(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 115266604 A
(43)申请公布日 2022.11.01

(21)申请号 202210794546.3
(22)申请日 2019.06.28
(30)优先权数据
62/692,248 2018.06.29 US
16/448,914 2019.06.21 US
(62)分案原申请数据
201910576094.X 2019.06.28
(71)申请人 唯亚威通讯技术有限公司
地址 美国亚利桑那州
(72)发明人 熊章民 孙岚
(74)专利代理机构 北京安信方达知识产权代理有限公司 11262
专利代理师 王红英 杨明钊
(51)Int.Cl.
G01N 21/25(2006.01)
G01J 3/28(2006.01)
G06K 9/62(2022.01)

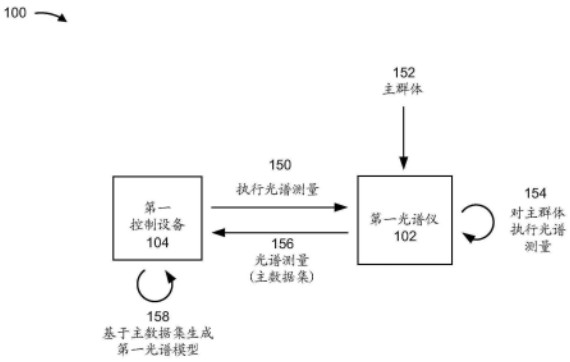
权利要求书3页 说明书13页 附图10页

(54)发明名称

对光谱模型的基于交叉验证的校准

(57)摘要

本公开涉及对光谱模型的基于交叉验证的校准。一种设备可以接收对于第一光谱模型的主数据集;接收对于与第一光谱模型相关联的目标群体的目标数据集以更新第一光谱模型;生成包括主数据集和来自目标数据集中的第一数据的训练数据集;生成包括来自目标数据集而不是主数据集的第二数据的验证数据集;使用交叉验证并使用训练数据集和验证数据集生成作为第一光谱模型的更新的第二光谱模型;以及提供第二光谱模型。



CN 115266604 A

1.一种方法,包括:

通过控制设备接收与第一光谱模型相关联的信息,所述第一光谱模型是基于来自第一光谱仪的主数据集生成的;

基于接收到与所述第一光谱模型相关联的所述信息,通过所述控制设备使得第二光谱仪执行一个或更多个光谱测量;

通过所述控制设备从所述第二光谱仪接收所述一个或更多个光谱测量作为目标数据集;

基于所述目标数据集,通过所述控制设备使用交叉验证生成第二光谱模型;以及

通过所述控制设备使用所述第二光谱模型执行光谱分析。

2.根据权利要求1所述的方法,其中,与所述第一光谱模型相关联的所述信息包括以下中的一个或更多个:

所述第一光谱模型,或

所述主数据集。

3.根据权利要求1所述的方法,

其中,所述主数据集包括由所述第一光谱仪在初始时间对初始群体执行的初始一组光谱测量;并且

其中,所述目标数据集包括由所述第二光谱仪在后续时间对后续群体执行的后续一组光谱测量。

4.根据权利要求1所述的方法,其中,使得所述第二光谱仪执行所述一个或更多个光谱测量包括:

向所述第二光谱仪传输指令,以使所述第二光谱仪执行所述一个或更多个光谱测量。

5.根据权利要求1所述的方法,其中,使得所述第二光谱仪执行所述一个或更多个光谱测量包括:

基于接收到与所述第一光谱模型相关联的所述信息,确定更新或校准所述第一光谱模型;以及

基于确定更新或校准所述第一光谱模型,触发所述第二光谱仪执行所述一个或更多个光谱测量。

6.根据权利要求1所述的方法,其中,生成所述第二光谱模型包括:

基于所述目标数据集生成验证集;以及

基于所述验证集,使用交叉验证生成所述第二光谱模型。

7.根据权利要求1所述的方法,其中,生成所述第二光谱模型包括:

基于所述主数据集和所述目标数据集确定多个训练集;

基于所述目标数据集和不是来自主数据集的数据确定多个验证集;

基于所述多个训练集和所述多个验证集,确定多个性能度量;

基于所述多个性能度量,使用交叉验证法确定最优偏最小二乘PLS因子;以及

基于所述最优PLS因子生成所述第二光谱模型。

8.根据权利要求1所述的方法,其中,生成所述第二光谱模型包括:

基于所述目标数据集,确定与交叉验证相关联的子类的性能度量;

确定所述子类的偏最小二乘PLS因子;

确定所述PLS因子的均方根误差RMSE值;

基于所述PLS因子和所述RMSE值确定最优PLS因子;以及

基于所述最优PLS因子生成所述第二光谱模型。

9.根据权利要求1所述的方法,其中,生成所述第二光谱模型包括:

通过合并所述主数据集和所述目标数据集生成合并的数据集;以及

使用所述合并的数据集和最优偏最小二乘PLS因子生成所述第二光谱模型。

10.一种控制设备,包括:

一个或更多个存储器;以及

一个或更多个处理器,所述一个或更多个处理器被通信地耦合到所述一个或更多个存储器,所述一个或更多个处理器被配置为:

接收与第一光谱模型相关联的信息,所述第一光谱模型是基于来自第一光谱仪的第一数据集生成的;

基于接收到与所述第一光谱模型相关联的所述信息,使得第二光谱仪执行一个或更多个光谱测量;

从所述第二光谱仪接收所述一个或更多个光谱测量作为第二数据集;

基于所述第二数据集,使用交叉验证生成第二光谱模型;以及

使用所述第二光谱模型执行光谱分析。

11.根据权利要求10所述的控制设备,其中,当使得所述第二光谱仪执行所述一个或更多个光谱测量时,所述一个或更多个处理器被配置为:

向所述第二光谱仪传输指令,以使所述第二光谱仪执行所述一个或更多个光谱测量。

12.根据权利要求10所述的控制设备,其中,当使所述第二光谱仪执行所述一个或更多个光谱测量时,所述一个或更多个处理器被配置为:

基于接收到与所述第一光谱模型相关联的所述信息,确定更新或校准所述第一光谱模型;以及

基于确定更新或校准所述第一光谱模型,触发所述第二光谱仪执行所述一个或更多个光谱测量。

13.根据权利要求10所述的控制设备,其中,当生成所述第二光谱模型时,所述一个或更多个处理器被配置为:

基于所述第二数据集生成验证集;以及

基于所述验证集,使用交叉验证法生成所述第二光谱模型。

14.根据权利要求10所述的控制设备,当生成所述第二光谱模型时,所述一个或更多个处理器被配置为:

基于所述第一数据集和所述第二数据集确定多个训练集;

基于所述多个训练集确定多个性能度量;

基于所述多个性能度量,使用交叉验证确定最优偏最小二乘PLS因子;以及

基于所述最优PLS因子生成所述第二光谱模型。

15.根据权利要求10所述的控制设备,当生成所述第二光谱模型时,所述一个或更多个处理器被配置为:

基于所述第一数据集,确定与所述交叉验证相关联的子类的性能度量;

确定所述子类的偏最小二乘PLS因子;

基于所述PLS因子确定最优PLS因子;以及

基于所述最优PLS因子生成所述第二光谱模型。

16.根据权利要求10所述的控制设备,其中,当生成所述第二光谱模型时,所述一个或更多个处理器被配置为:

通过合并所述第一数据集和所述第二数据集生成合并的数据集;以及

使用所述合并的数据集和最优偏最小二乘PLS因子生成所述第二光谱模型。

17.一种非临时性计算机可读介质,所述非临时性计算机可读介质存储一组指令,所述一组指令包括:

一个或更多个指令,所述一个或更多个指令当由控制设备的一个或更多个处理器执行时使所述控制设备:

接收与第一光谱模型相关联的信息,所述第一光谱模型是基于主数据集生成的;

基于接收到与所述第一光谱模型相关联的所述信息,使得光谱仪执行一个或更多个光谱测量;

从所述光谱仪接收所述一个或更多个光谱测量作为目标数据集;

基于所述目标数据集,使用交叉验证生成第二光谱模型;以及

使用所述第二光谱模型执行光谱分析。

18.根据权利要求17所述的非临时性计算机可读介质,其中,使得所述控制设备生成所述第二光谱模型的所述一个或更多个指令,使得所述控制设备:

基于所述目标数据集生成验证集;以及

基于所述验证集,使用交叉验证生成所述第二光谱模型。

19.根据权利要求17所述的非临时性计算机可读介质,其中,使得所述控制设备生成所述第二光谱模型的所述一个或更多个指令,使得所述控制设备:

基于所述目标数据集和不是来自所述主数据集的数据确定多个验证集;

基于所述多个验证集确定多个性能度量;以及

基于所述多个性能度量生成所述第二光谱模型。

20.根据权利要求17所述的非临时性计算机可读介质,其中,使得所述控制设备生成所述第二光谱模型的所述一个或更多个指令,使得所述控制设备:

通过合并所述主数据集和所述目标数据集生成合并的数据集;以及

使用所述合并的数据集和最优偏最小二乘PLS因子生成所述第二光谱模型。

对光谱模型的基于交叉验证的校准

[0001] 本申请是申请日为2019年6月28日,申请号为201910576094.X,发明名称为“对光谱模型的基于交叉验证的校准”的申请的分案申请。

[0002] 相关申请

[0003] 本申请基于美国法典第35条第119款要求于2018年6月29日提交的且题为“UPDATING CALIBRATION MODELS BASED ON NEAR-INFRARED(NIR)SPECTRA”的美国临时专利申请第62/692,248号的优先权,该美国临时申请的内容通过引用以其整体并入本文。

技术领域

[0004] 本申请涉及对光谱模型的基于交叉验证的校准。

背景技术

[0005] 原料识别可用于药物产品的质量控制。例如,可以对医用材料执行原料识别,以确定医用材料的组成成分是否对应于与医用材料相关联的包装标签。类似地,可以执行原料量化以确定特定样本中特定化学物的浓度。光谱法可以有助于进行无损的原料识别和/或量化,与其他化学计量技术相比,其准备和数据采集时间减少。

发明内容

[0006] 本申请的实现方式主要包括以下方面:

[0007] 1)一种设备,可以包括:一个或更多个存储器;以及一个或更多个处理器,所述一个或更多个处理器通信地耦合到所述一个或更多个存储器,所述一个或更多个处理器被配置为:接收对于第一光谱模型的主数据集;接收对于与第一光谱模型相关联的目标群体的目标数据集,以更新第一光谱模型;生成包括主数据集和来自目标数据集的第一数据的训练数据集;生成包括来自目标数据集而不是主数据集的第二数据的验证数据集;使用交叉验证并使用训练数据集和验证数据集生成作为第一光谱模型的更新的第二光谱模型;并提供第二光谱模型。

[0008] 2)根据1)所述的设备,其中,所述一个或更多个处理器被配置为:

[0009] 接收光谱测量;

[0010] 使用所述第二光谱模型来执行光谱测定;和

[0011] 提供识别所述光谱测定的输出。

[0012] 3)根据1)所述的设备,其中,所述训练数据集是多个训练数据集,并且所述验证数据集是多个验证数据集;以及

[0013] 其中,当生成所述第二光谱模型时,所述一个或更多个处理器被配置成:

[0014] 基于所述多个训练数据集和所述多个验证数据集生成多个性能度量;

[0015] 基于所述多个性能度量来确定总的性能度量;

