



(12)发明专利申请

(10)申请公布号 CN 108693009 A

(43)申请公布日 2018.10.23

(21)申请号 201810250655.2

G06T 7/90(2017.01)

(22)申请日 2018.03.26

G06T 7/00(2017.01)

(30)优先权数据

15/476,848 2017.03.31 US

(71)申请人 合度精密生物科技有限公司

地址 开曼群岛KY1-1 111邮政信箱2681杭
今道克来格广场

(72)发明人 谢煌彬 赖志铭 梅蕊 邵宏仁

王小明 吴仁家

(74)专利代理机构 上海浦一知识产权代理有限

公司 31211

代理人 郑权

(51)Int.Cl.

G01N 1/30(2006.01)

G01N 15/10(2006.01)

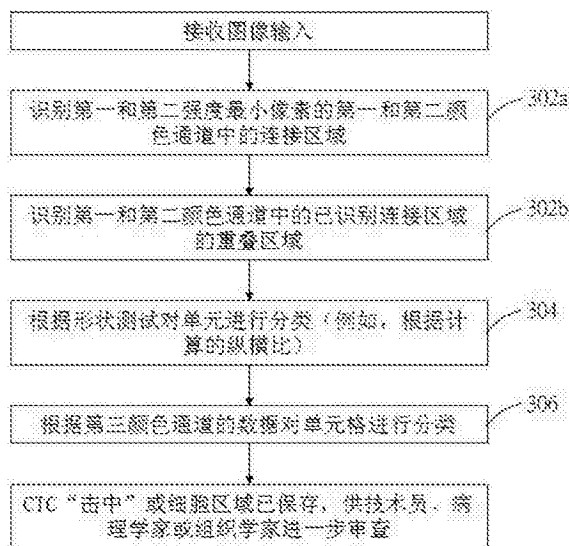
权利要求书3页 说明书16页 附图7页

(54)发明名称

使用图像分析识别候选细胞

(57)摘要

本发明公开了一种用于识别和计数生物流体样本中候选靶细胞的方法,包括获得生物流体样本,通过给生物流体样本中的细胞特征染色来制备生物流体样本,捕获具有多个颜色通道的生物流体样本的数字图像,并将图像分析应用于数字图像。此外,还说明了用于识别生物流体样本中候选靶细胞的计算机程序产品。所述计算机程序包括使处理器执行图像分析的指令。在可能包含感兴趣的候选细胞的样本区域内,可自动定位面积;可对此类面积进行标记,以供进一步评估;可明显减少需由操作者审查的样本面积数量;用于CTC检测和计数的自动化成像过程,可有助于预测治疗期间的疾病进展和总存活期,可实现对患者预后进行连续监测,从而成为了更明智的患者护理选择。



1. 一种用于识别生物流体样本中候选靶细胞的方法,其特征在于,包括:
 - 获得生物流体样本;
 - 使用具有第一颜色并可结合靶细胞的细胞核中核酸的第一生物缀合染料,给生物流体样本中的细胞核染色;使用具有第二颜色并可结合靶细胞的细胞骨架细胞特征的第二生物缀合染料,给生物流体样本中的细胞骨架细胞特征染色;使用具有第三颜色并可结合生物流体标本中人体白细胞抗原的第三生物缀合染料,给生物流体样本中的白血细胞染色,来制备生物流体样本;
 - 捕获生物流体样本的数字图像,其中所述数字图像具有多个颜色通道;
 - 并将图像分析应用于数字图像,包括:
 - 在计算机中接收数字图像;
 - 在计算机中识别多个颜色通道的第一通道中第一强度最小像素的第一连接区域;
 - 在计算机中识别多个颜色通道的第二通道中第二强度最小像素的第二连接区域;
 - 在计算机中确定空间上重叠的第一连接区域和第二连接区域;
 - 在计算机中,根据多个颜色通道中的颜色通道确定空间上重叠的第一连接区域和第二连接区域的纵横比;
 - 在计算机中识别空间上重叠且纵横比满足纵横比阈值的第一连接区域和第二连接区域;
 - 在计算机中确定空间上重叠的第二连接区域和第三连接区域;
 - 在计算机中,根据多个颜色通道中的两个颜色通道确定空间上重叠的第二连接区域和第三连接区域的强度比;
 - 在计算机中,对应空间上重叠且强度比未满足强度比阈值的第二连接区域和第三连接区域排除空间上重叠的第一连接区域和第二连接区域,作为候选区域;
 - 将对应剩余已识别的、空间上重叠的第一连接区域和第二连接区域的一部分图像提供到分类器,作为供分类的候选区域。
2. 根据权利要求1所述的方法,其特征在于,第一颜色为蓝色,第二颜色为红色或橙色,第三颜色为绿色。
3. 根据权利要求1所述的方法,其特征在于,
 - 第一生物缀合染料包括4',6-二脒基-2-苯基吲哚(DAPI),
 - 第二生物缀合染料包括与抗细胞角蛋白(CK)抗体缀合的红色或橙色荧光染料,
 - 第三生物缀合染料包括与抗CD45抗体缀合的绿色荧光染料,或第一抗体、抗CD45和与绿色荧光染料预缀合且靶向CD45的第二抗体的组合。
4. 根据权利要求1所述的方法,其特征在于,
 - 第二生物缀合染料包括与抗细胞角蛋白(CK)抗体缀合的红色或橙色荧光染料,或第一抗体、抗CK和与红色或橙色荧光染料预缀合且靶向CK的第二抗体的组合。
5. 一种用于识别生物流体样本中候选靶细胞的计算机程序产品,其特征在于,所述计算机程序产品可有形地体现在非临时性计算机可读介质中,包括使处理器执行以下程序的指令:
 - 接收生物流体样本的数字图像,其中所述数字图像具有多个颜色通道;
 - 识别多个颜色通道的第一通道中第一强度最小像素的第一连接区域;

识别多个颜色通道的第二通道中第二强度最小像素的第二连接区域；

确定空间重叠的第一和第二连接区域；

根据多个颜色通道中的颜色通道确定空间上重叠的第一连接区域和第二连接区域的纵横比；

识别空间上重叠且纵横比满足纵横比阈值的第一连接区域和第二连接区域；

确定空间上重叠的第二连接区域和第三连接区域；

根据多个颜色通道中的两个颜色通道确定空间上重叠的第二连接区域和第三连接区域的强度比；

对应空间上重叠且强度比未满足强度比阈值的第二连接区域和第三连接区域排除空间上重叠的第一连接区域和第二连接区域,作为候选区域；

将对应剩余已识别的、空间上重叠的第一连接区域和第二连接区域的一部分图像提供到分类器,作为供分类的候选区域。

6. 根据权利要求5所述的计算机程序产品,其特征在于,识别第一连接区域的指令包括:识别第一尺寸最小的第一连接区域的指令;识别第二连接区域的指令包括:识别第二尺寸最小的第二连接区域的指令。

7. 根据权利要求5所述的计算机程序产品,其特征在于,识别第一连接区域的指令包括:识别第一尺寸最大的第一连接区域的指令;识别第二连接区域的指令包括:识别第二尺寸最大的第二连接区域的指令。

8. 根据权利要求5所述的计算机程序产品,其特征在于,识别第一连接区域和第二连接区域的指令包括:最大稳定极值区域算法。

9. 根据权利要求5所述的计算机程序产品,其特征在于,识别第一连接区域的指令包括:将数字图像划分为多个部分,在每个部分搜索潜在的第一连接区域,以及识别以搜索到的潜在第一连接区域为中心的新数字图像部分的指令。

10. 根据权利要求5所述的计算机程序产品,其特征在于,确定空间上重叠的第一连接区域和第二连接区域的指令包括:确定第二连接区域的边界是否适合在第一连接区域的边界内或覆盖所述第一连接区域的边界的指令。

11. 根据权利要求5所述的计算机程序产品,其特征在于,包括确定第一连接区域和第二连接区域组合的指令。

12. 根据权利要求11所述的计算机程序产品,其特征在于,确定纵横比的指令包括:找到在组合的边界上距离最大的两点之间延伸的主轴,找到垂直于主轴且在所述主轴对侧边界上距离最大的两点之间延伸的副轴,计算副轴与主轴比的指令。

13. 根据权利要求12所述的计算机程序产品,其特征在于,纵横比阈值等于或小于0.4。

14. 根据权利要求11所述的计算机程序产品,其特征在于,进一步包括确定组合周围的边界框;确定组合的边界内第一数量的像素和一定范围中第二数量的像素;确定第一数量的像素与第二数量的像素比;将所述比率与范围阈值进行比较的指令。

15. 根据权利要求14所述的计算机程序产品,其特征在于,范围阈值间于0.4和0.85。

16. 根据权利要求11所述的计算机程序产品,其特征在于,所述组合为第一连接区域和第二连接区域的联合。

17. 根据权利要求5所述的计算机程序产品,其特征在于,确定增强比的指令包括:确定

第二连接区域的第一平均强度,确定第三连接区域的第二平均强度,确定第一平均强度与第二平均强度比的指令。

18. 根据权利要求17所述的计算机程序产品,其特征在于,若比率低于阈值,则排除空间上重叠的第二连接区域和第三连接区域。

19. 根据权利要求18所述的计算机程序产品,其特征在于,若I2是第一平均强度,I3是第二平均强度,则可排除空间上重叠的第二连接区域和第三连接区域。

20. 根据权利要求5所述的计算机程序产品,其特征在于,第一颜色为蓝色,第二颜色为红色或橙色,第三颜色为绿色。

21. 一种用于计数生物流体样本中候选靶细胞的方法,其特征在于,包括:

