失败”而继续。在执行比较740中,控制器模块处理受限(仅阶段A)的数据量,使得计算时间短到足以识别在生产批运行继续时仍然可以应用的控制参数。

[0273] 在相似性情况下,控制器模块使用参数(660-1)作为控制参数,用于在生产批运行(210)的持续期间控制技术设备(110)。在该情况下,参数被认为是得到“成功”的 $*R*(1)$ 的OFF参数660-1。

[0274] 换言之,使用参数660-1(OFF)被证明了促成了参考批运行(1)(就目标而言)的成功完成,并且参数660-2(ON)被证明了促成了失败。

[0275] 可以自动执行控制生产批运行(2)。在该情况下,控制器模块将生产批运行220的参数设定识别(下面的线)为ON(参阅平线),但将这变为OFF。换言之,将指示技术设备将(技术设备110的)特定组件切换为OFF。结果是潜在得到“成功”的批运行。说明被简化:可以使用更多参数。

[0276] 同样,来自正在进行的生产批运行(目前为止已经可用)的数据与参考数据(针对对应的阶段)之间的相似性的识别,结合参考的已知质量类别提供了作为技术系统的技术设备110的状态的指示符。

[0277] 在图12的示例中,状态指示符可以是以下:

[0278] ●技术设备执行与最终失败的参考批运行相似的生产批运行。

[0279] ●通过接管参数(ON到OFF),技术设备执行生产批运行,使得可以修改该生产批运行,从而与最终成功的批运行继续相似。

[0280] 指示符也可以被传递给操作者,其然后可以修改技术设备的操作(即,用于手动关闭组件)。

[0281] 在示例中,对于参考批运行的质量类别是“成功”对“失败”。然而,使用具有更细粒度的类别是可能的。例如,参考批运行在阶段A可以被归类为“有问题”并且在阶段B有条件地“成功”,其中条件是诸如将参数设置为OFF。

[0282] 或者,参考批运行在阶段A可以被归类为“不可校正的失败”。如果步骤740示出生产批运行220处于相同类别,则操作者将必须取消生产批运行220。

[0283] 着眼于步骤730的细节,可以以下面之中的多种方式执行阶段特定比较:

[0284] ●多变量时间系列 $\{\}$ 可以通过引用[1]的技术来比较,诸如通过时间规整(也参阅图2中的 \odot)。

[0285] ●多变量时间系列 $\{\}$ 可以被转换,使得步骤730使用如上文描述的经转换的时间系列 $\{\#\}$ (参阅步骤330/410/430中的转换)。

[0286] ●多变量时间系列 $\{\}$ 可以如在下面用图13描述的那样被预处理,因此比较是特征矩阵比较。

[0287] 为了通过示例总结该图的论述,不仅可以通知操作者关于与参考批的差异(经由警报、通知或以其他方式),而且还可以通知操作者如果将应用与参考相同的设定(或参数),则将生产批运行将如何继续。两种通知都反映了技术系统的技术状态。

[0288] 例如,参考中的通知(通过警报)将潜在地也在生产批运行中出现,除非操作者干预。干预由状态信息的支持,该状态信息潜在地包括参数(在参考批运行中已经应用的参数)。

[0289] 在另外的示例中,操作者可以估计批运行的潜在整体持续时间(即, $[t_1, t_k]$ 或剩余时间 $[t_{\Pi}, t_k]$)。这是可能的,因为参考批运行的持续时间是已知的并且已检测到与参考的相似性。此外,如果预计持续时间相对长,则操作者可以潜在地干预加速处理。因为状态信息可以包括参数(来自参考),所以操作者知道应用什么参数。控制参数的自动应用也是可能的。

[0290] 控制模块可以指示来自参考时间系列的参数作为对操作者的推荐。例如,当前运行的批将具有大约12小时的持续时间,并且具有特定的概率(即,高于预定义阈值)将出现特定警报。持续时间和警报对应于参考。因为参考批运行的操作者已通过特定指令控制设备,所以当前批运行的操作者知道这些特定指令并且可以相应地应用它们。

[0291] 13确定相似性

[0292] 在时间系列(待比较)通过转换已经获得单变量性质的条件下已经描述了比较430(图8),该描述描述了另外的选项。

[0293] 图13图示了时间图(左侧),其示出用于确定相似性指数(或其他指示符)的方法,以用于比较诸如参考时间系列 $\{\{R\}\}$ 之类的多变量时间系列和生产时间系列 $\{\{P\}\}$ 。在右侧,图13还图示正在被导出的特征矩阵。可以使用该方法来代替时间规整(参阅引用[1]和/或图1)。方法使用根据信号处理的方法论(methodology)。

[0294] 控制器模块执行针对生产过程200的多个先前执行的批运行的参考时间系列 $*R*$ 的步骤访问710(参阅图12),并且执行步骤接收720。在步骤比较740中,控制器模块对属于多变量时间系列(所谓的“特征”)的单变量时间系列的连续数据值之间的预定义改变的出現进行计数,并且将计数器存储到矩阵内。控制器模块在阶段A对(参考的)子系列并且对生产时间系列执行这样的预处理。要注意,预处理参考可以执行一次,使得控制器用被预处理到矩阵中的时间系列来执行访问710(当生产在进行中时节省计算时间)。

[0295] 然后控制器模块确定(参考的和生产时间系列的)矩阵之间的相似性度量。

[0296] 将解释用于预处理时间系列 $\{\{D\}\}$ 的方法并且方法将同样应用于 $\{\{R\}\}$ 和应用于 $\{\{P\}\}$ 。单独为 $\{D_v\}$ 识别特征,并且通过特征指数(j)来识别特征。对特征进行计数。一个 $\{D_v\}$ 可以示出在不同计数器数量中出現的不同特征。通过示例对图13的左侧上的时间图来解释特征。

[0297] 对于特征(1)——阈值达到特征, $\{D_v\}$ 的值d达到0.5的峰值(或更高,在具有水平线的图中是绝对值),其是: $d_{v\ k} < 0.5$ 并且 $d_{v\ k+1} \geq 0.5$ 。例如,对于具有 $K=10$ 的(更加简化的)时

间系列,值达到0.5、降到0.5以下、对于几个时间点保持在下面并且再次上升:0.1、0.3、0.4、0.5(出现的特征)、0.4、0.3、0.3、0.6(出现的特征)、0.7、0.9。特征的出现被计数为:计数器(1)=2。

[0298] 对于特征(2)——按增量跳转(jump-by-delta)特征, $\{D_v\}$ 的值d与它先前的值相比(在一个时间步骤中)上升0.5(或更高,相对值改变),即: $d_{v_{k+1}} \geq d_{v_k} + 0.5$ 。假设 $\{D_v\}$ 的值d可以在后面的时间点下降,则特征也可以多次出现,它可以被计数(计数器(2)=3),如在下列示例中的那样:0.1、0.2、0.6(出现的特征)、0.4、0.4、0.1、0.7(出现的特征)、0.1、0.8(出现的特征)、0.7、特征(3)。

[0299] 特征(3)是:按增量跳转特征是变化。这与特征(2)相似,但下降0.5(从k到k+1)。在示例中,计数器(3)=3。

[0300] 特征(4)——阈值跨越特征,可以被分配用于跨越预定义阈值,诸如0.5: ($d_{v_k} < 0.5$ 并且 $d_{v_{k+1}} \geq 0.5$) 或 ($d_{v_k} > 0.5$ 并且 $d_{v_{k+1}} < 0.5$)。存在计数器(4)=1的假设(参阅从 d_{v_7} 到 d_{v_8} 的跨越)。

[0301] 要注意,将数据值 $d_{v_{k+1}}$ 与在先的数据值 d_{v_k} (或与后来的值)比较,其中一个时隙 Δt (或多于一个)的不同在时间尺度上是不可知的。

[0302] 在用于多变量时间系列 $\{D\}$ 的应用中,计数器C(j)被写入矩阵:在示例中,行的数量等于多变量时间系列中的分量V(即源)的数量;并且列的数量等于所考虑的不同特征的数量(这里是4)。定义条目 $C(v, j)$ 来对单变量系列 $\{D_v\}$ 中的特征(j)的出现数量进行计数。

[0303] 表1针对第一双变量时间系列 $\{\{D_1\} \{D_2\}\}_R$ 提供示例矩阵 $C_R(v, j)$, 在示例中,对 $v=1$ 采用计数器。(第一行对应于示例,如刚解释的那样)。下标R指示由参考批运行产生的时间系列结果。

[0304]

| 变量/特征 | (1) | (2) | (3) | (4) |
|-------|-----|-----|-----|-----|
| v=1 | 2 | 2 | 3 | 1 |
| V=2 | 2 | 1 | 1 | 0 |

[0305] (表1)

[0306] 表2针对第二双变量时间系列 $\{\{D_1\} \{D_2\}\}_P$ 提供示例矩阵 $C_P(v, j)$ 。下标P指示由生产批运行产生的时间系列结果。

[0307]

| 变量/特征 | (1) | (2) | (3) | (4) |
|-------|-----|-----|-----|-----|
| v=1 | 2 | 2 | 1 | 1 |
| V=2 | 1 | 1 | 1 | 0 |

[0308] (表2)

[0309] 应用于矩阵空间的度量可以定义相似性度量。例如,度量是矩阵元素上的差的绝对值的总和。在简化示例中,相似性度量被计算为度量 = $\sum_v \sum_j |C_R(v,j) - C_P(v,j)|$

