



(12)发明专利

(10)授权公告号 CN 105283904 B

(45)授权公告日 2018.09.07

(21)申请号 201480031576.2

(22)申请日 2014.05.28

(65)同一申请的已公布的文献号
申请公布号 CN 105283904 A

(43)申请公布日 2016.01.27

(30)优先权数据
61/830620 2013.06.03 US

(85)PCT国际申请进入国家阶段日
2015.12.02

(86)PCT国际申请的申请数据
PCT/EP2014/061012 2014.05.28

(87)PCT国际申请的公布数据
W02014/195193 EN 2014.12.11

(73)专利权人 文塔纳医疗系统公司
地址 美国亚利桑那州

(72)发明人 J.布雷德诺 L.迪茨 J.F.马丁

(74)专利代理机构 中国专利代理(香港)有限公司 72001

代理人 马红梅 杜荔南

(51)Int.Cl.
G06T 7/11(2017.01)

(56)对比文件
WO 2009/006696 A1,2009.01.15,
JP 2012078716 A,2012.04.19,
Milan Gavrilovic 等.Blind Color
Decomposition of Histological Images.
《IEEE TRANSACTIONS ON MEDICAL IMAGING》
.2013,第32卷(第6期),摘要,第III节第A,B,C部分,
第IV节第A,C部分,第V节,第VI节.

Charles E.H.Berger 等.Color
deconvolution and support vector
machines.《International Workshop on
Computational Forensics》.2009,174-180.

审查员 郭妍妍

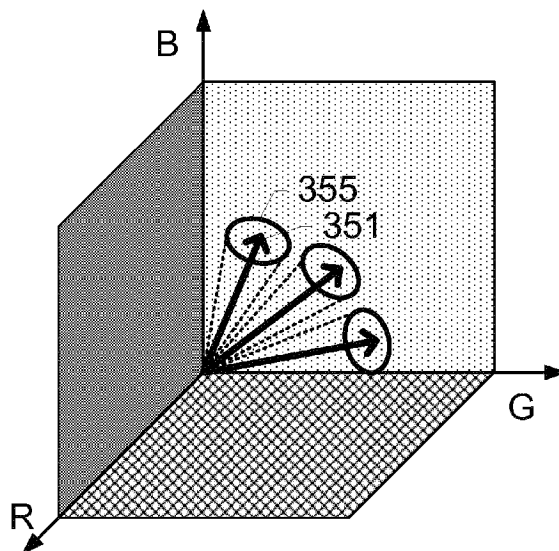
权利要求书2页 说明书13页 附图8页

(54)发明名称

图像自适应的生理上可信的颜色分离

(57)摘要

题述公开内容提出了用于通过基于特定于正在成像的化验物的信息自动且自适应地调节参考矢量来分离图像中的颜色的系统和方法,从而导致提供在物理和生理上可信的染色剂信息的经优化的解混过程。基于最小化非负颜色贡献、背景贡献、特定于背景的颜色通道中的高频或不想要的荧光、来自已知免疫组织化学标记物的信号以及被已知以承载生理上独立的信息的染色剂对,迭代地优化参考矢量。可以在基于测量来自多个输入图像的颜色而预先确定的范围内允许对参考矢量的调节。



1. 一种用于存储由处理器执行以执行包括以下各项的操作的计算机可执行指令的非暂时性计算机可读介质：

对从化验物提取的图像数据进行解混，该解混使用一个或多个初始参考矢量；

通过将多个规则应用于从图像数据的解混获得的一个或多个信号来确定质量度量；以及

基于质量度量来调节所述一个或多个初始参考矢量；

其中所述多个规则是基于化验物信息来选择的。

2. 根据权利要求1所述的计算机可读介质，其中所述操作还包括使用从所述一个或多个初始参考矢量的调节获得的一个或多个经优化的矢量来对图像数据进行解混。

3. 根据权利要求2所述的计算机可读介质，其中解混、确定和调节被迭代地重复直到质量度量通过阈值。

4. 根据权利要求1至3中任一项所述的计算机可读介质，其中化验物信息还包括染色剂强度。

5. 根据权利要求4所述的计算机可读介质，其中所述操作还包括基于染色剂强度来从多个参考矢量当中选择初始参考矢量。

6. 根据权利要求1至3中任一项所述的计算机可读介质，其中化验物信息还包括染色过程的过程参数。

7. 根据权利要求1至3中任一项所述的计算机可读介质，其中化验物信息还包括组织类型。

8. 一种用于优化参考矢量的方法，包括：

在耦合到工作站的处理器处接收包括图像数据的化验物信息，该图像数据包括从化验物信息中所标识的化验物检测到的多个信号；

由处理器基于所述多个信号中的一个或多个信号与存储在通信耦合到处理器的存储器上的多个规则中的一个或多个的相关来调节一个或多个参考矢量；以及

由处理器使用经调节的一个或多个参考矢量来对所述多个信号中的所述一个或多个信号进行解混。

9. 根据权利要求8所述的方法，其中多个规则包括最小化所述多个信号中的所述一个或多个信号中的负颜色贡献。

10. 根据权利要求8所述的方法，其中所述多个规则包括最小化不想要的背景颜色贡献。

11. 根据权利要求8所述的方法，其中所述多个规则包括最小化由不期望的元素导致的高频贡献。

12. 根据权利要求11所述的方法，其中不期望的元素包括来自未预期的小结构、强边缘、玻璃、嵌入材料或背景材料的荧光信号或色原信号。

13. 根据权利要求8至12中任一项所述的方法，还包括检测图像数据内的结构。

14. 根据权利要求13所述的方法，其中所述多个规则包括所述结构要具有所期望的大小或形状的要求。

15. 根据权利要求13所述的方法，其中所述多个规则包括所述结构要与另一结构协同定位的要求。

16. 根据权利要求8至12中任一项所述的方法,还包括通过预定义的搜索空间来调节初始参考矢量。

17. 一种用于存储由处理器执行以执行包括以下各项的操作的计算机可执行指令的非暂时性计算机可读介质:

生成从与包括染色剂的化验物相关联的图像数据导出的参考矢量空间内的颜色矢量的直方图,该染色剂与多个标记物和对应的多个参考矢量相关联;

比较所述直方图与针对化验物的已知直方图;以及

基于比较来确定是否要调节所述多个参考矢量中的一个;

其中一个或多个经调节的参考矢量被用于对图像数据进行解混。

18. 根据权利要求17所述的计算机可读介质,其中所述操作还包括基于所述直方图的多边形拟合。

19. 根据权利要求17或18所述的计算机可读介质,其中所述直方图包括大于图像所包括的通道数目的不同颜色矢量。

20. 根据权利要求17或18所述的计算机可读介质,其中所述操作还包括基于已知协同定位来分隔参考矢量空间的区段。

图像自适应的生理上可信的颜色分离

技术领域

[0001] 本题述公开内容涉及用于医学诊断的成像。更具体地,本题述公开内容涉及识别化验物内的生理上可信的染色剂矢量。

背景技术

[0002] 在诸如组织区段、血液、细胞培养等等之类的生物样品的分析中,利用一个或多个荧光团或色原对生物样品进行染色,并且随后对生物样品进行扫描或拍照以用于分析。观察从扫描生成的信号实现了多种过程,包括疾病的诊断、对治疗的响应的评价以及抗击疾病的新药物的开发。化验物包括生物样品,诸如利用包含与绑定到蛋白质、蛋白质片段或样品中的其它目标的抗体结合的荧光团或色原的染色剂而处理的来自人类主体的组织区段。当扫描化验物时,导出包括颜色通道的图像数据的多个通道,其中每一个所观察到的通道包括多个信号的混合物。

[0003] 一般地,颜色分离(或光谱解混(unmixing))被用于确定化验物的所观察到的一个或多个通道内的特定染色剂的浓度。这还可以被称为颜色去卷积。经扫描的图像的每一个像素由图像值的矢量或颜色矢量表示,并且每一个染色剂对应于还被称为参考光谱的参考矢量。染色剂的局部浓度由参考矢量的缩放因子表示。因此,针对包含具有不同浓度的多个协同定位的染色剂的像素的颜色矢量是所有目前染色剂的参考光谱的线性组合。典型地,荧光成像颜色通道直接提供图像矢量和参考光谱。在明场(透射)成像中,将由经染色的组织发射的光强度变换成光学密度空间,其中不同染色剂的混合由做出贡献的参考光谱的线性加权组合表示。

[0004] 解混过程使用对于标准类型的组织和染色剂组合而言众所周知的参考光谱来提取染色剂特定通道以确定各个染色剂的局部浓度。然而,用于纯染色剂的参考光谱往往随组织类型、染色期间的受控和不受控过程参数以及随年龄而变化。例如,总是存在基于组织年龄、染色剂年龄、组织被如何存储、脱水、固定、嵌入、切割等的组织类型内的变化。这些变化可能影响染色剂将看起来如何,并可能导致解混过程的结果中的不想要的伪像。现有方法不能在没有人指导的情况下处置这样的错误,并且,没有可靠的参考光谱可用于这样的变化。因此,不正确的分离和在生理或物理上不可信的结果继续出现。而且,关于包含3个颜色通道的明场图像,不能对大于3个染色剂的任何协同定位进行解混,或者没有明确的数学解存在。

发明内容

[0005] 题述公开内容提出了用于通过基于特定于正在成像的化验物的信息自动且自适应地调节参考矢量来分离图像中的颜色的系统和方法,从而导致提供在物理和生理上可信的染色剂信息的经优化的解混过程。迭代地对参考矢量进行优化,应用不受约束的颜色去卷积或解混,并且使所得到的颜色通道与基于关于化验物的信息而应用的多个规则相关。该多个规则包括例如最小化负面颜色贡献、背景贡献、特定于背景的颜色通道中的高频或