获得生物流体样本;

使用具有第一颜色并可结合靶细胞的细胞核中核酸的第一生物缀合染料,给生物流体样本中的细胞核染色;使用具有第二颜色并可结合靶细胞的细胞骨架细胞特征的第二生物缀合染料,给生物流体样本中的细胞骨架细胞特征染色;使用具有第三颜色并可结合生物流体标本中人体白细胞抗原的第三生物缀合染料,给生物流体样本中的非靶细胞染色,来制备生物流体样本;

捕获生物流体样本的数字图像,其中所述数字图像具有多个颜色通道和至少两个颜色通道;

并将图像分析应用于数字图像,包括:

在计算机中接收数字图像;

在计算机中识别多个颜色通道的第一通道中第一强度最小像素的第一连接区域;

在计算机中识别多个颜色通道的第二通道中第二强度最小像素的第二连接区域;

在计算机中识别多个颜色通道的第三通道中第三强度最小像素的第三连接区域;

在计算机中确定空间上重叠的第一连接区域和第二连接区域;

在计算机中,根据多个颜色通道中的两个颜色通道确定空间上重叠的第一连接区域和第二连接区域的纵横比;

在计算机中确定纵横比满足阈值的空间上重叠的第一连接区域和第二连接区域;

在计算机中确定空间上重叠的第二连接区域和第三连接区域;

在计算机中,根据多个颜色通道中的两个颜色通道确定空间上重叠的第二连接区域和第三连接区域的强度比;

在计算机中,对应空间上重叠且强度比未满足强度比阈值的第二连接区域和第三连接区域排除空间上重叠的第一连接区域和第二连接区域,作为候选区域;

根据对应剩余已识别的、空间上重叠的第一连接区域和第二连接区域的一部分图像,分类作为靶细胞或非靶向元素的候选细胞;

对分类为靶细胞的任何候选细胞进行计数,以生成计数值。

22. 一种用于确定人体受试者中癌症可能性的方法,其特征在于,包括:

根据权利要求21所述,将所得到的计数值与使用统计学确定的一组无肿瘤患者对照的循环上皮细胞计数进行比较;

根据与癌症患者循环上皮细胞计数的统计平均值进行比较的健康受试者循环上皮细胞计数的统计平均值,若总计数超过预定值,则可确定癌症发生的可能性。

使用图像分析识别候选细胞

技术领域

[0001] 本发明涉及在样本的图像中识别候选细胞(例如,循环肿瘤细胞)。

背景技术

[0002] 循环肿瘤细胞(CTC)是已从原发肿瘤脱落并进入脉管或淋巴系统的癌细胞,通常来源于上皮。当CTC侵入循环时,此类恶性细胞进入其他器官。CTC从固体块脱落后,可能会停留在血管壁上,渗入周围组织。血管生成会促进远离原始肿块的部位产生新的肿瘤。因此,CTC代表了其他肿瘤(转移灶)生长的种子。

[0003] 可理解到,外周血中CTC的数量与转移性疾病(包括乳腺癌、结肠直肠癌和前列腺癌)患者的无进展存活和总体存活期缩短有关。可使用来自血液或其他体液样本的CTC检测和计数,来评估肿瘤预后,并协助对癌症患者进行管理。

[0004] 肿瘤会脱落很多细胞。据估计,每克肿瘤组织中有100万个CTC进入外周血。但在24小时内,仅0.1%仍可存活。可存活的CTC应视为“稀有”细胞,原因是CTC在癌症患者的外周血中所观察到的浓度极低,例如,10⁶-10⁷个白细胞中有一个CTC(Sakurai等,2016年)。CTC也存在于高基质的造血细胞中,因而在转移性疾病患者中,CTC出现的频率为每1mL全血中约有1-10个CTC(Miller等,2010年)。因此,CTC难以准确检测和计数。

[0005] 生物染色可增强显微图像分析。可使用某些染料来突出生物细胞的特征和结构。若使用特定的染色,可使CTC在显微镜下具有明显的组织学特征。

发明内容

[0006] 一方面,本发明提供一种用于识别生物流体样本中候选靶细胞的方法包括:获得生物流体样本,通过给生物流体样本中的细胞核染色来制备生物流体样本,捕获具有多个颜色通道的生物流体样本的数字图像,并将图像分析应用于数字图像。

[0007] 另一方面,本发明提供一种用于识别生物流体样本中候选靶细胞的计算机程序产品可有形地体现在计算机可读介质中。所述计算机程序包括使处理器执行图像分析的指令。

[0008] 另一方面,本发明提供一种用于计数生物流体样本中靶细胞群的方法包括:识别候选靶细胞;再对应剩余已识别的、空间上重叠的第一连接区域和第二连接区域,根据一部分图像将候选细胞分类为靶细胞或非靶向元素;对分类为靶细胞的任何候选细胞进行计数,以生成计数值。

[0009] 另一方面,本发明提供一种用于确定人体受试者中癌症可能性的方法包括:将计数值与使用统计学确定的一组无肿瘤患者对照的循环上皮细胞计数进行比较;根据与癌症患者循环上皮细胞计数的统计平均值进行比较的健康受试者循环上皮细胞计数的统计平均值,若总计数超过预定值,则可确定癌症发生的可能性。

[0010] 制备生物流体样本包括:使用具有第一颜色并可结合靶细胞的细胞核中核酸的第一生物缀合染料,给生物流体样本中的细胞核染色;使用具有第二颜色并可结合靶细胞的

细胞骨架细胞特征的第二生物缀合染料,给生物流体样本中的细胞骨架细胞特征染色;使用具有第三颜色并可结合生物流体标本中人体白细胞抗原的第三生物缀合染料,给生物流体样本中的白血细胞染色。

[0011] 图像分析的应用包括:接收数字图像;识别多个颜色通道的第一通道中第一强度最小像素的第一连接区域;识别多个颜色通道的第二通道中第二强度最小像素的第二连接区域;确定空间上重叠的第一连接区域和第二连接区域;根据多个颜色通道中的颜色通道确定空间上重叠的第一连接区域和第二连接区域的纵横比;识别空间上重叠且纵横比满足纵横比阈值的第一连接区域和第二连接区域;确定空间上重叠的第二连接区域和第三连接区域;根据多个颜色通道中的两个颜色通道确定空间上重叠的第二连接区域和第三连接区域的强度比;对应空间上重叠且强度比未满足强度比阈值的第二连接区域和第三连接区域排除空间上重叠的第一连接区域和第二连接区域,作为候选区域;将对应剩余已识别的、空间上重叠的第一连接区域和第二连接区域的一部分图像提供到分类器,作为供分类的候选区域。

[0012] 实施方式可包括以下一个或多个特征:

[0013] 第一颜色可为蓝色,第二颜色可为红色或橙色,第三颜色可为绿色。

[0014] 第一染色剂或生物缀合染料可包括DAPI (4',6-二脒基-2-苯基吲哚)。第二染色剂或生物缀合染料可包括与抗细胞角蛋白 (CK) 抗体缀合的红色或橙色荧光染料。第三染色剂或生物缀合染料可包括与抗CD45抗体缀合的绿色荧光染料,或第一抗体、抗CD45和与绿色荧光染料预缀合且靶向CD45的第二抗体的组合。

[0015] 第二染色剂或生物缀合染料可包括与抗细胞角蛋白 (CK) 抗体缀合的红色或橙色荧光染料,或第一抗体、抗CK和与红色或橙色荧光染料预缀合且靶向CK的第二抗体的组合。

[0016] 在具体实施方式中,举例而言,第二染色剂或生物缀合染料可包括ALEXA568®缀合的抗细胞角蛋白抗体;第三染色剂或生物缀合染料可包括抗CD45-ALEXA488®抗体,或第一抗体、抗CD45和与ALEXA488®预缀合且靶向CD45的第二抗体的组合。

[0017] 第二种染色剂或生物缀合物可包括与ALEXA568®缀合的抗细胞角蛋白 (CK), 所述CK包括与ALEXA568®缀合的单克隆抗体,或第一抗体、抗CK和与ALEXA568®预缀合且靶向CK的第二抗体的组合。

[0018] 识别第一连接区域可包括:识别第一尺寸最小的第一连接区域;识别第二连接区域可包括:识别第二尺寸最小的第二连接区域。识别第一连接区域可包括:识别第一尺寸最小的第一连接区域;识别第二连接区域可包括:识别第二尺寸最小的第二连接区域。识别第一连接区域和第二连接区域包括:最大稳定极值区域 (MSER) 算法。

[0019] 识别第一连接区域可包括:将数字图像划分为多个部分,在每个部分搜索潜在的第一连接区域,以及识别以搜索到的潜在第一连接区域为中心的新数字图像部分。确定空间上重叠的第一连接区域和第二连接区域可包括:确定第二连接区域的边界是否适合在第一连接区域的边界内或覆盖所述第一连接区域的边界。

