

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
H03M 7/30 (2006.01)



[12] 发明专利申请公布说明书

[21] 申请号 200580040469.7

[43] 公开日 2007年11月7日

[11] 公开号 CN 101069353A

[22] 申请日 2005.9.27

[21] 申请号 200580040469.7

[30] 优先权

[32] 2004.9.28 [33] US [31] 10/951,566

[86] 国际申请 PCT/US2005/034694 2005.9.27

[87] 国际公布 WO2006/036988 英 2006.4.6

[85] 进入国家阶段日期 2007.5.25

[71] 申请人 通用电气公司

地址 美国纽约州

[72] 发明人 G·E·威廉斯

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 王岳 王小衡

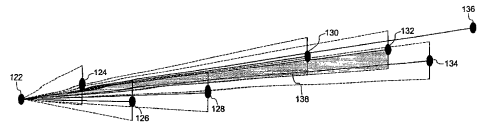
权利要求书6页 说明书12页 附图8页

[54] 发明名称

工业数据压缩系统和方法

[57] 摘要

提供一种用于压缩数据点集合的系统。该系统包括临界孔径压缩模块，其被配置成从数据点集合丢弃一个或多个数据点。丢弃的一个或多个数据点落在表示丢弃的一个或多个数据点趋势的直线周围的容差带内。剩余的数据点集合表示压缩的数据点集合。



1. 一种系统，包括：

临界孔径压缩模块，被配置成从数据点集合丢弃一个或多个数据点，其中丢弃的一个或多个数据点的容差带落在表示丢弃的一个或多个数据点趋势的直线周围，并且其中剩余的数据点集合表示压缩的数据点集合。

2. 权利要求1的系统，其中临界孔径压缩模块包括处理器组件。

3. 权利要求1的系统，其中临界孔径压缩模块包括临界孔径计算模块或临界孔径比较模块，或其组合。

4. 权利要求1的系统，其中容差带包括上线和下线，这两条线分别源自数据点集合中第一点附近的偏差上限和下限并延伸到数据点集合中第二点附近的偏差上限和下限。

5. 权利要求1的系统，还包括一个或多个传感器，被配置成收集数据点集合。

6. 权利要求5的系统，其中一个或多个传感器适于测量温度、压力、时间、流速、速率、电压、电流、振动、速度、加速度或它们的组合，或者基于传感器测量的计算。

7. 权利要求5的系统，其中一个或多个传感器监控工业系统中至少一个可测量的过程控制参数。

8. 权利要求5的系统，还包括数据获取模块，被配置成从一个或多个传感器接收数据点集合并将数据点集合提供给临界孔径压缩模块。

9. 权利要求1的系统，还包括数据存储模块，被配置成接收和存储来自临界孔径压缩模块的数据点集合。

10. 权利要求1的系统，还包括用户接口模块，被配置成将来自系统操作员的一个或多个输入提供给临界孔径压缩模块。

11. 权利要求1的系统，其中直线表示落在容差带内的一个或多个丢弃的数据点。

12. 权利要求1的系统，其中基本上实时地将数据点集合提供给临界孔径压缩模块。

13. 权利要求1的系统，其中临界孔径压缩模块被配置成通过在容差带内线性内插来重构丢弃的一个或多个数据点。

14. 一种用于数据压缩的方法，包括：

基于保存的数据点的容差带以及归档数据点和保存的数据点之间的中间点，来提供具有上斜率和下斜率的临界孔径，这两个斜率源自归档数据点并延伸到保存的数据点附近的偏差上限和下限，并且

如果源自归档数据点并延伸到后续新的数据点的直线斜率位于临界孔径内，则丢弃保存的数据点。

15. 权利要求14的方法，还包括将新的数据点设置成新的保存数据点，并基于新的保存数据点计算新的临界孔径。

16. 权利要求14的方法，其中提供临界孔径包括为归档数据点和后续新的数据点之间的数据点识别所有孔径的交集，每个孔径对于每个相应的点具有源自归档数据点并延伸到容差带的上限和下限的上斜率和下斜率。

17. 权利要求14的方法，还包括如果斜率不在临界孔径内，则归档保存的数据点，其中保存的数据点变成新的归档数据点。

18. 权利要求14的方法，还包括将虚拟数据点设置为新的保存数据点并基于新的保存数据点计算新的临界孔径，其中虚拟数据点不同于实际数据点并且其中虚拟数据点位于先前计算的临界孔径内和实际数据点的容差带内。

19. 权利要求14的方法，还包括基本上实时地获得一个或多个数据点，数据点包括保存的数据点和后续的新数据点。

20. 权利要求14的方法，还包括经由用户接口模块至少提供容差带。

21. 权利要求14的方法，还包括通过在容差带内线性内插来重构丢弃的保存数据点。

22. 一种有形媒体，包括：

数据过滤代码，包括：

配置成基于保存的数据点的容差带以及归档数据点和保存的数据点之间的中间点，来提供具有上斜率和下斜率的临界孔径的代码，这两个斜率源自归档数据点并延伸到保存的数据点，和

配置成如果源自归档数据点并延伸到后续新的数据点的直线斜率位于临界孔径内，则丢弃保存的数据点的代码。

23. 权利要求22的有形媒体,还包括被配置成将新的数据点设置成新的保存数据点,并基于新的保存数据点计算新的临界孔径的代码。

24. 权利要求22的有形媒体,还包括被配置成为归档数据点和后续新的数据点之间的数据点识别所有孔径的交集的代码,每个孔径对于每个相应的点具有源自归档数据点并延伸到容差带的上限和下限的上斜率和下斜率。

25. 权利要求22的有形媒体,还包括被配置成如果斜率不在临界孔径内,则归档保存的数据点的代码,其中保存的数据点变成新的归档数据点。

26. 权利要求24的有形媒体,还包括被配置成将一个或多个数据点提供给数据过滤算法以作为先前记录的数据点的代码。

27. 权利要求24的有形媒体,还包括被配置成通过不相交孔径之前的所有临界孔径的交集来确定新的临界孔径的代码。

28. 权利要求27的有形媒体,其中不相交孔径不具有与先前临界孔径一样的斜率。

29. 权利要求27的有形媒体,还包括被配置成至少对于容差带接收用户输入的代码。

30. 一种方法,包括:

基于归档数据点、保存的数据点、或容差带、或其组合来计算临界孔径;

如果所计算的从归档数据点到新数据点的斜率落在所计算的临界孔径内,则丢弃保存的数据点;并且

计算新的保存数据点的新临界孔径,其中新的保存数据点是新的数据点。

31. 权利要求30的方法,还包括如果斜率没有落在所计算的临界孔径内,则归档保存的数据点。

32. 权利要求30的方法,还包括计算源自归档数据点并延伸到虚拟数据点的直线斜率,其中虚拟数据点的容差带位于临界孔径内。

33. 权利要求30的方法,还包括基本上实时地提供归档数据点、保存的数据点、或容差带、或其组合。

34. 权利要求30的方法,还包括提供归档数据点、保存的数据点、

或容差带、或其组合作为先前记录的数据。

35. 权利要求30的方法，还包括通过在容差带内线性内插来重构丢弃的保存数据点。

36. 一种方法，包括：

将从工业系统采样的数据点集合提供给数据管理系统，其中数据点集合表示工业系统中一个或多个可测量的过程参数；

提供容差带给数据管理系统；并且

通过从数据点集合丢弃一个或多个数据点来获得压缩的数据点集合，其中丢弃的一个或多个数据点落在表示丢弃的一个或多个数据点趋势的直线周围的容差带内。

37. 权利要求36的方法，还包括经由设置在工业设备上的一个或多个传感器来测量数据点集合。

38. 权利要求36的方法，还包括基本上实时地将数据点集合从工业设备传递到数据管理系统。

39. 权利要求36的方法，还包括将数据点集合传递到数据管理系统作为记录的数据。

40. 权利要求36的方法，其中工业设备包括涡轮设备、发电机设备、热能设备、热交换设备、熔炉、液体压缩设备、或其组合。

41. 一种数据管理系统，包括：

用于基于保存的数据点的容差带以及归档数据点和保存的数据点之间的中间点，来提供具有上斜率和下斜率的临界孔径的装置，这两个斜率源自归档数据点并延伸到保存的数据点，并且