不想要的荧光、来自已知免疫组织化学标记物的信号以及被已知以承载生理上独立的信息的染色剂对。该相关可以用于确定结果的总体质量。如果质量是不可接受的,则可以调节参考矢量,并且可以利用经调节的参考矢量迭代地解混颜色通道,直到满足所有规则并且获得具有可接受质量的结果。可以在基于测量来自多个输入图像的染色剂参考矢量预先确定的范围内允许对参考矢量的调节。

[0006] 在附加实施例中,可以不仅针对特定图像而且针对特定图像内的指定区(包括各个像素)确定经优化的参考矢量。参考光谱调节和基于规则的对应质量度量可以在空间上遍及图像而变化。另外,可以从图像或图像的区确定染色剂的强度,并且可以在解混之前对应地调节参考矢量。此外,可以使用化验物信息来解混包括三个通道但具有多于3种染色剂的明场图像。可以将图像中的包括三个或更多颜色的光学密度空间分区成预定义的协同定位系统。化验物信息包括与手头上的化验物、利用化验物而染色的生物学结构、染色剂和复染剂的类型和数目及其线性混合物有关的细节、以及诸如化验物的年龄、组织类型和年龄等之类的附加元数据。

附图说明

[0007] 图1示出根据本题述公开内容的示例性实施例的用于优化参考矢量的系统。

[0008] 图2示出根据本题述公开内容的示例性实施例的用于优化参考矢量的方法。

[0009] 图3A和3B示出根据本题述公开内容的示例性实施例的光学密度空间中的参考矢量的调节。

[0010] 图4A-4D示出根据本题述公开内容的示例性实施例的被分区成针对协同定位的染色剂的区段的光学密度空间。

[0011] 图5示出根据题述公开内容的示例性实施例的基于协同定位的染色剂的层级而被分区成区段的光学密度图。

[0012] 图6A-6B示出根据题述公开内容的示例性实施例的图像数据的直方图。

[0013] 图7A-7B示出根据题述公开内容的示例性实施例的使用图像数据的直方图上的多边形拟合的参考矢量的优化。

具体实施方式

[0014] 所公开的系统和方法处理图像以使用迭代优化的参考矢量来分离或“解混”图像的分量信号。使来自化验物的图像数据与特定于化验物特性的预期或理想结果相关以确定质量度量。在与理想结果成对照的低质量图像或较差相关性的情况下,调节一个或多个参考矢量,并且使用经调节的参考矢量来迭代地重复解混,直到相关性示出与生理学和解剖学要求匹配的良好质量图像。解剖学、生理学和化验物信息可以用于定义被应用于所测量的图像数据以确定质量度量的规则。该信息包括:组织被如何染色;意图/不意图染色组织内的什么结构;以及特定于正在处理的化验物的结构、染色剂和标记物之间的关系。迭代过程导致可生成下述图像的染色剂特定矢量:该图像精确地标识感兴趣的结构和生物上相关的信息,没有任何噪声或不想要的光谱,且因此适合用于分析。将参考矢量调节到搜索空间内。搜索空间定义了参考矢量可取得以表示染色剂的值的范围。可以通过扫描包括已知或普遍发生的问题的多种代表性训练化验物并确定针对训练化验物的参考矢量的高质量集

合来确定搜索空间。

[0015] 在以下描述内,对光谱解混、颜色去卷积和颜色分离的任何引用是同义的,且涉及从包含这些染色剂的混合物的图像数据获得图像中的染色剂的局部浓度或量的过程。该混合物最常被假定为是线性的,并且利用线性解混方法。然而,非线性方法也可以被应用以执行解混。对于明场成像,化验物导致从光源透射通过经染色的样品的光的吸收和衰减。一种用于解混该图像数据的示例性方法利用光学密度空间,其中参考颜色和图像数据被转换成光衰减信号,例如使用朗伯-比尔(Lambert-Beer)定律。如本文所使用的图像数据和混合物模型的分析(即,图像数据和染色剂效果的数学描述)不必受限于已知方法,而是可以包括组织样本内的光散射的附加模型,包括可能衍射的建模,以及用于确定由检测器(例如荧光显微镜中的明场显微镜中的CCD或CMOS传感器,或者整个载玻片扫描仪)检测的光如何与在化验物中吸收或由化验物发射的光的量相关的其它方法。

[0016] 对于以下描述,可以假定,跨各图最为对应地加标签的结构(例如132和232等)拥有相同的特性并服从相同的结构和功能。如果在对应地加标签的元件之间存在未指出的差异,并且该差异导致针对特定实施例的元件的非对应结构或功能,则针对该特定实施例而给出的该冲突描述应当起支配作用。

[0017] 图1示出根据本题述公开内容的示例性实施例的用于优化参考矢量的系统。系统100包括用于生成化验物信息的源101。例如,源101可以是光谱相机、扫描仪中的CCD或CMOS传感器、明场显微镜、荧光显微镜、或用于对包括诸如利用一个或多个荧光或色原染色剂和标记物染色的生物样品之类的材料样本的化验物进行成像的整个载玻片扫描仪。源101与存储器110通信,存储器110包括由耦合到计算机120的处理器125执行的多个处理模块或逻辑指令。例如,诸如生物样品之类的样本可以被安装在载玻片或其它衬底或设备上以用于通过源101进行成像的目的,其中依照本公开内容,样本的图像的分析通过处理器125执行存储在存储器110上的多个模块中的一个或多个而执行。分析可以出于样本的识别和研究的目的。例如,生物学或病理学系统可以针对其解剖学结构、以及指示癌症或其它疾病的细胞、蛋白质、蛋白质片段或其它标记物的存在和组织构成、或者出于其它目的(例如基因组DNA检测、信使RNA检测、蛋白质检测、病毒的检测、基因的检测或其它)而分析样本。

[0018] 样本可以借助于包含一个或多个不同标记物、荧光团或色原染色剂的染色剂的应用而染色。荧光团可以包括一个或多个纳米晶半导体荧光团(即,量子点),每一个产生不同波长范围中的峰值发光响应。量子点是众所周知的,且可以从Invitrogen Corp.、Evident Technologies等等商业上可得到。被应用于样本的一个或多个荧光团可以是有机荧光团14(例如DAPI、德克萨斯红),其在本领域中是众所周知的,且在至少共同所有和转让的美国专利8,290,236中描述,该美国专利的内容通过引用以其整体并入本文。色原染色剂可以包括苏木精、曙红、固红或3,3'-二氨基联苯胺(DAB)。因此,系统100可以与利用仅量子点、利用量子点与常规有机荧光团相组合、仅常规有机荧光团、色原染色剂或者任何其它组合或染色剂和标记物而染色的样本一起使用。而且,在向样本应用染色剂的自动化染色/化验平台中处理典型的样本。在市场上存在适合用作染色/化验平台的多种商用产品,一个示例是受让人文塔纳医疗系统有限公司的Discovery.TM产品。例如,在荧光成像过程中,在初步组织处理和染色之后,将样本供应到包括光谱源(例如用于以意图产生来自被应用于样品的荧光团的发光响应的波长光照样本的光源)的相机系统。在量子点的情况下,光源可以是宽光

谱光源。可替换地,光源可以包括诸如激光器之类的窄带光源。相机平台还可以包括明场显微镜,一个示例是受让人文塔纳医疗系统公司的VENTANA iScan HT产品或者具有一个或多个物镜和数字成像器以及一组光谱过滤器的任何显微镜。可以使用用于在不同波长处捕获图像的其它技术。适合用于对经染色的生物样品进行成像的另外的相机平台在本领域中是已知的且从诸如Zeiss, Canon, Applied Spectral Imaging等等之类的公司商业上可得到,并且这样的平台可容易适配以供本题述公开内容的系统、方法和装置中使用。

[0019] 可以将经由源101从化验物获取的信息(包括颜色通道、强度和任何附加元数据)经由显微镜101与计算机120之间的线缆连接、经由计算机网络或使用常用于在计算机之间传送数字信息的任何其它介质供应到计算机可读介质110。化验物信息还可以通过网络而供应到网络服务器或数据库以用于存储和供计算机120稍后检索。除处理器125和存储器110外,计算机120还包括用户输入和输出设备,诸如键盘、鼠标、触控笔和显示器/触摸屏。如在以下讨论中将解释的,处理器125执行存储在存储器110上的逻辑指令,从而执行化验物信息的分析,执行一个或多个解混操作,检测图像中的结构,定量分析,以及向用户操作计算机120显示定量/图形结果。

[0020] 例如,如以上所描述的,在源101处扫描化验物以生成包括若干颜色通道的混合物的图像数据。例如,图像数据可以包括发射光谱、吸收光谱、荧光或由化验物包括的任何其它信号。图像数据还可以包括标准红色、绿色和蓝色通道。在源101是检测透射通过化验物的白光的明场显微镜的情况下,图像数据可以包括具有不同波长范围的多个通道或者也包括标准红色、绿色和蓝色通道。可以包括任何数目的分离颜色通道。在源101是荧光显微镜的情况下,图像数据可以包括量子点(Q点)通道以及用于染色剂和复染剂的通道。所成像的通道及其波长范围可以针对一般使用而选择或者被调节到手头上的化验物和组织类型。例如,解混对于专用显微镜(例如具有各种过滤器、激励光波长、明场光源波长等的显微镜)而言是可能的,以及对于标准成像设置(例如红绿蓝相机和白光源)而言是可能的。图像数据连同附加化验物信息一起由化验物信息提取模块111提取和解析。附加化验物信息可以包括染色剂标识、染色的过程参数(例如孵育(incubation)时间和试剂的浓度)、组织类型和其它物理或生理信息。附加化验物信息可以存储在由信息提取模块111接收且例如由用户、连接到系统100的实验室信息系统提供或者从粘附到承载了生物样品的载玻片的条形码读取的一个或多个数据分组的元数据中。鉴于本公开内容,将化验物信息提供到系统100中的其它方法将变得对本领域普通技术人员而言显而易见。

