



(10) 授权公告号 CN 110662962 B

(45) 授权公告日 2022.05.17

(21) 申请号 201880032114.0

(22) 申请日 2018.03.14

(65) 同一申请的已公布的文献号
申请公布号 CN 110662962 A

(43) 申请公布日 2020.01.07

(30) 优先权数据
62/471,274 2017.03.14 US(85) PCT国际申请进入国家阶段日
2019.11.14(86) PCT国际申请的申请数据
PCT/US2018/022293 2018.03.14(87) PCT国际申请的公布数据
WO2018/170035 EN 2018.09.20(73) 专利权人 沙特阿拉伯石油公司
地址 沙特阿拉伯宰赫兰(72) 发明人 李伟昌 塞巴斯蒂安·丘陶克
大卫·雅可比 马克斯·德芬鲍
蒂凡妮·道恩·麦卡尔平
香农·李·艾希曼(74) 专利代理机构 北京天昊联合知识产权代理
有限公司 11112
专利代理师 张娜 李荣胜(51) Int.Cl.
G01N 33/24 (2006.01)
G16C 20/30 (2019.01)
G16C 20/70 (2019.01)(56) 对比文件
US 2016349174 A1, 2016.12.01
US 6629086 B1, 2003.09.30
US 2011139989 A1, 2011.06.16
CN 105064984 A, 2015.11.18
US 2014365409 A1, 2014.12.11
US 2010313633 A1, 2010.12.16
US 2016069177 A1, 2016.03.10
CN 101949832 A, 2011.01.19
CN 1419209 A, 2003.05.21
CN 105758703 A, 2016.07.13
WO 2014006592 A2, 2014.01.09
CN 205049133 U, 2016.02.24 (续)

审查员 张红梅

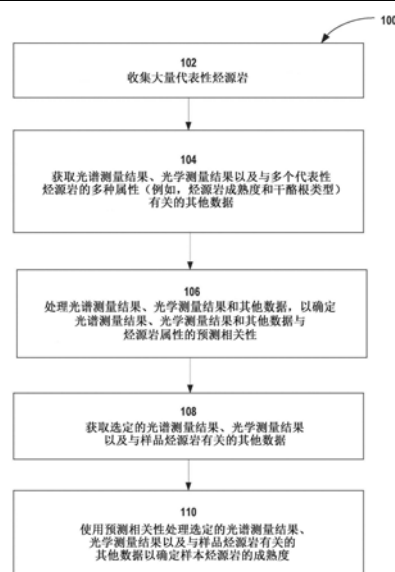
权利要求书3页 说明书21页 附图38页

(54) 发明名称

用于感测和预测烃源岩的成熟度的系统和
方法

(57) 摘要

提供了用于感测和预测烃源岩属性的系统、设备和计算机实施的方法。这里公开了一种预测烃源岩的成熟度的方法,该方法包括:从放置在样本烃源岩附近的多个数据获取装置获得样本烃源岩的多个数据、以及使用预测相关性来分析所接收的数据以确定样本烃源岩的成熟度。通过应用机器学习模型将从多个代表性烃源岩获得的多个数据与多个代表性烃源岩的多个属性相关联来生成预测相关性。



[接上页]

(56) 对比文件

CN 101839899 A, 2010.09.22
WO 2010063570 A1, 2010.06.10
CN 105247546 A, 2016.01.13
WO 2014113482 A1, 2014.07.24
CN 1820247 A, 2006.08.16
WO 2011005353 A1, 2011.01.13
CN 101689102 A, 2010.03.31
CN 105064984 A, 2015.11.18
CN 105242026 A, 2016.01.13
CN 104198677 A, 2014.12.10
CN 105242026 A, 2016.01.13
CN 102028454 A, 2011.04.27
GB 1330804 A, 1973.09.19
FR 2180288 A5, 1973.11.23
US 2016084081 A1, 2016.03.24
WO 2011133421 A2, 2011.10.27
WO 2011035391 A1, 2011.03.31
US 2010235154 A1, 2010.09.16
WO 2009094064 A1, 2009.07.30
US 2013064040 A1, 2013.03.14
EL SEBAKHY E A et al. Forecasting PVT

properties of crude oil systems based on support vector machines modeling scheme.

《Journal of petroleum science and engineering》.2009,第64卷(第1-4期),

Bahram Alizadeh et al. Artificial neural network modeling and cluster analysis for organic facies and burial

history estimation using well log data:

A case study of the South Pars Gas Field,

Persian Gulf, Iran.《Computers &

Geosciences》.2011,第45卷

Soodabeh Esmaili et al. Full field reservoir modeling of shale assets using advanced data-driven analytics.

《Geoscience Frontiers》.2015,第7卷(第1期),

Xian Shi et al. Application of extreme learning machine and neural networks in total organic carbon content prediction in organic shale with wire line logs. < Journal of Natural Gas Science and Engineering >.2016,第33卷

1. 一种确定样本烃源岩的成熟度的计算机实现的方法,该方法包括以下步骤:

通过数据分析引擎,建立与烃源岩数据库和放置在样本烃源岩附近的多个数据获取装置的通信链接,所述烃源岩数据库包含从多个代表性烃源岩获得的第一多个数据以及所述多个代表性烃源岩的多个属性;

通过所述数据分析引擎,从所述多个数据获取装置中获得样本烃源岩的第二多个数据;和

通过所述数据分析引擎,使用预测相关性分析所述第二多个数据,以确定所述样本烃源岩的成熟度,所述分析包括确定所述第一多个数据和所述多个属性中的至少一个的光谱波数波段以区分多个代表性烃源岩中的每一类型,所述确定包括计算所述光谱波数波段的多个波数点的权重,所述权重基于所述光谱波数波段的所述多个波数点对所述多个属性中的至少一个的灵敏度,

其中,通过所述数据分析引擎,通过应用机器学习模型将从所述多个代表性烃源岩获得的所述第一多个数据与所述多个代表性烃源岩的多个属性相关联,来生成所述预测相关性。

2. 根据权利要求1所述的计算机实现的方法,其中,所述多个数据获取装置包括光谱仪,所述光谱仪包括光源、热电检测器、以及用于反射来自所述样本烃源岩的光并将所反射的光引导至所述热电检测器的组件。

3. 根据权利要求2所述的计算机实现的方法,其中,所述热电检测器与可调滤波器集成在一起。

4. 根据权利要求2所述的计算机实现的方法,其中,用于反射来自所述样本烃源岩的光并将所反射的光引导至所述热电检测器的所述组件是衰减的全反射率单元。

5. 根据权利要求1或权利要求2所述的计算机实现的方法,还包括以下步骤:

在所述数据分析引擎通过实施离群值检测、基线校正、峰值增强和归一化中的一项或多项来分析所述第二多个数据之前,准备所述第二多个数据。

6. 根据权利要求1或权利要求2所述的计算机实现的方法,还包括以下步骤:

通过所述数据分析引擎,将样本烃源岩的所述第一多个数据和所确定的所述样本烃源岩的成熟度存储在烃源岩数据库中。

7. 根据权利要求1或权利要求2所述的计算机实现的方法,其中,所述多个代表性烃源岩的所述多个属性包括干酪根分型和元素组成。

8. 根据权利要求1或权利要求2所述的计算机实现的方法,其中,所述第一多个数据包括从所述多个代表性烃源岩获得的位置数据、光谱测量结果和光学测量结果中的两个或更多个。

9. 根据权利要求8所述的计算机实现的方法,其中,所述光谱测量结果包括从傅立叶变换红外光谱、电子自旋共振光谱、太赫兹光谱和紫外光谱获得的测量结果中的一个或多个。

10. 根据权利要求8所述的计算机实现的方法,其中,所述第一多个数据还包括热解数据。

11. 根据权利要求10所述的计算机实现的方法,其中,所述热解数据是通过所述多个代表性烃源岩进行Rock-Eval®热解分析来获得的。

12. 根据权利要求1或权利要求2所述的计算机实现的方法,其中,所述第二多个数据包

括从所述样本烃源岩获得的位置数据、光谱测量结果和光学测量结果中的两个或更多个。

13. 根据权利要求12所述的计算机实现的方法, 其中, 所述光谱测量结果包括从傅立叶变换红外光谱、电子自旋共振光谱、太赫兹光谱和紫外光谱获得的测量结果中的一个或多个。

14. 根据权利要求12所述的计算机实现的方法, 其中, 所述光学测量结果包括通过荧光显微镜和共聚焦激光扫描显微镜获得的测量结果中的一个或多个。

15. 根据权利要求1或权利要求2所述的计算机实现的方法, 其中, 所述机器学习模型基于支持向量机、Random Forest®、逻辑回归和自适应增强算法中的一种或多种。

16. 根据权利要求1或权利要求2所述的计算机实现的方法, 还包括以下步骤:

选择光谱波数波段, 以用于所述样本烃源岩附近的所述多个数据获取装置的操作。

17. 根据权利要求16所述的计算机实现的方法, 其中, 响应于通过所述数据分析引擎接收到来自用户界面的所述样本烃源岩的期望成熟度和期望有机相分布的一个或多个选择, 来选择所述样本烃源岩的光谱波数波段。

18. 一种确定样本烃源岩的成熟度的系统, 该系统包括:

多个数据获取装置, 其放置在样本烃源岩附近并通信地耦接到计算装置;

所述计算装置经由通信网络耦接到烃源岩数据库, 并且被配置为:

从所述多个数据获取装置获得样本烃源岩的第一多个数据; 和

使用预测相关性分析所述第一多个数据以确定所述样本烃源岩的成熟度, 其中, 分析所述第一多个数据的操作包括确定所述第一多个数据和多个属性中的至少一个的光谱波数波段以区分多个代表性烃源岩中的每一类型, 所述确定包括计算所述光谱波数波段的多个波数点的权重, 所述权重基于所述光谱波数波段的所述多个波数点对所述多个属性中的至少一个的灵敏度,

其中, 通过应用机器学习模型将从多个代表性烃源岩获得的第二多个数据与所述多个代表性烃源岩的多个属性相关联来生成所述预测相关性; 和

所述烃源岩数据库, 其包含与所述多个代表性烃源岩相关联的所述第二多个数据、所述多个代表性烃源岩的多个属性、以及所述预测相关性。

19. 根据权利要求18所述的系统, 其中, 所述多个数据获取装置被定位成从所述样本烃源岩的最佳感测波段获得数据。

20. 根据权利要求18或权利要求19所述的系统, 还包括: 样本烃源岩检索设备, 以获得所述样本烃源岩的一部分。

21. 根据权利要求18或权利要求19所述的系统, 其中, 所述多个数据获取装置被定位成获得位置数据、光谱测量结果和光学测量结果中的两个或更多个。

22. 根据权利要求21所述的系统, 其中, 所述光谱测量结果包括从傅立叶变换红外光谱、电子自旋共振光谱、太赫兹光谱和紫外光谱获得的测量结果中的一个或更多个。

23. 根据权利要求21所述的系统, 其中, 所述光学测量结果包括通过荧光显微镜和共聚焦激光扫描显微镜获得的测量结果中的一个或更多个。

24. 根据权利要求18或权利要求19所述的系统, 其中, 所述多个数据获取装置包括光谱仪, 所述光谱仪包括光源、热电检测器、以及用于反射来自所述样本烃源岩的光并将所反射的光引导至所述热电检测器的组件。

25. 根据权利要求24所述的系统,其中,所述热电检测器与可调滤波器集成在一起。

26. 根据权利要求24所述的系统,其中,用于反射来自所述样本烃源岩的光并且将所反射的光引导至所述热电检测器的所述组件是衰减的全反射率单元。

27. 一种确定样本烃源岩的成熟度的系统,该系统包括:

现场气体数据获取装置,其放置在样本烃源岩附近并通信地耦接到计算装置;

所述计算装置经由通信网络耦接到烃源岩数据库,并且被配置为:

从所述现场气体数据获取装置获得样本烃源岩的第一多个数据;和

使用预测相关性分析所述第一多个数据以确定所述样本烃源岩的成熟度,其中,分析所述第一多个数据的操作包括确定所述第一多个数据和多个属性中的至少一个的光谱波数波段以区分多个代表性烃源岩中的每一类型,所述确定包括计算所述光谱波数波段的多个波数点的权重,所述权重基于所述光谱波数波段的所述多个波数点对所述多个属性中的至少一个的灵敏度,

其中,通过应用机器学习模型将从多个代表性烃源岩获得的第二多个数据与所述多个代表性烃源岩的多个属性相关联来生成所述预测相关性;和

所述烃源岩数据库,其包含与所述多个代表性烃源岩相关联的所述第二多个数据、所述多个代表性烃源岩的多个属性、以及所述预测相关性。

28. 根据权利要求27所述的系统,其中,所述现场气体数据获取装置包括光谱仪,所述光谱仪包括光源、热电检测器、气体入口、气体出口和样本室。

29. 根据权利要求27或权利要求28所述的系统,其中,所述现场气体数据获取装置被部署为随钻测井组件的一部分。

30. 根据权利要求27或权利要求28所述的系统,其中,所述现场气体数据获取装置被部署为电缆测井组件的一部分。

用于感测和预测烃源岩的成熟度的系统和方法

技术领域

[0001] 这里公开了通常针对烃源岩属性的感测和预测的方法、设备和系统。

背景技术

[0002] 在常规储层中,从砂岩或石灰石中的地层或构造圈闭中回收油气。烃类是从盆地中一些较深的烃源岩中产生后迁移并聚集在这些储层中的。在非常规储层中,烃源岩既是岩石结构中的源也是储层。从每种类型中提取的烃类的价值高度依赖于与每种类型相关的烃源岩的属性。要了解,预测和解释每种类型的储层所产生的烃类的属性,就需要使用多种方法来分析烃源岩,以确定烃源岩的熟度和类型。熟度和烃源岩类型是对所产生的烃类的流体属性具有最大影响的参数,例如油气比 (Gas/oil ratio, GOR)、气的湿度或干度、以及粘度,所有这些都会影响烃类的流动性和质量。例如,在非常规储层中成功的油气生产由冷凝物/气体混合物支配。这主要与II型海洋烃源岩有关。这些烃源岩在后期油成熟阶段发生了转变,在该阶段,最大的干酪根 (kerogen) 内部孔隙度因该成熟周期而发展。在该成熟周期中产生的储层压力是由油气裂解引起的,这为驱出所产生的烃类提供了所需的气体驱动。该内部压力导致更大的储气能力,可以通过水力压裂来利用该储气能力使岩石破碎并回收烃类。产出流体是非常轻质的低粘度和高GOR的原油,其中包含大量的湿气 (wet gas)。湿气更容易被提炼成有价值的产品。相比之下,常规储层中的石油密度较高,其是从较不成熟的油源类型迁移而来的。由于高粘度和低GOR降低了储层中流体的流动性,因此这些储层在回收期间可能会带来昂贵的生产挑战。由于除去过量的树脂和沥青质并且在精炼过程中需要裂解较重的烃类,精炼这类石油也很昂贵,从而降低了最终生产产品的净值。

[0003] 商业开采取决于对有效烃源岩的识别,所述有效烃源岩包含期望的有机相分布 (organofacies profile) 和期望的成熟度、并且当前正在产生、或者具有产生烃类的能力。因此,储层和烃源岩的矿物学信息和有机地球化学信息对于烃类的评估和优化生产至关重要。获得烃源岩属性的标准方法是通过对均质、破碎/粉碎的样本进行大量测量。粉碎的岩石需要经过分离岩石样本中有机组分的提取过程。分析所提取的有机组分以确定烃源岩样本的成熟度和有机相分布。

发明内容

[0004] 在评估提取的和粉碎的烃源岩样本中认识到一些风险。这些风险包括但不限于分析过程中所提取的组分的改变以及提取过程中组分对分析的污染。此外,样本是从储层中提取的,稍后在地表进行分析,或者送到实验室之后再进行分析。此过程可能会花费很长时间,并且无法及时获得信息以告知有关收集位置处的井的完工决定。此过程也很昂贵,因此有时会在几个选定的样本位置收集矿物学信息和成熟度信息。由于采样有限,可能会遗漏有关岩石的重要信息。因此,需要一种电缆测井工具和其他数据获取装置,它们可以在数小时内提供有关储层间隔内的矿物学和成熟度的信息,并且可以获得沿井的完整属性分布,而不是选择位置处的一小组测量。还认识到需要套管测井工具,该套管测井工具可以测量

产出流体的属性,因为这些属性在生产期间随时间变化以指示储层的连通性以及是否可能错过了油。

[0005] 光谱学测量涉及光源、光的反射或光通过样本的透射、以及通过检测器对光强度的检测。此外,光源是单色的,或者使检测器具有波长选择性,因此可以观察到光(无论是通过反射还是透射)的衰减是波长的函数。当前用于中长红外(IR)波长的实验室光谱仪器通常使用半导体光电二极管检测器。这些检测器由诸如碲化汞镉、砷化镓或砷化镓等材料制成,必须(例如,用热电或液氮)将这些材料冷却至环境温度以下,以达到可用的信噪比。

[0006] 在某些实施例中,这些检测器被其他类型的检测器代替,以使得能够将利用适当尺寸的组件的快速光谱测量用在电缆工具中,并且可以在井下温度下操作。进一步认识到,可以利用一小部分波长确定储层岩石和流体的组成,而全光谱可能不是必需的。因此,可以将井下光谱仪设计为在少数选定波长下提供吸收。由于仅需要几个波长,因此使用该数据获取装置在每个波长上进行更长的观察时间是可能的,从而有时间平均出通过较高的井下温度所产生的较高噪声水平。井下光谱仪的一个实施例包括至少一个光源、一个检测器、一个反射来自感兴趣材料(流体或岩石)的光并将其反射到检测器的组件、将光源和检测器部署到井中的组件、以及检索光谱测量中获得的数据的组件。

[0007] 这里公开的是系统、计算机实现的方法以及具有存储的计算机程序的非暂时性计算机可读介质的实施例。这些实施例旨在解决本领域的缺点,包括用于烃源岩样本的光谱和光学测量以通过感测设备和专用算法确定其属性的特定方法。这些方法和系统提供了对勘探、开发和储层生产至关重要的空间准确及时的特征。

[0008] 这里的公开提供了用于确定样本烃源岩的成熟度的计算机实现的方法。一种这样的方法包括以下步骤:通过数据分析引擎,建立与烃源岩数据库和放置在样本烃源岩附近的多个数据获取装置的通信链接。烃源岩数据库包含从多个代表性烃源岩获得的第一多个数据以及多个代表性烃源岩的多个属性。该方法还包括以下步骤:通过数据分析引擎,从多个数据获取装置中获得样本烃源岩的第二多个数据;以及,通过数据分析引擎,使用预测相关性分析第二多个数据,以确定样本烃源岩的成熟度。数据分析引擎通过应用机器学习模型将从多个代表性烃源岩获得的第一多个数据与多个代表性烃源岩的多个属性相关联,从而生成预测相关性。多个数据获取装置可以包括光谱仪,该光谱仪具有光源、热电检测器、以及用于反射来自样本烃源岩的光并将反射的光引导至热电检测器的组件。热电检测器可以与可调滤波器集成在一起。在某些实施例中,用于反射来自样本烃源岩的光并将反射的光引导至热电检测器的组件是衰减的全反射率单元。在某些实施例中,该方法还可以包括以下步骤:在数据分析引擎通过实施离群值检测、基线校正、峰增强和归一化来分析第二多个数据的步骤之前,准备第二多个数据。该方法还包括以下步骤:通过数据分析引擎,将样本烃源岩的第一多个数据和确定的样本烃源岩的成熟度存储在烃源岩数据库中。多个代表性烃源岩的属性可以包括干酪根分型和元素组成。第一多个数据可以包括从多个代表性烃源岩获得的位置数据、光谱测量结果和光学测量结果中的两个或更多个。光谱测量结果可包括从傅立叶变换红外光谱(FTIR)、电子自旋共振光谱(ESR)、太赫兹光谱(THz)和紫外(UV)光谱获得的测量结果中的一个或更多个。在某些实施例中,第一多个数据还包括热解数据。可通过对多个代表性烃源岩进行Rock-Eval®热解分析来获得热解数据。在某些实施例中,第二多个数据包括从样本烃源岩获得的位置数据、光谱测量结果和光学测量结果

中的两个或更多个。光谱测量结果可以包括从傅立叶变换红外光谱、电子自旋共振光谱、太赫兹光谱和紫外光谱获得的测量结果中的一个或多个。光学测量结果可包括通过荧光显微镜和共聚焦激光扫描显微镜获得的测量结果中的一个或多个。某些实施例包括基于以下中的一种或多种的机器学习模型：支持向量机、RandomForest®、逻辑回归、以及自适应增强算法(Adaptive Boosting algorithms)。该方法的某些实施例还包括以下步骤：选择光谱波数波段，以用于样本烃源岩附近的多个数据获取装置的操作。可以响应于由数据分析引擎接收到来自用户界面的样本烃源岩的期望成熟度和期望有机相分布的一个或多个选择，来选择样本烃源岩的光谱波数波段。

[0009] 实施例包括确定样本烃源岩的成熟度的系统。一个这样的系统包括放置在样本烃源岩附近并通信地耦接到计算装置的多个数据获取装置。该计算装置经由通信网络耦接到烃源岩数据库。所述计算装置被配置为从所述多个数据获取装置获得样本烃源岩的第一多个数据；使用预测相关性分析第一多个数据以确定样本烃源岩的成熟度。烃源岩数据库包含与多个代表性烃源岩相关的第二多个数据、多个代表性烃源岩的多个属性、以及通过应用机器学习模型将从多个代表性烃源岩获得的第二多个数据与多个代表性烃源岩的多个属性相关联而生成的预测相关性。多个数据获取装置可以被定位成从样本烃源岩的最佳感测波段获得数据。该系统还可以包括样本烃源岩检索设备，以获得样本烃源岩的一部分。多个数据获取装置可以被定位成获得位置数据、光谱测量结果和光学测量结果中的两个或更多个。光谱测量结果可以包括从傅立叶变换红外光谱、电子自旋共振光谱、太赫兹光谱和紫外光谱获得的测量结果中的一个或多个。光学测量结果可以包括通过荧光显微镜和共聚焦激光扫描显微镜获得的测量结果中的一个或多个。数据获得装置可以包括光谱仪，该光谱仪包括光源、热电检测器、以及反射来自样本烃源岩的光并将反射光引导至热电检测器的组件。热电检测器可以与可调滤波器集成在一起。反射来自样本烃源岩的光并将反射光引导至热电检测器的组件可以是衰减的全反射率单元。

[0010] 确定样本烃源岩的成熟度的另一系统包括放置在样本烃源岩附近并通信地耦接到计算装置的现场气体数据获取装置。该计算装置经由通信网络耦接到烃源岩数据库，并且被配置为从现场气体数据获取装置获得样本烃源岩的第一多个数据并使用预测相关性来分析第一多个数据以确定样本烃源岩的成熟度。烃源岩数据库包含与多个代表性烃源岩相关的第二多个数据、多个代表性烃源岩的多个属性、以及通过应用机器学习模型将从多个代表性烃源岩获得的第二多个数据与多个代表性烃源岩的多个属性相关联而生成的预测相关性。现场气体数据获取装置可以包括光谱仪，该光谱仪包括光源、热电检测器、气体入口、气体出口、以及样本室。在某些实施例中，现场气体数据获取装置被部署为随钻测井组件的一部分。在某些实施例中，现场气体数据获取装置被部署为电缆测井组件的一部分。

[0011] 根据以下结合附图的详细描述，本公开的许多其他方面、特征和益处可以变得显而易见。根据所需的分析目标，系统可以包括更少的组件、更多的组件或不同的组件。

附图说明

[0012] 尽管本公开内容可以进行各种修改和替代形式，但是在附图中以示例的方式示出了具体实施方式，并且在此将对这些具体实施方式进行详细描述。附图可能未按比例绘制。然而，应理解，附图及其详细描述并非旨在将本公开限制为所公开的特定形式，而是相反，

其意图是涵盖落在如所附权利要求书所定义的本公开的精神和范围内的所有修改、等同形式和替代形式。

[0013] 图1是示出根据实施例的用于确定烃源岩样本的某些属性的方法的流程图。

[0014] 图2是示出根据另一实施例的用于确定烃源岩样本的某些属性的方法的流程图。

[0015] 图3是示出根据另一实施例的用于选择光谱波数波段以用于样本烃源岩附近的多个数据获取装置的操作的方法的流程图。

[0016] 图4A和图4B是示出从两个代表性烃源岩样本的热解获得的数据的样本热解图。

[0017] 图5是示出与相对于烃源岩的化学官能团相对应的各种波数的IR吸收光谱。

[0018] 图6是描绘随着烃源岩样本的成熟度增加而向较低波数的偏移的图形表示。

[0019] 图7A是根据各自的成熟度从上到下增加的四组烃源岩的FTIR光谱图。图7B是图7A中分析的四组大量烃源岩的FTIR光谱的放大图。图7C是在图7A中分析的四组大量烃源岩的FTIR光谱的进一步放大图。

[0020] 图8A是几种烃源岩粉末样本以及提取的干酪根的FTIR光谱。图8B是纯粘土粉末样本的FTIR光谱。

[0021] 图9A描绘了太赫兹基准(高密度聚乙烯(HDPE))的测量时间轨迹和根据各种烃源岩样本分析的脉冲时间波形,并且图9B是通过对波形进行傅里叶变换得到的相应光谱。

[0022] 图10A和图10B描绘了在太赫兹波段中在图9A和9B中分析的各种烃源岩样本的样本吸收率和透射率光谱。

[0023] 图11A是示出烃源岩的层理(红色区域富含有机物)的二维(2D)岩心截面图。图11B是相对于被标识为样本1A、样本2A、样本3A和样本4A的四种不同成熟度水平的烃源岩的烃指数绘制的镜质组反射率的图形表示。

[0024] 图12A至图12C提供了对烃源岩和油进行荧光测量的示例。

[0025] 图13是示出根据实施例的预测样本烃源岩的成熟度和有机相分布的方法的流程图。

[0026] 图14A至图14D是来自烃源岩样本3A的FTIR数据的预处理的表示。

[0027] 图15是跨太赫兹(THz)、IR和紫外(UV)波段的一种烃源岩样本的宽带输入光谱的表示。

[0028] 图16A和图16B描绘了降维空间中烃源岩样本聚类的两种表示。

[0029] 图17是分层的双簇图。

[0030] 图18是来自FTIR光谱测量的分级特征的权重与光谱波数波段的对准的表示,以区分处于不同成熟度水平的各种粘土、矿物和干酪根。

[0031] 图19A和图19B是来自投影在选定波数轴上的不同样本的FTIR光谱的簇图的表示。图19C是从18种不同类型的样本的样本FTIR光谱的分层聚类获得的示例树状图。