[0310] 其中 $v=1$ 到 V (在示例中 $V=2$), 其中 $j=1$ 到 J (在示例中 $J=4$)。

[0311] 在示例中,度量被计算为 $1+2=3$ 。换言之, $\{\{R\}\}$ 与 $\{\{P\}\}$ 之间的相似性指数已计算为 $S=3$ 。取决于预定义阈值,两个批运行相似(或不相似)。

[0312] 考虑图8的示例签名(阶段A), 矩阵方法可以用于检测相似性。阶段A的签名示出可

能由特定时间系列产生的振动。在这样的时间系列中,具有中值跨越的特征(4)将出现大约6次。

[0313] 为了确定矩阵之间的相似性度量,本领域内技术人员可以应用另外的方法,它们之中是计算曼哈顿距离(Manhattan distance)、余弦相似性以及莱文斯坦距离(Levenshtein distance)(从教科书已知)。要注意,矩阵也可以集群,用于识别相似的(并且导致相同的质量指示符的)参考(和生产)批运行。诸如k最近邻(k-nearest neighbours)之类的集群算法在本领域内是已知的,也可以使用上文提到的相似性度量。

[0314] 一旦检测到时间系列之间的相似性(或非相似性)(以及因此过程批运行之间的相似性),控制器模块就可以向操作者呈现结果。在下面,这通过示例来解释。模块可以针对相似的批显示轨迹。例如在图12中,计算机可以示出参考的阶段B轨迹,参考批运行210-1的阶段B轨迹示出潜在校正(参数660-1 OFF,可选地具有关于设备110的参数的历史数据),并且参考批运行210-2的阶段B轨迹可以示出潜在失败后果。例如在计算机确定生产批运行与具有“否定”质量类别(诸如“失败”)的参考批运行的相似性时,显示预警或警报是可能的。

[0315] 图12-13还图示一个或多个计算机程序或计算机程序产品。计算机程序产品在被加载到计算机的存储器内并且被该计算机的至少一个处理器执行时,执行计算机实现的方法的步骤。

[0316] 14通用计算机

[0317] 图14是示出通用计算机装置900和通用移动计算机装置950的示例的图,其可以与这里描述的技术一起使用。计算装置900旨在表示各种形式的数字计算机,诸如膝上型计算机、台式计算机、工作站、个人数字助理、服务器、刀片式服务器、大型机和其他适当的计算机。通用计算机装置900对应于图1的计算机系统200。计算装置950旨在表示各种形式的移动装置,诸如个人数字助理、蜂窝电话、智能电话和其他类似的计算装置。例如,计算装置950可以包括数据存储组件和/或装置的处理组件,如图1所示的那样。这里示出的组件、它们的连接和关系以及它们的功能仅仅意味着是示例性的,并且不意味着限制在本文档中描述和/或要求保护的发明的实现。

[0318] 计算装置900包括处理器902、存储器904、存储装置906、连接到存储器904和高速扩展端口910的高速接口908以及连接到低速总线914和存储装置906的低速接口912。组件902、904、906、908、910和912中的每一个都使用各种总线互连,并且可以被安装在公共母板上或者根据情况以其他方式安装。处理器902可以处理用于在计算装置900内执行的指令,包括存储在存储器904中或存储装置906上以在外部输入/输出装置(诸如耦合到高速接口908的显示器916)上显示GUI的图形信息的指令。在其他实现中,可以根据情况使用多个处理单元和/或多个总线,连同多个存储器和存储器类型。还有,可以连接多个计算装置900,其中每个装置提供必要操作的部分(例如,作为服务器库、刀片服务器的群组或多处理器系统)。

[0319] 存储器904在计算装置900内存储信息。在一个实现中,存储器904是一个或多个易失性存储单元。在另一个实现中,存储器904是一个或多个非易失性存储单元。存储器904也可以是另一种形式的计算机可读介质,诸如磁盘或光盘。

[0320] 存储装置906能够为计算装置900提供大容量存储。在一个实现中,存储装置906可以是或者包含计算机可读介质,诸如软盘装置、硬盘装置、光盘装置或磁带装置、闪存或其

他类似的固态存储器装置、或者装置阵列,包括存储区域网络或其他配置中的装置。计算机程序产品可以有形地体现在信息载体中。计算机程序产品还可以包含指令,所述指令当被执行时,执行一种或多种方法,诸如上述的那些方法。信息载体是计算机或机器可读介质,诸如存储器904、存储装置906或处理器902上的存储器。

[0321] 高速控制器908管理用于计算装置900的带宽密集型操作,而低速控制器912管理较低带宽密集型操作。这样的功能分配仅仅是示例性的。在一个实现中,高速控制器908耦合到存储器904、显示器916(例如,通过图形处理器或加速器)以及耦合到高速扩展端口910,其可以接受各种扩展卡(未示出)。在实现中,低速控制器912耦合到存储装置906和低速扩展端口914。可以包括各种通信端口(例如,USB、蓝牙、以太网、无线以太网)的低速扩展端口,可以例如通过网络适配器耦合到一个或多个输入/输出装置,诸如键盘、定点装置、扫描仪或联网装置,诸如交换机或路由器。

[0322] 计算装置900可以以多种不同的形式实现,如图所示的那样。例如,它可以被实现为标准服务器920,或者在这样的服务器的群组中多次实现。它也可以被实现为机架(rack)服务器系统924的部分。此外,它可以在诸如膝上型计算机922之类的个人计算机中实现。备选地,来自计算装置900的组件可以与移动装置(诸如装置950)中的其他组件(未示出)组合。这样的装置中每一个可以包含计算装置900、950中的一个或多个,并且整个系统可以由彼此通信的多个计算装置900、950组成。

[0323] 计算装置950包括处理器952、存储器964、诸如显示器954的输入/输出装置、通信接口966和收发器968,以及其他组件。装置950还可以被提供有存储装置,诸如微驱动器或其他装置,以提供附加存储设备。组件950、952、964、954、966和968中的每一个都使用各种总线来互连,并且组件中的若干组件可以被安装在公共母板上或者根据情况以其他方式安装。

[0324] 处理器952可以执行计算装置950内的指令,包括存储在存储器964中的指令。处理器可以被实现为芯片的芯片集,其包括单独的和多个模拟和数字处理单元。处理器可以提供例如装置950的其他组件的协调,诸如用户界面的控制、通过装置950的应用运行以及通过装置950的无线通信。

[0325] 处理器952可以通过耦合到显示器954的控制接口958和显示接口956与用户通信。显示器954可以是例如TFT LCD(薄膜晶体管液晶显示器)或OLED(有机发光二极管)显示器,或者其他适当的显示技术。显示接口956可以包括用于驱动显示器954从而向用户呈现图形和其他信息的适当电路。控制接口958可以接收来自用户的命令,并对它们进行转换以用于提交给处理器952。此外,可以提供与处理器952通信的外部接口962,以便能够实现装置950与其他装置的近域通信。外部接口962可以例如在一些实现中提供有线通信,或者在其他实现中提供无线通信,并且也可以使用多个接口。

[0326] 存储器964在计算装置950内存储信息。存储器964可以被实现为一个或多个计算机可读介质、一个或多个易失性存储单元或一个或多个非易失性存储单元中的一个或多个。还可以提供扩展存储器984,并通过扩展接口982连接到装置950,所述扩展接口982可以包括例如SIMM(单列直插存储器模块)卡接口。这种扩展存储器984可以为装置950提供额外的存储空间,或者也可以为装置950存储应用或其他信息。特别地,扩展存储器984可以包括用于执行或补充上述过程的指令,并且还可以包括安全信息。因此,例如,扩展存储器984可

以充当装置950的安全性模块,并且可以用准许装置950安全使用的指令进行编程。此外,可以经由SIMM卡提供安全应用连同附加信息,诸如以不可破解的方式将标识信息置于SIMM卡。

[0327] 存储器可以包括例如闪存和/或NVRAM存储器,如下所述的那样。在一个实现中,计算机程序产品有形地体现在信息载体中。计算机程序产品包含指令,所述指令当被执行时,执行一种或多种方法,诸如上述的那样方法。信息载体是计算机可读介质或机器可读介质,诸如存储器964、扩展存储器984或处理器952上的存储器,其可以例如通过收发器968或外部接口962接收。

[0328] 装置950可以通过通信接口966无线通信,所述通信接口966在必要时可以包括数字信号处理电路。通信接口966可以提供在各种模式或协议下的通信,诸如GSM语音呼叫、SMS、EMS或MMS消息传递、CDMA、TDMA、PDC、WCDMA、CDMA2000或GPRS等。这种通信可以例如通过射频收发器968发生。此外,可以诸如使用蓝牙、WiFi或其他这种收发器(未示出)发生短程通信。此外,GPS(全球定位系统)接收器模块980可以向装置950提供附加的导航和位置相关的无线数据,所述无线数据可以由在装置950上运行的应用根据情况使用。

[0329] 装置950还可以使用音频编解码器960可听地通信,所述音频编解码器960可以从用户接收口头信息,并将其转换成可用的数字信息。音频编解码器960同样地可以为用户生成可听声音,诸如通过扬声器,例如在装置950的手持机(handset)中。这种声音可以包括来自语音电话呼叫的声音,可以包括录制的声音(例如,语音消息、音乐文件等),并且还可以包括由在装置950上操作的应用生成的声音。

[0330] 计算装置950可以以多种不同的形式实现,如图所示的那样。例如,它可以被实现为蜂窝电话980。它也可以被实现为智能电话982、个人数字助理或其他类似移动装置的部分。