用于如果源自归档数据点并延伸到后续新的数据点的直线斜率位于临界孔径内，则丢弃保存的数据点的装置。

42. 一种系统，包括：

经由一个或多个传感器监控的工业设备，其中传感器提供表示工业设备中任何可测量活动的的数据点集合；和

数据管理系统，被配置成从所述一个或多个传感器接收数据点集合，其中数据管理系统包括临界孔径压缩模块，被配置成从数据点集合丢弃一个或多个数据点，其中丢弃的一个或多个数据点落在表示丢弃的一个或多个数据点趋势的直线周围的容差带内，并且其中剩余的数据点集合表示压缩的数据点集合。

43. 权利要求42的系统, 其中工业设备包括涡轮设备、熔炉设备、热交换设备、液体压缩设备、发电机设备、或其组合中的至少一个。

44. 权利要求42的系统, 其中一个或多个传感器适于从工业设备测量温度、压力、时间、流速、速率、电压、电流、振动、速度、加速度或它们的组合。

45. 权利要求42的系统, 其中临界孔径压缩模块包括处理器组件、临界孔径计算模块或临界孔径比较模块。

46. 权利要求42的系统, 其中数据管理系统包括数据获取模块, 被配置成经由一个或多个传感器接收数据点集合并将数据点集合提供给临界孔径压缩模块。

47. 权利要求42的系统, 还包括数据存储模块, 被配置成接收和存储来自临界孔径压缩模块的具有一个或多个数据点的集合。

48. 权利要求42的系统, 其中临界孔径压缩模块基本上实时地接收数据点集合或将其接收作为先前记录的数据。

49. 权利要求42的系统, 其中临界孔径压缩模块被配置成从数据点集合丢弃一个或多个数据点, 其中丢弃的一个或多个数据点落在表示丢弃的一个或多个数据点趋势的直线周围的容差带内, 并且其中剩余的数据点集合表示压缩的数据点集合。

50. 权利要求42的系统, 其中数据管理系统包括用户接口模块, 被配置成将来自系统操作员的一个或多个输入提供给临界孔径压缩模块, 该一个或多个输入包括最小可允许误差、最大可允许误差、压缩之前的最小时间、强迫归档前的最大时间以及压缩使能切换中的至少一个。

51. 一种操作工业设备中的数据管理系统的方法, 包括:

测量表示来自工业设备的一个或多个处理参数的数据点集合;

基于容差带和归档数据点中的至少一个, 来经由数据过滤算法计算保存的数据点的至少一个临时孔径; 和

当所计算的从归档数据点到后续新数据点得出的直线斜率落在临界孔径内时, 丢弃保存的数据点。

52. 权利要求51的方法, 包括当所计算的斜率没有落在临界孔径内时, 归档保存的数据点。

53. 权利要求51的方法, 包括经由放置在工业设备附近的一个或

多个传感器来基本上实时地测量来自工业设备的数据点集合，其中工业设备包括涡轮、发电机、电动机、压缩器、和热交换器中的至少一个。

54. 权利要求51的方法，包括经由数据过滤算法通过在不相交孔径之前的所有临界孔径的交集来确定临界孔径，其中不相交孔径不具有与先前计算的临界孔径一样的斜率。

工业数据压缩系统和方法

技术领域

本发明总的涉及数据管理系统,且更特别地涉及用于在数据管理系统中有效的数据压缩和数据存储的系统和方法。

背景技术

在许多企业中,数据管理系统对于有效监控、采样、存储、维护、处理和分析与特定装备、过程等有关的数据是重要的。例如,某些数据管理系统用于通过收集操作数据或参数来在不同的时间和位置监控各种过程,这些数据或参数用于所监控过程的后续分析。在某些过程中,操作数据或参数包括压力、温度、声音、速率、液体流速、或任何其它可测量的物理、化学、或生物参数。随后的分析可包括性能评估、误差分析、成本分析、故障预测或预期寿命分析、或对各种过程的另一所期望的评估。

在许多这些分析中,基于数据采样率、数据存储容量、数据压缩效率和其它瓶颈来限制数据管理系统。这些分析的精确度通常取决于数据采样的程度,即数据采样水平越高导致更高的分析精确度。不幸的是,数据存储容量和压缩效率的限制通常导致比所期望的数据采样率更低的采样率,因此降低了特定分析的精确度。在过程控制应用中,例如以非常频繁的时间间隔从一个或多个并行过程收集数据(例如每秒一个或每分钟一个),这导致了非常大量的所收集数据。因此,数据压缩对于降低所收集数据的存储消耗是重要的,同时还维持了所收集数据的重要部分以便在所期望的分析中使用。

现有的数据压缩技术降低了所收集数据的存储消耗,而这些技术为了便于存储而牺牲了数据精确度。例如,一个现有的数据压缩技术涉及组合和平均一组顺序数据以创建一个平均数据点。但是,通过平均,数据管理系统丢失了平均数据点之上和之下的变化,这样基于平均数据可导致不精确的预测。类似地,其它现有的数据压缩技术可丢弃有效和显著的数据以为后续进入的数据保存或打开存储空间。

因此,存在对用于精确数据分析和预测的有效数据压缩技术的需要。

发明内容

根据本发明的某些实施例,提供一种用于压缩数据点集合的系统。该系统包括临界孔径压缩模块,其被配置成从数据点集合丢弃一个或多个数据点。丢弃的一个或多个数据点落在表示丢弃的一个或多个数据点趋势的直线周围的容差带内。剩余的数据点集合表示压缩的数据点集合。

根据本发明的某些其它实施例,提供一种实现数据压缩的方法。该方法涉及提供临界孔径,其具有源自归档数据点并向保存的数据点延伸的上斜率和下斜率。基于保存的数据点的容差带以及归档数据点和保存的数据点之间的中间点来获得上斜率和下斜率。该方法还涉及如果源自归档数据点并延伸到后续新的数据点的直线斜率位于临界孔径内,则丢弃保存的数据点。

根据本发明的又一方面,提供一种用于压缩从工业设备采样的数据点集合的系统。经由一个或多个传感器来监控工业设备,该一个或多个传感器提供表示工业设备中任何可测量活动的数据点集合。该系统还包括数据管理系统,被配置成从一个或多个传感器接收数据点集合并从该数据点集合丢弃一个或多个数据点。