[0032] 图20A和图20B以混淆矩阵的形式呈现了使用两种不同的机器学习算法--自适应增强方法(如图20A所示)和Random Forest®方法(如图20B所示)对烃源岩样本进行分类后得出的评估结果。

[0033] 图21表示与通过传统方法处理各种烃源岩获得的成熟度指数相比通过两种不同的机器学习算法(自适应增强方法和Random Forest®方法)根据烃源岩样本的FTIR光谱预测的成熟度指数(氢指数(HI)与镜质组反射率(Ro%)的比率)。

[0034] 图22A表示根据实施例的简化的热模型,并且图22B表示根据该实施例的热电检测器的等效电路。

[0035] 图23是根据实施例的包含与可调滤波器集成的商用热电检测器的设备的示意图。

[0036] 图24是根据实施例的衰减全反射率(ATR)单元的单反射配置的示意图。

[0037] 图25是根据实施例的衰减全反射率(ATR)单元的多反射配置的示意图。

[0038] 图26是根据实施例的使用热电传感器的GIP数据获取装置的示意图。

[0039] 图27A、图27B和图27C分别是根据实施例的GIP数据获取装置、光源和热电检测器的实验室原型的照片。

[0040] 图28A和图28B分别是根据实施例的使用乙烷获得的IR光谱和相关的系统校准光谱的表示。

[0041] 图29是根据实施例的采样装置的示意图。

具体实施方式

[0042] 在以下描述中,阐述了许多具体细节以便提供对各种实施例的透彻理解。在其他情况下,可能不会特别详细地描述众所周知的过程和方法,以免不必要地混淆这里描述的实施例。另外,这里的实施例的图示可以省略某些特征或细节,以免模糊这里描述的实施例。

[0043] 在下面的详细描述中,参考形成说明书一部分的附图。在不脱离本公开的范围的情况下,可以利用其他实施例,并且可以进行逻辑改变。因此,下面的详细描述将不具有限制意义。

[0044] 描述可以使用短语“在一些实施例中”、“在各种实施例中”、“在某些实施例中”或“在实施例中”,其可以各自指代相同或不同实施例中的一个或多个。此外,关于本公开的实施例使用的术语“包括”、“包含”、“具有”等是同义的。

[0045] 开采和提取最大价值的烃类的成功需要了解给定盆地内非常规储层和常规储层的烃源岩。系统和方法提供了高等级的目标烃源岩,这需要开发涉及机器学习方法和化学计量学方法的智能系统。在此框架中,用于确定成熟度和源类型的所有分析数据都经过整合和分析,以得出有关成熟度和来源的确定。该系统可用于预测可能的API重力指标(API gravity)、油气比和烃类的粘度,然后评估提取、生产和精炼目标烃类的成本,并为这些预测值提供不确定性量化。

[0046] 如本文所用,术语“烃源岩”是指能够产生或已经产生可活动量的烃类的岩石。在数据库中记录了大量代表性烃源岩样本、它们的位置(例如井位的种类)、它们的有机相分布、它们的年龄和成熟度、光谱测量结果以及其他代表性烃源岩测量结果,并将其用于建立光谱测量与烃源岩属性的相关性。

[0047] “数据分析引擎”是指一个或多个软件模块,该模块处理数据,制定模型和规则,并通过使用适当的逻辑和准则执行数据匹配、训练和交叉验证,包括但不限于用于预测烃源岩的属性的软件。在一些实施例中,数据分析引擎可以被实现为服务器、用户计算装置等的一部分。数据分析引擎的适当实现的示例包括服务器、授权的用户计算装置、智能手机、台式计算机、膝上型计算机、平板计算机、PDA、以及其他类型的接收、处理或传输数字数据的处理器控制的装置。

[0048] 岩石中矿物质含量和有机物成熟度的表征对于确定页岩储层中的分支的最佳深度、页岩储层中的井的水力压裂的最佳位置以及精炼盆地模型和对烃类系统的地质理解而言是重要的。在页岩储层中,有机物的成熟度和数量可指示干酪根的孔隙度和原处的气体。当物质超过一定的成熟度水平时,孔隙度可能会崩溃并且可能不存在气体。矿物学表明页岩可以如何有效地破裂。产出流体的组成的表征可指示流体的物理属性,这对于设计完井、人工举升和地面设施可能很重要。两个生产区之间或两个井之间的流体组成不同可指示这些区或井不连通。它还可指示,相应的流体来自不同的烃源岩,或者在注入后具有不同的温度、生物降解或其他变化。

[0049] 本公开的某些方面包括用于基于光谱测量和光学测量来预测烃源岩属性的方法、系统和软件产品。某些实施例包括特定的方法、系统和软件产品,包括使用光谱测量和光学测量以及传感技术对烃源岩样本进行多次测量的集成,以对烃源岩进行地球化学表征和矿物学表征,例如基于烃源岩的成熟度和有机相分布对烃源岩进行分类。

[0050] 本公开中的方法、系统和软件产品涉及用于基于选定的频谱测量来确定选定的烃源岩属性的技术。这里还包括烃源岩属性数据库,该数据库用于生成光谱测量结果与选定烃源岩属性的预测相关性。烃源岩数据库是一个这样的数据库,其包含代表不同成熟度水平和沉积成因的大量烃源岩样本的若干特征的测量结果。该数据库包括代表性烃源岩的测量结果,例如选定的光谱测量结果、光学测量结果和烃源岩属性的标准测量结果。这些属性包括烃源岩的成熟度和烃源岩中存在的干酪根类型。使用计算机算法处理光谱测量结果、光学测量结果和标准测量结果,以开发光谱测量结果的选定方面与烃源岩属性的相关性。在光谱测量结果与烃源岩属性之间开发的相关性用于基于烃源岩的光谱测量结果和光学测量结果来预测未知烃源岩的烃源岩属性。

[0051] 烃源岩数据库可以被配置为外部源和内部数据库的组合,并且被实现为面向对象的、网络的、半结构的或其他灵活的数据库,或者它们的组合,以提供获取、索引和存储数据的功能。在这些实施例中,烃源岩数据库向系统中的一个或多个软件模块提供数据以及其他存储的数据和文件。外部数据源可以是例如单个数据库、多个数据库或虚拟数据库,其中包括来自多个源(例如,万维网上的服务器)的数据。在这些实施例中,可以使用应用协议来实现烃源岩数据库,该应用协议用于访问和维护代表性烃源岩的光谱测量结果、光学测量结果和标准测量结果等等。本领域技术人员将理解,本公开中描述的存储在数据库中的数据也可以存储或维护在非暂时性存储器中,并且可以在子例程、函数、模块、对象、程序产品或过程之间进行访问,例如,根据这些子例程、函数、模块、对象、程序产品或过程的对象或变量进行访问。

[0052] 本公开中的方法、系统和软件产品的某些实施例涉及仅针对对成熟度、干酪根类型和其他地质属性(例如有机相)具有重要性的波数波段的子集的新的感测技术。基于由大量代表性样本和多种光谱测量类型组成的光谱数据库,已开发出多种新的算法来识别多种不同测量(例如FTIR、花期(Florescence)、THz和ESR数据)之间的信息重叠和空白。这是通过对适用于多种光谱测量的算法进行协作学习(特别是借助高维特征学习算法、分类算法和预测算法)来实现的。得出了最佳的感测波段和预测因子作为约束优化问题的解决方案,该解决方案在现场可访问的约束条件下最大化了与烃源岩属性(成熟度和干酪根类型)有关的总信息。

[0053] 当必须表征未知的烃源岩样本时,对整个岩石样本进行光谱测量,并将该光谱测量与从烃源岩样本的测试获得的烃源岩属性相关联。在某些实施例中,仅基于未知样本的整个岩石部分的光谱测量结果来确定未知样本的烃源岩属性。在某些实施例中,在烃源岩的现场进行测量。这些实施例提供了成本有效且可靠的多光谱原位感测、烃源岩成熟度和有机相分布的准确预测、准确的储量和生产率估算,这导致了烃类的产量增加、优化的采收方法和更大的风险缓解。

[0054] 在某些实施例中,确定样本烃源岩属性的计算机实现的方法包括由数据分析引擎执行的以下步骤:从放置在样本烃源岩附近的多个数据获取装置获得样本烃源岩的第一多个数据,并使用预测相关性分析所接收的数据以确定样本烃源岩的成熟度。所接收的数据通过进行离群值检测、基线校正、峰值增强和归一化来进行处理,然后再由数据分析引擎进行分析。通过数据分析引擎,通过应用机器学习模型将从多个代表性烃源岩获得的数据与多个代表性烃源岩的多个属性相关联来生成预测相关性。烃源岩数据库存储从多个代表性烃源岩获得的数据、各种代表性烃源岩的属性、接收的样本烃源岩的数据以及确定的样本烃源岩的成熟度。

[0055] 在某些实施例中,确定样本烃源岩的成熟度的计算机实现的方法可以包括选择光谱波数波段,以用于样本烃源岩附近的多个数据获取装置的操作。响应于从用户界面接收样本烃源岩的期望的成熟度和期望的有机相分布的一个或多个选择,来选择样本烃源岩的光谱波数波段。

[0056] 在某些实施例中,光谱技术包括FTIR、THz、ESR和荧光光谱中的一种或多种。所测试的属性的示例包括镜质组反射率、烃指数、元素组成、热解和锂皂石荧光。这些测试提供了烃源岩成熟度和干酪根类型的指示。诸如高效液相色谱(HPLC)、气相色谱(GC)和GC质谱(GC-MS)之类的其他分析方法可用于表征代表性烃源岩样本的干酪根。为了诸如各种光谱测量的子带之类的原位场感测和预测的目的,导出了一组最佳感测属性。在某些实施例中,使用FTIR光谱、荧光光谱、THz光谱、ESR光谱、能量色散X射线(EDS或EDX)光谱中的一种或多种来评估岩石样本。也使用Rock-Eval®热解来评估岩石样本。

[0057] 这里公开的实施例还包括与数据获取装置相关联的计算机系统,该计算机系统包括存储器、处理器以及一个或多个输入/输出(I/O)接口。存储器可以包括非易失性存储器(例如,闪存、只读存储器(ROM)、可编程只读存储器(PROM)、可擦除可编程只读存储器(EPROM)、电可擦除可编程只读存储器(EEPROM))、易失性存储器(例如,随机存取存储器(RAM)、静态随机存取存储器(SRAM)、同步动态RAM(SDRAM))、大容量存储存储器(例如,CD-ROM、DVD-ROM或硬盘)驱动器)、以及它们的组合。存储器可以包括其中存储有程序指令的非暂时性计算机可读存储介质。程序指令可以包括可由计算机处理器执行以引起这里描述的功能操作的程序模块,包括关于确定烃源岩样本的成熟度所描述的那些功能操作。

[0058] 处理器可以是能够运行/执行程序指令的任何合适的处理器。该处理器可以包括中央处理单元(CPU),该中央处理单元运行程序指令(例如,用于图1、图2、图3和图13中所示方法的程序指令)以执行在此描述的算术操作、逻辑操作和输入/输出操作。该处理器可以包括一个或多个处理器。处理器可以通信地链接到一个或多个I/O装置,例如操纵杆、计算机鼠标、键盘、显示屏(例如,用于显示图形用户界面(GUI)的电子显示器)、触摸或语音响应装置等。I/O装置可以包括一个或多个用户输入装置、一个或多个数据获取装置、一个或多

个数据处理仪器及其组合。这些装置可以通过有线连接或无线连接而连接到处理器和I/O接口。I/O接口可以提供用于与一个或多个外部装置(例如,其他计算机、网络、数据获取装置、采样装置、以及它们的组合)通信的接口。在一些实施例中,I/O接口包括天线、收发器、以及通信地耦接到其他装置所需的其他组件。在一些实施例中,外部装置包括上游设施。结合这里公开的实施例描述的各种说明性的逻辑块、模块、电路和算法步骤可以被实现为电子硬件、计算机软件或两者的组合。当使这些装置能够在它们之间交换数据、控制信号或其他信息时,将在该系统中的两个或多个装置之间建立通信链接。为了清楚地说明硬件和软件的这种互换性,已经就其功能性总体地描述了各种说明性的组件、框、模块、电路和步骤。将这种功能性实现为硬件还是软件取决于特定的应用和施加在整个系统上的设计约束。技术人员可以针对每个特定应用以各种方式来实现所描述的功能,但是这种实现决定不应被解释为导致脱离本公开的范围。

[0059] 图1是示出根据实施例的用于确定烃源岩样本的某些属性的方法100的流程图。在步骤102处,收集大量代表性烃源岩。在步骤104处,数据分析引擎获取光谱测量结果、光学测量结果以及与代表性烃源岩的多种属性(例如,烃源岩成熟度和干酪根类型)有关的其他数据。在步骤106处,数据分析引擎处理光谱测量结果、光学测量结果和其他数据,以确定光谱测量结果、光学测量结果和其他数据与烃源岩属性之间的预测相关性。在步骤108处,数据分析引擎获取选定的光谱测量结果、光学测量结果和其他有关样品源岩石的数据。在步骤110处,数据分析引擎使用预测相关性处理选定的光谱测量结果、光学测量结果以及其他与样品源岩石有关的数据以确定样本烃源岩的成熟度。

[0060] 图2是示出根据实施例的用于确定烃源岩样本的某些属性的方法200的流程图。在步骤202处,获得新样本的FTIR光谱。在步骤204处,对FTIR光谱进行预处理方法,例如,离群值检测、基线校正和归一化。在步骤206处,通过加权选择波数/频段或者从烃源岩数据库训练数据中选择波数/频段。在步骤208处,针对数据库样本,对处理后的FTIR光谱进行聚类或分类。在步骤210处,基于烃源岩数据库训练模型对成熟度/有机相分布进行预测。在步骤212处,如果FTIR图形是大量样本,则在步骤214处,将类和预测属性映射到样本表面,并与其他图像(例如,能量色散X射线光谱术(EDS或EDX)或荧光)共注册和交叉验证。在步骤216处,如果共注册/交叉验证成功,则在步骤218处,将新样本以及所有获取的数据和分析添加到烃源岩数据库220。在步骤212处,如果FTIR图像不是大量样本,则在步骤218处,将来自烃源岩样本的数据添加到烃源岩数据库220中。在步骤216处,如果共注册/交叉验证未成功,则从步骤208开始重复处理。

[0061] 图3是示出根据实施例的用于确定烃源岩样本的某些属性的方法300的流程图。在步骤302处,诸如成熟度和有机相分布之类的目标属性被设置为感测或预测烃源岩样本的属性。在步骤304处,从策划的烃源岩数据库中选择数据库样本的子集作为输入。在步骤306处,特征选择算法被用来计算来自选定的数据库样本的波数/频段的权重。在步骤308处,在选择最佳波数/频段时考虑传感器设计和部署的可行性约束。在步骤310处,通过特征分级选择这些波段并将这些波段用作感测波段。这些感测波段对选定的目标属性最有帮助,或者最有助于区分烃源岩样本中的差异。

[0062] 某些实施例包括一种设备,该设备用于在井眼中的选定子带中执行选定的光谱测量,并将光谱传送到地面计算机以进行处理或可能自己执行该处理并传送成熟度和干酪根

类型的结果。已经开发了用于从一般的和选定的感测中的多次测量来预测烃源岩属性(成熟度和有机相)的一组方法,并且在某些实施例中,这些多次测量是从原位场感测来的。对评估和生产石油系统很重要的与储层和烃源岩有关的岩石矿物学信息和有机地球化学信息可从多种类型的光谱测量方法(例如FTIR、ESR、THz和UV光谱、气相色谱(GC)和质谱(MS))中得出。建立了一个数据库,其包含从代表性烃源岩样本的收集得到的这些测量结果。通过利用数据库确定新样本或未知样本的烃源岩分类。从这些测量结果的全部或子集中得出用于预测烃源岩成熟度和有机相类型的最佳属性,并将其与诸如镜质组反射率和烃类指数之类的指标进行比较。还将分子结构的行为和由不同成熟度水平引起的变化与数据库中的值进行比较。出于原位场感测和预测的目的,导出了一组最佳感测属性。确定了与最佳属性密切相关的光谱波数波段,包括来自不同类型的物理测量的那些光谱波数波段以及每种测量类型的不同子带的那些光谱波数波段。这为现场应用中为用于原位感测的波段选择提供了指南。还对所选的感测波段及其在预测成熟度和有机相类型中的准确性和确定性进行了量化。这使用基于未知样本的可用光谱检查数据的最佳属性或从原位选择的感测波段数据得出的属性来实现。

[0063] 烃源岩成熟过程的属性是通过埋藏烃源岩引起的热成因过程的干酪根结构的物理和化学转化。随后的温度梯度导致干酪根热分解成烃。通常,通过多种不同的时间密集型方法(例如,元素组成、热解、镜质组反射率和类脂组荧光剂)来测量该成熟度的进程。镜质组反射率(%Ro)是对来自干酪根中镜质组煤岩组分的光的反射率的目视显微镜估计。通过适当的校准,反射率值的范围可以反映出烃源岩的成熟度。表1中提供了指示烃源岩年龄的镜质组反射率(%Ro)值范围的示例。这些反射率值范围在不同的烃源岩和不同的干酪根之间会有所不同。

[0064] 表1

反射率值	烃源岩的年龄
0.25-0.6	不成熟
0.6-0.8	早熟
0.8-1.1	峰值成熟度
1.1-1.35	晚熟
1.35>	干气体-过成熟

[0066] 烃源岩元素组成的测量有助于确定干酪根成熟度,并表示为从那些元素的浓度测量的氢相对于碳和氧的损失。如组成图(称为Van Krevelen图)所示,氢碳比相对于氧碳比的系统下降提供了成熟度的估算。此外,H/C比相对于O/C比的行为还提供了干酪根中的煤岩组分组成估算以及未成熟烃源岩是否被分类为I型、II型、III型或IV型。这四种类型基于烃源岩样本中碳、氢和氧的相对含量。由于在烃源岩中产生了油气,因此干酪根中的氢和氧相对于碳变得贫乏。干酪根中的氢含量与烃源岩的生油潜力成正比。

[0067] 烃源岩的属性和分类与有机相有关,该有机相既反映了有机物的类型又反映了与沉积环境有关的沉积物的组成。这是根据煤岩组分类型的相对丰度和组成基质的沉积物来源来定义的。煤岩组分是保存的藻类和植物残渣的蓝图残余物(blueprint remnant),并且包含有机物,该有机物包括三种主要类型:类脂组、镜质组及惰性组。这些有机相具有诊断意义,并且提供对某些烃源岩有机相所期望的潜在烃产率和流体质量的洞悉。例如,II-S型

碳酸盐有机相(富含硫的干酪根)包含类脂组煤岩组分,其高度易挥发且将产生油。然而,与同样包含易挥发的类脂组煤岩组分且产生油的标准II型干酪根相比,其成熟度要低得多。发生这种情况的原因是与硫链接的活性能较低,硫起自由基作用用于提高反应活性。相比之下,III型硅质有机相往往会产生更多的气体,因为干酪根组成的挥发性较小,并且由于富含镜质组而具有较低的H/C比。以惰性组(inertinite)为主的IV型有机相挥发性最小,H/C更低,并且不具有烃生成潜力。因此,它不被认为是有吸引力的烃源岩。这些只是各种盆地的岩石记录中遇到的几种有机相类型中的几种。

[0068] 已经从各个井场收集了大量的烃源岩样本,对这些烃源岩样本进行这里描述的所有测量或选择的一组测量。除了这些光谱测量结果和光学测量结果,样本还具有其他相关数据,包括但不限于样本位置、深度、Rock-Eval®分析结果、FTIR光谱测量和图像、元素组成图、共聚焦荧光图像、X射线荧光图像、ESR测量、太赫兹图像和其他数据。Rock-Eval®热解技术是法国石油研究所(Institut Français du Pétrole)开发的一种广泛使用的技术,用于测量岩石样本中有机物的数量、质量和热成熟度,并且包括确定烃生成潜力、有机特征、以及热成岩的程度(成熟度)。FTIR数据包括原始光谱和预处理光谱、FTIR成像数据、以及提取样本的FTIR光谱。在执行进一步分析(例如聚类、分类、预测和解释)之前,通常需要对FTIR光谱进行预处理,以去除基线、抑制噪声并将光谱归一化。这些预处理步骤的输入光谱和输出光谱分别称为原始光谱和预处理光谱。元素组成数据包括烃源岩样本的单个矿物图和矿物分布,并使用诸如基于EDS的化学微分析之类的技术获得。

[0069] 图4A和图4B是示出从未成熟的志留纪II型干酪根烃源岩样本(例如表2中的样本3A)获得的Rock-Eval®数据的样本热解图。热解图中的峰值代表烃源岩中的可移动和不可移动的烃馏分,并且记录为热解数据(S1值和S2值)。这些值用于计算成熟指标。S1峰值表示在330°C的温度下从干酪根释放出的可移动的、可蒸馏的烃类或沥青,而S2值表示在550°C至600°C的温度下演变的干酪根的剩余生成潜力。S1值和S2值二者都以岩石的毫克每克记录。S3峰对应于由于干酪根的热裂解产生的二氧化碳,并且也以烃源岩的毫克每克记录。Tmax值是达到S2值的最高温度,并且提供了烃源岩成熟度的某种量度。表2列出了本研究中使用的烃源岩样本、以及Rock-Eval®参数和记录的热解数据。这些样本都是来自同一地质盆地中不同位置和深度的志留纪年代的II型干酪根烃源岩样本。因此,由于地理位置和埋藏深度的差异,样本经历了不同的成熟历史。选择这些样本是为了提供从不成熟到峰值气体生成的范围广泛的烃成熟度,。

[0070] 表2列出了一些烃源岩示例及其Rock-Eval®参数。这些代表了II型志留纪年代的烃源岩,涵盖了从早期石油到气体成熟度的一系列成熟度。S2值除以总有机碳(TOC)提供了氢指数(HI),该氢指数可用于估计成熟度(%Ro-HI),尤其是当T-max不易从热解图中解析时。未解析的T-max(N/A值)发生于过成熟的样本,例如样本4A和样本4D,其中的生成电势已耗尽,通过低S2值可看出,这相当于气体成熟度的低HI-%Ro(2.0-2.5)特征。同样,II型样本3A、样本3B和样本3C代表早期油窗中的样本,通过0.62至0.64的HI-%Ro可看出。相反,“峰值油(peak oil)”成熟度是样本1A的特征(HI-%Ro=1.02)以及样本2A的最终冷凝物成熟度(HI-%Ro=1.59)。

[0071] 表2

[0072] 当在实验室或现场环境中预测新烃源岩样本的成熟度或对新烃源岩样本的有机相进行分类时,可以从感兴趣的新样本中收集相似的光谱或光学数据,然后针对数据库中的烃源岩样本的数据对相似的光谱或光学数据进行分析。

[0073] 电磁波的不同频段与不同的分子响应相关联,这形成了各种光谱学和光学测量技术的物理/化学基础。UV光响应与电子跃迁的能隙相关联,并且红外波段产生的吸收光谱是分子键(例如C-H和O-H)的拉伸和弯曲模式的特征。在分子系统中导致太赫兹吸收的机制主要受分子间以及分子内振动的激发,例如大分子链中的扭转模式或相邻分子之间的分子间振动,这具有更强的聚集性,而微波响应通常与分子旋转有关。不同波长的光谱和光学测量表征了分子行为的不同方面,这些方面与由成熟过程和有机相变化引起的结构和组成变化有关。因此,当进行最佳组合时,来自不同类型测量的数据将提供更全面的表征,从而导致更准确的预测和分类。这些协作式机器学习和预测方法有助于对未知烃源岩样本进行分类。

[0074] 以下是用于表征数据库中各种烃源岩样本的某些方法的描述。FTIR是一种非破坏性分析方法,其可为识别和定量矿物学以及岩石基质中所含有机物的组成提供重要信息。特别地,它具有区分与干酪根相关的脂族碳结构和芳香族碳结构的能力。这使得能够根据干酪根类型以及生热过程产生的干酪根结构的物理转变和化学转变来对烃源

[0075]

样本	深度 (pt)	年龄	油母岩 类型	S1(mg/g)	S2(mg/g)	Tmax(°C)	HI	OI	PI	TOC(%)	%Ro- Tmax	HI-%Ro	HC范围
1A	13,5xx	志留纪	类型II	2.64	9.85	439	185	3	0.21	5.33	0.74	1.02	峰油
2A	15,1xx	志留纪	类型II	1.49	1.73	456	45	3	0.46	3.87	1.05	1.59	峰油
3A	6,3xx	志留纪	类型II	2.14	27.24	422	485	12	0.07	5.62	0.44	0.63	未成熟
3B	6,3xx	志留纪	类型II	2.2	29.31	422	476	9	0.07	6.16	0.44	0.64	未成熟
3C	6,3xx	志留纪	类型II	2.82	35.4	421	504	8	0.07	7.02	0.42	0.62	未成熟
4A	6,5xx	志留纪	类型II	0.08	0.09	N/A	4.21	0.93	0.47	2.14	N/A	2.54	气体
4B	6,5xx	志留纪	类型II	0.16	1.07	523	8.36	1.09	0.13	12.8	2.25	2.26	气体
4C	6,6xx	志留纪	类型II	0.09	0.31	534	10.84	0.35	0.23	2.86	2.45	2.16	气体
4D	6,6xx	志留纪	类型II	0.02	0.06	N/A	13.99	2.33	0.25	0.43	N/A	2.06	气体

[0076] 类。FTIR光谱通过使用振动能测量组成和成熟度,该振动能对应于被称为构成干酪根的官能团的结构部分。这些主要由单键结构和双键结构组成。尤其是,与芳香族组分以及羰基和羧基相关的C=C和C=O双键的变化以及与亚甲基(CH₂)和甲基(CH₃)相关的单键C-H的变化可诊断成熟度和干酪根中的组成。图5是示出相对于与烃源岩的化学官能团相对应的各种波数的IR吸收光谱。例如,波数2930cm⁻¹和波数2860cm⁻¹主要与CH₂和CH₃的C-H键的不对称拉伸和对称拉伸有关。可以使用峰强度或面积来确定相对于与波数1710cm⁻¹和波数1630cm⁻¹相对应的C=O和C=C键的组成。如果在1710cm⁻¹和1450cm⁻¹处的C=O大于C-H键的CH₂-CH₃,则表明地面有机物输入,而相反的关系表明海洋输入更多。这种关系表明,这些结构之间的连续体可用于确定所含有机物是I型、II型或III型干酪根。