[0331] 这里描述的系统和技术各种实现可以在数字电子电路、集成电路、专门设计的ASIC(专用集成电路)、计算机硬件、固件、软件和/或其组合中实现。这些各个实现可以包括在一个或多个计算机程序中的实现,所述一个或多个计算机程序在可编程系统上可执行和/或可解译,所述可编程系统包括至少一个可编程处理器,所述可编程处理器可以是专用的或通用的,被耦合以从存储系统、至少一个输入装置和至少一个输出装置接收数据和指令,并将数据和指令传送到存储系统、至少一个输入装置和至少一个输出装置。

[0332] 这些计算机程序(也称为程序、软件、软件应用或代码)包括用于可编程处理器的机器指令,并且可以采用高级面向过程和/或面向对象的编程语言和/或采用汇编/机器语言来实现。如本文所使用的,术语“机器可读介质”和“计算机可读介质”是指用于向可编程处理器提供机器指令和/或数据的任何计算机程序产品、设备和/或装置(例如,磁盘、光盘、存储器、可编程逻辑器件(PLD)),其包括接收机器指令作为机器可读信号的机器可读介质。术语“机器可读信号”是指用于向可编程处理器提供机器指令和/或数据的任何信号。

[0333] 为了提供与用户的交互,这里描述的系统和技术可以在计算机上实现,该计算机具有用于向用户显示信息的显示装置(例如,CRT(阴极射线管)或LCD(液晶显示器)监视器)以及用户可以通过其向计算机提供输入的键盘和定点装置(例如,鼠标或轨迹球)。也可以使用其他种类的装置来提供与用户的交互;例如,提供给用户的反馈可以是任何形式的传感反馈(例如,视觉反馈、听觉反馈或触觉反馈);并且来自用户的输入可以包括声音、语

音或触觉输入的任何形式来接收。

[0334] 这里描述的系统和技术可以在计算装置中实现,该计算装置包括后端组件(例如,作为数据服务器),或者该计算装置包括中间件组件(例如,应用服务器),或者该计算装置包括前端组件(例如,具有图形用户界面或网络浏览器的客户端计算机,用户可以通过该图形用户界面或网络浏览器与这里描述的系统和技术的实现交互),或者这种后端、中间件或前端组件的任何组合。系统的组件可以通过数字数据通信(例如,通信网络)的任何形式或介质来互连。通信网络的示例包括局域网(“LAN”)、广域网(“WAN”)以及因特网。

[0335] 计算系统可包括客户端和服务端。客户端和服务端一般彼此远离,并且通常通过通信网络交互。客户端和服务端的关系凭借运行在相应计算机上并且彼此具有客户端-服务端关系的计算机程序而出现。

[0336] 已经描述了多个实施例。不管怎样,将理解,在不脱离本发明的精神和范围的情况下,可以进行各种修改。

[0337] 此外,图中描绘的逻辑流程不要求所示的特定次序或顺序次序来获得期望的结果。此外,可以从所描述的流程提供其他步骤,或者可以消除这些步骤,并且可以向所描述的系统添加其他组件,或者从所描述的系统移除其他组件。因此,其他实施例在下面权利要求书的范围内。

