



(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 I759577 B

(45)公告日：中華民國 111 (2022) 年 04 月 01 日

(21)申請案號：108101270

(22)申請日：中華民國 108 (2019) 年 01 月 11 日

(51)Int. Cl. : G06K9/62 (2006.01)

G01C3/00 (2006.01)

(30)優先權：2018/01/26 美國

62/622,641

2018/07/13 美國

16/034,901

(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, CHANGMENG (US)

(74)代理人：閻啓泰；林景郁

(56)參考文獻：

TW 201709110A

Sun, L., Hsiung, C., Pederson, C. G., Zou, P., Smith, V., von Gunten, M., & O'Brien, N. A, Pharmaceutical Raw Material Identification Using Miniature Near-Infrared (MicroNIR) Spectroscopy and Supervised Pattern Recognition Using Support Vector Machine, Applied Spectroscopy, 西元 2016 年, 第 70 卷第 5 期, Pages 816~825。

審查人員：黃偉程

申請專利範圍項數：20 項 圖式數：7 共 50 頁

(54)名稱

用於在光譜量化期間避免誤報識別的設備和方法以及相關的非暫時性電腦可讀取媒體

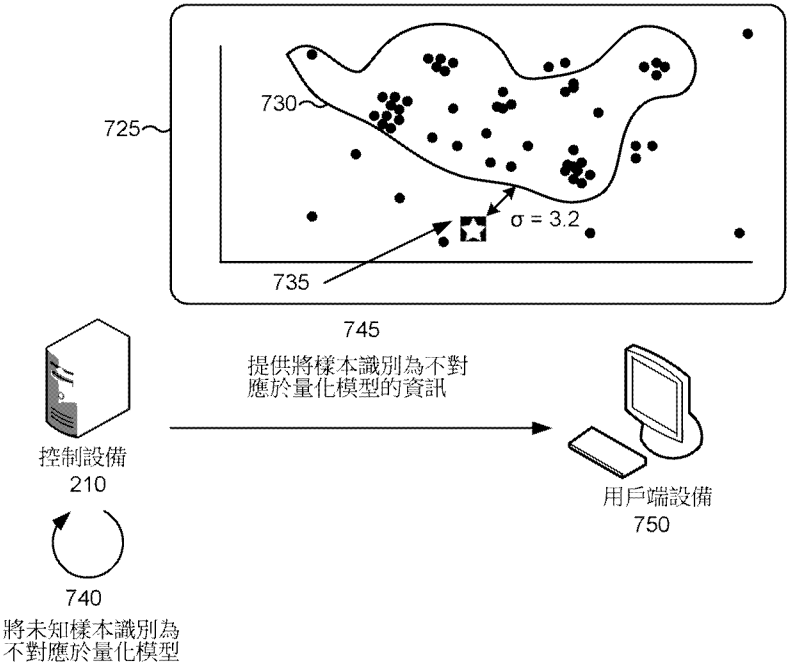
(57)摘要

一種設備，其可以接收識別對未知樣本執行的光譜測量的結果的資訊。設備可以基於可配置參數確定用於量化模型的決策邊界，使得量化模型的第一多個訓練集樣本在決策邊界內，並且量化模型的第二多個訓練集樣本不在決策邊界內。設備可以確定對未知樣本執行的光譜測量相對於決策邊界的距離度量。設備可以確定量化模型的第二多個訓練集樣本相對於決策邊界的多個距離度量。設備可以提供指示對未知樣本執行的光譜測量是否對應於量化模型的資訊。

A device may receive information identifying results of a spectroscopic measurement performed on an unknown sample. The device may determine a decision boundary for a quantification model based on a configurable parameter, such that a first plurality of training set samples of the quantification model is within the decision boundary and a second plurality of training set samples of the quantification model is not within the decision boundary. The device may determine a distance metric for the spectroscopic measurement performed on the unknown sample relative to the decision boundary. The device may determine a plurality of distance metrics for the second plurality of training set samples of the quantification model relative to the decision boundary. The device may provide information indicating whether the spectroscopic measurement performed on the unknown sample corresponds to the quantification model.

指定代表圖：

700 →



符號簡單說明：

- 210:控制設備
- 700:實施方式
- 725:量化模型
- 730:單個類別
- 735:樣本
- 740:將未知樣本識別為不對應於量化模型
- 745:提供將樣本識別為不對應於量化模型的資訊
- 750:用戶端設備

圖7B



I759577

【發明摘要】

【中文發明名稱】 用於在光譜量化期間避免誤報識別的設備和方法以及相關的非暫時性電腦可讀取媒體

【英文發明名稱】 DEVICE AND METHOD FOR AVOIDING FALSE
POSITIVE IDENTIFICATION DURING
SPECTROSCOPIC QUANTIFICATION AND RELEVANT
NON-TRANSITORY COMPUTER-READABLE
MEDIUM

【中文】

一種設備，其可以接收識別對未知樣本執行的光譜測量的結果的資訊。設備可以基於可配置參數確定用於量化模型的決策邊界，使得量化模型的第一多個訓練集樣本在決策邊界內，並且量化模型的第二多個訓練集樣本不在決策邊界內。設備可以確定對未知樣本執行的光譜測量相對於決策邊界的距離度量。設備可以確定量化模型的第二多個訓練集樣本相對於決策邊界的多個距離度量。設備可以提供指示對未知樣本執行的光譜測量是否對應於量化模型的資訊。

【英文】

A device may receive information identifying results of a spectroscopic measurement performed on an unknown sample. The device may determine a decision boundary for a quantification model based on a configurable parameter, such that a first plurality of training set samples of the quantification model is within the decision boundary and a second plurality of training set samples of the quantification

model is not within the decision boundary. The device may determine a distance metric for the spectroscopic measurement performed on the unknown sample relative to the decision boundary. The device may determine a plurality of distance metrics for the second plurality of training set samples of the quantification model relative to the decision boundary. The device may provide information indicating whether the spectroscopic measurement performed on the unknown sample corresponds to the quantification model.

【指定代表圖】 圖7B

【代表圖之符號簡單說明】

- 210 控制設備
- 700 實施方式
- 725 量化模型
- 730 單個類別
- 735 樣本
- 740 將未知樣本識別為不對應於量化模型
- 745 提供將樣本識別為不對應於量化模型的資訊
- 750 用戶端設備

【特徵化學式】

無

【發明說明書】

【中文發明名稱】 用於在光譜量化期間避免誤報識別的設備和方法以及相關的非暫時性電腦可讀取媒體

【英文發明名稱】 DEVICE AND METHOD FOR AVOIDING FALSE
POSITIVE IDENTIFICATION DURING
SPECTROSCOPIC QUANTIFICATION AND RELEVANT
NON-TRANSITORY COMPUTER-READABLE
MEDIUM

【技術領域】

【0001】 本發明大致涉及用於光譜量化的減少的誤報識別。

【先前技術】

【0002】 原料識別可用於藥物產品的質量控制。例如，可以對醫用材料執行原料識別，以確定醫用材料的成分組成物是否對應於與醫用材料相關聯的包裝標籤。類似地，可以執行原料量化以確定特定樣本中特定成分的濃度。例如，可以執行原料量化以確定藥物中活性成分的濃度。與其他化學計量技術相比，光譜法可以有助於以減少的準備和資料獲取時間進行原料的無損識別及/或量化。

【發明內容】

【0003】 根據一些可能的實施方式，一種設備可以包括通信地耦合到一個或多個處理器的一個或多個記憶體。一個或多個記憶體和一個或多個處理器可以被配置為接收識別對未知樣本執行的光譜測量的結果的資訊。一個或多個

記憶體和一個或多個處理器可以被配置為基於可配置參數來確定用於量化模型的決策邊界，使得量化模型的第一多個訓練集樣本在決策邊界內，並且量化模型的第二多個訓練集樣本不在決策邊界內。一個或多個記憶體和一個或多個處理器可以被配置為確定對未知樣本執行的光譜測量相對於決策邊界的距離度量。一個或多個記憶體和一個或多個處理器可以被配置為確定量化模型的第二多個訓練集樣本相對於決策邊界的多個距離度量。一個或多個記憶體和一個或多個處理器可以被配置為基於光譜測量的距離度量和第二多個訓練集樣本的多個距離度量來確定對未知樣本執行的光譜測量是否對應於量化模型。一個或多個記憶體和一個或多個處理器可以被配置為提供指示對未知樣本執行的光譜測量是否對應於量化模型的資訊。

【0004】 根據一些可能的實施方式，一種非暫時性電腦可讀取媒體可以儲存一個或多個指令。當由一個或多個處理器執行時，一個或多個指令可以使一個或多個處理器獲得與特定類型的感興趣的材料相關的量化模型。量化模型可以被配置用於確定特定類型的感興趣的材料的樣本中的特定成分的濃度。當由一個或多個處理器執行時，一個或多個指令可以使一個或多個處理器接收識別對未知樣本執行的特定光譜測量的結果的資訊。當由一個或多個處理器執行時，一個或多個指令可以使一個或多個處理器將量化模型的訓練集樣本的其他光譜測量聚集到用於量化模型的單個類別中。當由一個或多個處理器執行時，一個或多個指令可以使一個或多個處理器將訓練集樣本的其他光譜測量細分為第一組和第二組。第一組其他光譜測量可以在決策邊界內。第二組其他光譜測量可以不在決策邊界內。當由一個或多個處理器執行時，一個或多個指令可以使一個或多個處理器確定對未知樣本執行的特定光譜測量的度量相對於第二組其他光譜測量的相應的度量滿足臨界值。當由一個或多個處理器執行時，一個或多個指令可以使一個或多個處理器提供指示未知樣本不是特定類型的感興趣

的材料的資訊。

【0005】 根據一些可能的實施方式，一種方法可以包括由設備接收識別對未知樣本執行的近紅外光（NIR）光譜測量的結果的資訊。方法可以包括由設備確定用於量化模型的決策邊界，其中，決策邊界將量化模型的單個類別劃分為在決策邊界內的量化模型的第一多個訓練集樣本以及不在決策邊界內的量化模型的第二多個訓練集樣本。方法可以包括由設備確定對未知樣本執行的NIR光譜測量的特定距離度量相對於第二多個訓練集樣本的其他距離度量滿足臨界值。方法可以包括由設備基於確定對未知樣本執行的NIR光譜測量的特定距離度量相對於第二多個訓練集樣本的其他距離度量滿足臨界值，提供指示對未知樣本執行的NIR光譜測量不對應於量化模型的資訊。

【圖式簡單說明】

【0006】 圖1A和圖1B是本文中所述的示例實施方式的概況的圖式；

【0007】 圖2是本文中所述的系統及/或方法可以在其中被實現的示例環境的圖式；

【0008】 圖3是圖2的一個或多個設備的示例元件的圖式；

【0009】 圖4是產生用於光譜量化的量化模型的示例流程的流程圖；

【0010】 圖5是與圖4中所示的示例流程相關的示例實施方式的圖式；

【0011】 圖6是用於在光譜量化期間避免誤報（false positive）識別的示例流程的流程圖；以及

【0012】 圖7A和圖7B是與圖6中所示的示例流程有關的示例實施方式的圖式。

【實施方式】

【0013】 以下針對示例實施方式的詳細描述係參考附圖。不同附圖中的相同元件符號可以指代相同或類似的元件。

【0014】 原料識別（**RMID**）是用於識別特定樣本的成分（例如，組成物）以用於識別、驗證等的技術。例如，**RMID**可用於驗證藥物材料中的組成物是否對應於標籤上確定的一組組成物。類似地，原料量化是用於對特定樣本執行量化分析的技術，例如確定特定樣本中特定成分材料的濃度。光譜儀（**spectrometer**）可用於對樣本（例如，藥物材料）執行光譜法，以確定樣本的成分、樣本的成分的濃度等。光譜儀可確定樣本的一組測量，並可提供一組測量以用於光譜測定。光譜分類技術（例如，分類器）可以有助於基於樣本的一組測量來確定樣本的成分或樣本的成分的濃度。