[0077] 由于热生成过程,随着成熟度的增加,在1630cm⁻¹处的C=C键也稳定地移动。这表示芳香族环的融合增加。可以从%Ro和波数从1630cm⁻¹到1570cm⁻¹的变化之间的关系看到这一点,如图6所示,其中1630cm⁻¹波数随着烃源岩中的干酪根成熟度的增加而变化。此外,使用振动能2930cm⁻¹和2860cm⁻¹和1630cm⁻¹的关系的组合可以用来绘制Van Krevelen图。除了有机物的组成和分类外,FTIR还可用于鉴定和定量矿物组成。利用识别烃源岩矿物学的能力,然后将这一信息与FTIR衍生的干酪根组成信息一起用于定义有机相分布。有机相反反映了与沉积环境、煤岩组分类型和地壳构造地域有关的烃源岩的复杂性。

[0078] 太赫兹光谱法是另一种用于分析烃源岩的非破坏性分析方法。太赫兹辐射是位于

0.3THz至10THz (1THz=10¹²Hz) 的频率范围内的电磁波,其位于微波和近红外之间的频谱中。太赫兹吸收主要受远距离振动激发,可以是长链分子中的分子内扭转模式或相邻分子之间的分子间振动模式。这些低频振动模式是高度集中的,并且可能会大量混合。结果,将各个吸收带或峰与特定模式以及分子结构和组成相关联通常是困难的任务。结果,在烃源岩表征的背景下对太赫兹光谱数据的解释在文献中很少。在某些实施例中,分析烃源岩样本的方法包括将针对具有相同成熟度和有机相类型的样本的THz数据与FTIR数据并置,从而从已得到更好表征的FTIR数据与分子属性的关联中推断出THz数据与分子属性的关联。产生的太赫兹关联存储在数据库中,并用于使用其自己的THz数据预测新样本的属性。

[0079] 电子自旋共振范围内的ESR光谱提供了特征的测量,例如颜色强度、自旋数和每次自旋的芳香族碳原子数。石油的树脂和沥青质馏分的强烈深色通常被认为表示与芳香族结构有关的有机自由基。

[0080] 荧光显微镜(FM)是一种用于对有机材料和无机材料进行识别、分类和定量的光学工具。荧光是一种发光属性,其中光子用于激发矿物、有机物和地质材料的原子结构内的敏感电子。这些激发的电子暂时跳到矿物的原子结构内的更高轨道。当这些电子落入其原始轨道时,能量以光的形式释放。发出的荧光比激发光(斯托克位移(Stoke's shift))具有更长的波长和更低的能量。与烃链长度有关的具有不同热成熟度的烃源岩具有特征性的斯托克位移。根据爱因斯坦-普朗克方程,降低吸收能会导致更长的发射波长、伴随着有机物的热成熟而众所周知的荧光红移。较高成熟度时荧光发射强度的降低是由于分子芳构化作用增强所引起的浓度猝灭效应(非辐射弛豫),这降低了荧光团的密度。烃源岩中存在的某些有机物是一种生物聚合物,其主要包含长链饱和脂族烃。荧光图像与光谱数据相结合的诊断潜力为快速烃源岩和有机相分析提供了机会。

[0081] 收集气相色谱法(GC)和气相色谱质谱法(GC-MS)数据以分析与从烃源岩中提取的石油或沥青有关的烃分子化学组成。在某些实施方案中,烃源岩样本经历有机地球化学分析的五个主要步骤。第一步,对烃源岩进行Rock-Eval®热解,然后进行第二步,即几种溶剂提取方法(例如,索氏提取、超声提取、液-液提取)中的一种,以提取沥青。在第三步中,对沥青进行高效液相色谱分析,以分析饱和物、芳烃、NSO化合物和沥青质。在第四步中,使用气相色谱法进一步分析饱和物和芳烃,以了解正构烷烃(C₂₂-C₃₅)和芳烃(C₂₂-C₃₅)的分布。通过GC-MS对这些化合物进行进一步分析,以获得生物标志物谱(biomarker profile)。这些数据可用于如下的大量评估,例如,用于确定沉积环境的降植烷分子和植烷分子的评估,使用异庚烷不变性以确定产出流体是否源自盆地内的相同烃源岩的油与油的相关性,以及用于确定有机物类型的偶数-奇数烷烃碳数比。API重力(美国石油协会重力)是使用油的比重计算的,该API重力是其密度与水的密度的比值(油的密度/水的密度)。API计算的比重始终以60华氏度(oF)来确定。GC-MS数据可用于检测由于油的蒸发分馏或生物降解而引起的API重力的变化。GC-MS数据还可用于检测和评估油的属性和生物降解程度。这里公开的方法利用产出流体的分子化学组成来开发一种预测烃源岩储层和常规储层跨盆地的API重力的方法。气相色谱/质谱数据也可用于分析油类和从饱和馏分和芳香族馏分提取的用于生物标志物的沥青。GC-MS数据可用于检测某些生物标记物的异构化和芳构化,以估算烃源岩中流体的成熟度,并确定烃类是原位生成还是从其他位置迁移。GC-MS数据还可用于使用沉积生物标记物建立有机相。GC-MS数据可用于金刚烷分析以确定烃源岩中油气裂解的百分比及

其与干酪根内部孔隙度发展的关系,并用于确定油是来自相同还是不同的烃源岩和/或储层。与从沥青中提取的生物标志物相比,GC-MS数据可用于产出流体中生物标志物的时间序列地质化学,以确定在受激岩石体积内生产起源于何处。一旦建立了来自代表性样本的多次测量结果的数据库,并且获得了新的目标样本的测量结果,就可以采用一种或多种用于预测成熟度和有机相分布的分析方法。一种方法包括数据预处理、特征提取、用于实验室设置和现场设置的最佳成熟度和有机相属性的确定、烃源岩样本分类、以及成熟度/有机相分布的预测和解释。

[0082] 本公开还提供一种确定样本烃源岩的成熟度的系统。该系统包括放置在样本烃源岩附近并通信地耦接到一个或多个处理器的多个数据获取装置。该系统还包括非暂时性计算机可读介质,其被定位成与一个或多个处理器通信,并且其上存储有计算机程序,该计算机程序包括一组指令,当该一组指令由一个或多个处理器执行时,该指令使得所述一个或多个处理器执行以下操作:建立与放置在样本烃源岩附近的多个数据获取装置的通信链路、从所述多个数据获取装置获得样本烃源岩的第一多个数据、以及使用预测相关性分析第一多个数据以确定样本烃源岩成熟度。通过数据分析引擎,通过应用机器学习模型将从多个代表性烃源岩获得的第二多个数据与多个代表性烃源岩的多个属性相关联来生成预测相关性。该系统还包括烃源岩数据库,该烃源岩数据库至少包含与多个代表性烃源岩相关联的第二多个数据、多个代表性烃源岩的多个属性以及预测相关性。数据获取装置可被定位成从样本烃源岩的最佳感测波段获得数据。数据获取装置被定位成获得位置数据、光谱测量结果和光学测量结果中的两个或更多个。该系统可包括样本烃源岩检索设备,以获得样本烃源岩的一部分。光谱测量结果可包括从傅立叶变换红外光谱、电子自旋共振光谱、太赫兹光谱和紫外光谱获得的测量结果中的一个或多个。光学测量结果可包括通过荧光显微镜和共聚焦激光扫描显微镜获得的测量结果中的一个或多个。例如,在某些实施例中,当仪器具有200nm的分辨率时,来自共聚焦激光扫描显微镜分析的图像具有400纳米(nm)的像素大小。

[0083] 另外,这里公开了专用采样装置。这些装置包括采样入口和样本获得室。采样入口被设计成从钻孔环境中提取样本。样本可包括流体、固体或者它们的组合。在某些实施例中,样本是烃源岩样本,其不是钻井液流的组分。样本获得室可被配备成存储样本,而不会损失其完整性,直到进一步处理为止。样本获得室可被配备成使样本与原位数据获取装置接触以进行进一步处理。采样入口可被设计成包含一个或多个开口,以允许样本检索装置、样本或两者通过。样本检索装置包括样本提取组件和样本收集组件。样本提取组件可包括机械操纵器(例如钻头)、或者使用流体(例如水或空气)或波(例如超声波)提取样本的装置。可以使用机械操纵器或基于流体的钻井工具提取烃源岩样本。样本收集组件可被集成到样本提取组件中,或者也可以单独存在,例如,在样本与环境分离后收集样本的容器或抽吸装置。

[0084] 数据获取装置的示例是井下光谱仪,并且包括至少一个光源、检测器、用于反射来自感兴趣材料的源(无论是流体还是岩石)的光并在该检测器处返回的组件、将源和检测器部署到井中的组件、以及检索在光谱测量中获得的数据的组件。在一个实施例中,光源是一个或多个单芯片激光器(例如铋化镓分布反馈激光器或量子级联激光器),每个激光器被调谐以产生不同的波长并且各自被供电,使得一次只有一个这样的激光器被激活。在该

实施例中,检测器是由例如钛酸镧(LaTiO_3)制成的热电检测器。热电检测器不会通过吸收产生载流子。相反,由于入射光子的加热,它们通过机电膨胀产生电荷。因此,它们对很宽的频率范围敏感,并且一个这样的热电检测器可用于测量来自任何单芯片激光器的光强度。在该实施例中,一次激活一个激光器,并且当每个激光器激活时,由热电检测器检测到的光强度被测量并记录在计算机存储器中,从而提供了在每个激光器波长下光衰减的量度。

[0085] 在一个实施例中,光源可以是铁丝,该铁丝在加热时辐射红外波段中的广谱光。在该实施例中,检测器是多个热电检测器,其中每个检测器在其顶部具有一个永久性滤波器,因此它仅对滤波器选择的波长的光敏感。当光源处于活动状态时,将记录每个热电检测器上的光强度,以提供每个滤波器选择的波长处的光衰减量度。在一个实施例中,有多个热电检测器,其中滤波器被调谐到每个波长。每个这样的检测器都比光束的宽度小得多,并且所有检测器以棋盘图案放置在光源的光束内。将具有相同波长的滤波器的所有检测器的光强度相加,以使光束强度的变化在许多检测器位置上平均。在一个实施例中,将光反射离开材料的设备包括将岩石或流体样本引入其中的腔室。源和检测器在腔室外部,并且光从光源穿过窗进入腔室,然后通过同一窗返回到腔室的外部,在该外部处由检测器对其进行测量。在另一个实施例中,源和检测器被包含在传感器封装中,并通过窗与井眼流体或岩壁分开。在一个实施例中,窗在岩石样本或流体侧上涂覆有不粘涂层,例如聚对二甲苯,该不粘涂层排斥井眼流体,使得所述流体可通过窗看到但不粘附到窗上。在一个实施例中,窗被设置在机械爪或刮刀附近,该机械爪或刮刀从岩壁去除钻井泥浆,使得光可从源传播到岩壁,并且可通过检测器观察到来自岩壁的反射光。在另一个实施例中,将透明流体的射流喷射在岩壁上,以在窗和岩壁之间保持透明的流体层。

[0086] 在数据获取装置的某些实施例中,光源和检测器被部署到井中,并且包含在电缆工具、钻柱、油滑工具、钻井接头或井底钻具组件内,或者他们的组合中。在某些实施例中,原始光谱数据(a)被存储在工具中的计算机存储器中,以便在从井中回收工具之后进行下载,或者(b)通过电缆传输到地面,或者(c)使用无线遥测(包括声学遥测,电磁遥测或泥浆脉冲遥测)设备遥测到地面。

[0087] 储层分析的不同方面可被并入数据获取装置中。某些实施例具有可调光学元件,其可直接测量矿物和沥青组分原位组成,并提供有关页岩和碳酸盐中烃类储量的信息。

[0088] 某些实施例包括用于获得井下红外光谱测量结果的设备和方法。红外光谱学可提供有关岩石中有机物和矿物含量的信息,以及区分和表征产出流体的信息。但是,由于需要的仪器体积庞大且标准光谱仪器的设计要求不能在井下温度下操作的检测器,因此目前在实验室中对回收的样本进行此类测量。此处描述的设备使用小型坚固的组件来进行光谱测量,这些组件可以在井下温度下操作。在一个实施例中,装置包括广谱红外源和热电检测器,该热电检测器具有在衰减反射光谱法方案中特定固定集成光学滤波器。在另一个实施例中,装置包括使用选定频率的激光芯片和热电或石英增强的光声光谱检测器。借助如此众多的小型单波长衰减测量,即使在井下温度下,也可以在井中的所有深度快速获得关键波长的光谱信息,而费用仅为实验室测量的一小部分。这些红外吸收日志有多种应用。在关键波长处井眼中岩壁的红外吸收与矿物学和成熟度有关,以指导页岩气井的钻井和水力压裂。井的垂直部分中的这种日志将指示分支的最佳深度,而分支中的此类日志将指示水力压裂的最佳位置。这样的日志还将使得随着新井的钻探,对油气系统的盆地模型和地质认

识得以完善和更新。生产井内流体的红外吸收日志提供了有关流体组分的信息,该信息可以指示流体的来源,从而指示储层的分隔。

[0089] 示例1

[0090] 烃源岩数据库包含烃源岩样本的各种特征的测量。例如,无限制地,烃源岩数据库包含与各种烃源岩样本相关联的参数(例如,位置和深度)以及其Rock-Eval®参数(例如S1、S2、Tmax、HI、氧指数(OI)、产量指数(PI)、TOC、%Ro-Tmax和HI-%Ro)。烃源岩数据库还可包含FTIR数据(例如原始数据和预处理数据、FTIR成像数据和提取样本的FTIR数据)、以及元素组成数据(包括烃源岩样本的单个矿物图和矿物分布)。烃源岩数据库还包含光谱测量结果和图像、共聚焦荧光图像、X射线荧光图像、ESR测量结果、太赫兹图像和其他数据。下面提供了用于获得这些度量的各种方法以及数据库中包含的数据示例。

[0091] 在波数波段 $500-4000\text{cm}^{-1}$ 上获得了所有代表性烃源岩样本的FTIR光谱。测量是从与表2相同的粉末形式的烃源岩样本中进行的,这些样本具有不同的粒度并且包括提取的干酪根以及粘土部分。使用粒度范围在45微米至250微米之间的高岭石、伊利石、蒙脱石和钠型蒙脱石(NA-MONT)粘土样本,例如蒙脱石-45、蒙脱石-63、蒙脱石-75、高岭石-106、高岭石-150、伊利石-150、伊利石-150、蒙脱石-150、和250微米。通过选择性蚀刻,提取的干酪根约为100%(体积),粘土也为100%(微量)。以粉末形式分析,粘土部分与提取的干酪根相比似乎对烃源岩成熟度的依赖性较小。

[0092] 图7A是根据各自的成熟度从上到下增加的四组烃源岩的FTIR光谱。图7B是图7A中分析的四组大量烃源岩的FTIR光谱的放大图,以便更好地可视化与有机物含量相关联的波段。 2930cm^{-1} 和 2860cm^{-1} 的波数主要与 CH_2 和 CH_3 的C-H键的不对称拉伸和对称拉伸有关。在图7B中,显然, 2930cm^{-1} 和 2860cm^{-1} 处的峰值从样本3、样本1A、样本1B到样本2减小,而成熟度增加(表2)。还可以使用峰强度或峰面积来确定相对于与 1710cm^{-1} 和 1630cm^{-1} 的波数相对应的C=O和C=C键的组成(图7B)。图5以透射率(T)示出了FTIR光谱。在本公开中提供的大多数其他FTIR光谱,包括图7A、图7B和图7C及其之外的光谱,都使用吸光度(A)表示。注意, $A = \log_{10}(1/T)$ 解释了与图5相比的峰反转。图7C是图7A中分析的四组大量烃源岩的FTIR光谱的放大图,以更好地可视化与粘土和矿物含量相关联的波段。与图7B中提到的波段不同,粘土波段和矿物波段的变化并不具有已知的对成熟度的直接依赖关系,至少从FTIR光谱中看不明显。本文公开的方法的目的之一是通过使同一组样本上不同波段之间的这些变化相关联来推断和捕获这些关系。 2930cm^{-1} 和 2860cm^{-1} 的波数主要与 CH_2 和 CH_3 的C-H键的不对称拉伸和对称拉伸有关。在图7B中,显然, 2930cm^{-1} 和 2860cm^{-1} 处的峰值从大量样本3、大量样本1A和大量样本1B到大量样本2自上而下减小,而成熟度增加。还可以使用峰强度或峰面积来确定相对于与 1710cm^{-1} 和 1630cm^{-1} 的波数相对应的C=O和C=C键的组成。

[0093] 图8A是多个烃源岩粉末样本以及提取的干酪根的FTIR光谱,而图7A和图7B是来自大量样本的ATR FTIR测量结果。如预期,从样本1B、样本2A、样本2B、样本3A、样本4A和样本4B提取的干酪根的FTIR光谱在粘土和矿物带中的响应明显比在粘土粉末中的响应低。另外,在C-O和C-C波段附近的峰更明显。图7A、图7B和图7C所示的光谱是从大量烃源岩样本中获得的衰减的全反射率(ATR)FTIR测量结果,而图8A和图8B中的光谱是来自烃源岩粉末样本和基准粘土样本的测量结果。

[0094] 图8B是纯粘土粉末样本的FTIR光谱的曲线图,其中粘土样本在C-H波段处显示出

很小的响应。相反, 3600cm^{-1} 附近的特征信号是明显的, 尤其是对于高岭土, 像峰一样的兔子耳朵显得突出。所有粘土样本在与矿物响应相关联的 1000cm^{-1} 附近也显示出强烈的峰值。使用粒度范围在45微米至250微米之间的高岭石、伊利石、蒙脱石和钠型蒙脱石粘土样本, 例如蒙脱石-45、蒙脱石-63、蒙脱石-75、高岭石-106、高岭石-150、伊利石-150、伊利石-150、蒙脱石-150、以及250微米。

[0095] 使用THz时域光谱 (THz-TDS) 方法获得烃源岩数据库中的THz光谱数据。图9A、图9B、图10A和图10B是烃源岩样本的THz时域光谱数据。图9A示出了THz基准 (HDPE) 的测量时间迹线以及根据各种烃源岩样本的分析的脉冲时间波形, 并且图9B是通过波形的傅立叶变换获得的对应光谱。基于此, 计算了各种烃源岩样本的样本吸收和透射光谱, 并分别将其示为图10A中的烃源岩样本的THz吸收光谱和图10B中的烃源岩样本的透射光谱。如图10A和图10B所示, THz光频波段与有机分子结构的关联不是那么直接并且定义得不太好。相信大多数特征可能与大链中的分子内扭转模式或相邻分子之间的分子间振动模式有关。通过这里公开的协作学习方法, 对于相同的成熟度和有机相分布的样本, THz数据与其他测量结果 (例如FTIR光谱) 并存。产生的太赫兹关联是经过训练的模型, 其可用于通过其代表性太赫兹数据来识别新岩石样本的特征结构, 例如有机组分。

[0096] ESR光谱法用于分析四种不同成熟度水平的烃源岩。图11A是示出透射中的层理 (红色区域富含有机物) 的2D岩心截面图。色标表示自由基的水平, 因为它们在整个样本中都不同, 红色为更高的水平。图11B是四种不同成熟度水平 (如表2所述的样本4A、样本2A、样本1A和样本3A) 的烃源岩的自由基、铁和烃类指数与镜质组反射率的关系的图形表示。自由基和黄铁矿的响应是通过大量测量获得的。黄铁矿/自由基浓度与成熟度呈抛物线关系。在ESR图像中, 烃源岩样本为样本3A, 示出了有机富集度 (橙色), 并且图11B中的曲线图以自由基 (蓝色为RD曲线)、黄铁矿 (红色为Fe曲线) 和氢指数 (绿色为HI/100) 的形式示出了空间平均组分差异。

[0097] 图12A至图12C示出了不同成熟度的富含有机物的烃源岩样本的相对荧光响应。有机物的荧光强度随着成熟度而降低。共聚焦图像是在清洁、切割和离子铣削的样本上拍摄的。样本光滑。这些图像是用激光扫描共聚焦显微镜获得的。用一系列三个激光 (405nm、488nm和633nm) 激发样本。对于405nm激发, 发射范围是400nm至513nm。对于488nm激光器, 发射范围是490nm至633nm。最后, 对于633nm激光器, 发射范围为638nm至747nm。图像对应于发出的光的荧光强度。颜色表示通过一系列滤波器在大约490nm、505nm和560nm处检测到的发射光。如图12A所示, 样本3A是氢指数 (HI) 为485的未成熟样本。相对于较成熟的样本, 样本3A具有很强的荧光响应。如图12B所示, 样本1A在峰值油窗内, 其中HI为185。如图12C所示, 样本4B正好处于气窗内, HI为~8。存在与荧光强度的降低相对应的成熟度的增加 (氢的损失, 因此HI降低)。尽管这些示例显示出质量下降, 但也可以定量捕获此响应。

[0098] 使用来自代表性样本的多次测量结果和感兴趣的新样本的测量结果的烃源岩数据库, 成熟度和有机相预测的分析过程包括几个步骤, 包括预处理、特征提取、用于实验室设置和现场设置的最佳成熟度和有机相属性、烃源岩分类、以及成熟度/有机相的预测和解释。在某些实施例中, 可以在进一步呈现用于分析之前, 对样本进行检索和处理。例如, 样本可以显示为薄片。样本也可以呈现为大碎片, 该大碎片可以通过机械抛光或离子铣削来固定并弄平。样本也可以呈现为核的小碎片, 该小碎片可以通过机械抛光或离子铣削来固定

和平滑。某些分析方法可能需要将样本清洗作为样本制备的一部分,因为外部化合物(例如钻井泥浆)的存在会影响分析。

[0099] 图13是根据实施例的预测样本烃源岩的成熟度和有机相分布的方法1300的框图。该方法开始于通过先前描述的方法处理从烃源岩样本中获得的数据。数据处理步骤1301包括用于每种测量类型的数据清理和准备步骤,例如,离群值检测、基线校正、峰增强和归一化。离群值检测可以通过许多技术(例如,主组分分析或基于接近度的方法)来实现。来自烃源岩样本3A的FTIR数据的示例受这些处理步骤的影响,如图14A至图14D所示。图14A是来自烃源岩样本3A的原始FTIR数据。图14B示出了经受小波去噪和基线校正的FTIR数据。图14C示出了五阶多项式基线校正之后的FTIR数据。图14D示出了经受向量归一化的图14C的FTIR数据。

[0100] 对各个测量结果进行相应预处理的多个测量结果可以在频段的波数上进行积分,从而构成一个高维标称输入空间。这样的宽带输入光谱的示例在图15中示出,其示出了跨越THz波段、IR波段和UV波段(示出的每个波段中的幅度被居中并且归一化)的一个烃源岩样本的全光谱。

[0101] 参照图13,处理后的数据经历特征提取步骤1302,其包括稀疏波段选择和降维。可以使用来自数据库的带注释的样本对光谱进行聚类或分类,然后可以应用特征分级算法以在缩小的特征空间中获得信息量最大的光谱图案。详细地,将光谱测量结果映射到一些缩小的特征空间中,该缩小的特征空间捕获显著的光谱模式。然后,将目标属性(例如成熟度和有机相分布)设置为感测或预测烃源岩样本的属性。从策划的烃源岩数据库中选择数据库样本的子集作为输入。最后,特征选择算法用于根据选定的数据库样本来计算波数/频段的权重。然后,通过三个单独的步骤处理来自步骤1302的数据。开发感测波段以预测成熟度和有机相分布的步骤1303包括通过特征分级的波段选择和优化。该步骤被概括为在用于预测或分类的特定目标下代表图3中的步骤306、步骤308和步骤310的数学表达式。出于成熟度和有机相预测的目的,值西塔(theta) θ 代表光谱波段的最小子集,可为具有已知成熟度或有机相的数据库样本选择光谱波段的最小子集,它们将成功预测的可能性最大化。最小集是根据向量 p_norm 小于 r 的约束表示的。成功预测的可能性由 $\log p(\text{成熟度,有机相}|\theta)$ 表示。选定的波段对选定的目标属性(例如,特定的成熟度或有机相分布)最有帮助,或者最有助于区分目标烃源岩样中的差异。选定的波段还将受到传感器设计和部署的可行性约束。在步骤1304中确定最佳成熟度或有机相属性。在步骤1305中,烃源岩分类包括利用几种分类算法,例如无监督聚类、有监督潜在狄利克雷分配(supervised Latent Dirichlet Allocation, sLDA) 或支持向量机建模。在步骤1306中,通过建模方法(例如,支持向量回归)将源于步骤1303、步骤1304和步骤1305的数据和约束条件用于预测烃源岩样本的成熟度或有机相分布。

[0102] 图16A和图16B描绘了降维空间中烃源岩样本聚类的两种表示。在基于聚类的成熟度预测模型中,光谱被映射到缩小的特征空间中,该特征空间捕获显著的光谱模式。样本被分为几组,其中烃源岩样本在每个组中彼此相似,但在组之间不同。推导了成熟度特定的特征和簇结构,以确定相关的波数波段和代表性色谱图。然后从一个或多个簇中预测新样本的成熟度,其中现有光谱与新样本的光谱最相似。该过程可以在波段的可访问子集上完成,或者融合在众多测量结果中。作为示例,通过K-均值聚类算法获得图16A和图16B。在图17以

及图19A和图19B中提供了使用分层聚类的结果。

[0103] 在基于分层双聚类的成熟度预测模型中,根据光谱相似性对样本进行聚类。光谱波段基于它们在整个样本数据集中的振幅分布进行聚类。热图与有助于聚类的光谱图(在已排序的坐标中)相对应,可以概括为包括各种类型的测量结果,例如荧光、太赫兹和ESR的测量结果。图17是分层的双簇图。主热图的每一行对应一个样本,并且每列对应光谱中的特定波数。色标是根据预处理后的光谱绘制的。以预处理的宽带光谱为输入,该算法根据选择的特定距离度量对样本和光谱频率进行分组和排序。顶部的树形结构对应于不同光谱频率或波数的聚类。左侧的树形结构表示不同烃源岩样本的聚类。类似的样本被分组在树的相邻叶子下,具有相同的光谱频率。对于图17中所示的特定示例,光谱频率通过算法进行分组,以使其形成与样本烃源岩中各种矿物和有机组分以及有机物的分子结构相关联的波段,如图17中用颜色编码的框中所示。通过将新样本包含到数据库中,这些结果可以将新样本与具有已知成熟度和有机相的现有样本以及有助于聚类结构的特定子带一起聚类。该信息可用于确定更多的颗粒信息,例如与特定成熟度或有机相相关联的矿物含量和分子结构。

[0104] 这些烃源岩预测方法有助于非常规储层的地球化学表征,并提供预测工具,这些预测工具将允许野外作业人员更准确地评估储层质量和开采潜力,并评估是否可以提高给定油田的产量。对从井中获得的样本和岩心样本进行的分析还可为勘探和开发提供关键信息,以更好地评估储层潜力。