[0021] 解混模块112可以被调用以解混图像数据、图像数据的所选部分或隔离于图像数据的特定区的信号的混合物(例如,线性的混合物或基本上线性的混合物)以获得染色剂特定矢量。出于本公开内容的目的,解混与光谱解混、颜色去卷积和颜色分离同义。然而,可以使用用于分离混合物的任何其它已知或将来的方法。在示例性实施例中,使用线性最小二乘法。例如,基于化验物信息,解混模块112可以利用已知的参考光谱,以解混特定像素中的信号的混合物以便获得对应于该像素中的染色剂或结构的分量信号或矢量。解混模块112可以从存储器110上的参考光谱数据库(诸如数据库118)检索一个或多个已知参考矢量。例如,线性光谱解混过程可以获得对应于被协同定位在单个像素上的一个或多个染色剂的矢量的线性组合,其中每一个矢量是通过其强度或浓度来加权的。针对每一个标记物的用于对组合进行解混的参考矢量通过本文所描述的依赖于化验物的相关和分析而迭代

地优化。

[0022] 负抑制模块113被用于识别由解混过程产生的任何负矢量。负值的存在指示像素可以包含具有至少一种染色剂的混合物,所述至少一种染色剂具有负浓度,这在物理上是不可信的。负抑制模块113使用负值的存在来推断解混结果具有由于一个或多个不正确的参考矢量所致的较低质量。尽管现有方法已经在解混过程期间使用非负约束,其迫使所有结果在数学上为正,但是这些方法在很大程度上忽视了负值被返回这一事实。取而代之,负抑制模块113辨别错误,并可以调节质量度量以指示这样的解剖学或物理上的不可信性。这触发了参考矢量的自动调节或优化以再次解混图像数据。解混可以从负值的辨别触发的,或者解混可以被延迟直到其它模块被处理以确定质量度量。

[0023] 染色剂强度确定模块114执行下述逻辑操作:根据图像数据(例如整个图像、图像中的区或图像中的各个像素)确定染色剂的强度,并且在对图像数据进行解混之前适当选择或调节参考矢量。染色剂强度确定可以简单地从图像数据所包括的颜色矢量执行,而不需要任何具体化验物信息。例如,可以处理来自明场显微镜的信号以比较源光强度与已经经过组织的所检测到的光强度以在颜色和染色剂的任何分离之前指示化验物、区或图像中的个体像素上的一个或多个染色剂的强度。可以处理荧光图像以确定从一片组织发射的总强度或总体亮度,使得能够确定平均染色剂强度而不要求染色剂、颜色、色调和其它贡献的分离。在两种情况中的任一种情况下,给定针对图像或像素的染色的平均或总体强度,与染色剂强度相关联的参考矢量的预定义集合可以用于对图像数据进行解混。例如,可以针对非常浅、浅、中度、强和非常强的染色剂强度而预定义初始参考矢量的不同集合。染色剂强度确定模块114将所测量的强度分类成这些类别中的一个,并选择初始参考矢量或光谱的适当集合。

[0024] 结构确定模块115识别了图像数据内的结构,并使这些结构与结构和/或染色剂的已知组合相关以识别和消除已知或明显的不可信性。例如,可能已知的是,两个特定量子点不能够在某个样本材料中共存。不一致、不现实或不可能的一个或多个信号可以由结构确定模块115来辨别,并通过最小化或消除违犯信号而被计及。已知的不一致、不现实或不可能的一个或多个信号可以从数据库118或者与系统或系统的有技能的操作者(诸如病理学家或知识丰富的技术人员)通信的任何其它数据储存器检索的。结构可以通过解析图像数据或经解混的结果而确定以辨别特定大小、形状或颜色的结构。例如,可以辨别利用棕色染料而染色的小圆细胞,并且可以将它们的分量矢量与理想结果进行比较以确定它们的存在和/或结构是否可信。另外,可以识别和消除残余染色剂的量。例如,可以将不对应于小圆形状的任何棕色染色剂识别为噪声,或者仅仅不想要的信号。此外,结构确定模块115最小化或消除由不期望的元素导致的高频贡献,诸如来自未预期的小结构、强边缘、玻璃和化验物的其它部分、嵌入材料或背景材料的荧光或色原信号。

[0025] 背景确定模块116可以基于与背景信号相关联的光谱签名或由包括至少一个背景信号的信号混合物包括的像素的位置来辨别信号混合物内的一个或多个背景信号。背景信号可以通过其唯一签名(例如与载玻片的玻璃、用于盖玻片的胶、组织的自发荧光相关联)和经过图像的普遍存在的色散而辨别。图像的某些区可以被确定为主导地或仅仅包含诸如自发荧光等之类的信号。当确定具有与背景相关联的签名(例如宽带签名)的分量信号时,可以将分量信号与特定于正在分析的样本材料的已知背景签名进行比较。例如,诸如用于

解剖或临床病理学的图像分析的系统可以将组织样本的经扫描的载玻片与包含具有已知背景签名的类似组织样本的校准载玻片的图像进行比较,以识别经扫描的图像中的背景信号。数据库118可以包括已知签名。可以将已知背景签名与图像的区进行比较以主导地辨别所述区内的宽带信号。对于由背景、玻璃和/或全局结构(即,遍及图像而存在的结构)引起的信号,小结构或高频图像内容(例如精细纹理)在经解混的图像中的存在指示针对该信号的不正确的参考矢量。这样的小结构的存在和强度是对质量度量的负面贡献。因此,通过检测具有正确参考矢量的信号以及基于来自并非背景的结构(例如,组织中的细胞)的阴影识别不正确的参考矢量,来从解混结果去除不想要的信号。这指示了针对这样的背景信号的参考矢量需要调节。

[0026] 调节模块117基于来自每一个模块111-116的结果来迭代地调节数据库118中的输入或初始参考矢量。另外,每一个模块的结果可以被组合以生成质量度量。例如,结构确定模块可以指示结构的不可信的组合,导致质量的负面指示或较低质量度量。可替换地或此外,在由处理器125执行的解混过程之后,经解混的染色剂特定图像的互相关可能导致较低质量度量。模块之间的任何冲突,诸如提供在每一个染色剂特定通道中个体可信的结果但其中结果在不同染色剂特定通道之间冲突的解混过程,类似地指示低质量结果。例如,可能已知的是,两个染色剂对于手头上的组织类型而言相互排斥,使得相同位置中这些染色剂的高浓度不是可信的。所得到的质量度量可以触发初始参考矢量在其搜索空间或所允许的范围内的调节以及重复的解混过程。经解混的结果可以再次由每一个模块评价,其中结果用于生成新的质量度量。可以将质量度量与针对理想化验物的已知质量度量进行比较,并且当确定质量度量充分接近于理想化验物时,调节模块117可以停止调节参考矢量,并指示经解混的信号接近于理想。多种优化策略(例如单纯形下降优化策略(即,通过自动调节参考矢量并迭代地检查所得到的质量度量最大化质量度量的策略))可以被采用以便以增加质量度量的方式调节参考矢量。数据库118可以利用与特定化验物信息相关联的字段中的新的经优化的参考矢量而更新。另外,该操作序列可以由方法迭代地执行,如图2中所示。

[0027] 另外,质量确定可以特定于用户查询。例如,用户输入可以将解混或结构确定过程隔离于搜索特定颜色通道中的单元,或者,可以提交请求特定结构的质量的查询。结构确定和解混模块可以处理服从查询要求的图像数据,并且调节模块117生成适当的质量度量。可以在定义参考矢量可被改变多少和可在哪个方向上被改变的针对每一个参考矢量的搜索空间内调节参考矢量。搜索空间可以是预定义且固定的。可以收集和分析具有来自不同图像的已知参考矢量的训练数据以提供初始或缺省参考矢量,连同用于定义搜索空间的参考矢量的所允许的变化的范围。在一些示例性实施例中,主成分分析(即,将训练数据的均值识别为针对参考矢量和方向(诸如本征矢量)以及距离(诸如本征值)的初始值的分析,其中这些在优化期间可以被修改)可以用于根据训练示例确定有效搜索空间。

[0028] 另外,诸如调节经解混的图像数据中的染色剂浓度的最小值或最大值之类的其它改善操作可以被应用以突出显示特定范围并消除该范围外的信号。可以调节由信号的经解混的集合产生的图像以用于对照来查看更加动态的范围。例如,在光谱解混之后获得的数据在其动态范围方面可能具有不足的分辨率,并且因此,亮度或对比度调节(其人为地增加了针对经解混的通道的图像内容的动态范围)可以使得在视觉上更容易感知到经解混的通道在图像中的不同像素处有多强。这样的调节使得能够研究来自经解混的通道的输出并改