附图说明

当参考附图阅读下面的详细叙述时,可更好地理解本发明的这些和其它特征、方面和优点,在整个附图中相同的符号表示相同的部件,其中:

图1是说明根据本技术的实施例的示例系统的图示,该系统具有多个传感器和临界孔径压缩模块;

图2是说明图1所示的临界孔径压缩模块的示例实施例的图示;

图3是说明临界孔径压缩模块在评估数据点集合中示例操作的图解;

图4是说明根据本技术的实施例的孔径的图解,该孔径用于识别何时基于归档数据点、后续的新数据点、和可允许的偏差来保存数

据点；

图5是说明根据本技术的实施例的过程的流程图，该过程用于定义如图4所示的临界孔径；

图6是说明根据本技术的实施例的图1所示的临界孔径压缩模块20的数据过滤过程的流程图；

图7是说明图6所示的数据过滤过程的一个示例实施例的图解；和图8是说明图6所示的数据过滤过程的另一个示例实施例的图解。

具体实施方式

这里公开了用于提供信息存储减小的系统和方法的各种实施例，所述信息由示例数据管理系统收集。这些实施例有效地减小数据点集合，该数据点集合表示从示例工业工厂记录的一个或多个可测量的操作参数。例如，如下面进一步详细讨论的，本技术的某些实施例定义了与先前的归档数据点有关的临界孔径（例如，上斜率和下斜率）。如果后续的新数据点在该临界孔径内，则将其丢弃，而如果后续的新数据点在该临界孔径外，则将其保存。以这种方式，这些实施例提供了有效和节省成本的数据压缩技术，该技术使存储空间的使用减少。

现在转向附图，并首先参考图1，示出一个示例数据管理系统10，用于压缩来自数据点集合的一个或多个数据点。数据管理系统10用于经由多个传感器14从工业设备12收集数据16。数据管理系统10包括数据获取模块18，用于从传感器14获得数据16、临界孔径压缩模块20，用于过滤或压缩数据16、用户存储模块22，用于存留过滤或压缩的数据、以及用户接口模块24，用于与系统10交互并将用户输入26提供到系统10中。在下面讨论的某些实施例中，临界孔径压缩模块20有利于通过丢弃落在与先前的归档数据点有关的临界孔径（例如，上斜率和下斜率）内的数据点来降低数据16的存储消耗。换句话说，临界孔径压缩模块20丢弃临界孔径内一系列的数据点，直到后续数据点落在临界孔径之外。在该外侧的数据点处，临界孔径压缩模块20存留、存储或一般地保存外侧的数据点，以作为下一个归档数据点，并且过程重复。在随后的部分中会更详细地讨论前

述的每个模块，包括临界孔径压缩模块20。

工业设备12包括各种执行或促进工业任务的装置。工业设备12的例子包括液压设备；发电装备，比如燃气涡轮、流式涡轮、发电机；燃气涡轮飞机引擎；热传输装备，比如锅炉、空气预热器、过热器、或其它热交换装置；以及热生成装备，比如熔炉、装置运转监视，比如造纸厂；和/或生产设施，比如汽车配件工厂。传感器14被配置为监控工业单元12中操作参数的变化。

操作参数可包括温度、压力、容量、重量、质量、数量、速率（比如线性或旋转）、加速度、流速（比如，液体流、空气流等）、位置、张力、电流、电压、功率、工作输出、时间、颜色、硬度、粘性、湿度/湿气等。由此，传感器14包括各种传感电路和机构，比如致动器、开关、标准尺、仪表、光学装置、温度测量装置（例如，温度计、热电偶、热敏电阻、高温计等）、压力测量装置（例如，气压计、压力计等）、刻度尺、计数器、计时器、位移测量装置、速度计、频率测量装置、流量计、测力计、容量计、安培计、欧姆表、电位计、转速计、色度计、分光光度计、应变仪、粘度计、光度计、反射计、浊度计、湿度计、热量计等。在本示例实施例中，传感器14便于监控、获取、采样、分析和响应控制工业设备12。例如，如果传感器14识别速率、温度或其它操作参数的不期望状态或变化，则系统10可通过调整适当的操作参数来应答。在其它示例实施例中，传感器不直接提供感测的数据。相反，感测的数据可从归档系统或数据获取系统获得，这些系统被配置为通过调整信号增益、降低噪声来修改感测数据，或基于某些测量标准来转换感测数据。修改的感测数据可由临界孔径压缩模块一前一后地或与实际感测数据相分离地处理。

在操作中，数据管理系统10便于监控、获取、压缩、存储和分析来自多个传感器14的数据16。例如数据获取模块18与传感器14交互以便以规定的时间间隔监控和获取（或采样）与多个操作参数对应的数据16。数据获取模块18的实施例提供对采样率、数据类型、工程测量单元等的控制。在一些实施例中，系统10将数据16传送到数据存储模块22，以作为数据点28的集合或流。

在某些示例实施例中，系统10将数据16传送到孔径压缩模块20，

其基于如下面进一步详细讨论的数据过滤算法来分析和压缩数据。在操作中，临界孔径压缩模块20使用该数据过滤算法来减小数据点28的集合，由此降低数据点28的存储消耗。数据过滤算法的操作机制在下面更详细地描述。在某些实施例中，数据获取模块18可被配置成基本实时地提供数据点28的集合给临界孔径压缩模块20，其在从数据获取模块18接收数据时执行数据过滤或压缩。以这样的方式，数据点28的集合可在它们被存储在数据存储模块22时被过滤或压缩。在其它实施例中，临界孔径压缩模块20可接收数据点28的集合作为先前记录的数据。在这样的场景下，数据获取模块18可提供数据点28的集合给数据存储模块22，其中临界孔径压缩模块20可访问数据点28的集合用于后续分析和压缩。

当在临界孔径压缩模块20中应用数据过滤算法之后，数据存储模块22为采样的数据点28的集合和/或压缩的数据点集合提供可靠的存储空间。在某些实施例中，数据存储模块22包括本地存储装置、远程存储装置、计算机网络或数据存储装置、或其它适合的存储结构。但是，数据存储模块22可包括一个或多个磁性媒体驱动器、硬盘驱动器、光驱动器、或其组合。在一些实施例中，数据存储模块22包括独立（或廉价）磁盘冗余阵列（RAID），比如冗余硬盘驱动器的阵列。

用户接口模块24由系统操作员使用以便与临界孔径压缩模块20交互并提供输入26给该模块。输入26包括影响由临界孔径压缩模块20执行的过滤或压缩量的各种标准。例如，输入26可包括容差带或偏差限制，临界孔径压缩模块20使用该容差带或偏差限制来确定数据点何时被保存或从数据点28的集合丢弃。在某些实施例中，容差带或偏差限制包括上限和下限，比如与在数据点28的集合中最后归档的点有关的上斜率和下斜率。如下面详细讨论的，这些上斜率和下斜率定义与最后归档的点有关的临界孔径。如下所讨论的，临界孔径压缩模块20的过滤算法使用该临界孔径来确定何时存留或丢弃数据点28集合中的后续点。用户接口还可用于在超时发生之前启用或禁用数据压缩、设置最小时间间隔、和设置最大时间间隔。