[0105] 示例2

[0106] 以下示例是用于选择和优化感测波段的方法。输入通常来自策划的原始岩石数据库。利用特定的目标属性(例如特定的烃源岩成熟度范围或特定的有机相分布)来选择和优化感测波段。选择波段以为所选目标属性提供最多的信息,或者最大程度地区分目标烃源岩样本中的差异。波段选择/优化可以通过特征排名来实现,并且还可能受到传感器设计和部署的可行性约束。

[0107] 在感测波段最优化的该示例中,选择波段以在不同成熟度水平下将粘土与干酪根区分开。应用特征选择算法来计算每个频率/波数点的权重。图18表示来自FTIR光谱测量的分级特征的权重与光谱波数波段的对准,以区分不同成熟度水平下的各种粘土、矿物和干酪根。图18的顶部面板示出了样本的FTIR波数光谱的所得权重。图18的底部面板示出了各种样本的FTIR光谱。结果包括每个频率/波数点的一组权重;在图18的顶部面板中,Y轴上的权重越高,所选的波段对目标分布越敏感。权重最高的波段与表示烃、粘土和矿物的已知波段一致,如图18的底部面板所示。在作者为Jundong Li, Kewei Cheng, Suhang Wang, Fred Morstatter, Trevino Robert, Jiliang Tang, and Huan Liu于2016年公开的“特征选择:数据透视图 (Feature Selection: A Data Perspective)”(其从网址<https://arxiv.org/pdf/1601.07996v4>.PDF可获得)中描述了几种用于通过特征排名算法实现目标聚类、分类或预测目标的方法。从具有最高灵敏度的感测波段中选择最佳感测波段,并且对于传感器设计和部署也是可行的。例如,对于粘土的感测波段被确定为高于 3500cm^{-1} ,而对于矿物的感测波段为大约 450cm^{-1} 至 1050cm^{-1} 。

[0108] 示例3

[0109] 以下示例是用于通过分层聚类对烃源岩样本进行聚类的方法。样本被分为几组,

其中烃源岩样本在每个组中彼此相似,但在组之间不同。推导了成熟度特定的特征和簇结构,以确定相关的波数波段和代表性谱图。图19A和图19B是来自投影在选定波数轴上的不同样本的FTIR光谱的簇图的表示。聚类是在由波数波段定义的高维空间中完成的。图19A示出了使用光谱波数 1490.9cm^{-1} (x轴)和 2806.6cm^{-1} (y轴)的聚类结果。图19B示出了在两个不同的光谱波数位置方面(对于x轴是 3697cm^{-1} ,并且对于y轴是 2862.5cm^{-1})相同的聚类结果。当在这两个不同的坐标系中投影时,由于来自这些波数波段的不同影响,相同的聚类结果将具有不同的外观。这里的样本包括已知的粘土材料(伊利石(C1)、高岭石(C2)、蒙脱石(C3))以及按升序以不同成熟度提取的干酪根(样本3A(C7)、样本1A(C4)、样本2A(C5和C6)和样本4A(C8))。图19C是从18种不同类型的样本的样本FTIR光谱的分层聚类获得的示例树状图。

[0110] 然后,将这些样本中的每个样本作为测试样本来运行,以确定模型的预测能力。使用不同的学习算法(例如,支持向量机(SVM)回归、Random Forest®、最近邻(Nearest Neighbor)分析和Adaptive Boosting®(也称为AdaBoost®))进行分类,并使用四种误差度量法评估了它们的性能,如下所示在表3中。

[0111] 表3

[0112]	方法	均方差 (MSE)	均方根差 (RMSE)	平均绝对值 差 (MAE)	确定系数 (R2)
	最近邻	0.139	0.373	0.200	0.791
	AdaBoost®	0.049	0.220	0.088	0.927
	Random Forest® 回归	0.064	0.252	0.128	0.904
	SVM 回归	0.708	0.841	0.738	-0.067

[0113] 选择两个性能最高的预测因子Adaptive Boosting®和Random Forest®进行成熟度预测,结果如图20A和20B所示。将Adaptive Boosting®方法(如图20A所示)和Random Forest®方法(如图20B所示)的预测结果相对于粘土材料(伊利石(C1)、高岭石(C2)、蒙脱石(C3))以及按升序以不同成熟度提取的干酪根(样本3A(C7)、样本1A(C4)、样本2A(C5和C6)和样本4A(C8))的已知值来绘制。

[0114] 如先前所讨论的,烃源岩和干酪根的成熟度水平由诸如镜质组反射率($\text{Ro}\%$)、氢指数(HI)、热解 T_{max} 和原子H/C比率的几个量表示。在这样的示例中,选择氢指数(HI)与镜质组反射率($\text{Ro}\%$)的比率作为成熟度的预测指标。Adaptive Boosting®和Random Forest®回归模型用于预测各种样本的HI- $\text{Ro}\%$ 指数。图21显示了与通过传统方法处理各种烃源岩获得的成熟度指数相比通过两种不同的学习算法(自适应Boosting®方法和Random Forest®方法)从烃源岩样本的FTIR光谱预测的成熟度指数(HI- $\text{Ro}\%$ 比)。如图21所示,这些模型提供了极高的样本成熟度的可预测性。

[0115] 尽管已经使用FTIR数据描述了这些示例,但是这些方法可以应用于从THz光谱、ESR光谱和荧光光谱以及烃源岩的其他光谱和光学测量中获得的其他数据。

[0116] 示例4

[0117] 在图22A和图22B中描绘了数据获取装置、热电检测器的组件的示例。热电检测器

对高背景温度不敏感；而且它们仅对温度变化做出响应，使其适合井下应用。图22A表示简化的热模型，而图22B表示等效电路，其中， α 是吸收系数， H_p 是热容， G_T 是热导率， T_A 是环境温度。随着光量由于反射/透射到样本中而发生变化，产生的热量将发生变化，从而使热电材料（如钽酸锂）中的电荷变化很小。如图22A所示，辐射通量 Φ_s 被吸收并引起热电元件中的温度 ΔT_p 的变化。热电转换是由于热电效应引起的，因此温度变化 ΔT_p 改变了电极上的电荷密度 ΔQ_p 。如图22B所示，随后经常进行电转换，其中信号被放大并转换成电压。等式1给出了图22B所示的小电路的温度差。

$$[0118] \quad \Delta T_p = \frac{\alpha \tau_F \Phi_s}{\sqrt{G_T^2 + \omega H_p^2}} \quad \text{等式 1}$$

[0119] 当 $\alpha \tau_F$ 最大化而 G_T 和 H_p 最小化时， ΔT_p 将最大化。实际上，将需要折中。在某些应用中，迹线检测需要大的信噪比，而在其他应用中，需要高光谱分辨率，例如用于粘土识别。根据应用的不同，报告的信噪比为10000至100，其中准直几何结构中的可调滤波器可实现高信噪比，并且聚焦光学器件可实现高光谱分辨率。

[0120] 图23示出了包含与可调滤波器集成的商用热电检测器的设备2300。可以将滤波器2301调整为感兴趣的特定波长：高岭石为2703nm，脂族为3496nm，蒙脱石为2923nm，等等。在一个实施例中，滤波器2301可以是宽带滤波器。该设备还包括可移动反射器2302和固定反射器2303。该设备包括热电检测器2304，该热电检测器2304由几个热电元件的阵列组成，每个热电元件具有针对特定波长调整的固定滤波器，因此不需要用于调谐的相对脆弱的MEMS机械结构。这种脆弱的机械结构将使传感器不适合部署在井下工具中，尤其是在工具内部经过压力补偿并充满绝缘液体的情况下。热电检测器2304耦接到电路2305。该设备还包括控制电极2306和弹簧悬架2307。

[0121] 图24和图25描绘了样本测量的两种配置。在该实施例中，选择衰减全反射率（ATR）方法。在图24中描绘了单个反射ATR单元2400，其中来自光谱仪的IR光束2401在ATR晶体2402中被内部反射并且在晶体界面处与样本2403相互作用。分析反射的IR光束2404以确定消逝波与样本的相互作用并获得吸收光谱。如图25所示，在ATR单元2500中发生了具有多次反射的波导技术。入射光束2501在由特殊棱镜构成的ATR元件2502内部经历多次内部反射。宽阔的样本界面可提供与样本2503的更大接触，并且可用于弱吸收剂和稀样本。出射的红外光束2504被引导至红外光谱仪中的检测器。检测器记录IR光束并生成红外光谱。图8A和图8B是几种烃源岩样本和纯粘土的红外光谱的示例。高岭石、伊利石和钠蒙脱石（蒙脱石）有独特的峰。通过查看 3800cm^{-1} 至 3000cm^{-1} 范围内的每个特定粘土波数，可以区分这些粘土。高岭石在 3400cm^{-1} 附近有两个明显的峰，而伊利石在 3450cm^{-1} 处有一个尖峰，然后在 3200cm^{-1} 处有一个宽峰。

[0122] 示例5

[0123] 原位气体（GIP）是在特定的大量储层岩石体积内存储的气体体积。随着钻屑被输送到地面，一些气体逸出到钻井泥浆中。这种气体损失量会导致GIP估计错误。图26是使用热电传感器的GIP数据获取装置的示意图。设备2600具有气体入口2601和气体出口2602，用于管理气体样本通过气体样本室2603的进入和离开。光源2604放置在气体样本室2603的一端，并且热电检测器2605放置在气体样本室2603的另一端。选定的传感器在高温下表现良好，并且该设备的占地面积已减小以装入随钻测井（LWD）工具或生产测井工具（PLT）中。与

当前的GCMS装置相比,该识别的成本也显著降低,并且某些GCMS装置较大且不适合高温使用。可以对从钻井液或钻屑中释放出的气体进行原位测量,以估算适当的储量并确定源层的成熟度。图27A、图27B和图27C分别是GIP数据获取装置2600、光源2604和热电检测器2605的实验室原型的照片。图28A和图28B分别是使用乙烷获得的IR光谱和相关联的系统校准光谱。初步结果表明,该系统的灵敏度为百万分之三(3毫克/升),并且还可以检测 C_1-C_4 及其比率。

[0124] 在这些示例中描述的装置提供了与原位和在接近储层条件下的井下岩石和流体有关的IR光谱信息。该信息包括矿物学、粘土中的游离水对约束水、成熟度。它可以检测沥青和焦油沥青。它可以区分芳香烃和脂肪烃。它可以检测井筒流体的组分。由于光电检测器中使用的半导体材料的高温性能差,因此很难进行原位测量。尺寸要求不允许使用大型实验室规模的装置进行原位测量。本公开通过使用热电传感器解决了温度问题,并且通过使用集成的可调滤波器使光谱仪小型化。该封装可以安装在电缆工具中,并且价格低廉,因此可以部署在永久性感测应用中。

[0125] 示例6

[0126] 这里描述了用于促进对LWD应用以及测井中适用的钻井液/固体进行采样的装置。图29是采样装置2900的示意图。该装置2900包括采样入口2902和样本获得室2904。采样入口2902被设计为允许样本取回装置和来自该装置的环境的样本通过。采样入口2902可被设计成容纳过滤器以获得某些尺寸的样本或流体。样本获得室2904可被装备成在不损失完整性的情况下存储样本2906,直到进一步处理为止。样本获得室2904可被装备成使样本与原位数据采集装置接触以进行进一步处理。

[0127] 鉴于此描述,本文公开的设备和方法的各个方面的进一步修改和替代实施例对于本领域技术人员将是显而易见的。因此,该描述仅应被解释为说明性的,并且是为了教导本领域技术人员实施实施例的一般方式。应当理解,这里示出和描述的实施例的形式将被视为实施例的示例。元件和材料可以代替这里图示和描述的那些,部件和过程可以颠倒或省略,并且实施例的某些特征可以独立地利用,在受益于实施例的描述之后,所有这些对于本领域技术人员将是显而易见的。在不脱离如所附权利要求书中所描述的实施例的精神和范围的情况下,可以对这里描述的元件进行改变。

[0128] 前述方法、设备以及使用它们获得的结果的描述仅作为示例说明。方法的描述不旨在要求或暗示各种实施例的步骤必须以所呈现的顺序执行。如本领域普通技术人员将理解的,前述实施例中的步骤可以以任何顺序执行。诸如“然后”之类的词不旨在限制步骤的顺序。这些词仅用于指导读者进行方法的描述。许多操作可以并行或同时执行。另外,可以重新安排操作顺序。过程可以对应于方法、函数、过程、子例程、子程序等。当过程对应于函数时,其终止可以对应于该函数返回到调用函数或主函数。对这些实施例的各种修改对于本领域技术人员将是显而易见的,并且在不脱离本公开的精神或范围的情况下,这里定义的一般原理可以应用于其他实施例。