进图像理解。可以执行其它成像操作,其中在计算机120的显示器上描绘任何所得到的图像以及用于执行和操纵存储在存储器110中的模块的接口。

[0029] 如以上所描述的,模块包括由处理器125执行的逻辑。如本文中和遍及本公开内容所使用的,“逻辑”是指具有可被应用以影响处理器的操作的指令信号和/或数据的形式任何信息。软件是这样的逻辑的一个示例。处理器的示例是计算机处理器(处理单元)、微处理器、数字信号处理器、控制器和微控制器等。逻辑可以由存储在诸如存储器110之类的计算机可读介质上的信号形成,所述计算机可读介质在示例性实施例中可以是随机存取存储器(RAM)、只读存储器(ROM)、可擦除/电可擦除可编程只读存储器(EPROM/EEPROM)、闪存存储器等。逻辑还可以包括数字和/或模拟硬件电路,例如包括逻辑AND(与)、OR(或)、XOR(异或)、NAND(与非)、NOR(或非)和其它逻辑运算的硬件电路。逻辑可以由软件和硬件的组合形成。在网络上,逻辑可以在服务器或服务器的综合体上被编程。特定逻辑单元不限于网络上的单个逻辑位置。

[0030] 图2示出根据本题述公开内容的示例性实施例的用于优化参考矢量的方法。图2的方法可以由与图1中所描绘的那些计算机执行模块类似的计算机执行模块执行。该方法以从包括扫描仪或光谱相机或者与其相关联的诸如源101之类的源或者可捕获一频率范围处的图像内容的任何源接收样本的图像(S230)而开始。样本可以借助于由相机捕获的图像以及由例如光源光照的包含一个或多个不同荧光团或色原染色剂的染色剂的应用而染色,如以上所描述的那样。将图像供应到执行存储在存储器上的逻辑指令以用于执行示例性方法中所描述的操作的计算机。

[0031] 化验物信息包括图像数据以及附加化验物信息。图像数据包括由如以上详述的源在具有不同波长范围的多个通道中检测的多个颜色通道。提取图像数据连同附加化验物信息并且将其提供给解混模块以进行解混(S231),以便获得由图像数据包括的分量信号或染色剂特定矢量。解混利用基于化验物信息从参考数据库219检索的已知参考光谱/矢量。在其它实施例中,可以在解混之前迭代地优化参考矢量。例如,在解混之前,可以针对总体图像、针对图像中的区、或各个像素调用强度确定(S233)以确定最优参考光谱。染色剂强度确定(S233)可以被简单地从图像数据所包括的颜色矢量执行,而不需要任何特定化验物信息。在给定来自S233的针对图像、图像中的区或像素的染色剂强度的情况下,可以选择与染色剂强度相关联的参考矢量的预定义或经调节的集合(S241)以用于对图像数据进行解混。强度确定(S233)的结果可以用于确定(例如,通过使参考矢量移位以更接近地对应于实际强度值)新的参考光谱(S241),或者该方法可以进行到负抑制和背景确定(S235)。识别由解混过程产生的任何负矢量,并且做出解混结果具有由于不正确的参考矢量所致的较低质量的确定。

[0032] 再一次,该方法可以进行到自动调节质量度量以指示这样的解剖学或物理上的不可信性,选择经调节或经优化的参考矢量(S241)以再次对混合物进行解混,或者该方法可以进行到生理上可信的结构的确定的确定(S237)。该步骤识别图像数据内的结构,例如通过解析图像数据或经解混的结果以辨别特定大小、形状或颜色的结构。感兴趣的结构可以由用户输入到系统中,或者从化验物信息提取。可以使这些结构与存储在染色剂数据库221中的染色剂和结构的已知组合相关以识别和消除已知或明显的不可信性。例如,可能已知的是,两个特定标记物不能够在某个样本材料中共存或者仅能够出现在已知形状和大小的结构中。

可以通过最小化或消除不一致或不现实的信号来辨别和计及这样的“不可能”。例如,RGB明场图像提供针对利用蓝色染料、棕色染料和红色染料而染色的化验物的红色、绿色和蓝色强度。当针对每一种染料或染色剂对通道进行解混时,可以通过使经解混的结果与关于化验物而已知的信息相关来确定结构的存在。基于例如小圆细胞利用棕色染料而染色、没有特定形状的较大细胞(诸如巨噬细胞)利用红色染料而染色并且背景和所有细胞核被染色成蓝色这一知识,指示被着色成红色的小圆结构的任何信号是生理上不可信性。类似地,任何大棕色区是不可信的,因为其不与在化验物信息中所标识的关于特定染色剂/结构组合而已知的内容相关。

[0033] 因此,图像的质量受到损害,从而触发对用于对图像进行解混的参考矢量的调节(S241)或除相关步骤的所有结果外还影响质量度量(S239)。质量度量可以附加地比较解混结果的质量与在给定化验物类型的情况下的已知或理想结果,并且,做出是要调节参考矢量(S241)并重复过程还是要在质量可接受的情况下结束循环的确定。可以基于预定义的阈值来确定可接受的质量。当确定结果质量良好或者所得到的矢量在生理上可信时,可以调节针对矢量的搜索空间(S240)以包括可接受的质量度量,并且将该搜索空间存储在与参考矢量数据库219相关联的空间数据库222中。如果质量度量不可接受或者处于阈值以下,则可以在定义参考矢量可被改变多少和可在哪个方向上被改变的针对每一个参考矢量的搜索空间内调节参考矢量(S241)。

[0034] 这些步骤可以被迭代地执行,其中,调节数据库219中的初始参考矢量,并且对图像数据进行解混(S231),并且生成新的质量度量,直到确定质量度量充分接近于理想化验物或者确定不可能有进一步的改进。当确定最优质量时,该方法可以基于化验物信息调节针对矢量的搜索空间(S240)以包括或突出显示当前配置。搜索空间可以随对图像数据的每一个应用而动态地扩张或收缩,从而使得能够连续地训练系统。

[0035] 图3A和3B示出根据本题述公开内容的示例性实施例的光学密度空间中的参考矢量的调节。在将本题述公开内容应用于明场图像时使用光学密度空间。光学密度是与在透射成像中所吸收的光的量相对应的染色剂或颜色的性质。对于荧光图像,该空间被称为参考矢量空间并反映由扫描仪获取的不同波长范围中的荧光团或量子点的发射。如本文所描述的,可以基于化验物信息来在预定义的搜索空间内调节或优化每一个参考矢量。图3中所描绘的光学密度空间为了易于可视化而包括三个颜色。出于本公开内容的目的,图像可以包括任何数目的颜色,其中其对应光学密度空间被分区成预定义的协同定位系统,如本实施例中所示。

[0036] 图3A示出光学密度空间345中的像素350的值,连同其对应的参考矢量351、352和353,该参考矢量可以被线性地混合以获得像素值。在所提出的示例中,光学密度空间345包括基于图像的三个颜色(即蓝色346、绿色347和红色348)的三个轴。将像素350解混成其分量矢量351、352和353。例如,像素的分量矢量指示蓝色矢量351、红色矢量352和棕色矢量353,每一个矢量具有不同的方向和强度,该不同的方向和强度使得这三个矢量的混合物能够得出像素350。

[0037] 图3B示出每一个分量参考矢量如何在其搜索空间355内可调节。可以将搜索空间355预定义成涵盖多个已知解剖学、化学和生物学假设和变化。搜索空间355可以是参考矢量351的所允许的改变的预定义范围,其中该范围基于如本文所描述的训练数据。在一个示

例性实施例中,可以针对每一个参考矢量准许百分之十的变化,并且可以在这样的百分之十的变化内迭代地调节参考矢量直到实现令人满意的结果。在附加实施例中,可以解混具有纯的且经隔离的染色剂的若干训练化验物,并且结果中的变化用于估计平均参考光谱连同平均变化以确定搜索空间355。

[0038] 如本文所描述的,利用三色图像示出在这些实施例中描绘的光学密度空间,但是本文所描述的原理可以被应用于来自明场和荧光成像器的多维和多通道图像。对于包括三个颜色通道的较简单的图像,例如由明场显微镜扫描的RGB图像,可以利用已知方法解混任何单个染色剂或多达三种染色剂的协同定位。本文所描述的实施例提供了用于解混利用大于图像数据中的维度或颜色的数目的染色剂而染色的任何类型的图像(包括RGB图像)的手段。化验物特定信息和规则可以被应用以将任何光学密度空间或参考矢量空间分区成预定义的协同定位系统。这些系统是基于手头上的化验物、利用化验物而染色的生物医学结构以及这些结构在如图像像素那样小的区中的共存来定义的。基于化验物包括什么染色剂、结构和标记物的在先知识以及已知可能协同定位的层级,可以对参考矢量空间(例如光学密度空间)进行分区以确定生理上可信的协同定位。这样的图形表示可以用于优化用于对图像或像素进行解混的参考矢量。例如,特定区可以单独专用于光学密度空间的一个隔间或区,并且因而,该区的解混将很可能导致纯染色剂矢量。在另一示例中,复染剂苏木精或DAPI可以用于将每个细胞核染色成蓝色,并可以仅与特定染色剂或标记物协同定位,导致使用适当参考矢量的解混过程。在解混之前,可以类似地评价包括附加染色剂的光学密度空间的其它区域并将其与已知协同定位进行比较,从而通过排除所有不太可能的协同定位来导致更高效且精确的解混过程。