【0015】 然而，一些要進行光譜量化的未知樣本實際上不被包括在量化模型用以量化的材料的類別中。例如，對於被訓練來確定魚肉的樣本中的特定類型的蛋白質的濃度的量化模型，使用者可能無意中提供了牛肉的樣本用於量化。在該情況下，控制設備可能對牛肉的樣本執行光譜量化，並且可能提供該牛肉的樣本作為具有特定濃度的特定類型的蛋白質的識別。然而，由於牛肉和魚肉及其蛋白質的光譜特徵之間的差異，識別可能不準確，並且可以被稱為誤報識別。

【0016】 作為另一示例，量化模型可以被訓練以量化不同類型的糖（例如，葡萄糖、果糖、半乳糖等）和在未知樣本中的相對濃度。然而，光譜儀和控制設備的使用者可能無意中試圖基於不正確地使用光譜儀執行測量來對糖的未知樣本進行分類。例如，使用者可能在離未知樣本不正確的距離處、在不同於校準條件（其中執行光譜法以訓練量化模型）的環境條件下、及/或類似的情況下操作光譜儀，導致不正確地獲得測量。在該情況下，當未知樣本實際上包括第二濃度下的第二類型的糖時，控制設備可能接收對於未知樣本的不準確光

譜，導致未知樣本作為具有第一濃度下的第一類型的糖的誤報識別。

【0017】 本文中所述的一些實施方式可以使用單類支持向量機（SC-SVM）技術來減少光譜量化中的誤報識別的可能性。例如，接收未知樣本的光譜測量的控制設備可以確定未知樣本的光譜測量是否對應於光譜模型用以量化的材料的類別。在一些實施方式中，控制設備可以確定未知樣本不與光譜模型用以量化的材料的類別相關聯，並且可以提供指示未知樣本不與該材料的類別相關聯的資訊，從而避免未知樣本的誤報識別。可替代地，基於確定未知樣本與光譜模型用以量化的材料的類別相關聯，控制設備可以分析未知樣本的光譜以提供例如對濃度、分類等的光譜測定。此外，控制設備可以利用置信度量（諸如機率估計、決策值等）來濾除誤報識別。

【0018】 以這種方式，光譜法的準確度相對於在不識別潛在錯誤樣本（例如，與光譜模型未被配置的材料的類別相關聯的樣本或光譜測量被不正確地獲得的樣本）及/或沒有置信度量度的情況下執行的光譜法而得到提高。此外，當基於已知光譜樣本的訓練集產生量化模型時，可以使用材料是否與光譜模型被配置的類別相關聯的測定。例如，控制設備可以確定訓練集的樣本不是對應於訓練集的其餘部分的類型（例如，基於導致錯誤樣本被引入訓練集的人為誤差），並且可以確定在產生量化模型時不包括關於該樣本的資料。以這種方式，控制設備提高了用於光譜法的量化模型的準確度。

【0019】 圖1A和圖1B是本文中所述的示例實施方式100的概況的圖式。如圖1A所示，示例實施方式100可以包括控制設備和光譜儀。

【0020】 如圖1A中進一步示出的，控制設備可以使光譜儀對訓練集和驗證集（例如，用於分類模型的訓練和驗證的已知樣本集）執行一組光譜測量。訓練集和驗證集可以被選擇以包括要為其訓練量化模型的成分的臨界值數量的樣本。成分可以出現在其中並可用於訓練量化模型的材料可以被稱為感興趣的

材料。在該情況下，訓練集和驗證集可以包括例如，表示感興趣的材料的第一濃度的第一組樣本、表示感興趣的材料第二濃度的第二組樣本等，以使得能夠訓練量化模型來識別未知樣本中感興趣的材料濃度。

【0021】 如圖1A中進一步示出的，光譜儀可以基於從控制設備接收到指令，對訓練集和驗證集執行一組光譜測量。例如，光譜儀可以確定關於訓練集和驗證集的每個樣本的光譜，以使得控制設備能夠產生一組類別，以用於將未知樣本分類為用於量化模型的感興趣的材料之一。

【0022】 光譜儀可以向控制設備提供一組光譜測量。控制設備可以使用特定的測定技術並基於該組光譜測量來產生量化模型。例如，控制設備可以使用支持向量機（SVM）技術（例如，用於資訊測定的機器學習技術），諸如單類SVM（SC-SVM）技術，來產生量化模型。量化模型可以包括與將特定光譜分配給感興趣的材料特定濃度的成分（例如，感興趣的材料中的特定濃度水平的成分）相關聯的資訊。以這種方式，控制設備可以基於將未知樣本的光譜分配給對應於特定濃度的量化模型的濃度的特定類別來提供識別未知樣本中的成分的濃度的資訊。

【0023】 如圖1B所示，控制設備可以（例如，從儲存器、從產生量化模型的另一控制設備等）接收量化模型。控制設備可以使光譜儀對未知樣本（例如，要對其執行分類或量化的未知樣本）執行一組光譜測量。光譜儀可以基於從控制設備接收到指令來執行一組光譜測量。例如，光譜儀可以確定未知樣本的光譜。光譜儀可以向控制設備提供一組光譜測量。控制設備可以試圖基於量化模型量化未知樣本（例如，將未知樣本分類到與未知樣本中的特定成分的特定濃度或特定量相關聯的特定類別中）。例如，控制設備可以嘗試確定未知樣本（例如，藥丸的樣本）內布洛芬的特定濃度、未知樣本（例如，糖基產品的樣本）內葡萄糖的單位的特定量等。

【0024】 關於圖1B，控制設備可以嘗試確定未知樣本是否對應於量化模型。例如，控制設備可以確定對應於未知樣本屬於感興趣的材料（例如，以對於其使用訓練集和驗證集配置量化模型的一組濃度中的任何濃度）的可能性的置信度度量。作為示例，對於被配置為識別布洛芬藥丸的樣本內布洛芬的濃度的量化模型，控制設備可以確定未知樣本是否是布洛芬藥丸（而不是另一類型的藥丸，例如對乙醯氨基酚藥丸、乙醯水楊酸藥丸等）。作為另一示例，對於被配置為識別魚肉中的鹽的濃度的量化模型，控制設備可以確定未知樣本是否是魚肉（而不是雞肉、牛肉、豬肉等）。

【0025】 在該情況下，基於控制設備確定置信度度量（諸如機率估計、支持向量機的決策值輸出等）滿足臨界值（例如，如本文中所述的標準偏差臨界值），控制設備可以確定未知樣本不是感興趣的材料（例如，這可以對應於未知樣本是不同的材料、未知樣本的光譜測量被錯誤地執行等）。在該情況下，控制設備可以報告未知樣本不能夠使用量化模型被準確地量化，從而降低未知樣本經受未知樣本屬感興趣的材料中的特定濃度的成分的誤報識別的可能性。

【0026】 以這種方式，基於降低報告未知樣本作為感興趣的材料中的特定濃度的成分的誤報識別的可能性，控制設備相對於其他量化模型以提高的準確度實現了對於未知樣本的光譜法。

【0027】 如上所指示，圖1A和圖1B僅作為示例被提供。其它示例是可能的，並且可以不同於關於圖1A和圖1B描述的示例。

【0028】 圖2是本文中所述的系統及/或方法可以在其中被實現的示例環境200的圖。如圖2所示，環境200可以包括控制設備210、光譜儀220和網路230。環境200的設備可經由有線連接、無線連接或有線連接和無線連接的組合來互連。

【0029】 控制設備210包括能夠儲存、處理及/或路由與光譜量化相關聯的資訊的一個或多個設備。例如，控制設備210可以包括伺服器、電腦、可穿戴設備、雲端計算設備、及/或基於訓練集的一組測量產生量化模型、基於驗證集的一組測量驗證量化模型及/或利用量化模型來基於未知集的一組測量執行光譜量化的類似設備。在一些實施方式中，如本文中所述，控制設備210可以利用機器學習技術來確定未知樣本的光譜測量是否將被分類為不對應於用於量化模型的感興趣的材料。在一些實施方式中，控制設備210可以與特定光譜儀220相關聯。在一些實施方式中，控制設備210可以與多個光譜儀220相關聯。在一些實施方式中，控制設備210可以從環境200中的另一設備（例如光譜儀220）接收資訊及/或向環境200中的另一設備（例如光譜儀220）傳輸資訊。

【0030】 光譜儀220包括能夠對樣本執行光譜測量的一個或多個設備。例如，光譜儀220可以包括執行光譜法（例如，振動光譜法，諸如近紅外光（NIR）光譜儀、中紅外光光譜法（mid-IR）、拉曼（Raman）光譜法等）的光譜儀設備。在一些實施方式中，光譜儀220可以被結合到可穿戴設備中，例如可穿戴光譜儀及/或類似設備。在一些實施方式中，光譜儀220可以從環境200中的另一設備（例如控制設備210）接收資訊及/或向環境200中的另一設備（例如控制設備210）傳輸資訊。

【0031】 網路230可包括一個或多個有線網路及/或無線網路。例如，網路230可包括蜂巢式網路（例如，長期演進（LTE）網路、3G網路、分碼多重存取（CDMA）網路等）、公用陸上行動網路（PLMN）、區域網路（LAN）、廣域網路（WAN）、都會區網路（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、雲端計算網路等、及/或這些或其它類型的網路的組合。

【0032】 圖2中顯示的設備和網路的數量和佈置作為示例被提供。實際

上，與圖2中顯示的那些設備及/或網路相比，可以有附加的設備及/或網路、更少的設備及/或網路、不同的設備及/或網路、或不同地佈置的設備及/或網路。此外，圖2中顯示的兩個或更多個設備可在單個設備內實現，或圖2中顯示的單個設備可被實現為多個分散式設備。例如，儘管控制設備210和光譜儀220在本文中描述為兩個獨立的設備，但是控制設備210和光譜儀220可以在單個設備內實現。另外或可替代地，環境200的一組設備（例如，一個或多個設備）可以執行被描述為由環境200的另一組設備執行的一個或多個功能。

【0033】 圖3是設備300的示例元件的圖式。設備300可以對應於控制設備210及/或光譜儀220。在一些實施方式中，控制設備210及/或光譜儀220可以包括一個或多個設備300及/或設備300的一個或多個元件。如圖3所示，設備300可包括匯流排310、處理器320、記憶體330、儲存元件340、輸入元件350、輸出元件360、以及通信介面370。

【0034】 匯流排310包括允許在設備300的元件當中通信的元件。處理器320在硬體、韌體、或硬體和軟體的組合中實現。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）、現場可程式閘陣列（FPGA）、專用積體電路（ASIC）、或另一類型的處理元件。在一些實施方式中，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設備（例如，快閃記憶體、磁記憶體、及/或光學記憶體）。

【0035】 儲存元件340儲存與設備300的操作和使用相關的資訊及/或軟體。例如，儲存元件340可以包括硬碟（例如，磁片、光碟、磁光碟、及/或固態磁碟）、壓縮光碟（CD）、數位通用光碟（DVD）、軟碟、盒式磁帶

（cartridge）、磁帶、及/或另一類型的非暫時性電腦可讀取媒體連同相應的驅動器。

【0036】 輸入元件350包括允許設備300例如經由使用者輸入（例如，觸控式螢幕顯示器、鍵盤、小鍵盤、滑鼠、按鈕、開關、及/或麥克風）來接收資訊的元件。另外或可替代地，輸入元件350可以包括用於感測資訊的傳感器（例如，全球定位系統（GPS）元件、加速度計、陀螺儀、及/或致動器）。輸出元件360包括提供來自設備300的輸出資訊的元件（例如，顯示器、揚聲器、及/或一個或多個發光二極管（LED））。

【0037】 通信介面370包括使設備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設備通信的類似收發機元件（例如，收發機及/或單獨的接收機和發射機）。通信介面370可以允許設備300從另一設備接收資訊及/或將資訊提供給另一設備。例如，通信介面370可以包括以太網路介面、光學介面、同軸介面、紅外光介面、射頻（RF）介面、通用串列匯流排（USB）介面、無線區域網路介面、蜂巢式網路介面等。

【0038】 設備300可以執行本文中所述的一個或多個流程。設備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/或儲存元件340）儲存的軟體指令來執行這些流程。電腦可讀取媒體在本文中被定義為非暫時性記憶體設備。記憶體設備包括單個實體儲存設備內的記憶體空間或遍佈於多個實體儲存設備的記憶體空間。

【0039】 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設備被讀取到記憶體330及/或儲存元件340中。儲存在記憶體330及/或儲存元件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流程。另外或可替代地，硬連線電路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流程。因此，本文中所述的實施方式不限於硬體