[0020] 可确定第一连接区域和第二连接区域的组合。确定纵横比可包括:找到在组合的边界上距离最大的两点之间延伸的主轴,找到垂直于主轴且在所述主轴对侧边界上距离最大的两点之间延伸的副轴,计算副轴与主轴比。纵横比阈值可等于或小于0.4。

[0021] 可确定组合周围的边界框;可确定组合的边界内第一数量的像素和一定范围中第

二数量的像素;可确定第一数量的像素与第二数量的像素比,所述比率可与范围阈值进行比较。范围阈值可介于0.4和0.85。所述组合可为第一连接区域和第二连接区域的联合。

[0022] 确定增强比可包括:确定第二连接区域的第一平均强度,确定第三连接区域的第二平均强度,确定第一平均强度与第二平均强度比。若比率低于阈值,则可排除空间上重叠的第二连接区域和第三连接区域。若 $(I2/I3) < 1$,则可排除空间上重叠的第二连接区域和第三连接区域,其中I2是第一平均强度,I3是第二平均强度。

[0023] 优点可包括以下一项或多项。

[0024] 在可能包含感兴趣的候选细胞的样本区域内,可自动定位面积。可对此类面积进行标记,以供进一步评估。如此可明显减少需由操作者审查的样本面积数量。此类用于CTC检测和计数的自动化成像过程,可有助于预测治疗期间的疾病进展和总存活期,可实现对患者预后进行连续监测,从而成为了更明智的患者护理选择。

附图说明

[0025] 图1是CTC识别过程的示意图。

[0026] 图2是说明染色过程的示意图。

[0027] 图3是CTC识别系统的示意图。

[0028] 图4是数字图像中候选细胞识别的计算机实施过程的流程图。

[0029] 图5是数字图像中重叠“斑点”识别的计算机实施过程的流程图。

[0030] 图6A为说明了纵横比的确定方法的示意图一。

[0031] 图6B为说明了纵横比的确定方法的示意图一。

[0032] 图6C为说明了纵横比的确定方法的示意图一。

[0033] 图7为说明了数字图像中围绕“斑点”的边界框的示意图。

[0034] 图8是确定图像中对象是否应分类为候选细胞的思考流程图。

具体实施方式

[0035] 下面将结合附图对本发明的技术方案进行清楚、完整的描述,显然,所描述的实施例是本发明的一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动的前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0036] 实体肿瘤取样是癌症诊断中的常规程序。如今,新一代测序技术能够对癌细胞或其成分DNA中的肿瘤DNA进行灵敏、快速和低成本的检测和分析,所述癌细胞已超出其原始组织,进入细胞之间的流体成分(例如,间质液、淋巴液、血液、唾液、脑脊液、滑液、尿液、粪便等分泌物)。从原发肿瘤中采集的癌细胞碎片可作为监测疾病进展的标记,且可在症状出现之前辅助癌症诊断。

[0037] 在样本载玻片上确定CTC的过程,首先是缩小样本载玻片上存在候选CTC的区域,以便进一步人工评估。CTC的常规测量值范围为 $8\mu\text{m}$ 至 $25\mu\text{m}$,常规样本面积范围约为 50mm^2 至 1200mm^2 。通过肉眼观察肿瘤细胞形态或其他细胞特征来确认“阳性因子”是必要的。给定CTC须定位的大样本区域,在整个载玻片区域中,人工识别候选细胞十分困难、费力;对于操作者而言,评估成像载玻片样本的效率很低。还可能忽略一些阳性因子,特别是图像中的候

选细胞出现的频率很低时。因此,CTC的检测和量化极具挑战性。本文结合高分辨率显微镜对数字图像分析进行了说明,该图像分析可用于有效识别感兴趣的候选CTC。

[0038] 图1是CTC识别过程100一个示例的示意图。如图1所示,从临床点接收生物流体(例如,血液、淋巴液、脑脊髓液、唾液、滑液、尿液、粪便或其他分泌物等体液)的样本10(步骤115)。例如,医生可能希望让患者接收测试,例如,在症状出现之前检测癌症、诊断特定的癌症、监测癌症的进展或表征癌细胞的DNA,以便选择适当的治疗选择方案。

[0039] 患者的血液或其他体液样本可在医生办公室或医疗诊所收集,并将样本发送给系统200的操作者。在其他实施方式中,可在系统200点收集样本。

[0040] 样本也可经历浓缩过程(步骤120至步骤125)。对血液样本进行评估时,富集癌细胞样本特别有用。

[0041] 有几种CTC浓缩技术可减少须分析的细胞总数。示例包括抗体功能化的微流体设备、基于细胞大小的过滤、被动细胞分选和免疫磁性分离。用于分离感兴趣的癌细胞的其他方法、组分和系统包括PCT/US2015/023956、主张的是2014年4月1日提交的申请号为61/973,348的美国临时专利申请案以及2014年4月4日提交的申请号为61/975,699的美国临时专利申请案的权益,序列号为14/065,265(公开号为U.S.2014/0120537)的美国申请案主张的是2012年10月29日提交的编号为61/719,491的美国临时申请案的权益,序列号为14/836,390(公开号为U.S.2016/0059234)的美国申请案主张的是2014年8月26提交的编号为62/042,079的美国临时申请案的权益,全部内容在此通过引用的方式明确地并入本文。

[0042] 例如,如上文引用的参考文献中所述,靶CTC可流过包括玻璃等表面的微流体通道(图1,步骤125)。表面可包含结合部分,感兴趣的CTC附着于所述结合部分(EpCAM示意性结合如图1,步骤125所示)。表面可包含降低非特异性颗粒结合能力的非污染组分(例如,脂质组分、生物活性组分和/或官能团)。因此,感兴趣的CTC纯度可通过降低非特异性颗粒的结合能力而增加。一旦感兴趣的CTC在该表面(例如,可包含脂质双层)上捕获,则可使用温和的清扫力,用一组抗体来洗涤并染色,以保持细胞完整性(图1,步骤127至步骤130)。所述清扫力可以是例如气泡剪力、空气泡沫剪力、乳化流体、超声波振动或油相剪力。在一个具体实例中,包含气泡的泡沫组分可流过表面,以除去结合的细胞和/或不结垢组分(图1,步骤127)。在另一示例中,如PCT/US2012/044701以及序列号为14/128,354(公开号为U.S.2014/0255976,主张的是2011年6月29日提交的申请号为61/502,844的美国临时申请案以及2012年3月2日提交的申请号为61/606,220的美国临时申请案的权益,全部内容在此通过引用的方式明确地并入本文)的美国申请案中所述,“可释放”组分可起到润滑表面的作用,从而仅需低流速的剪切应力即可从表面涂层去除或释放非特异性细胞或血液成分。

[0043] 举例而言,如PCT/US2015/023956(要求的是2014年4月1日提交的编号为61/973,348的美国临时申请案以及2014年4月4日提交的编号为61/975,699美国临时申请案的权益,全部内容在此通过引用的方式明确地并入本文)中所述,可通过使泡沫流过微流体表面,来释放靶CTC,以进行成像和分析,上述方法可提高效率和细胞生存力。

[0044] 较普通、简单的例子中,通过使用RBC裂解缓冲液以及阳性免疫磁性选择,可富集有核细胞的外周血。可通过加入红细胞裂解缓冲液、倒置混合、孵育来裂解红细胞。

[0045] 富集过程的另一示例使用了超过度表达的细胞表面生物标志物,所述生物标志物对CTC(例如,上皮细胞粘附分子(EpCAM))具有高特异性和敏感性。CELLSEARCHSYSTEM®

(Veridex公司)利用抗EpCAM抗体包层的磁性纳米颗粒来捕获和富集CTC,再进行细胞角蛋白免疫染色。ADNATEST®(德国AdnaGen公司)是另一种市场上可获得的CTC检测系统,采用了利用抗EpCAM和Mucin 1(MUC1)缀合磁珠的类似免疫磁性方法。最近,人们开发了基于抗EpCAM抗体包层微流控芯片的“CTC芯片”,用于CTC检测和富集(Nagrath等,《自然》2007, 450:1235-9)。上文引用的专利申请案阐述了血细胞与抗EpCAM抗体的非特异性结合。

[0046] 接下来,应将染色过程(步骤130)应用于样本。在某些情况下,可使用免疫学方法(其中,抗体针对特征性细胞成分)给感兴趣的细胞染色。可使用单克隆抗体进行细胞染色,所述单克隆抗体可识别细胞群体内的特定细胞类型和特征。可使用荧光化合物直接标记抗体,或使用例如识别第一抗体的荧光标记的第二抗体间接标记抗体。可使用一组抗体来分析多标记成像方法中的细胞群体。例如,可使用不同的颜色标记不同的抗体,再进行成像。在一些情况下,多标记成像方法可增加检测CTC的灵敏度。

[0047] 在体液样本中进行CTC的检测和计数应根据通常将来源于上皮的细胞定义为核酸+、CD45-和细胞角蛋白+(CK)的前提。可使用荧光标记的抗体或抗体片段来进行任何数量的不同细胞角蛋白(CK)的免疫细胞化学染色。可将细胞固定在冰冷的甲醇中,在PBS中漂洗,并与抗细胞角蛋白抗血清或针对泛细胞角蛋白(包括所有类型的细胞角蛋白)、I类或II类细胞角蛋白的单克隆抗体或抗体片段或抗个体细胞角蛋白同种型(例如,细胞角蛋白1至细胞角蛋白20)或任何数量的细胞角蛋白同种型的组合进行孵育。