[0039] 图4A-4D示出根据本题述公开内容的示例性实施例的被分区成协同定位染色剂的区段的光学密度空间。图4A中描绘的3维光学密度空间445包括当被混合时产生像素值450的两个染色剂特定矢量451和453。在该光学密度空间中,相同的两个染色剂的任何可能混合物将必然落在由两个染色剂参考矢量横跨的圆形区段内。另外,三个或更多染色剂的任何可能混合物将落到由三个或更多参考矢量横跨的金字塔形或圆锥形区中。不论每一个染色剂中的多少被使用,这些特定染色剂的任何组合将总是落到光学密度或参考矢量空间的有限区中。该原理可以用于预测或确定矢量的组合的任何可能值的位置。例如,将另一染色剂添加到两个染色剂的系统可以导致金字塔形或圆锥形区以及三个染色剂的所有可能变化将落在金字塔体或圆锥体内的逻辑确定。类似地,添加第四染色剂将导致具有四边形底座的金字塔体,其中该金字塔体必然包括四个染色剂的所有可能组合。

[0040] 而且,可以使光学密度空间变平成2维表示460,如图4B中所示。在该平面表示中,两个染色剂451和453的所有可能混合物落在将染色剂相连接的区461内。图4C和4D分别示出涵盖被包含在单个图像信号或像素中的三个染色剂451、453和454以及四个染色剂451、453、454和456的任何可能组合或协同定位的区461。尽管针对源自RGB相机的三个颜色通道描述该示例性实施例,但是本文所描述的原理适用于可能不能够在视觉上描绘但可由计算机处理的任何组合中的任何类型的图像数据,诸如多光谱图像、荧光图像数据等,如本文所描述的那样。

[0041] 用于分隔光学密度空间的区段的应用偏好基于针对特定化验物的已知可能的或生理上可信的协同定位。这些已知协同定位包括一般生物学知识以及什么生物标记物作为

手头上的化验物的目标的特定知识,并可以通过优先级或者重要性或可能性的层级而排序。例如,苏木精复染剂对样本中的所有细胞核进行染色,并典型地独自出现而没有与其它信号协同定位。另一标记物Ki67仅对增殖的癌症细胞的核进行染色。Ki67必然连同苏木精一起出现,因此独立的Ki67标记物的任何出现将被视为在生理上不可信,且因此被忽略或抑制。通过迭代地将光学密度空间与作为手头上的化验物的目标的生物标记物的这样的列表和与这些生物标记物相关联的染色剂的参考矢量进行比较,可以使用仅导致生理上可信的矢量的特定矢量来解混图像数据。

[0042]

#	混合系统	描述
1	仅复染剂	复染剂(大多数为苏木精)对组织中的每个细胞的细胞核进行染色。其可以独自出现而没有协同定位(且频繁地这样做)
2	复染剂 + Ki67/染色剂 1	Ki67是针对增殖的肿瘤细胞的核的标记物。该标记物是利用染色剂1来染色的,并且其仅可以与标记所有细胞核的复染剂一起出现。
3	仅CD20/染色剂 2	CD20标记B细胞的细胞膜。染色剂2可以在没有协同定位的情况下出现在这些细胞的膜中。
4	仅CD3/染色剂3	CD3标记T细胞的细胞膜。染色剂3也可以在没有协同定位的情况下出现。没有细胞可以同时为T细胞和B细胞,因此CD3和CD20不能协同定位。
5	CD3/染色剂3 + CD8/染色剂4	一些CD3阳性T细胞此外是CD8阳性的。这两种标记物位于T细胞的细胞膜上,并且每个CD8阳性细胞也是CD3阳性的。染色剂4可以因此仅与染色剂3协同定位出现。
6	复染剂 + CD20/染色剂 2	当在显微镜下对B细胞进行成像时,细胞膜和细胞核频繁地一起处于一个像素中——它们不能位于“彼此顶上”。这些像素包含复染剂和染色剂2的混合物
7	复染剂 + CD3/染色剂 3	同样情况在对T细胞进行成像时发生——利用染色剂3而染色的细胞膜和利用复染剂而染色的核可以出现在相同像素中
8	复染剂 + CD3/染色剂 3 + CD8 / 染色剂 4	最后,CD3阳性且CD8阳性的T细胞的膜也可以与该细胞的核一起被成像。

[0043] 表1。

[0044] 表1示出按针对一个示例化验物的可能性的次序布置的协同定位系统的一个示例和可能实现。题述公开内容的示例性实施例引用该层级以在解混之前或之后消除不可信的协同定位。该列表仅仅是示例性的而不是详尽的——基于特定化验物的规则的许多附加组合可以由本领域普通技术人员鉴于本公开内容而设想。

[0045] 图5示出被划分成协同定位染色剂的区段的光学密度空间560,该区段是基于涉及化验物中的染色剂的规则或偏好的层级来定义的。类似于本文所描述的其它实施例,调用特定于可用化验物信息的规则以创建和定义具有协同定位染色剂的光学密度空间的区段。规则可以具有如以上所描述的层级的形式。区段可以被分离地解混以基于化验物信息和偏

好来标识染色剂的特定组合的协同定位。

[0046] 每一个区段利用多达3个参考染色剂的系统而解混。图像中的染色剂包括551、553、554和556。例如，区A、B、C和D排他性地对应于染色剂551、553、554和556中的每一个，而没有任何协同定位。区E标识染色剂554和556的任何可能的协同定位。区F标识染色剂554和551的任何可能的协同定位。区G标识染色剂554和553的任何可能的协同定位。区H标识染色剂553和551的任何可能的协同定位。区J标识染色剂551、554和556的任何可能的协同定位。最后，区K标识染色剂554、553和551的任何可能的协同定位。基于规则集合，进一步确定没有附加染色剂协同定位在生理上可信。而且，某些区比其它区更有可能，如其重叠所示。例如，由区F表示的染色剂554和551的协同定位比由区G表示的染色剂554和553的协同定位更可能出现（且因此重叠）。区K最不可能出现且因此被所有其它区重叠。这使得能够解混具有大于颜色通道数目的数目的染色剂的图像。区的重叠和与染色剂协同定位相关联的区段的大小反映了某些协同定位出现的先验概率。当利用该方法进行的解混被应用于组织的图像时，这些协同定位及其在图像中的结构的后验概率形成重要输入以对经染色的组织进行诊断。

[0047] 图6A-6B示出根据题述公开内容的示例性实施例的用于确定化验物上的协同定位染色剂的直方图，即，对染色剂参考矢量的调节以获得针对新图像数据的高质量解混结果。在该实施例中，生成表示图像数据的直方图和/或直方图数据。将该所生成的直方图数据与包括例如目标组织类型、生物标记物集合和关联的色原染色剂或荧光团的针对化验物的预期直方图数据进行比较。基于该比较，可以调节参考矢量，使得预期直方图数据更接近地对应于表示图像数据的直方图数据。例如，如果直方图数据中的许多像素落到表示纯红色的组区(bin)中，但预期直方图数据不具有这些组区中的条目，则可以通过将针对红色染色剂的参考矢量修改为更纯来修正该失配。

[0048] 针对不同的图像数据类创建多个类别或组区。这些类可以包括例如染色剂的色调。例如，在不需要关于总吸收率或每一个染色剂的总强度的任何信息的情况下，可以向特定组区分配针对每一个经染色的像素的颜色或色调，诸如黄色、品红色、蓝色、红色等。在示例性实施例中，在直方图中存在比化验物中的染色剂更多的组区。而且，可以监视针对每一个颜色通道的阈值以确定像素中的该颜色通道内染色剂的存在。组区和阈值可以基于化验物信息。直方图可以映射出所有所观察到的色调在图像中的存在，并通过相对暗度来表示每一个色调内的像素的数目，如图6A-6B中所示。

[0049] 图6A描绘了针对具有复染剂654以及两个IHC染色剂651和653的系统的典型直方图。解剖学复染剂标记物这里可以是蓝色的，其中若干蓝色像素与第一IHC染色剂653混合进来，并且其它像素与第二IHC染色剂651混合，并且中间的每个组合也是可信的。图6B描绘了针对具有四个染色剂654、653、651和新添加的染色剂656的图像的直方图。这些染色剂的存在指示可与解剖学可能性比较的图案。观察到，尽管复染剂654非常可能与染色剂653和651并置，但是存在与个体染色剂656的分离，染色剂656与其它染色剂无关。因此，这可以被视为不可信的协同定位，且在解混过程中被忽略。针对该示例化验物的预期直方图数据将不填充表示染色剂651和653的混合物的直方图组区。该方法使得能够优化针对具有大于颜色通道数目的数目的染色剂的图像的参考矢量。

[0050] 图7A-B图示了用于执行预期直方图数据与针对图像数据而生成的直方图数据的

匹配的示例性实施例。在该实施例中,可以将像平面布置、3D空间中的布置或直方图组区的任何其它几何空间之类的几何布置与从该几何空间中的染色剂的参考矢量创建的单纯形(例如针对平面布置的多边形)进行比较。参考矢量可以被用于在直方图之上拟合多边形以充当视觉参考。图7A示出形成多边形765的初始未经优化的矢量。初始参考矢量定义了直方图组区的平面布置中的“角点”。图7B示出从经优化的参考矢量生成的理想多边形拟合。参考矢量优化循环可以优化图7A的参考光谱以创建最类似于图7B的所观察到的直方图的模型直方图。例如,图7A中所描绘的多边形拟合与图7B中所示的模型多边形拟合匹配,并且确定可以存在更好的匹配或者从匹配生成的质量度量低于阈值。作为结果,可以执行优化循环以创建搜索空间内的可能参考矢量的不同集合,直到找到最佳地匹配于直方图的参考矢量的集合。