[0016] 基于所述总的性能度量来确定最优偏最小二乘(PLS)因子;和

[0017] 基于所述最优PLS因子和合并的数据集确定所述第二光谱模型,其中,所述合并的

数据集包括所述主数据集和所述目标数据集。

[0018] 4)根据1)所述的设备,其中,所述第一光谱模型和所述第二光谱模型是量化模型。

[0019] 5)根据1)所述的设备,其中,所述主数据集基于由主光谱仪执行的第一组光谱测量,并且所述目标数据集基于由不同于所述主光谱仪的目标光谱仪执行的第二组光谱测量。

[0020] 6)根据1)所述的设备,其中,所述主数据集基于由特定光谱仪执行的第一组光谱测量,并且所述目标数据集基于由所述特定光谱仪执行的第二组光谱测量。

[0021] 7)一种方法,可以包括:通过设备接收对于与第一光谱模型相关联的目标群体的目标数据集;基于接收到所述目标数据集通过所述设备来获得对于所述第一光谱模型的主数据集;通过所述设备使用交叉验证确定最优偏最小二乘(PLS)因子,其中,所述最优PLS因子基于多个(a plurality of)训练数据集且基于多个验证数据集确定,每个训练数据集包括目标数据集的相应部分和主数据集的全部,每个验证数据集包括目标数据集的相应部分并且不包括主数据集的数据;通过所述设备合并所述目标数据集和所述主数据集,以生成合并的数据集;通过所述设备并使用合并的数据集和最优PLS因子生成第二光谱模型,其中,所述第二光谱模型是所述第一光谱模型的更新;以及通过所述设备提供第二光谱模型来代替第一光谱模型。

[0022] 8)根据7)所述的方法,其中,确定所述最优PLS因子包括:

[0023] 确定关于所述多个训练数据集中的每一个训练数据集和所述多个验证数据集中的每一个验证数据集的偏最小二乘(PLS)性能度量;

[0024] 基于所述PLS性能度量确定总的PLS性能度量;和

[0025] 基于所述总的PLS性能度量,优化PLS因子以确定对于所述第二光谱模型的最优PLS因子。

[0026] 9)根据8)所述的方法,其中,所述总的PLS性能度量与均方根误差(RMSE)值相关联,并且

[0027] 其中,优化所述PLS因子包括:

[0028] 优化所述PLS因子以最小化RMSE值。

[0029] 10)根据8)所述的方法,其中,所述多个验证数据集包括与所述多个训练数据集不同的所述目标数据集中的数据。

[0030] 11)根据8)所述的方法,其中,确定所述总的PLS性能度量包括:

[0031] 聚合所述PLS性能度量。

[0032] 12)根据7)所述的方法,其中,所述目标数据集关联于在与所述主数据集相关联的测量之后执行的对所述目标群体的一组测量。

[0033] 13)根据12)所述的方法,其中,所述第二光谱模型是所述第一光谱模型的校准更新模型。

[0034] 14)根据7)所述的方法,其中,所述目标数据集与由特定光谱仪执行的一组测量相关联,所述特定光谱仪不同于执行与所述主数据集相关联的测量的一个或更多个光谱仪。

[0035] 15)根据14)所述的方法,其中,所述第二光谱模型是所述第一光谱模型的校准传递模型。

[0036] 16)根据14)所述的方法,其中,提供所述第二光谱模型包括:

[0037] 提供所述第二光谱模型以用于与通过所述特定光谱仪进行的后续测量结合使用。

[0038] 17)一种非临时性计算机可读介质,可以存储一个或更多个指令。当被设备的一个或更多个处理器执行时,所述一个或更多个指令可以使所述一个或更多个处理器:接收对于第一光谱模型的主数据集;接收对于与第一光谱模型相关联的目标群体的目标数据集,以更新第一光谱模型;基于主数据集和目标数据集生成多个训练数据集;基于目标数据集生成多个验证数据集,其中,所述多个验证数据集不包括主数据集的数据;基于所述多个训练数据集和所述多个验证数据集并使用交叉验证来确定模型设置;基于模型设置、目标数据集和主数据集生成第二光谱模型;并提供第二光谱模型。

[0039] 18)根据17)所述的非临时性计算机可读介质,其中,所述模型设置是以下中的至少一项,

[0040] 偏最小二乘(PLS)模型的PLS因子,

[0041] 主成分回归(PCR)模型的成分的量,

[0042] 支持向量回归(SVR)模型的SVR参数,或

[0043] 预处理设置。

[0044] 19)根据17)所述的非临时性计算机可读介质,其中,使所述一个或更多个处理器确定所述模型设置的所述一个或更多个指令使所述一个或更多个处理器:

[0045] 生成关于所述多个训练数据集中的每一个训练数据集和所述多个验证数据集中的对应的验证数据集的多个部分性能度量;

[0046] 聚合所述多个部分性能度量以生成总的性能度量;和

[0047] 确定所述模型设置,以最小化所述总的性能度量的误差值。

[0048] 20)根据17)所述的非临时性计算机可读介质,其中,所述一个或更多个指令在由所述一个或更多个处理器执行时使所述一个或更多个处理器:

[0049] 接收测量;

[0050] 基于所述测量并使用所述第二光谱模型来执行光谱测定;和

[0051] 提供识别所述光谱测定的输出。

附图说明

[0052] 图1A-1E是本文中描述的示例实现方式的纵览的图示。

[0053] 图2是在其中可以实现本文中描述的系统和/或方法的示例环境的图示。

[0054] 图3是图2中的一个或更多个设备的示例部件的图示。

[0055] 图4-6是对于光谱模型的基于交叉验证的校准的示例过程的流程图。

具体实施方式

[0056] 示例实现方式的以下详细描述参考了附图。不同附图中的相同参考数字可以标识相同或类似的元素。以下描述使用光谱仪(spectrometer)作为示例,然而,本文描述的校准原理、过程和方法可以与任何传感器一起使用,包括但不限于其他光学传感器和光谱传感器。

[0057] 原料识别(RMID)是用于识别特定样本的组分(例如,成分)以用于识别、验证等的技术。例如,RMID可用于验证药物材料中的成分是否对应于标签上标识的成分集。类似地,

原料量化是用于对特定样本执行量化分析的技术，例如确定特定样本中特定材料的浓度。光谱仪可用于对样本(例如，药物材料)执行光谱法，以确定样本的组分、样本的组分的浓度等。光谱仪可确定对样本的一组测量，并可提供该一组测量用于光谱测定。光谱分类技术(例如，分类器)可以有助于基于对样本的该一组测量而确定样本的组分。

[0058] 为了执行光谱分类或量化，光谱模型可被用于评估对未知样本的一个或更多个测量。例如，控制设备可以尝试将对未知样本的一个或更多个测量分类为对应于光谱模型的特定类别、与光谱模型相关联的特定水平和/或量，等等。然而，随着时间的推移，原料可能会发生变化，这可能会导致光谱模型不准确。例如，对于被应用于农产品的光谱分类，与不同年份相关联的不同收获可能具有不同的光谱。因此，关于主数据集(例如，在初始时间对初始群体的初始一组光谱测量)训练的光谱模型在应用于目标数据集(例如，在后续时间对后续群体的后续一组光谱测量)时可能不准确。

[0059] 在另一种情况下，使用对于每个光谱仪的主数据集来针对每个光谱仪训练光谱模型可能是不切实际的。因此，控制设备可以在主数据集上训练单个光谱模型，并且部署该单个光谱模型以用于许多不同的光谱仪。然而，不同的光谱仪可以与不同的校准相关联和/或可以在不同的环境条件下操作。因此，当被应用于由第二光谱仪执行的光谱测量的目标数据集时，使用由第一光谱仪执行的光谱测量的主数据集训练的光谱模型可能不准确。

[0060] 本文所述的一些实现方式能够使用交叉验证技术对光谱模型进行校准更新和校准传递(calibration transfer)。例如，来自目标数据集的数据可以与来自主数据集的数据合并，以能够生成新的光谱模型。在这种情况下，来自主数据集的数据用于训练光谱模型的训练集，以及来自目标数据集的数据用于训练集和验证光谱模型的验证集两者。以这种方式，相对于用于模型生成和/或模型更新的其他技术，光谱模型的准确度得到提高。此外，基于提高传递的光谱模型的准确度，针对每个光谱仪获得主数据集的必要性被降低，从而降低了与部署光谱仪相关联的成本。

[0061] 图1A-1E是本文描述的示例实现方式100的图示。如图1A所示，示例实现方式100包括第一光谱仪102和第一控制设备104。

[0062] 又如图1A并由参考数字150所示，第一控制设备104可以向第一光谱仪102传输指令，以使第一光谱仪102对主群体152执行一组光谱测量。例如，第一控制设备104可以使第一光谱仪102针对要使用分类模型分类的每个类别、针对要使用量化模型量化的每个量等等对样本执行测量。分类模型的类别可以指共享一种或更多种共同特征的相似材料的分组，例如(在药物背景下的)乳糖材料、果糖材料、对乙酰氨基酚材料、布洛芬(ibuprophen)材料、阿司匹林材料等。用于训练分类模型的以及对于其要使用分类模型来执行原料识别的材料可以被称为感兴趣材料。

[0063] 又如图1A并且由参考数字154和156所示，第一光谱仪102可以执行一组光谱测量，并且可以将该一组光谱测量提供给第一控制设备104进行处理。例如，第一光谱仪102可以确定对于主群体152的每个样本的光谱，以使第一控制设备104能够生成一组类别，用于将未知样本分类为对于量化模型的感兴趣材料之一或者分类为具有与量化模型相关的特定量。

[0064] 又如图1A并且由参考数字158所示，第一控制设备104可以基于主数据集生成第一光谱模型。例如，第一控制设备104可以使用特定的测定技术并基于该一组光谱测量来生成

第一光谱模型。在一些实现方式中，第一控制设备104可以使用支持向量机(SVM)技术(例如，用于信息确定的机器学习技术)生成量化模型。附加地或替代地，第一控制设备104可以使用另一种类型的量化技术来生成量化模型。

[0065] 量化模型可以包括与将特定光谱分配给感兴趣材料的量的特定类别相关联的信息。在一些实现方式中，量化模型可以包括与识别与量的特定类别相关联的感兴趣材料的类型相关联的信息。以该方式，第一控制设备104可以基于将未知样本的光谱分配给量化模型的量的特定类别来提供识别未知样本的材料的量的信息作为光谱输出。

[0066] 如图1B并且由参考数字160所示，第二控制设备104'可以接收与第一光谱模型相关联的信息。例如，第二控制设备104'可以接收第一光谱模型、主数据集等。在一些实现方式中，相比于第一控制设备104，第二控制设备104'可以与不同的光谱仪相关联。例如，在校准传递情况下，第二控制设备104'可以与第二光谱仪102'(例如，目标光谱仪)结合使用，并且可以接收与第一光谱模型相关联的信息，以实现从第一光谱仪102(例如，主光谱仪)到第二光谱仪102'的校准传递。在这种情况下，如本文更详细描述的那样，第二控制设备104'和第二光谱仪102'可以执行对目标群体的测量，并生成第二光谱模型。可选地，如本文更详细描述的那样，在校准更新的情况下，第一控制设备104和第一光谱仪102可以执行对目标群体的测量并生成第二光谱模型，而不是将第一光谱模型传递到第二控制设备104'。