图2是说明根据本技术的某些实施例的具有图1所示的临界孔径压缩模块20的系统30的图示。在所示的系统30中，临界孔径压缩模

块20包括处理器组件32、临界孔径计算模块34、和临界孔径比较模块36。在操作中，处理器组件32分别用数据存储模块22、临界孔径计算模块34、和临界孔径比较模块36来接收、处理、交换、和传送数据和控制信号38、40和42。此外，处理器组件32执行数据过滤算法，以便压缩数据点28的集合，数据点28的集合是由数据获取模块18从图1所示的工业设备12处的多个传感器14获得的。处理器组件32还被配置为经由用户接口模块24从系统操作员（未示出）接收用户输入26。并且，用户输入26可包括各种压缩标准，比如上容差带或偏置限制和下容差带或偏置限制，以为临界孔径压缩模块20定义临界孔径。

在临界孔径压缩模块20的操作中，以与先前归档的数据点相关的临界孔径，例如上斜率和下斜率看来，数据点28的系列或集合被相继评估。先前归档点之后的下一个点被临时存留为保存点。如果保存点之后的下一个点在临界孔径之外，则临界孔径压缩模块20永久地将保存点存留为下一个归档点。否则，如果保存点之后的下一个点在临界孔径内，则临界孔径压缩模块20丢弃现有的保存点并临时将该下一个点存留为新的保存点。模块20将继续用下一个点来替换临时的保存点，直到下一个点落在临界孔径之外。在到达落在临界孔径之外的下一个点时，模块20将保存点永久地存留为下一个归档点，并且还临时地将下一个点存留为新的保存点。如下面进一步详细讨论的，该过程重复以过滤出数据点28的集合内的点，从而减小数据点28的存储消耗。

在前面过程的每个阶段，临界孔径计算模块34基于先前存档的数据点、当前保存的数据点和下一个数据点的可允许偏差的分析为下一个数据点计算临界孔径。在某些实施例中，临界孔径压缩模块20为每个数据点提供可允许上偏差和下偏差，并且临界孔径计算模块34定义每个数据点的孔径以作为从先前存档的数据点延伸到相应数据点附近上偏差和下偏差的上斜率和下斜率。换句话说，归档数据点和下一个数据点之间的每个数据点包括一个孔径。依次地，临界孔径计算模块34定义下一个数据点的临界孔径以作为归档数据点和下一个数据点之间所有孔径的交集或公共部分。结果，当临界孔径计算模块34分析后续新的点时临界孔径一般收缩。临界孔径，作为

这里使用的术语，可被定义为对于所有先前计算的孔径公共的斜率集合。下面参考图4进一步详细地讨论临界孔径的计算。

现在转向临界孔径比较模块36，该模块被配置为比较临界孔径和下一个数据点。更具体地，临界孔径比较模块36的实施例计算归档数据点和下一个数据点之间的斜率，并接着将该斜率与临界孔径的上斜率和下斜率比较。如果临界孔径比较模块36确定该斜率落在临界孔径的上斜率和下斜率之外，则这是导致先前保存点被归档且下一个数据点变为保存点的触发事件。每个归档点被传送到数据存储模块22，其根据该示例的数据过滤过程来存储压缩的数据44。

作为一个例子考虑，一般由引用数字48到62表示说明一系列数据点的图解（如图3所示）。临界孔径压缩模块20考虑每个数据点，以便从通常为归档数据点的数据点48开始以顺序的方式进行可能的压缩。容差带63与连接点48和诸如点62的后续数据点的直线相交的任何点可通过在规定的限制内线性内插而被重构。对于其重构值超过真实值的部分比可允许的误差还大的数据点60，不是这样的情况。因此，点62不能是归档线段的端点。由于数据点60是最后的点，将其连接到点48的直线与所有插进的容差带相交，数据点60是定义该线段端点且将被归档的数据点。为了验证每个前述的容差带与每条将归档数据点（数据点48）连接到新引入的数据点的后续直线相交，该数据系列将通常需要被保存，直到识别结束的数据点，但是下面描述的临界孔径技术允许结束的数据点在不必要保存多于单个保存数据点的情况下而被找到。本技术对于数据压缩高度有效，因为其评估每个数据点自身的价值而不是要求一组数据点被缓存和比较，直到获得结果。

在本文中，图4是说明根据本技术的实施例的孔径的图解，该孔径用于识别何时基于归档数据点A、后续的新数据点C、和可允许的偏差来保存数据点B。在所示的实施例中，点B、A和C由具有坐标 (x_1, y_1) 的保存数据点64、具有坐标 (x_0, y_0) 的归档数据点66、和具有坐标 (x_2, y_2) 的后续下一个数据点68表示。可允许的偏差 e 由可允许的偏差70来表示，其包括先前部分描述的上容差带或偏差限制和下容差带或偏差限制。如所示的，保存点64的孔径72包括由三角形包括的阴影区域，该三角形由归档数据点66、可允许偏差70的上端

点74和可允许偏差70的下端点76形成。数学上，保存数据点64的孔径72由线段78和80界定，它们具有如下计算的斜率：

$$\text{线段78的斜率} = \frac{y1 + e - y0}{x1 - x0}$$

$$\text{线段80的斜率} = \frac{y1 - e - y0}{x1 - x0}$$

对于具有坐标 (x_2, y_2) 的数据点68，将数据点68与归档数据点66联接的线段82的斜率可如下计算：

$$\text{线段82的斜率} = \frac{y2 - y0}{x2 - x0}$$

在某些示例实施例中，在所计算的保存数据点64的孔径72是仅存的所计算孔径的情况下，孔径72变为临界孔径。并且，在当前实施例中，如果线段82的斜率位于临界孔径72的线段78和80的上斜率和下斜率之内，则保存数据点64可被丢弃并且数据点68变为新的保存数据点。现在可为新的保存数据点68计算新的孔径。如果线段78的斜率大于或等于线段82的斜率且如果线段82的斜率大于或等于线段80的斜率，则线段82位于临界孔径内。当线段82不位于临界孔径72内时，保存的数据点64可被归档且变为新的归档数据点。还可在超过用户规定的最大超时且在最大超时持续期间没有点被归档时归档保存的数据点64。

图5是说明根据本技术的实施例的过程84的流程图，该过程用于定义如图4所示的临界孔径。过程84涉及作为输入来在步骤86提供最后归档的数据点(A)、保存数据点(B)的坐标和允许偏差(e)。在步骤88，基于所提供的输入，上端点(U)和下端点(L)通过将允许偏差(e)外插在保存数据点(B)的坐标上而被计算。在步骤90，由图4中的线段78和80表示的上直线AU和下直线AL的斜率被计算，且这些斜率定义孔径。在步骤92，过程84为先前保存的数据点确定孔径(例如，上斜率和下斜率)与所有先前计算的孔径的交集。换句话说，过程84的步骤92从最近存档的数据点(A)直到当前保存的数据点(B)来识别所有孔径的公共部分。结果，步骤92识别所有孔径共同具有的最大上斜率和最小下斜率。这些斜率将是来自先前临界孔径和当前孔径的两个上斜率中的较低者以及来自先前临界孔径和当前孔径的两个下斜率中的较高者。在步骤92处确定的公共上斜率和下斜率定义如框94所示的临界孔径。

图6是说明根据本技术的实施例的图1所示的临界孔径压缩模块

20的数据过滤过程96的流程图。数据过滤过程96被配置成计算具有一个或多个临界孔径的集合，并且基于相应的临界孔径通过从数据点28的集合丢弃一个或多个数据点来归档数据压缩。应当注意，尽管丢弃一个或多个数据点，数据过滤算法还便于重构机制，通过该重构机制，丢弃的数据点可基于所存储的有关归档数据点的信息来通过线性内插被重构。