[0048] 还可将细胞与另一种第一(主要)抗体(例如,针对WBC的CD45)和/或针对主要CD45的第二抗体进行孵育。再使用0.5 μ g/ml的DAPI,将样本在室温下的PBS中复染10分钟,并装在甘油明胶中。

[0049] 可将标本固定在中性缓冲甲醛中,再进行渗透(步骤135)。或可将载玻片干燥,并用三乙酸纤维素薄膜或网眼滑动覆盖,以防褪色。在步骤135中,每个载玻片上应用的细胞总数范围可为100至1.5 \times 10⁶。载玻片上的粘合剂区域可由一至三个单独的图像分析圈组成,总共100至530mm²。

[0050] 染色过程包括至少两种不同颜色的染色剂:将生物流体样本中的细胞核染成第一颜色,将生物流体中的细胞骨架细胞特征染成第二颜色;任选地,将生物流体中的白细胞或其他非靶细胞染成第三种颜色。一种或两种此类染色剂可例如使用能特异性识别并结合细胞表面标志物或细胞角蛋白的抗体,优先结合感兴趣的细胞。在一些实施方式中,可使用第一生物缀合染料给细胞核染色,所述第一生物缀合染料可在成像时提供第一颜色并结合靶细胞的细胞核中的核酸。可使用第二种生物缀合染料给细胞骨架细胞特征染色,所述第二生物缀合染料可在成像时提供第二种颜色并结合靶细胞的细胞骨架细胞特征。具体地,第二染色剂可包括通过直接免疫荧光来结合细胞骨架细胞特征(例如,细胞角蛋白)的抗体或抗体片段。抗体可缀合荧光蛋白、第二抗体或可在经光激发后再发光的其他荧光化学化合物。以这种方式,可使用两种抗体,通过间接免疫荧光来实现放大作用。在识别细胞骨架细胞特征(例如,细胞角蛋白)时,第二种染色可标记具有细胞骨架的任何数目的细胞,包括但不限于上皮细胞、内皮细胞、内皮祖细胞、“癌症干细胞”和播散的肿瘤细胞等。可使用第三种生物缀合染料(例如,绿色荧光蛋白(GFP))给白血细胞染色,所述第三种生物缀合染料可在成像并结合人类白细胞抗原时提供第三种颜色。间接免疫荧光也可与第三种染色或生物缀合染料一起使用,以扩增信号。

[0051] 现如图2所示,在一个具体实施方式中,第一染色剂可包括核染色剂(例如,DAPI(4',6-二脒基-2-苯基吡啶))。CTC应对核酸染料(例如,DAPI)染色呈阳性,表明细胞核包含在细胞质内并小于细胞质。存在细胞核表明细胞非红细胞,而是 α 核。

[0052] 第二种染色剂可包括一种或多种染料缀合的抗细胞角蛋白(CK)。所述CK可包含与藻蓝蛋白(APC)、藻红蛋白(PE)或市场上可获得的任何数量的荧光分子(例如,ALEXA FLUOR®或DYLIGHT®染料)缀合的细胞角蛋白特异性单克隆抗体。在图2所示的具体示例中,第二种染色剂包括对细胞角蛋白(CK)(小椭圆形)具有特异性并与ALEXA568®(一种具有荧光红色发射光谱的小分子有机染料)缀合的抗体,因此能标记和分化上皮细胞。CTC应对ALEXA568®缀合的抗细胞角蛋白染色呈阳性,呈圆形、椭圆形或多边形,具有完好的膜,尺寸至少约为4 μ m。

[0053] 第三种染色剂可包括抗CD45、CD45特异性的单克隆抗体、存在于白细胞表面的抗原(例如,与绿色荧光蛋白(GFP)缀合),或市场上可获得的具有荧光绿色发射光谱的任何数量有机染料(例如,ALEXA488®或DYLIGHT488®)。CTC不对CD45染色呈阳性,原因是此类染色剂可识别白细胞上存在的抗原,且CTC不能是白血细胞。虽然本文讨论了特定的染料,但也可考虑本领域中已知的其他类似染料。

[0054] 第一种颜色可为红色或橙色,第二种颜色可为蓝色,第三种颜色可为绿色,但也可使用其他颜色组合。染料可为在成像过程中应用的、在光(例如,UV光或可见光或红外光)下发光的荧光染料。或者,染料可为吸收性染料。

[0055] 一旦应用了染色剂,则将样本转移到观察载玻片20(图1,步骤135)。例如,载玻片20可包括过滤器22(见图1和图3)(例如,多孔膜或网);可将样本分配到过滤器上,使过滤器捕获候选细胞(例如,候选CTC),同时允许其他流体流过滤网。过滤器也可捕获其他细胞(例如,白细胞或其他非靶细胞)。过滤器22可安装在观察载玻片20的顶部。观察载玻片20可由玻璃、有机玻璃或本领域中已知的类似合适材料制成。过滤器22的直径可约为5-25mm,平均孔径(例如,网的间距)可达10 μ m(例如,1至3 μ m或2至5 μ m);平均孔径可小于2 μ m。过滤器22可由塑料(例如,聚碳酸酯)制成。

[0056] 现可分析样本(图1,步骤140-150)。具体地,可使用下文中的系统200来识别候选细胞(例如,候选CTC),而无需技术人员输入(见图3)。可放置载玻片20用于成像(步骤140),可对样本进行成像(步骤145),可分析图像以识别候选细胞(步骤150)。

[0057] 图3是CTC识别系统的示意图。如图3所示,系统200包括成像显微镜210和至少一台计算机250,所述计算机250可控制显微镜的捕获机构、控制台架和显微镜之间的相对运动(例如,沿X、Y和/或Z方向)和/或控制光源的激活和/或光学滤波器的移动,以激发和捕获不同波长的荧光。计算机250还可分析来自显微镜210的图像并识别样本中的候选细胞(例如,候选CTC)。

[0058] 成像显微镜210包括数码摄像头212和光学组件214(例如,透镜等),以将摄像头212聚焦在固定于台架220的载玻片上的点。在计算机250的控制下,所述台架可沿X、Y和/或Z方向进行机动运动。由显微镜210捕获的数字图像具有至少两个颜色通道,例如,三个通道(例如,红色通道、绿色通道和蓝色通道)。每个颜色通道可对应其中一种染料产生的颜色,但无需对应精确的波长。可选择成像显微镜的分辨率和放大率,使单个像素在一侧上对应0.3至1.3 μ m(例如,约0.648 μ m²)。在一个示例中,成像显微镜210可使用10倍物镜和数字摄

像头,所述数字摄像头可生成具有三个颜色通道的1392x1040像素和每像素每通道12位的数字图像。

[0059] 摄像头可耦合或包括存储器232,以存储来自摄像头212的数字图像。所述存储器232可为控制器230(例如,运行应用程序的通用计算机)的一部分,以便控制显微镜210。

[0060] 台架220可由致动器222(例如,三轴致动器)支撑,所述致动器222可沿平行于固定载玻片20平面的两个垂直水平轴线移动台架220,垂直轴线垂直于所述平面。或者,致动器222可在台架220保持固定的同时,移动摄像头212和光学部件214。

[0061] 致动器222可连接到控制器230。控制器230可在台架220和摄像头212之间产生相对运动,以自动扫描横跨载玻片20的显微镜210成像的区域,并控制摄像头212捕获图像的定时,以便生成可覆盖样本所放置的区域(例如,过滤区域22)的数字图像阵列。控制器230可允许操作者配置显微镜210或调整扫描参数。

[0062] 成像显微镜210还可包括光源240。假设染色剂包括荧光染料,则光源240可产生波长适合于使染料发荧光的光(例如,UV),对红外光可见。因此,成像显微镜210可产生染色剂(例如核、细胞角蛋白和CD45染色剂)的数字图像。假设使用荧光染料且滤光器22由在来自光源240的光下不发荧光的材料组成,则滤光器不应出现在数字图像中。或者,若使用被动染料,则光源240可产生白光。

[0063] 再将数字图像传送到计算机250,以便存储和分析。数字图像可以tiff等无损格式进行存储和/或传输。例如,计算机250可通过USB连接等串行总线连接或以太网或因特网等网络耦合到显微镜210,且计算机250可自动检索来自存储器232的数字图像。或者,计算机250本身可提供控制器230;这种情况下,存储器232可为计算机250的一部分。可选地,存储器232可为便携式设备(例如,闪存驱动器),可由操作者从显微镜210物理地移除并插入计算机250中,所述计算机250可为与操作显微镜的计算机分离的独立计算机。

[0064] 任何情况下,所述计算机250可接收数字图像,并对图像进行分析,以识别候选细胞(例如,候选CTC和/或白血细胞(WBC))。通常,由计算机250识别的候选细胞的数字图像需由技术人员审查,以确认每个候选细胞实际上是感兴趣的细胞(例如,CTC)。但通过自动排除数字图像中的大部分无关物体,可明显减少技术人员评估的候选细胞数量,从而提高效率并缩短生成测试结果所需的时间。