[0067] 又如图1B并且由参考数字162所示，第二控制设备104'可以向第二光谱仪102'传输指令，以使第二光谱仪102'执行对目标群体164的一组光谱测量。例如，第二控制设备104'可以使第二光谱仪102'基于接收到第一光谱模型来执行对目标群体164的光谱测量。在一些实现方式中，第二控制设备104'可以确定更新或校准第一光谱模型，并且可以触发第二光谱仪102'执行该一组光谱测量。在这种情况下，第二控制设备104'可以与第一控制设备104通信，以获得识别主数据集的信息，从而能够生成第二光谱模型。

[0068] 在一些实现方式中，目标群体164可以对应于主群体152。例如，目标群体164可以是与主群体152中包括的相同类别的另外的样本。在这种情况下，目标群体164可以与主群体152关于样本被收集或测量的时间、位置、环境条件等方面不同。此外，或者替代地，基于使用不同的光谱仪进行测量(例如，由第二光谱仪102'而不是由第一光谱仪102对主群体152进行测量)，目标群体164可以不同于主群体152。

[0069] 又如图1B并且由参考数字166和168所示，第二光谱仪102'可以执行一组光谱测量并且可以向第二控制设备104'提供识别该一组光谱测量的信息。例如，第二光谱仪102'可以执行对目标群体164的光谱测量，并且可以将识别光谱测量(例如，作为目标数据集)的信息提供给第二控制设备104'以进行处理。

[0070] 如图1C并且由参考数字170所示，第二控制设备104'可以确定总的性能度量。例如，第二控制设备104'可以基于将数据划分成多个子类(folds)、确定对于多个子类的多个性能度量、聚合多个性能度量以确定均方根误差(RMSE)值、以及优化偏最小二乘(PLS)因子(其可以被称为最优PLS因子)以最小化RMSE值，从而确定总的性能度量。子类可以指用于交叉验证的数据子组，其包括生成候选模型的训练集和评估候选模型在预测数据时的准确度的验证集。在另一个示例中，第二控制设备104'可以确定另一种类型的优化模型设置，例如与主成分回归(PCR)因子、支持向量回归(SVR)因子等相关的模型设置。在一些实现方式中，第二控制设备104'可以执行预处理优化。例如，第二控制设备104'可以确定优化的预处理

参数作为模型设置的一部分。

[0071] 在一些实现方式中，第二控制设备104’可以针对每个子类向训练集或验证集分配数据。例如，第二控制设备104’可以确定对于N个子类的多个训练集1至N以及对于N个子类的多个对应的验证集1至N。在一些实现方式中，训练集可以包括通过合并主数据集和目标数据集生成的合并数据。例如，训练集(例如，训练集1)可以包括来自主数据集(例如，MDS)的所有数据和来自目标数据集(例如，$TDS_{1,TS}$)的一部分数据。在这种情况下，对应的验证集可以包括来自目标数据集(例如，$TDS_{1、VS}$)的数据和不是来自主数据集的数据的对应部分。对应的验证集可以省略从与被包括在训练集中的相同物理样本的副本扫描中导出的数据。

[0072] 基于将数据分配到多个子类，第二控制设备104’可以确定对于每个子类的性能度量。例如，第二控制设备104’对于每个子类可以确定并可以聚合性能度量，以确定总的性能度量。例如，第二控制设备104’可以确定对于每个子类的PLS因子，并且可以确定对于每个子类的每个PLS因子的RMSE值。基于确定对于每个子类的每个PLS因子的RMSE值，第二控制设备104’可以确定总RMSE值。例如，第二控制设备104’可以根据所有子类的所有PLS因子来确定RMSE值。在这种情况下，基于确定总RMSE值，第二控制设备104’可以确定最优PLS因子，其可以是具有最低RMSE值的PLS因子。

[0073] 在这种情况下，基于在交叉验证期间将主数据集和目标数据集包括在N个子类的训练集中，但是仅将目标数据集包括在对应的验证集中，第二光谱模型的准确度相对于其他技术被提高。例如，相对于在不更新的情况下使用第一光谱模型、相对于仅使用目标数据集来确定PLS性能度量、相对于合并所有的目标数据集数据和所有的主数据集数据来生成合并的数据集以及使用将合并的数据集划分在训练集和验证集两者中等等，这种技术可以导致提高的准确度。

[0074] 如图1D并且由参考数字172所示，第二控制设备104’可以生成第二光谱模型。例如，第二控制设备104’可以使用主数据集(MDS)、目标数据集(TDS)、和最优PLS因子来生成第二光谱模型。以这种方式，第二控制设备104’可以实现校准的光谱模型、更新的光谱模型、传递的光谱模型等的生成。

[0075] 在一些实现方式中，第二控制设备104’可以合并主数据集和目标数据集以生成合并的数据集(例如，用于训练第二光谱模型的最终训练集)。例如，第二控制设备104’可以聚合主数据集和目标数据集以生成合并的数据集。基于生成合并的数据集，第二控制设备104’可以使用合并的数据集和最优PLS因子(例如，具有最低RMSE值)生成第二光谱模型。例如，第二控制设备104’，可以使用量化模型生成技术以结合合并的数据集(例如，其可以是对于第二光谱模型的训练集)和最优PLS因子来生成第二光谱模型。以这样的方式，通过在不使用合并的数据集的情况下确定最优PLS因子，然后将最优PLS因子与合并的数据集相组合，第二控制设备104’实现了比其他技术更准确的光谱模型。

[0076] 在一些实现方式中，第二控制设备104’可以基于生成第二光谱模型来提供第二光谱模型。例如，第二控制设备104’可以提供第二光谱模型，该第二光谱模型用于经由数据结构进行存储、用于部署在一个或更多个其他光谱仪上等等。附加地或替代地，第二控制设备104’可以基于生成第二光谱模型来提供与第二光谱模型相关的输出。例如，如本文更详细描述的，第二控制设备104’可以基于使用第二光谱模型分析未知样本来提供量化未知样本的信息。

[0077] 如图1E且由参考数字174所示，第二控制设备104’可以向第二光谱仪102’传输指令，以使第二光谱仪102’对未知样本176执行一组光谱测量。例如，第二控制设备104’可以使第二光谱仪102’在已经生成第二光谱模型之后对未知样本176执行光谱测量。

[0078] 又如图1E且由参考数字178和180所示，第二光谱仪102’可以执行一组光谱测量，并且可以向第二控制设备104’提供识别该一组光谱测量的信息。例如，第二光谱仪102’可以确定未知样本176的光谱，并且可以向第二控制设备104’提供识别光谱的信息，以用于分类和/或量化。

[0079] 又如图1E且由参考数字182所示，第二控制设备104’可以使用第二光谱模型来执行对该一组光谱测量的光谱分析。例如，第二控制设备104’可以使用第二光谱模型来确定未知样本176的分类和/或未知样本176的量化。在这种情况下，第二控制设备104’可以提供识别分类和/或量化的输出。以这样的方式，第二控制设备104’基于生成第二光谱模型来使用第二光谱模型。

[0080] 如上面指示的，图1A-1E仅作为一个或更多个示例被提供。其它示例可以不同于关于图1A-1E描述的示例。

[0081] 图2是在其中可以实现本文中所述的系统和/或方法的示例环境200的图示。如图2所示，环境200可以包括控制设备210、光谱仪220、网络230等。环境200中的设备可经由有线连接、无线连接或有线连接和无线连接的组合来互连。

[0082] 控制设备210包括能够存储、处理和/或路由与光谱分类相关联的信息的一个或更多个设备。例如，控制设备210可以包括服务器、计算机、可穿戴设备、云计算设备等，其基于训练集的一组测量生成光谱模型(如，分类模型或量化模型)、基于验证集的一组测量验证光谱模型和/或利用光谱模型来基于对未知样本的一组测量执行光谱分析。在一些实现方式中，控制设备210可以与特定光谱仪220相关联。在一些实现方式中，控制设备210可以与多个光谱仪220相关联。在一些实现方式中，控制设备210可以从环境200中的另一设备(例如，光谱仪220)接收信息和/或向环境200中的另一设备(例如光谱仪220)传输信息。

[0083] 光谱仪220包括能够对样本执行光谱测量的一个或更多个设备。例如，光谱仪220可以包括执行光谱法(例如，振动光谱法，诸如近红外(NIR)光谱法、中红外光谱法(mid-IR)、拉曼(Raman)光谱法等)的光谱设备。在一些实现方式中，光谱仪220可以被结合到可穿戴设备中，例如可穿戴光谱仪等。在一些实现方式中，光谱仪220可以从环境200中的另一设备(例如控制设备210)接收信息和/或向环境200中的另一设备(例如控制设备210)传输信息。

[0084] 网络230包括一个或更多个有线和/或无线网络。例如，网络230可包括蜂窝网络(例如，长期演进(LTE)网络、3G网络、码分多址(CDMA)网络等)、公共陆地移动网络(PLMN)、局域网(LAN)、广域网(WAN)、城域网(MAN)、电话网络(例如，公共交换电话网络(PSTN))、专用网络、自组织网络、内联网、互联网、基于光纤的网络、云计算网络等，和/或这些或其它类型的网络的组合。

[0085] 图2中显示的设备和网络的数量和布置作为一个或更多个示例被提供。实际上，与图2中显示的那些设备和/或网络相比，可以有另外的设备和/或网络、更少的设备和/或网络、不同的设备和/或网络或不同地布置的设备和/或网络。此外，图2中显示的两个或更多个设备可在单个设备内实现，或图2中显示的单个设备可被实现为多个分布式设备。另外或

可替代地，环境200中的一组设备(例如，一个或更多个设备)可以执行被描述为由环境200中的另一组设备执行的一个或更多个功能。

[0086] 图3是设备300的示例部件的图示。设备300可以对应于控制设备210和/或光谱仪220。在一些实现方式中，控制设备210和/或光谱仪220可以包括一个或更多个设备300和/或设备300的一个或更多个部件。如图3所示，设备300可包括总线310、处理器320、存储器330、储存部件340、输入部件350、输出部件360以及通信接口370。

[0087] 总线310包括允许在设备300的多个部件当中通信的部件。处理器320以硬件、固件、和/或硬件和软件的组合实现。处理器320是中央处理单元(CPU)、图形处理单元(GPU)、加速处理单元(APU)、微处理器、微控制器、数字信号处理器(DSP)、现场可编程门阵列(FPGA)、专用集成电路(ASIC)或另一类型的处理部件。在一些实现方式中，处理器320包括能够被编程以执行功能的一个或更多个处理器。存储器330包括随机存取存储器(RAM)、只读存储器(ROM)和/或存储信息和/或指令以用于由处理器320使用的另一类型的动态或静态储存设备(例如，闪存、磁存储器和/或光学存储器)。

[0088] 储存部件340存储与设备300的操作和使用相关的信息和/或软件。例如，储存部件340可以包括硬盘(例如，磁盘、光盘和/或磁光盘)、固态驱动器(SSD)、压缩光盘(CD)、数字通用盘(DVD)、软盘、盒式磁带(cartridge)、磁带(magnetic tape)和/或另一类型的非临时性计算机可读介质，连同对应的驱动器。