電路和軟體的任何特定組合。

【0040】 圖3中顯示的元件的數量和佈置作為示例被提供。實際上，與圖3中顯示的那些元件相比，設備300可以包括附加的元件、更少的元件、不同的元件、或不同地佈置的元件。另外或可替代地，設備300的一組元件（例如，一個或多個元件）可以執行被描述為由設備300的另一組元件執行的一個或多個功能。

【0041】 圖4是用於產生用於光譜量化的量化模型的示例流程400的流程圖。在一些實施方式中，圖4的一個或多個流程方塊可以由控制設備210執行。在一些實施方式中，圖4的一個或多個流程方塊可以由另一設備或與控制設備210分離的或者包括控制設備210的設備組（諸如光譜儀220）執行。

【0042】 如圖4所示，流程400可包括針對訓練集及/或驗證集執行一組光譜測量（方塊410）。例如，控制設備210可以（例如，使用處理器320、通信介面370等）使光譜儀220對樣本的訓練集及/或驗證集執行一組光譜測量，以確定對於訓練集及/或驗證集的每個樣本的光譜。訓練集可以指具有一組濃度的成分的一種或多種已知材料的樣本集，其用於產生該成分的量化模型。類似地，驗證集可以指具有一組濃度的成分的一種或多種已知材料的樣本集，其用於驗證量化模型的準確性。例如，訓練集及/或驗證集可以包括一組不同濃度下的特定材料的一個或多個版本（例如，由不同製造商製造以控制製造差異的一個或多個版本）。

【0043】 在一些實施方式中，訓練集及/或驗證集可以基於對其要使用量化模型來執行光譜量化的預期的一組感興趣的材料集來被選擇。例如，當預期對藥物材料執行光譜量化以確定藥物材料的特定成分的濃度時，訓練集及/或驗證集可以包括針對特定成分的存在而被測試的一組藥物材料中的一組不同可能濃度下的特定成分的樣本集。

【0044】 在一些實施方式中，訓練集及/或驗證集可以被選擇以包括針對材料的每個濃度的特定數量的樣本。例如，訓練集及/或驗證集可以被選擇以包括特定濃度的多個樣本（例如，5個樣本、10個樣本、15個樣本、50個樣本等）。以這種方式，控制設備210可以被提供有與特定類型的材料相關聯的臨界值數量的光譜，從而有助於產生及/或驗證用於量化模型的類別（例如，對應於特定濃度的成分的一組樣本），未知樣本可以被準確地分配給該類別（例如，基於未知樣本具有特定濃度的成分）。

【0045】 在一些實施方式中，控制設備210可以使多個光譜儀220執行一組光譜測量以應對一個或多個物理條件。例如，控制設備210可以使第一光譜儀220和第二光譜儀220使用NIR光譜法來執行一組振動光譜測量。另外或可替代地，控制設備210可以使一組光譜測量在多個時間、在多個位置、在多個不同的實驗室條件等被執行。以這種方式，控制設備210減少了由於相對於使一組光譜測量由單個光譜儀220執行的物理條件而導致的光譜測量不準確的可能性。

【0046】 以這種方式，控制設備210使一組光譜測量針對訓練集及/或驗證集而執行。

【0047】 如圖4中進一步示出的，流程400可包括接收識別一組光譜測量的結果的資訊（方塊420）。例如，控制設備210可以（例如，使用處理器320、通信介面370等）接收識別該組光譜測量的結果的資訊。在一些實施方式中，控制設備210可以接收識別對應於訓練集及/或驗證集的樣本的一組光譜的資訊。例如，控制設備210可以接收識別在光譜儀220對訓練集執行光譜法時觀察到的特定光譜的資訊。在一些實施方式中，控制設備210可以同時接收識別訓練集樣本和驗證集樣本的光譜的資訊。在一些實施方式中，控制設備210可以接收識別關於訓練集樣本的光譜的資訊，可以產生量化模型，並且可以在產

生量化模型之後接收識別驗證集樣本的光譜的資訊，以便能夠測試量化模型。

【0048】 在一些實施方式中，控制設備210可以從多個光譜儀220接收識別一組光譜測量的結果的資訊。例如，控制設備210可以通過接收由多個光譜儀220執行的、在多個不同時間執行的、在多個不同位置處等執行的光譜測量來控制物理條件（例如多個光譜儀220之間的差異、實驗室條件中的潛在差異等）。

【0049】 在一些實施方式中，控制設備210可以從產生量化模型的利用中移除一個或多個光譜。例如，控制設備210可以執行光譜量化並且可以確定光譜不對應於量化模型用以量化的材料的類型，並且可以確定對應於光譜的樣本無意地是不感興趣的材料（例如，基於正確地執行光譜法中的人為誤差、識別訓練集的光譜的資訊中的誤差等）。在該情況下，控制設備210可以確定從訓練集中移除該光譜。以這種方式，控制設備210可以通過減少使用關於訓練集或驗證集的不正確或不準確資訊產生量化模型的可能性來提高量化模型的準確性。

【0050】 以這種方式，控制設備210接收識別一組光譜測量的結果的資訊。

【0051】 如圖4中進一步示出的，流程400可以包括基於識別一組光譜測量的結果的資訊產生量化模型（方塊430）。例如，控制設備210可以基於識別一組光譜測量的結果的資訊來（例如，使用處理器320、記憶體330、儲存元件340等）產生與SVM分類器技術相關聯的量化模型。

【0052】 SVM可以指執行模式識別並使用置信度度量以進行量化的監督學習模型。在一些實施方式中，當使用SVM技術產生量化模型時，控制設備210可以利用特定類型的核心函數來確定兩個或更多個輸入（例如，光譜）的類似性。例如，控制設備210可以利用徑向基底函數（RBF）（例如，稱為

SVM-rbf) 類型的核心函數，其可以被表示為對於光譜 x 和 y 的 $k(x,y) = \exp(-\|x-y\|^2)$ ；線性函數（例如，當用於多階段測定技術時，稱為SVM-線性和稱為層次(hier)-SVM線性) 類型的核心函數，其可以被表示為 $k(x,y) = \langle x \cdot y \rangle$ ；S型函數(sigmoid function) 類型的核心函數；多項式函數類型的核心函數；指數函數類型的核心函數；及/或類似的函數。在一些實施方式中，控制設備210可以使用單類SVM(SC-SVM) 分類器技術來產生量化模型。例如，控制設備210可以聚集對應于訓練集中的成分的多個濃度的多個類別以產生表示量化模型的單個類別。在該情況下，如本文中所述，控制設備210可以利用置信度量來確定未知樣本是量化模型用以分析的類型的可能性。

【0053】 在一些實施方式中，控制設備210可以針對SVM利用特定類型的置信度量，諸如基於機率值的SVM（例如，基於確定樣本是（可能的濃度的）一組類別中的類別的成員的機率的測定）、基於決策值的SVM（例如，利用決策函數投票給一組類別中的類別作為樣本是其成員的類別的測定）等。例如，在利用基於決策值的SVM的量化模型的使用期間，控制設備210可以基於未知樣本的光譜的測繪來確定未知樣本是否位於組成類別的邊界內（例如，未知樣本的成分的特定量或濃度），並且可以基於未知樣本是否位於組成類別的邊界內來將樣本分配給類別。以這種方式，控制設備210可以確定是否將未知光譜分配給特定類別以進行量化。

【0054】 儘管本文中所述的一些實施方式是以一組特定的機器學習技術的方式來描述的，但是其他技術也可能用於確定關於未知光譜的資訊，例如材料的分類等等。

【0055】 在一些實施方式中，控制設備210可以從一組量化技術中選擇要用於產生量化模型的特定分類器。例如，控制設備210可以產生對應於多個分類器的多個量化模型，並且可以測試多個量化模型，例如通過確定每個模型的

可傳遞性（例如，基於在第一光譜儀220上執行的光譜測量產生的量化模型當應用於在第二光譜儀220上執行的光譜測量時準確的程度）、大規模測定的準確度（例如，量化模型可被用於同時確定滿足臨界值的一定量的樣本的濃度的準確度）等等。在該情況下，控制設備210可以基於確定分類器相對於其他分類器與優越的可轉移性及/或大規模測定的準確度相關聯來選擇分類器，諸如SVM分類器（例如，層次-SVM線性分類器、SC-SVM分類器等）。

【0056】 在一些實施方式中，控制設備210可以基於識別訓練集的樣本的資訊來產生量化模型。例如，控制設備210可以利用識別由訓練集的樣本表示的材料的類型或濃度的資訊來識別具有材料的類型或濃度的光譜的類別。在一些實施方式中，當產生量化模型時，控制設備210可以訓練量化模型。例如，控制設備210可以使用一組光譜測量的一部分（例如，與訓練集相關的測量）來訓練量化模型。另外或可替代地，控制設備210可以執行量化模型的評估。例如，控制設備210可以利用該組光譜測量的另一部分（例如，驗證集）驗證量化模型（例如，針對預測強度）。

【0057】 在一些實施方式中，控制設備210可以使用多階段測定技術來驗證量化模型。例如，對於基於原位（in-situ）局部建模的量化，控制設備210可以確定量化模型當與一個或多個局部量化模型相關聯地被利用時是準確的。以這種方式，控制設備210確保在提供量化模型以用於例如由控制設備210、由與其他光譜儀220相關聯的其他控制設備210等利用之前以臨界值準確度產生量化模型。

【0058】 在一些實施方式中，控制設備210可以在產生量化模型之後向與其他光譜儀220相關聯的其他控制設備210提供量化模型。例如，第一控制設備210可以產生量化模型，並且可以將量化模型提供給第二控制設備210用於利用。在該情況下，對於基於原位局部建模的量化，第二控制設備210可以儲存

量化模型（例如，全域量化模型），並且可以利用量化模型來產生一個或多個原位局部量化模型，用於確定未知集的一個或多個樣本中的材料的成分的濃度。另外或可替代地，控制設備210可以儲存量化模型以用於由控制設備210在執行量化、在產生一個或多個局部量化模型（例如，對於基於原位局部建模的量化）等時利用。以這種方式，控制設備210提供用於在未知樣本的光譜量化中利用的量化模型。

【0059】 以這種方式，控制設備210基於識別一組光譜測量的結果的資訊產生量化模型。

【0060】 雖然圖4示出了流程400的示例方塊，但在一些實施方式中，與圖4中描繪的那些方塊相比，流程400可以包括附加的方塊、更少的方塊、不同的方塊或不同地佈置的方塊。另外或可替代地，流程400的兩個或更多個方塊可平行地被執行。

【0061】 圖5是與圖4中所示的示例流程400相關的示例實施方式500的圖式。圖5示出了產生量化模型的示例。

【0062】 如圖5中由元件符號505示出的，控制設備210-1向光譜儀220-1傳輸資訊，以指導光譜儀220-1對訓練集和驗證集510執行一組光譜測量。假設訓練集和驗證集510包括第一組訓練樣本（例如，其測量用於訓練量化模型）和第二組驗證樣本（例如，其測量用於驗證量化模型的準確性）。如由元件符號515示出的，光譜儀220-1基於接收到指令來執行一組光譜測量。如由元件符號520示出的，控制設備210-1接收用於訓練樣本的第一組光譜和用於驗證樣本的第二組光譜。在該情況下，訓練樣本和驗證樣本可以包括用於量化的感興趣的材料組中的成分的多個濃度的樣本。例如，控制設備210-1可以接收與產生全域模型（例如，全域分類模型或量化模型）相關的光譜，以使用全域模型和原位局部建模技術（用於產生局部模型，例如局部分類模型或量化模型）來識別

肉的類型，以及量化該類型的肉中的特定蛋白質的濃度。在該情況下，控制設備210-1可以被配置成產生多個局部量化模型（例如，用於確定使用原位局部建模識別的第一類型的肉中的特定蛋白質的濃度的第一量化模型、用於確定使用原位局部建模識別的第三類型的肉中的特定蛋白質的濃度的第三量化模型等）。假設控制設備210-1儲存識別訓練集和驗證集510的每個樣本的資訊。