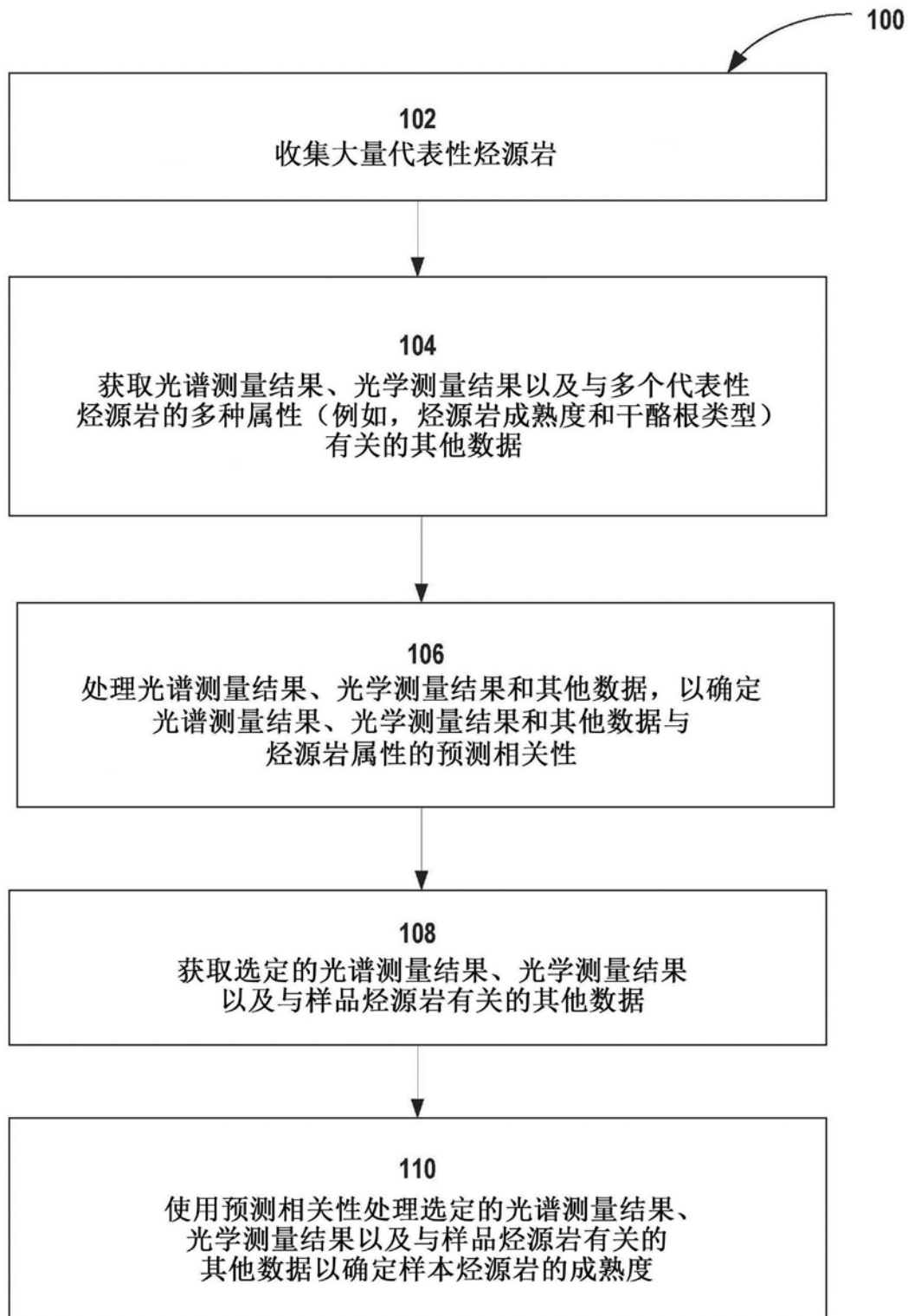


图1

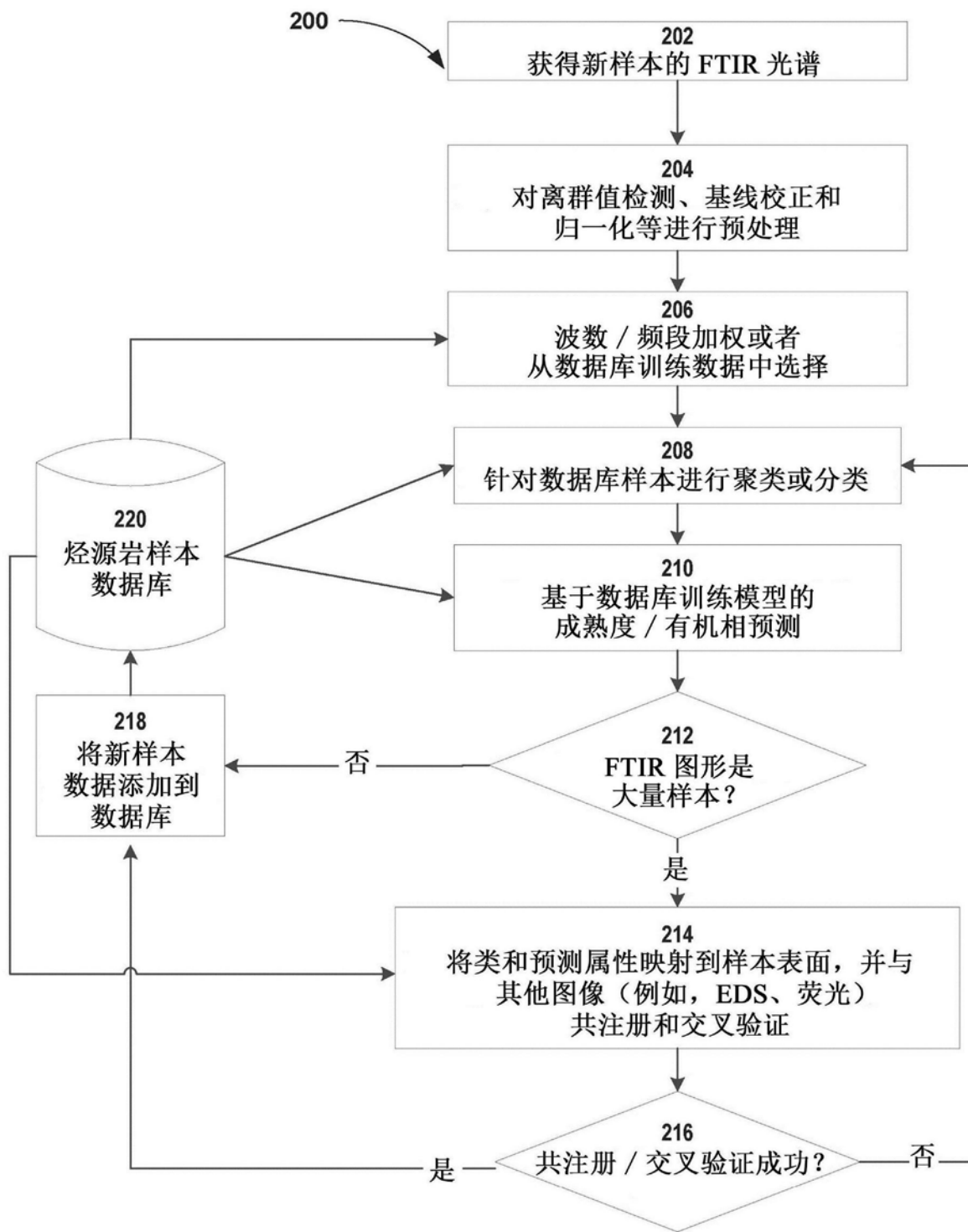


图2

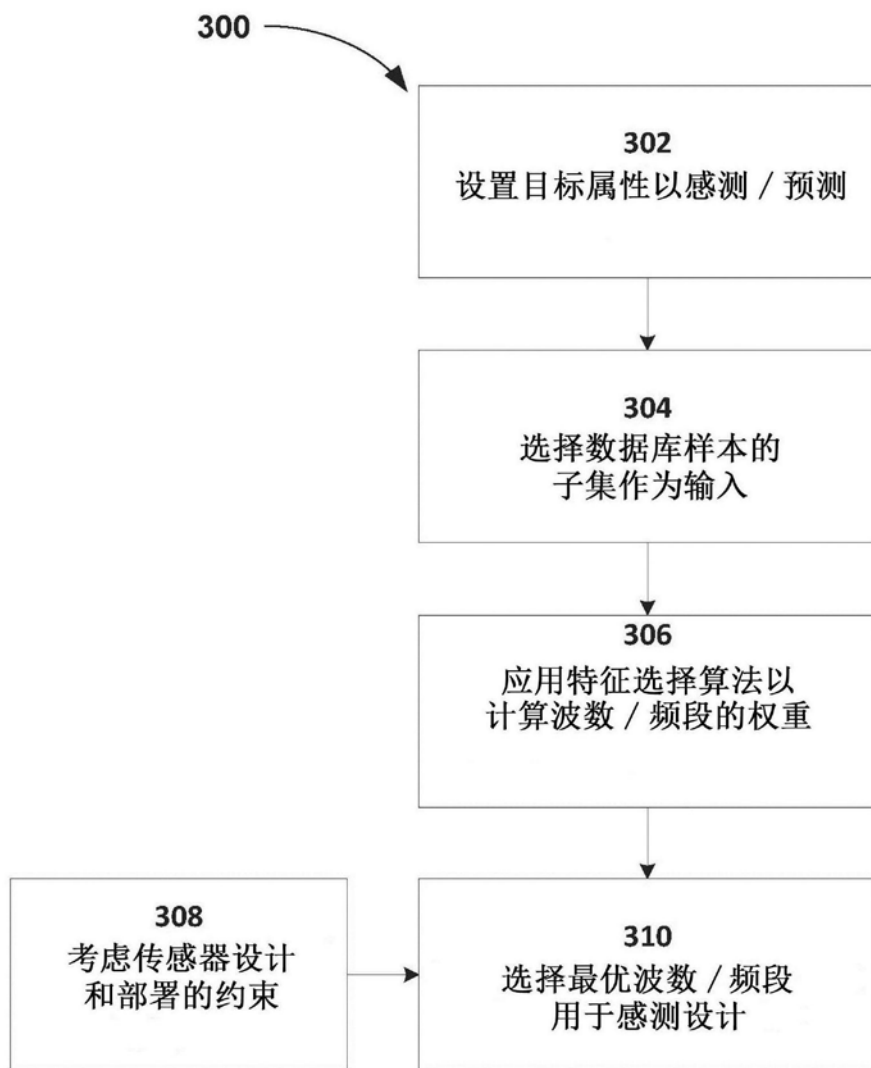


图3

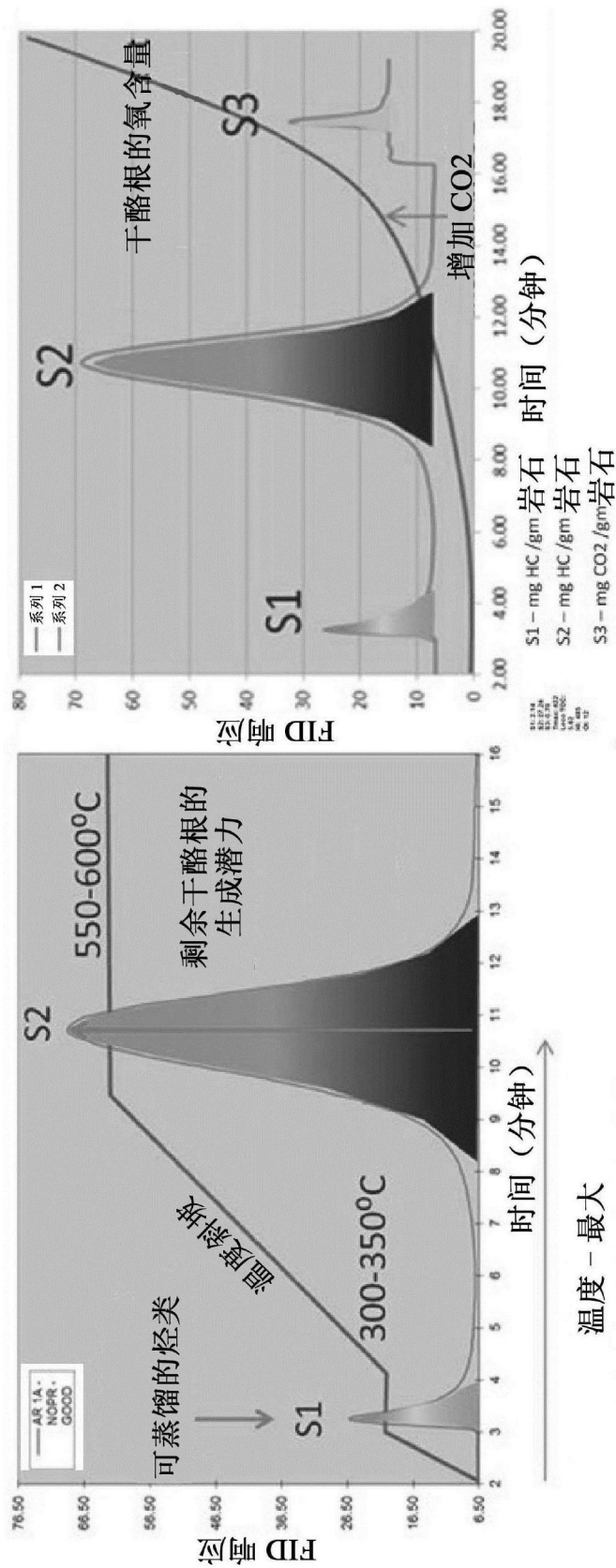


图 4A

图 4B

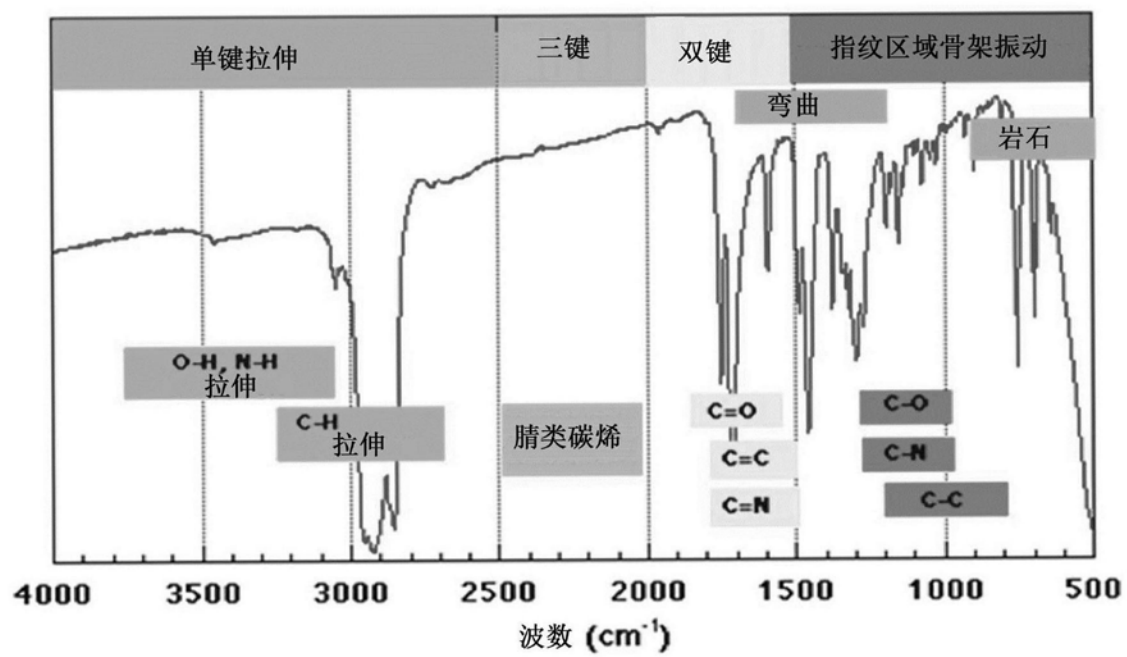


图5

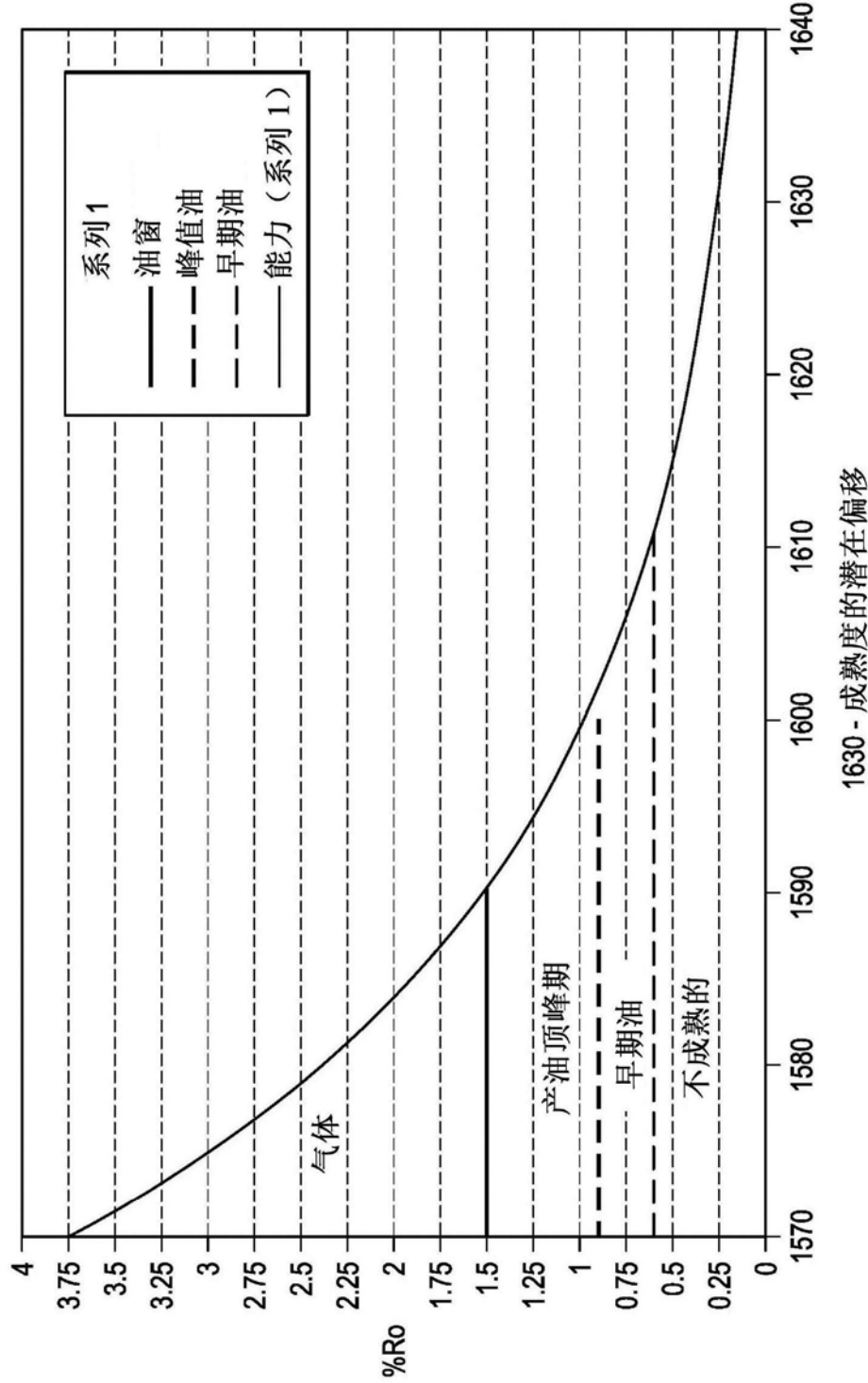


图6

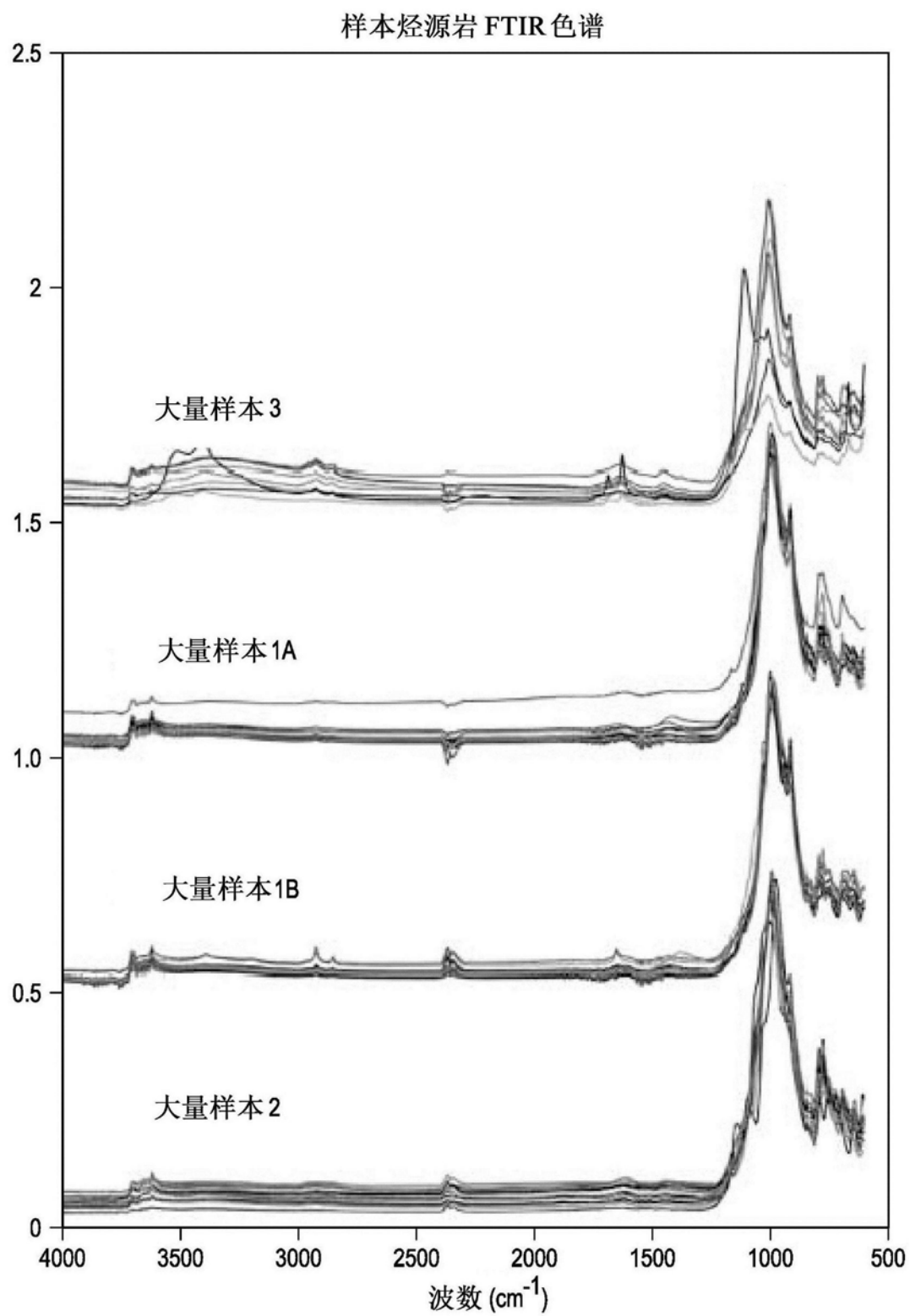


图7A

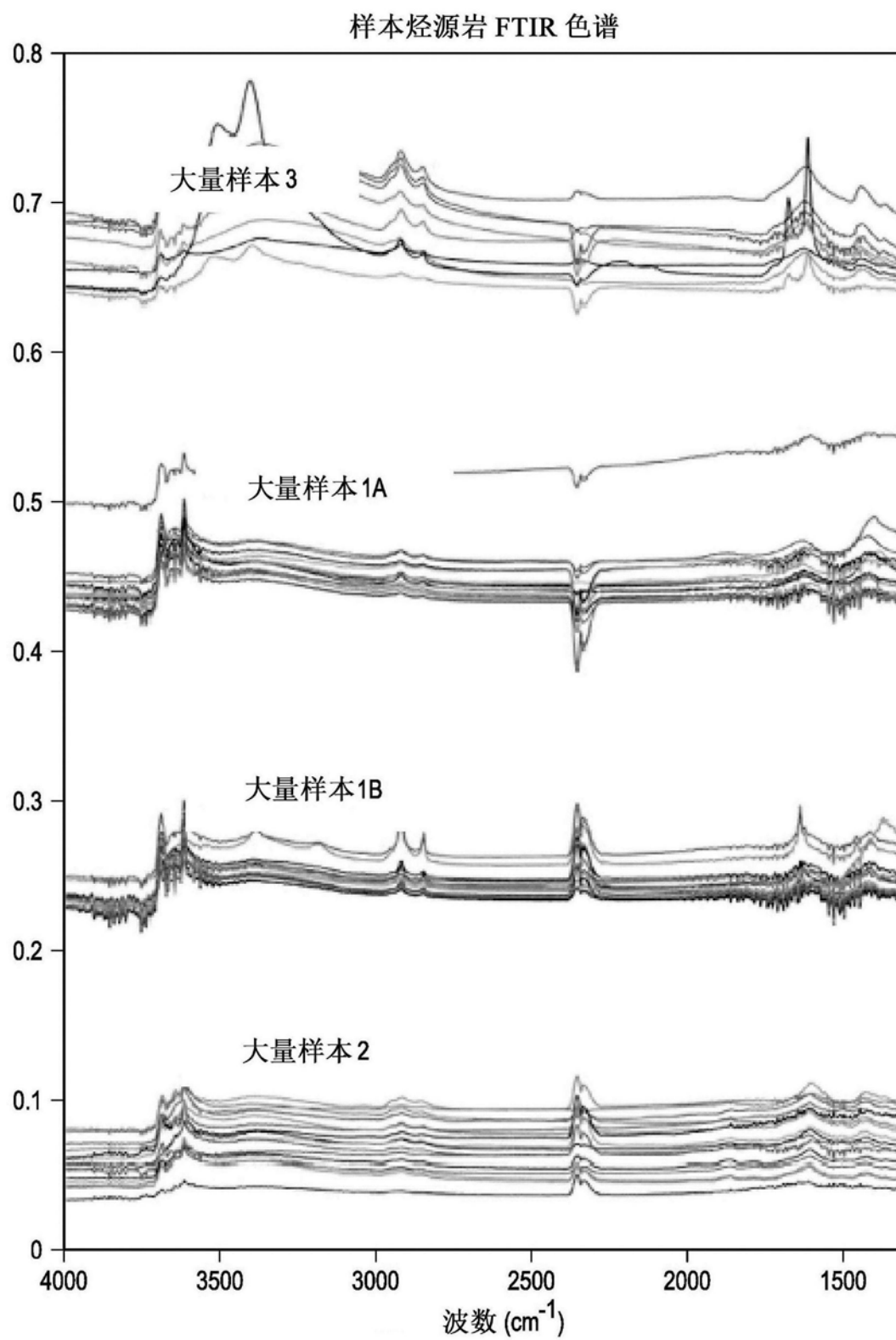


图7B

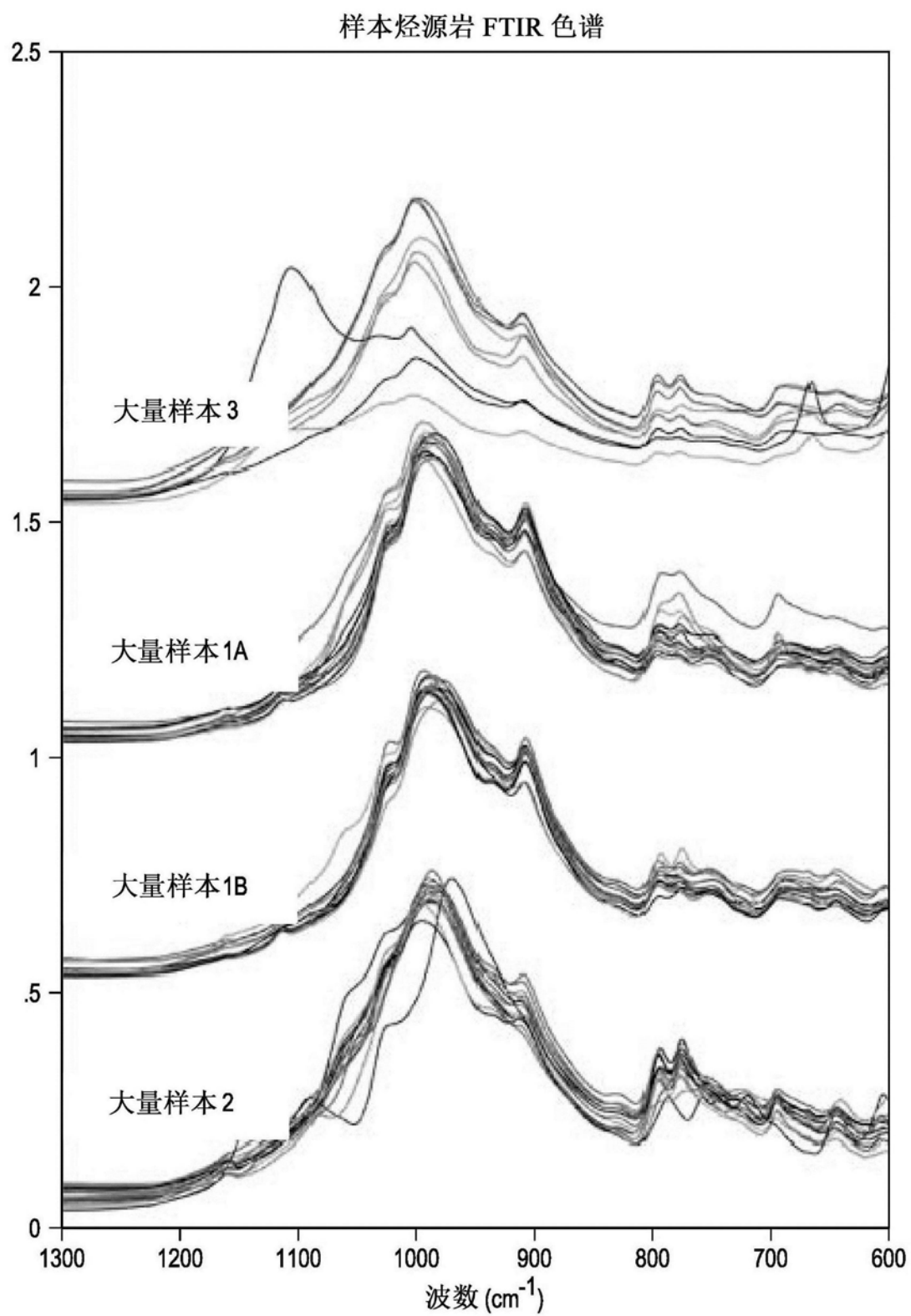


图7C

