I783159

(19)中華民國智慧財產局

(12)發明說明書公告本 **(11)證書號數：TW I783159 B**

(45)公告日： 中華民國 111 (2022) 年 11 月 11 日

(21)申請案號：108122370 (22)申請日： 中華民國 108 (2019) 年 06 月 26 日

(51)Int. Cl.： *G01J3/02 (2006.01)* *G16C20/30 (2019.01)*
G16C60/00 (2019.01)

(30)優先權：2018/06/29 美國 62/692,248
2019/06/21 美國 16/448,914

(71)申請人：美商菲爾薇解析公司 (美國) VIAVI SOLUTIONS INC. (US)
美國

(72)發明人：熊 章民 HSIUNG, CHANG MENG (US)；孫嵐 SUN, LAN (CN)

(74)代理人：閻啓泰；林景郁

(56)參考文獻：

KR 10-2005-0078924A US 8914312B2

期刊 Alister J.Tencate,John H.Kalivas,Alexander J.White Fusion strategies for selecting multiple tuning parameters for multivariate calibration and other penalty based processes: A model updating application for pharmaceutical analysis Analytica Chimica Acta 921 ELSEVIER 19 May 2016 28~37

期刊 Erik Andries Penalized eigendecompositions: motivations from domain adaptation for calibration transfer journal of chemometrics 31 wiley April 2017 e2818

審查人員：吳耿榮

申請專利範圍項數：20 項 圖式數：6 共 45 頁

(54)名稱

對光譜模型的基於交叉驗證的校準

(57)摘要

本公開涉及對光譜模型的基於交叉驗證的校準。一種裝置可以接收對於第一光譜模型的主數據集；接收對於與第一光譜模型相關聯的目標群體的目標數據集以更新第一光譜模型；產生包括主數據集和來自目標數據集中的第一數據的訓練數據集；產生包括來自目標數據集的第二數據而不是主數據集的驗證數據集；使用交叉驗證並使用訓練數據集和驗證數據集產生第二光譜模型，該第二光譜模型為第一光譜模型的更新；以及提供第二光譜模型。

The present application relates to cross-validation based calibration of a spectroscopic model. A device may receive a master data set for a first spectroscopic model; receive a target data set for a target population associated with the first spectroscopic model to update the first spectroscopic model; generate a training data set that includes the master data set and first data from the target data set; generate a validation data set that includes second data from the target data set and not the master data set; generate, using cross-validation and using the training data set and the validation data set, a second spectroscopic model that is an update of the first spectroscopic model; and provide the second spectroscopic model.

指定代表圖：

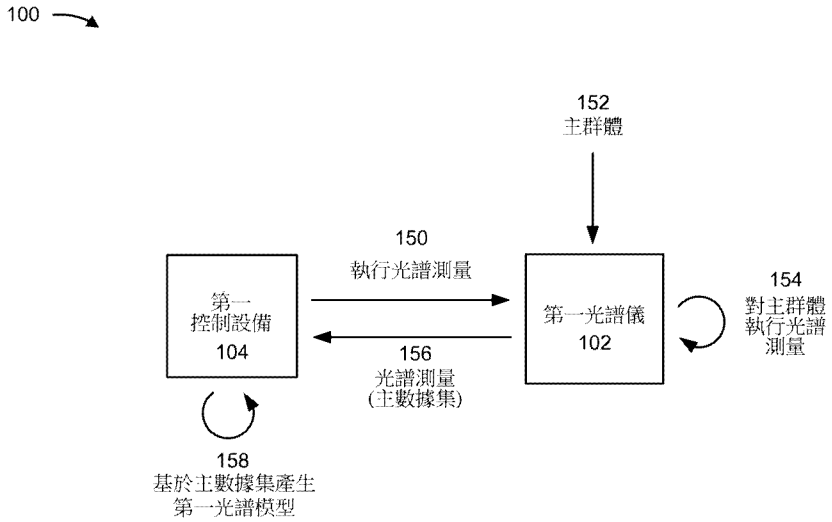


圖1A

符號簡單說明：

100:方式

102:光譜儀

104:控制裝置

150:執行光譜測量

152:主群體

154:對主群體執行光譜測量

156:光譜測量(主數據集)

158:基於主數據集產生第一光譜模型

專利案號：108122370

公告本

I783159

【發明摘要】

【中文發明名稱】 對光譜模型的基於交叉驗證的校準

【英文發明名稱】 CROSS-VALIDATION BASED CALIBRATION OF A SPECTROSCOPIC MODEL

【中文】

本公開涉及對光譜模型的基於交叉驗證的校準。一種裝置可以接收對於第一光譜模型的主數據集；接收對於與第一光譜模型相關聯的目標群體的目標數據集以更新第一光譜模型；產生包括主數據集和來自目標數據集中的第一數據的訓練數據集；產生包括來自目標數據集的第二數據而不是主數據集的驗證數據集；使用交叉驗證並使用訓練數據集和驗證數據集產生第二光譜模型，該第二光譜模型為第一光譜模型的更新；以及提供第二光譜模型。

【英文】

The present application relates to cross-validation based calibration of a spectroscopic model. A device may receive a master data set for a first spectroscopic model; receive a target data set for a target population associated with the first spectroscopic model to update the first spectroscopic model; generate a training data set that includes the master data set and first data from the target data set; generate a validation data set that includes second data from the target data set and not the master data set; generate, using cross-validation and using the training data set and the validation data set, a second spectroscopic model that is an update of the first spectroscopic model; and provide the second spectroscopic model.

【指定代表圖】　圖1A

【代表圖之符號簡單說明】

100　方式

102　光譜儀

104　控制裝置

150　執行光譜測量

152　主群體

154　對主群體執行光譜測量

156　光譜測量（主數據集）

158　基於主數據集產生第一光譜模型

【特徵化學式】

無

111 年 07 月 13 日修正替換頁

【發明說明書】

【中文發明名稱】 對光譜模型的基於交叉驗證的校準

【英文發明名稱】 CROSS-VALIDATION BASED CALIBRATION OF A SPECTROSCOPIC MODEL

【技術領域】

【0001】 本申請涉及對光譜模型的基於交叉驗證的校準。

【0002】 相關申請

【0003】 本申請基於美國法典第35條第119款主張於2018年6月29日提交的且標題為“UPDATING CALIBRATION MODELS BASED ON NEAR-INFRARED（NIR）SPECTRA”的美國臨時專利申請案第62/692,248號的優先權，該美國臨時申請案的內容藉由引用以其整體併入本文。

【先前技術】

【0004】 原料識別可用於藥物產品的品質控制。例如，可以對醫用材料執行原料識別，以確定醫用材料的組成成分是否對應於與醫用材料相關聯的包裝標簽。類似地，可以執行原料定量以確定特定樣本中特定化學物的濃度。光譜法可以有助於進行無損的原料識別及/或定量，與其他化學計量技術相比，其準備和數據採集時間減少。

【發明內容】

【0005】 本申請的實現方式主要包括以下方面：

1）一種裝置，可以包括：一個或多個記憶體；以及一個或多個處理器，所述一個或多個處理器通信地耦合到所述一個或多個記憶體，所述一個或多個處

理器被配置為：接收對於第一光譜模型的主數據集；接收對於與第一光譜模型相關聯的目標群體的目標數據集，以更新第一光譜模型；產生包括主數據集和來自目標數據集的第一數據的訓練數據集；產生包括來自目標數據集的第二數據而不是主數據集的驗證數據集；使用交叉驗證並使用訓練數據集和驗證數據集產生作為第一光譜模型的更新的第二光譜模型；並提供第二光譜模型。

2）根據1）所述的裝置，其中，所述一個或多個處理器被配置為：

接收光譜測量；

使用所述第二光譜模型來執行光譜測定；和

提供識別所述光譜測定的輸出。

3）根據1）所述的裝置，其中，所述訓練數據集是複數個訓練數據集，並且所述驗證數據集是複數個驗證數據集；以及

其中，當產生所述第二光譜模型時，所述一個或多個處理器被配置成：

基於所述複數個訓練數據集和所述複數個驗證數據集產生複數個性能度量；

基於所述複數個性能度量來確定總體性能度量；

基於所述總體性能度量來確定最優部分最小平方（partial least squares，PLS）因子；和

基於所述最優PLS因子和合併的數據集確定所述第二光譜模型，

其中，所述合併的數據集包括所述主數據集和所述目標數據集。

4）根據1）所述的裝置，其中，所述第一光譜模型和所述第二光譜模型是定量模型。

5）根據1）所述的裝置，其中，所述主數據集基於由主光譜儀執行的第一組光譜測量，並且所述目標數據集基於由不同於所述主光譜儀的目標光譜儀執行的第二組光譜測量。

6）根據1）所述的裝置，其中，所述主數據集基於由特定光譜儀執行的第一組光譜測量，並且所述目標數據集基於由所述特定光譜儀執行的第二組光譜測量。

7）一種方法，可以包括：藉由裝置接收對於與第一光譜模型相關聯的目標群體的目標數據集；基於接收到所述目標數據集藉由所述裝置來獲得對於所述第一光譜模型的主數據集；藉由所述裝置使用交叉驗證確定最優部分最小平方（PLS）因子，其中，所述最優PLS因子基於複數個（a plurality of）訓練數據集且基於複數個驗證數據集確定，每個訓練數據集包括目標數據集的各別部分和主數據集的全部，每個驗證數據集包括目標數據集的各別部分並且不包括主數據集的數據；藉由所述裝置合併所述目標數據集和所述主數據集，以產生合併的數據集；藉由所述裝置並使用合併的數據集和最優PLS因子產生第二光譜模型，其中，所述第二光譜模型是所述第一光譜模型的更新；以及藉由所述裝置提供第二光譜模型來代替第一光譜模型。

8）根據7）所述的方法，其中，確定所述最優PLS因子包括：

確定對於所述複數個訓練數據集中的每一個訓練數據集和所述複數個驗證數據集中的每一個驗證數據集的部分最小平方（PLS）性能度量；

基於所述PLS性能度量確定總體PLS性能度量；和

基於所述總體PLS性能度量，優化PLS因子以確定對於所述第二光譜模型的最優PLS因子。

9）根據8）所述的方法，其中，所述總體PLS性能度量與均方根誤差（RMSE）值相關聯，並且

其中，優化所述PLS因子包括：

優化所述PLS因子以最小化RMSE值。

10）根據8）所述的方法，其中，所述複數個驗證數據集包括與所述複數個

訓練數據集不同的所述目標數據集中的數據。

11）根據8）所述的方法，其中，確定所述總體PLS性能度量包括：

聚合所述PLS性能度量。

12）根據7）所述的方法，其中，所述目標數據集關聯於在與所述主數據集相關聯的測量之後執行的對所述目標群體的一組測量。

13）根據12）所述的方法，其中，所述第二光譜模型是所述第一光譜模型的校準更新模型。

14）根據7）所述的方法，其中，所述目標數據集與由特定光譜儀執行的一組測量相關聯，所述特定光譜儀不同於執行與所述主數據集相關聯的測量的一個或複數個光譜儀。

15）根據14）所述的方法，其中，所述第二光譜模型是所述第一光譜模型的校準傳遞模型。

16）根據14）所述的方法，其中，提供所述第二光譜模型包括：

提供所述第二光譜模型以用於與藉由所述特定光譜儀進行的後續測量結合使用。

17）一種非暫時性電腦可讀取媒體，可以儲存一個或多個指令。當被裝置的一個或多個處理器執行時，所述一個或多個指令可以使所述一個或多個處理器：接收對於第一光譜模型的主數據集；接收對於與第一光譜模型相關聯的目標群體的目標數據集，以更新第一光譜模型；基於主數據集和目標數據集產生複數個訓練數據集；基於目標數據集產生複數個驗證數據集，其中，所述複數個驗證數據集不包括主數據集的數據；基於所述複數個訓練數據集和所述複數個驗證數據集並使用交叉驗證來確定模型設定；基於模型設定、目標數據集和主數據集產生第二光譜模型；並提供第二光譜模型。