[0089] 输入部件350包括允许设备300例如经由用户输入端(例如，触摸屏显示器、键盘、小键盘(keypad)、鼠标、按钮、开关和/或麦克风)来接收信息的部件。附加地或可替换地，输入部件350可以包括用于确定位置的部件(例如，全球定位系统(GPS)部件)和/或传感器(例如，加速度计、陀螺仪、致动器、另一种类型的位置或环境传感器等)。输出部件360包括(经由例如显示器、扬声器、触觉反馈部件、声音或视觉指示器等)提供来自设备300的输出信息的部件。

[0090] 通信接口370包括使设备300能够例如经由有线连接、无线连接或有线和无线连接的组合与其它设备通信的类似收发机的部件(例如，收发机、单独的接收机、单独的发射机等)。通信接口370可以允许设备300从另一设备接收信息和/或将信息提供给另一设备。例如，通信接口370可包括以太网接口、光学接口、同轴接口、红外接口、射频(RF)接口、通用串行总线(USB)接口、Wi-Fi接口、蜂窝网络接口等。

[0091] 设备300可以执行本文中所述的一个或更多个过程。设备300可以基于处理器320执行由非临时性计算机可读介质(例如存储器330和/或储存部件340)存储的软件指令来执行这些过程。如本文所使用的，术语“计算机可读介质”指的是非临时性存储器设备。存储器设备包括在单个物理储存设备内的存储器空间或遍布于多个物理储存设备的存储器空间。

[0092] 软件指令可以经由通信接口370从另一计算机可读介质或从另一设备被读取到存储器330和/或储存部件340中。存储在存储器330和/或储存部件340中的软件指令当被执行时可使处理器320执行本文中所述的一个或更多个过程。另外或可替代地，硬件电路可代替软件指令或与软件指令组合来使用，以执行本文中所述的一个或更多个过程。因此，本文中所述的实现方式不限于硬件电路和软件的任何特定组合。

[0093] 图3中显示的部件的数量和布置作为示例被提供。实际上，与图3中显示的那些部件相比，设备300可以包括另外的部件、更少的部件、不同的部件或不同地布置的部件。另外

或可替代地，设备300的一组部件(例如，一个或更多个部件)可以执行被描述为由设备300的另一组部件执行的一个或更多个功能。

[0094] 图4是用于光谱模型的基于交叉验证的校准的示例过程400的流程图。在一些实现方式中，图4的一个或更多个过程块可由控制设备(例如控制设备210)执行。在一些实现方式中，图4的一个或更多个过程块可由与控制设备分离的或者包括控制设备的另一设备或一组设备(诸如光谱仪(如光谱仪220)等)执行。

[0095] 如图4中所示，过程400可以包括接收对于第一光谱模型的主数据集(块410)。例如，如上所述，控制设备(例如，使用处理器320、存储器330、储存部件340、输入部件350、输出部件360、通信接口370等)可以接收对于第一光谱模型的主数据集。

[0096] 如图4中进一步所示，过程400可以包括接收对于与第一光谱模型相关联的目标群体的目标数据集，以更新第一光谱模型(块420)。例如，如上所述，控制设备(例如，使用处理器320、存储器330、储存部件340、输入部件350、输出部件360、通信接口370等)可以接收对于与第一光谱模型相关联的目标群体的目标数据集，以更新第一光谱模型。

[0097] 如图4中进一步所示，过程400可以包括生成训练数据集，该训练数据集包括主数据集和来自目标数据集的第一数据(块430)。例如，如上所述，控制设备(例如，使用处理器320、存储器330、储存部件340、输入部件350、输出部件360、通信接口370等)可以生成包括主数据集和来自目标数据集的第一数据的训练数据集。

[0098] 如图4中进一步所示，过程400可以包括生成验证数据集，该验证数据集包括来自目标数据集而不是主数据集的第二数据(块440)。例如，如上所述，控制设备(例如，使用处理器320、存储器330、储存部件340、输入部件350、输出部件360、通信接口370等)可以生成包括来自目标数据集而不是主数据集的第二数据的验证数据集。

[0099] 如图4进一步所示，过程400可以包括使用交叉验证并使用训练数据集和验证数据集生成作为第一光谱模型的更新的第二光谱模型(块450)。例如，如上所述，控制设备(例如，使用处理器320、存储器330、储存部件340、输入部件350、输出部件360、通信接口370等)可以使用交叉验证并使用训练数据集和验证数据集生成作为第一光谱模型的更新的第二光谱模型。

[0100] 如在图4中进一步所示，过程400可以包括提供第二光谱模型(块460)。例如，如上所述，控制设备(例如，使用处理器320、存储器330、储存部件340、输入部件350、输出部件360、通信接口370等)可以提供第二光谱模型。

[0101] 过程400可以包括另外的实现方式，诸如以下描述的和/或结合本文其他地方描述的一个或更多个其他过程的任何单个实现方式或实现方式的任何组合。

[0102] 在第一实现方式中，过程400包括接收光谱测量；使用第二光谱模型执行光谱测定；以及提供识别光谱测定的输出。

[0103] 在第二实现方式中，单独或与第一实现方式相结合，训练数据集是多个训练数据集，且验证数据集是多个验证数据集，并且过程400包括基于多个训练数据集和多个验证数据集生成多个性能度量，基于多个性能度量确定总的性能度量，基于总的性能度量确定最优偏最小二乘(PLS)因子，并且基于最优PLS因子和合并的数据集确定第二光谱模型，其中合并的数据集包括主数据集和目标数据集。

[0104] 在第三实现方式中，单独或与第一实现方式和第二实现方式中的一个或更多个结

合，第一光谱模型和第二光谱模型是量化模型。

[0105] 在第四实现方式中，单独地或者与第一实现方式至第三实现方式中的一个或更多个相结合，主数据集基于由主光谱仪执行的第一组光谱测量，并且目标数据集基于由不同于主光谱仪的目标光谱仪执行的第二组光谱测量。

[0106] 在第五实现方式中，单独或与第一实现方式至第四实现方式中的一个或更多个相结合，主数据集基于由特定光谱仪执行的第一组光谱测量，而目标数据集基于由特定光谱仪执行的第二组光谱测量。

[0107] 虽然图4示出了过程400的示例块，但在一些实现方式中，与图4中描绘的那些块相比，过程400可包括附加的块、更少的块、不同的块或不同地布置的块。另外或可替代地，过程400的两个或更多个块可并行地被执行。

[0108] 图5是用于光谱模型的基于交叉验证的校准的示例过程500的流程图。在一些实现方式中，图5的一个或更多个过程块可由控制设备(如控制设备210)执行。在一些实现方式中，图5的一个或更多个过程块可由与控制设备210分离的或者包括控制设备的另一设备或一组设备(诸如光谱仪(如光谱仪220))执行。

[0109] 如图5所示，过程500可以包括接收对于与第一光谱模型相关联的目标群体的目标数据集(块510)。例如，如上所述，控制设备(例如，使用处理器320、存储器330、储存部件340、输入部件350、输出部件360、通信接口370等)可以接收对于与第一光谱模型相关联的目标群体的目标数据集。

[0110] 如图5中进一步所示，过程500可以包括基于接收目标数据集来获得对于第一光谱模型的主数据集(块520)。例如，如上所述，控制设备(例如，使用处理器320、存储器330、储存部件340、输入部件350、输出部件360、通信接口370等)可以基于接收目标数据集来获得对于第一光谱模型的主数据集。

[0111] 如图5中进一步所示，过程500可以包括使用交叉验证确定最优偏最小二乘(PLS)因子，其中最优PLS因子基于多个训练数据集且基于多个验证数据集来确定，每个训练数据集包括目标数据集的相应部分和主数据集的全部，并且每个验证数据集包括目标数据集的相应部分并且不包括主数据集的数据(块530)。例如，如上所述，控制设备(例如，使用处理器320、存储器330、储存部件340、输入部件350、输出部件360、通信接口370等)可以使用交叉验证确定最优偏最小二乘(PLS)因子。在一些方面，最优PLS因子基于多个训练数据集且基于多个验证数据集来确定，每个训练数据集包括目标数据集的相应部分和主数据集的全部，每个验证数据集包括目标数据集的相应部分并且不包括主数据集的数据。

[0112] 如图5中进一步所示，过程500可以包括合并目标数据集和主数据集以生成合并的数据集(块540)。例如，如上所述，控制设备(例如，使用处理器320、存储器330、储存部件340、输入部件350、输出部件360、通信接口370等)可以合并目标数据集和主数据集以生成合并的数据集。

[0113] 如图5中进一步所示，过程500可以包括使用合并的数据集和最优PLS因子来生成第二光谱模型，其中第二光谱模型是第一光谱模型的更新(块550)。例如，如上所述，控制设备(例如，使用处理器320、存储器330、储存部件340、输入部件350、输出部件360、通信接口370等)可以使用合并的数据集和最优PLS因子来生成第二光谱模型。在一些方面，第二光谱模型是第一光谱模型的更新。

[0114] 如在图5中进一步所示，过程500可以包括提供第二光谱模型以代替第一光谱模型(块560)。例如，如上所述，控制设备(例如，使用处理器320、存储器330、储存部件340、输入部件350、输出部件360、通信接口370等)可以提供第二光谱模型以代替第一光谱模型。

[0115] 过程500可以包括另外的实现方式，诸如以下描述的和/或结合本文其他地方描述的一个或更多个其他过程的任何单个实现方式或实现方式的任何组合。

[0116] 在第一实现方式中，确定最优PLS因子包括确定对于多个训练数据集的每一个训练数据集和多个验证数据集的每一个验证数据集的偏最小二乘(PLS)性能度量；基于PLS性能度量确定总的PLS性能度量；以及基于总的PLS性能度量来优化用于第二光谱模型的PLS因子。

[0117] 在第二实现方式中，单独地或与第一实现方式结合，总的PLS性能度量与均方根误差(RMSE)值相关联，并且优化PLS因子包括优化PLS因子以最小化RMSE值。

[0118] 在第三实现方式中，单独地或与第一实现方式和第二实现方式中的一个或更多个结合，多个验证数据集包括与多个训练数据集相比的目标数据集的不同数据。

[0119] 在第四实现方式中，单独地或与第一实现方式至第三实现方式中的一个或更多个结合，确定总的PLS性能度量包括聚合PLS性能度量。

[0120] 在第五实现方式中，单独地或与第一实现方式至第四实现方式中的一个或更多个结合，目标数据集与在与主数据集相关联的测量之后执行的对目标群体的一组测量相关联。

[0121] 在第六实现方式中，单独或与第一实现方式至第五实现方式中的一个或更多个结合，第二光谱模型是第一光谱模型的校准更新模型。

[0122] 在第七实现方式中，单独地或者与第一实现方式至第六实现方式中的一个或更多个相结合，目标数据集与由特定光谱仪执行的一组测量相关联，该特定光谱仪不同于执行与主数据集相关联的测量的一个或更多个光谱仪。

[0123] 在第八实现方式中，单独或与第一实现方式至第七实现方式中的一个或更多个结合，第二光谱模型是第一光谱模型的校准传递模型。

[0124] 在第九实现方式中，单独地或与第一实现方式至第八实现方式中的一个或更多个相结合，提供第二光谱模型包括提供第二光谱模型以结合由特定光谱仪进行的后续测量来进行使用。

[0125] 虽然图5示出了过程500的示例块，但在一些实现方式中，与图5中描绘的那些块相比，过程500可包括另外的块、更少的块、不同的块或不同地布置的块。另外或可替代地，过程500的两个或更多个块可并行地被执行。