【0063】 關於圖5，假設控制設備210-1已被選擇以利用層次-SVM線性分類器來產生分類模型，並利用SC-SVM分類器來產生多個量化模型。如由元件符號525所示，控制設備210-1使用層次-SVM線性分類器和第一組光譜訓練全域分類模型，並使用層次-SVM線性分類器和第二組光譜驗證全域分類模型。此外，控制設備210-1訓練並驗證多個局部量化模型（例如，對應於全域分類模型的每個類別及/或基於全域分類模型產生的局部分類模型的每個類別的局部量化模型）。假設控制設備210-1確定量化模型滿足驗證臨界值（例如，具有超過驗證臨界值的準確度）。如由元件符號530示出的，控制設備210-1將量化模型提供給控制設備210-2（例如，以用於在對由光譜儀220-2執行的光譜測量執行量化時利用）並提供給控制設備210-3（例如，以用於在對由光譜儀220-3執行的光譜測量執行量化時利用）。

【0064】 如上所指示，圖5僅作為示例被提供。其它示例是可能的，並且可以不同於關於圖5描述的示例。

【0065】 以這種方式，控制設備210有助於基於選擇的分類技術（例如，基於模型可傳遞性、大規模量化的準確度等所選擇的技術）的量化模型的產生和量化模型的分配以供與一個或多個光譜儀220相關聯的一個或多個其他控制設備210利用。

【0066】 圖6是用於在原料量化期間避免誤報識別的示例流程600的流程圖。在一些實施方式中，圖6的一個或多個流程方塊可由控制設備210執行。在

一些實施方式中，圖6的一個或多個流程方塊可由另一設備或與控制設備210分離的或者包括控制設備210的設備組（諸如光譜儀220）執行。

【0067】 如圖6所示，流程600可以包括接收識別對未知樣本執行的一組光譜測量的結果的資訊（方塊610）。例如，控制設備210可以（例如，使用處理器320、通信介面370等）接收識別對未知樣本執行的一組NIR光譜測量的結果的資訊。在一些實施方式中，控制設備210可以接收識別對（例如，多個樣本的）未知集的一組光譜測量的結果的資訊。未知集可以包括要對其執行測定（例如，光譜量化）的一組樣本（例如，未知樣本）。例如，控制設備210可以使光譜儀220對該組未知樣本執行一組光譜測量，並且可以接收識別對應於該組未知樣本的一組光譜的資訊。

【0068】 在一些實施方式中，控制設備210可以從多個光譜儀220接收識別結果的資訊。例如，控制設備210可以使多個光譜儀220對未知集（例如，相同的樣本集）執行一組光譜測量，並且可以接收識別對應於未知集的樣本的一組光譜的資訊。另外或可替代地，控制設備210可以接收識別在多個時間、在多個位置等執行的一組光譜測量的結果的資訊，並且可以基於在多個時間、在多個位置等執行的一組光譜測量（例如，基於平均該組光譜測量或基於另一技術）對特定樣本進行量化。以這種方式，控制設備210可以應對可能影響該組光譜測量的結果的物理條件。

【0069】 另外或可替代地，控制設備210可以使第一光譜儀220對未知集的第一部分執行一組光譜測量的第一部分，並且可以使第二光譜儀220對未知集的第二部分執行該組光譜測量的第二部分。以這種方式，相對於使所有光譜測量都由單個光譜儀220執行，控制設備210可以減少一定量的時間來執行一組光譜測量。

【0070】 以這種方式，控制設備210接收識別對未知樣本執行的一組光譜

測量的結果的資訊。

【0071】 如圖6中進一步所示，流程600可以包括確定未知樣本是否對應於量化模型（方塊620）。例如，控制設備210可以嘗試（例如，使用處理器320、記憶體330、儲存元件340等）確定未知樣本是否是量化模型用以量化的材料及/或包括材料中量化模型用以量化的成分。

【0072】 在一些實施方式中，控制設備210可以使用SC-SVM分類器技術來確定未知光譜是否對應於量化模型。例如，控制設備210可以確定可配置參數值 nu 以用於使用SC-SVM技術。參數值可以對應於被確定在SC-SVM技術的決策邊界內的訓練集樣本與被確定不在決策邊界內的訓練集樣本的比率。在一些實施方式中，控制設備210可以基於參數值來確定決策邊界。在一些實施方式中，控制設備210可以使用交叉驗證流程來設置多個可能的決策邊界，並且可以（例如經由平均）組合使用多個可能的決策邊界的結果來確定未知光譜是否對應於量化模型。

【0073】 在一些實施方式中，基於設置決策邊界以滿足參數值（例如，對於0.5的參數值，設置決策值使得訓練集的一半測量位於決策邊界內，並且訓練集的一半測量位於決策邊界外部），控制設備210可以確定決策值，該決策值可以對應於從測量到決策邊界的距離度量。例如，控制設備210可以確定未知樣本的光譜在一組軸上的位置，並且可以確定該位置和決策邊界的最近點之間的距離。儘管本文中所述的一些實施方式是根據圖或一組軸來描述的，但是本文中所述的實施方式可以在不使用圖或一組軸的情況下確定，例如使用與未知光譜相關的資料的另一表示。

【0074】 在一些實施方式中，控制設備210可以確定用於未知光譜的決策值。例如，控制設備210可以確定從未知光譜到決策邊界的距離。在一些實施方式中，控制設備210可以確定用於位於決策邊界外部的、其他測量的決策

值。在該情況下，控制設備210可以確定統計度量，以表示未知光譜的決策值相對於決策邊界外部的其他測量的決策值的標準偏差的數量。例如，控制設備210可以基於對數常態分配確定對數常態標準偏差，並且可以確定標準偏差是否滿足臨界值（例如，1個標準偏差、2個標準偏差、3個標準偏差等）。在該情況下，基於未知樣本的光譜的測量相對於決策邊界外部的其他測量大於相對於決策邊界的標準偏差的臨界值數量（例如，相對於決策邊界的3個標準偏差），控制設備210可以確定未知樣本不對應於量化模型（方塊620-否）。可替代地，基於測量小於相對於決策邊界的臨界值量標準偏差，控制設備210可以確定未知樣本確實對應於量化模型（方塊620-是）。儘管本文中按照特定統計技術及/或標準偏差的特定臨界值數量進行了描述，但是也可以使用其他統計技術及/或臨界值。

【0075】 以這種方式，控制設備210能夠識別未知光譜與量化模型相差的臨界值量，而無需使用與未知樣本類似的樣本（例如，也與感興趣的材料的訓練集樣本相差的臨界值量）來訓練量化模型。另外，控制設備210減少了要收集用於產生量化模型的樣本量，從而相對於獲得、儲存和處理其他樣本以確保準確識別與感興趣材料及/或其濃度相差的臨界值量的樣本，減少了成本、時間和計算資源利用率（例如，處理資源和記憶體資源）。

【0076】 以這種方式，控制設備210確定未知樣本是否對應於量化模型。

【0077】 如圖6中進一步示出的，基於確定未知樣本對應於量化模型（方塊620-是），流程600可以包括基於一組光譜測量的結果執行一個或多個光譜測定（方塊630）。例如，控制設備210可以基於一組光譜測量的結果來（例如，使用處理器320、記憶體330、儲存元件340等）執行一個或多個光譜測定。在一些實施方式中，控制設備210可以將未知樣本分配給特定類別（例如，表示感興趣的材料中的成分的一組濃度中的特定濃度）。

【0078】 在一些實施方式中，控制設備210可以基於置信度度量來分配特定樣本。例如，控制設備210可以基於量化模型確定特定光譜與量化模型的每個類別（例如，每個候選濃度）相關聯的機率。在該情況下，控制設備210可以基於類別的特定機率超過與類別相關聯的其他機率而將未知樣本分配給該類別（例如，特定濃度）。以這種方式，控制設備210確定與樣本相關聯的感興趣的材料中的成分的濃度，從而量化樣本。

【0079】 在一些實施方式中，為了執行原位局部建模，例如對於具有大於臨界值數量的類別的量化模型，控制設備210可以基於第一測定產生局部量化模型。局部量化模型可以指基於與第一測定相關聯的置信度度量使用SVM測定技術（例如SVM-rbf、SVM-線性等的核心函數；基於機率值的SVM、基於決策值的SVM等；及/或類似技術）產生的原位量化模型。

【0080】 在一些實施方式中，控制設備210可以基於使用全域分類模型執行第一測定來產生局部量化模型。例如，當控制設備210被用於確定未知樣本中的成分的濃度，並且多個未知樣本與用於確定成分的濃度的不同量化模型相關聯時，控制設備210可以利用第一測定來選擇類別的子集作為用於未知樣本的局部類別，並且可以產生與用於未知樣本的局部類別相關聯的局部量化模型。以這種方式，控制設備210利用層次測定和量化模型來改進光譜分類。在該情況下，控制設備210可以基於確定未知樣本相對於局部量化模型的其他測量的子集的距離度量來確定未知樣本是否對應於該局部量化模型。

【0081】 作為示例，當執行原料識別以確定植物材料中特定化學物的濃度時，其中植物材料與多個量化模型相關聯（例如，與植物是在室內還是室外、在冬季還是夏季生長等相關），控制設備210可以執行一組分類測定以識別特定量化模型。在該情況下，控制設備210可以基於執行一組測定來確定植物在冬季室內生長，並且可以選擇與在冬季室內生長的植物相關的量化模型以

用於確定特定化學物的濃度。基於選擇量化模型，控制設備210可以確定未知樣本對應於量化模型，並且可以使用量化模型量化未知樣本。

【0082】 在一些實施方式中，控制設備210可能無法使用量化模型量化未知樣本。例如，基於一個或多個決策值或其他置信度度量未能滿足臨界值，控制設備210可以確定未知樣本不能夠使用量化模型被準確地量化（方塊630-A）。可替代地，控制設備210可以基於一個或多個決策值或其他置信度度量滿足臨界值而成功量化未知樣本（方塊630-B）。

【0083】 以這種方式，控制設備210基於一組光譜測量的結果執行一個或多個光譜測定。

【0084】 如圖6中進一步示出的，基於確定未知樣本不對應於量化模型（方塊620-否），或者基於執行一個或多個光譜測定時的測定失敗（方塊630-A），流程600可以包括提供指示未知樣本不對應於量化模型的輸出（方塊640）。例如，控制設備210可以例如經由使用者介面來（例如，使用處理器320、記憶體330、儲存元件340、通信介面370等）提供指示未知樣本不對應於量化模型的資訊。在一些實施方式中，控制設備210可以提供與識別未知樣本相關聯的資訊。例如，基於嘗試量化特定植物中的特定化學物的量，並且確定未知樣本不是該特定植物（而是另一植物，例如基於人為誤差），控制設備210可以提供識別另一植物的資訊。在一些實施方式中，基於確定未知樣本不對應於量化模型，控制設備210可以獲得另一量化模型，並且可以使用另一量化模型以識別未知樣本。

【0085】 以這種方式，控制設備210減少了基於對未知樣本的誤報識別提供不正確資訊的可能性，並且通過提供資訊以輔助確定未知樣本例如是另一種植物而不是特定植物來使技術人員能夠進行誤差校正。

【0086】 以該方式，控制設備210提供指示未知樣本不對應於量化模型的

輸出。

【0087】 如圖6中進一步示出的，基於當執行一個或多個光譜測定時的分類成功（方塊630-B），流程600可以包括提供識別與未知樣本相關的分類的資訊（方塊650）。例如，控制設備210可以（例如，使用處理器320、記憶體330、儲存元件340、通信介面370等）提供識別與未知樣本相關的量化的資訊。在一些實施方式中，控制設備210可以提供識別用於未知樣本的特定類別的資訊。例如，控制設備210可以提供指示與未知樣本相關聯的特定光譜被確定為與對應於感興趣的材料中的成分的特定濃度的特定類別相關聯的資訊，從而識別未知樣本。

【0088】 在一些實施方式中，控制設備210可以提供指示與將未知樣本分配給特定類別相關聯的置信度度量的資訊。例如，控制設備210可以提供識別未知樣本與特定類別相關聯的機率的資訊等。以該方式，控制設備210提供指示特定光譜被準確地分配給特定類別的可能性的資訊。

【0089】 在一些實施方式中，控制設備210可以基於執行一組分類來提供量化。例如，基於識別與未知樣本的類別相關的局部量化模型，控制設備210可以提供識別未知樣本中的物質的濃度的資訊。在一些實施方式中，控制設備210可以基於執行一組量化來更新量化模型。例如，控制設備210可以基於將未知樣本的量化確定為感興趣的材料中的成分的特定濃度來產生包括作為訓練集的樣本的未知樣本的新的量化模型。