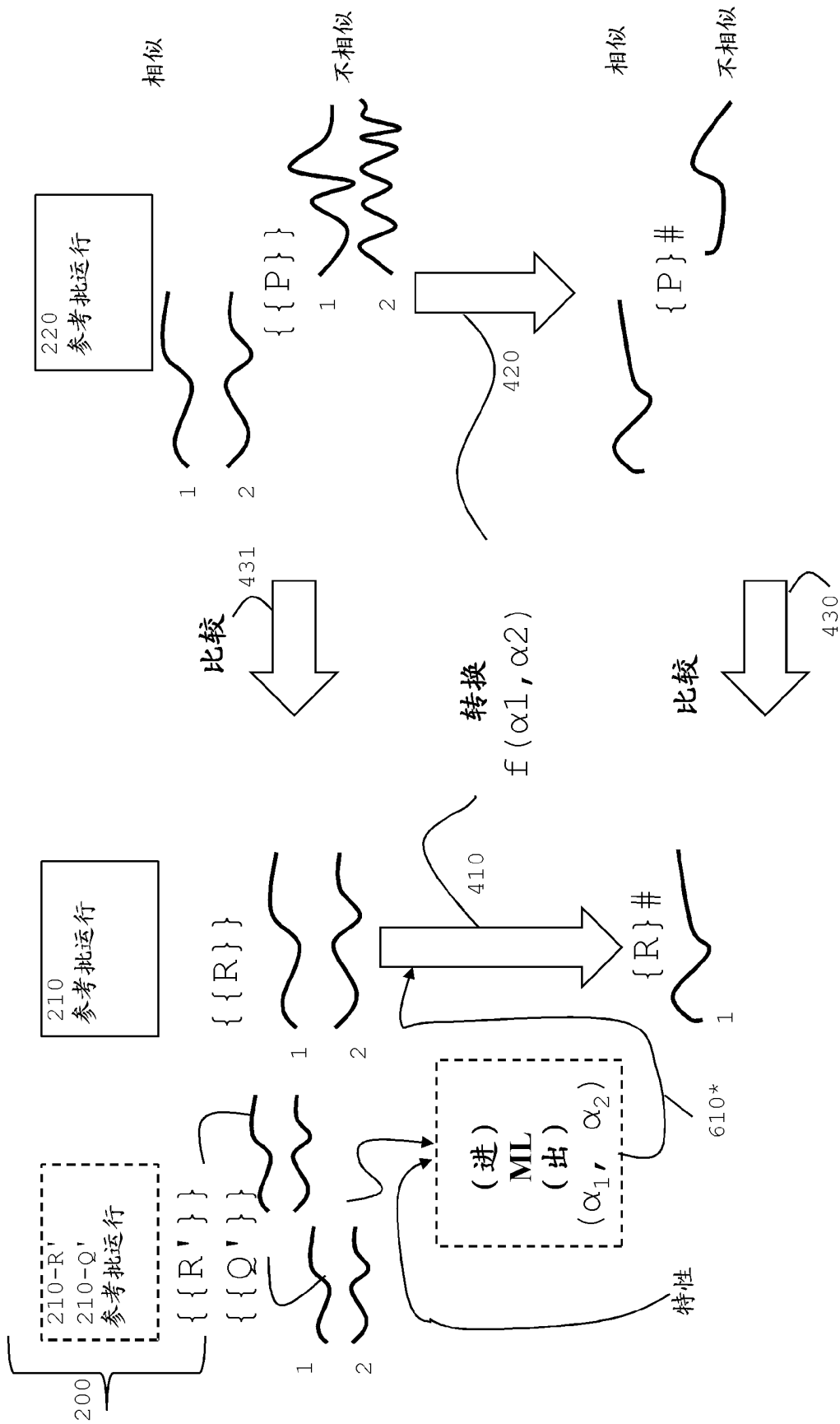


图 1

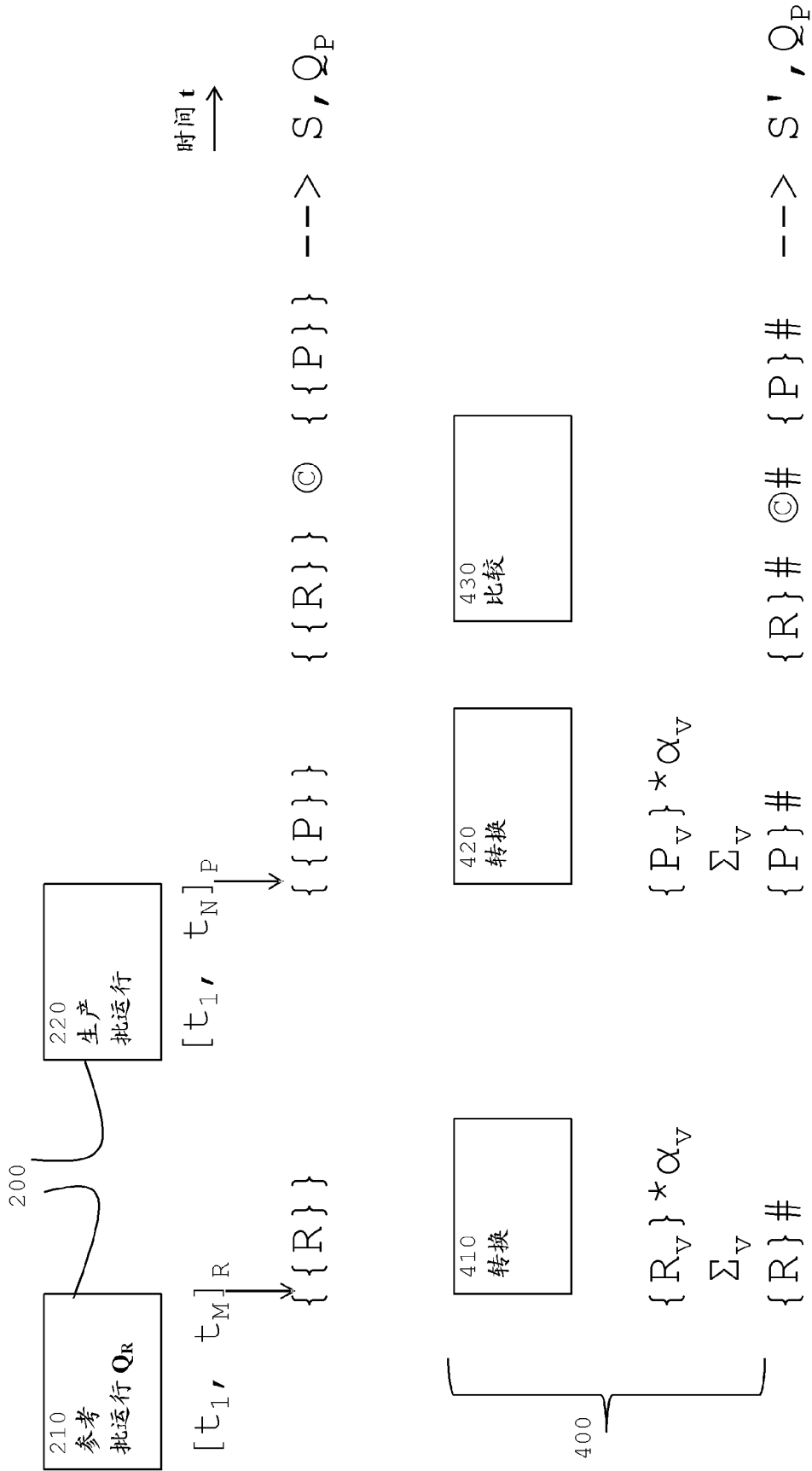


图 2

| 一般 | 参考 | 生产 |
|---------------------|---------------------|---------------------|
| 数据 | 参考数据 | 生产数据 |
| D | R, Q', R' | P |
| d_{kv} | r_{kv} | P_{kv} |
| $\{D_v\}$ | $\{R_v\}$ | $\{P_v\}$ |
| 时间 | | |
| $1 \dots k \dots K$ | $1 \dots m \dots M$ | $1 \dots n \dots N$ |
| $[t_1, t_K]$ | $[t_1, t_M]$ | $[t_1, t_N]$ |
| 源/变量 | | |
| $1 \dots v \dots V$ | $1 \dots v \dots V$ | $1 \dots v \dots V$ |