18）根據17）所述的非暫時性電腦可讀取媒體，其中，所述模型設定是以

下中的至少一項，

部分最小平方（PLS）模型的PLS因子，

主成分回歸（PCR）模型的成分的數量，

支持向量回歸（SVR）模型的SVR參數，或

預處理設定。

19）根據17）所述的非暫時性電腦可讀取媒體，其中，使所述一個或多個處理器確定所述模型設定的所述一個或多個指令使所述一個或多個處理器：

產生關於所述複數個訓練數據集中的每一個訓練數據集和所述複數個驗證數據集中的對應的驗證數據集的複數個部分性能度量；

聚合所述複數個部分性能度量以產生總體性能度量；和

確定所述模型設定，以最小化所述總體性能度量的誤差值。

20）根據17）所述的非暫時性電腦可讀取媒體，其中，所述一個或多個指令在由所述一個或多個處理器執行時使所述一個或多個處理器：

接收測量；

基於所述測量並使用所述第二光譜模型來執行光譜測定；和

提供識別所述光譜測定的輸出。

【圖式簡單說明】

【0006】 圖1A-1E是本文中描述的示例實現方式的概觀的示意圖。

【0007】 圖2是在其中可以實現本文中描述的系統及/或方法的示例環境的示意圖。

【0008】 圖3是圖2中的一個或多個裝置的示例部件的示意圖。

【0009】 圖4-6是對於光譜模型的基於交叉驗證的校準的示例過程的流程圖。

【實施方式】

【0010】 示例實現方式的以下詳細描述參考了附圖。不同附圖中的相同參考數字可以標識相同或類似的元素。以下描述使用光譜儀（spectrometer）作為示例，然而，本文描述的校準原理、過程和方法可以與任何感測器一起使用，包括但不限於其他光學感測器和光譜感測器。

【0011】 原料識別（raw material identification，RMID）是用於識別特定樣本的組成（例如，成分）以用於識別、驗證及/或類似物的技術。例如，RMID可用於驗證藥物材料中的成分是否對應於標簽上標識的成分集。類似地，原料定量是用於對特定樣本執行定量分析的技術，例如確定特定樣本中特定材料的濃度。光譜儀可用於對樣本（例如，藥物材料）執行光譜法，以確定樣本的組成、樣本的組成的濃度及/或類似物。光譜儀可確定對樣本的一組測量，並可提供該一組測量用於光譜測定。光譜分類技術（例如，分類器）可以有助於基於對樣本的該一組測量而確定樣本的組成。

【0012】 為了執行光譜分類或定量，光譜模型可被用於評估對未知樣本的一個或多個測量。例如，控制裝置可以嘗試將對未知樣本的一個或多個測量分類為對應於光譜模型的特定類別、與光譜模型相關聯的特定水平及/或數量，及/或類似物。然而，隨著時間的推移，原料可能會發生變化，這可能會導致光譜模型不準確。例如，對於被應用於農產品的光譜分類，與不同年份相關聯的不同收穫可能具有不同的光譜。因此，在主數據集（例如，在初始時間對初始群體的初始一組光譜測量）訓練的光譜模型當應用於目標數據集（例如，在後續時間對後續群體的後續一組光譜測量）時可能不準確。

【0013】 在另一種情況下，使用對於每個光譜儀的主數據集來針對每個光譜儀訓練光譜模型可能是不切實際的。因此，控制裝置可以在主數據集上訓練

單個光譜模型，並且部署該單個光譜模型以用於許多不同的光譜儀。然而，不同的光譜儀可以與不同的校準相關聯及/或可以在不同的環境條件下操作。因此，當被應用於由第二光譜儀執行的光譜測量的目標數據集時，使用由第一光譜儀執行的光譜測量的主數據集訓練的光譜模型可能不準確。

【0014】 本文所述的一些實現方式能夠使用交叉驗證技術對光譜模型進行校準更新和校準傳遞（calibration transfer）。例如，來自目標數據集的數據可以與來自主數據集的數據合併，以能夠產生新的光譜模型。在這種情況下，來自主數據集的數據用於訓練光譜模型的訓練集，以及來自目標數據集的數據用於訓練集和驗證光譜模型的驗證集兩者。以這種方式，相對於用於模型產生及/或模型更新的其他技術，光譜模型的準確度得到提高。此外，基於提高傳遞的光譜模型的準確度，針對每個光譜儀獲得主數據集的必要性被降低，從而降低了與部署光譜儀相關聯的成本。

【0015】 圖1A-1E是本文描述的示例實現方式100的示意圖。如圖1A所示，示例實現方式100包括第一光譜儀102和第一控制裝置104。

【0016】 又如圖1A並由參考數字150所示，第一控制裝置104可以向第一光譜儀102傳輸指令，以使第一光譜儀102對主群體152執行一組光譜測量。例如，第一控制裝置104可以使第一光譜儀102針對將要使用分類模型而被分類的每個類別、針對將要使用定量模型而被定量的每個數量及/或類似物，對樣本執行測量。分類模型的類別可以指共享一種或更多種共同特徵的相似材料的分組，例如（在藥物背景下的）乳糖材料、果糖材料、對乙醯氨酚材料、布洛芬（ibuprophen）材料、阿斯匹靈材料及/或類似物。用於訓練分類模型的以及對於其要使用分類模型來執行原料識別的材料可以被稱為感興趣材料。

【0017】 又如圖1A並且由參考數字154和156所示，第一光譜儀102可以執行一組光譜測量，並且可以將該一組光譜測量提供給第一控制裝置104進行處

理。例如，第一光譜儀102可以確定對於主群體152的每個樣本的光譜，以使第一控制裝置104能夠產生一組類別，用於將未知樣本分類為對於定量模型的感興趣材料之一或者分類為具有與定量模型相關的特定量。

【0018】 又如圖1A並且由參考數字158所示，第一控制裝置104可以基於主數據集產生第一光譜模型。例如，第一控制裝置104可以使用特定的測定技術並基於該一組光譜測量來產生第一光譜模型。在一些實現方式中，第一控制裝置104可以使用支持向量機（SVM）技術（例如，用於資訊確定的機器學習技術）產生定量模型。附加地或替代地，第一控制裝置104可以使用另一種類型的定量技術來產生定量模型。

【0019】 定量模型可以包括與將特定光譜分配給感興趣材料的數量的特定類別相關聯的資訊。在一些實現方式中，定量模型可以包括與識別與數量的特定類別相關聯的感興趣材料的類型相關聯的資訊。以該方式，第一控制裝置104可以基於將未知樣本的光譜分配給定量模型的數量的特定類別，來提供識別未知樣本的材料的數量的資訊作為光譜輸出。

【0020】 如圖1B並且由參考數字160所示，第二控制裝置104可以接收與第一光譜模型相關聯的資訊。例如，第二控制裝置104可以接收第一光譜模型、主數據集及/或類似物。在一些實現方式中，相比於第一控制裝置104，第二控制裝置104可以與不同的光譜儀相關聯。例如，在校準傳遞情況下，第二控制裝置104可以與第二光譜儀102（例如，目標光譜儀）結合使用，並且可以接收與第一光譜模型相關聯的資訊，以實現從第一光譜儀102（例如，主光譜儀）到第二光譜儀102的校準傳遞。在這種情況下，如本文更詳細描述的那樣，第二控制裝置104和第二光譜儀102可以執行對目標群體的測量，並產生第二光譜模型。可選地，如本文更詳細描述的那樣，在校準更新的情況下，第一控制裝置104和第一光譜儀102可以執行對目標群體的測量並產生第二光譜模型，而不是將第一光

譜模型傳遞到第二控制裝置104。

【0021】 又如圖1B並且由參考數字162所示，第二控制裝置104可以向第二光譜儀102傳輸指令，以使第二光譜儀102執行對目標群體164的一組光譜測量。例如，第二控制裝置104可以使第二光譜儀102基於接收到第一光譜模型來執行對目標群體164的光譜測量。在一些實現方式中，第二控制裝置104可以確定更新或校準第一光譜模型，並且可以觸發第二光譜儀102執行該一組光譜測量。在這種情況下，第二控制裝置104可以與第一控制裝置104通信，以獲得識別主數據集的資訊，從而能夠產生第二光譜模型。

【0022】 在一些實現方式中，目標群體164可以對應於主群體152。例如，目標群體164可以是如同已被包括在主群體152中的相同類別之另外的樣本。在這種情況下，目標群體164可以與主群體152關於樣本被收集或測量的時間、位置、環境條件及/或類似物方面不同。此外，或者替代地，基於使用不同的光譜儀進行測量（例如，由第二光譜儀102而不是由第一光譜儀102對主群體152進行測量），目標群體164可以不同於主群體152。

【0023】 又如圖1B並且由參考數字166和168所示，第二光譜儀102可以執行一組光譜測量，並且可以向第二控制裝置104提供識別該一組光譜測量的資訊。例如，第二光譜儀102可以執行對目標群體164的光譜測量，並且可以將識別光譜測量（例如，作為目標數據集）的資訊提供給第二控制裝置104以進行處理。

【0024】 如圖1C並且由參考數字170所示，第二控制裝置104可以確定總體性能度量。例如，第二控制裝置104可以基於將數據劃分成多個等分（folds）、確定對於多個等分的多個性能度量、聚合多個性能度量以確定均方根誤差（RMSE）值、以及優化部分最小平方（PLS）因子（其可以被稱為最優PLS因子）以最小化RMSE值，從而確定總體性能度量。等分可以指用於交叉驗證的數

據子組，其包括產生候選模型的訓練集和評估候選模型在預測數據時的準確度的驗證集。在另一個示例中，第二控制裝置104可以確定另一種類型的優化模型設定，例如與主成分回歸（PCR）因子、支持向量回歸（SVR）因子及/或類似物相關的模型設定。在一些實現方式中，第二控制裝置104可以執行預處理優化。例如，第二控制裝置104可以確定優化的預處理參數作為模型設定的一部分。

【0025】 在一些實現方式中，第二控制裝置104可以針對每個等分向訓練集或驗證集分配數據。例如，第二控制裝置104可以確定對於N個等分的多個訓練集1至N，以及對於N個等分的多個對應的驗證集1至N。在一些實現方式中，訓練集可以包括藉由合併主數據集和目標數據集產生的合併數據。例如，訓練集（例如，訓練集1）可以包括來自主數據集（master data set）（例如，MDS）的所有數據和來自目標數據集（target data set）的一部分數據（例如，$TDS_{1,TS}$）。在這種情況下，對應的驗證集可以包括來自目標數據集的一對應部分數據（例如，$TDS_{1,VS}$）和不是來自主數據集的數據。對應的驗證集可以省略從相同物理樣本的副本掃描中導出的數據，因為該相同物理樣本被包括在訓練集中。

【0026】 基於將數據分配到多個等分，第二控制裝置104可以確定對於每個等分的性能度量。例如，第二控制裝置104對於每個等分可以確定並可以聚合性能度量，以確定總體性能度量。例如，第二控制裝置104可以確定對於每個等分的PLS因子，並且可以確定對於每個等分的每個PLS因子的RMSE值。基於確定對於每個等分的每個PLS因子的RMSE值，第二控制裝置104可以確定總體RMSE值。例如，第二控制裝置104可以根據所有等分的所有PLS因子來確定RMSE值。在這種情況下，基於確定總體RMSE值，第二控制裝置104可以確定最優PLS因子，其可以是具有最低RMSE值的PLS因子。

【0027】 在這種情況下，在交叉驗證期間，基於將主數據集和目標數據集包括在N個等分的訓練集中，但是僅將目標數據集包括在對應的驗證集中，第二

光譜模型的準確度相對於其他技術被提高。例如，相對於在不更新的情況下使用第一光譜模型、相對於僅使用目標數據集來確定PLS性能度量、相對於合併所有的目標數據集數據和所有的主數據集數據來產生合併的數據集以及使用將合併的數據集劃分在訓練集和驗證集兩者中及/或類似物，這種技術可以導致提高的準確度。

【0028】 如圖1D並且由參考數字172所示，第二控制裝置104可以產生第二光譜模型。例如，第二控制裝置104可以使用主數據集（MDS）、目標數據集（TDS）、和最優PLS因子來產生第二光譜模型。以這種方式，第二控制裝置104可以實現校準的光譜模型、更新的光譜模型、傳遞的光譜模型及/或類似物的產生。