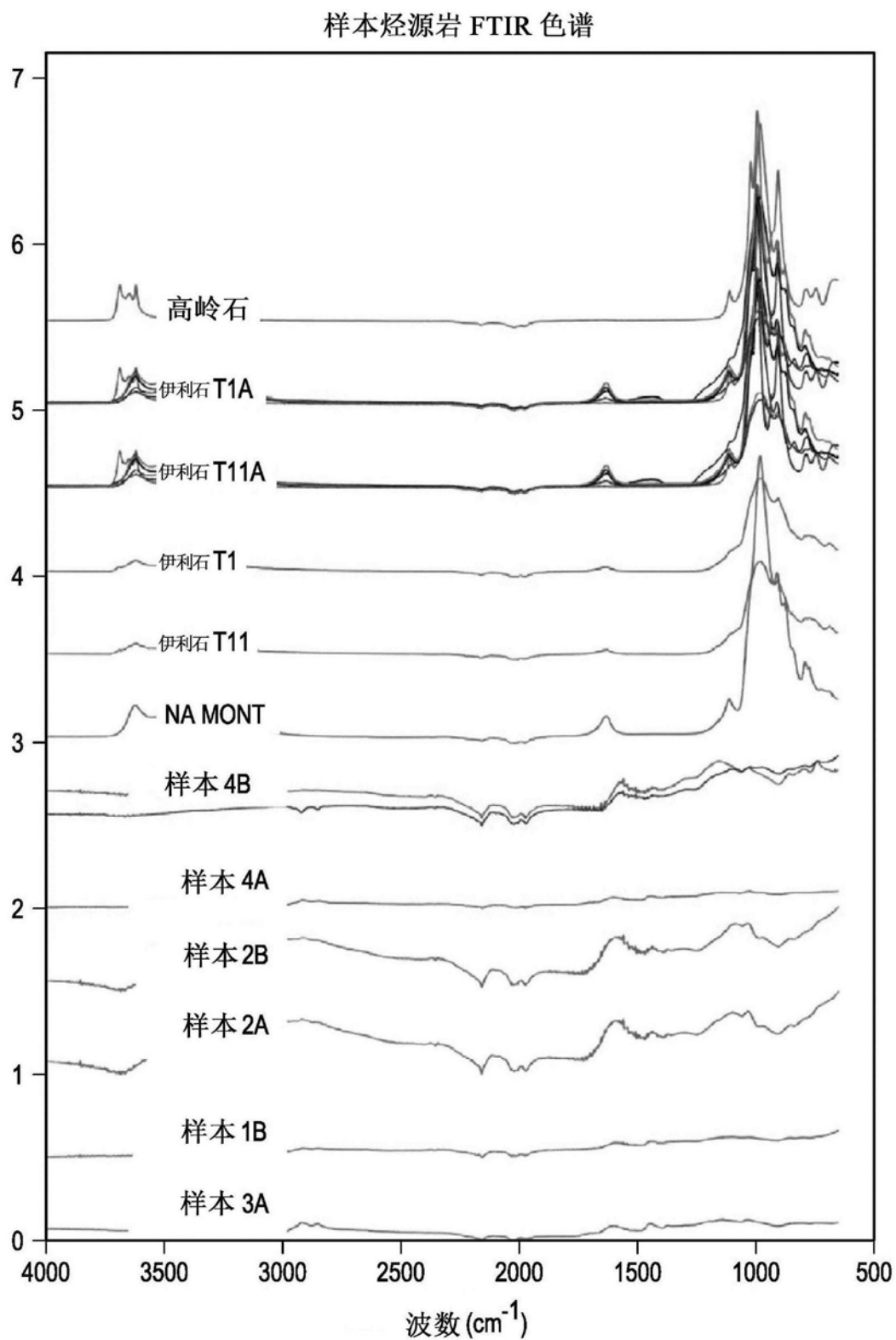


图8A

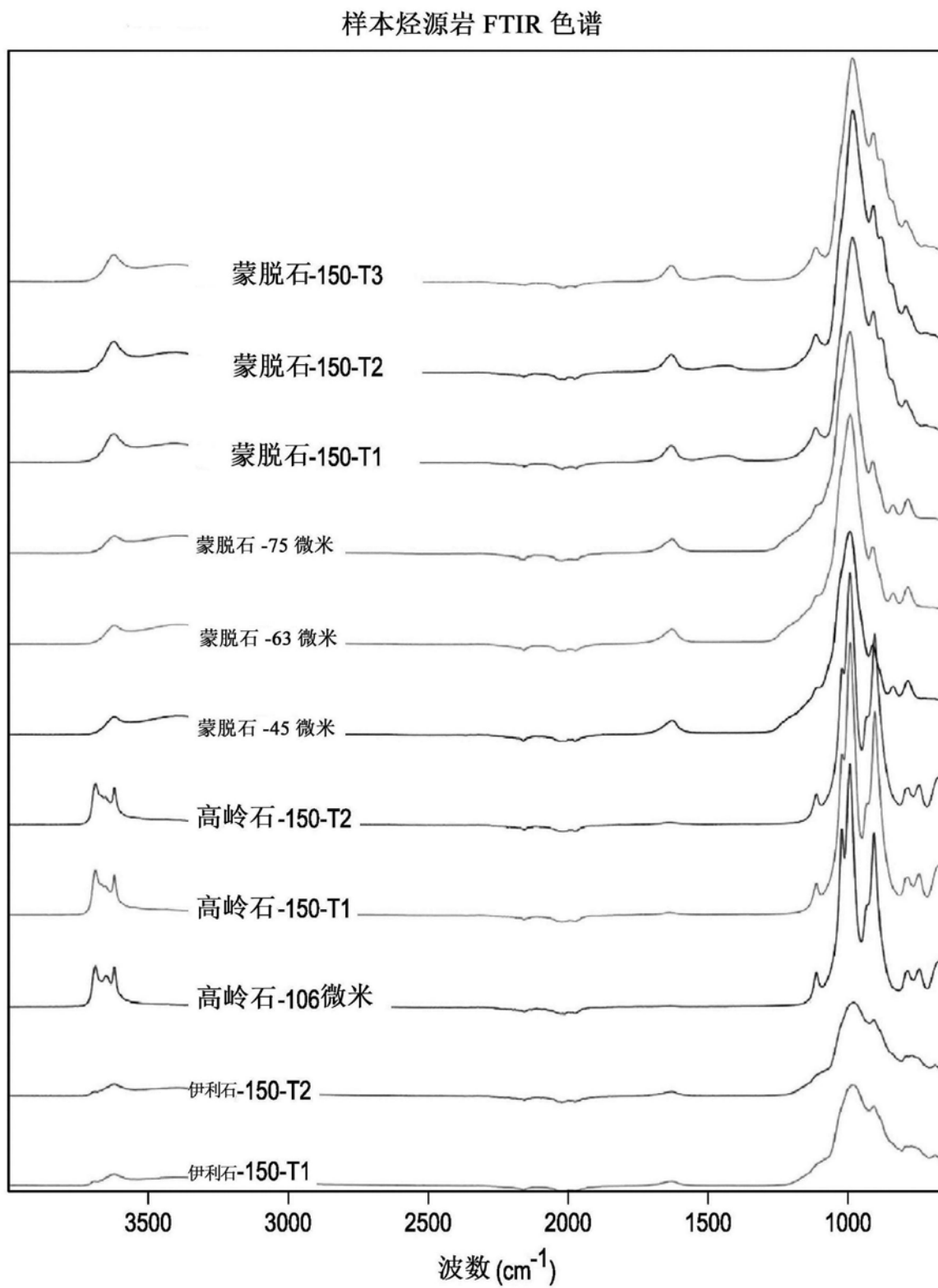


图8B

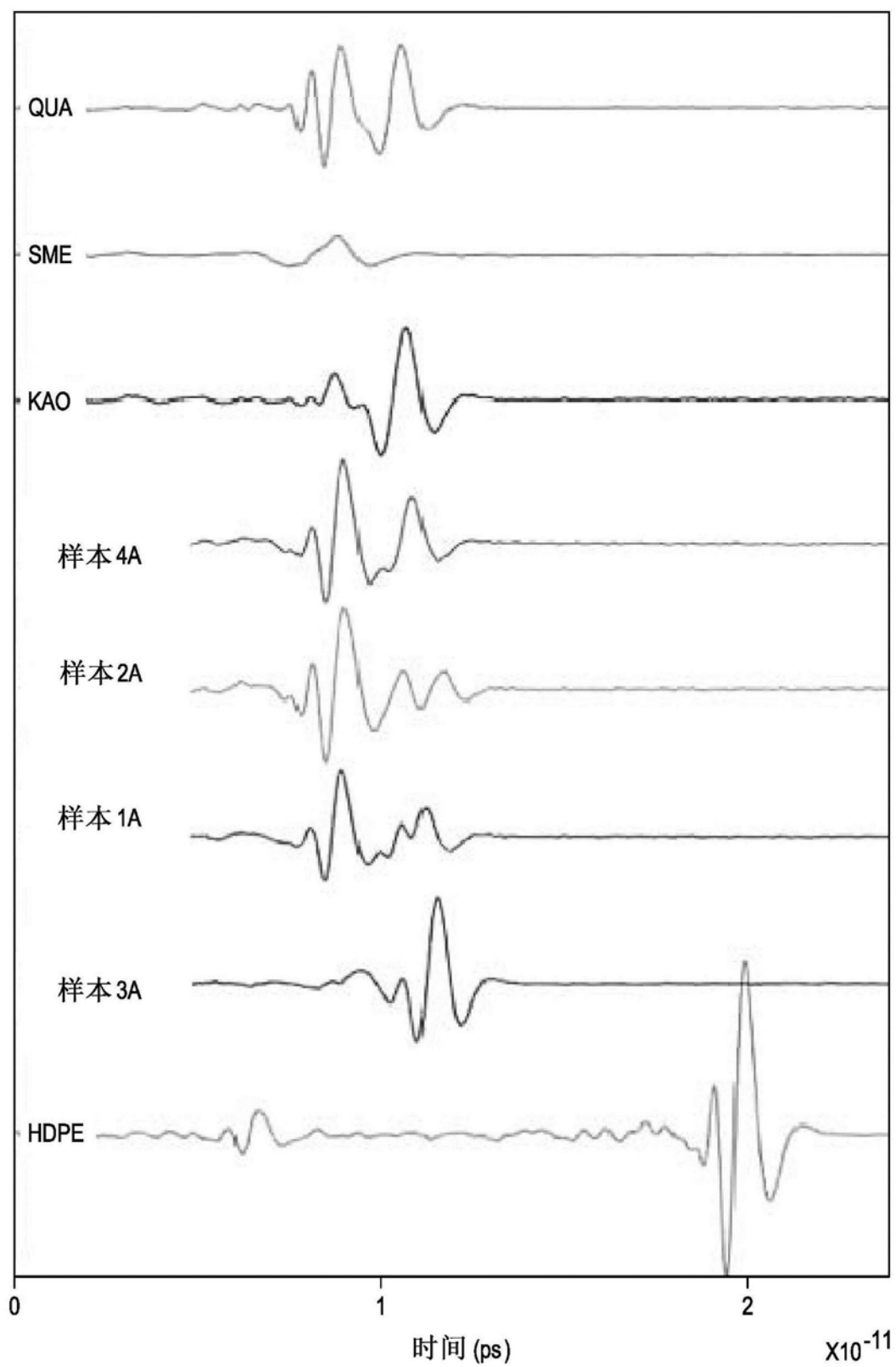


图9A

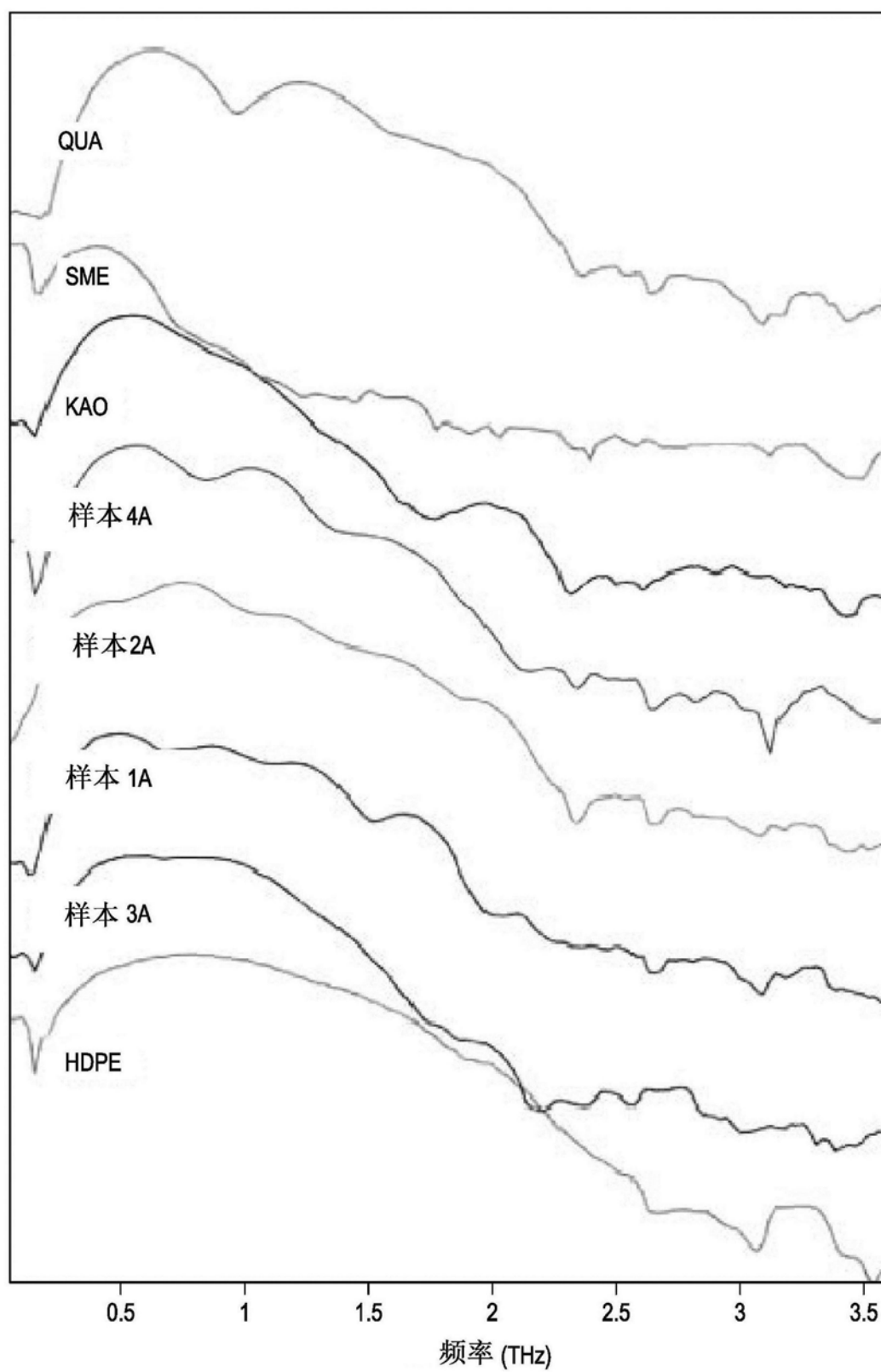


图9B

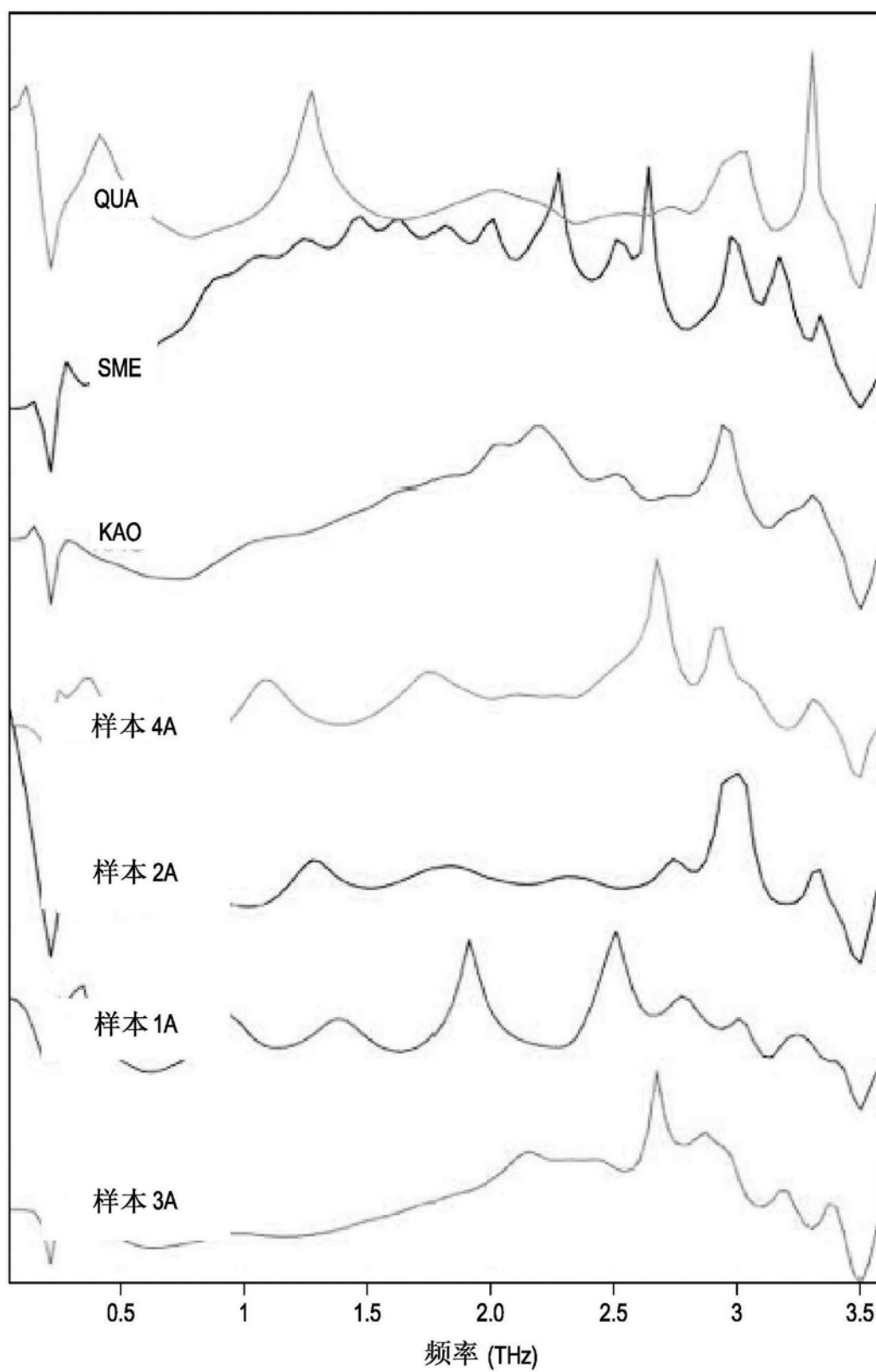


图10A

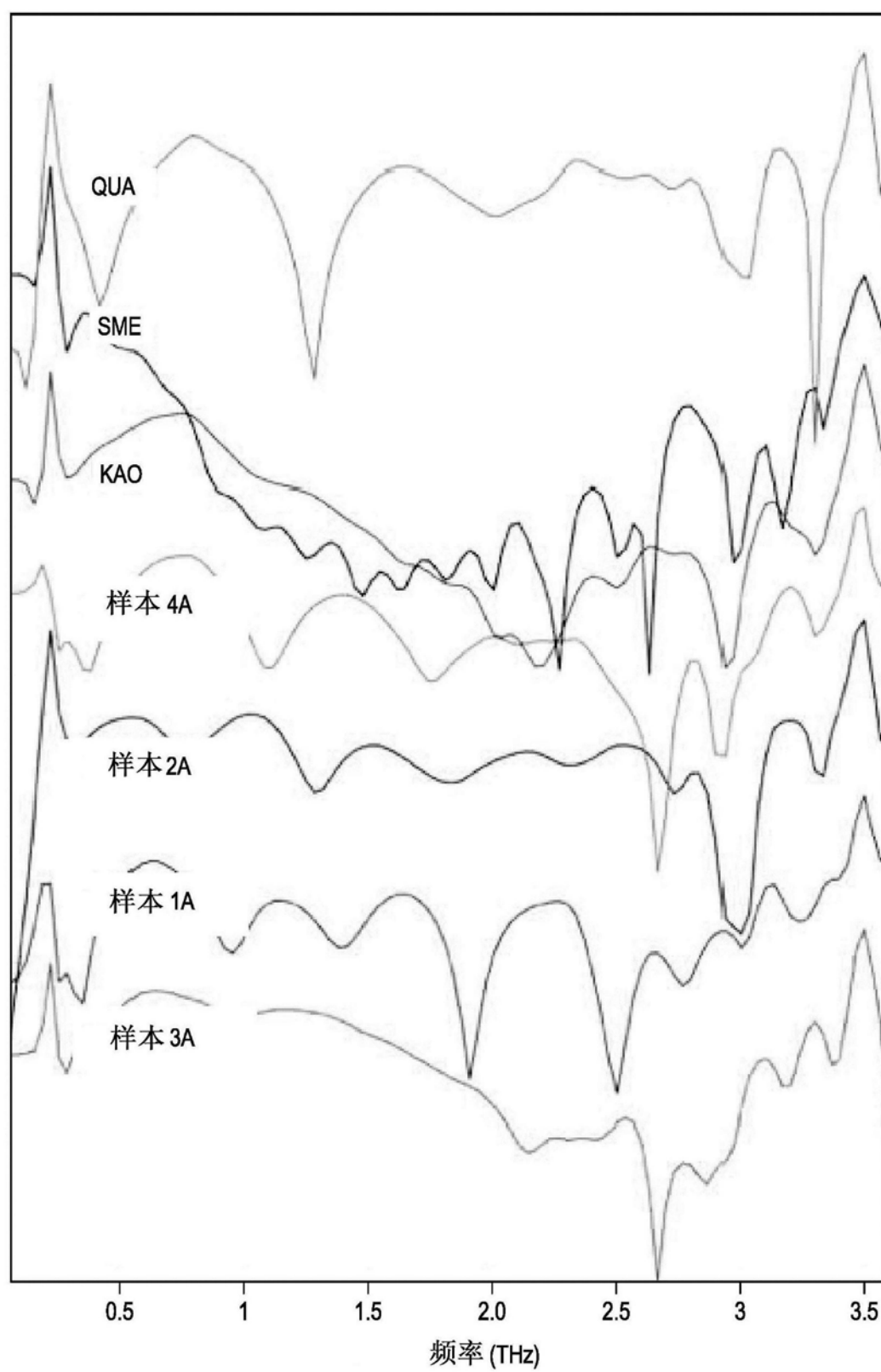


图10B

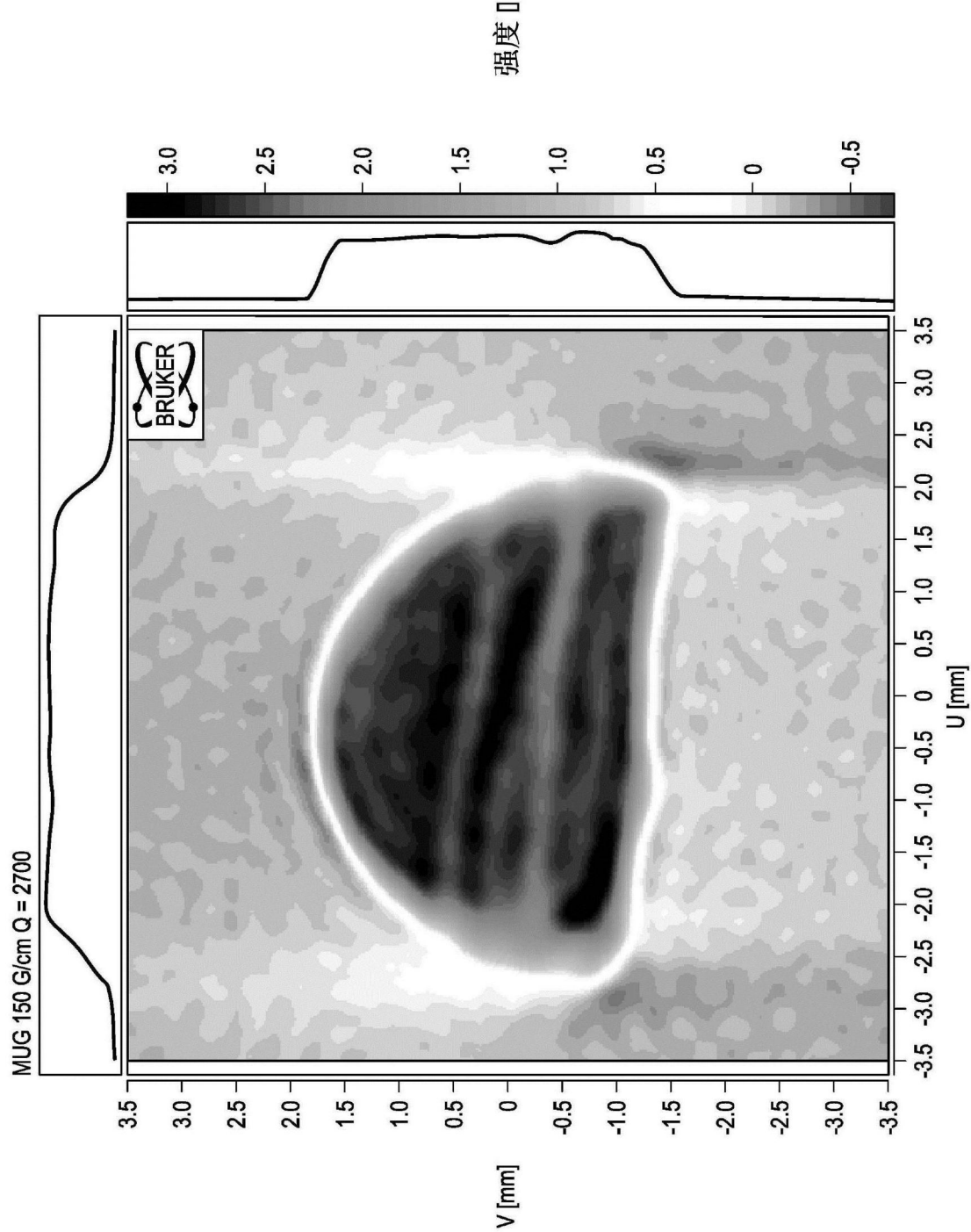


图11A

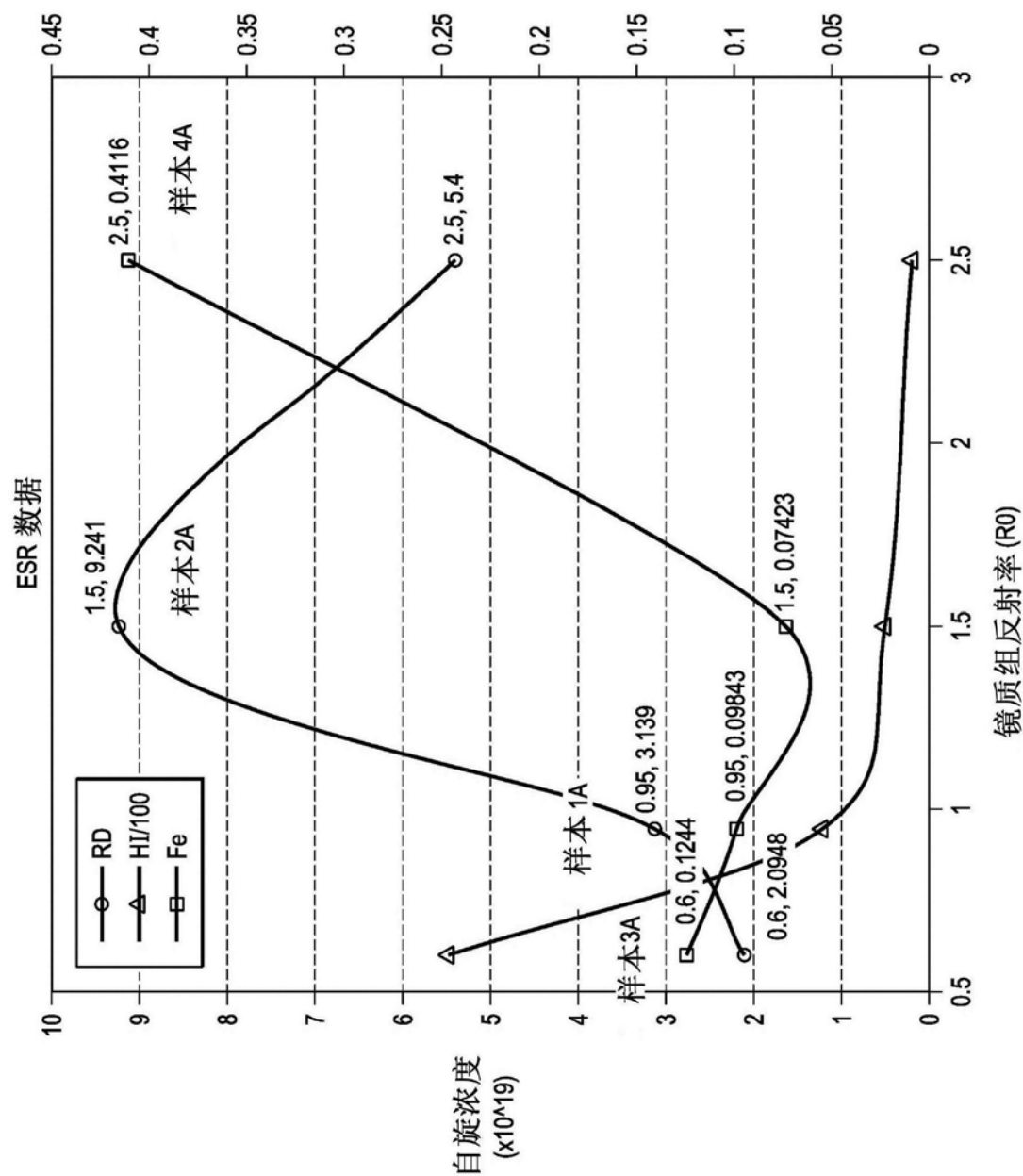
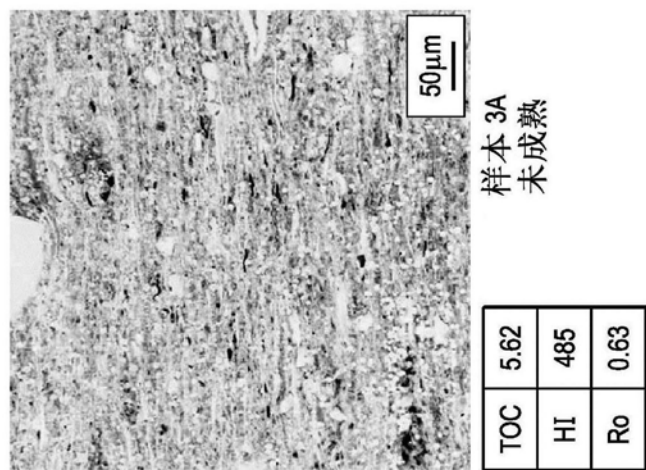
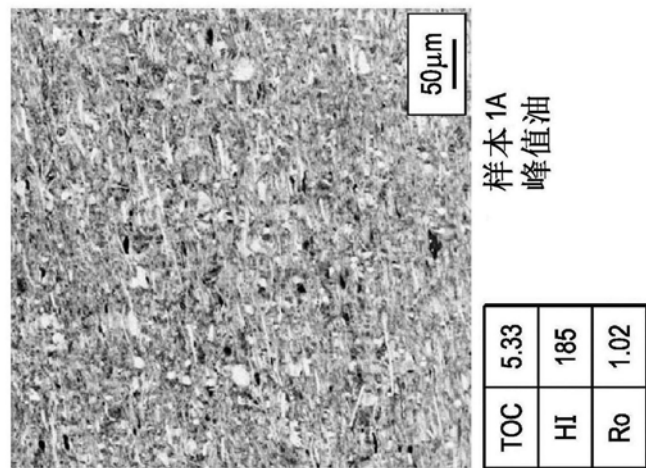