图 3

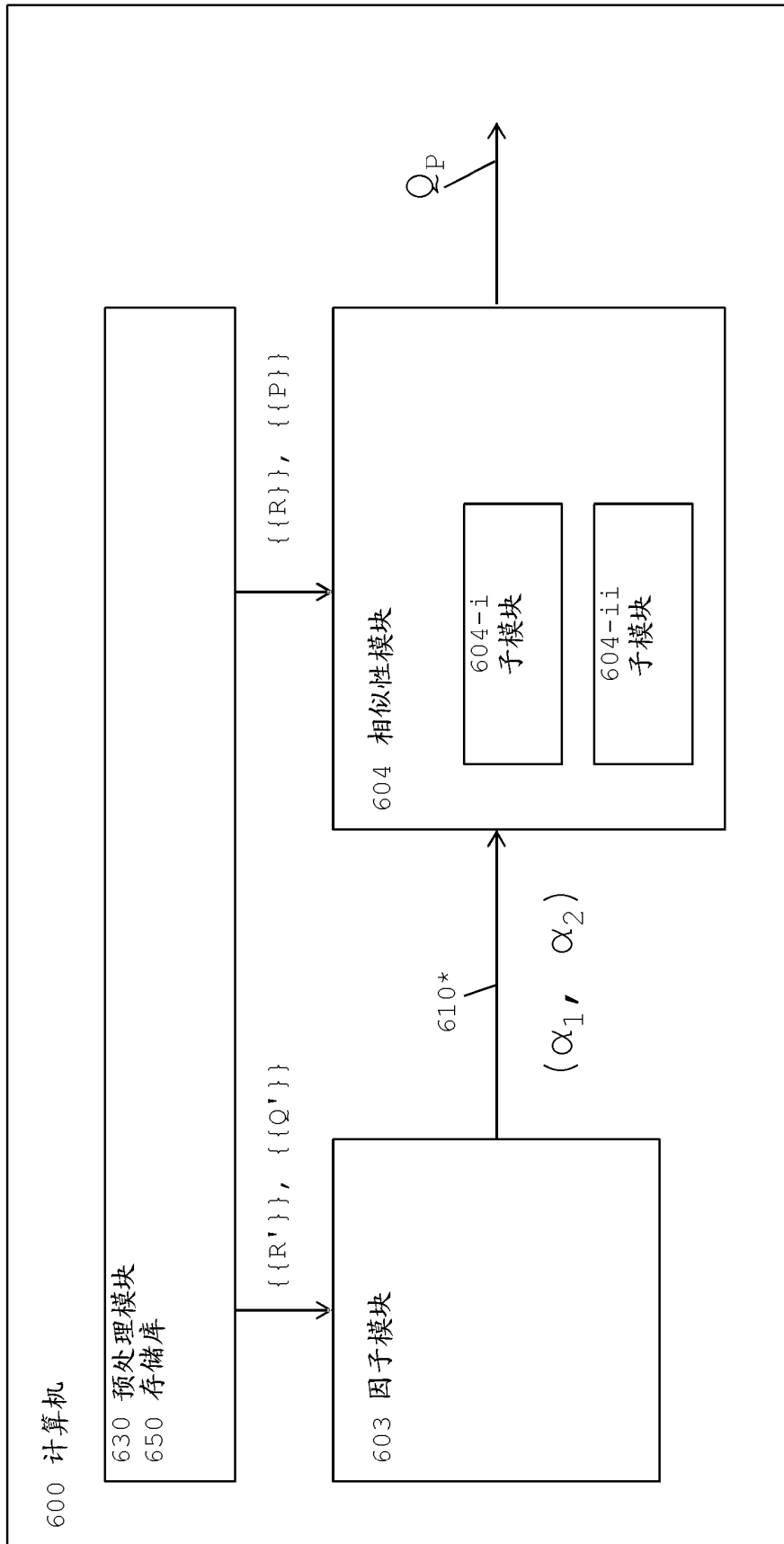


图 4

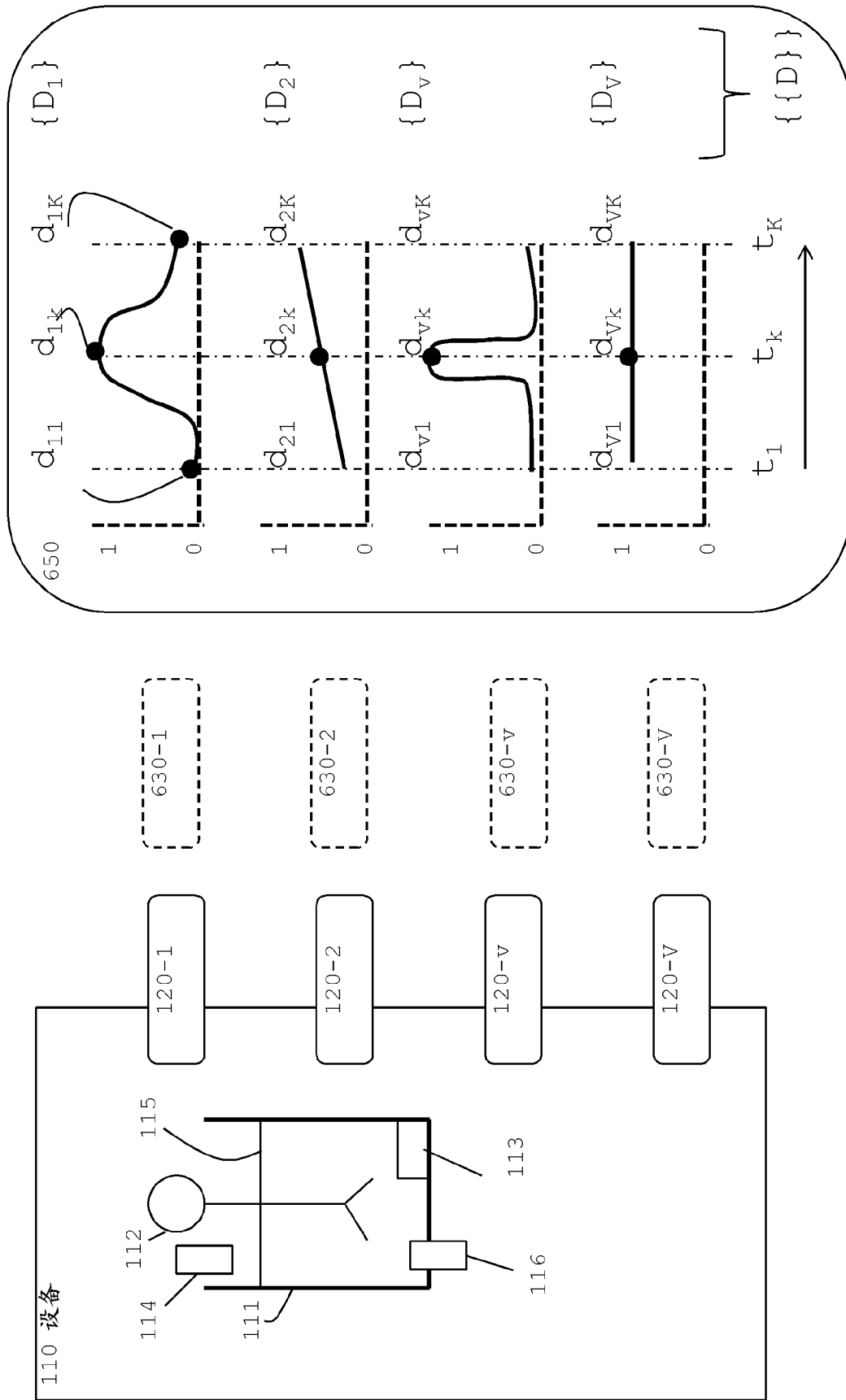


图 5

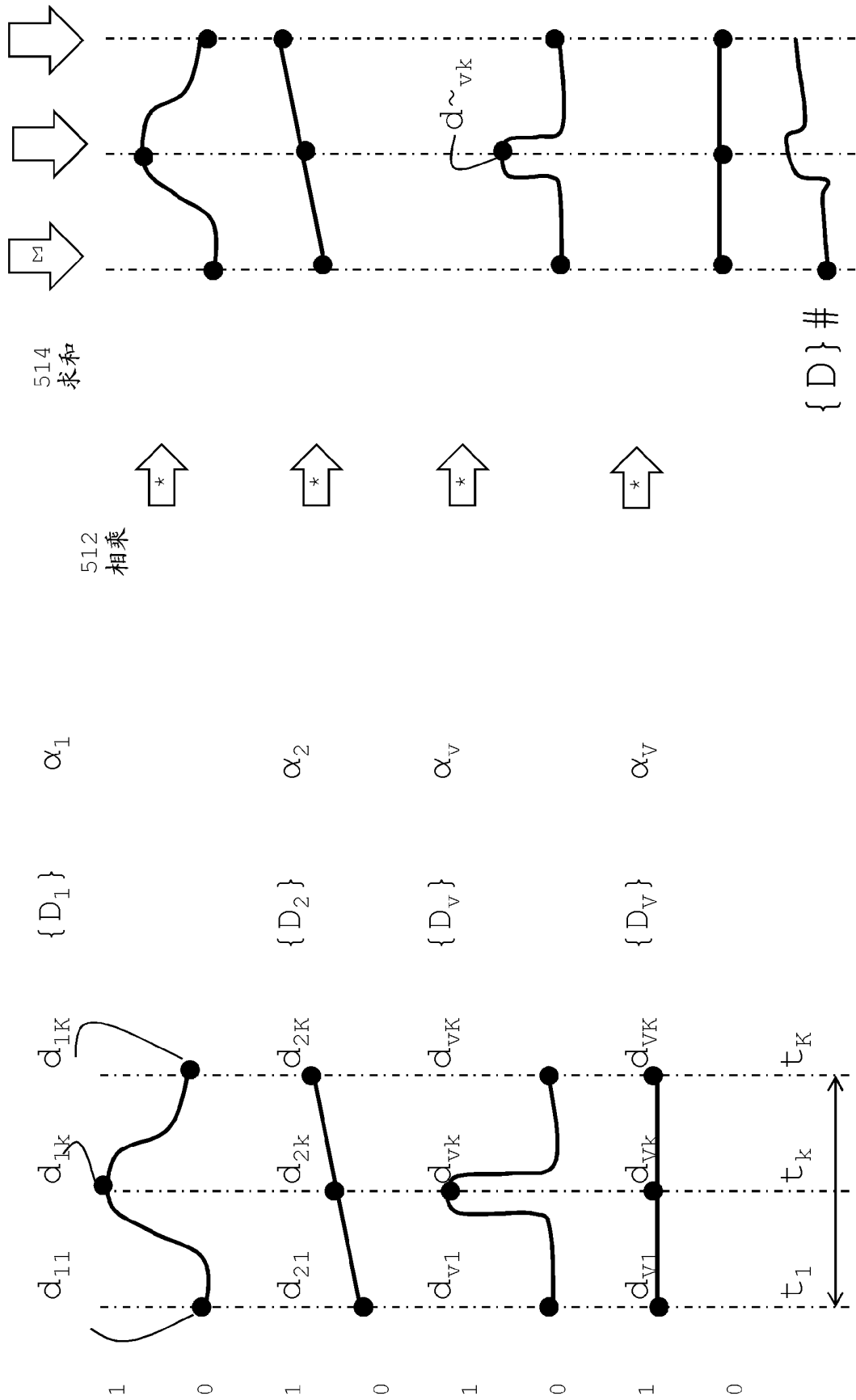


图 6