[0126] 图6是用于光谱模型的基于交叉验证的校准的示例过程600的流程图。在一些实现方式中，图6的一个或更多个过程块可由控制设备(例如，控制设备210)执行。在一些实现方式中，图6的一个或更多个过程块可由与控制设备分离的或者包括控制设备的另一设备或一组设备(诸如光谱仪(如光谱仪220)等)执行。

[0127] 如图6所示，过程600可以包括接收对于第一光谱模型的主数据集，接收对于与第一光谱模型相关联的目标群体的目标数据集以更新第一光谱模型，基于主数据集和目标数据集生成多个训练数据集，基于目标数据集生成多个验证数据集，并且其中多个验证数据集不包括主数据集的数据。例如，如上所述，控制设备(例如，使用处理器320、存储器330、储

存部件340、输入部件350、输出部件360、通信接口370等)可以接收对于第一光谱模型的主数据集,接收对于与第一光谱模型相关联的目标群体的目标数据集以更新第一光谱模型,基于主数据集和目标数据集生成多个训练数据集,基于目标数据集生成多个验证数据集,并且其中,多个验证数据集不包括主数据集的数据。在一些方面,多个验证数据集不包括主数据集的数据。

[0128] 如图6中所示,过程600可以包括接收对于第一光谱模型的主数据集(块610)。例如,如上所述,控制设备(例如,使用处理器320、存储器330、储存部件340、输入部件350、输出部件360、通信接口370等)可以接收对于第一光谱模型的主数据集。

[0129] 如图6所示,过程600可以包括接收对于与第一光谱模型相关联的目标群体的目标数据集以更新第一光谱模型(块620)。例如,如上所述,控制设备(例如,使用处理器320、存储器330、储存部件340、输入部件350、输出部件360、通信接口370等)可以接收对于与第一光谱模型相关联的目标群体的目标数据集以更新第一光谱模型。

[0130] 如图6所示,过程600可包括基于主数据集和目标数据集生成多个训练数据集(块630)。例如,如上所述,控制设备(例如,使用处理器320、存储器330、储存部件340、输入部件350、输出部件360、通信接口370等)可以基于主数据集和目标数据集生成多个训练数据集。

[0131] 如图6所示,过程600可以包括基于目标数据集生成多个验证数据集,并且其中,多个验证数据集不包括主数据集的数据(块640)。例如,如上所述,控制设备(例如,使用处理器320、存储器330、储存部件340、输入部件350、输出部件360、通信接口370等)可以基于目标数据集生成多个验证数据集。在一些方面,多个验证数据集不包括主数据集的数据。

[0132] 如图6中进一步所示,过程600可以包括基于多个训练数据集和多个验证数据集并使用交叉验证来确定模型设置(块650)。例如,如上所述,控制设备(例如,使用处理器320、存储器330、储存部件340、输入部件350、输出部件360、通信接口370等)可以基于多个训练数据集和多个验证数据集来确定模型设置。

[0133] 如图6进一步所示,过程600可以包括基于模型设置、目标数据集和主数据集生成第二光谱模型(块660)。例如,如上所述,控制设备(例如,使用处理器320、存储器330、储存部件340、输入部件350、输出部件360、通信接口370等)可以基于模型设置、目标数据集和主数据集生成第二光谱模型。

[0134] 如在图6中进一步所示,过程600可以包括提供第二光谱模型(块670)。例如,如上所述,控制设备(例如,使用处理器320、存储器330、储存部件340、输入部件350、输出部件360、通信接口370等)可以提供第二光谱模型。

[0135] 过程600可以包括另外的实现方式,诸如以下描述的和/或结合本文其他地方描述的一个或更多个其他过程的任何单个实现方式或实现方式的任何组合。

[0136] 在第一实现方式中,模型设置是偏最小二乘(PLS)模型的PLS因子、主成分回归(PCR)模型的成分的量、支持向量回归(SVR)模型的SVR参数或预处理设置中的至少一个。

[0137] 在第二实现方式中,单独地或与第一实现方式结合,过程600包括生成关于多个训练数据集中的每一个训练数据集和多个验证数据集的对应的验证数据集的多个部分性能度量(a plurality of partial performance metrics),聚合该多个部分性能度量以生成总的性能度量,以及确定模型设置以最小化总的性能度量的误差值。

[0138] 在第三实现方式中,单独地或与第一实现方式和第二实现方式中的一个或更多个

结合，过程600包括基于测量和使用第二光谱模型来执行光谱测定，并提供识别光谱测定的输出。

[0139] 虽然图6示出了过程600的示例块，但在一些实现方式中，与图6中描绘的那些块相比，过程600可包括另外的块、更少的块、不同的块或不同地布置的块。另外或可替代地，过程600的两个或更多个块可并行地被执行。

[0140] 前述公开提供了说明和描述，但并不旨在穷举或将实现方式限制到所公开的精确形式。修改和变型可以根据以上公开内容进行，或者可以从实现方式的实践中获得。

[0141] 如在本文使用的，术语“部件”被规定为广泛地被解释为硬件、固件、和/或硬件和软件的组合。

[0142] 如本文所使用的，满足阈值可以根据上下文指大于阈值、多于阈值、高于阈值、大于或等于阈值、小于阈值、少于阈值、低于阈值、小于或等于阈值、等于阈值等的值。

[0143] 将明显的是，本文中所述的系统和/方法可在硬件、固件、或硬件和软件的组合的不同形式中实现。用于实现这些系统和/或方法的实际专用控制硬件或软件代码不是实现方式的限制。因此，在本文中描述系统和/或方法的操作和行为而不参考特定的软件代码，应理解的是，软件和硬件可以被设计成基于本文中的描述来实现系统和/或方法。

[0144] 尽管在权利要求中陈述和/或在说明书中公开了特征的特定组合，但是这些组合并不旨在限制各种实现方式的公开内容。事实上，这些特征中的许多可以以权利要求中未具体陈述和/或说明书中未公开的方式组合。尽管以下所列出的每个从属权利要求可以直接从属于仅仅一个权利要求，但是各种实现方式的公开内容包括每个从属权利要求与权利要求组中的每个其他权利要求相结合。

[0145] 本文中使用的任何元素、动作或指令都不应当被解释为关键或必要的，除非明确这样描述。另外，如本文中所使用的，冠词“一(a)”和“一(an)”旨在包括一个或更多个项，并且可以与“一个或更多个(one or more)”互换使用。此外，如本文所用，冠词“该(the)”旨在包括与冠词“该”相结合的一个或更多个项，并且可以与“该一个或更多个”互换使用。此外，如本文所使用的，术语“集(set)”旨在包括一个或更多个项(例如，相关项、不相关项、相关项和不相关项的组合等)，并且可以与“一个或更多个”互换使用。在意指仅一个项的情况下，使用术语“仅一个(only one)”或类似的语言。此外，如本文所用，术语“具有(has)”、“具有(have)”、“具有(having)”等旨在是开放式的术语。此外，除非另有明确地说明，否则短语“基于”旨在表示“至少部分地基于”。此外，如本文所使用的，术语“或”在串联使用时旨在包括在内，并且可以与“和/或”互换使用，除非另有明确说明(例如，与“任一”或“仅其中之一”结合使用的情况)。

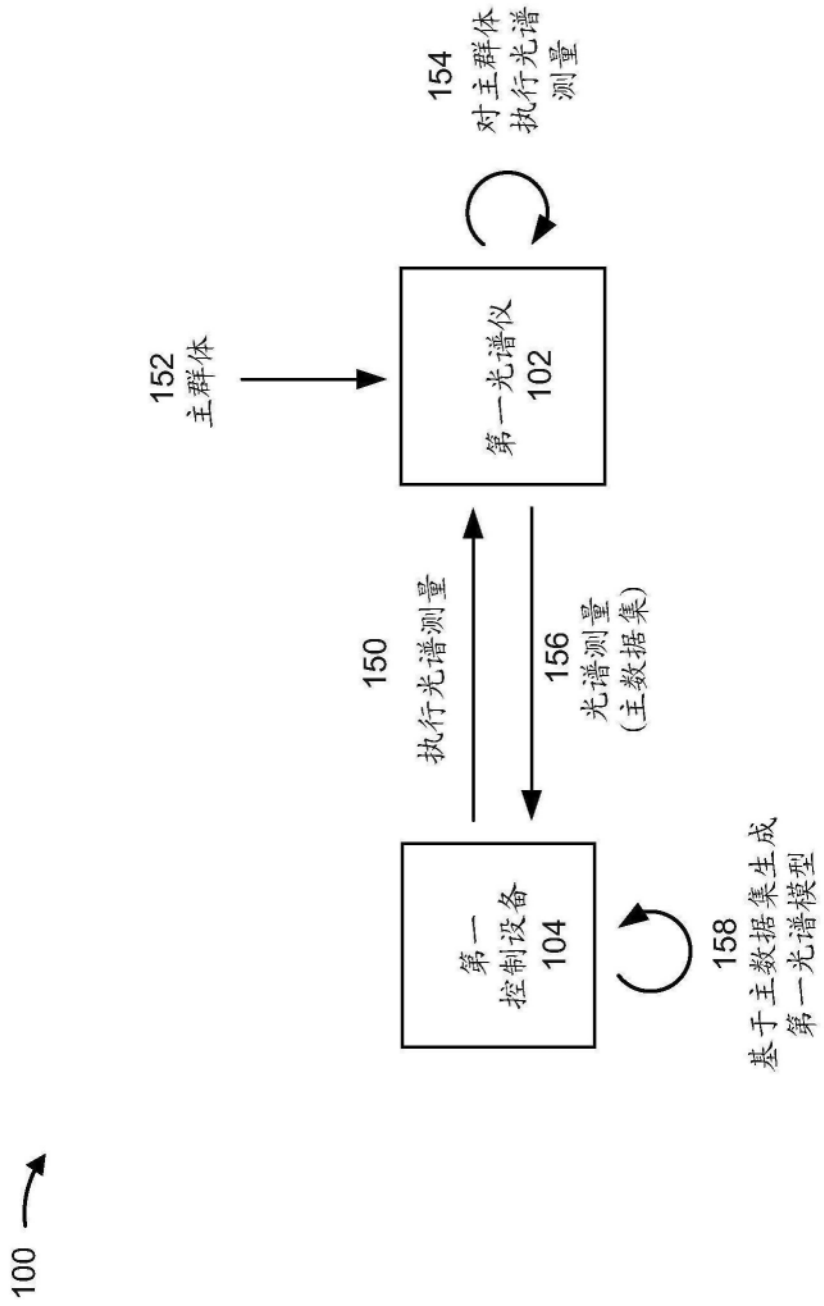

100
第一
控制设备
104
第一光谱仪
102
150
执行光谱测量
156
光谱测量
(主数据集)
152
主群体
154
对主群体
执行光谱
测量
158
基于主数据集生成
第一光谱模型