【0029】 在一些實現方式中，第二控制裝置104可以合併主數據集和目標數據集以產生合併的數據集（例如，用於訓練第二光譜模型的最終訓練集）。例如，第二控制裝置104可以聚合主數據集和目標數據集以產生合併的數據集。基於產生合併的數據集，第二控制裝置104可以使用合併的數據集和最優PLS因子（例如，具有最低RMSE值）產生第二光譜模型。例如，第二控制裝置104，可以使用定量模型產生技術以結合合併的數據集（例如，其可以是對於第二光譜模型的訓練集）和最優PLS因子來產生第二光譜模型。以這樣的方式，藉由在不使用合併的數據集的情況下確定最優PLS因子，然後將最優PLS因子與合併的數據集相組合，第二控制裝置104實現了比其他技術更準確的光譜模型。

【0030】 在一些實現方式中，第二控制裝置104可以基於產生第二光譜模型來提供第二光譜模型。例如，第二控制裝置104可以提供第二光譜模型，該第二光譜模型用於經由數據結構進行儲存、用於部署在一個或多個其他光譜儀及/或類似物上。附加地或替代地，第二控制裝置104可以基於產生第二光譜模型來提供與第二光譜模型相關的輸出。例如，如本文更詳細描述的，第二控制裝置

104可以基於使用第二光譜模型分析未知樣本來提供定量未知樣本的資訊。

【0031】 如圖1E且由參考數字174所示，第二控制裝置104可以向第二光譜儀102傳輸指令，以使第二光譜儀102對未知樣本176執行一組光譜測量。例如，第二控制裝置104可以使第二光譜儀102在已經產生第二光譜模型之後對未知樣本176執行光譜測量。

【0032】 又如圖1E且由參考數字178和180所示，第二光譜儀102可以執行一組光譜測量，並且可以向第二控制裝置104提供識別該一組光譜測量的資訊。例如，第二光譜儀102可以確定未知樣本176的光譜，並且可以向第二控制裝置104提供識別光譜的資訊，以用於分類及/或定量。

【0033】 又如圖1E且由參考數字182所示，第二控制裝置104可以使用第二光譜模型來執行對該一組光譜測量的光譜分析。例如，第二控制裝置104可以使用第二光譜模型來確定未知樣本176的分類及/或未知樣本176的定量。在這種情況下，第二控制裝置104可以提供識別分類及/或定量的輸出。以這樣的方式，第二控制裝置104基於產生第二光譜模型來使用第二光譜模型。

【0034】 如上面指示的，圖1A-1E僅作為一個或多個示例被提供。其它示例可以不同於關於圖1A-1E描述的示例。

【0035】 圖2是在其中可以實現本文中所述的系統及/或方法的示例環境200的示意圖。如圖2所示，環境200可以包括控制裝置210、光譜儀220、網路230及/或類似物。環境200中的裝置可經由有線連接、無線連接或有線連接和無線連接的組合來互連。

【0036】 控制裝置210包括能夠儲存、處理及/或路由與光譜分類相關聯的資訊的一個或多個裝置。例如，控制裝置210可以包括伺服器、電腦、可穿戴裝置、雲計算裝置及/或類似物，其基於訓練集的一組測量產生光譜模型（如，分類模型或定量模型）、基於驗證集的一組測量驗證光譜模型，及/或利用光譜模型

來基於對未知樣本的一組測量執行光譜分析。在一些實現方式中，控制裝置210可以與特定光譜儀220相關聯。在一些實現方式中，控制裝置210可以與多個光譜儀220相關聯。在一些實現方式中，控制裝置210可以從環境200中的另一裝置（例如，光譜儀220）接收資訊及/或向環境200中的另一裝置（例如光譜儀220）傳輸資訊。

【0037】 光譜儀220包括能夠對樣本執行光譜測量的一個或多個裝置。例如，光譜儀220可以包括執行光譜法的光譜裝置，該光譜法例如，振動光譜法，諸如近紅外（NIR）光譜法、中紅外光譜法（mid-IR）、拉曼（Raman）光譜法及/或類似物）的光譜裝置。在一些實現方式中，光譜儀220可以被結合到可穿戴裝置中，例如可穿戴光譜儀及/或類似物。在一些實現方式中，光譜儀220可以從環境200中的另一裝置（例如控制裝置210）接收資訊及/或向環境200中的另一裝置（例如控制裝置210）傳輸資訊。

【0038】 網路230包括一個或多個有線及/或無線網路。例如，網路230可包括蜂窩網路（例如，長期演進（LTE）網路、3G網路、分碼多工存取（CDMA）網路等）、公共陸地移動網路（PLMN）、區域網路（LAN）、廣域網路（WAN）、城域網域（MAN）、電話網路（例如，公共交換電話網路（PSTN））、專用網路、自組織網路、內部網路、網際網路、基於光纖的網路、雲計算網路及/或類似物，及/或這些或其它類型的網路的組合。

【0039】 圖2中顯示的裝置和網路的數量和佈置作為一個或多個示例被提供。實際上，與圖2中顯示的那些裝置及/或網路相比，可以有另外的裝置及/或網路、更少的裝置及/或網路、不同的裝置及/或網路或不同地佈置的裝置及/或網路。此外，圖2中顯示的兩個或多個裝置可在單個裝置內實現，或圖2中顯示的單個裝置可被實現為多個分布式裝置。另外或可替代地，環境200中的一組裝置（例如，一個或多個裝置）可以執行被描述為由環境200中的另一組裝置執行的

一個或多個功能。

【0040】 圖3是裝置300的示例部件的示意圖。裝置300可以對應於控制裝置210及/或光譜儀220。在一些實現方式中，控制裝置210及/或光譜儀220可以包括一個或多個裝置300及/或裝置300的一個或多個部件。如圖3所示，裝置300可包括匯流排310、處理器320、記憶體330、儲存部件340、輸入部件350、輸出部件360以及通信介面370。

【0041】 匯流排310包括允許在裝置300的複數個部件當中通信的部件。處理器320以硬體、韌體、及/或硬體和軟體的組合實現。處理器320是中央處理單元（CPU）、圖形處理單元（GPU）、加速處理單元（APU）、微處理器、微控制器、數位信號處理器（DSP）、場可編程閘陣列（FPGA）、特殊應用積體電路（ASIC）或另一類型的處理部件。在一些實現方式中，處理器320包括能夠被編程以執行功能的一個或多個處理器。記憶體330包括隨機存取記憶體（RAM）、唯讀記憶體（ROM）及/或儲存資訊及/或指令以用於由處理器320使用的另一類型的動態或靜態儲存裝置（例如，閃存、磁記憶體及/或光學記憶體）。

【0042】 儲存部件340儲存與裝置300的操作和使用相關的資訊及/或軟體。例如，儲存部件340可以包括硬碟（例如，磁碟、光碟及/或磁光碟）、固態驅動器（SSD）、壓縮光碟（CD）、數位多功能光碟（DVD）、軟碟、盒式磁帶（cartridge）、磁帶（magnetic tape）及/或另一類型的非暫時性電腦可讀取媒體，連同對應的驅動器。

【0043】 輸入部件350包括允許裝置300例如經由用戶輸入端（例如，觸控螢幕顯示器、鍵盤、小鍵盤（keypad）、鼠標、按鈕、開關及/或麥克風）來接收資訊的部件。附加地或可替換地，輸入部件350可以包括用於確定位置的部件（例如，全球定位系統（GPS）部件）及/或感測器（例如，加速度計、陀螺儀、致動器、另一種類型的位置或環境感測器及/或類似物）。輸出部件360包括（經由

例如顯示器、揚聲器、觸覺反饋部件、聲音或視覺指示器及/或類似物）提供來自裝置300的輸出資訊的部件。

【0044】 通信介面370包括使裝置300能夠例如經由有線連接、無線連接或有線和無線連接的組合與其它裝置通信的類似收發機的部件（例如，收發機、單獨的接收機、單獨的發射機及/或類似物）。通信介面370可以允許裝置300從另一裝置接收資訊及/或將資訊提供給另一裝置。例如，通信介面370可包括乙太網路介面、光學介面、同軸介面、紅外線介面、射頻（RF）介面、通用串列匯流排（USB）介面、Wi-Fi介面、蜂窩網路介面及/或類似物。

【0045】 裝置300可以執行本文中所述的一個或多個過程。裝置300可以基於處理器320執行由非暫時性電腦可讀取媒體（例如記憶體330及/或儲存部件340）儲存的軟體指令來執行這些過程。如本文所使用的，術語“電腦可讀取媒體”指的是非暫時性記憶體裝置。記憶體裝置包括在單個物理儲存裝置內的記憶體空間或遍佈於多個物理儲存裝置的記憶體空間。

【0046】 軟體指令可以經由通信介面370從另一電腦可讀取媒體或從另一裝置被讀取到記憶體330及/或儲存部件340中。儲存在記憶體330及/或儲存部件340中的軟體指令當被執行時，可使處理器320執行本文中所述的一個或多個過程。另外或可替代地，硬體電路系統可代替軟體指令或與軟體指令組合來使用，以執行本文中所述的一個或多個過程。因此，本文中所述的實現方式不限於硬體電路系統和軟體的任何特定組合。

【0047】 圖3中顯示的部件的數量和佈置作為示例被提供。實際上，與圖3中顯示的那些部件相比，裝置300可以包括另外的部件、更少的部件、不同的部件或不同地佈置的部件。另外或可替代地，裝置300的一組部件（例如，一個或多個部件）可以執行被描述為由裝置300的另一組部件執行的一個或多個功能。

【0048】 圖4是用於光譜模型的基於交叉驗證的校準的示例過程400的流

程圖。在一些實現方式中，圖4的一個或多個過程區塊可由控制裝置（例如控制裝置210）執行。在一些實現方式中，圖4的一個或多個過程區塊可由與控制裝置分離的或者包括控制裝置的另一裝置或一組裝置（諸如光譜儀（如光譜儀220）及/或類似物）執行。

【0049】 如圖4中所示，過程400可以包括接收對於第一光譜模型的主數據集（區塊410）。例如，如上所述，控制裝置（例如，使用處理器320、記憶體330、儲存部件340、輸入部件350、輸出部件360、通信介面370及/或類似物）可以接收對於第一光譜模型的主數據集。

【0050】 如圖4中進一步所示，過程400可以包括接收對於與第一光譜模型相關聯的目標群體的目標數據集，以更新第一光譜模型（區塊420）。例如，如上所述，控制裝置（例如，使用處理器320、記憶體330、儲存部件340、輸入部件350、輸出部件360、通信介面370及/或類似物）可以接收對於與第一光譜模型相關聯的目標群體的目標數據集，以更新第一光譜模型。

【0051】 如圖4中進一步所示，過程400可以包括產生訓練數據集，該訓練數據集包括主數據集和來自目標數據集的第一數據（區塊430）。例如，如上所述，控制裝置（例如，使用處理器320、記憶體330、儲存部件340、輸入部件350、輸出部件360、通信介面370及/或類似物）可以產生包括主數據集和來自目標數據集的第一數據的訓練數據集。

【0052】 如圖4中進一步所示，過程400可以包括產生驗證數據集，該驗證數據集包括來自目標數據集的第二數據而不是主數據集（區塊440）。例如，如上所述，控制裝置（例如，使用處理器320、記憶體330、儲存部件340、輸入部件350、輸出部件360、通信介面370及/或類似物）可以產生包括來自目標數據集的第二數據而不是主數據集的驗證數據集。

【0053】 如圖4進一步所示，過程400可以包括使用交叉驗證並使用訓練數

據集和驗證數據集，產生作為第一光譜模型的更新的第二光譜模型（區塊450）。例如，如上所述，控制裝置（例如，使用處理器320、記憶體330、儲存部件340、輸入部件350、輸出部件360、通信介面370及/或類似物）可以使用交叉驗證並使用訓練數據集和驗證數據集產生作為第一光譜模型的更新的第二光譜模型。

【0054】 如在圖4中進一步所示，過程400可以包括提供第二光譜模型（區塊460）。例如，如上所述，控制裝置（例如，使用處理器320、記憶體330、儲存部件340、輸入部件350、輸出部件360、通信介面370及/或類似物）可以提供第二光譜模型。

【0055】 過程400可以包括另外的實現方式，諸如以下描述的及/或結合本文其他地方描述的一個或多個其他過程的任何單個實現方式或實現方式的任何組合。

【0056】 在第一實現方式中，過程400包括接收光譜測量；使用第二光譜模型執行光譜測定；以及提供識別光譜測定的輸出。

【0057】 在第二實現方式中，單獨或與第一實現方式相結合，訓練數據集是複數個訓練數據集，且驗證數據集是複數個驗證數據集，並且過程400包括基於複數個訓練數據集和複數個驗證數據集產生複數個性能度量，基於複數個性能度量確定總體性能度量，基於總體性能度量確定最優部分最小平方（PLS）因子，並且基於最優PLS因子和合併的數據集確定第二光譜模型，其中合併的數據集包括主數據集和目標數據集。