[0051] 而且,可以以针对通道的多个维度更加抽象的方式来生成和处理针对三通道或RGB图像的所描绘的直方图。尽管不能够被可视化或以图形方式描绘,但是多维直方图将基于解剖学知识和化验物信息而继续供给与模型直方图的有用比较,并且参考矢量被调节以找到所预期的内容与在图像数据中所观察到的内容之间的最佳拟合。

[0052] 因此,与以上所描述的参考矢量优化循环相结合,并且使用已知的生理和化验物信息,与理想直方图的直方图比较提供了哪些参考矢量提供质量图像以及哪些参考矢量需要被调节的有用指示。数值优化器可以用于找到所定义的搜索空间内的解以确定用于对图像进行解混的最优矢量。

[0053] 题述公开内容因此提供了用于优化用于对图像数据进行解混以获得理想结果的参考矢量的系统和方法,例如计算机实现的系统和方法。化验物信息连同图像数据与取决于化验物信息的规则的规则的相关一起用于优化审阅与图像相关联、由用户输入等等的标签或元数据的参考矢量。可以通过在搜索空间内调节参考矢量并且确定后续解混的质量度量来计及否则类似的化验物之间的细小或微小改变。换言之,关于被分析的化验物而已知的内容可以用于消除噪声、不可能性并增强感兴趣的目标结构,生成适合用于后续分析或诊断的清晰图像。所公开的系统和方法因而使得能够生成基本上包括期望的或精确的信号而没有任何噪声或不期望的伪像的图像。可以使用图像自适应参考矢量来迭代地对不期望的签名进行最小化,从而仅留下生物学相关且生理上可信的信息。而且,除了诸如解剖或临床病理学、前列腺/肺癌诊断等之类的医学应用之外,还可以执行相同的方法以分析其它类型的样本,诸如地质学或天文学数据的远程感测等。另外,所公开的重复迭代使得能够精确地分析大型或多个载玻片/图像分析,或者用于分析一个或多个图像立方体,并可以被端口连接到硬件图形处理单元(GPU)中,从而使得能够进行多线程并行实现。

[0054] 已经出于说明和描述的目的提出了本题述公开内容的示例性实施例的前述公开内容。其不意图是详尽的或者将题述公开内容限制到所公开的确切形式。本文所描述的实施例的许多变型和修改鉴于以上公开内容而将对本领域普通技术人员而言显而易见。题述公开内容的范围应仅由随附于此的权利要求且由其等同物限定。

[0055] 另外,在描述本题述公开内容的代表性实施例时,说明书可能已将本题述公开内容的方法和/或过程作为步骤的特定顺序而提出。然而,在该方法或过程不依赖于本文所阐述的步骤的特定次序的程度上,该方法或过程不应当受限于所描述的步骤的特定顺序。如本领域普通技术人员将领会到的,步骤的其它顺序可以是可能的。因此,在说明书中阐述的

步骤的特定次序不应当解释为对权利要求的限制。此外,涉及本题述公开内容的方法和/或过程的权利要求不应当受限于其步骤按所撰写的次序的执行,并且本领域技术人员可以容易地领会到,顺序可以变化并仍旧保持在本题述公开内容的精神和范围内。