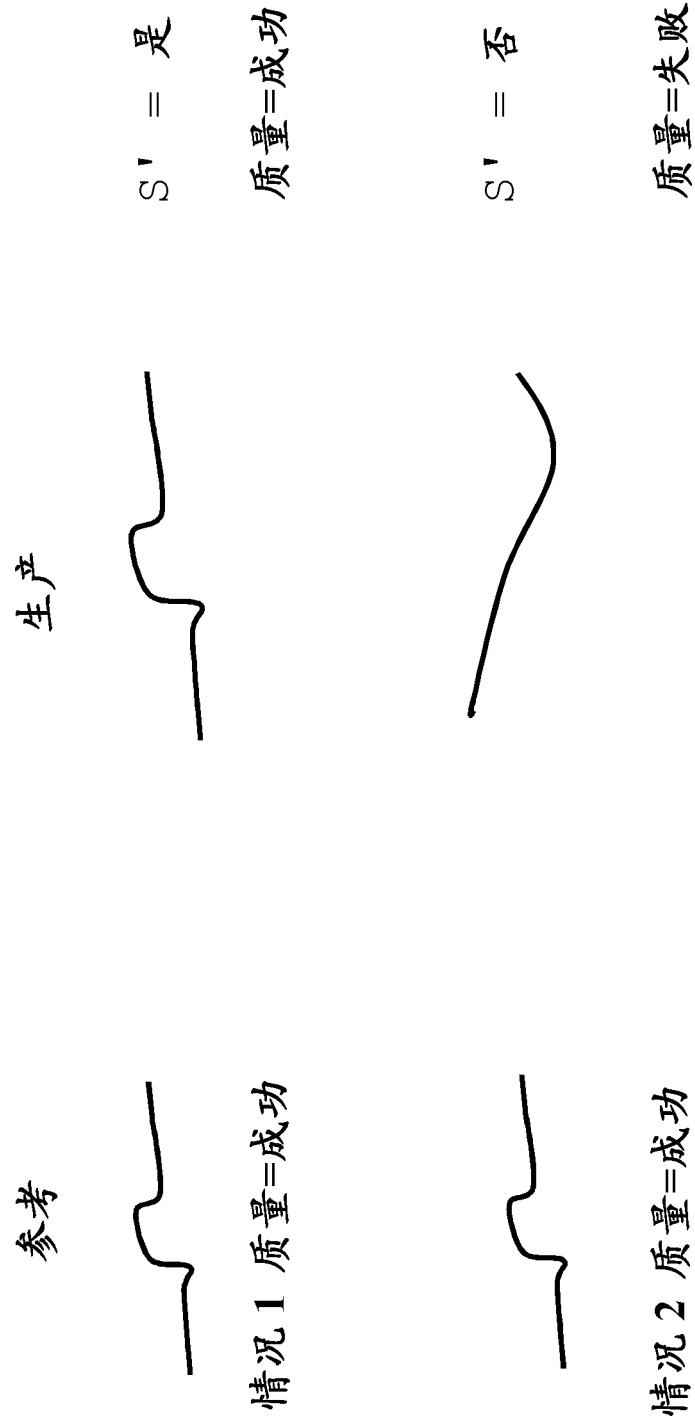


图 7

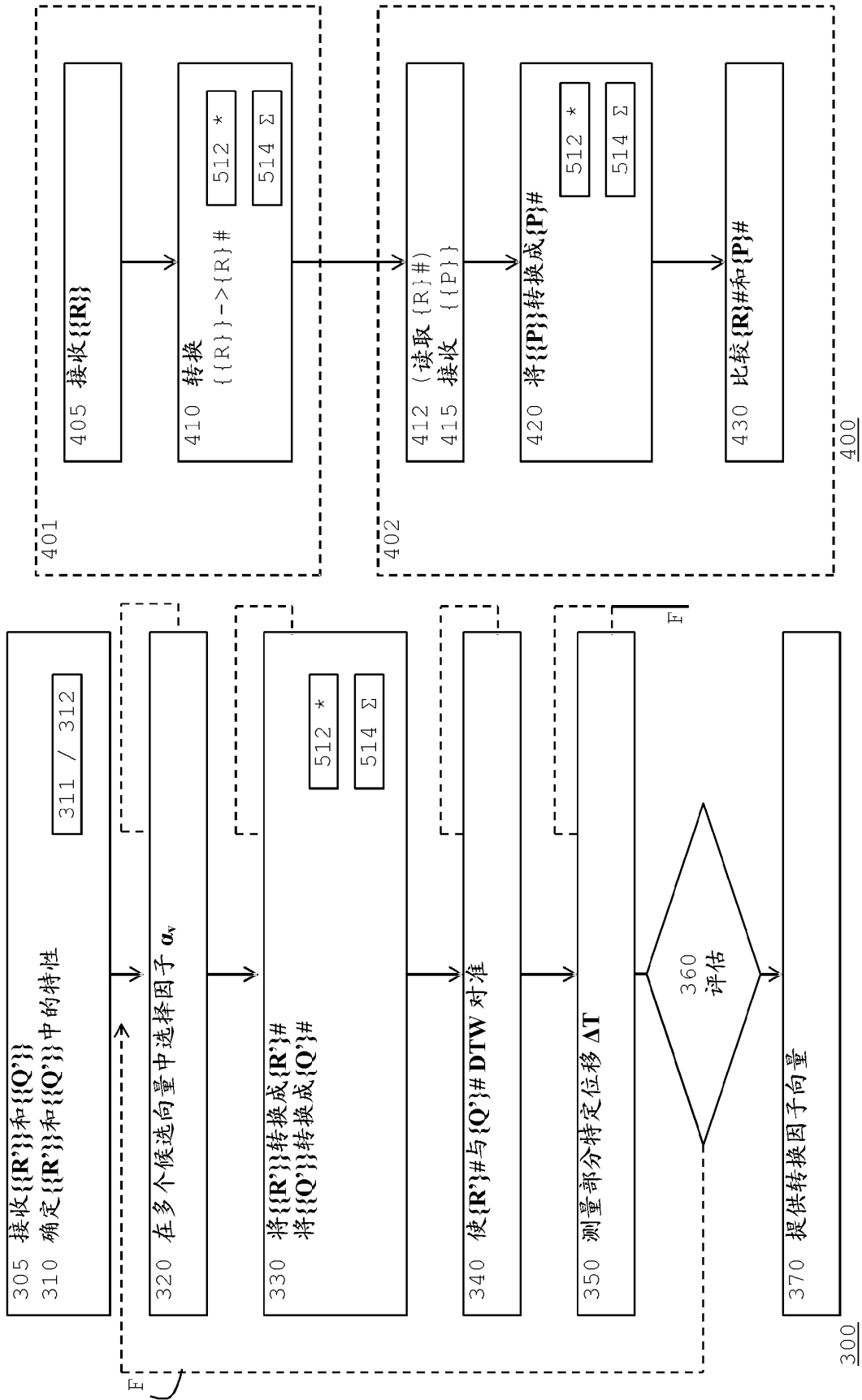


图 8

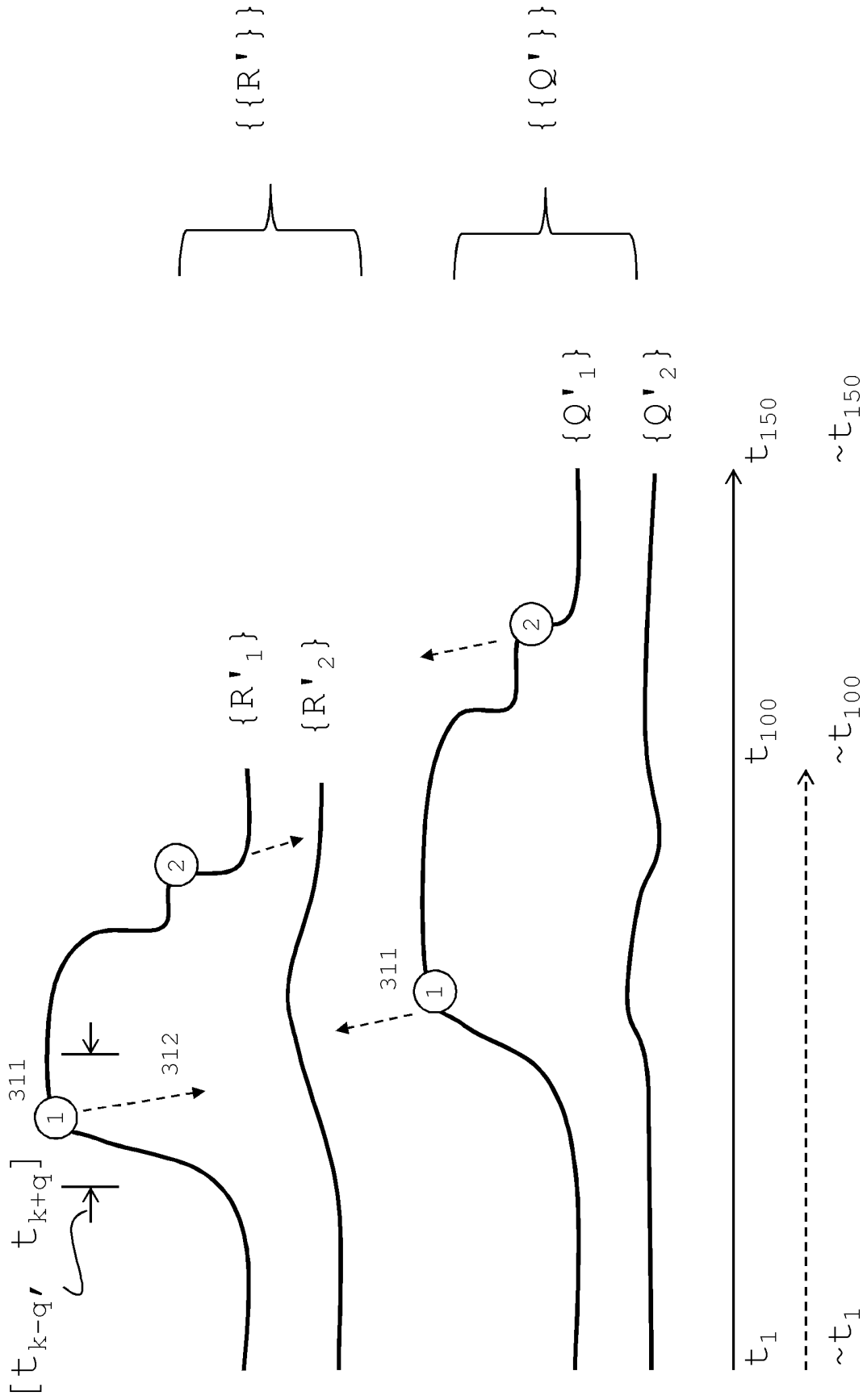


图 10A