【0058】 在第三實現方式中，單獨或與第一實現方式和第二實現方式中的一個或多個結合，第一光譜模型和第二光譜模型是定量模型。

【0059】 在第四實現方式中，單獨地或者與第一實現方式至第三實現方式中的一個或多個相結合，主數據集基於由主光譜儀執行的第一組光譜測量，並且目標數據集基於由不同於主光譜儀的目標光譜儀執行的第二組光譜測量。

【0060】 在第五實現方式中，單獨或與第一實現方式至第四實現方式中的一個或多個相結合，主數據集基於由特定光譜儀執行的第一組光譜測量，而目標數據集基於由特定光譜儀執行的第二組光譜測量。

【0061】 雖然圖4示出了過程400的示例區塊，但在一些實現方式中，與圖4中描繪的那些區塊相比，過程400可包括附加的區塊、更少的區塊、不同的區塊或不同地佈置的區塊。另外或可替代地，過程400的兩個或多個區塊可並行地被執行。

【0062】 圖5是用於光譜模型的基於交叉驗證的校準的示例過程500的流程圖。在一些實現方式中，圖5的一個或多個過程區塊可由控制裝置（如控制裝置210）執行。在一些實現方式中，圖5的一個或多個過程區塊可由與控制裝置210分離的或者包括控制裝置的另一裝置或一組裝置（諸如光譜儀（如光譜儀220））執行。

【0063】 如圖5所示，過程500可以包括接收對於與第一光譜模型相關聯的目標群體的目標數據集（區塊510）。例如，如上所述，控制裝置（例如，使用處理器320、記憶體330、儲存部件340、輸入部件350、輸出部件360、通信介面370及/或類似物）可以接收對於與第一光譜模型相關聯的目標群體的目標數據集。

【0064】 如圖5中進一步所示，過程500可以包括基於接收目標數據集來獲得對於第一光譜模型的主數據集（區塊520）。例如，如上所述，控制裝置（例如，使用處理器320、記憶體330、儲存部件340、輸入部件350、輸出部件360、通信介面370及/或類似物）可以基於接收目標數據集來獲得對於第一光譜模型的主數據集。

【0065】 如圖5中進一步所示，過程500可以包括使用交叉驗證確定最優部分最小平方（PLS）因子，其中最優PLS因子基於複數個訓練數據集且基於複數

個驗證數據集來確定，每個訓練數據集包括目標數據集的各別部分和主數據集的全部，並且每個驗證數據集包括目標數據集的各別部分並且不包括主數據集的數據（區塊530）。例如，如上所述，控制裝置（例如，使用處理器320、記憶體330、儲存部件340、輸入部件350、輸出部件360、通信介面370及/或類似物）可以使用交叉驗證確定最優部分最小平方（PLS）因子。在一些方面，最優PLS因子基於複數個訓練數據集且基於複數個驗證數據集來確定，每個訓練數據集包括目標數據集的各別部分和主數據集的全部，每個驗證數據集包括目標數據集的各別部分並且不包括主數據集的數據。

【0066】 如圖5中進一步所示，過程500可以包括合併目標數據集和主數據集以產生合併的數據集（區塊540）。例如，如上所述，控制裝置（例如，使用處理器320、記憶體330、儲存部件340、輸入部件350、輸出部件360、通信介面370及/或類似物）可以合併目標數據集和主數據集以產生合併的數據集。

【0067】 如圖5中進一步所示，過程500可以包括使用合併的數據集和最優PLS因子來產生第二光譜模型，其中第二光譜模型是第一光譜模型的更新（區塊550）。例如，如上所述，控制裝置（例如，使用處理器320、記憶體330、儲存部件340、輸入部件350、輸出部件360、通信介面370及/或類似物）可以使用合併的數據集和最優PLS因子來產生第二光譜模型。在一些方面，第二光譜模型是第一光譜模型的更新。

【0068】 如在圖5中進一步所示，過程500可以包括提供第二光譜模型以代替第一光譜模型（區塊560）。例如，如上所述，控制裝置（例如，使用處理器320、記憶體330、儲存部件340、輸入部件350、輸出部件360、通信介面370及/或類似物）可以提供第二光譜模型以代替第一光譜模型。

【0069】 過程500可以包括另外的實現方式，諸如以下描述的及/或結合本文其他地方描述的一個或多個其他過程的任何單個實現方式或實現方式的任何

組合。

【0070】 在第一實現方式中，確定最優PLS因子包括確定對於複數個訓練數據集的每一個訓練數據集和複數個驗證數據集的每一個驗證數據集的部分最小平方（PLS）性能度量；基於PLS性能度量確定總體PLS性能度量；以及基於總體PLS性能度量來優化用於第二光譜模型的PLS因子。

【0071】 在第二實現方式中，單獨地或與第一實現方式結合，總體PLS性能度量與均方根誤差（RMSE）值相關聯，並且優化PLS因子包括優化PLS因子以最小化RMSE值。

【0072】 在第三實現方式中，單獨地或與第一實現方式和第二實現方式中的一個或多個結合，複數個驗證數據集包括與複數個訓練數據集相比的目標數據集的不同數據。

【0073】 在第四實現方式中，單獨地或與第一實現方式至第三實現方式中的一個或多個結合，確定總體PLS性能度量包括聚合PLS性能度量。

【0074】 在第五實現方式中，單獨地或與第一實現方式至第四實現方式中的一個或多個結合，目標數據集與在與主數據集相關聯的測量之後執行的對目標群體的一組測量相關聯。

【0075】 在第六實現方式中，單獨或與第一實現方式至第五實現方式中的一個或多個結合，第二光譜模型是第一光譜模型的校準更新模型。

【0076】 在第七實現方式中，單獨地或者與第一實現方式至第六實現方式中的一個或多個相結合，目標數據集與由特定光譜儀執行的一組測量相關聯，該特定光譜儀不同於執行與主數據集相關聯的測量的一個或多個光譜儀。

【0077】 在第八實現方式中，單獨或與第一實現方式至第七實現方式中的一個或多個結合，第二光譜模型是第一光譜模型的校準傳遞模型。

【0078】 在第九實現方式中，單獨地或與第一實現方式至第八實現方式中

的一個或多個相結合，提供第二光譜模型包括提供第二光譜模型以結合由特定光譜儀進行的後續測量來進行使用。

【0079】 雖然圖5示出了過程500的示例區塊，但在一些實現方式中，與圖5中描繪的那些區塊相比，過程500可包括另外的區塊、更少的區塊、不同的區塊或不同地佈置的區塊。另外或可替代地，過程500的兩個或多個區塊可並行地被執行。

【0080】 圖6是用於光譜模型的基於交叉驗證的校準的示例過程600的流程圖。在一些實現方式中，圖6的一個或多個過程區塊可由控制裝置（例如，控制裝置210）執行。在一些實現方式中，圖6的一個或多個過程區塊可由與控制裝置分離的或者包括控制裝置的另一裝置或一組裝置（諸如光譜儀（如光譜儀220）及/或類似物）執行。

【0081】 如圖6所示，過程600可以包括接收對於第一光譜模型的主數據集，接收對於與第一光譜模型相關聯的目標群體的目標數據集以更新第一光譜模型，基於主數據集和目標數據集產生複數個訓練數據集，基於目標數據集產生複數個驗證數據集，並且其中複數個驗證數據集不包括主數據集的數據（區塊610）。例如，如上所述，控制裝置（例如，使用處理器320、記憶體330、儲存部件340、輸入部件350、輸出部件360、通信介面370及/或類似物）可以接收對於第一光譜模型的主數據集，接收對於與第一光譜模型相關聯的目標群體的目標數據集以更新第一光譜模型，基於主數據集和目標數據集產生複數個訓練數據集，基於目標數據集產生複數個驗證數據集，並且其中複數個驗證數據集不包括主數據集的數據。在一些方面，複數個驗證數據集不包括主數據集的數據。

【0082】 如圖6中所示，過程600可以包括接收對於第一光譜模型的主數據集（區塊610）。例如，如上所述，控制裝置（例如，使用處理器320、記憶體330、

儲存部件340、輸入部件350、輸出部件360、通信介面370及/或類似物)可以接收對於第一光譜模型的主數據集。

【0083】 如圖6所示，過程600可以包括接收對於與第一光譜模型相關聯的目標群體的目標數據集以更新第一光譜模型（區塊620）。例如，如上所述，控制裝置（例如，使用處理器320、記憶體330、儲存部件340、輸入部件350、輸出部件360、通信介面370及/或類似物）可以接收對於與第一光譜模型相關聯的目標群體的目標數據集以更新第一光譜模型。

【0084】 如圖6所示，過程600可包括基於主數據集和目標數據集產生複數個訓練數據集（區塊630）。例如，如上所述，控制裝置（例如，使用處理器320、記憶體330、儲存部件340、輸入部件350、輸出部件360、通信介面370及/或類似物）可以基於主數據集和目標數據集產生複數個訓練數據集。

【0085】 如圖6所示，過程600可以包括基於目標數據集產生複數個驗證數據集，並且其中複數個驗證數據集不包括主數據集的數據（區塊640）。例如，如上所述，控制裝置（例如，使用處理器320、記憶體330、儲存部件340、輸入部件350、輸出部件360、通信介面370及/或類似物）可以基於目標數據集產生複數個驗證數據集。在一些方面，複數個驗證數據集不包括主數據集的數據。

【0086】 如圖6中進一步所示，過程600可以包括基於複數個訓練數據集和複數個驗證數據集並使用交叉驗證來確定模型設定（區塊650）。例如，如上所述，控制裝置（例如，使用處理器320、記憶體330、儲存部件340、輸入部件350、輸出部件360、通信介面370及/或類似物）可以基於複數個訓練數據集和複數個驗證數據集來確定模型設定。

【0087】 如圖6進一步所示，過程600可以包括基於模型設定、目標數據集和主數據集產生第二光譜模型（區塊660）。例如，如上所述，控制裝置（例如，使用處理器320、記憶體330、儲存部件340、輸入部件350、輸出部件360、通信

介面370及/或類似物）可以基於模型設定、目標數據集和主數據集產生第二光譜模型。

【0088】 如在圖6中進一步所示，過程600可以包括提供第二光譜模型（區塊670）。例如，如上所述，控制裝置（例如，使用處理器320、記憶體330、儲存部件340、輸入部件350、輸出部件360、通信介面370及/或類似物）可以提供第二光譜模型。

【0089】 過程600可以包括另外的實現方式，諸如以下描述的及/或結合本文其他地方描述的一個或多個其他過程的任何單個實現方式或實現方式的任何組合。

【0090】 在第一實現方式中，模型設定是部分最小平方（PLS）模型的PLS因子、主成分回歸（PCR）模型的成分的數量、支持向量回歸（SVR）模型的SVR參數或預處理設定中的至少一個。

【0091】 在第二實現方式中，單獨地或與第一實現方式結合，過程600包括關於複數個訓練數據集中的每一個訓練數據集和複數個驗證數據集的對應的驗證數據集而產生複數個部分性能度量（a plurality of partial performance metrics），聚合該複數個部分性能度量以產生總體性能度量，以及確定模型設定以最小化總體性能度量的誤差值。

【0092】 在第三實現方式中，單獨地或與第一實現方式和第二實現方式中的一個或多個結合，過程600包括基於測量和使用第二光譜模型來執行光譜測定，並提供識別光譜測定的輸出。

【0093】 雖然圖6示出了過程600的示例區塊，但在一些實現方式中，與圖6中描繪的那些區塊相比，過程600可包括另外的區塊、更少的區塊、不同的區塊或不同地佈置的區塊。另外或可替代地，過程600的兩個或多個區塊可並行地被執行。

【0094】 前述公開提供了說明和描述，但並不旨在窮舉或將實現方式限制到所公開的精確形式。修改和變型可以根據以上公開內容進行，或者可以從實現方式的實踐中獲得。

【0095】 如在本文使用的，術語“部件”被規定為廣泛地被解釋為硬體、韌體、及/或硬體和軟體的組合。

【0096】 如本文所使用的，滿足閾值可以根據上下文指大於閾值、多於閾值、高於閾值、大於或等於閾值、小於閾值、少於閾值、低於閾值、小於或等於閾值、等於閾值及/或類似物的值。