【0090】 以這種方式，控制設備210提供識別未知樣本的資訊。

【0091】 雖然圖6示出了流程600的示例方塊，但在一些實施方式中，與圖6中描繪的那些方塊相比，流程600可包括附加的方塊、更少的方塊、不同的方塊或不同地佈置的方塊。另外或可替代地，流程600的兩個或更多個方塊可平行地被執行。

【0092】 圖7A和圖7B是與圖6中所示的示例流程600相關聯的預測成功率相關的示例實施方式700的圖式。圖7A和圖7B示出了使用基於層次支持向量機（層次-SVM線性）的技術的原料識別的示例結果。

【0093】 如圖7A中由元件符號705所示，控制設備210可以使光譜儀220執行一組光譜測量。例如，控制設備210可以提供指令，以使光譜儀220獲得未知樣本的光譜，從而確定未知樣本中的成分的濃度。如由元件符號710和元件符號715示出的，光譜儀220可以接收未知樣本，並且可以對未知樣本執行一組光譜測量。如由元件符號720示出的，控制設備210可以基於光譜儀220對未知樣本執行一組光譜測量而接收未知樣本的光譜。

【0094】 如圖7B所示，控制設備210可以使用量化模型725來執行光譜量化。量化模型725包括基於參數值 nu 確定的單個類別730，使得單個類別730的決策邊界產生決策邊界內的訓練集的樣本與不在決策邊界內的訓練集的樣本的臨界值比率。在該情況下，量化模型725可以與對應於訓練集的樣本中的成分的多個不同濃度的多個子類別相關聯。如由元件符號735和740示出的，基於樣本到決策邊界的距離的標準偏差值（例如， $\sigma=3.2$ ）滿足臨界值（例如，3），確定未知樣本的光譜不對應於量化模型。如由元件符號745所示，控制設備210向用戶端設備750提供輸出，指示未知樣本不對應於量化模型，而不是提供未知樣本作為感興趣的材料中的特定濃度的成分的誤報識別。

【0095】 如上所指示，圖7A和圖7B僅作為示例被提供。其它示例是可能的，並且可以不同於關於圖7A和圖7B描述的示例。

【0096】 以這種方式，控制設備210基於避免未知樣本作為量化模型被訓練來量化的感興趣的材料中特定濃度的成分的誤報識別來降低了提供光譜法的不準確結果的可能性。

【0097】 前述揭示提供了說明和描述，但並不旨在窮舉或將實施方式限

制到所揭露的精確形式。根據以上揭示，修改和變化是可能的或者可以從實施方式的實踐中獲得。

【0098】 本文中結合臨界值描述了一些實施方式。如本文中所使用的，滿足臨界值可以指值大於臨界值、多於臨界值、高於臨界值、大於或等於臨界值、小於臨界值、少於臨界值、低於臨界值、小於或等於臨界值、等於臨界值等。

【0099】 明顯的是，本文中所述的系統和/方法可在硬體、軟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碼不是實施方式的限制。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碼，應理解，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。

【0100】 儘管在申請專利範圍中陳述及/或在說明書中揭露了特徵的特定組合，但是這些組合並不旨在限制可能的實施方式的揭露內容。事實上，這些特徵中的許多可以以申請專利範圍中未具體陳述及/或說明書中未揭露的方式組合。儘管所附的每個附屬項可以直接附屬於僅僅一個申請專利範圍，但是可能的實施方式的揭露內容包括與申請專利範圍中的每個其他申請專利範圍相結合的每個附屬項。

【0101】 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”和“一個”旨在包括一個或多個項目，並且可以與“一個或多個（one or more）”互換使用。此外，如本文中所使用的，術語“集（set）”旨在包括一個或多個項目（例如，相關項目、不相關項目、相關項目和不相關項目的組合等），並且可以與“一個或多個”互換使用。在意指僅一個項目的情況下，使用術語“一個（one）”或類似的語言。另外，如本文中所使用的，術語“包含”、“包

括”、“具有”、及/或類似詞語旨在是開放式的術語。此外，除非另有明確地說明，否則術語“基於”旨在表示“至少部分基於”。

【0102】 此外，本申請案包括根據下列條目的實施例：

條目1：一種設備，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用以：

接收識別對未知樣本執行的光譜測量的結果的資訊；

基於可配置參數來確定用於量化模型的決策邊界，使得所述量化模型的第一多個訓練集樣本在所述決策邊界內，並且所述量化模型的第二多個訓練集樣本不在所述決策邊界內；

確定對所述未知樣本執行的所述光譜測量相對於所述決策邊界的距離度量；

確定所述量化模型的所述第二多個訓練集樣本相對於所述決策邊界的多個距離度量；

基於用於所述光譜測量的所述距離度量和用於所述第二多個訓練集樣本的所述多個距離度量，確定對所述未知樣本執行的所述光譜測量是否對應於所述量化模型；以及

提供指示對所述未知樣本執行的所述光譜測量是否對應於所述量化模型的資訊。

條目2：根據條目1所述的設備，其中，當確定對所述未知樣本執行的所述光譜測量是否對應於所述量化模型時，所述一個或多個處理器用以：

確定所述光譜測量不對應於所述量化模型；以及

其中，當提供指示對所述未知樣本執行的所述光譜測量是否對應於所述量

化模型的資訊時，所述一個或多個處理器用以提供指示所述光譜測量不對應於所述量化模型的資訊。

條目3：根據條目1所述的設備，其中，當確定對所述未知樣本執行的所述光譜測量是否對應於所述量化模型時，所述一個或多個處理器用以：

確定所述光譜測量對應於所述量化模型；以及

其中，當提供指示對所述未知樣本執行的所述光譜測量是否對應於所述量化模型的資訊時，所述一個或多個處理器用以提供指示所述光譜測量對應於所述量化模型的資訊。

條目4：根據條目1所述的設備，其中，當確定對所述未知樣本執行的所述光譜測量是否對應於所述量化模型時，所述一個或多個處理器用以：

確定所述距離度量相對於所述多個距離度量的統計度量；以及

基於所述統計度量確定對所述未知樣本執行的所述光譜測量是否對應於所述量化模型。

條目5：根據條目4所述的設備，其中，所述統計度量是對數常態標準偏差；並且

其中，當基於所述統計度量確定對所述未知樣本執行的所述光譜測量是否對應於所述量化模型時，所述一個或多個處理器用以：

確定所述對數常態標準偏差滿足臨界值；以及

基於確定所述對數常態標準偏差滿足所述臨界值，確定對所述未知樣本執行的所述光譜測量是否對應於所述量化模型。

條目6：根據條目1所述的設備，其中，所述量化模型與單類支持向量機（SC-SVM）分類器相關聯。

條目7：根據條目1所述的設備，其中，所述一個或多個處理器進一步用以：

接收與所述第一多個訓練集樣本和所述第二多個訓練集樣本相關的多個光譜測量；

基於所述多個光譜測量確定所述量化模型；

基於對多個驗證集樣本的另一多個光譜測量來驗證所述量化模型；

儲存所述量化模型；以及

其中，當確定所述決策邊界時，所述一個或多個處理器用以：

從儲存器獲得所述量化模型；以及

在從儲存器獲得所述量化模型之後，確定所述決策邊界。

條目8：一種儲存指令的非暫時性電腦可讀取媒體，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：

獲得與特定類型的感興趣的材料相關的量化模型，所述量化模型被配置用於確定所述特定類型的感興趣的材料的樣本中的特定成分的濃度；

接收識別對未知樣本執行的特定光譜測量的結果的資訊；

將所述量化模型的訓練集樣本的其他光譜測量聚集到用於所述量化模型的單個類別中；

將所述訓練集樣本的所述其他光譜測量細分為第一組和第二組，所述其他光譜測量的所述第一組在決策邊界內，所述其他光譜測量的所述第二組不在所述決策邊界內；

確定對所述未知樣本執行的所述特定光譜測量的度量相對於所述其他光譜測量的所述第二組的對應的度量滿足臨界值；以及

提供指示所述未知樣本不是所述特定類型的感興趣的材料的資訊。

條目9：根據條目8所述的非暫時性電腦可讀取媒體，其中，所述未知樣本是與所述特定類型的感興趣的材料不同類型的材料。

條目10：根據條目8所述的非暫時性電腦可讀取媒體，其中，所述未知樣本是所述特定類型的感興趣的材料，並且是不正確地獲得的測量。

條目11：根據條目8所述的非暫時性電腦可讀取媒體，其中，所述度量和所述對應的度量是決策值。

條目12：根據條目8所述的非暫時性電腦可讀取媒體，其中，所述臨界值是所述度量相對於所述對應的度量的平均值的標準偏差的臨界值數量。

條目13：根據條目8所述的非暫時性電腦可讀取媒體，其中，所述度量和所述對應的度量是使用單類支持向量機技術來確定的。

條目14：根據條目8所述的非暫時性電腦可讀取媒體，其中，所述量化模型是局部模型，

其中，所述一個或多個指令當由所述一個或多個處理器執行時還使所述一個或多個處理器：

使用與所述特定類型的感興趣的材料相關的全域模型，執行與所述未知樣本的所述特定光譜測量相關的第一測定；

基於所述第一測定的特定結果並使用原位局部建模技術來產生所述局部模型；並且

其中，使所述一個或多個處理器獲得所述量化模型的所述一個或多個指令使所述一個或多個處理器基於產生所述局部模型來獲得所述量化模型。

條目15：一種方法，包括：

由設備接收識別對未知樣本執行的近紅外光（NIR）光譜測量的結果的資訊；

由所述設備確定用於量化模型的決策邊界，其中，所述決策邊界將所述量化模型的單個類別劃分為在所述決策邊界內的所述量化模型的第一多個訓練集樣本以及不在所述決策邊界內的所述量化模型的第二多個訓練集樣本；

由所述設備確定對所述未知樣本執行的所述NIR光譜測量的特定距離度量相對於所述第二多個訓練集樣本的其他距離度量滿足臨界值；以及

由所述設備基於確定對所述未知樣本執行的所述NIR光譜測量的所述特定距離度量相對於所述第二多個訓練集樣本的所述其他距離度量滿足所述臨界值，提供指示對所述未知樣本執行的所述NIR光譜測量不對應於所述量化模型的資訊。

條目16：根據條目15所述的方法，進一步包括：

基於使用分類模型的所述NIR光譜測量並基於確定對所述未知樣本執行的所述特定距離度量相對於所述第二多個訓練集樣本的所述其他距離度量滿足所述臨界值，確定所述未知樣本的類型；以及

提供識別所述未知樣本的類型的資訊。

條目17：根據條目15所述的方法，進一步包括：

基於核心函數確定所述決策邊界。

條目18：根據條目17所述的方法，其中，所述核心函數是下列中的至少一個：

徑向基底函數，

多項式函數，

線性函數，或

指數函數。

條目19：根據條目15所述的方法，其中，所述臨界值大於下列中的至少一個：

1個標準偏差，

2個標準偏差，或

3個標準偏差。

條目20：根據條目15所述的方法，其中，所述第一多個訓練集樣本和所述第二多個訓練集樣本與成分的一組濃度相關聯，並且

其中，所述成分的一組濃度中的所述成分的每個濃度與所述第一多個訓練集樣本和所述第二多個訓練集樣本的訓練集樣本的臨界值數量相關聯。

【符號說明】

【0103】

- 100 實施方式
- 200 環境
- 210 控制設備
- 210-1 控制設備
- 210-2 控制設備
- 210-3 控制設備
- 220 光譜儀
- 220-1 光譜儀
- 220-2 光譜儀
- 220-3 光譜儀
- 230 網路
- 300 設備
- 310 匯流排
- 320 處理器
- 330 記憶體
- 340 儲存元件
- 350 輸入元件

- 360 輸出元件
- 370 通信介面
- 400 流程
- 410 方塊
- 420 方塊
- 430 方塊
- 500 實施方式
- 505 執行光譜測量
- 510 訓練集和驗證集
- 515 對訓練樣本和驗證樣本執行光譜測量
- 520 用於訓練樣本的第一光譜和用於驗證樣本的第二光譜
- 525 訓練和驗證量化模型
- 530 提供量化模型
- 600 流程
- 610 方塊
- 620 方塊
- 630 方塊
- 640 方塊
- 650 方塊
- 700 實施方式
- 705 執行光譜測量
- 710 未知樣本
- 715 對未知樣本執行光譜測量
- 720 未知樣本的光譜