图11B



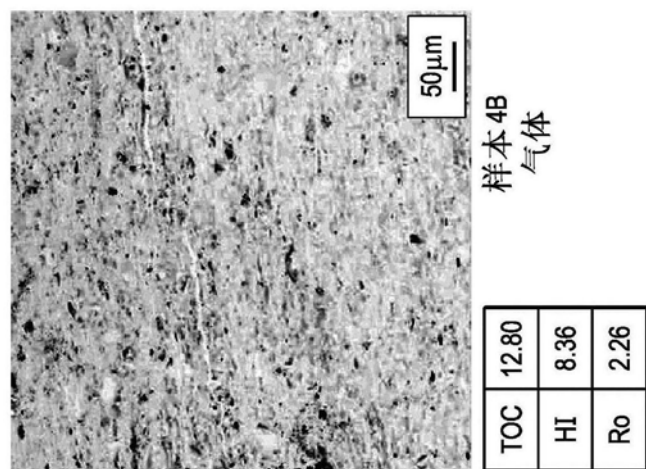
样本 3A
未成熟

图12A



样本 1A
峰值油

图12B



样本 4B
气体

图12C

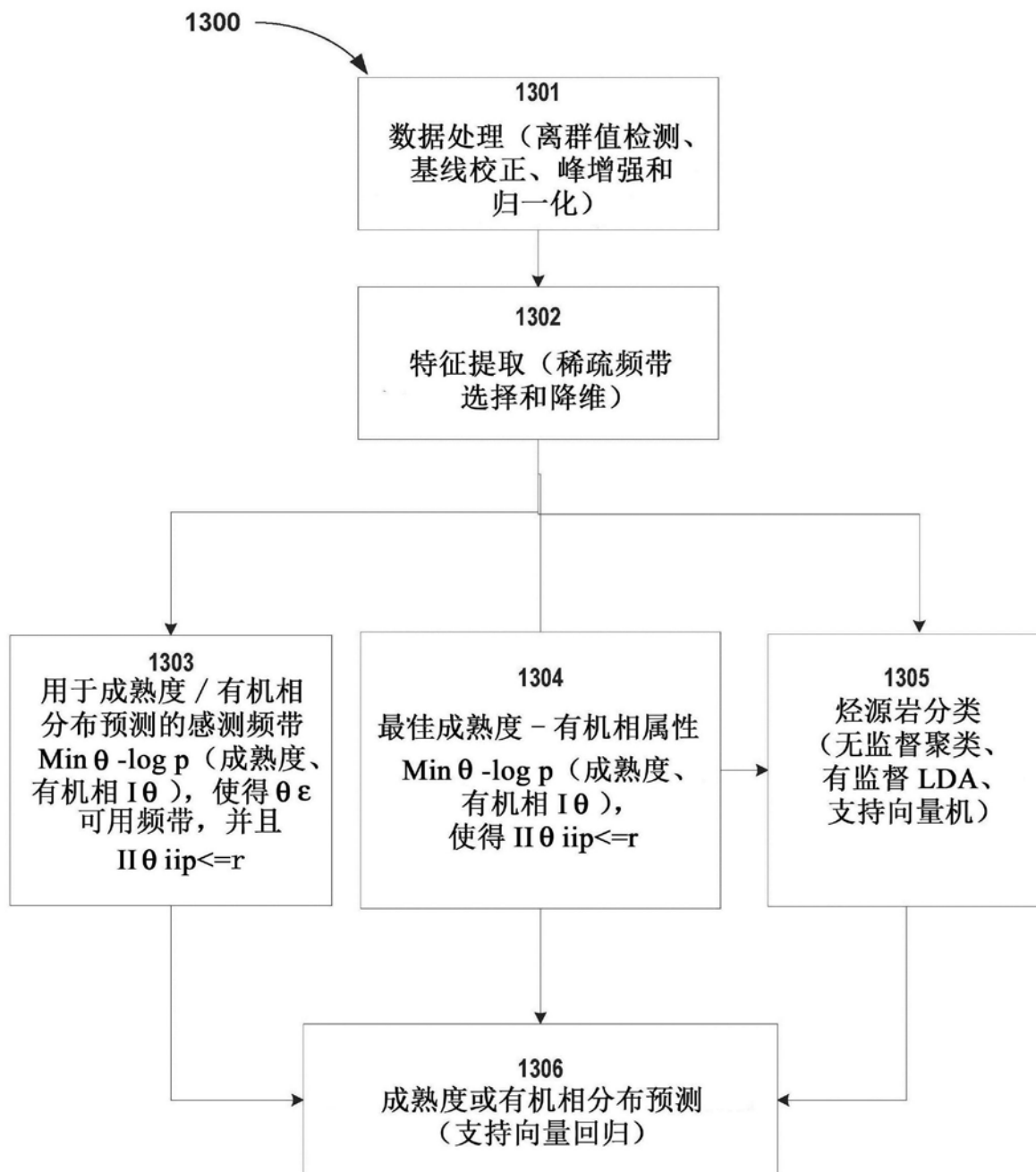


图13

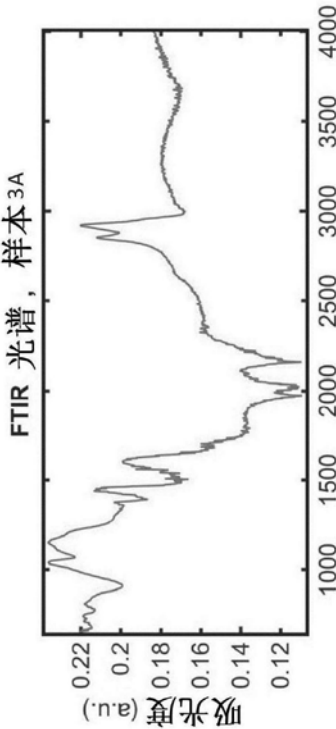


图14A



图14B

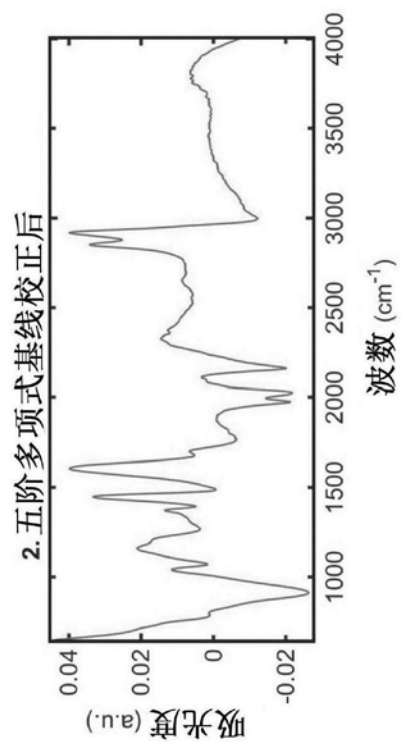


图14C

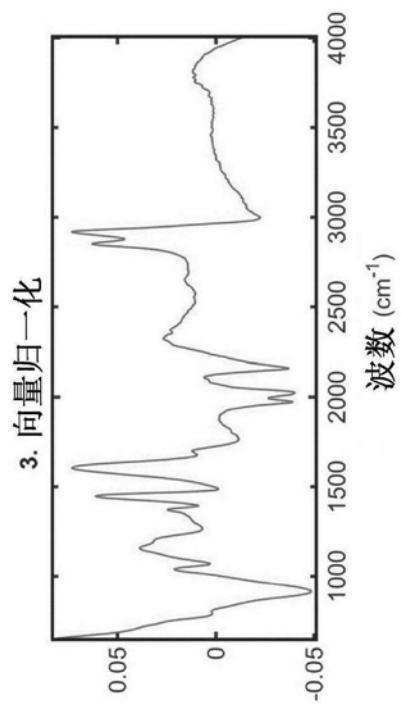


图14D

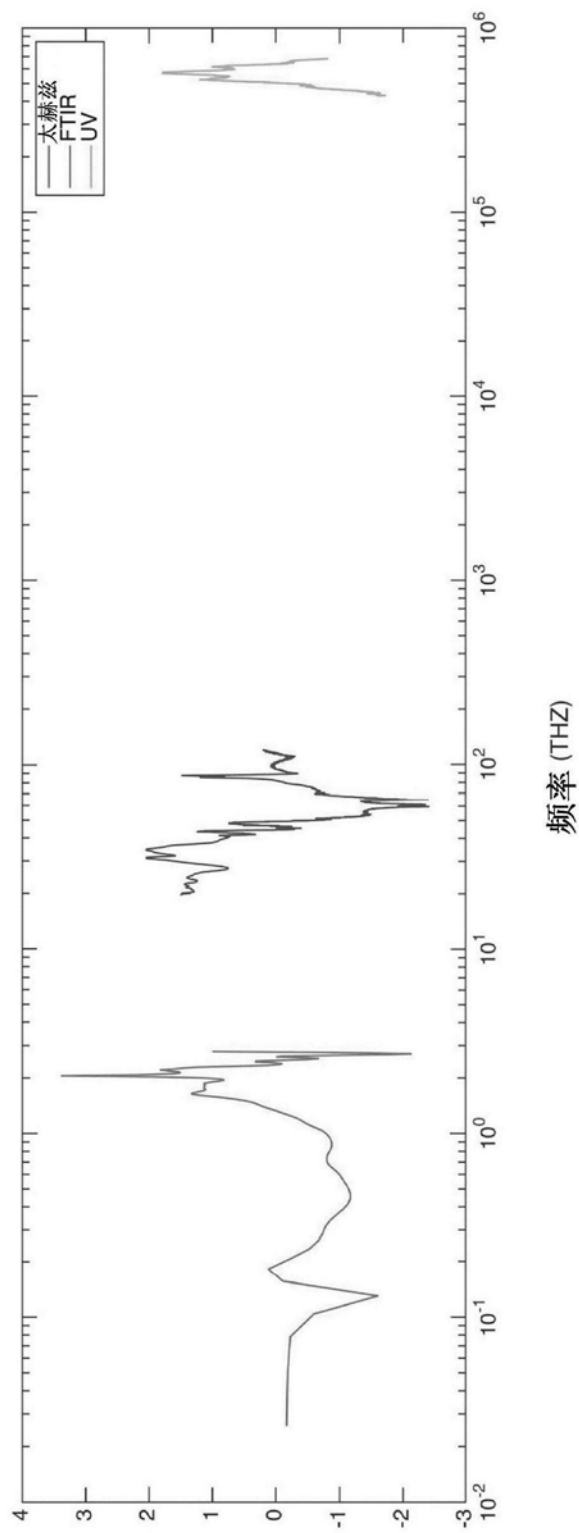


图15

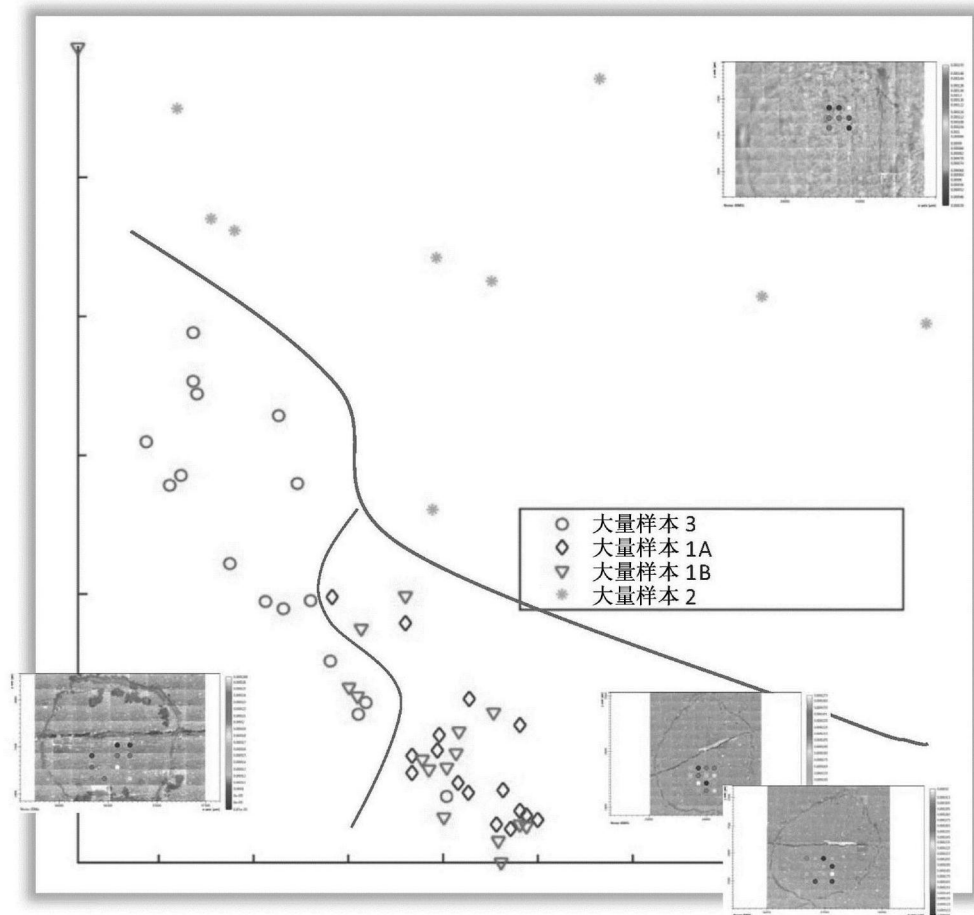


图16A

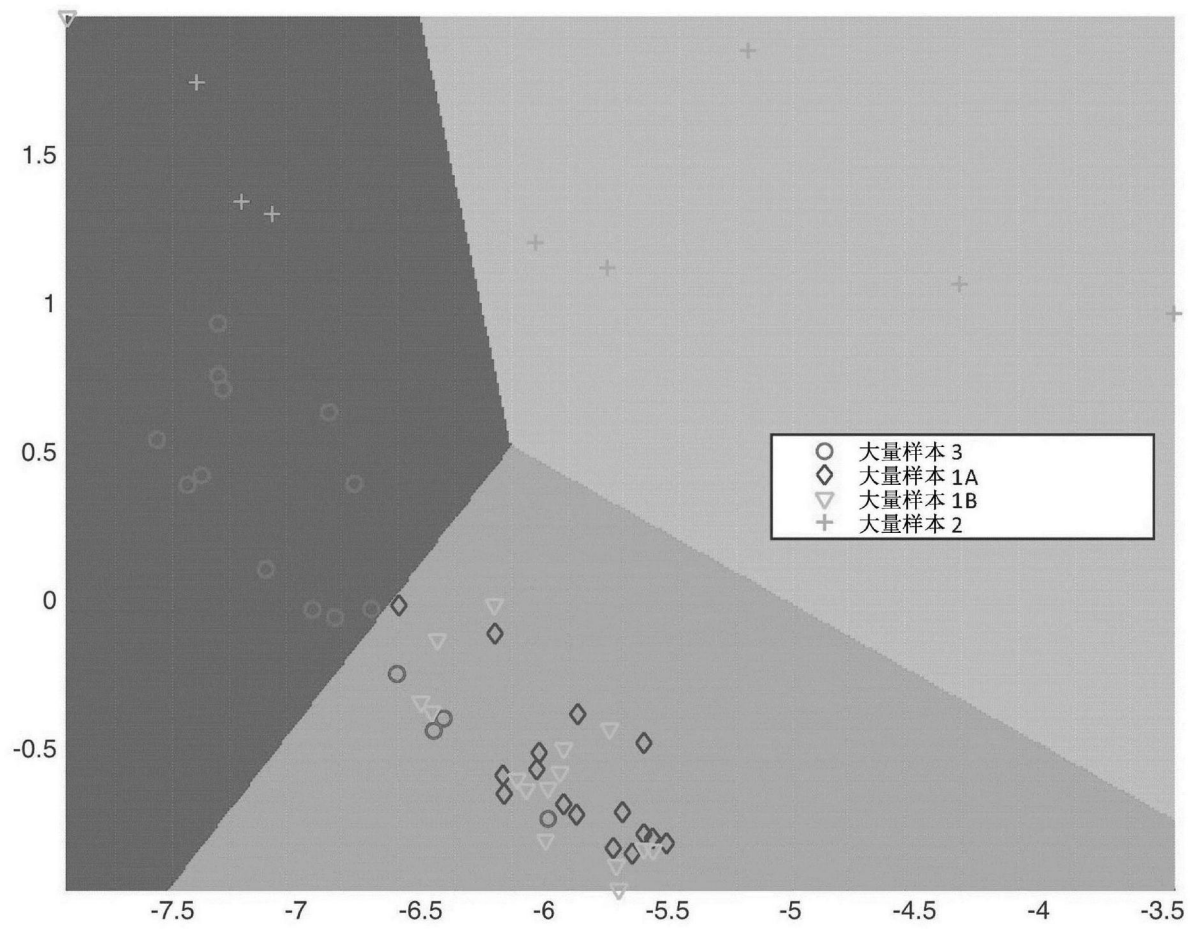


图16B

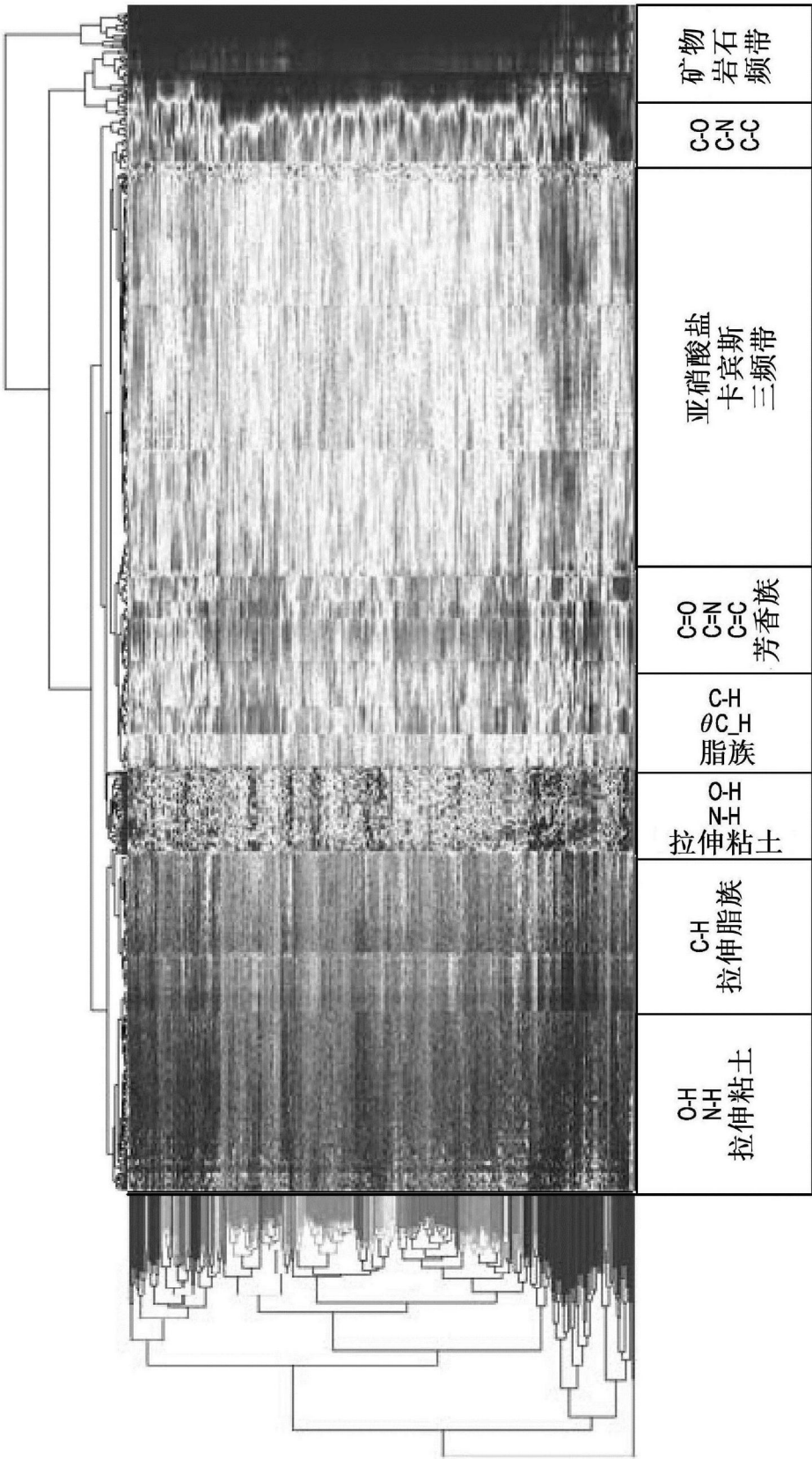


图17

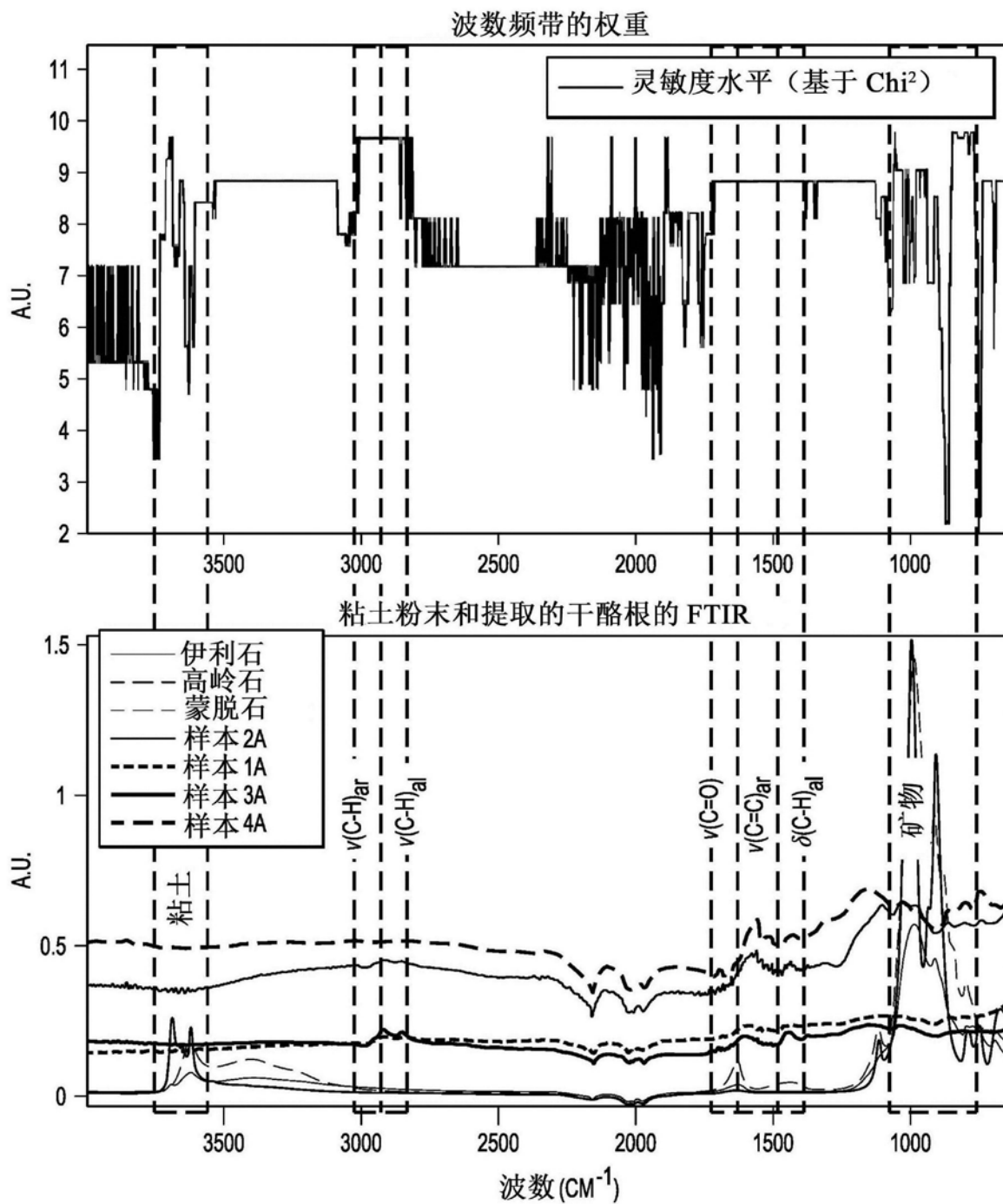


图18

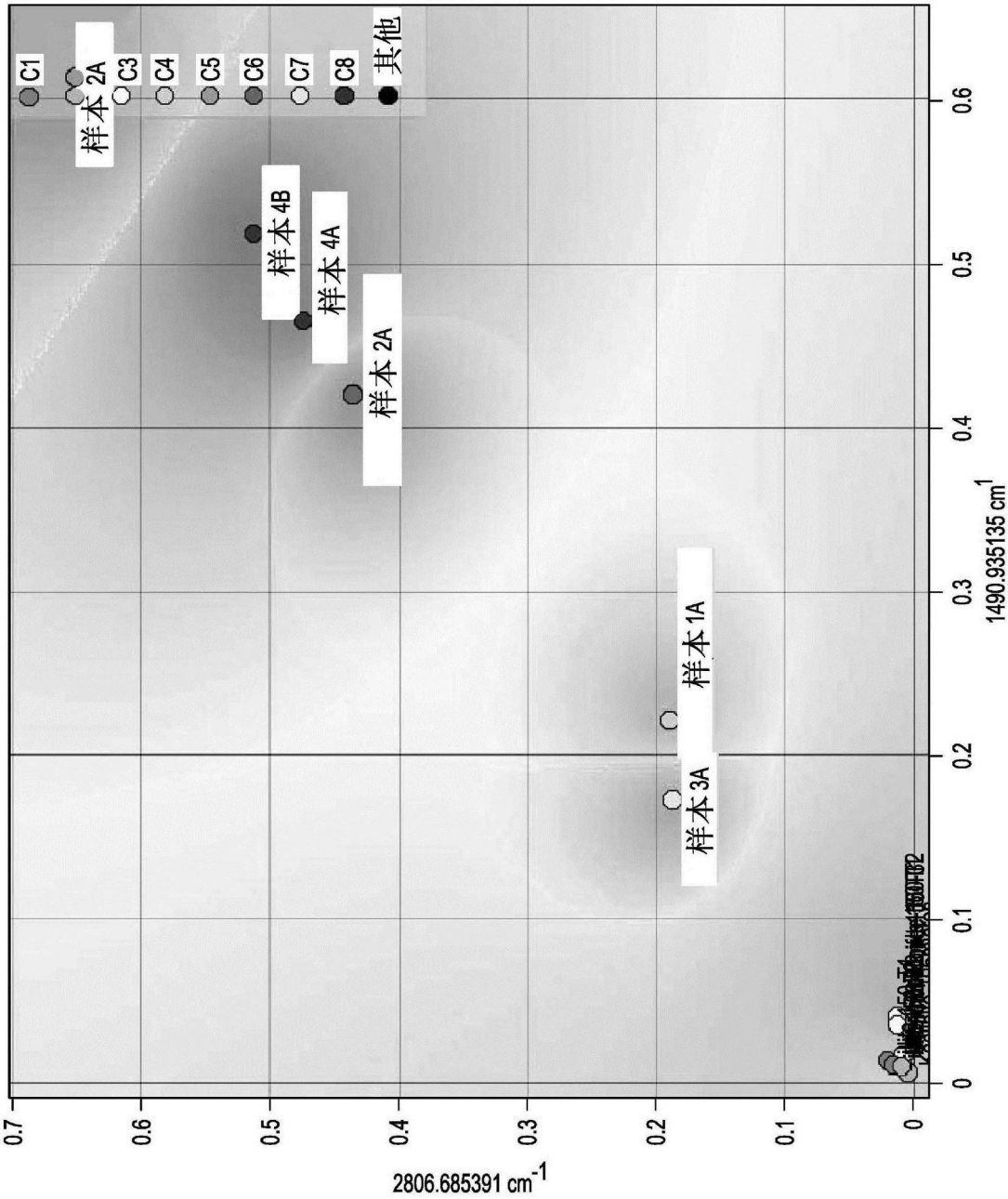


图19A

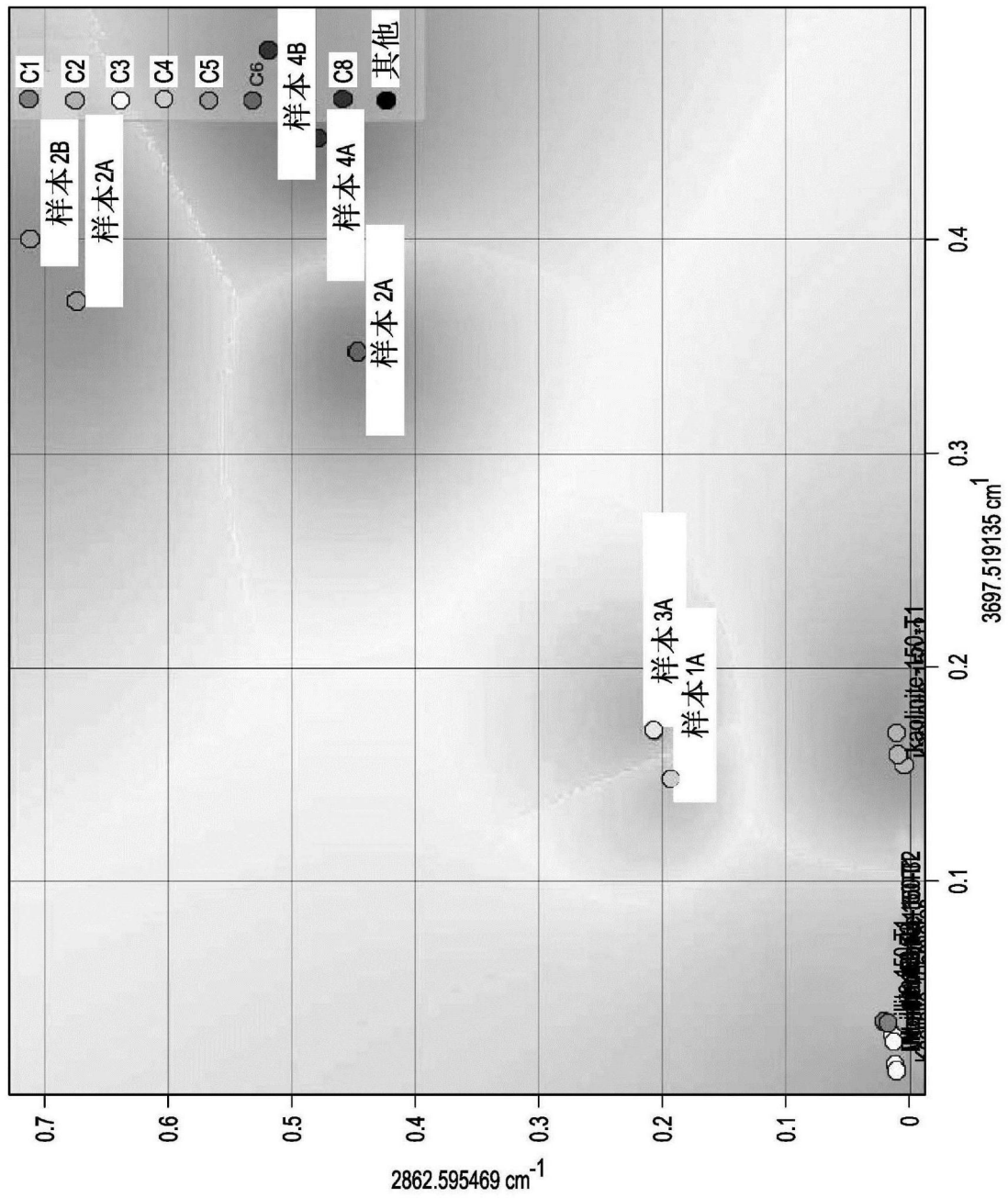


图19B

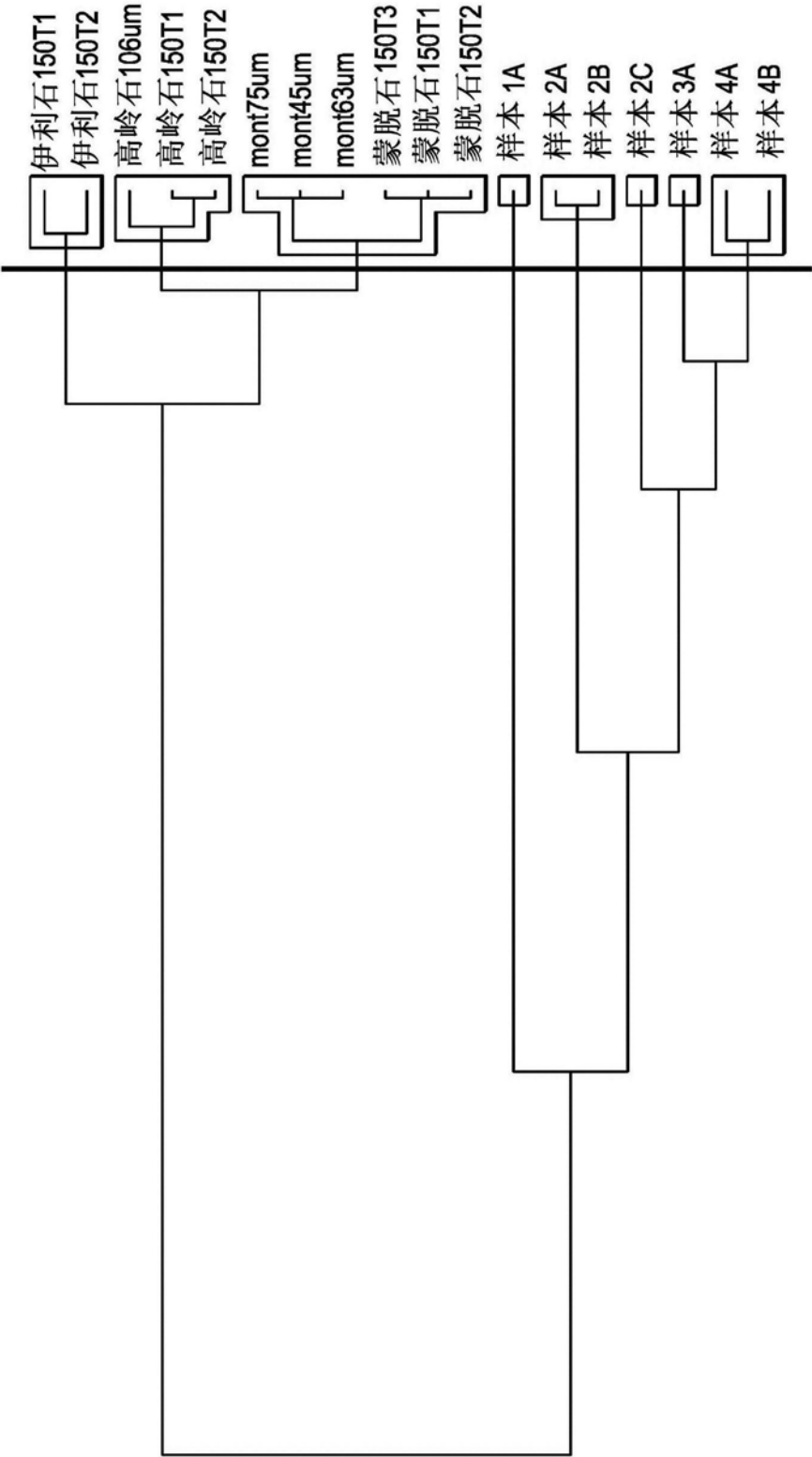


图19C

AdaBoost		预测的							
	C1	C2	C3	C4	C5	C6	C7	C8	Σ
C1	40.0%	0.0%	0.0%	NA	0.0%	NA	NA	0.0%	2
C2	0.0%	100.0%	0.0%	NA	0.0%	NA	NA	0.0%	3
C3	20.0%	0.0%	100.0%	NA	0.0%	NA	NA	0.0%	6
C4	20.0%	0.0%	0.0%	NA	0.0%	NA	NA	0.0%	1
C5	0.0%	0.0%	0.0%	NA	66.7%	NA	NA	0.0%	2
C6	0.0%	0.0%	0.0%	NA	33.3%	NA	NA	0.0%	1
C7	20.0%	0.0%	0.0%	NA	0.0%	NA	NA	0.0%	1
C8	0.0%	0.0%	0.0%	NA	0.0%	NA	NA	100.0%	2
Σ	5	3	5	0	3	0	0	2	18

图20A

实际的

RANDOM FOREST		预测的								
	C1	C2	C3	C4	C5	C6	C7	C8	Σ	
C1	100.0%	0.0%	14.3%	NA	NA	0.0%	NA	0.0%	2	
C2	0.0%	75.0%	0.0%	NA	NA	0.0%	NA	0.0%	3	
C3	0.0%	0.0%	85.7%	NA	NA	0.0%	NA	0.0%	6	
C4	0.0%	0.0%	0.0%	NA	NA	100.0%	NA	0.0%	1	