如图6所示，过程96涉及作为输入在步骤98接收与最后归档的数据点、保存的数据点、和容差带有关的信息。在步骤100，过程96涉及通过图4和5所示和在前面部分描述的示例技术来计算临界孔径。一旦临界孔径被确定，与新数据点有关的信息在步骤102被提供。在步骤104，从最后归档数据点到新数据点的斜率被计算。在步骤106，将在步骤100计算的临界孔径与在步骤104确定的新数据点的斜率相比较。

在询问步骤106，过程96将新数据点的斜率（步骤104）与临界孔径（步骤100）比较。该比较产生两个相互排斥的结果之一：（1）新数据点的斜率落在临界孔径的上斜率和下斜率内，或（2）新数据点的斜率落在临界孔径的上斜率和下斜率之外。当所计算的新数据点的斜率落在临界孔径的斜率内时，步骤108跟进。在此时，保存的数据点被丢弃。在步骤110，新的数据点变为新的保存点。在步骤112，过程96按如上详细讨论的来计算新的孔径。过程96接着按如上详细讨论的在步骤114计算新的临界孔径。步骤108到步骤114的序列可进一步重复，直到考虑了所有数据点。

如果在询问步骤106新数据点的斜率没有落在临界孔径内，则过程96进行到步骤116。该事件可被定义为不相交孔径。在该步骤116，保存的数据点被归档且变为新的归档数据点。在步骤118，新的数据点变为新的保存数据点，并且在步骤120，新的临界孔径被计算。以类似方式，上述的步骤116到120的序列可对于后续新的数据点进一步重复，直到来自数据点28的集合的所有数据点已经被考虑。因此，通过在一个时间段上重复该方法，并且在从如图1所示的工业设备12采样的数据点28的集合中从一个数据点进行到另一个数据点，可从数据存储模块22（如图1所示）丢弃大量的数据点。这导致数据点28的集合所消耗的存储空间减小。

图7是说明图6所示的数据过滤过程的一个示例实施例的图解。在该示例实施例中，数据点122到136被考虑为通过数据过滤过程96来压缩。数据点122是归档的数据点，并且数据点124到132是由数据过滤过程96丢弃的中间数据点，因为这些数据点的斜率在临界孔径138的斜率内。如前所解释的，中间数据点126被丢弃，因为联接数据点126到归档数据点122的直线穿过所计算的临界孔径138直到先前的数据点124。类似地，中间数据点128到132被丢弃，因为联接每个中间数据点到归档数据点122的直线穿过所计算的临界孔径直到每个相应中间数据点。应当注意，在特殊数据点处的临界孔径138完全由每个所计算的孔径封闭，直到达到了该特殊数据点。但是，由于联接数据点136到归档数据点122的直线位于临界孔径之外，数据点134变为新的归档数据点且点136变为新的保存点。应当注意，联接数据点134到归档数据点122的直线穿过所计算的临界孔径138直到数据点132。此外，还应当注意，在所示的图解中，中间数据点122到132位于联接数据点134到归档数据点122的直线线段中和周围，位于的方式是每个被丢弃的数据点的容差带与该直线段相交。因此，可以理解，联接归档数据点122和保存数据点134的直线段是被丢弃的数据点124到132趋势的良好表示，并且这些丢弃的数据点在所期望时通过在固定的误差限制(ϵ)内线性内插直线段的坐标来被重构。

图8是说明根据技术的实施例的另一个示例数据过滤过程的图解。数据点140到158的集合包括归档数据点140和中间数据点142到150。在该示例实施例中，具有一般由数字160、162和164表示的一个或多个虚拟数据点的集合被如此选择，使得这些虚拟数据点位于实际数据点的容差带内，在该示例情况中实际数据点由数字156来表示。现在可基于虚拟数据点而不是实际数据点来选择保存的数据点152。计算从归档数据点140并迭代地从数据点142到156开始的临界孔径166。但是数据点158位于临界孔径之外，因为联接数据点158的上端点和下端点的直线不与临界孔径166相交。因此，如此选择一个或多个虚拟数据点使得虚拟数据点位于实际数据点的容差带内，实际数据点在本示例情况中由数字156表示，并且使得所述虚拟数据点与临界孔径166相交。由该示例实施例提供的优点在于，彼此靠近的某些数据点可被丢弃，以便能够更好地考虑位于更远距离的数据

点。这还进一步使得能够在更大的数据点集合进行更高的压缩，如模拟所指示的。

在前面部分描述的各个实施例和方面可包括可执行指令（例如，顺序或面向对象的代码），用于实现各种逻辑功能。可执行指令可体现在任何计算机可读介质中以供或联合基于计算机的系统来使用，基于计算机的系统可取得指令并执行它们。在本申请的上下文中，计算机可读介质可以是包含、存储、传递、传播、传输或传送该指令的任何装置。计算机可读介质可以是电、磁、光、电磁、或红外系统、设备、或装置。如所示的，有形媒体的非穷尽列表可包括具有一条或多条导线的电连接（电的）、便携式计算机磁盘（磁的）、随机访问存储器（RAM）（磁的）、只读存储器（ROM）（磁的）、可擦除可编程只读存储器（EPROM或闪速存储器）（磁的）、光纤（光的）、和便携式压缩盘只读存储器（CDROM）（光的）。注意到，有形媒体还可包含纸件或另一适当的介质，在其之上可借助机械或电子的方式或手写的方式来印刷指令。例如，指令可经由光扫描纸件或其它介质而电子地获取，接着被编译、解释或如果需要以适当的方式处理，并且接着被存储在计算机的可读存储器中。

如本领域技术人员将理解的，上述的数据管理系统和技术通过从数据点集合丢弃一个或多个数据点来提供有效的数据压缩，同时允许利用来自所存储的数据点的信息借助线性内插来容易地重构丢弃的数据点。丢弃的数据点的重构对于任何类型的数据分析或计算可在任何稍后时间及时使用，该重构允许数据管理系统维持提高的与数据点集合代表的工业设备操作参数有关的精确度水平。数据管理系统还通过诸如平均的方法，在经由数据过滤算法通过考虑来自数据点集合的每个数据点计算数据点集合并且评估特殊数据点是否应当被丢弃时，来保持提高的精确度水平。数据管理系统因此允许对数据集合使用减小的存储容量。此外，数据管理系统允许对于给定存储容量存储数据点更长的时间段，这导致增加的成本节约。此外，这种类型的数据压缩不会将更多的数据打包成较少的位。它仅仅丢弃其可重构的数据，只留下可在线性内插数据点重构方法中使用的线段端点。并且，数据压缩技术不可总执行数据数据压缩。数据压缩不能被执行的示例情形包括当数据点不能由任何趋势或模式定义

时的情形。

尽管只有本发明的某些特征在这里被说明和描述,但对于本领域技术人员可出现许多修改和变化。因此,应当理解,所附权利要求旨在覆盖所有这种落在本发明真实精神内的修改和变化。