- 725 量化模型
- 730 單個類別
- 735 樣本
- 740 將未知樣本識別為不對應於量化模型
- 745 提供將樣本識別為不對應於量化模型的資訊
- 750 用戶端設備

【發明申請專利範圍】

【第1項】一種用於在光譜量化期間避免誤報識別的設備，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用以：

接收識別對未知樣本執行的光譜測量的結果的資訊；

基於可配置參數來確定用於量化模型的決策邊界，所述決策邊界將所述量化模型的單個類別劃分為在所述決策邊界內的所述量化模型的第一多個訓練集樣本以及不在所述決策邊界內的所述量化模型的第二多個訓練集樣本；

確定對所述未知樣本執行的所述光譜測量相對於所述決策邊界的距離度量；

確定所述量化模型的所述第二多個訓練集樣本相對於所述決策邊界的多個距離度量；

基於用於所述光譜測量的所述距離度量和用於所述第二多個訓練集樣本的所述多個距離度量，確定對所述未知樣本執行的所述光譜測量是否對應於所述量化模型；以及

提供指示對所述未知樣本執行的所述光譜測量是否對應於所述量化模型的資訊。

【第2項】根據請求項1所述的設備，其中，當確定對所述未知樣本執行的所述光譜測量是否對應於所述量化模型時，所述一個或多個處理器用以：

確定所述光譜測量不對應於所述量化模型；以及

其中，當提供指示對所述未知樣本執行的所述光譜測量是否對應於所述量化模型的資訊時，所述一個或多個處理器用以：

提供指示所述光譜測量不對應於所述量化模型的資訊。

【第3項】根據請求項1所述的設備，其中，當確定對所述未知樣本執行的所述光譜測量是否對應於所述量化模型時，所述一個或多個處理器用以：

確定所述光譜測量對應於所述量化模型；以及

其中，當提供指示對所述未知樣本執行的所述光譜測量是否對應於所述量化模型的資訊時，所述一個或多個處理器用以：

提供指示所述光譜測量對應於所述量化模型的資訊。

【第4項】根據請求項1所述的設備，其中，當確定對所述未知樣本執行的所述光譜測量是否對應於所述量化模型時，所述一個或多個處理器用以：

確定所述距離度量相對於所述多個距離度量的統計度量；以及

基於所述統計度量確定對所述未知樣本執行的所述光譜測量是否對應於所述量化模型。

【第5項】根據請求項4所述的設備，其中，所述統計度量是對數常態標準偏差；並且

其中，當基於所述統計度量確定對所述未知樣本執行的所述光譜測量是否對應於所述量化模型時，所述一個或多個處理器用以：

確定所述對數常態標準偏差滿足臨界值；以及

基於確定所述對數常態標準偏差滿足所述臨界值，確定對所述未知樣本執行的所述光譜測量是否對應於所述量化模型。

【第6項】根據請求項1所述的設備，其中，所述量化模型與單類支持向量機（SC-SVM）分類器相關聯。

【第7項】根據請求項1所述的設備，其中，所述一個或多個處理器進一步用以：

接收與所述第一多個訓練集樣本和所述第二多個訓練集樣本相關的多個光譜測量；

基於所述多個光譜測量確定所述量化模型；

基於對多個驗證集樣本的另一多個光譜測量來驗證所述量化模型；

儲存所述量化模型；以及

其中，當確定所述決策邊界時，所述一個或多個處理器用以：

從儲存器獲得所述量化模型；以及

在從儲存器獲得所述量化模型之後，確定所述決策邊界。

【第8項】一種儲存指令的非暫時性電腦可讀取媒體，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：

獲得與特定類型的感興趣的材料相關的量化模型，所述量化模型被配置用於確定所述特定類型的感興趣的材料的樣本中的特定成分的濃度；

接收識別對未知樣本執行的特定光譜測量的結果的資訊；

將所述量化模型的訓練集樣本的其他光譜測量聚集到用於所述量化模型的單個類別中；

將所述訓練集樣本的所述其他光譜測量細分為第一組和第二組，

所述其他光譜測量的所述第一組在決策邊界內，且

所述其他光譜測量的所述第二組不在所述決策邊界內；

確定對所述未知樣本執行的所述特定光譜測量的度量相對於所述其他光譜測量的所述第二組的對應的度量滿足臨界值；以及

提供指示所述未知樣本不是所述特定類型的感興趣的材料的資訊。

【第9項】根據請求項8所述的非暫時性電腦可讀取媒體，其中，所述未知樣本是與所述特定類型的感興趣的材料不同類型的材料。

【第10項】根據請求項8所述的非暫時性電腦可讀取媒體，其中，所述未知樣本是所述特定類型的感興趣的材料，並且所述未知樣本的光譜測量是不正

確地獲得的測量。

【第11項】根據請求項8所述的非暫時性電腦可讀取媒體，其中，所述度量和所述對應的度量是決策值。

【第12項】根據請求項8所述的非暫時性電腦可讀取媒體，其中，所述臨界值是所述度量相對於所述對應的度量的平均值的標準偏差的臨界值數量。

【第13項】根據請求項8所述的非暫時性電腦可讀取媒體，其中，所述度量和所述對應的度量是使用單類支持向量機技術來確定的。

【第14項】根據請求項8所述的非暫時性電腦可讀取媒體，其中，所述量化模型是局部模型，

其中，所述一個或多個指令當由所述一個或多個處理器執行時還使所述一個或多個處理器：

使用與所述特定類型的感興趣的材料相關的全域模型，執行與所述未知樣本的所述特定光譜測量相關的第一測定；

基於所述第一測定的特定結果並使用原位局部建模技術來產生所述局部模型；並且

其中，使所述一個或多個處理器獲得所述量化模型的所述一個或多個指令使所述一個或多個處理器：

基於產生所述局部模型來獲得所述量化模型。

【第15項】一種用於在光譜量化期間避免誤報識別的方法，包括：

由設備接收識別對未知樣本執行的近紅外光（NIR）光譜測量的結果的資訊；

由所述設備確定用於量化模型的決策邊界，其中，所述決策邊界將所述量化模型的單個類別劃分為在所述決策邊界內的所述量化模型的第一多個訓練集樣本以及不在所述決策邊界內的所述量化模型的第二多個訓練集樣本；

由所述設備確定對所述未知樣本執行的所述NIR光譜測量的特定距離度量相對於所述第二多個訓練集樣本的其他距離度量滿足臨界值；以及

由所述設備基於確定所述特定距離度量相對於所述第二多個訓練集樣本的所述其他距離度量滿足所述臨界值，提供指示對所述未知樣本執行的所述NIR光譜測量不對應於所述量化模型的資訊。

【第16項】根據請求項15所述的方法，進一步包括：

由所述設備基於使用分類模型的所述NIR光譜測量並基於確定所述特定距離度量相對於所述第二多個訓練集樣本的所述其他距離度量滿足所述臨界值，確定所述未知樣本的類型；以及