图1A

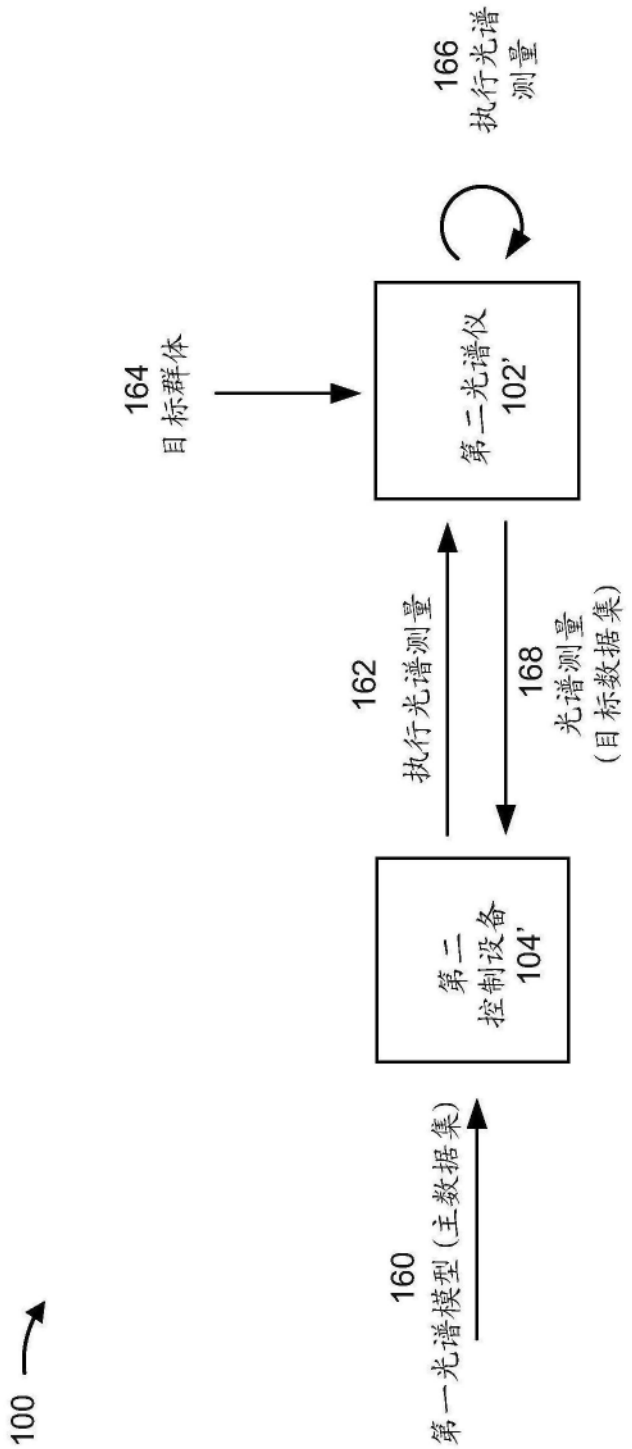

100
160
第一光谱模型(主数据集)
第二
控制设备
104'
162
执行光谱测量
168
光谱测量
(目标数据集)
第二光谱仪
102'
164
目标群体
166
执行光谱
测量

图1B

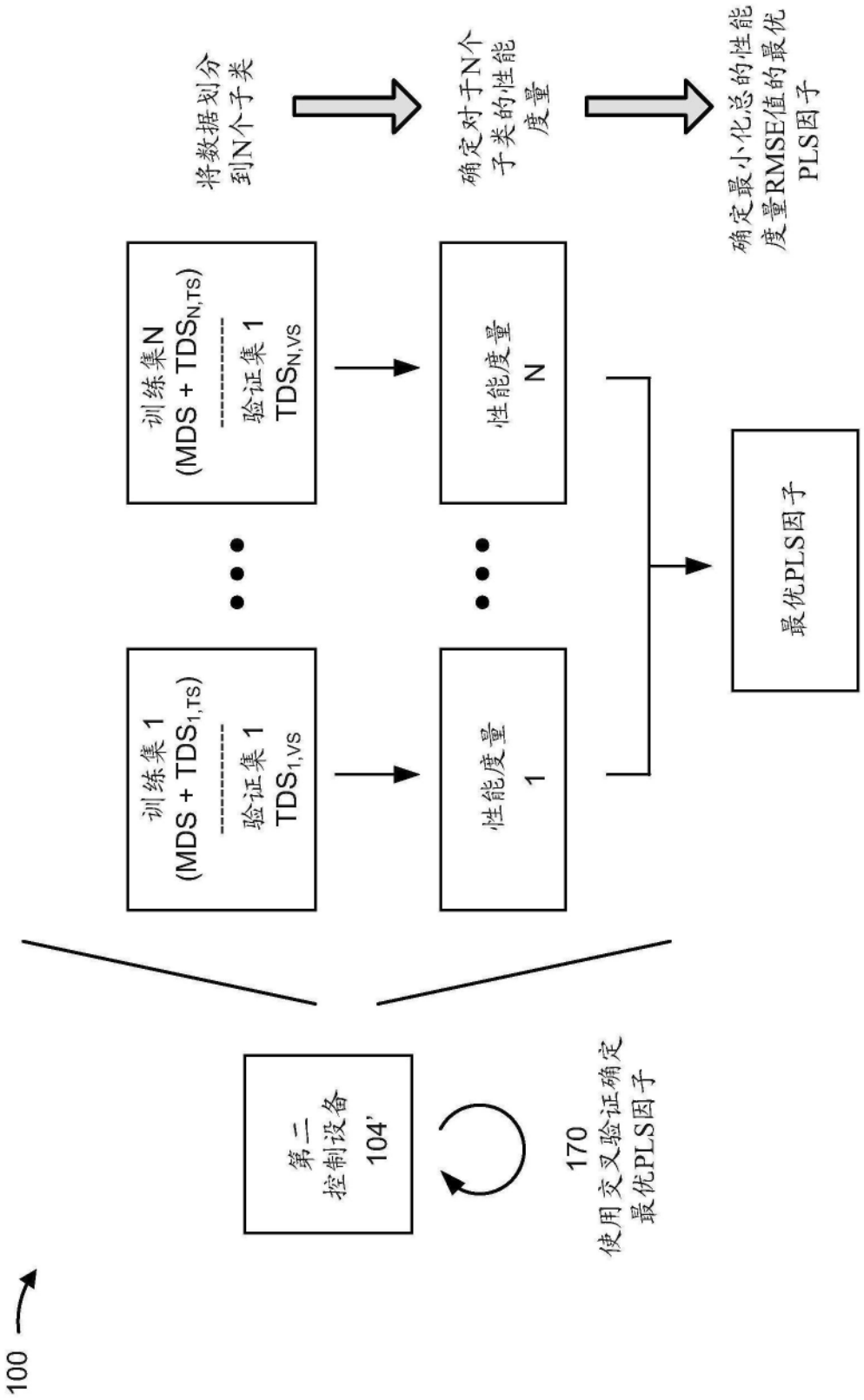

100
第二
控制设备
104'
170
使用交叉验证确定
最优PLS因子
训练集 1
(MDS + TDS1,TS)
验证集 1
TDS1,VS
训练集N
(MDS + TDSN,TS)
验证集 1
TDSN,VS
性能度量
1
性能度量
N
最优PLS因子
将数据划分
到N个子类
确定对于N个
子类的性能
度量
确定最小化总的性能
度量RMSE值的最优
PLS因子

图1C

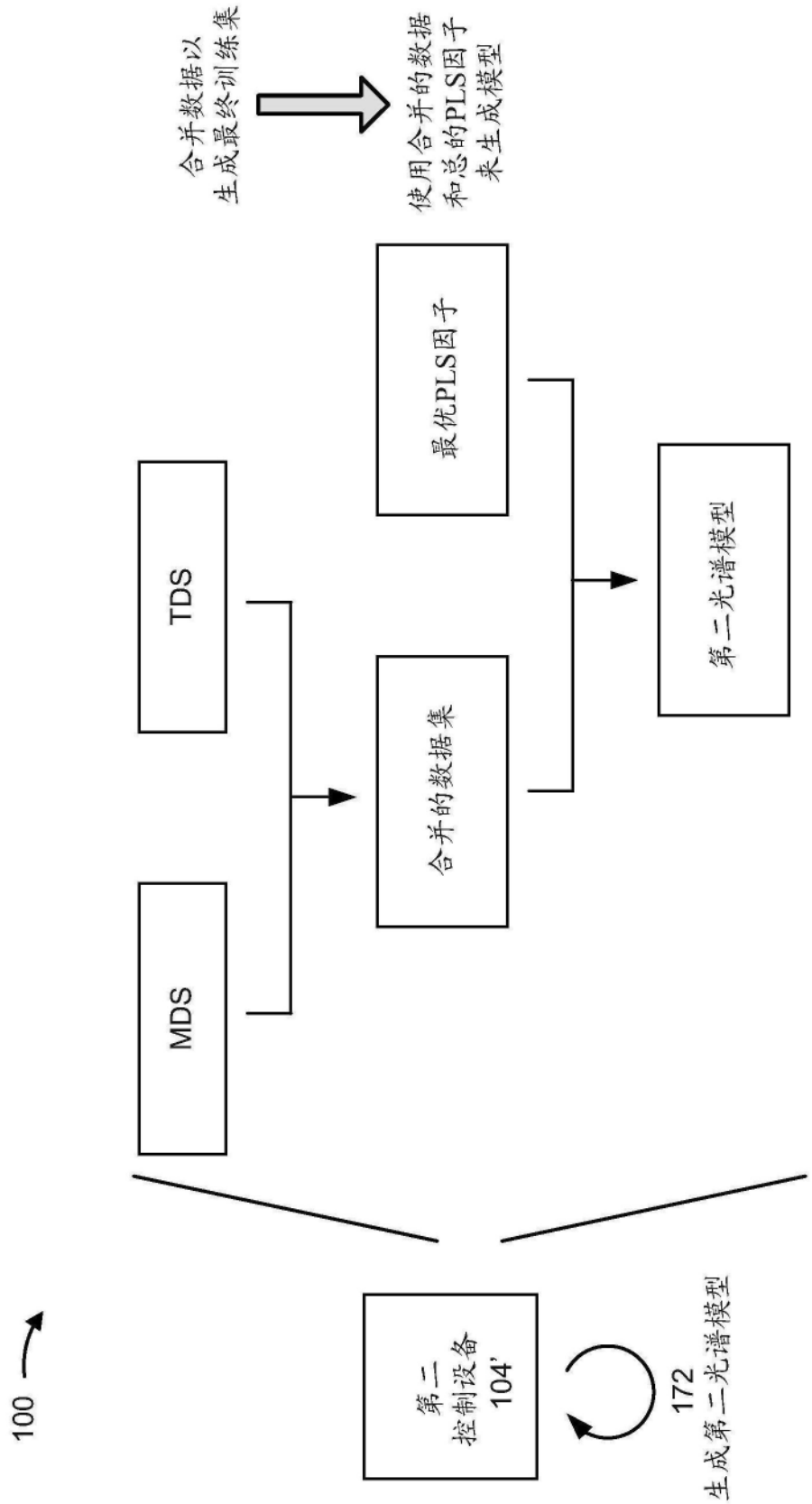

100
第二
控制设备
104'
172
生成第二光谱模型
MDS
TDS
合并的数据集
最优PLS因子
第二光谱模型
合并数据以
生成最终训练集
使用合并的数据
和总的PLS因子
来生成模型