[0065] 如图4所示,识别候选细胞的计算机实施过程包括三个主要步骤:首先,计算机250可识别数字图像两个颜色通道中的重叠“斑点”(即,强度最小的像素连接区域)(见步骤302a和302b)。第一颜色通道可对应第一颜色(例如,对应靶细胞的细胞核中核酸的染色),第二颜色通道可对应第二染料的颜色,例如,对应靶细胞的细胞骨架细胞特征的染色)。其次,计算机250对重叠斑点组合中产生的形状进行一个或多个形状测试(步骤304)。如此,可排除图像伪像以及某些类型的细胞。第三,计算机250可根据对第三颜色通道的评估来排除形状(步骤306)。第三颜色通道可对应第三颜色(例如,对应白血细胞的染色),如此可排除白细胞。若先前的过程(例如,富集过程)可有效地从样本中去除白血细胞,则所述第三步可任选。下文将更详细地讨论每个步骤。

[0066] 如图3所示,计算机250识别候选细胞时,可存储每个候选细胞的识别信息。例如,计算机250可将标签(例如,书签)插入到数字图像中,可将数字图像中候选细胞的坐标存储在数据库中,或可剪切与候选细胞对应的一部分数字图像并将该部分保存在单独的文件或

数据库中。

[0067] 若需要,可将识别信息和数字图像转发到分类器以进行后处理或离线检查,以确定每个候选细胞是否应分类为靶细胞。候选细胞的识别也可由分类器执行。

[0068] 分类器可为检查数字图像的技术人员。例如,计算机250可通过网络260(例如局域网(LAN)或因特网)连接到计算机270。

[0069] 技术人员可使用计算机270来查看与每个候选细胞相对应的该部分数字图像,例如,查看该部分的数字图像并确定该候选细胞是否实际上是靶细胞(例如,CTC)。例如,计算机270可自动接收识别信息,根据识别信息确定一部分数字图像,并按控制的顺序将所确定的部分通过计算机270显示器显示给技术人员。例如,计算机270可接收候选细胞的坐标,选择以坐标为中心的一部分数字图像,再自动显示所选部分。—如此可减少技术人员需通过数字图像搜索候选细胞的要求。在另一实施方式中,技术人员可接收候选列表,所述列表中的每个元素可链接到一部分数字图像。用户可从列表中选择元素,使计算机向技术人员呈现相应部分的数字图像。

[0070] 或者,若将来能完全自动确定候选细胞是否是靶细胞,则可通过向计算机272提供适当的指令来执行该任务。这种情况下,分类器为计算机272。例如,计算机250可通过网络260连接到计算机272。

[0071] 在一些实施方式中,计算机270或272可例如通过技术人员或自动执行来计数确定为靶细胞的细胞数量。该计数可用于生成评分(例如,CTC的总数、百分比、相对于健康个体的比率等)。在其他实施方式中,可通过回归方程式生成评分,除了CTC计数之外还包括年龄、性别、体重指数、家族癌症史、酒精使用情况、身体活动或其他生活方式等其他风险因素。最终将测试结果返回给实体(例如,进行测试的医生)。测试结果可包括评分、与靶细胞相对应的数字图像部分,或两者。根据这些测试结果,医生可通过例如将评分与来自一组无肿瘤患者对照的循环上皮细胞的统计学所确定的评分进行比较,来确定癌症发生率或复发可能性。根据此比较,若所得的总分超过根据来自健康受试者循环上皮细胞计数的统计平均值的预定值,则医生可确定癌症发生率或复发可能性。

[0072] 该评分还可用于筛查健康个体中未检测到的癌症,诊断患有症状的患者的癌症或检测患者状态的推定改变。在一个实施方式中,该评分有助于确定侵袭性或惰性形式癌症的适当治疗过程。就此而言,可使用新一代测序技术对从患者取样的CTC内的DNA进行测序,以确定癌症驱动突变。可对照视为与特定靶向疗法相关的遗传标志物组,来评估突变。靶向治疗可攻击特定类型的癌细胞,对正常细胞的损害较小。例如,一种抗体药物HERCEPTIN®(美国基因泰克(Genentech)),可靶向人类表皮生长因子受体2(HER2),所述HER2可在癌性卵巢和乳房组织的表面上过度表达。

[0073] 可对患者样本中的CTC进行其他基因检测。若出现肺癌的情况,可通过抽血来评估CTC,如此有利于病情严重而不能提供肺肿瘤活检的患者。根据简单的抽血和CTC测序,CTC中DNA的基因检测可揭示患者是否患有表皮生长因子受体(EGFR)基因突变的非小细胞肺癌。了解肿瘤的基因组成,可帮助医生决定患者是否可受益于常规化疗或靶向抗癌药物(例如,TARCEVA®),所述靶向抗癌药物可抑制EGF受体的特定活化突变形式。

[0074] 通过CTC的检测和计数,并进一步测序CTC中的DNA,医生可较好地了解患者的特定癌症亚型。如此可得出治疗选择,并改善结果。就乳腺癌而言,例如,70%的肿瘤细胞可表达

过量的激素受体,所述激素受体可结合雌激素或黄体酮并刺激细胞生长。此类肿瘤最好用荷尔蒙疗法进行治疗。在其他乳腺癌亚型中,约20%具有与人类表皮生长因子2 (HER2) 结合的过多受体。此类细胞最好用TYKERB® (美国诺华 (Novartis)) 和HERCEPTIN® (美国基因泰克 (Genentech)) 等靶向HER2受体的药物进行治疗。其他约10%的乳腺癌细胞是“三重阴性”,即此类细胞不具有任何过多的受体,而可能产生BRCA肿瘤抑制基因中的突变。HERCEPTIN® (美国基因泰克 (Genentech)) 不适用于此类肿瘤的治疗,且若对通过CTC的检测、定量和测序所提供的信息无效果的患者进行治疗,可能会导致不理想的结果并增加潜在的死亡率。

[0075] 参见计算机250的实施过程,图5是数字图像中重叠“斑点”识别的计算机实施过程的流程图。为了识别每个“斑点”,计算机可识别强度最小的像素连接区域。连接区域是数字图像中相邻像素的连续区域。可对每个连接区域进行尺寸测试,例如,该区域的尺寸范围是否介于总像素数量的上限阈值和下限阈值之间。可根据图像的分辨率(即,每微米像素的数量)来预先确定上限阈值和下限阈值,以对应可与CTC对应的物理尺寸(例如,以微米为单位)。

[0076] 具体地,为了识别第一颜色通道中的每个“斑点”,计算机可识别第一颜色通道中第一强度最小的像素第一连接区域(步骤312)。

[0077] 可确定第一连接区域的第一尺寸是否最小(步骤314)。例如,可计算第一连接区域中的像素总数。若第一连接区域的像素总数小于第一下限阈值(例如,五十像素),则可将其作为候选细胞排除。由于细胞核的尺寸最小,若第一连接区域过小,则表示第一通道中的“斑点”非细胞核,因而非候选细胞。

[0078] 另外,可确定第一连接区域的第一尺寸是否最大。例如,若第一连接区域的像素总数大于第一上限阈值(例如,一千五百像素、或一千像素),则可将其作为候选细胞排除。由于细胞核的尺寸最大,若第一连接区域过大,则表示“斑点”非细胞核,因而非候选细胞。

[0079] 可对第二颜色通道执行类似的处理。为了识别第二颜色通道中的每个“斑点”,计算机可识别第二颜色通道中第二强度最小的像素的第二连接区域(步骤316)。

[0080] 可确定第二连接区域的第二尺寸是否最小(步骤318)。例如,可计算第二连接区域中的像素总数,例如,通过确定该区域之后对图像进行栅格化并对标记为处于第二连接区域中的像素进行计数。若第二连接区域的像素总数小于第二下限阈值(例如,一百像素),则可将其作为候选细胞排除。由于细胞的尺寸最小,若第二连接区域过小,则表示第二颜色通道中的“斑点”非细胞,因而非候选细胞。

[0081] 另外,可确定第二连接区域的第二尺寸是否最大。例如,若第二连接区域的像素总数大于第二上限阈值(例如,一千五百像素),则可将其作为候选细胞排除。由于细胞的尺寸最大,若第二连接区域过大,则表示“斑点”非细胞,因而非候选细胞。

[0082] 可将像素数量除以分辨率(例如,每单位面积的像素数量),再与表示尺寸(以面积为单位,例如,平方微米)的阈值进行比较,来代替将像素数量与阈值数量进行比较。

[0083] 识别强度最小的像素连接区域的一种实施方式是通过设置阈值(例如,低于最小强度的像素设为0,大于最大强度的像素设为1)将灰度图像转换为二进制图像。该过程可检查二进制图像的像素,例如,在数字图像中从上到下和从左到右运行,确定像素是否与已分配给斑点的另一像素相邻。若像素与现有斑点的像素相邻,则可将该像素分配给已识别的

斑点。否则,将创建并存储新的斑点数据记录(例如,给该像素分配新的斑点记录)。可通过经验研究来选择最小强度作为区分细胞与噪声的值,例如,操作者在显示样本图像时调整每个颜色通道中的最小强度,并通过视觉检查确定阈值是否可靠地生成并可可靠地从噪声中区分细胞。