由所述設備提供識別所述未知樣本的類型的資訊。

【第17項】根據請求項15所述的方法，進一步包括：

由所述設備基於核心函數確定所述決策邊界。

【第18項】根據請求項17所述的方法，其中，所述核心函數是下列中的至少一個：

徑向基底函數，

多項式函數，

線性函數，或

指數函數。

【第19項】根據請求項15所述的方法，其中，所述臨界值大於下列中的至少一個：

1個標準偏差，

2個標準偏差，或

3個標準偏差。

【第20項】根據請求項15所述的方法，其中，所述第一多個訓練集樣本和

所述第二多個訓練集樣本與成分的一組濃度相關聯，並且

其中，在所述成分的所述一組濃度中，所述成分的每個濃度與所述第一多個訓練集樣本和所述第二多個訓練集樣本的訓練集樣本的臨界值數量相關聯。

【發明圖式】

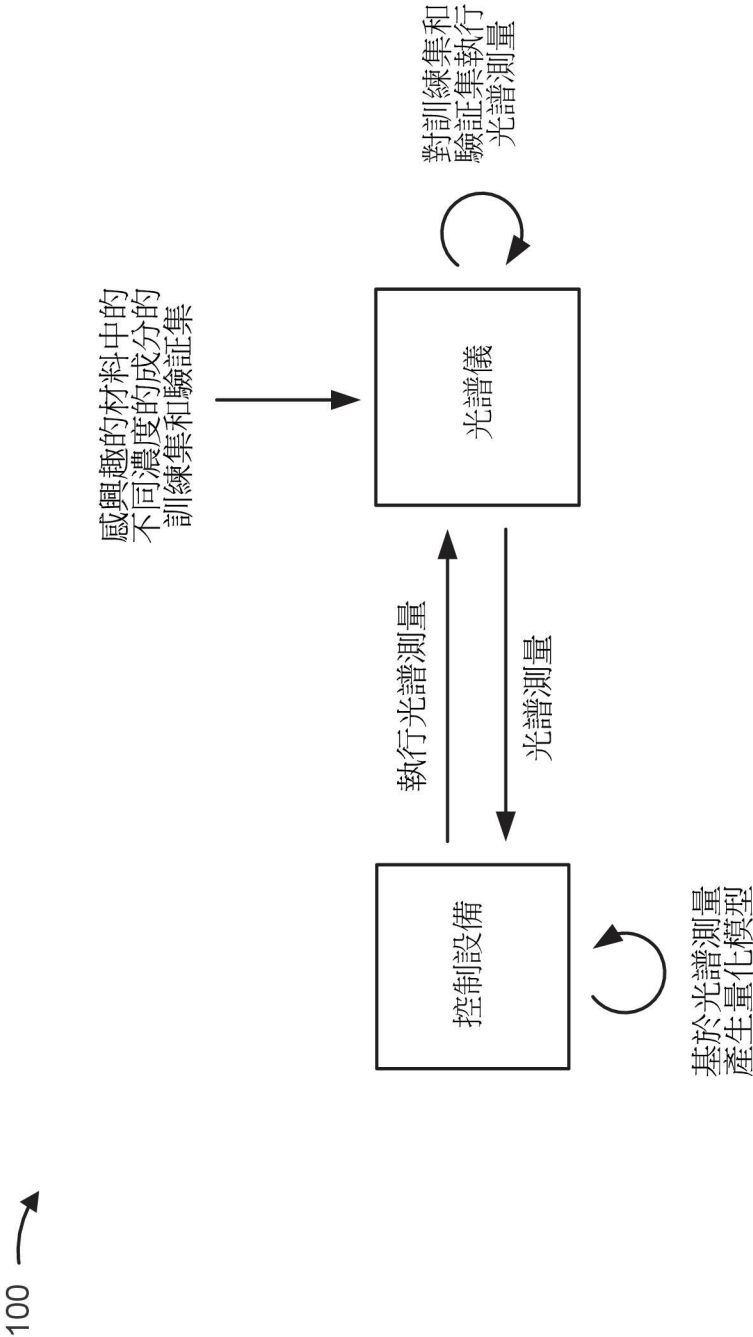


圖1A