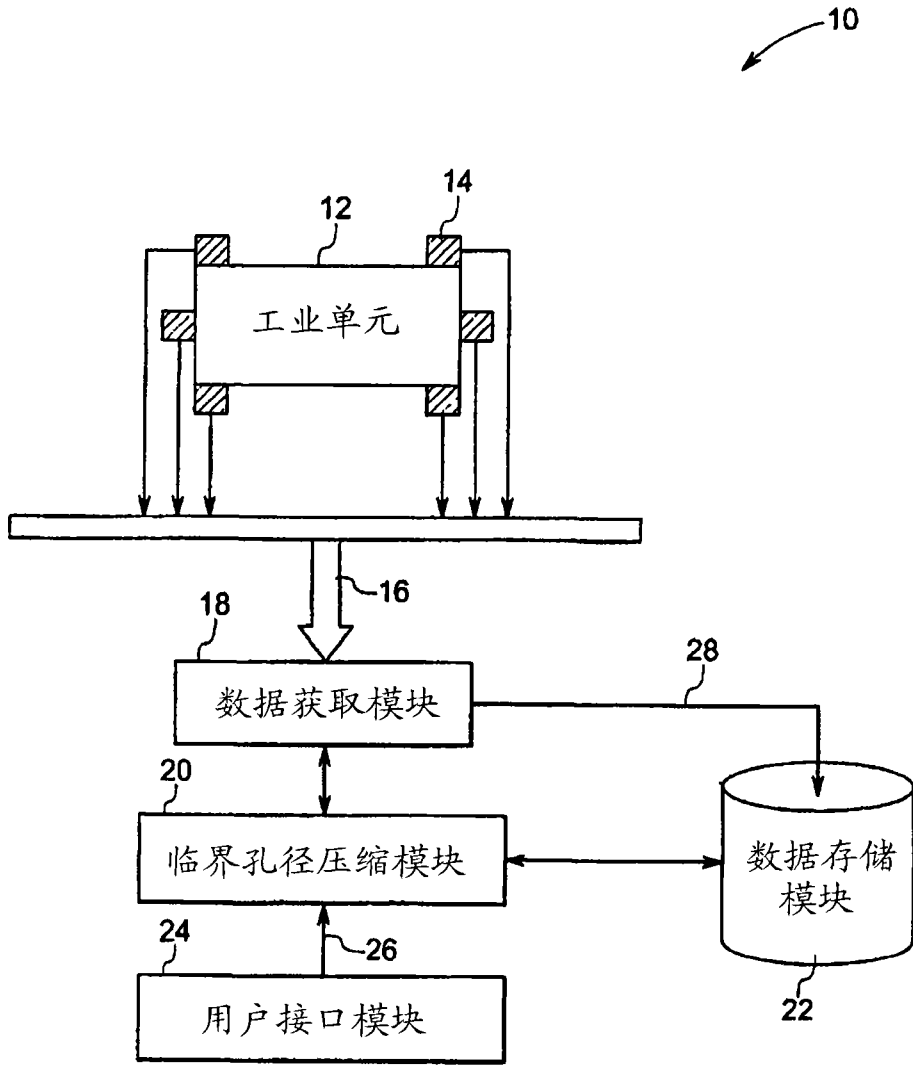


图 1

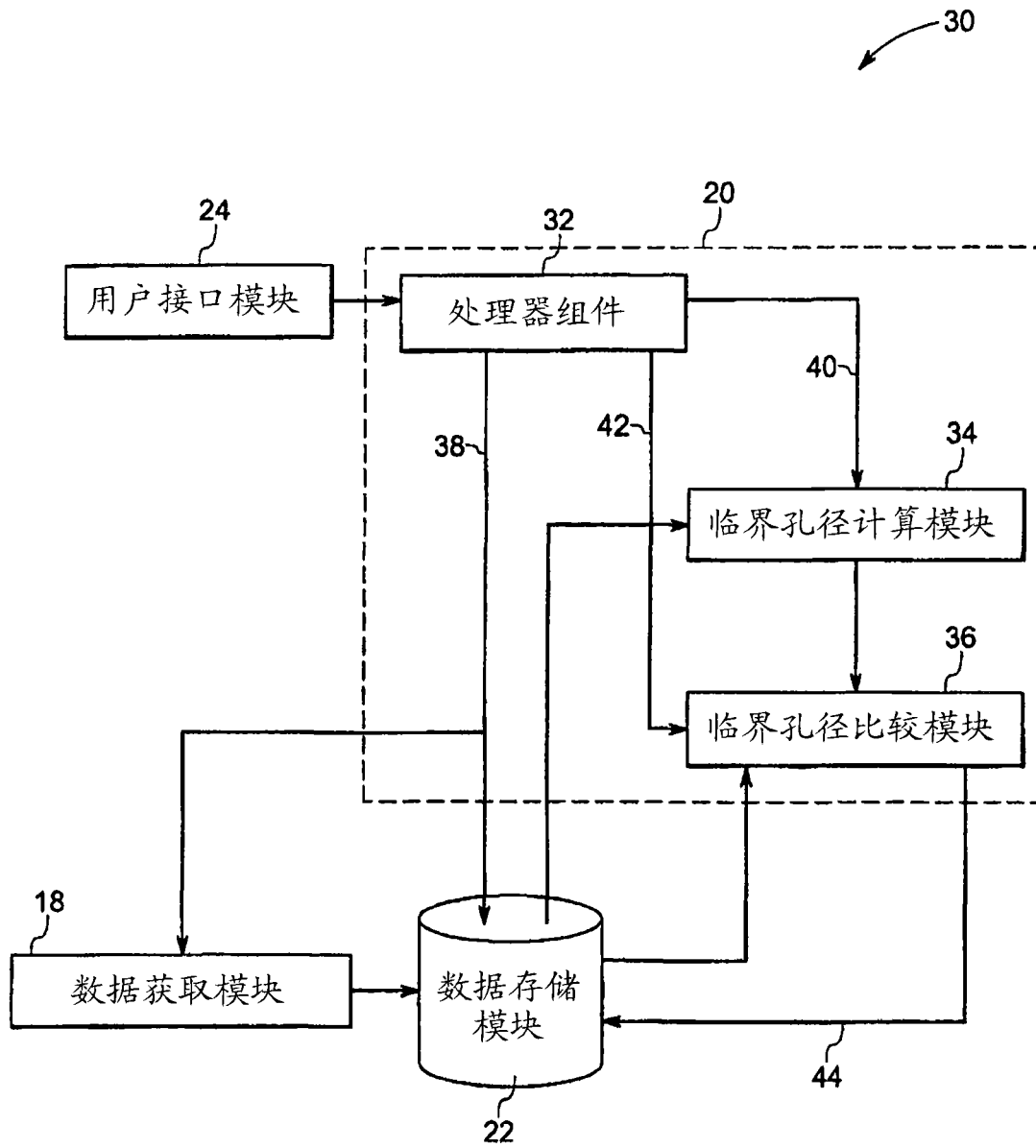


图 2

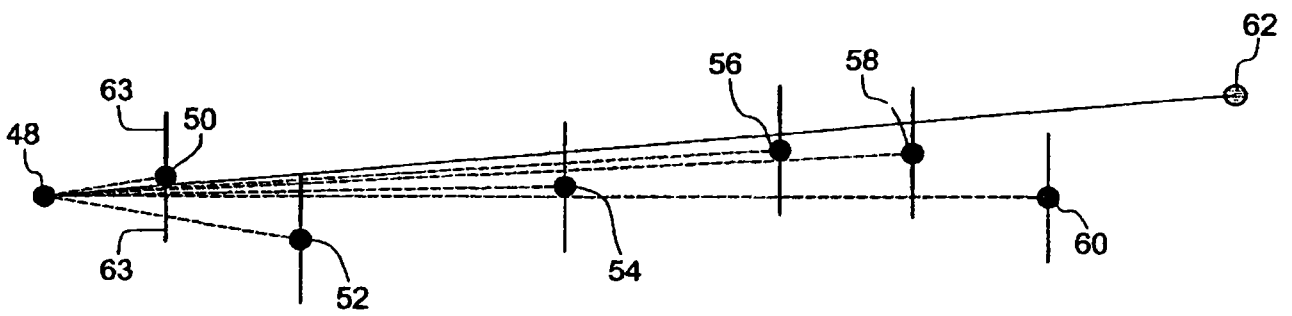