【0097】 將明顯的是，本文中所述的系統和/方法可在硬體、韌體、或硬體和軟體的組合的不同形式中實現。用於實現這些系統及/或方法的實際專用控制硬體或軟體代碼不是實現方式的限制。因此，在本文中描述系統及/或方法的操作和行為而不參考特定的軟體代碼，應理解的是，軟體和硬體可以被設計成基於本文中的描述來實現系統及/或方法。

【0098】 儘管在請求項中陳述及/或在說明書中公開了特徵的特定組合，但是這些組合並不旨在限制各種實現方式的公開內容。事實上，這些特徵中的許多可以以請求項中未具體陳述及/或說明書中未公開的方式組合。儘管以下所列出的每個附屬請求項可以直接附屬僅僅一個請求項，但是各種實現方式的公開內容包括每個附屬請求項與請求項組中的每個其他請求項相結合。

【0099】 本文中使用的任何元素、動作或指令都不應當被解釋為關鍵或必要的，除非明確這樣描述。另外，如本文中所使用的，冠詞“一（a）”和“一（an）”旨在包括一個或多個項，並且可以與“一個或多個（one or more）”互換使用。此外，如本文所用，冠詞“該（the）”旨在包括與冠詞“該”相結合的一個或多個項，並且可以與“該一個或多個”互換使用。此外，如本文所使用的，術語“集（set）”旨在包括一個或多個項（例如，相關項、不相關項、相關項和不相關項的組合

及/或類似物），並且可以與“一個或多個”互換使用。在意指僅一個項的情況下，使用術語“僅一個（only one）”或類似的語言。此外，如本文所用，術語“具有（has）”、“具有（have）”、“具有（having）”及/或類似物旨在是開放式的術語。此外，除非另有明確地說明，否則片語“基於”旨在表示“至少部分地基於”。此外，如本文所使用的，術語“或”在串聯使用時旨在包括在內，並且可以與“及/或”互換使用，除非另有明確說明（例如，與“任一”或“僅其中之一”結合使用的情況）。

【符號說明】

【0100】

100	方式
102	光譜儀
104	控制裝置
150	執行光譜測量
152	主群體
154	對主群體執行光譜測量
156	光譜測量（主數據集）
158	基於主數據集產生第一光譜模型
160	第一光譜模型 (主數據集)
162	執行光譜測量
164	目標群體
166	執行光譜測量
168	光譜測量（目標數據集）
170	使用交叉驗證確定最優PLS因子
172	產生第二光譜模型

174 執行光譜測量

176 未知樣本

178 執行光譜測量

180 光譜測量

182 使用第二光譜模型對未知樣本的光譜測量執行光譜分析

200 環境

210 控制裝置

220 光譜儀

230 網路

300 裝置

310 匯流排

320 處理器

330 記憶體

340 儲存部件

350 輸入部件

360 輸出部件

370 通信介面

400 過程

410 區塊

420 區塊

430 區塊

440 區塊

450 區塊

460 區塊

500 過程

510 區塊

520 區塊

530 區塊

540 區塊

550 區塊

560 區塊

600 過程

610 區塊

620 區塊

630 區塊

640 區塊

650 區塊

660 區塊

670 區塊

111 年 07 月 13 日修正替換頁

【發明申請專利範圍】

【第1項】一種用於模型產生的裝置，包括：

一個或多個記憶體；以及

一個或多個處理器，所述一個或多個處理器被通信地耦合到所述一個或多個記憶體，所述一個或多個處理器被配置為：

接收由第一控制裝置產生之對於第一光譜模型的主數據集，

所述裝置是不同於所述第一控制裝置的第二控制裝置，以及

所述主數據集是基於由第一光譜儀執行的第一組光譜測量；

接收對於與所述第一光譜模型相關聯的目標群體的目標數據集以更新所述第一光譜模型，

所述目標數據集是基於由不同於所述第一光譜儀的第二光譜儀執行的第二組光譜測量；

產生包括所述主數據集和來自所述目標數據集的第一數據的訓練數據集；

產生包括來自所述目標數據集的第二數據而不是所述主數據集的驗證數據集；

使用交叉驗證並使用所述訓練數據集和所述驗證數據集產生作為所述第一光譜模型的更新的第二光譜模型，

相對於使用所述第一光譜模型，使用與所述第二光譜儀相關聯之所述第二光譜模型的準確度被提高；以及

提供所述第二光譜模型。

【第2項】根據請求項1所述的裝置，其中，所述一個或多個處理器被配置為：

接收光譜測量；

使用所述第二光譜模型來執行光譜測定；和

提供識別所述光譜測定的輸出。

【第3項】根據請求項1所述的裝置，其中，所述訓練數據集是複數個訓練數據集，並且所述驗證數據集是複數個驗證數據集；以及

其中，當產生所述第二光譜模型時，所述一個或多個處理器被配置成：

基於所述複數個訓練數據集和所述複數個驗證數據集產生複數個性能度量；

基於所述複數個性能度量來確定總體性能度量；

基於所述總體性能度量來確定最優部分最小平方（PLS）因子；和

基於所述最優PLS因子和合併的數據集確定所述第二光譜模型，

其中，所述合併的數據集包括所述主數據集和所述目標數據集。

【第4項】根據請求項1所述的裝置，其中，所述第一光譜模型和所述第二光譜模型是定量模型。

【第5項】根據請求項1所述的裝置，

其中，所述第一光譜儀是主光譜儀，並且

其中，所述第二光譜儀是目標光譜儀。

【第6項】一種用於模型產生的方法，包括：

藉由裝置，接收由第一控制裝置產生之對於與第一光譜模型相關聯的目標群體的目標數據集，

其中，所述裝置是不同於所述第一控制裝置的第二控制裝置；

藉由所述裝置，獲得對於所述第一光譜模型的主數據集，

其中，所述主數據集是基於由第一光譜儀執行的第一組測量，以及

其中，所述目標數據集是基於由不同於所述第一光譜儀的第二光譜儀執行的第二組測量；

藉由所述裝置，使用交叉驗證來確定最優部分最小平方（PLS）因子，

其中，所述最優PLS因子基於複數個訓練數據集且基於複數個驗證數據集來確定，每個訓練數據集包括所述目標數據集的各別部分和所述主數據集的全部，每個驗證數據集包括所述目標數據集的各別部分並且不包括所述主數據集的數據；