100 

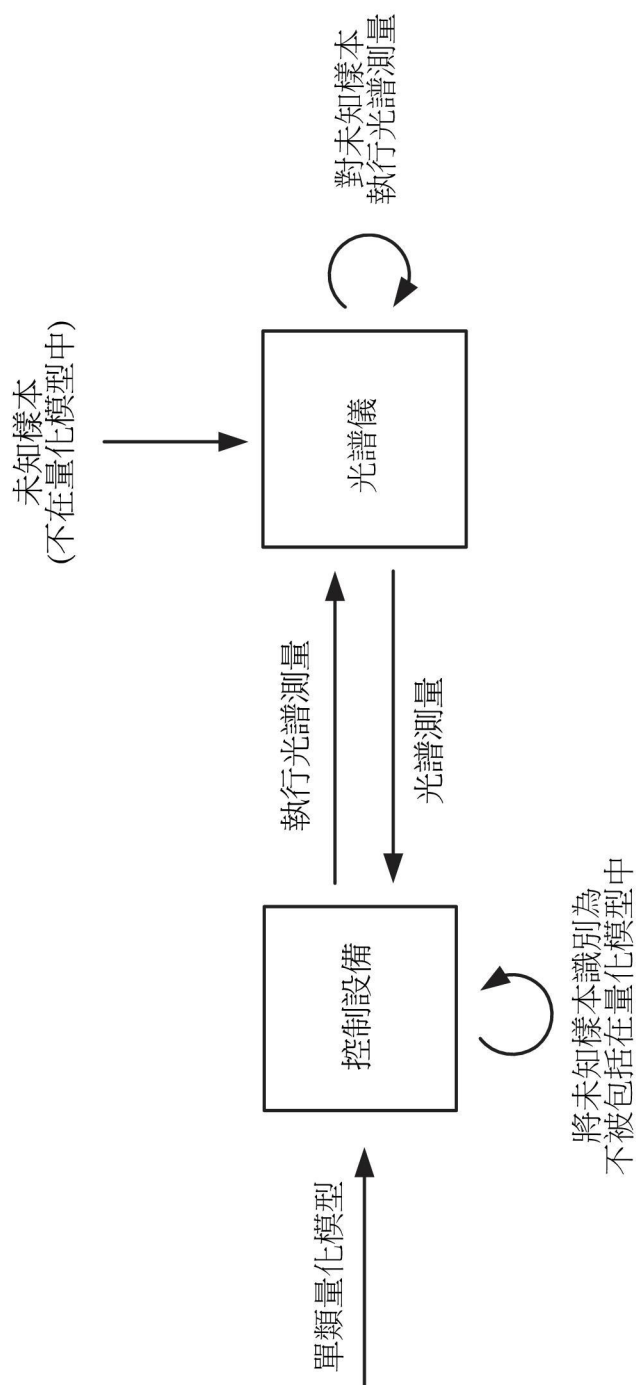


圖 1B

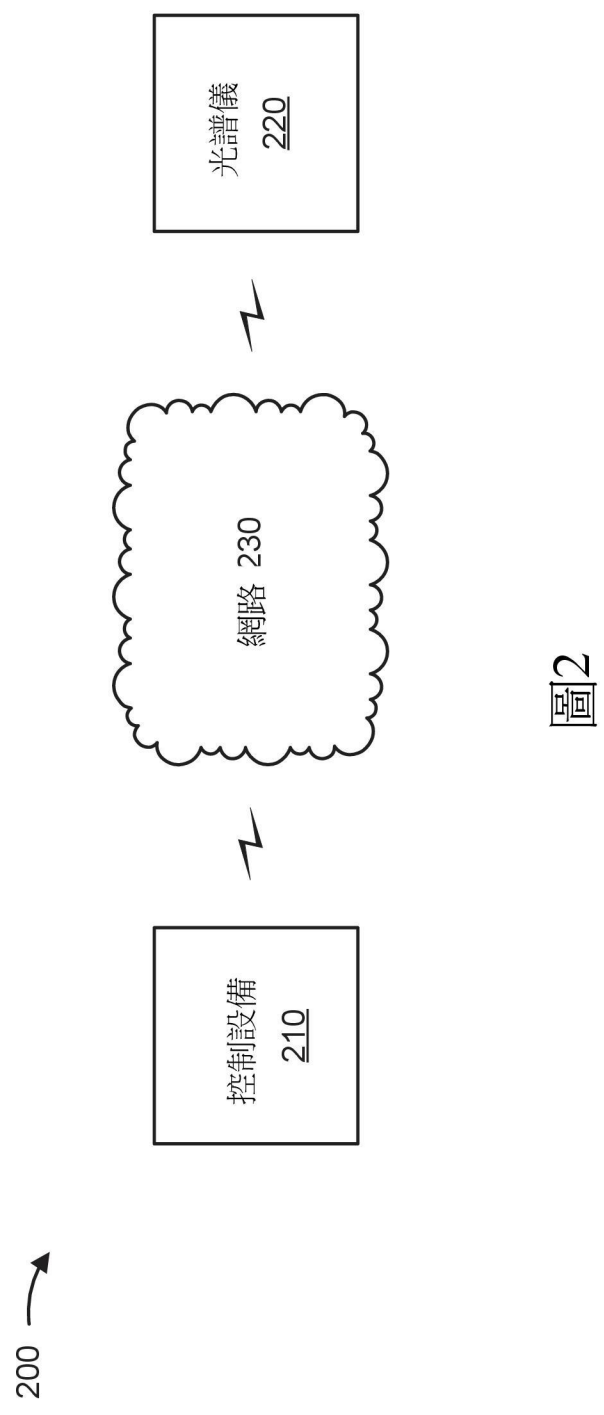


圖2

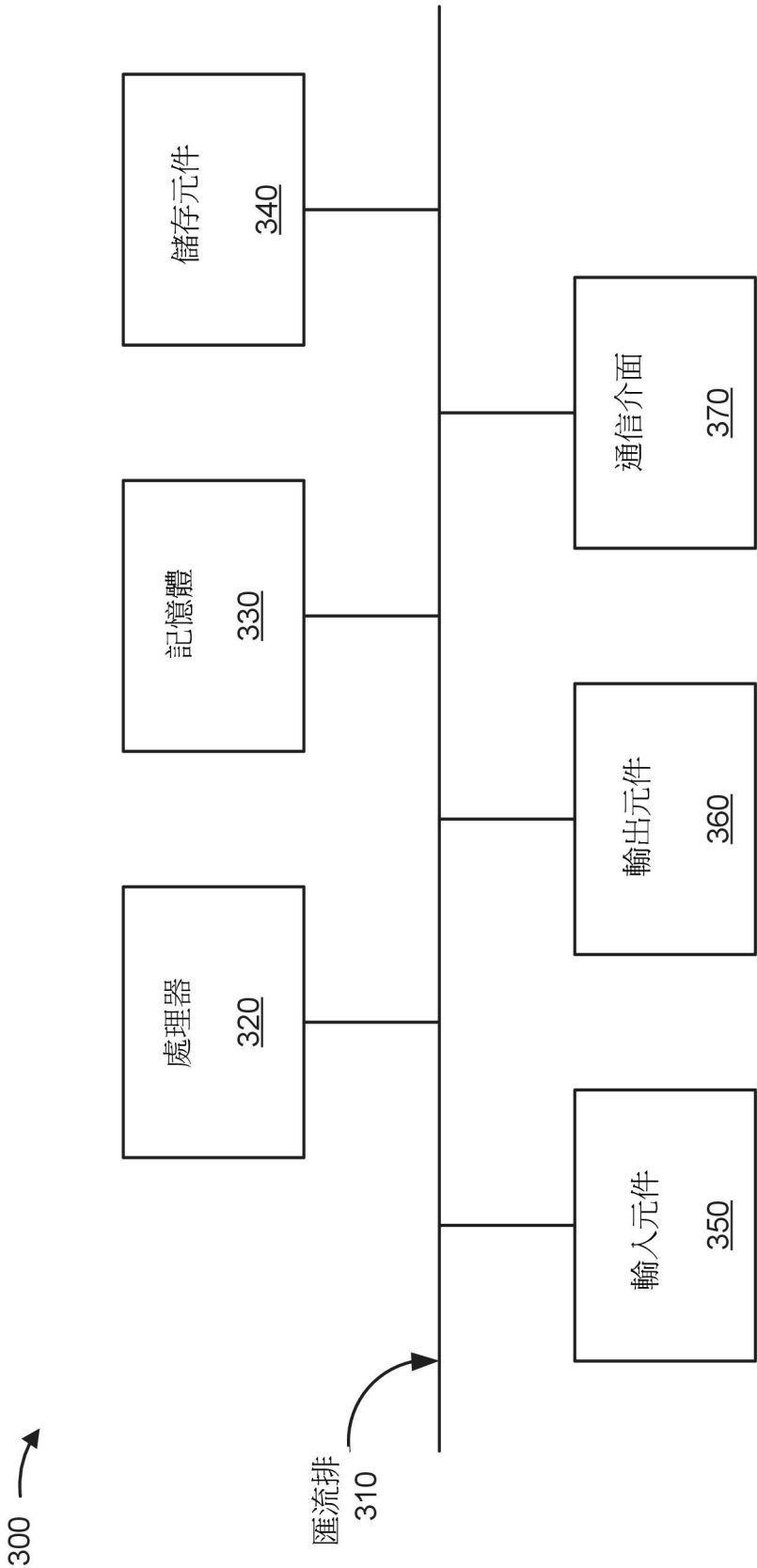


圖3

400 →

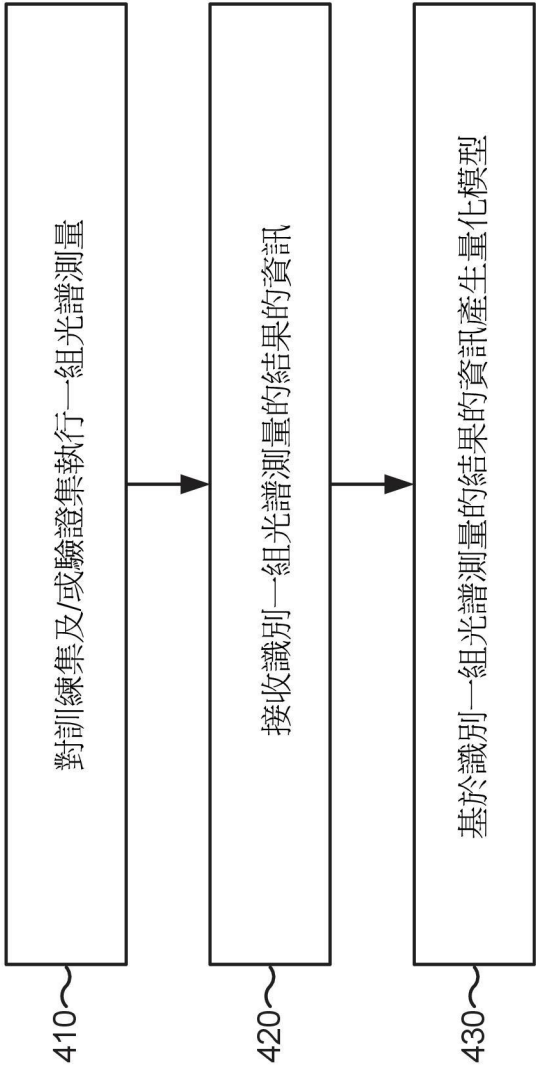


圖4

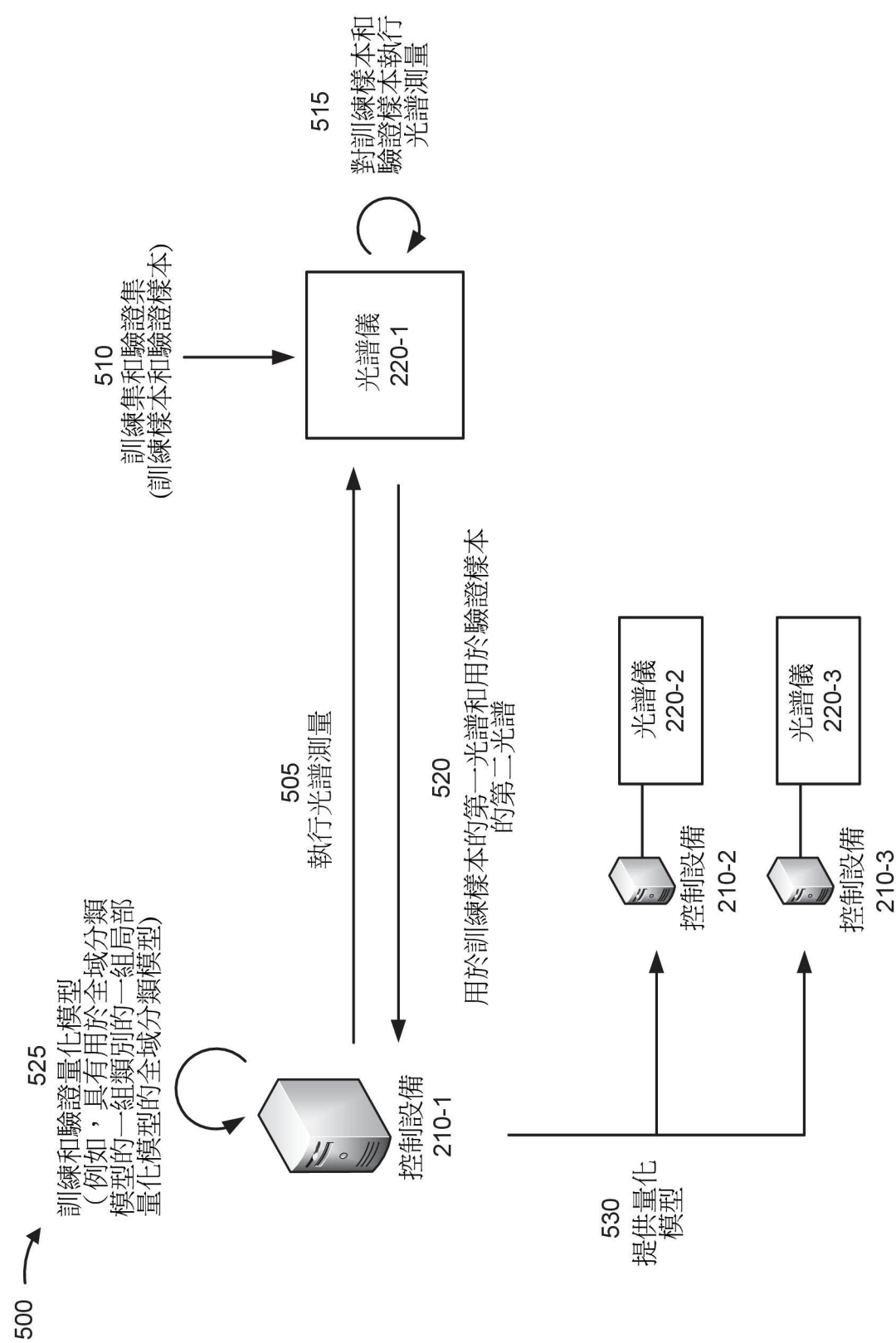


圖5

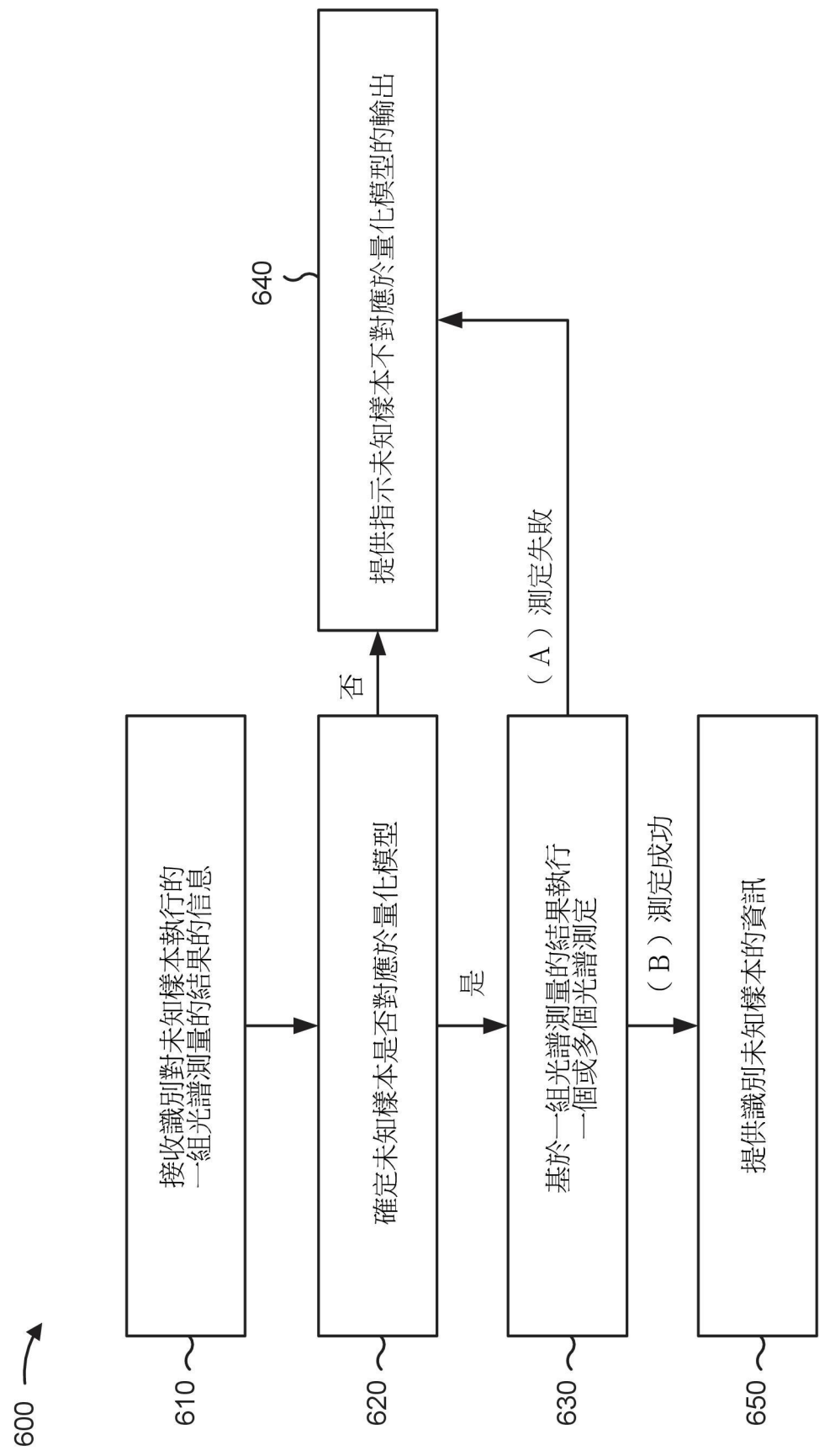


圖6

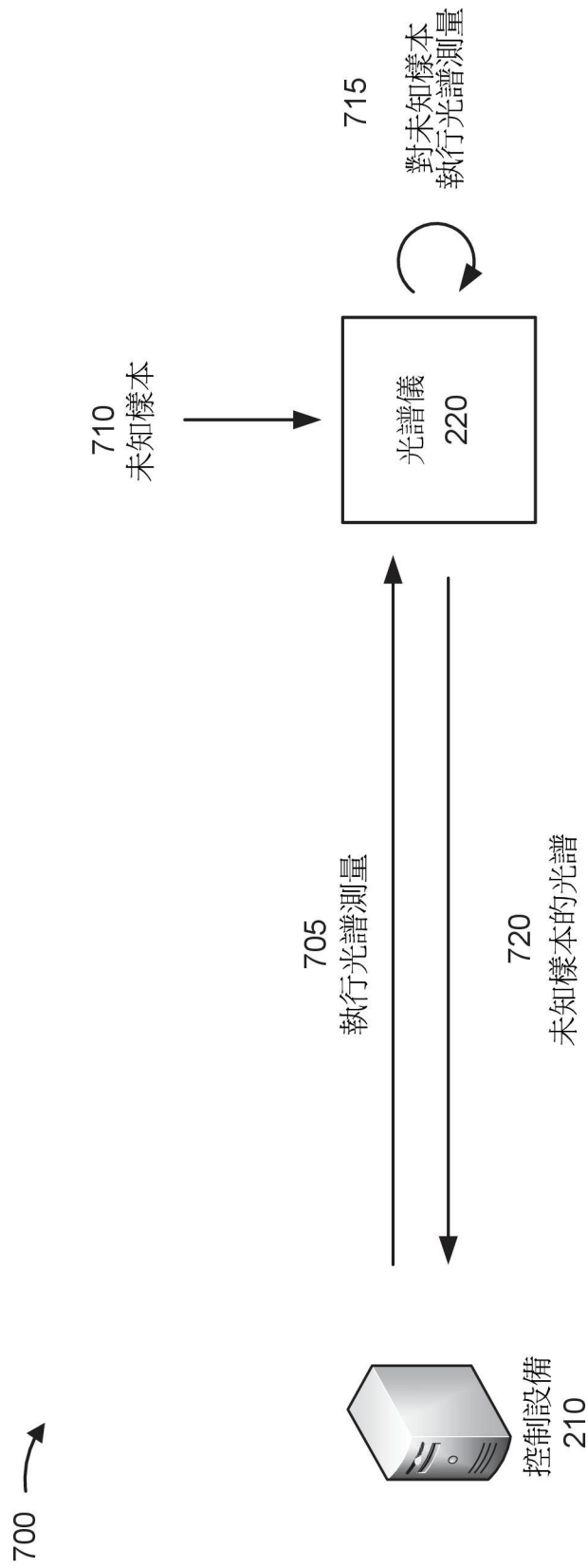


圖7A

700 →

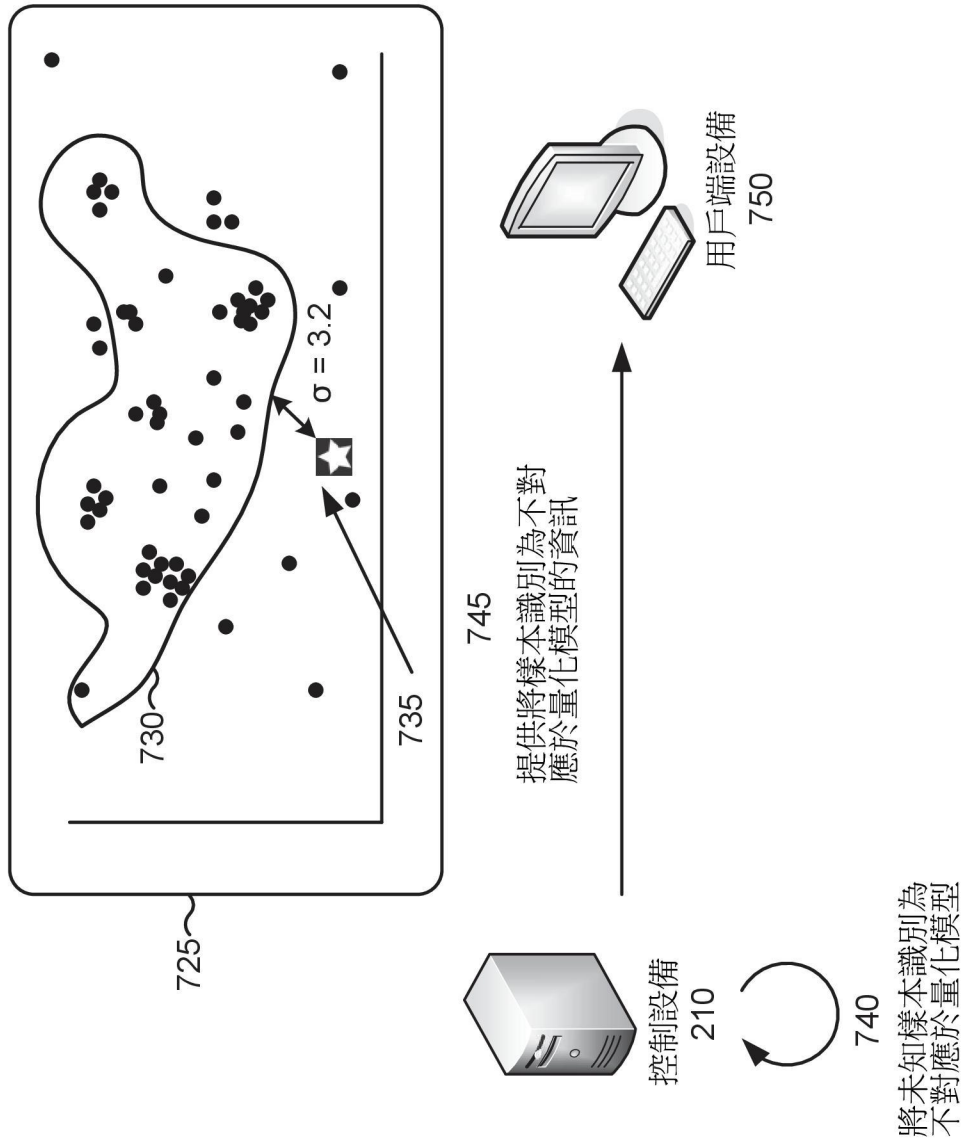


圖7B