[0084] 可将例如根据强度平滑的其他图像滤波技术应用于图像。噪声过滤可避免错误的“开”像素,从而防止产生新的斑点数据记录,例如,斑点足够小的情况。噪声过滤还可避免错误的“关”像素,以改善第二连接区域中像素总数的计算。

[0085] 识别强度最小的像素连接区域的另一实施方式是最大稳定极值区域(MSER)算法。在这种技术中,数字图像中的像素可按照亮度顺序排序。将排序的像素逐一放置到空白图像上,以生成斑点。具体地,该过程包括一个try-catch子句,该子句包括作为迭代变量、带有排序强度指数的for循环。在该for循环中,首先检查结束条件,以查看是否已达到图像背景强度。结束条件可包括两个方面:图像中斑点的数量达到20,图像中的最大斑点达到3000像素。对于在给定图像位置具有给定像素强度的给定迭代,若未达到结束条件,则该迭代将通过将给定像素强度放置在给定图像位置来进行。

[0086] 像素放置之后,应检查相邻像素,查看已放置的像素是否存在于图像中。若存在,应针对该像素的位置对现有的斑点数据记录进行更新(例如,将像素分配到已记录的斑点),否则应创建并存储新的斑点数据记录(例如,给像素分配新的斑点记录)。继续该像素放置过程,直到达到结束条件。若达到结束条件,迭代将停止,操作将退出for循环,并完成try-catch子句。第十三届英国机器视觉会议电子会议(2002)上,J.Matas等人的《基于最大稳定极值区域的鲁棒宽基线立体》中讨论了极大稳定极值区域。

[0087] 在一些实施方式中,数字图像可分成多个区域,并分析每个区域。可选择区域的像素数量,使该区域可大于靶细胞的预期尺寸(例如,约3到10倍)。例如,数字图像可分成矩形区域,例如,一侧为50至200个像素,形成 100×100 像素的正方形。假设一侧的像素为 $0.648 \mu\text{m}$,则该 100×100 像素区域可表示观察载玻片上的 $64.8 \mu\text{m} \times 64.8 \mu\text{m}$ 区域。由于CTC的测量范围为 $8 \mu\text{m}$ 至 $20 \mu\text{m}$,该区域足够大,完全适合将CTC置于该区域。

[0088] 在一些实施方式中,若识别连接区域包括多于阈值数量的像素,则选择以连接区域为中心的新的重新居中区域,并对新的重新居中区域进行重叠“斑点”的识别。重新居中包括图像的形成,该图像包括分割图像阵列中多个相邻区域部分。

[0089] 接下来,确定每个第一连接区域是否与第二区域重叠(步骤302b)。由于循环肿瘤细胞既包括细胞核和细胞骨架,若第一或第二颜色通道中存在连接区域,而其他第一和第二颜色通道中不存在对应连接区域,则表示该连接区域不属于靶细胞的一部分,因而非靶细胞,并标记出。因此,计算机250可确定第一颜色通道和第二颜色通道中“斑点”的重叠区域;若无重叠,则可将斑点作为候选细胞排除。

[0090] 在一些实施方式中,若第二区域重叠和/或适合置于第一连接区域中,则可认为第二连接区域与第一连接区域重叠。在一些实施方式中,第一连接区域需围绕第二连接区域,例如,第一连接区域围绕第二连接区域,形成连续环形。可对两个通道(例如,蓝色和红色通道)进行二进制滤波,作为确定蓝色斑点和红色斑点之间重叠过程的一部分。

[0091] 现可对第一和第二颜色通道中重叠连接区域所指示的潜在候选细胞进行一次或多次形状测试。

[0092] 第一形状测试可确定重叠连接区域的纵横比是否落入阈值纵横比范围内。靶细胞(例如,CTC)通常非长形。假定纵横比是较短测量值与较长测量值的比值,若重叠连接区域的纵横比低于阈值,则表示重叠连接区域太长且非候选细胞。

[0093] 在一些实施方式中,如图6A所示,为了确定纵横比,可确定第一连接区域和重叠的第二连接区域相联合的组合区域,再确定组合区域的外边界。外边界可为组合区域外周边上的一组像素。

[0094] 如图6B所示,可计算组合区域的主轴。主轴可为连接边界上距离最大的两个像素的线段。例如,为了确定主轴,对边界上每个像素进行函数计算,可确定该像素与边界上每个其他像素之间的距离;距离最大的一对像素为定义主轴的两个像素。

[0095] 如图6C所示,可针对组合区域计算副轴。主轴将组合区域分成两半。副轴可为与主轴垂直的最长线段,所述主轴可连接主轴两侧边界上的两个像素。例如,为了确定副轴,对主轴一侧边界上的每个像素,可确定在垂直于主轴的线上,该像素与主轴另一侧边界上像素之间的距离;距离最大的一对像素为定义副轴的两个像素。

[0096] 纵横比可计算为副轴长度与主轴长度的比值。纵横比可与阈值进行比较。例如,若纵横比小于0.4(例如,小于0.35,小于0.3),则可将重叠连接区域作为候选细胞排除。

[0097] 第二形状测试可比较重叠连接区域的填充因子。如此提供了另一种检测细长区域或消除形状极端不规则的方法。

[0098] 在一些实施方式中,如图7所示,为了确定填充因子,可在组合区域周围建立边界框。边界框呈矩形,具有与组合区域的最上面和最下面像素匹配的上边界和下边界,以及与组合区域的最右边和最左边像素相匹配的右边界和左边界。

[0099] 可对组合区域的外边界内的像素数量进行计数,并可对组合区域之外但在边界框内(也称为“范围”)的像素数量进行计数。将外边界内的像素数量与该范围内像素数量的比率与填充率阈值进行比较。若比率小于填充率阈值,则组合区域非候选细胞。填充率阈值可约为0.4至0.85(例如,0.60)。

[0100] 假设所有先前的测试均通过,则可将组合区域确定为候选细胞,但非靶细胞和白血细胞也有可能通过该测试。因此,可使用附加过程来筛选白血细胞。通常,该步骤包括确定组合区域中两个不同通道之间的强度比。

[0101] 其中一个通道可为第二通道。另一通道可为第三频道。这种情况下,可对第一通道和第二通道确定第三通道中的第三连接区域(例如,如上文所述),且可确定第二连接区域与第三连接区域的重叠区域。重叠区域中两个通道的强度值可用于确定强度比。可排除该强度比不满足强度比阈值的空间上重叠的第二连接区域和第三连接区域,作为候选细胞。

[0102] 简而言之,白血细胞应含有较多第三染色剂,因而在第三种颜色通道中应显示出较大强度。因此,也可排除第三颜色通道中同样表现出高强度的候选区域。

[0103] 在一些实施方式中,可通过计算第二连接区域的平均强度 I_2 来执行该测试。另外,可例如使用上述用于第一和第二连接区域的技术之一,来识别第三颜色通道中第三强度最小的像素第三连接区域。还可计算第三连接区域的平均强度 I_3 。

[0104] 在一些实施方式中,可计算颜色通道内的平均背景强度,例如,范围区域内或包括组合区域的所选图像子区域内的强度。从通道连接区域的平均强度中减去颜色通道的平均背景强度,以提供调整后的平均强度。

[0105] 可使用以下公式来确定细胞是否是候选细胞(例如,CTC(组1)、模糊(组2)或白血细胞(组3))。

$$[0106] \text{ Group 1: } \frac{I2}{I2-I3} \leq 2$$

$$[0107] \text{ Group 2: } \frac{I2}{I2-I3} > 2 \text{ and } \frac{I2}{I3} > 1$$

$$[0108] \text{ Group 3: } \frac{I2}{I3} < 1$$

[0109] 其中,I2和I3是如上文所述的平均值或调整后的平均值。

[0110] 可排除对应组3的组合区域;满足 $I2/I3 \geq 2$ 的区域可标记为由技术人员评估的候选对象。

[0111] 使用本文中的算法来测试,计算机250可确定数字图像中第一和第二颜色通道中重叠的第一连接区域和第二连接区域。再根据一个或多个颜色通道确定空间上重叠的第一连接区域和第二连接区域的纵横比。再确定纵横比满足阈值的一个或多个空间上重叠的第一连接区域和第二连接区域。再将与识别的空间上重叠的第一连接区域和第二连接区域对应的图像部分显示给操作者,作为分类的候选区域。

[0112] 图8是确定图像中对象是否应分类为候选细胞的思考流程图。该流程图不一定与计算机250上软件的操作顺序相同。

[0113] 示例

[0114] CTC检测在癌症诊断和预后中具有重要的应用性。在外周血中存在表面表达“上皮细胞粘附分子”(EpCAM,一种几乎可在所有癌上表达的泛上皮(包括全部)分化抗原)的CTC与接受转移性乳腺癌治疗的患者的无进展存活期和总存活期下降相关。表达EpCAM的CTC缩写为“EpCAM+CTC”。对于某些平台,每7.5ml血液中EpCAM+CTC计数等于或大于5,表示无进展存活期和总存活期缩短。

[0115] 在另一应用中,每个样本可处理2至10mL血液,且可使用确定的循环肿瘤细胞数目来预测疾病的风险或复发。在另一实施方式中,从血液中确定的CTC数量可能仅是包含其他作为变量的危险因素(除CTC之外)的回归方程中数个变量之一。