藉由所述裝置，合併所述目標數據集和所述主數據集，以產生合併的數據集；

藉由所述裝置並使用所述合併的數據集和所述最優PLS因子，產生第二光譜模型，

其中，所述第二光譜模型是所述第一光譜模型的更新，以及

其中，相對於使用所述第一光譜模型，使用與所述第二光譜儀相關聯之所述第二光譜模型的準確度被提高；以及

藉由所述裝置，提供所述第二光譜模型來代替所述第一光譜模型。

【第7項】根據請求項6所述的方法，其中，確定所述最優PLS因子包括：

確定對於所述複數個訓練數據集中的每一個訓練數據集和所述複數個驗證數據集中的每一個驗證數據集的部分最小平方（PLS）性能度量；

基於所述PLS性能度量確定總體PLS性能度量；和

基於所述總體PLS性能度量，優化PLS因子以確定對於所述第二光譜模型的所述最優PLS因子。

【第8項】根據請求項7所述的方法，其中，所述總體PLS性能度量與均方根誤差（RMSE）值相關聯，並且

其中，優化所述PLS因子包括：

優化所述PLS因子以最小化所述RMSE值。

【第9項】根據請求項7所述的方法，其中，所述複數個驗證數據集包括與

所述複數個訓練數據集不同的所述目標數據集中的數據。

【第10項】根據請求項7所述的方法，其中，確定所述總體PLS性能度量包括：

聚合所述PLS性能度量。

【第11項】根據請求項6所述的方法，其中，所述第二組測量是在所述第二組測量之後執行。

【第12項】根據請求項6所述的方法，其中，所述第二光譜模型是所述第一光譜模型的校準傳遞模型。

【第13項】根據請求項6所述的方法，其中，提供所述第二光譜模型包括：

提供所述第二光譜模型以用於與藉由所述第二光譜儀進行的後續測量結合使用。

【第14項】一種非暫時性電腦可讀取媒體，所述非暫時性電腦可讀取媒體儲存指令，所述指令包括：

一個或多個指令，所述一個或多個指令當由一個或多個處理器執行時使所述一個或多個處理器：

接收由第一控制裝置產生之對於第一光譜模型的主數據集，

其中，所述一個或多個處理器是在不同於所述第一控制裝置的第二控制裝置裡，以及

所述主數據集是基於由第一光譜儀執行的第一組光譜測量；

接收對於與所述第一光譜模型相關聯的目標群體的目標數據集以更新所述第一光譜模型，

其中，所述目標數據集是基於由不同於所述第一光譜儀的第二光譜儀執行的第二組光譜測量；

基於所述主數據集和所述目標數據集產生複數個訓練數據集；

基於所述目標數據集產生複數個驗證數據集，

其中，所述複數個驗證數據集不包括所述主數據集的數據；

基於所述複數個訓練數據集和所述複數個驗證數據集並使用交叉驗證來確定模型設定；

基於所述模型設定、所述目標數據集和所述主數據集，產生第二光譜模型，

其中，相對於使用所述第一光譜模型，使用與所述第二光譜儀相關聯之所述第二光譜模型的準確度被提高；和

提供所述第二光譜模型。

【第15項】根據請求項14所述的非暫時性電腦可讀取媒體，其中，所述模型設定是以下中的至少一項，

部分最小平方（PLS）模型的PLS因子，

主成分回歸（PCR）模型的成分的數量，

支持向量回歸（SVR）模型的SVR參數，或

預處理設定。

【第16項】根據請求項14所述的非暫時性電腦可讀取媒體，其中，使所述一個或多個處理器確定所述模型設定的所述一個或多個指令使所述一個或多個處理器：

產生關於所述複數個訓練數據集中的每一個訓練數據集和所述複數個驗證數據集中的對應的驗證數據集的複數個部分性能度量；

聚合所述複數個部分性能度量以產生總體性能度量；和

確定所述模型設定，以最小化所述總體性能度量的誤差值。

【第17項】根據請求項14所述的非暫時性電腦可讀取媒體，其中，所述一個或多個指令在由所述一個或多個處理器執行時使所述一個或多個處理器：

接收測量；

基於所述測量並使用所述第二光譜模型來執行光譜測定；和

提供識別所述光譜測定的輸出。

【第18項】根據請求項14所述的非暫時性電腦可讀取媒體，其中，所述第一光譜模型和所述第二光譜模型是定量模型。

【第19項】根據請求項14所述的非暫時性電腦可讀取媒體，其中，使所述一個或多個處理器提供所述第二光譜模型的所述一個或多個指令使所述一個或多個處理器：

提供所述第二光譜模型以用於與藉由所述第二光譜儀進行的後續測量結合使用。

【第20項】根據請求項14所述的非暫時性電腦可讀取媒體，其中，使所述一個或多個處理器接收所述目標數據集的所述一個或多個指令使所述一個或多個處理器：

接收來自所述第二光譜儀的所述目標數據集。

【發明圖式】

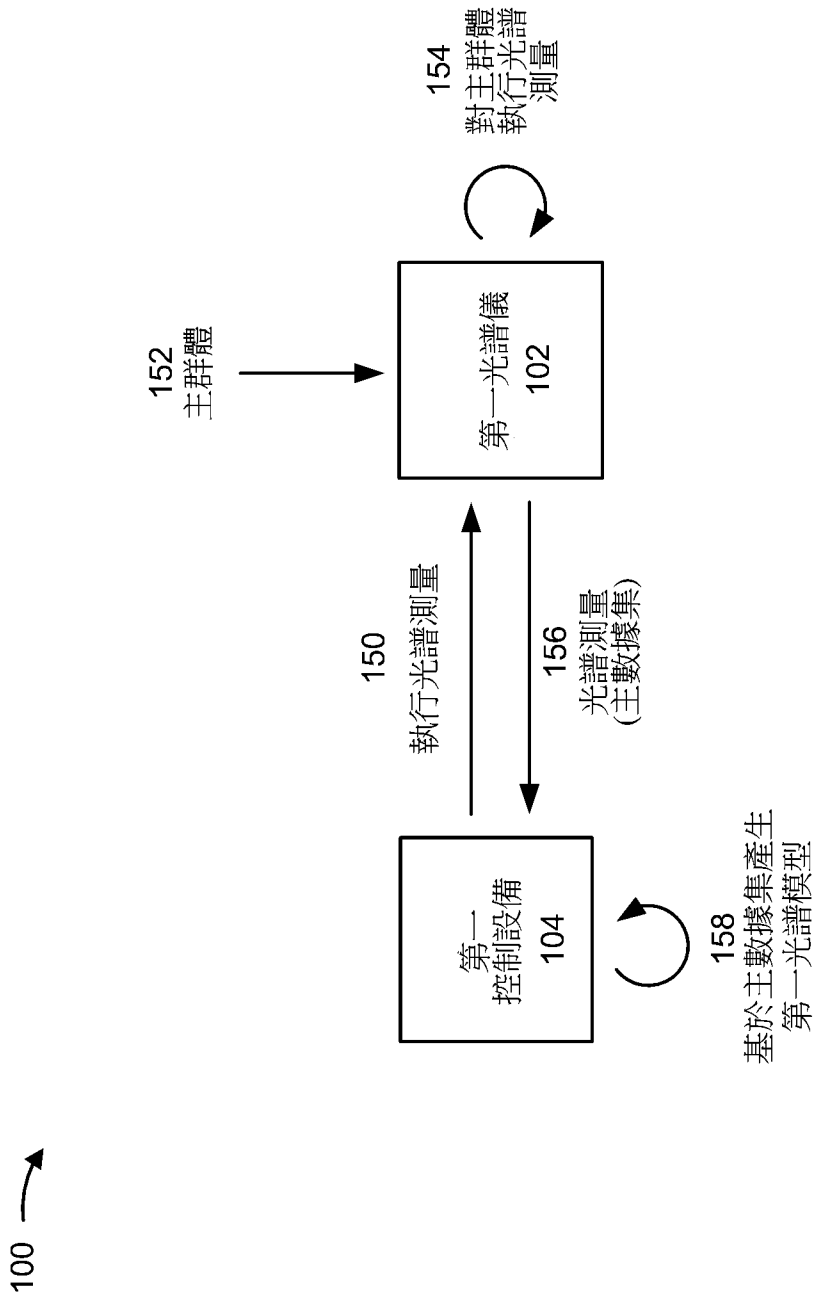

100
152
主群體
第一光譜儀
102
154
對主群體
執行光譜
測量
150
執行光譜測量
156
光譜測量
(主數據集)
第一
控制設備
104
158
基於主數據集產生
第一光譜模型

圖1A

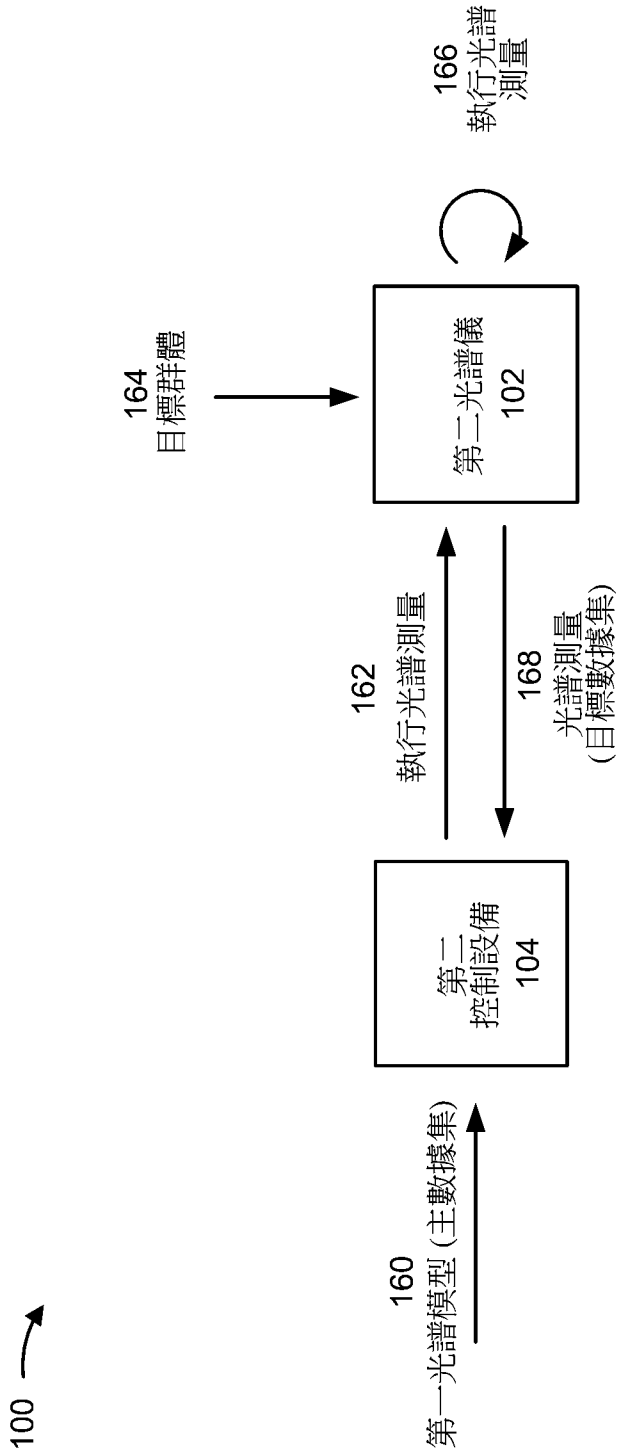

100
160
第一光譜模型 (主數據集)
第二
控制設備
104
162
執行光譜測量
168
光譜測量
(目標數據集)
第二光譜儀
102
164
目標群體
166
執行光譜
測量

圖1B

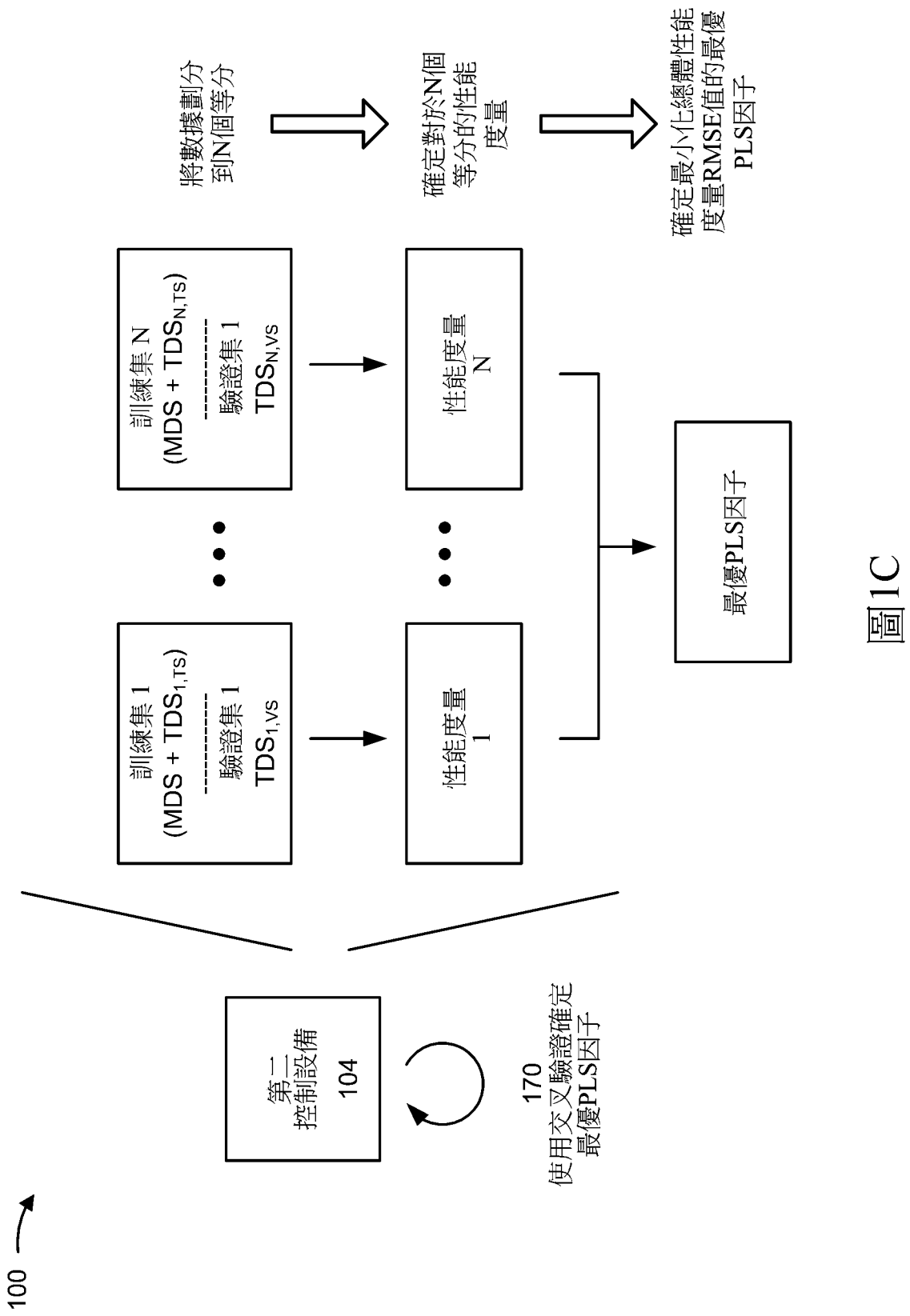

100
第二控制設備 104
170 使用交叉驗證確定最優PLS因子
訓練集 1 (MDS + TDS1,TS)
驗證集 1 TDS1,VS
訓練集 N (MDS + TDSN,TS)
驗證集 1 TDSN,VS
性能度量 1
性能度量 N
最優PLS因子
將數據劃分到N個等分
確定對於N個等分的性能度量
確定最小化總體性能度量RMSE值的最優PLS因子

圖1C

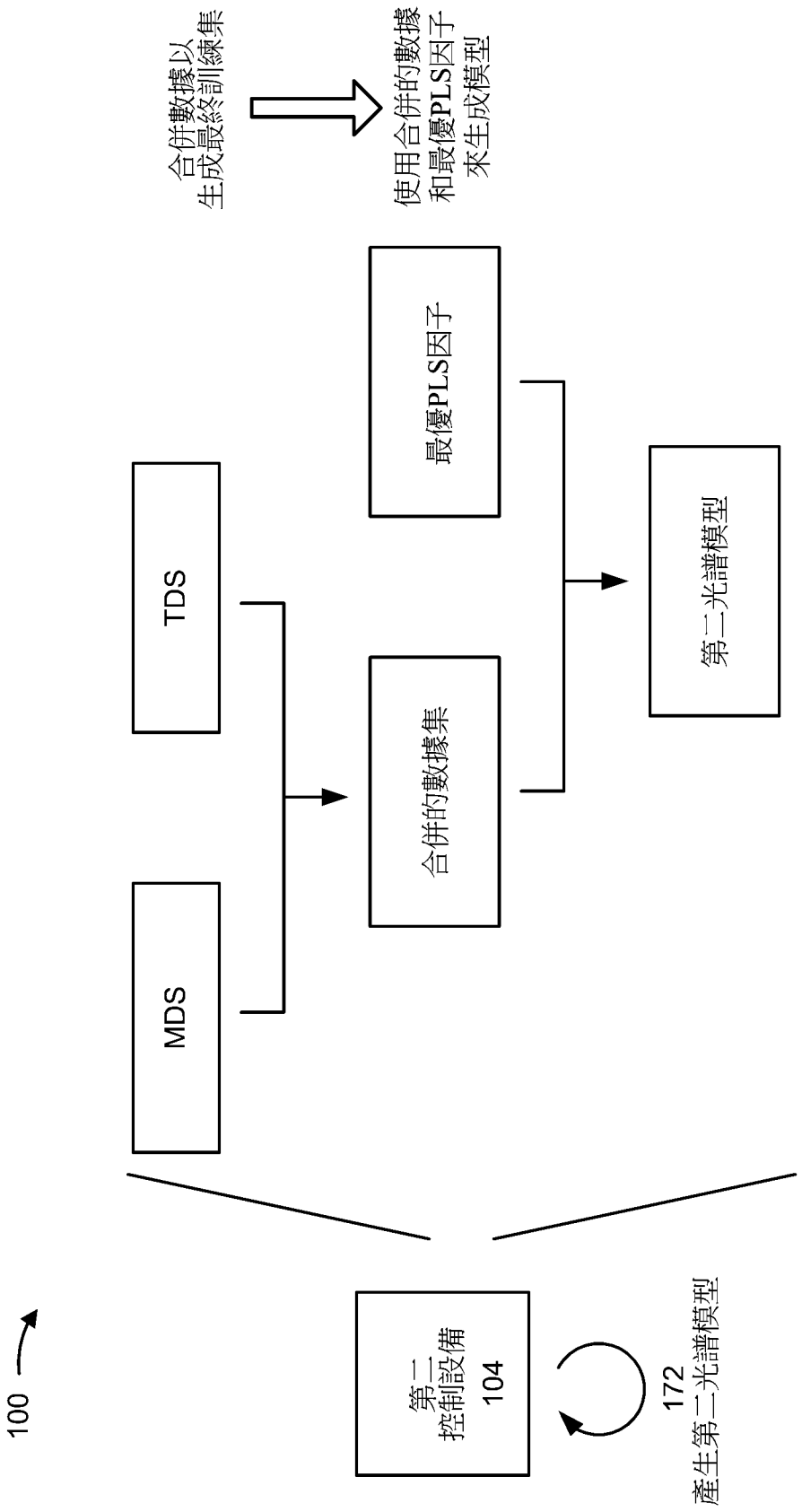

100
第二
控制設備
104
172
產生第二光譜模型
MDS
TDS
合併的數據集
最優PLS因子
第二光譜模型
合併數據以
生成最終訓練集
使用合併的數據
和最優PLS因子
來生成模型