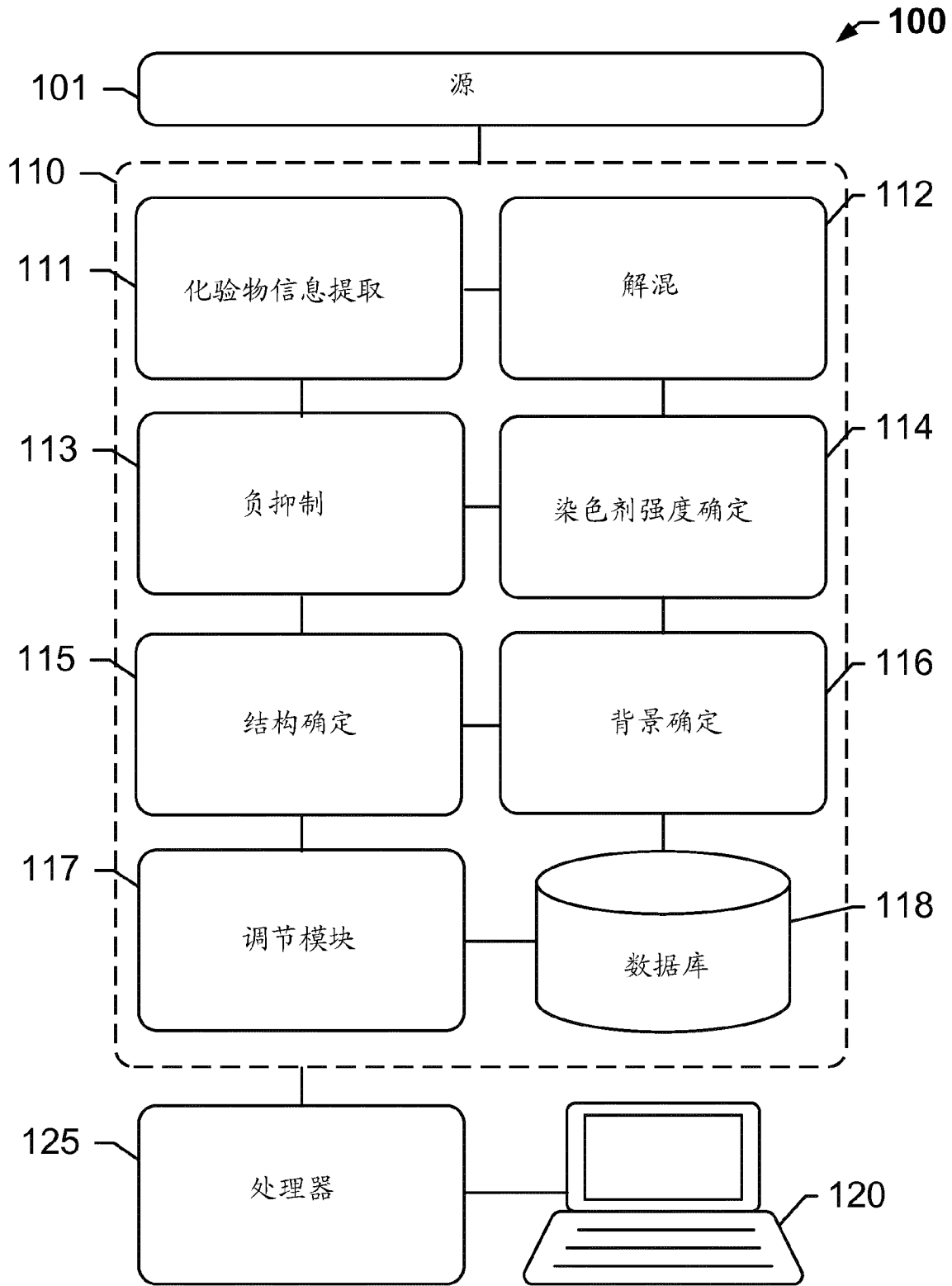


图 1

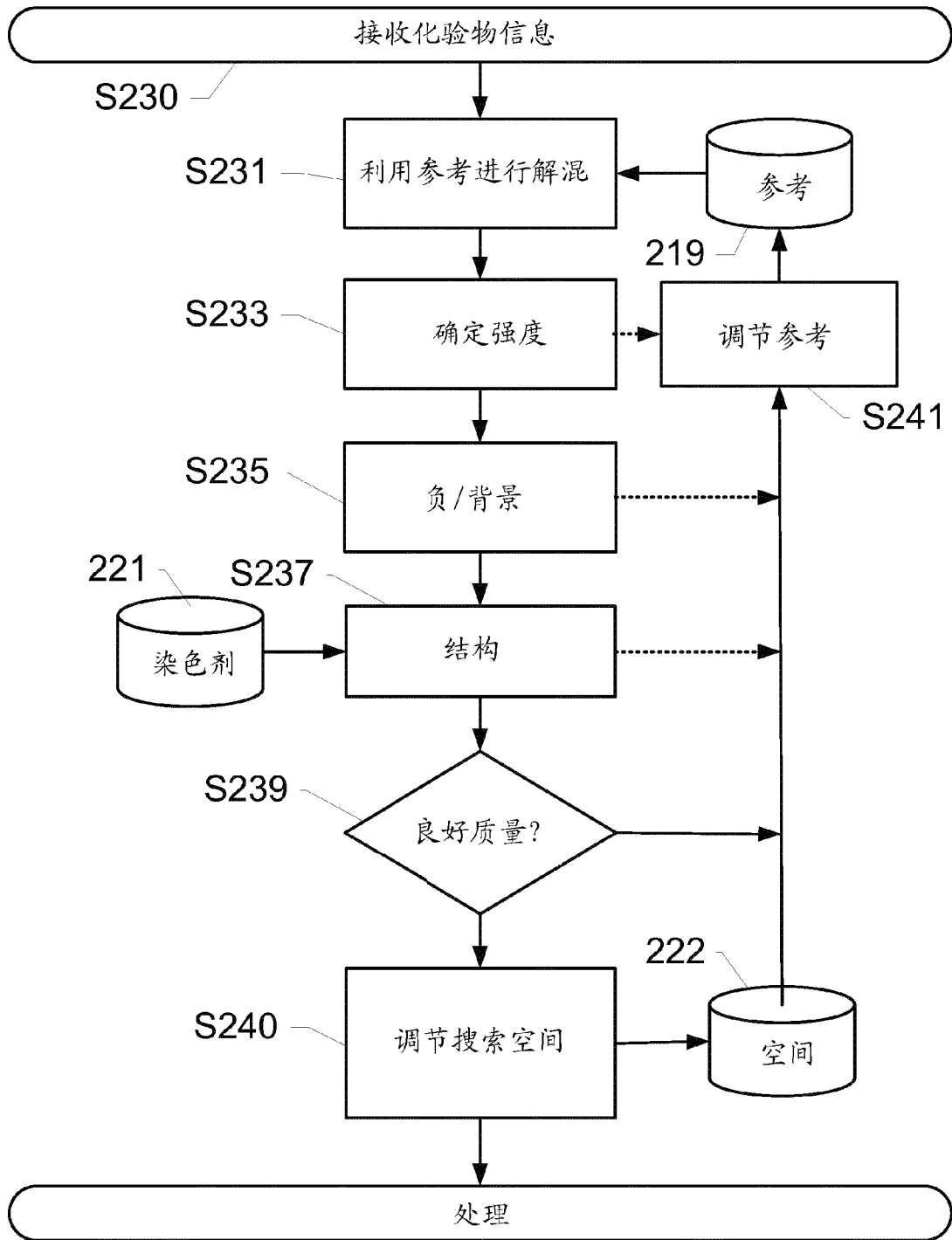


图 2

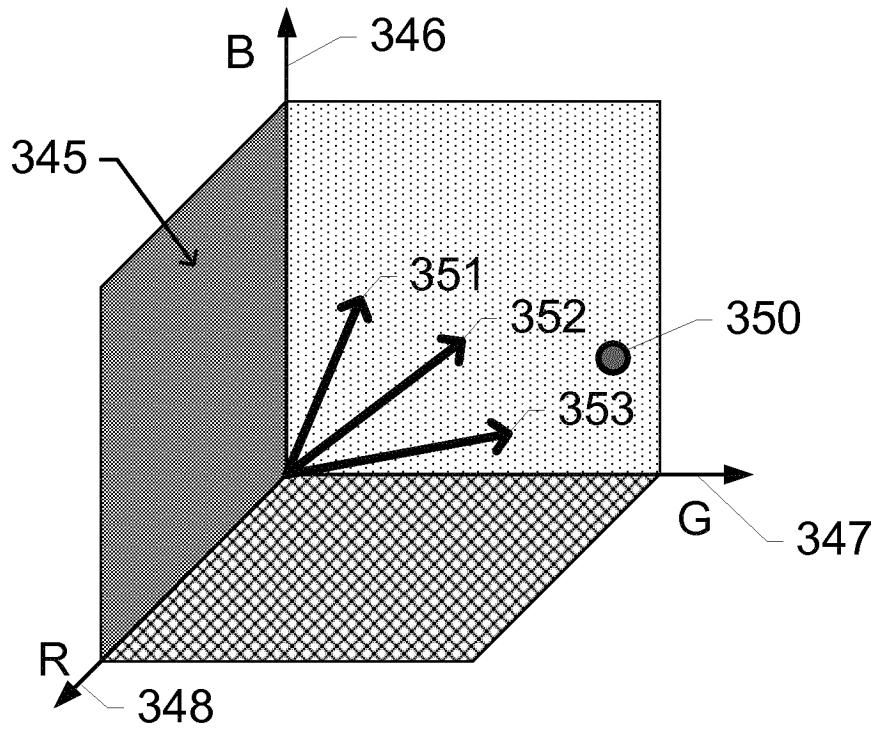


图 3A

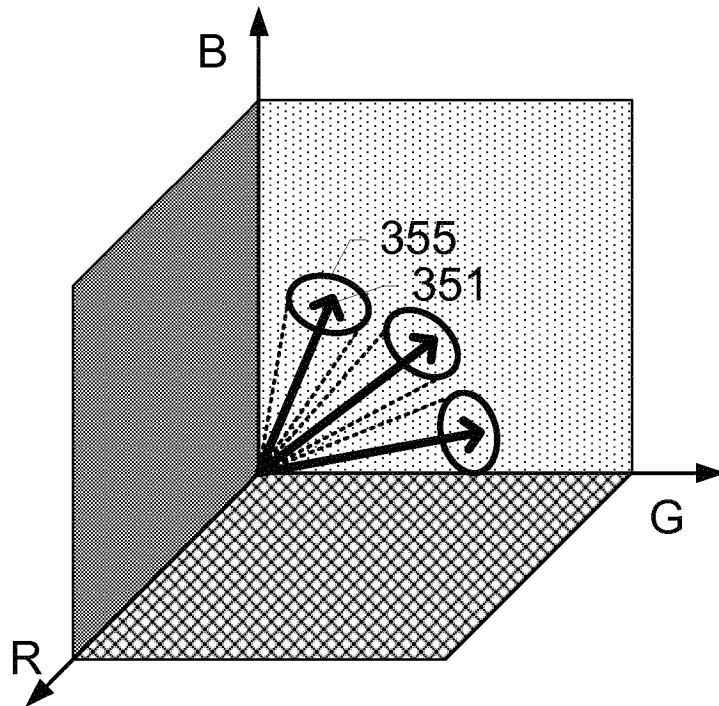


图 3B

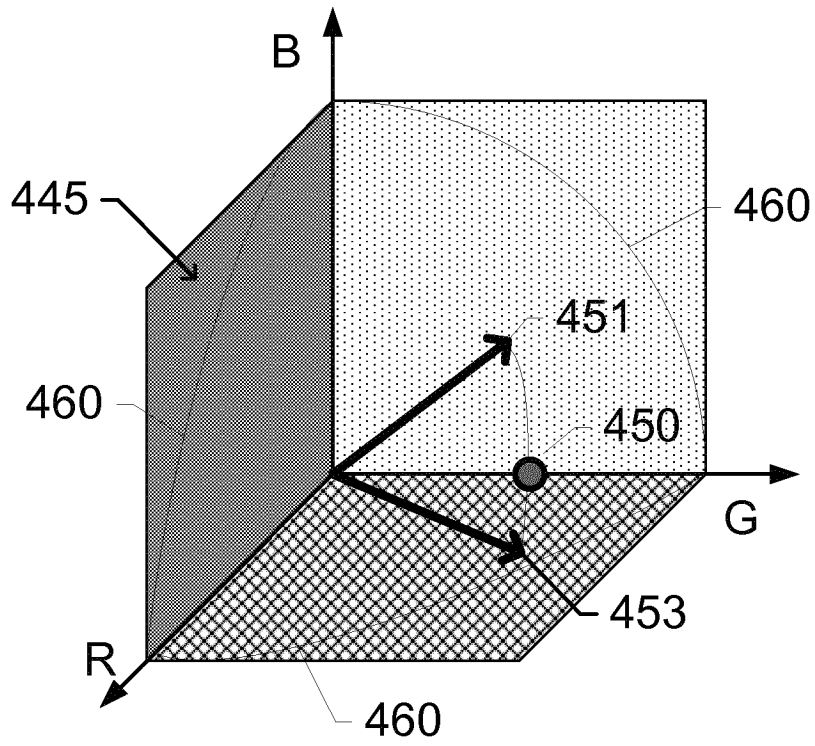


图 4A

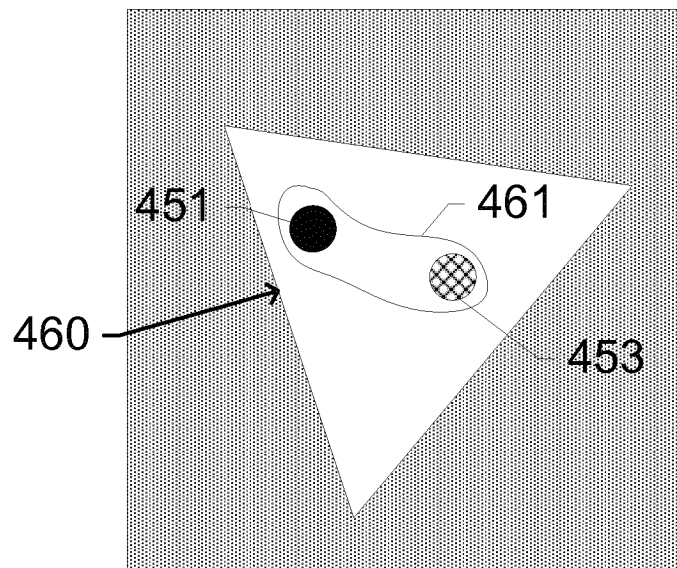


图 4B

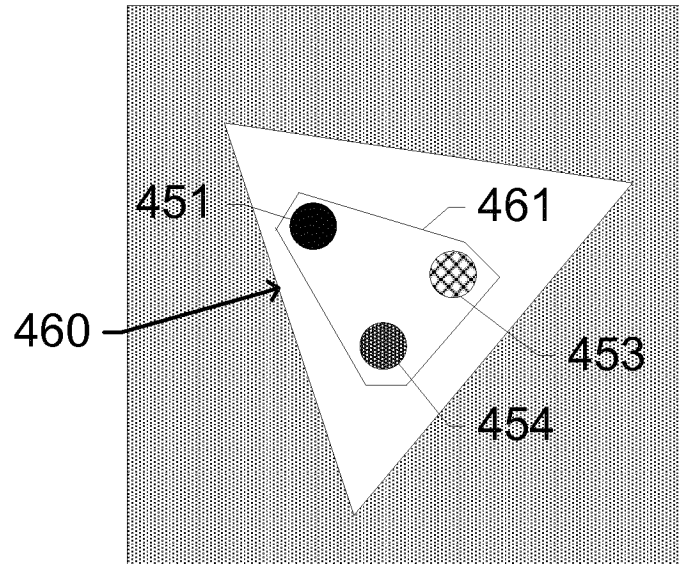