[0116] 在一些实施方案中,CTC应该对CK-ALEXA568®(红色染料)染色呈阳性,呈具有完整膜的圆形、椭圆形或多边形,尺寸至少为4 μ m。CTC不应该对CD45-ALEXA488®(绿色染料)染色呈阳性,原因是此类染色剂可识别白细胞上存在的抗原,而CTC非白血细胞。CTC也应该对DAPI核染色(蓝色染料)染色呈阳性,表明细胞核包含在细胞质内且比胞质小至少30%。存在细胞核表明细胞非红细胞,而是 α 核。

[0117] CTC可从肿瘤块脱离流入血液。早期对外周血中的CTC跟踪和计数,可改变治疗方案并可减缓转移。CTC难以在其他细胞中分类。如图6A至6C所示,在光学观察之前,对视野中来自外周血的示意性细胞进行染色。细胞的图像集可用于审查,例如,其中每个图像集包含CK-ALEXA568®(细胞角蛋白)柱、DAPI核复染剂以及用于WBC(或其他非靶细胞)的CD45-ALEXA488®染色剂。技术专家可根据特定的标准检查每条线,并选择有资格作为肿瘤细胞的细胞。在一些实施方式中,例如,技术专家可将图像库中的对象解释为循环肿瘤细胞(CTC)、白细胞(WBC)、鳞状细胞、在同一帧中具有白细胞的肿瘤细胞、非靶细胞或双阳性细胞。观察图像对象时的各种考虑因素包括:有核细胞、像素化细胞、分离的细胞核和非细胞碎片(例如,伪像和计算机噪声)。

[0118] 使用光控制,可用红光先捕获第一图像,用蓝光捕获第二图像,用绿光捕获第三图像。

[0119] 显微镜可具有机动平台,所述平台可扫描幻灯片上的样本,并分段捕捉图像。使用自动化步骤,本文中的计算机执行过程可解释重叠的颜色和形状(圆形/椭圆形与非蜂窝形)。在某些实施方式中,机动平台可重新定向幻灯片并解析哪些部分具有CTC。可将此类图像存储为感兴趣部分,以供审阅。

[0120] 大型、非典型出现的细胞不应作为CTC计数。像素化、模糊和不完整的细胞不应作为CTC计数。由于CK-ALEXA568®或DAPI信号的过度放大,因此可能会出现少量污染物。此类图像不符合CTC的标准。

[0121] 在一个实施方式中,图像捕获装置(例如,显微镜)可识别第一强度最小的蓝色像素第一连接区域。第一强度最小的连接蓝色像素可对应推定的WBC或CTC核。蓝色像素连接区域的识别可在100x100像素块中进行。一旦识别,则系统将重新对中该像素块,使蓝色像素连接区域在100x100像素块中位于中心。

[0122] 图像俘获装置再识别第二强度最小的红色像素第二连接区域。可识别空间上重叠的蓝色像素第一连接区域和红色像素第二连接区域,并确定空间上重叠的第一连接区域和第二连接区域的纵横比。一个或多个空间重叠区域中,若纵横比满足阈值,则可识别该区域。再将与空间上重叠的第一连接区域和第二连接区域对应的图像部分显示给操作者,作为候选细胞以供分类。

[0123] 定义

[0124] 本文中,术语“循环肿瘤细胞”(CTC)是指非白细胞循环液(优选外周血)中的有核细胞。

[0125] CTC是遗留下原发性肿瘤,从而进入血流或淋巴系统的罕见细胞。若患膀胱癌,CTC可从肿瘤块脱离并进入尿液。若患唾液腺癌,CTC可从肿瘤块脱离并进入唾液。因此,本文中的方法和过程将CTC定义为具有细胞核(例如,对DAPI染色呈阳性)、具有上皮细胞特征(例如,对细胞角蛋白染色呈阳性),且非白细胞(例如,不对CD45染色呈阳性)。该物体须大于 $4 \times 4 \mu\text{m}^2$,并具有细胞状形态。

[0126] 本文中,术语“DAPI”是指4',6-二脒基-2-苯基吲哚,一种具有荧光蓝色发射光谱的染色剂,可强烈地与DNA中富含A-T的区域结合。用于荧光显微镜时,DAPI可在活细胞和固定细胞中通过完整的细胞膜,从而能标记和区分有核细胞和无核细胞(例如,红细胞)。

[0127] 本文中,术语“CK”是指细胞角蛋白,是在上皮细胞的胞质内细胞骨架中发现的含角蛋白的中间丝。术语“抗CK/ALEXA568®”是指特异于细胞角蛋白的单克隆抗体,与ALEXA568®(一种具有荧光红色发射光谱的有机染料)缀合,因此能标记和分化上皮细胞。应注意到,抗CK/ALEXA568®也指可识别CK的第一抗体和靶向第一抗体、与染料(例如,ALEXA568®)缀合的第二抗体的组合。抗细胞角蛋白(CK)可为抗泛(包括全部)细胞角蛋白、抗个体细胞角蛋白同种型(例如,细胞角蛋白1至细胞角蛋白20),或任何数量的细胞角蛋白同种型的组合。更普遍地,“抗CK/PE”和“抗CK/APC”是指特异于细胞角蛋白的单克隆抗体,可缀合到具有荧光红色发射光谱的有机染料类别(例如,别藻蓝蛋白(APC)和藻红蛋白(PE))。

[0128] 本文中,术语“CD45”是指簇分化45,一种存在于白细胞表面上的抗原。-抗

CD45-ALEXA488®是CD45特异性单克隆抗体,与ALEXA488®(一种具有荧光绿色发射光谱的有机染料)缀合,因此能标记和分化白细胞(也称为白血细胞(WBC))。应注意到,抗CD45/ALEXA488®也可指靶向第一抗体的第一(CD45)和第二抗体(ALEXA488®)的组合。

[0129] 本文中,术语“体液”包括腹水、唾液、尿液、滑液、腹膜液、羊水、脑脊髓液、浆膜液和/或脊髓液。

[0130] 术语“核染色剂”是指用于指示细胞中存在细胞核的染料化合物。核染色剂包括吖啶橙、溴化乙锭、单叠氮乙锭、Hoechst染料、碘化丙锭和DAPI等嵌入染料。

[0131] 本文中,术语“荧光标记”是指可共价连接至另一分子(例如,蛋白质或核酸)的荧光团,该连接过程通常通过使用荧光团的反应性衍生物来完成,所述反应性衍生物可选择性结合包含在靶分子中的官能团。荧光标记包括但不限于别藻蓝蛋白(APC)、荧光素(FITC)、罗丹明(FAM、R6G、TET、TAMRA、JOE、HEX、钙红、VIC和ROX)、德克萨斯红、BODIPY、香豆素、花青染料(噻唑橙[TO]、恶唑黄[YO]、TOTO、YOYO;Cy3、Cy5)、ALEXAFLUOR®染料、DYLIGHT®染料、绿色荧光蛋白(GFP)和藻红蛋白(PE)。

[0132] 本文中,术语“生物样本”可最广义地使用为含有核酸或其蛋白质转录产物。样本可包含体液(例如,血液);细胞制剂的可溶部分或细胞生长的培养基的等分试样;染色体、细胞器或从细胞中分离或提取的膜;溶液中或与基质结合的染色体DNA、RNA或cDNA;细胞;组织;组织印迹;指纹;细胞;皮肤等。在优选的实施方式中,该术语是指来自含有细胞的受试者的生物材料,且包括可检测到CTC的任何材料。样本可例如为全血、血浆、唾液或其他含有细胞的体液或组织。优选的样本是全血,更优选为外周血,更优选为外周血细胞部分,更优选为从血液中分离或富集的CTC。

[0133] 本文中,术语“抗体”是指通常存在于体内或响应所消耗的抗原而产生的多种蛋白质中的任何一种,从而可产生免疫应答。抗体优选为包含IgG亚型的免疫球蛋白。

[0134] 本文中,术语“与...特异性反应”是指具有比相同抗原和非特异性抗体之间的结合更好的特异性(通常还具有亲和性)抗体与抗原之间的结合。

[0135] 使用本文所述系统和方法进行分析,所需的血液(或其他身体分泌物)量可约等于或小于25 μ L、50 μ L、75 μ L、100 μ L、0.2mL、0.5mL、1mL、1.5mL、2mL、2.5mL、3mL、3.5mL、4mL、4.5mL、5mL、5.5mL、6mL、6.5mL、7mL、7.5mL或8mL、9mL、10mL、11mL、12mL、13mL、14mL、15mL或16mL。使用本文所述系统和方法进行分析,所需的血液量可等于或高达25 μ L、50 μ L、75 μ L、100 μ L、0.2mL、0.5mL、1mL、1.5mL、2mL、2.5mL、3mL、3.5mL、4mL、4.5mL、5mL、5.5mL、6mL、6.5mL、7mL、7.5mL或8mL、9mL、10mL、11mL、12mL、13mL、14mL、15mL或16mL。本文中,术语“约”是指后面所跟值 \pm 1、2、3、4、5、6、7、8、9或10%内的量。