圖1D

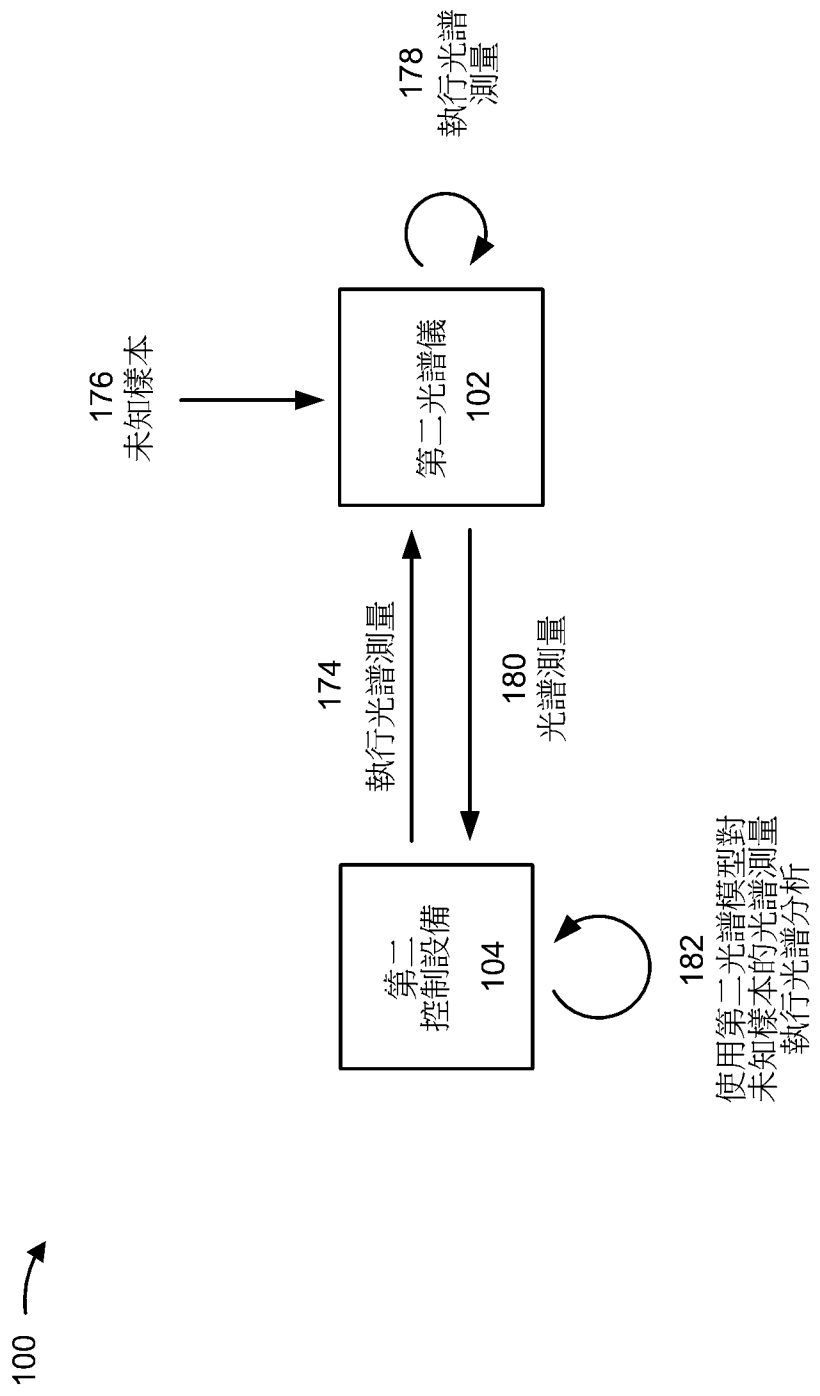

100
176
未知樣本
第二光譜儀
102
178
執行光譜
測量
174
執行光譜測量
180
光譜測量
第二
控制設備
104
182
使用第二光譜模型對
未知樣本的光譜測量
執行光譜分析