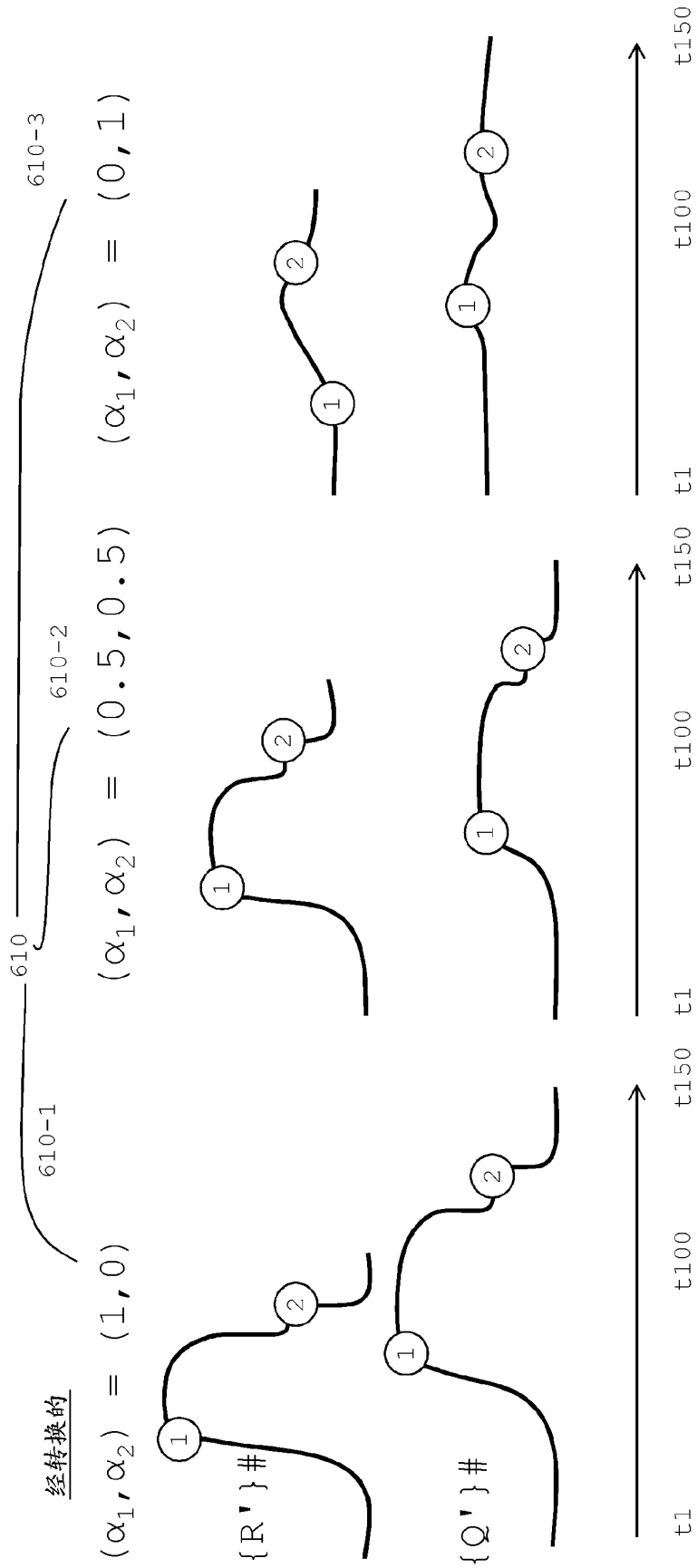


图 10B

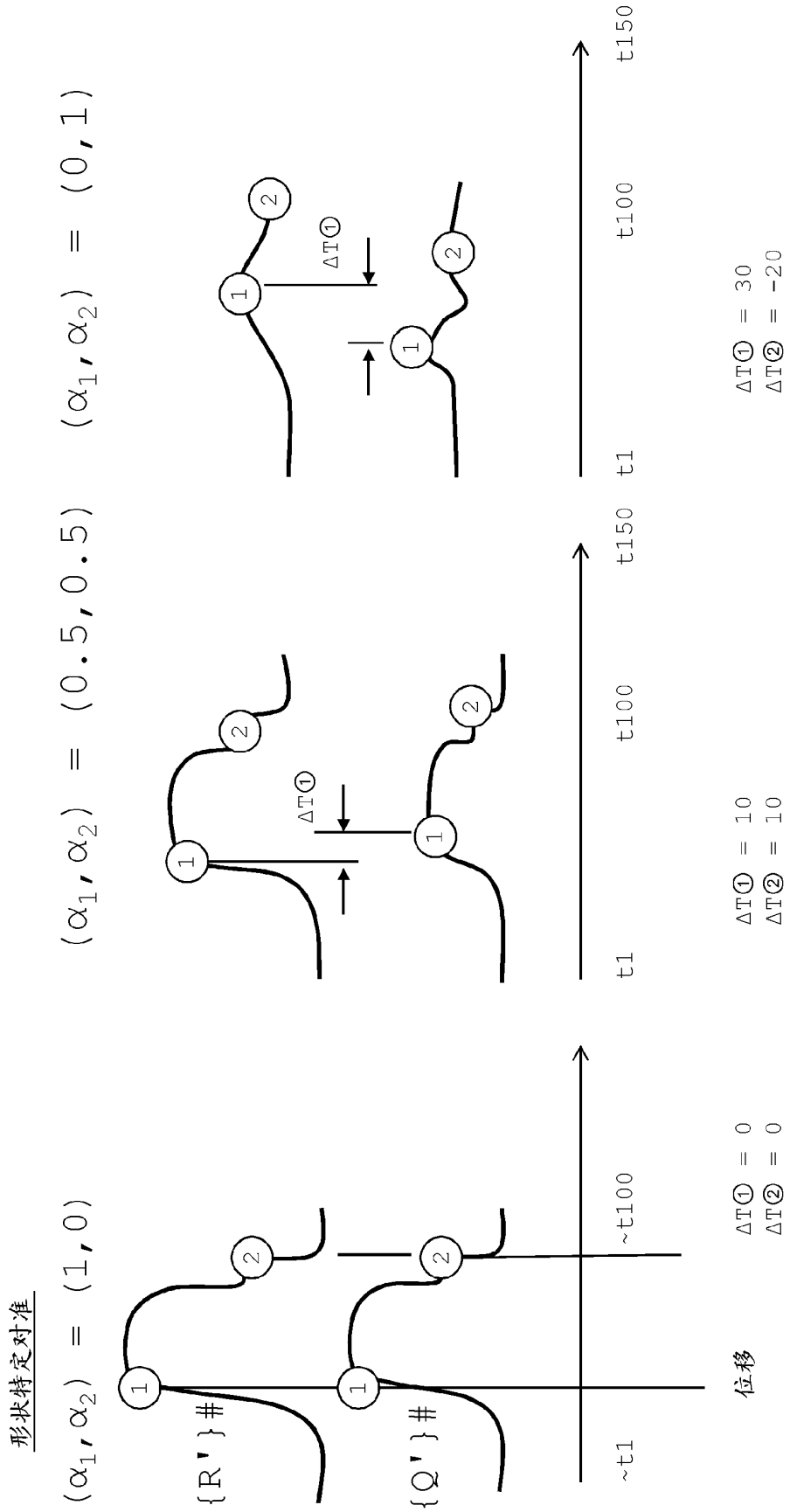


图 10C

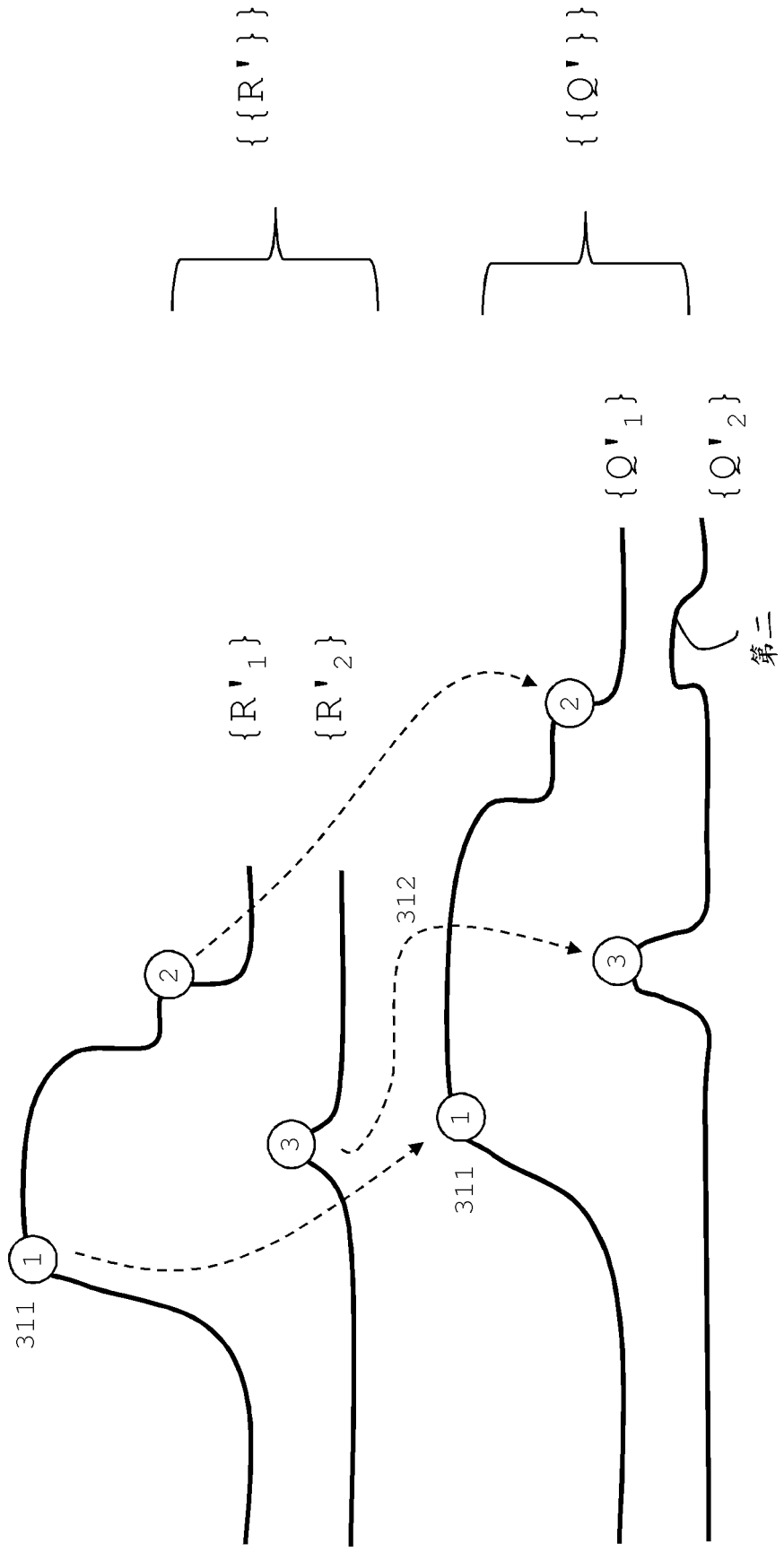


图 11