图 4C

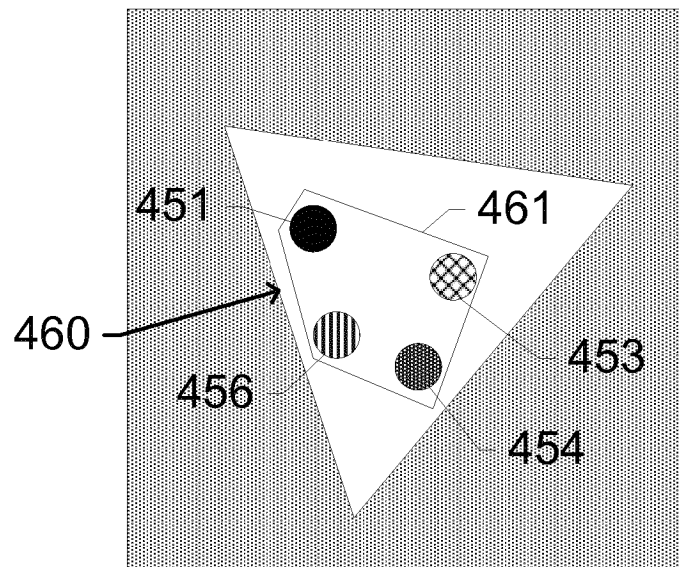


图 4D

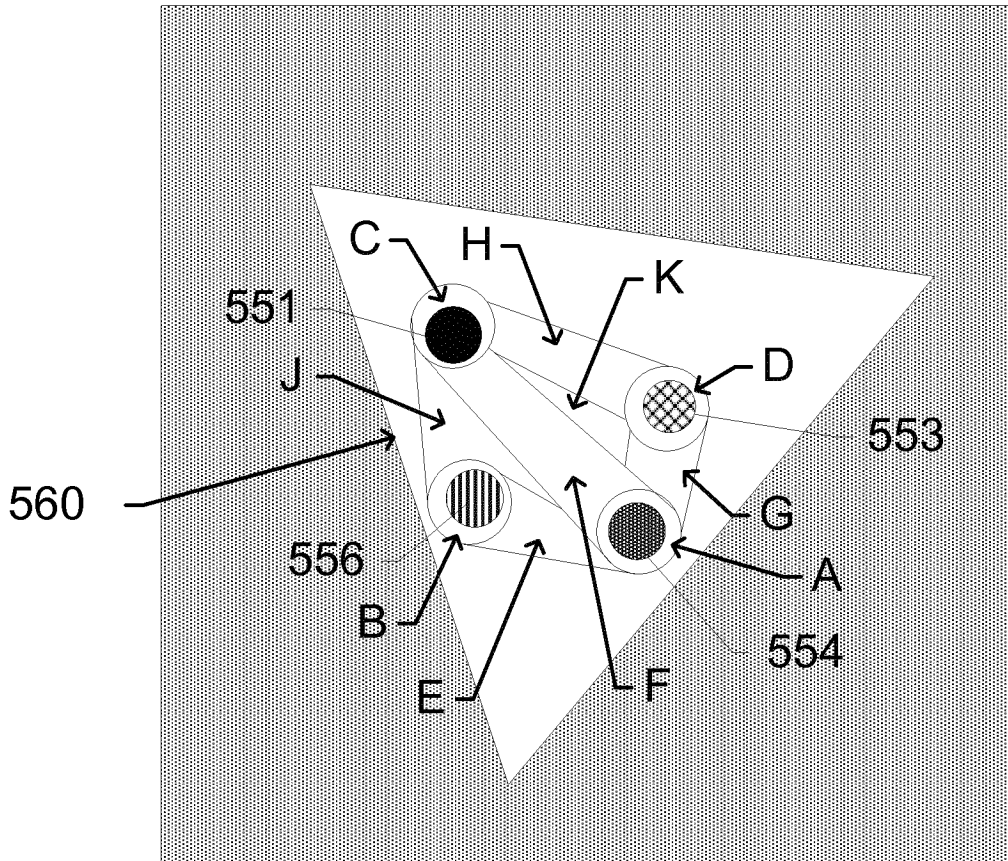


图 5

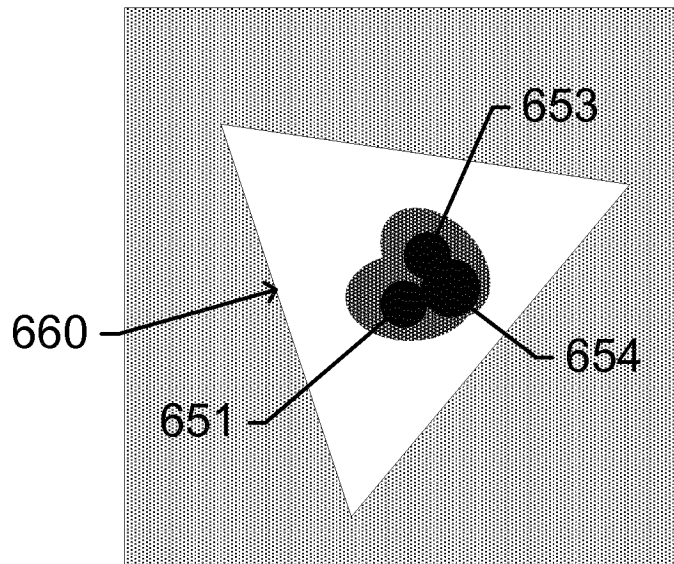


图 6A

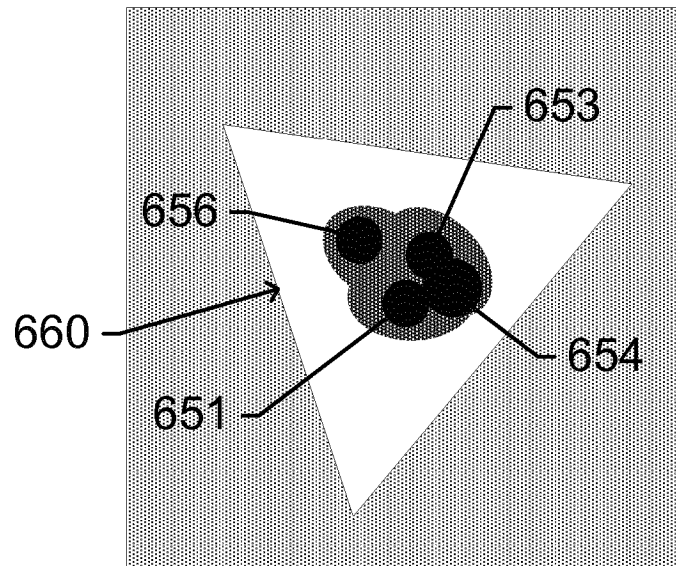


图 6B

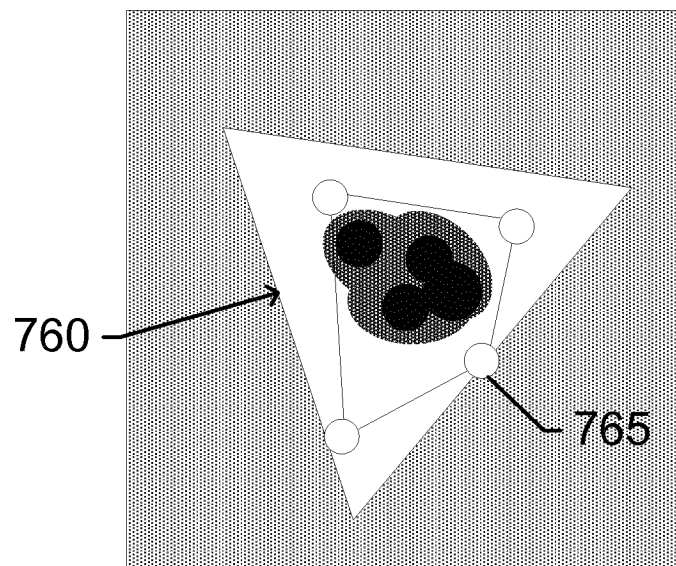


图 7A

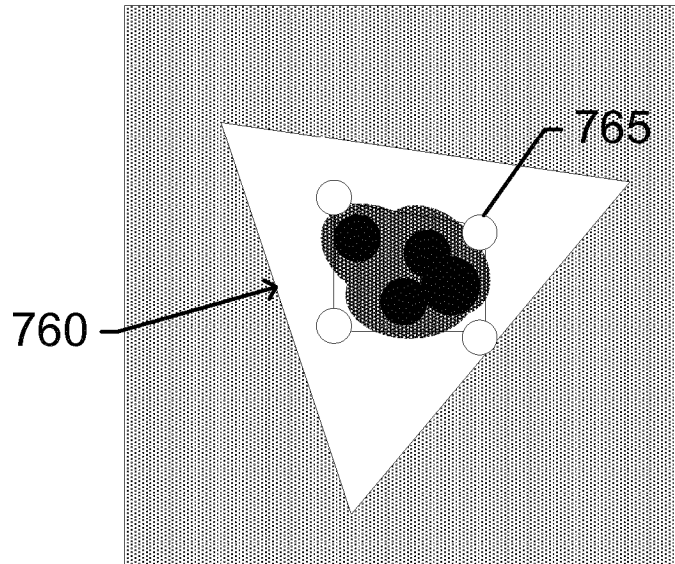


图 7B