图 3

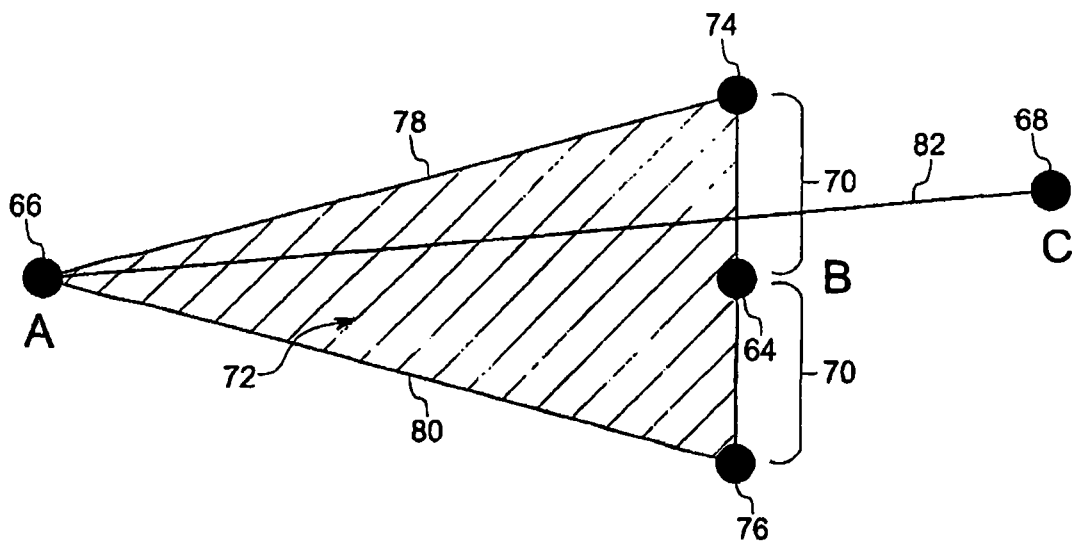


图 4

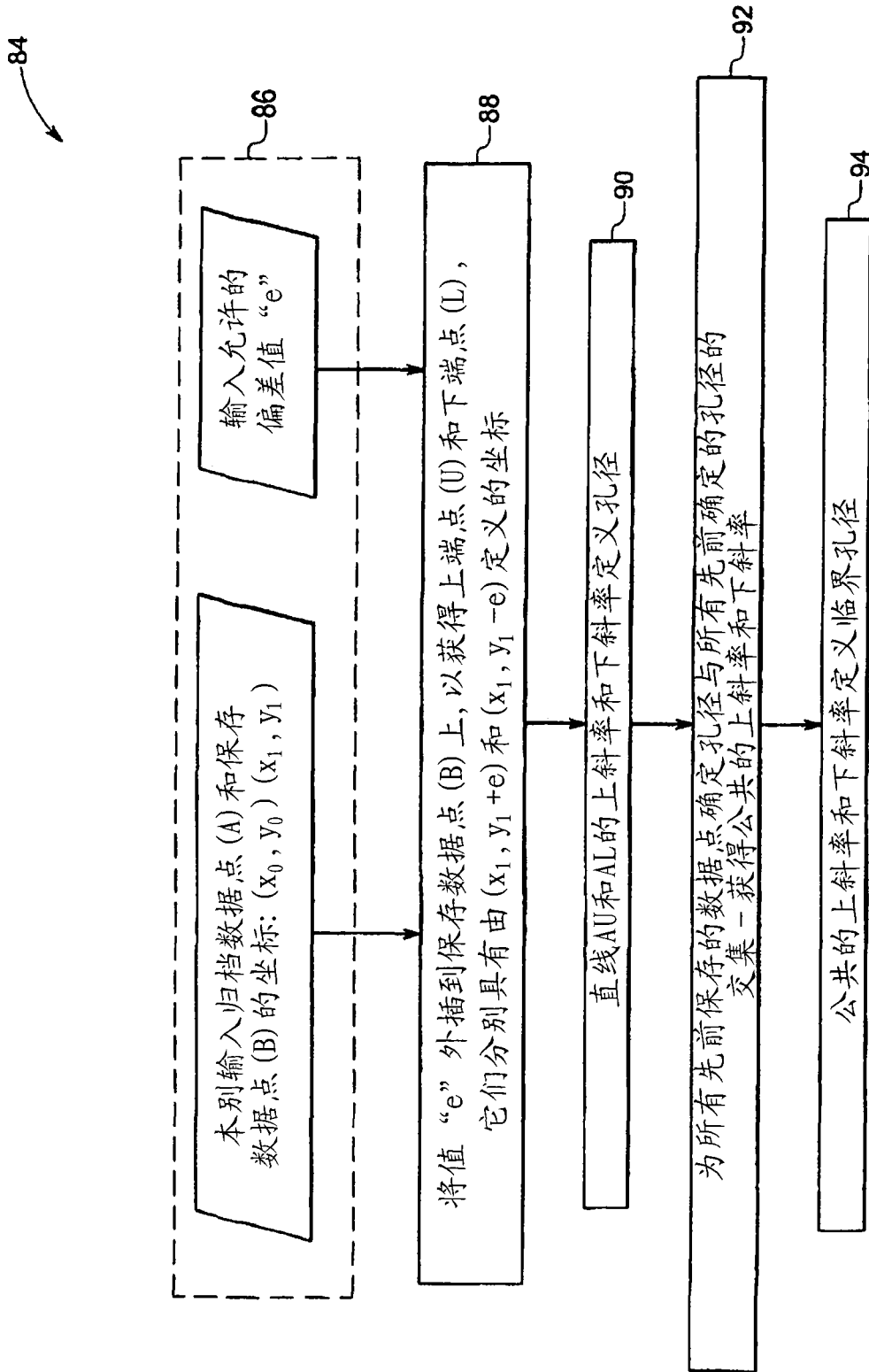


图 5