C5	0.0%	0.0%	0.0%	NA	NA	0.0%	NA	40.0%	2	
C6	0.0%	0.0%	0.0%	NA	NA	0.0%	NA	20.0%	1	
C7	0.0%	25.0%	0.0%	NA	NA	0.0%	NA	0.0%	1	
C8	0.0%	0.0%	0.0%	NA	NA	0.0%	NA	40.0%	2	
Σ	1	4	7	0	0	1	0	5	18	

图20B

数据和预测				
	AdaBoost	RANDOM FOREST REGRESSION	HI-%Ro	样本名称
1	0.000	0.000	0.000	伊利石 150T1
2	0.000	0.000	0.000	伊利石 150T2
3	0.000	0.000	0.000	高岭石 106um
4	0.000	0.000	0.000	高岭石 150T1
5	0.000	0.000	0.000	高岭石 150T2
6	0.000	0.000	0.000	mont45um
7	0.000	0.000	0.000	mont63um
8	0.000	0.000	0.000	mont75um
9	0.000	0.000	0.000	蒙脱石 150T1
10	0.000	0.000	0.000	蒙脱石 150T2
11	0.000	0.000	0.000	蒙脱石 150T3
12	1.590	1.658	1.590	样本 1A
13	1.590	1.658	1.590	样本 2A
14	1.590	1.658	1.590	样本 2B
15	1.020	1.137	1.020	样本 2C
16	0.630	0.910	0.630	样本 3A
17	2.160	1.791	2.160	样本 4A
18	2.160	1.791	2.160	样本 4B

图21

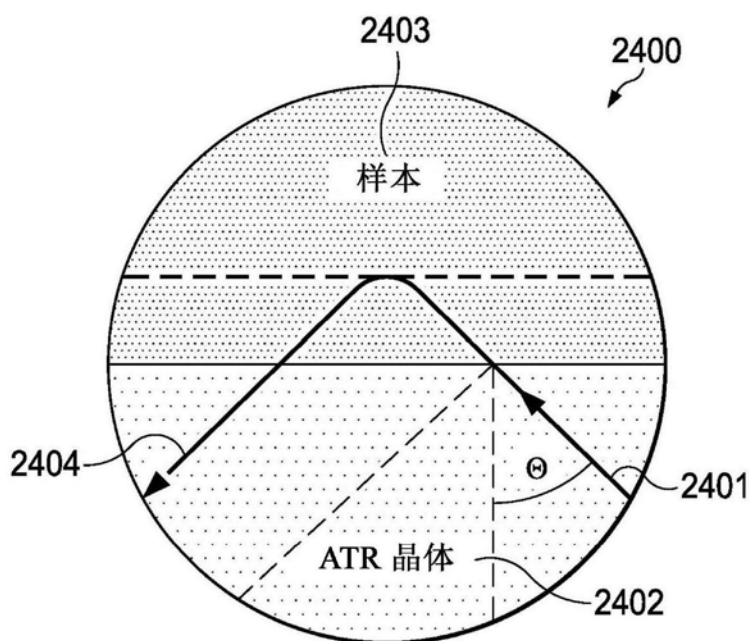


图24

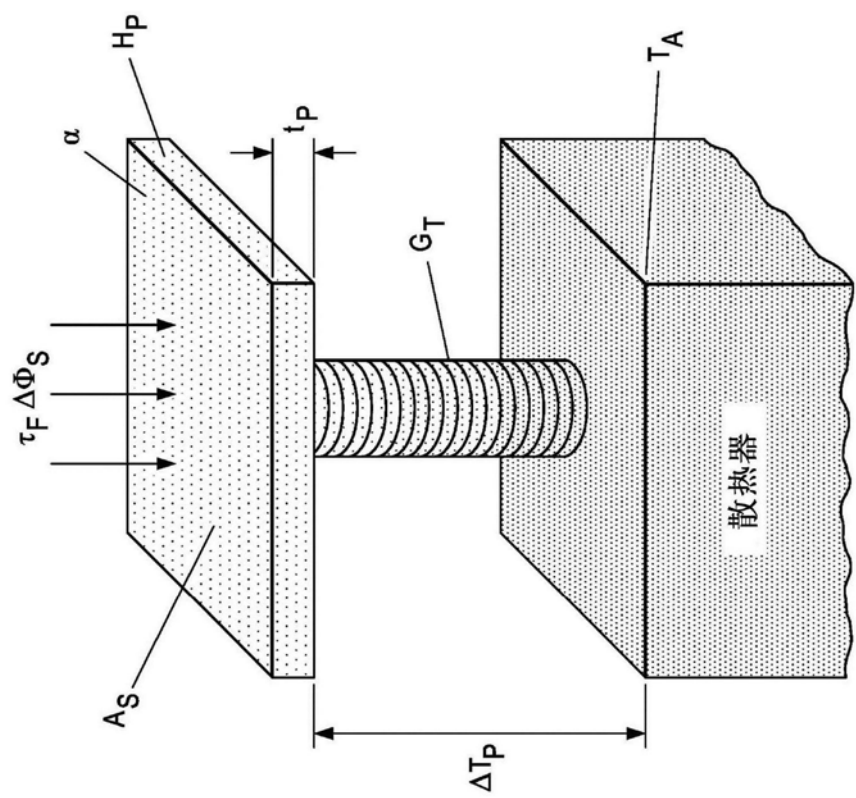


图22A

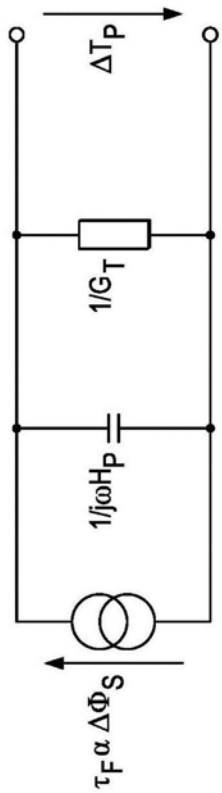


图22B

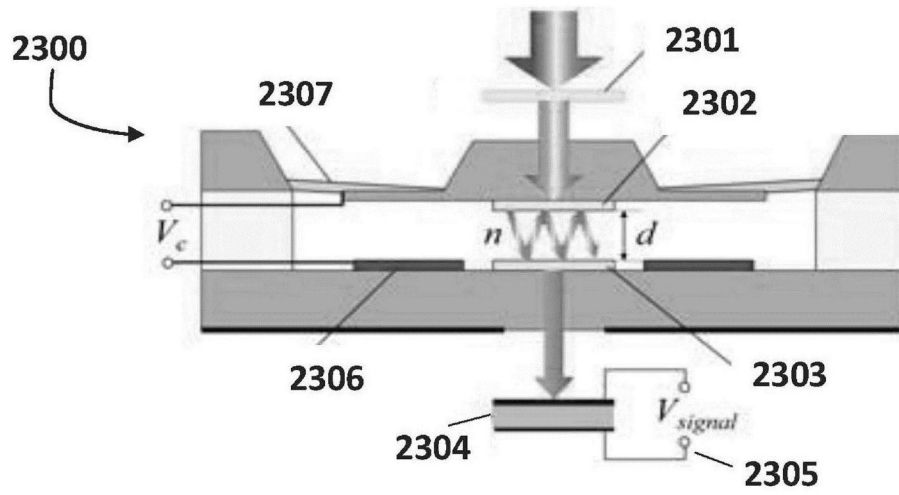


图23

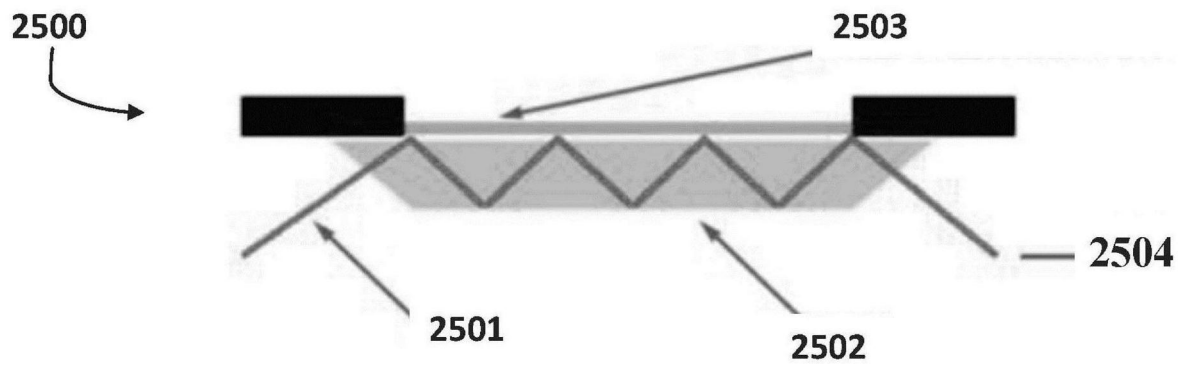


图25

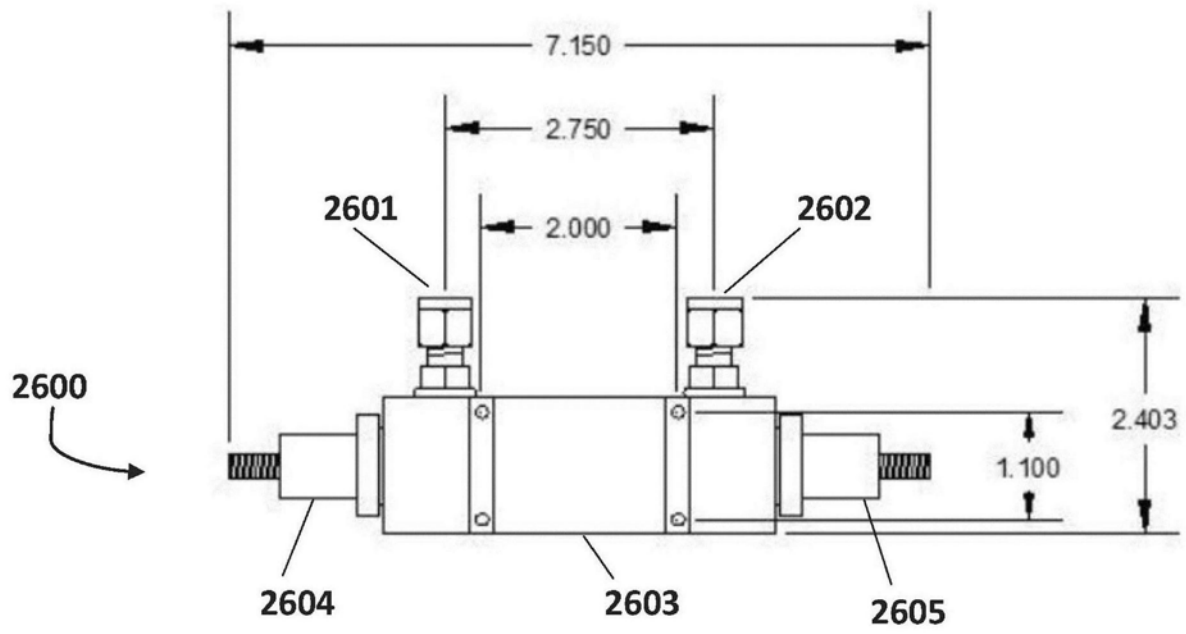


图26

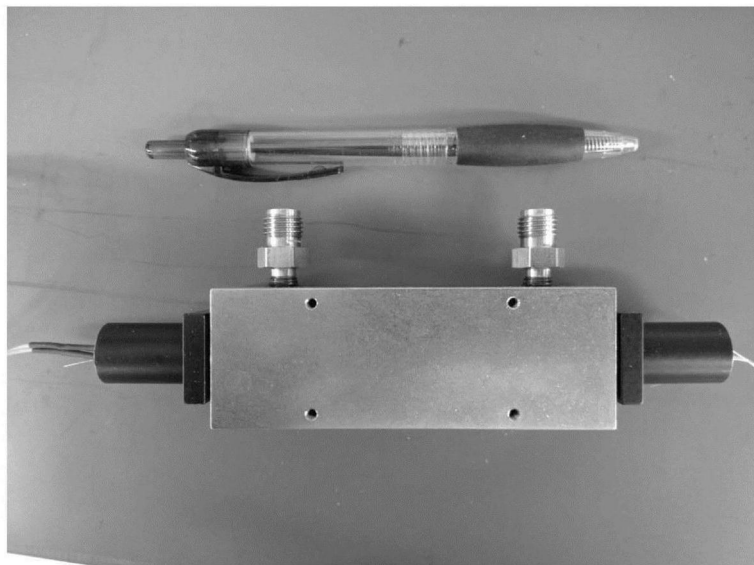


图27A



图27B

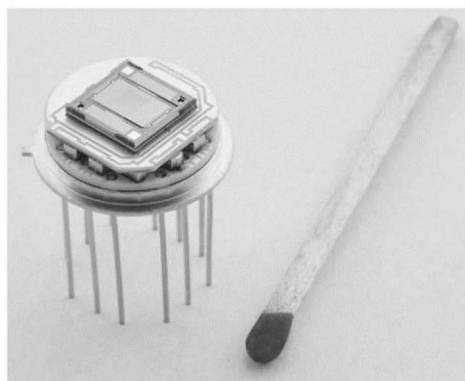


图27C

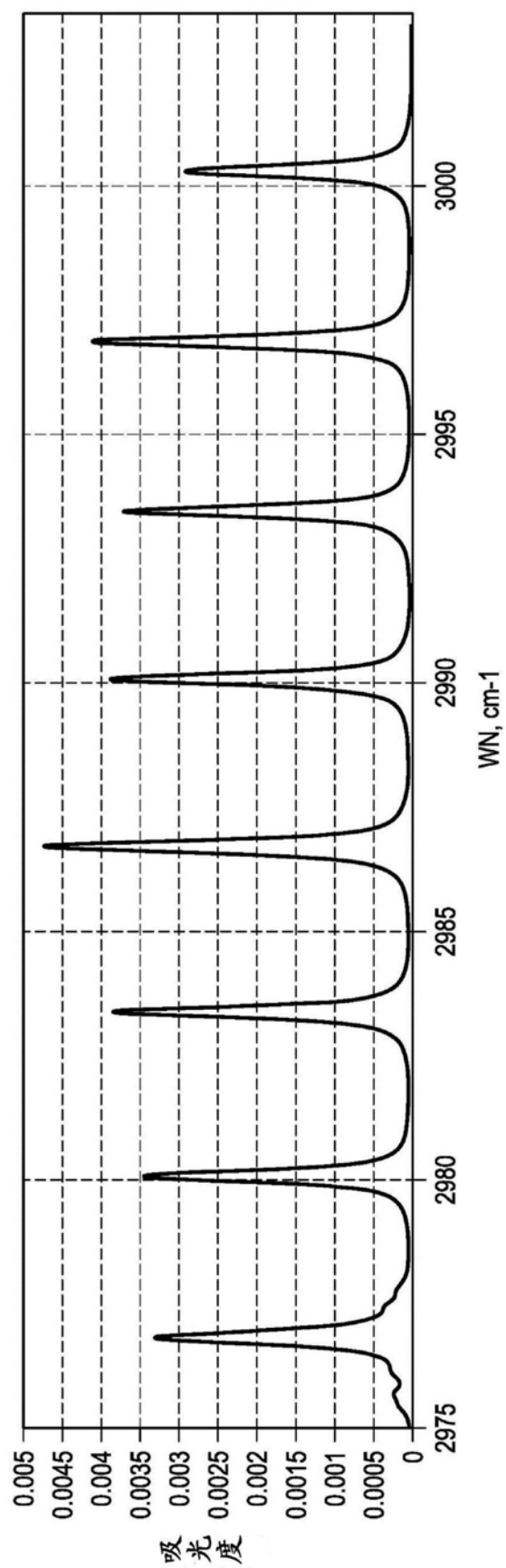


图28A

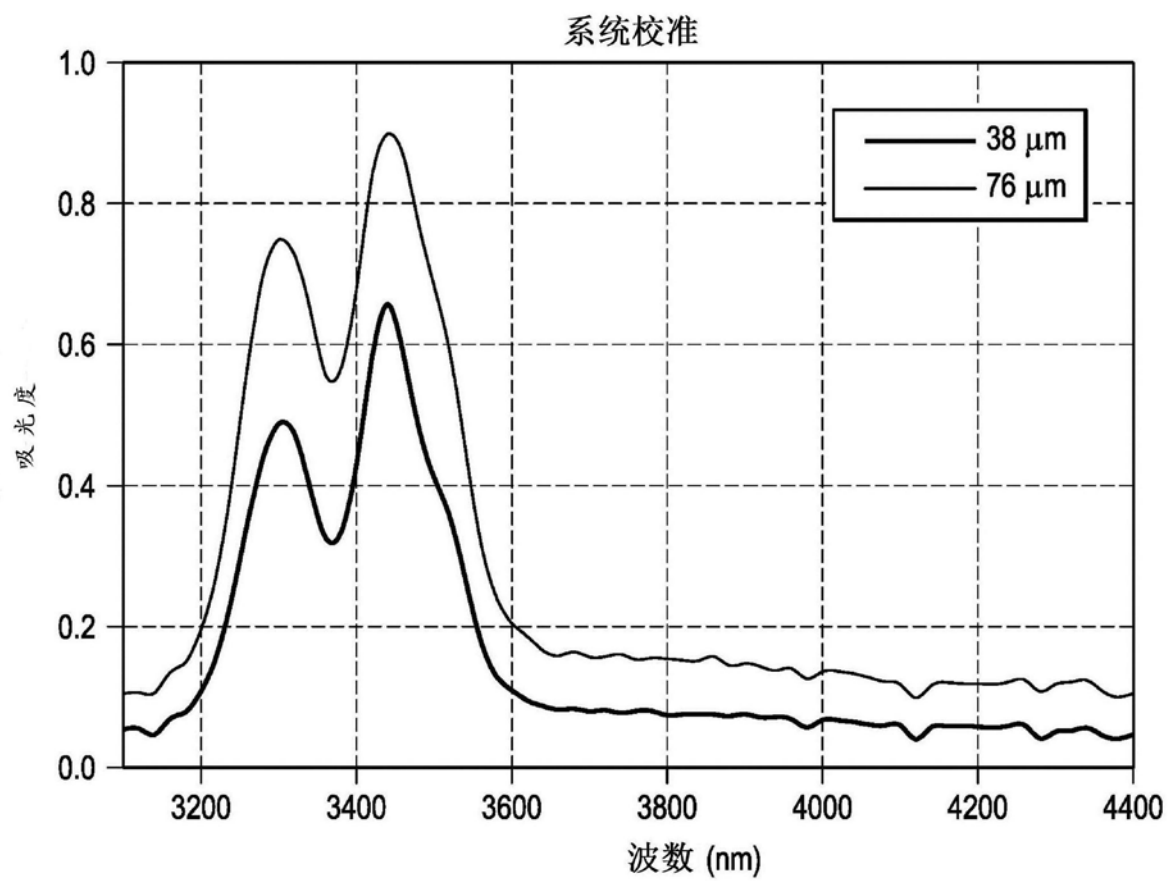


图28B

2900 ↘

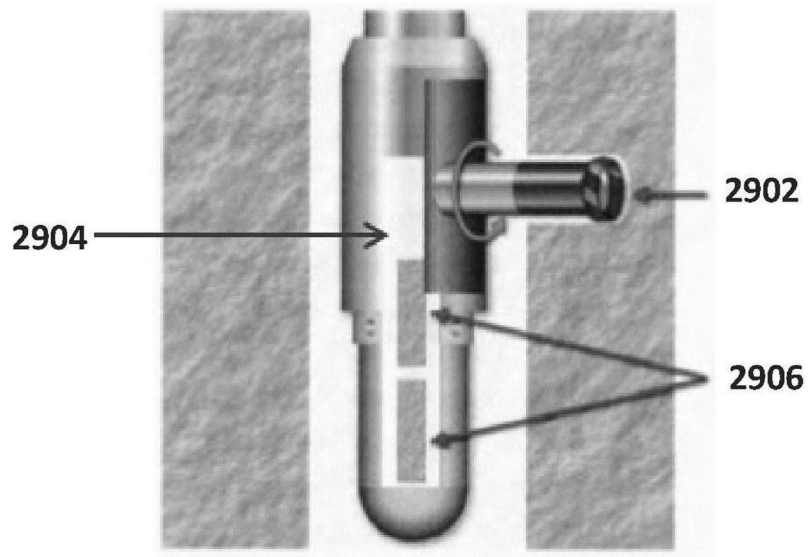


图29