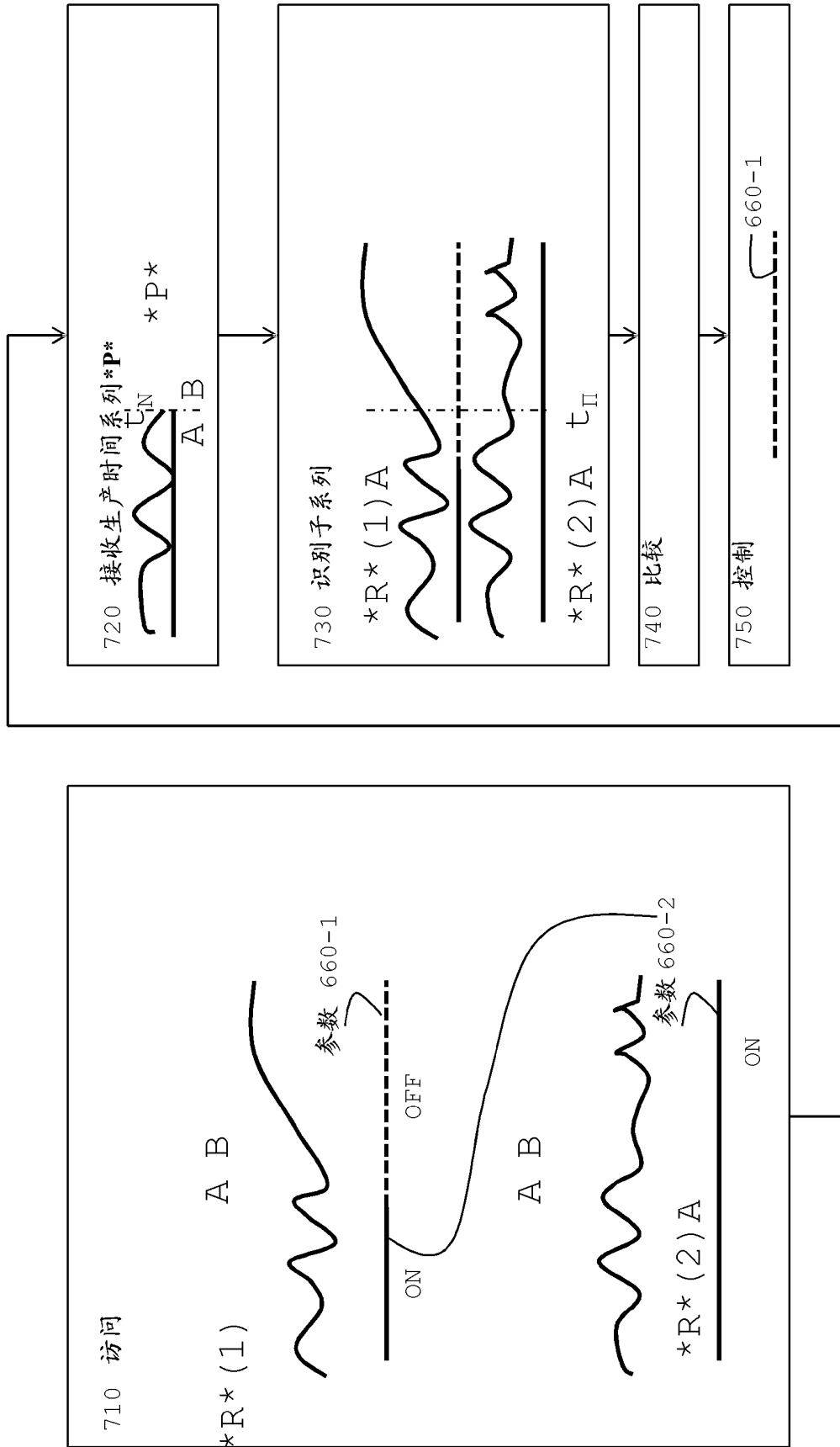


图 12

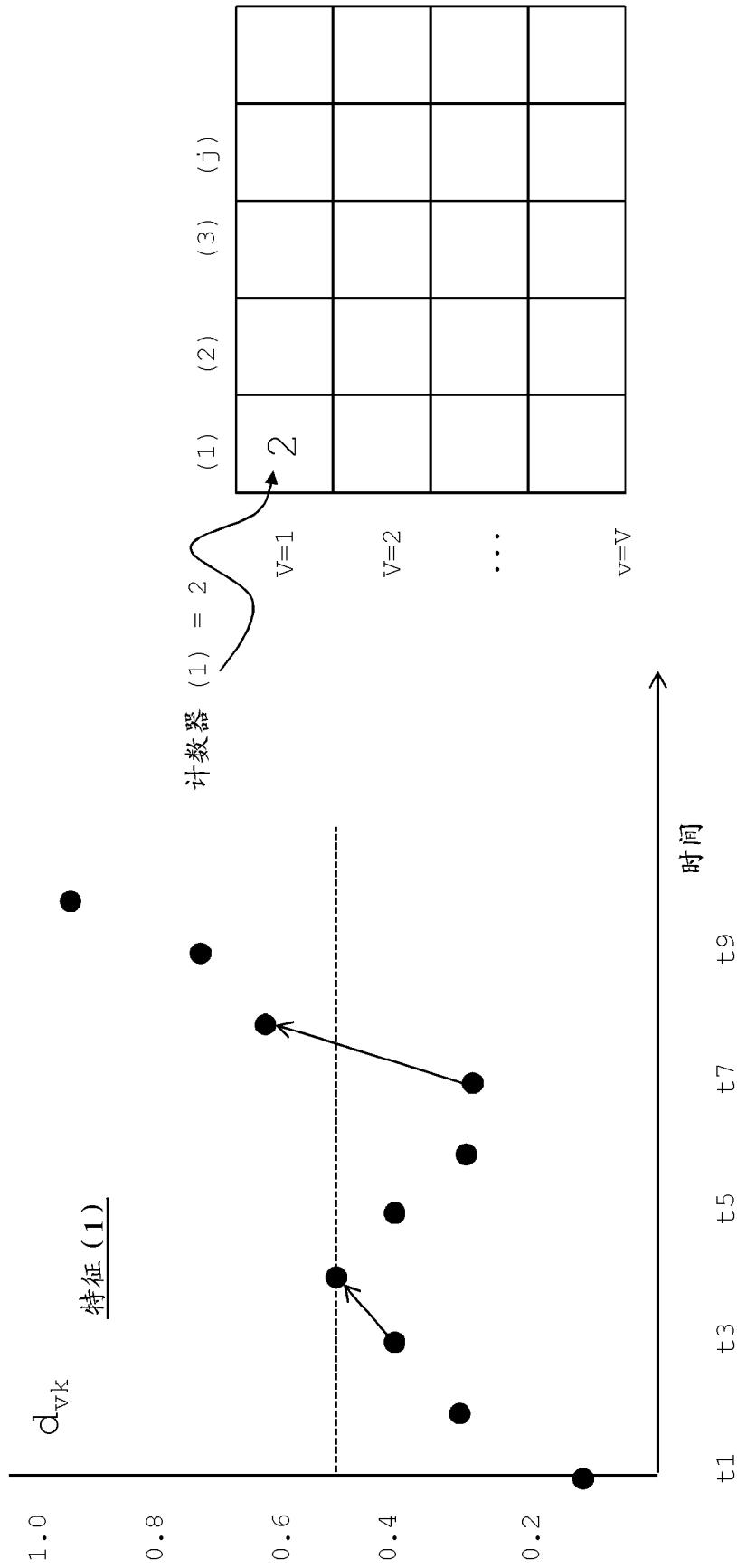


图 13

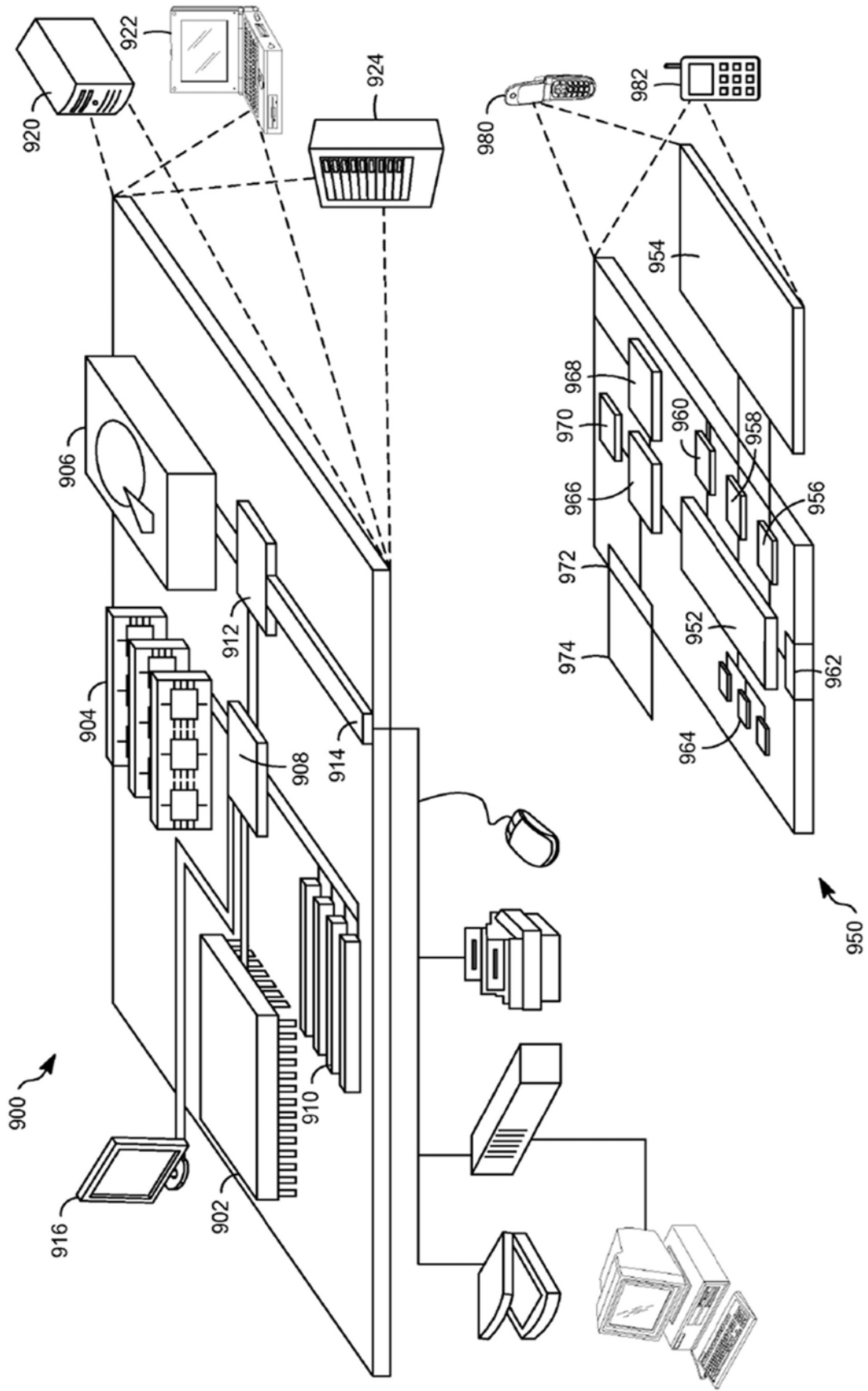


图 14