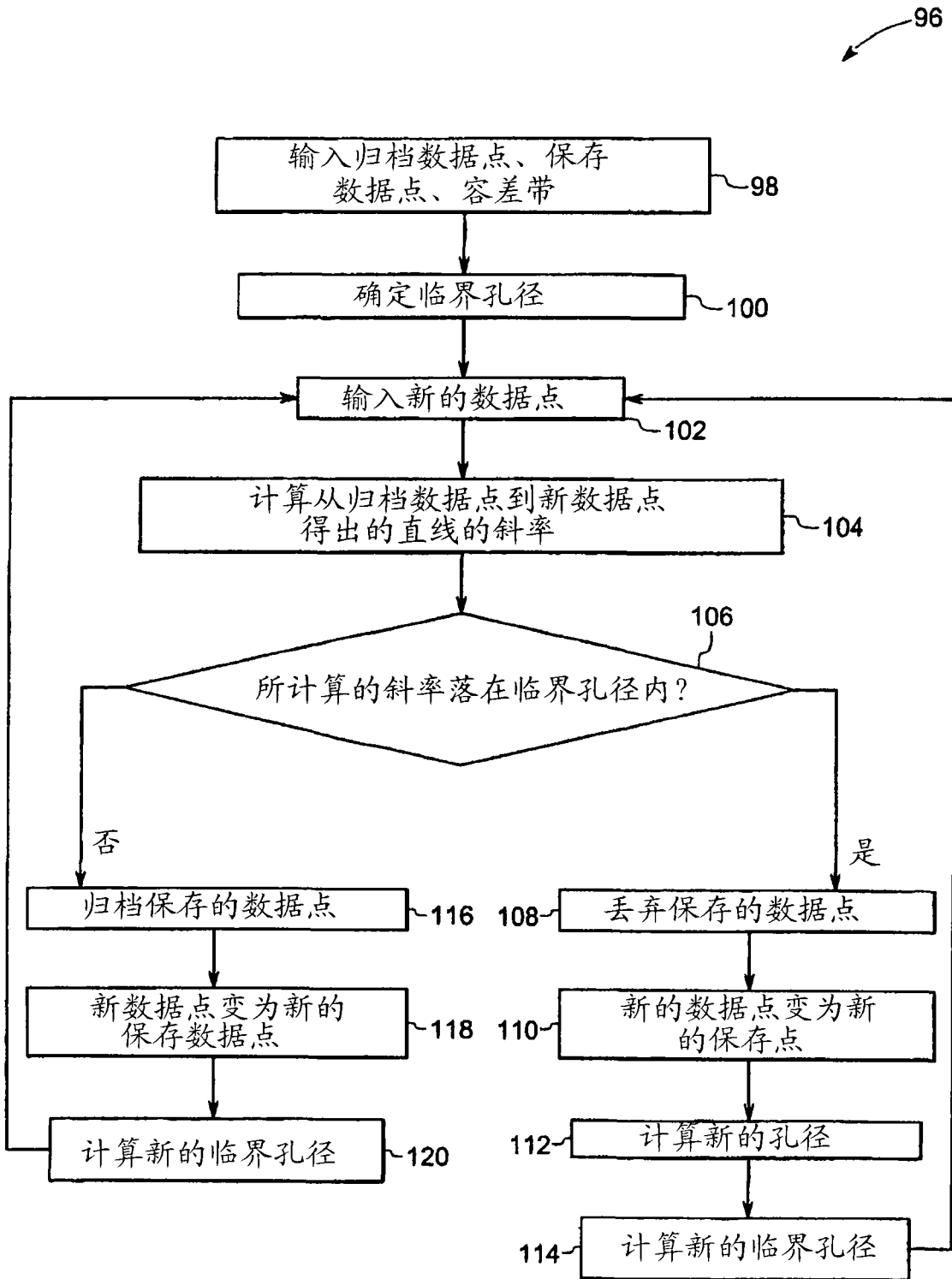


图 6

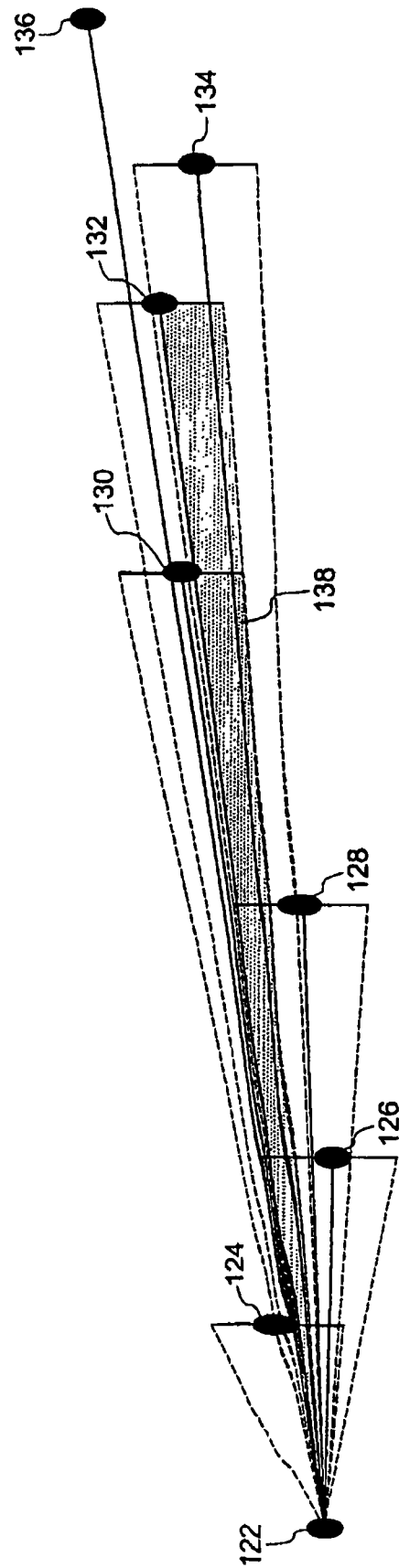


图 7

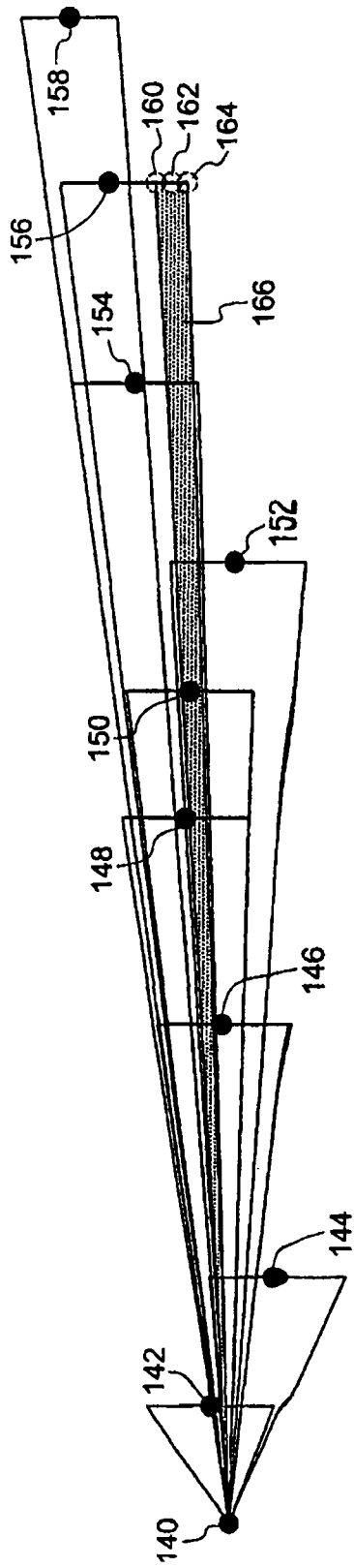


图 8