圖1E

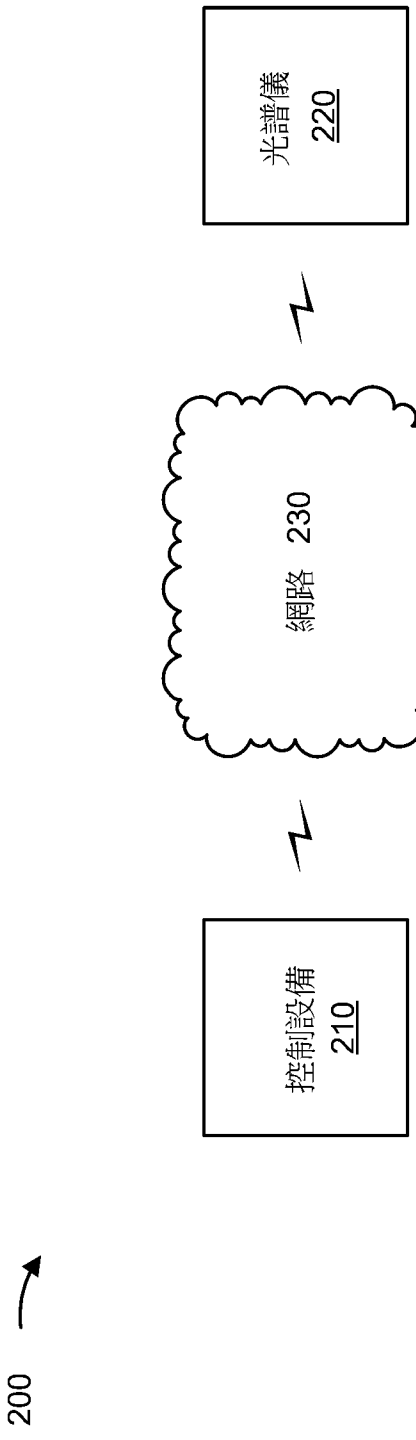


圖2

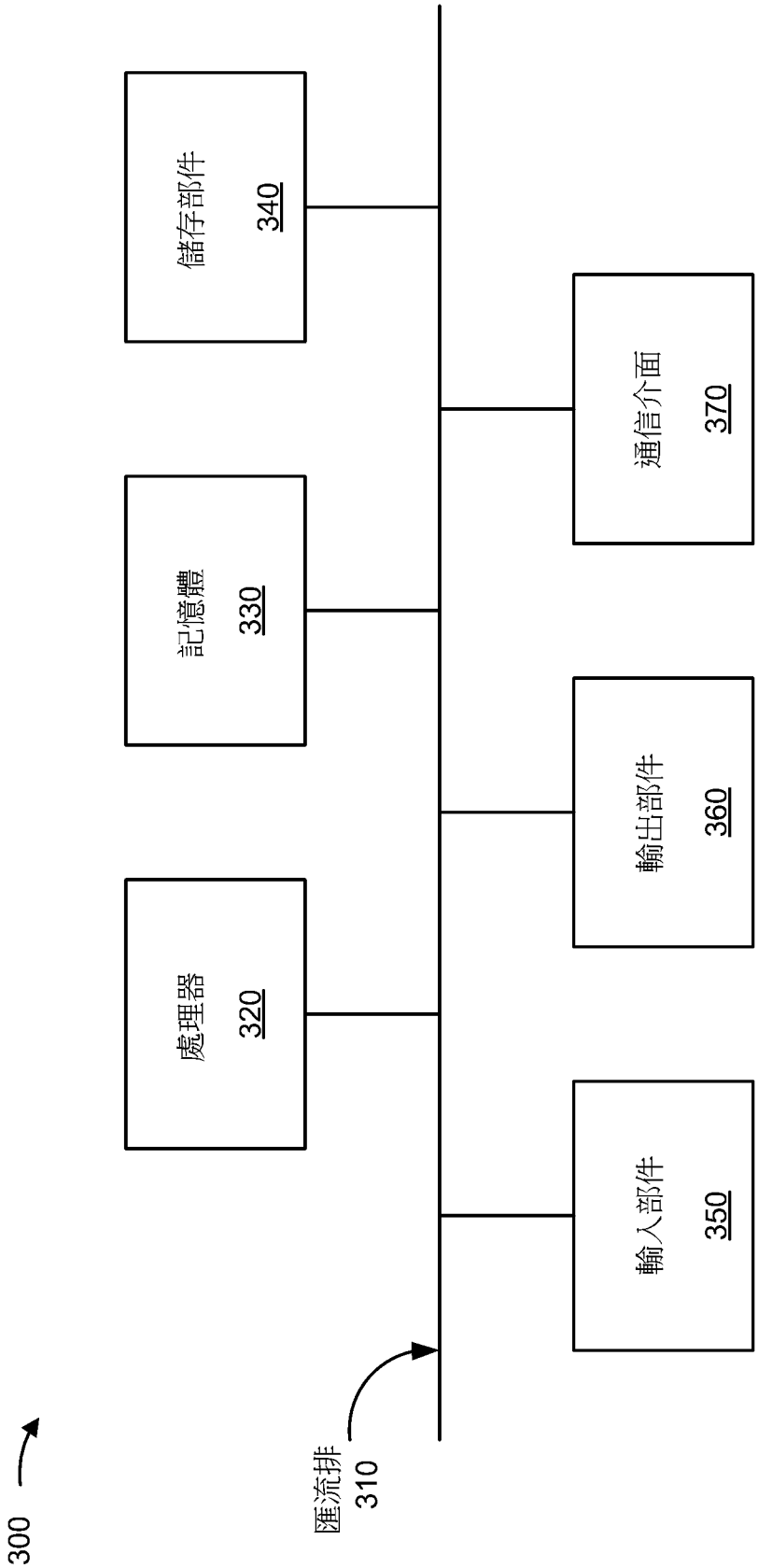


圖3

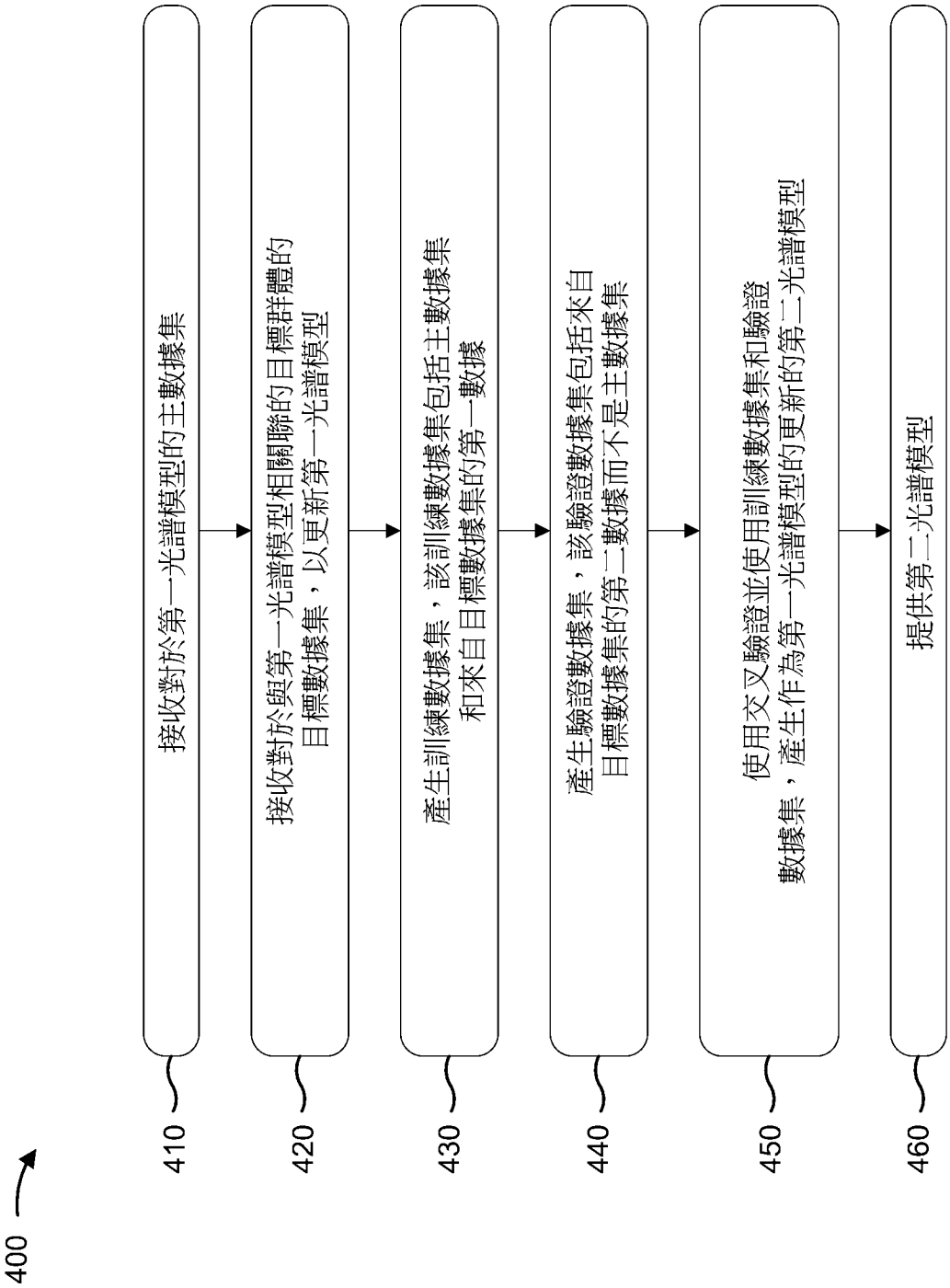

400
410
接收對於第一光譜模型的主數據集
420
接收對於與第一光譜模型相關聯的目標群體的
目標數據集，以更新第一光譜模型
430
產生訓練數據集，該訓練數據集包括主數據集
和來自目標數據集的第一數據
440
產生驗證數據集，該驗證數據集包括來自
目標數據集的第二數據而不是主數據集
450
使用交叉驗證並使用訓練數據集和驗證
數據集，產生作為第一光譜模型的更新的第二光譜模型
460
提供第二光譜模型

圖4

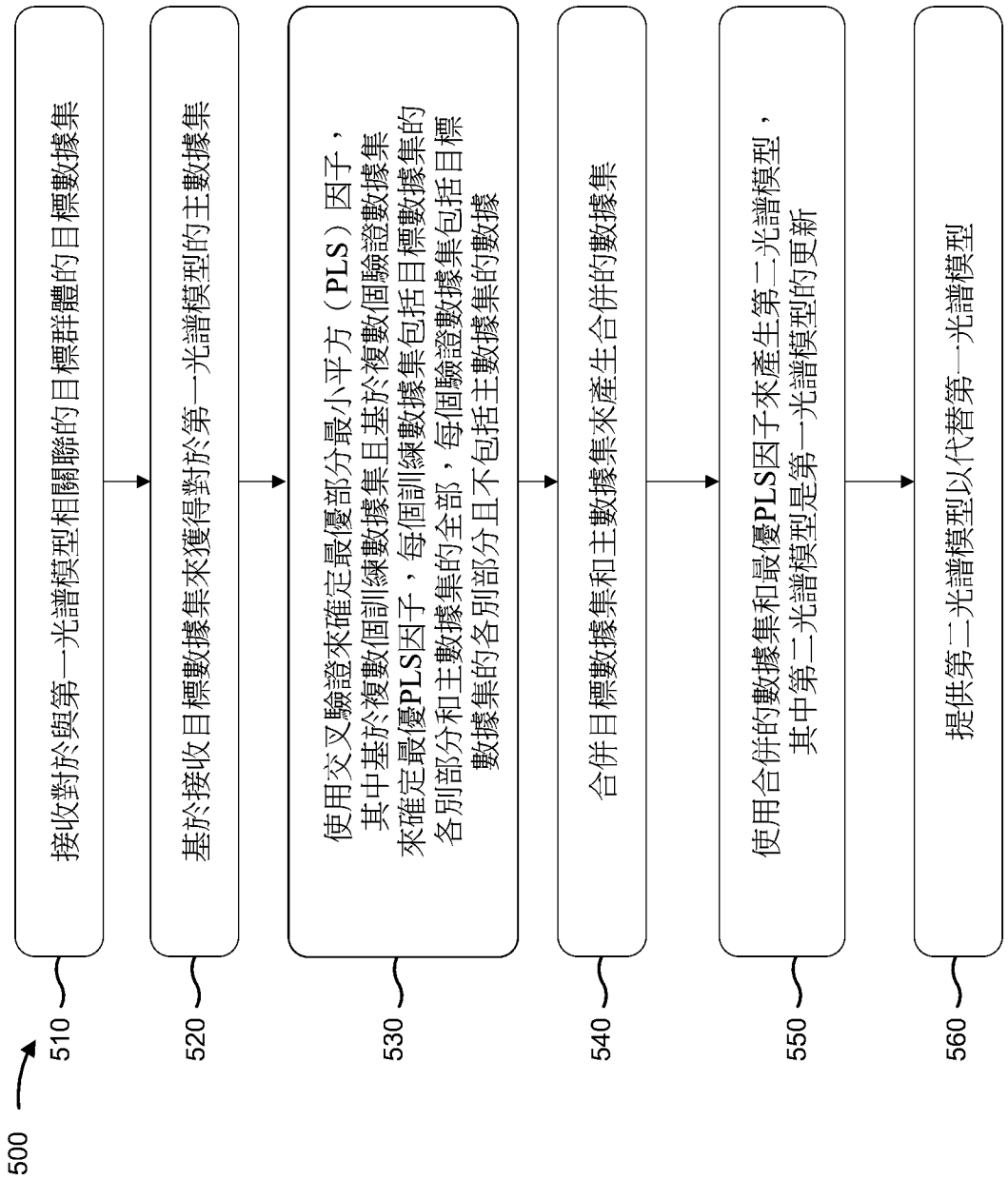

500
510
接收對於與第一光譜模型相關聯的目標群體的目標數據集
520
基於接收目標數據集來獲得對於第一光譜模型的主數據集
530
使用交叉驗證來確定最優部分最小平方（PLS）因子，其中基於複數個訓練數據集且基於複數個驗證數據集來確定最優PLS因子，每個訓練數據集包括目標數據集的各別部分和主數據集的全部，每個驗證數據集包括目標數據集的各別部分且不包括主數據集的數據
540
合併目標數據集和主數據集來產生合併的數據集
550
使用合併的數據集和最優PLS因子來產生第二光譜模型，其中第二光譜模型是第一光譜模型的更新
560
提供第二光譜模型以代替第一光譜模型

圖5

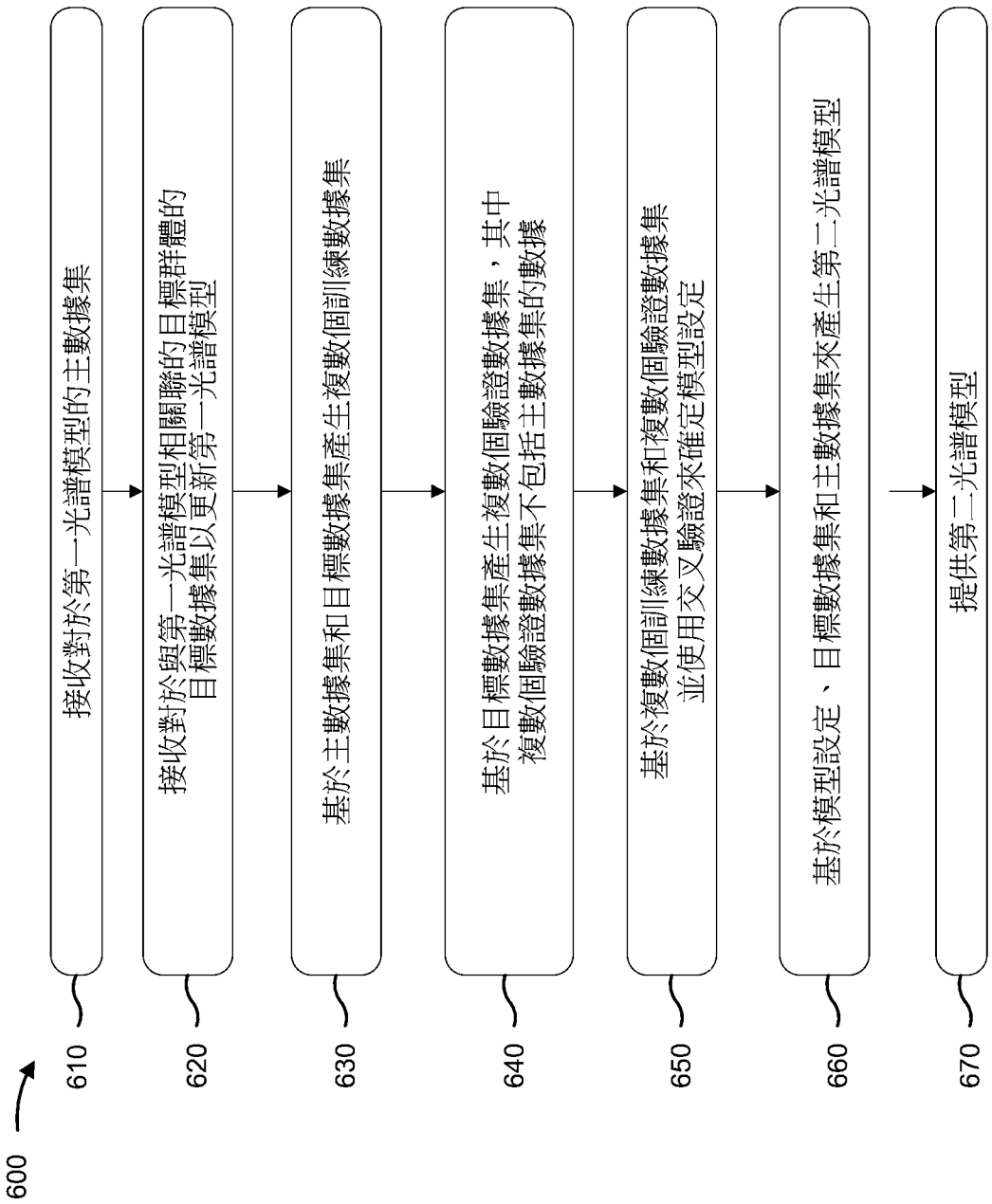

600
610
接收對於第一光譜模型的主數據集
620
接收對於與第一光譜模型相關聯的目標群體的目標數據集以更新第一光譜模型
630
基於主數據集和目標數據集產生複數個訓練數據集
640
基於目標數據集產生複數個驗證數據集，其中複數個驗證數據集不包括主數據集的數據
650
基於複數個訓練數據集和複數個驗證數據集並使用交叉驗證來確定模型設定
660
基於模型設定、目標數據集和主數據集來產生第二光譜模型
670
提供第二光譜模型

圖6