图1D

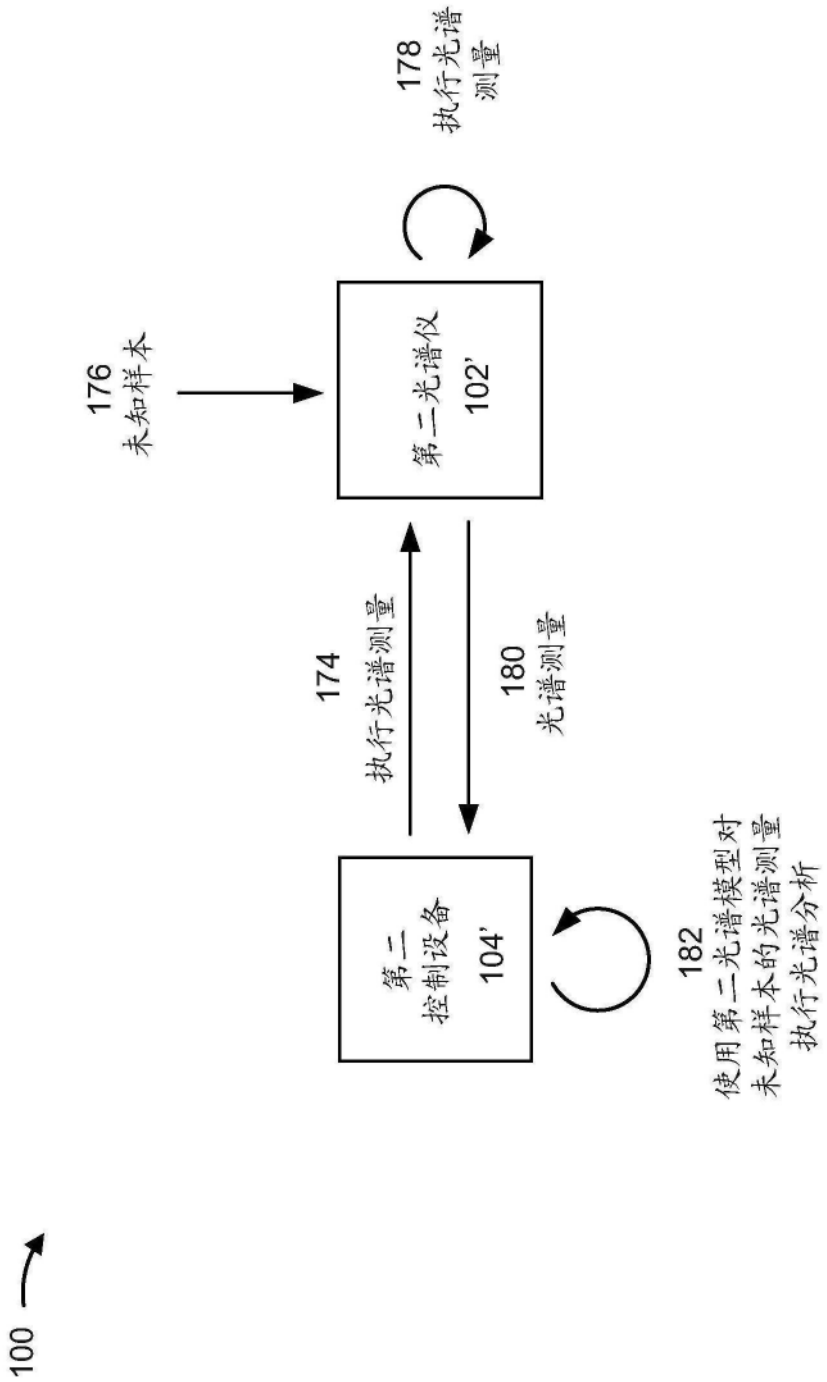

178
执行光谱
测量
176
未知样本
第二光谱仪
102'
174
执行光谱测量
180
光谱测量
第二
控制设备
104'
182
使用第二光谱模型对
未知样本的光谱测量
执行光谱分析
100

图1E

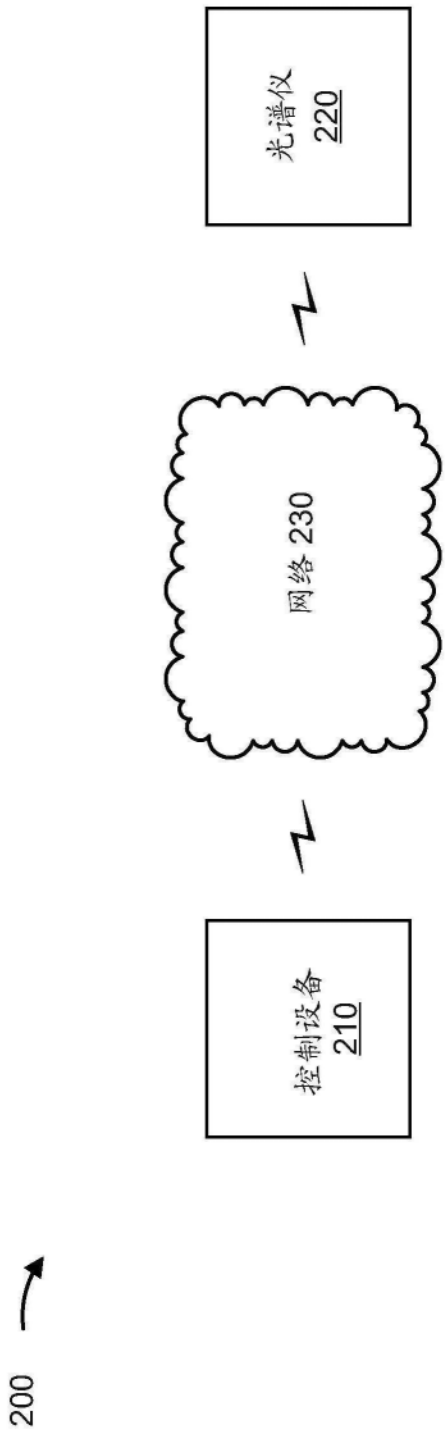


图2

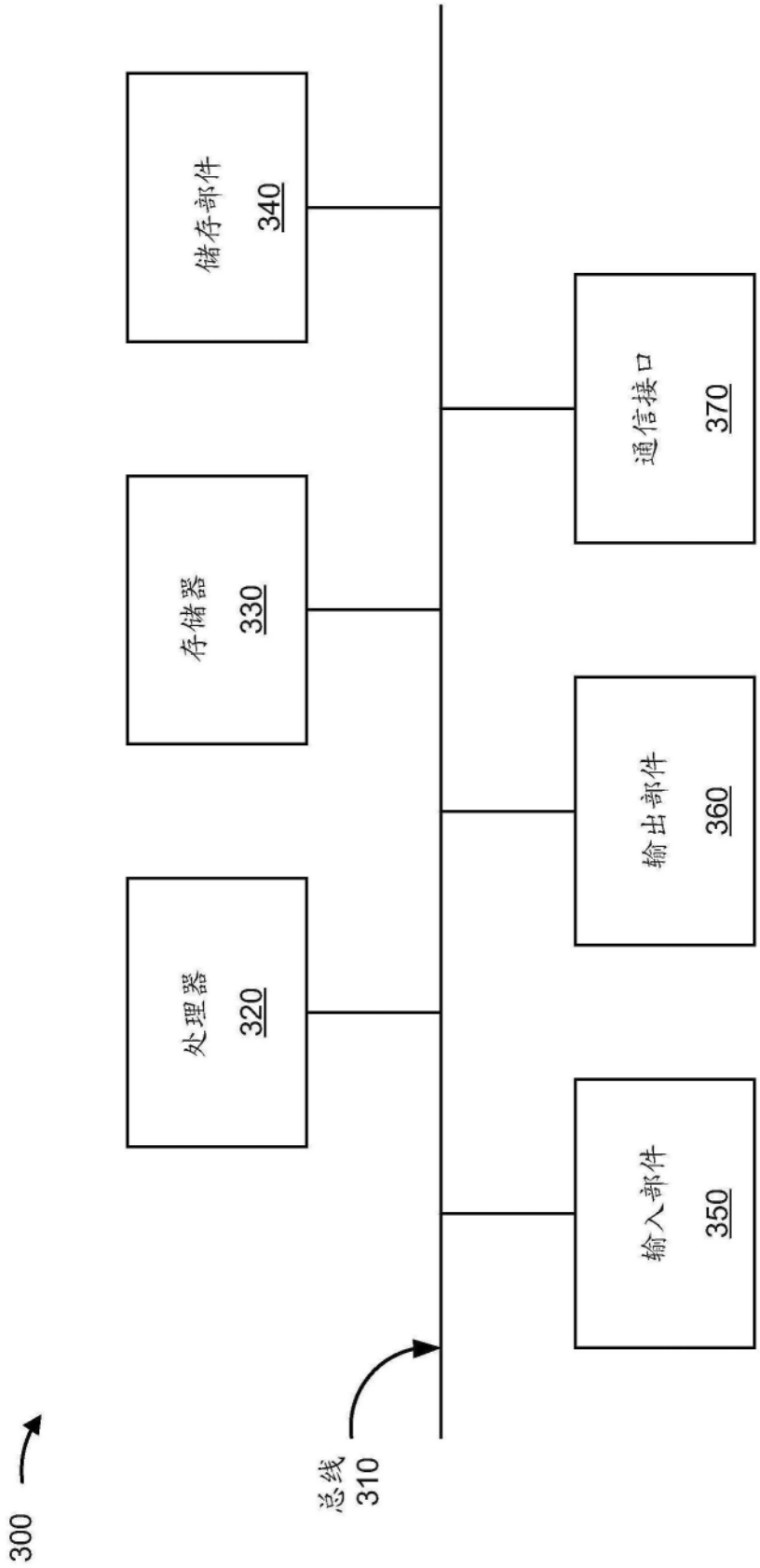


图3

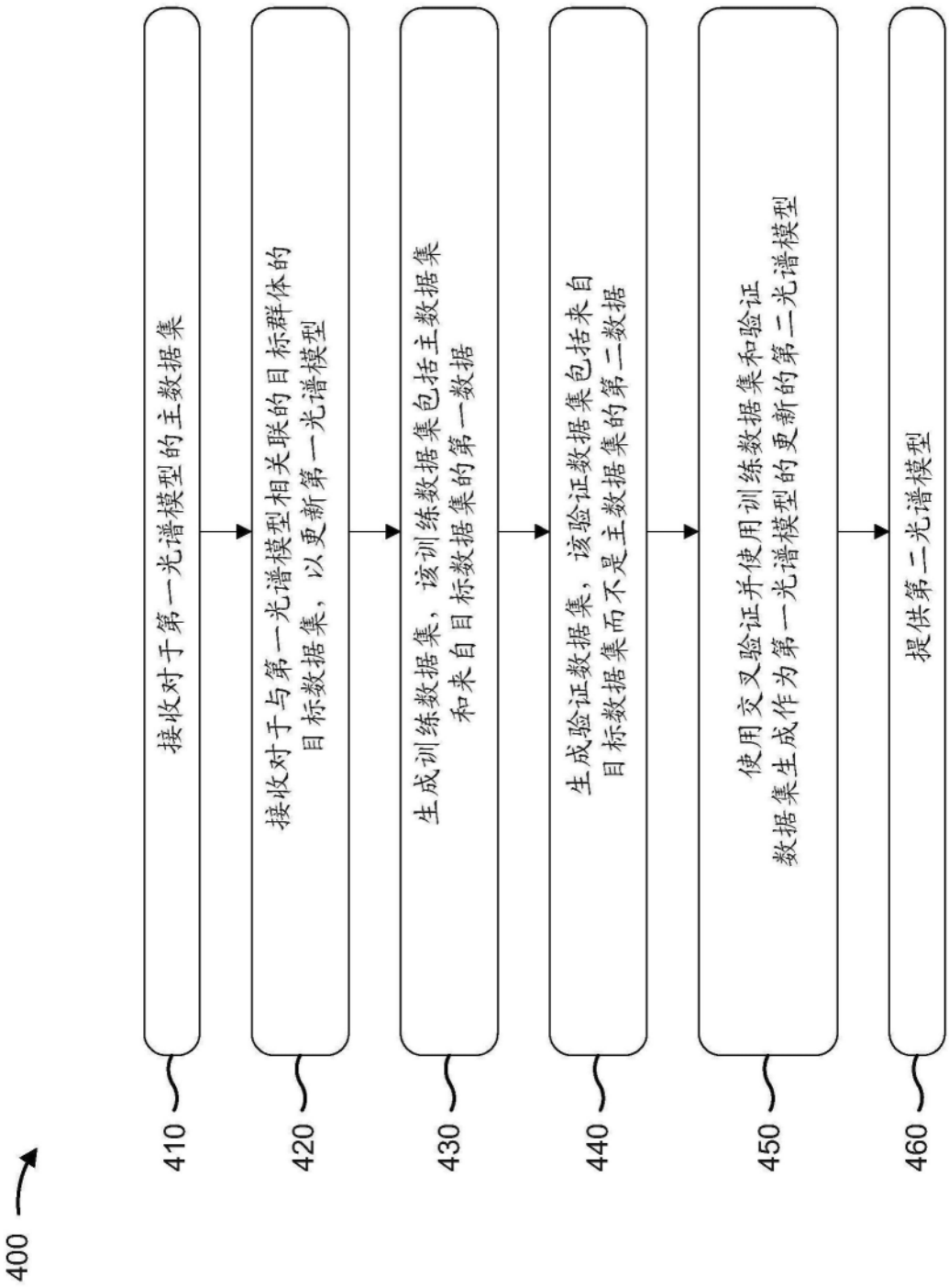

400
410
接收对于第一光谱模型的主数据集
420
接收对于与第一光谱模型相关联的目标群体的目标数据集，以更新第一光谱模型
430
生成训练数据集，该训练数据集包括主数据集和来自目标数据集的第一数据
440
生成验证数据集，该验证数据集包括来自目标数据集而不是主数据集的第二数据
450
使用交叉验证并使用训练数据集和验证数据集生成作为第一光谱模型的更新的第二光谱模型
460
提供第二光谱模型

图4

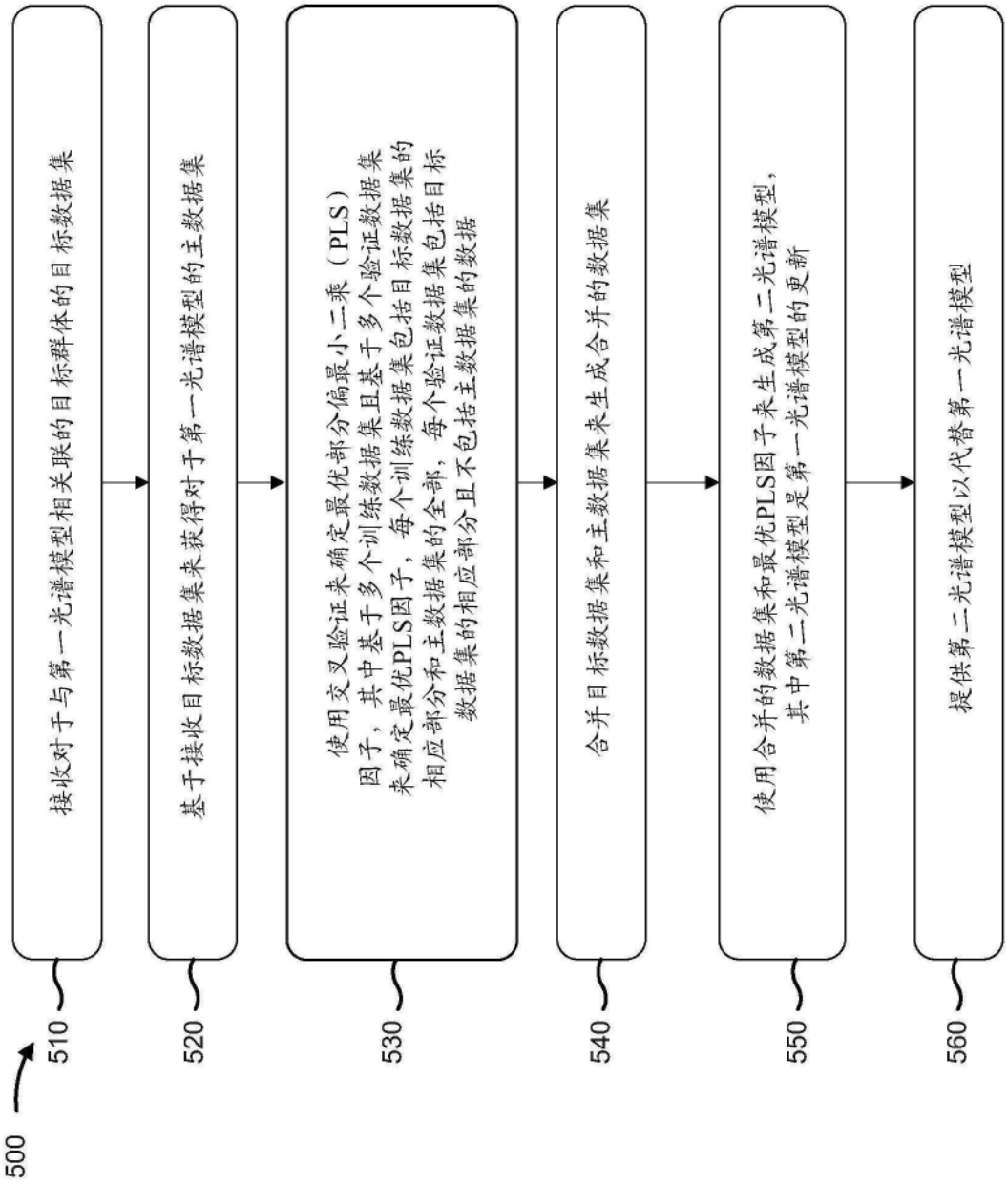

500
510
接收对于与第一光谱模型相关联的目标群体的目标数据集
520
基于接收目标数据集来获得对于第一光谱模型的主数据集
530
使用交叉验证来确定最优部分偏最小二乘（PLS）因子，其中基于多个训练数据集且基于多个验证数据集来确定最优PLS因子，每个训练数据集包括目标数据集的相应部分和主数据集的全部，每个验证数据集包括目标数据集的相应部分且不包括主数据集的数据
540
合并目标数据集和主数据集来生成合并的数据集
550
使用合并的数据集和最优PLS因子来生成第二光谱模型，其中第二光谱模型是第一光谱模型的更新
560
提供第二光谱模型以代替第一光谱模型

图5

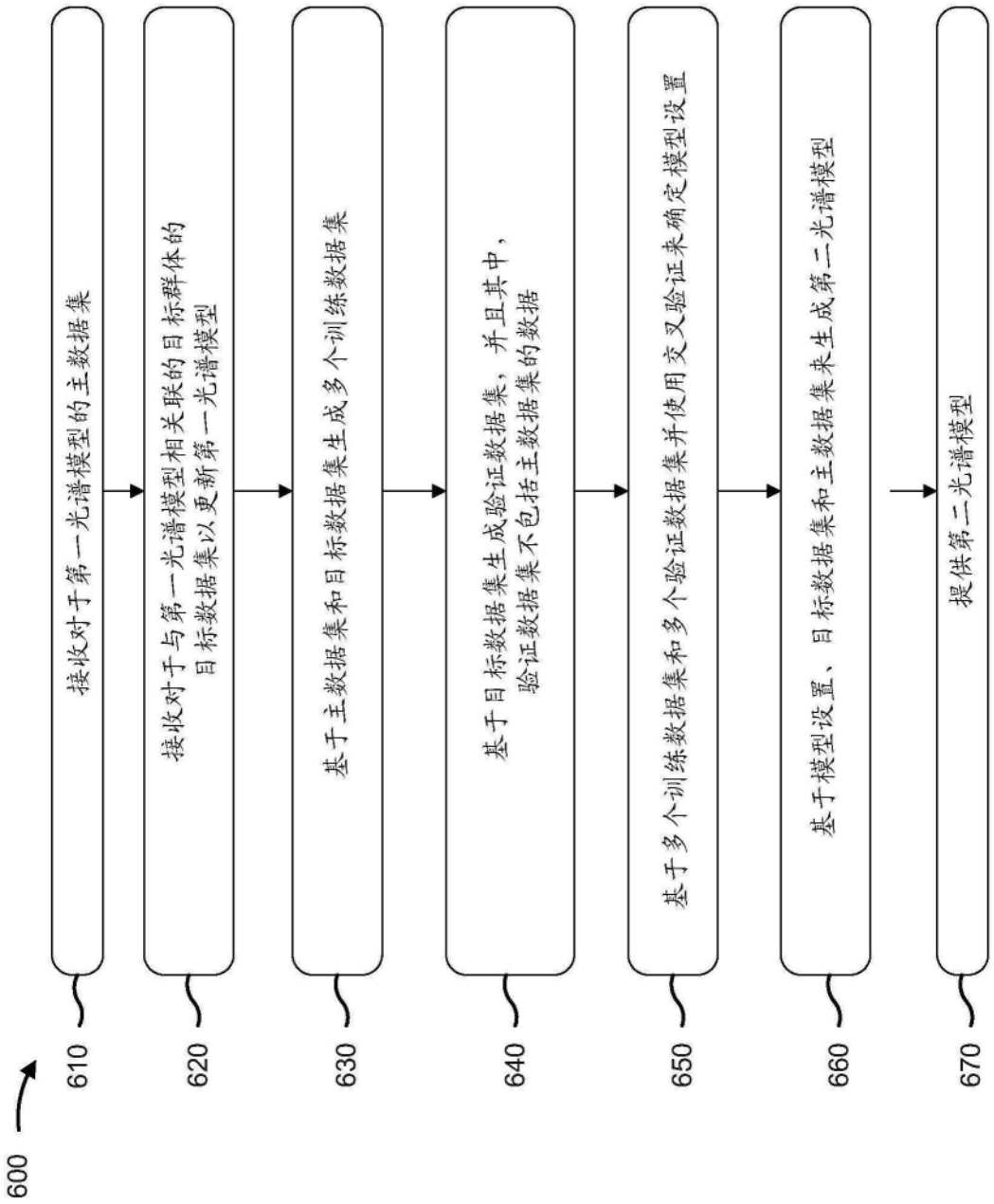

600
610
接收对于第一光谱模型的主数据集
620
接收对于与第一光谱模型相关联的目标群体的目标数据集以更新第一光谱模型
630
基于主数据集和目标数据集生成多个训练数据集
640
基于目标数据集生成验证数据集，并且其中，验证数据集不包括主数据集的数据
650
基于多个训练数据集和多个验证数据集并使用交叉验证来确定模型设置
660
基于模型设置、目标数据集和主数据集来生成第二光谱模型
670
提供第二光谱模型

图6