[0136] 例如,11mL血液的样本也可称为约等于10mL的血液样本。应理解到,若提供数值范围,除非上下文清楚地另行指明下限的十分之一单位,否则还应具体说明该范围的上限和下限之间每个中间值。所述范围内的任何规定值或中间值与该规定范围内任何其他规定值或中间值之间的每个较小范围均包含在本发明内。该较小范围的上限和下限可独立地包括或不包括在该范围内,且每个范围(其中任意一个或两个限值均包括或不包括在较小范围内)也包含在本发明内,受限于规定的范围。若所述范围包括一个或两个限值,则范围(排除所包括的限值中的任意一个或两个)也应包括在本发明中。以下示例将更全面地说明本发明的优选实施方式,但决不应视为对本发明广泛范围的限制。

[0137] 计算机系统

[0138] 上述一个或多个计算机(例如,计算机250)包括处理器252、存储器254、存储设备256以及一个或多个输入/输出接口设备258。例如,可使用系统总线259将组件252、254、256和258中的每一个互连。

[0139] 处理器252能处理在系统250内执行的指令。本文中,术语“执行”是指程序代码可使处理器执行一个或多个处理器指令的技术。在一些实施方式中,处理器252是单线程处理器。在一些实施方式中,处理器252是多线程处理器。在一些实施方式中,处理器252是量子计算机。处理器252能处理存储在存储器254或存储设备256中的指令。

[0140] 存储器254可存储系统250中的信息。在一些实施方式中,存储器254是计算机可读介质。在一些实施方式中,存储器254是易失性存储器单元。在一些实施方式中,存储器254是非易失性存储单元。

[0141] 存储设备256能为系统250提供大容量存储。在一些实施方式中,存储设备256是非临时性计算机可读介质。在各种不同实施方式中,存储设备256可包括例如硬盘设备、光盘设备、固态驱动器、闪存驱动器、磁带或一些其他大容量存储设备。在一些实施方式中,存储设备256可为云存储设备,例如,包括分布在网络上并使用网络(例如,图3中的网络260)访问的一个或多个物理存储设备的逻辑存储设备。在一些示例中,存储设备256可存储长期数据(例如,数字图像)。

[0142] 输入/输出接口设备258可为系统250提供输入/输出操作。在一些实施方式中,输入/输出接口设备258可包括接口设备的网络(例如,以太网接口)、串行通信设备(例如,RS-232接口)和/或无线接口设备中的一种或多种(例如,802.11接口、3G无线调制解调器、4G无线调制解调器等);网络接口设备可例如使用网络260使计算机系统250传送和接收数据(例如,数字图像)。在一些实施方式中,输入/输出设备可包括用于接收输入数据并将输出数据发送到其他输入/输出设备(例如键盘、打印机和显示设备)的驱动器设备。在一些实施方式中,可使用移动计算设备、移动通信设备和其他设备。

[0143] 用于执行系统200的图像分析和其他操作的软件可在执行时,通过使一个或多个处理设备执行上述过程和功能的指令来实现。该指令可包括例如解释指令(例如,脚本指令),或可执行代码,或存储在计算机可读介质中的其他指令。

[0144] 尽管本文已说明了示例性处理系统,但可在其他类型的数字电子电路或计算机软件、固件或硬件中执行上述主题和固定操作的实施方式,包括本说明书中公开的结构及其结构等同物或一个或多个结构的组合。本说明书中所述的存储、维护和显示伪像等主题的实施方式可作为一个或多个计算机程序产品来实施,即在有形程序载体上进行编码的计算机程序指令的一个或多个模块(例如,计算机可读介质),以便处理系统完成执行任务或控制处理系统的操作。计算机可读介质可为机器可读存储设备、机器可读存储基质、存储设备或一个或多个上述设备的组合。

[0145] 术语“系统”可包括用于处理数据的所有装置、设备和机器(例如,包括可编程处理器、计算机或多个处理器或计算机)。除了硬件之外,处理系统还可包括为所涉及的计算机程序创建执行环境的代码(例如,构成处理器固件、协议栈、数据库管理系统、操作系统或一个或多个系统组合的代码)。

[0146] 计算机程序(也称为程序、软件、软件应用程序、脚本、可执行逻辑或代码)可以任

何形式的编程语言(包括编译或解释语言,或声明性或过程性语言)进行编写,且可以任何形式部署,包括作为独立程序或模块、组件、子程序或适用于计算环境的其他单元的形式。计算机程序不一定对应文件系统中的文件。程序可存储在保存其他程序或数据(例如,存储在标记语言文档中的一个或多个脚本)的一部分文件、专用于所述程序的单个文件,或多个协调文件(例如,存储一个或多个模块、子程序或部分代码的文件)中。计算机程序可部署在一台计算机上或多台计算机上执行,此类计算机位于一个站点或跨多个站点分布并通过通信网络互连。

[0147] 适用于存储计算机程序指令和数据的计算机可读介质包括所有形式的非易失性或易失性存储器、介质和存储设备,包括半导体存储设备(例如,EPROM、EEPROM和闪存设备)、磁盘(例如,内部硬盘或可移动磁盘或磁带)、磁光盘和CD-ROM、DVD-ROM和蓝光光盘等。处理器和存储器可由专用逻辑电路补充或并入专用逻辑电路。服务器有时可为一台通用计算机,有时可为定制的专用电子设备,有时可为上述设备的组合。实施方式可包括后端组件(例如,数据服务器)、中间件组件(例如,应用服务器)或前端组件(例如,具有图形用户界面或Web浏览器的客户端计算机,通过所述图形用户界面或Web浏览器,用户可与本说明书中所述主题的实施方式进行互动),或一个或多个此类后端、中间件或前端组件的任何组合。系统的组件可通过数字数据通信的任何形式或介质(例如,通信网络)互连。通信网络的示例包括局域网(“LAN”)和广域网(“WAN”) (例如,因特网)。

[0148] 尽管上文的讨论集中于上皮细胞CTC的检测,但原则上,本文所述的技术可应用于其他种类的细胞(例如,其他循环稀有细胞(CRC)、播散性癌细胞、干细胞(例如,肿瘤干细胞和骨髓干细胞)、胎儿细胞、细菌、内皮细胞等)。

[0149] 综上所述,上述各实施例及附图仅为本发明的较佳实施例而已,并不用以限定本发明的保护范围,凡在本发明的精神和原则之内,所做的任何修改、等同替换、改进等,皆应包含在本发明的保护范围内。

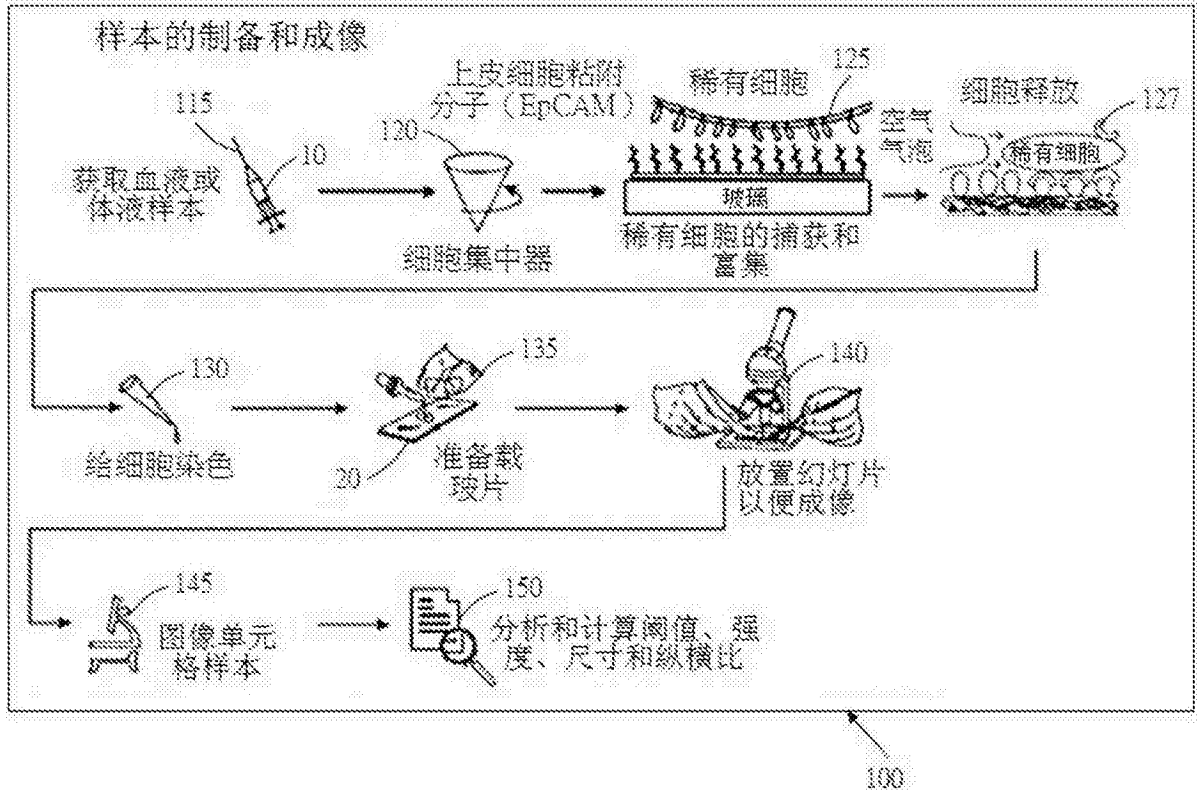


图1

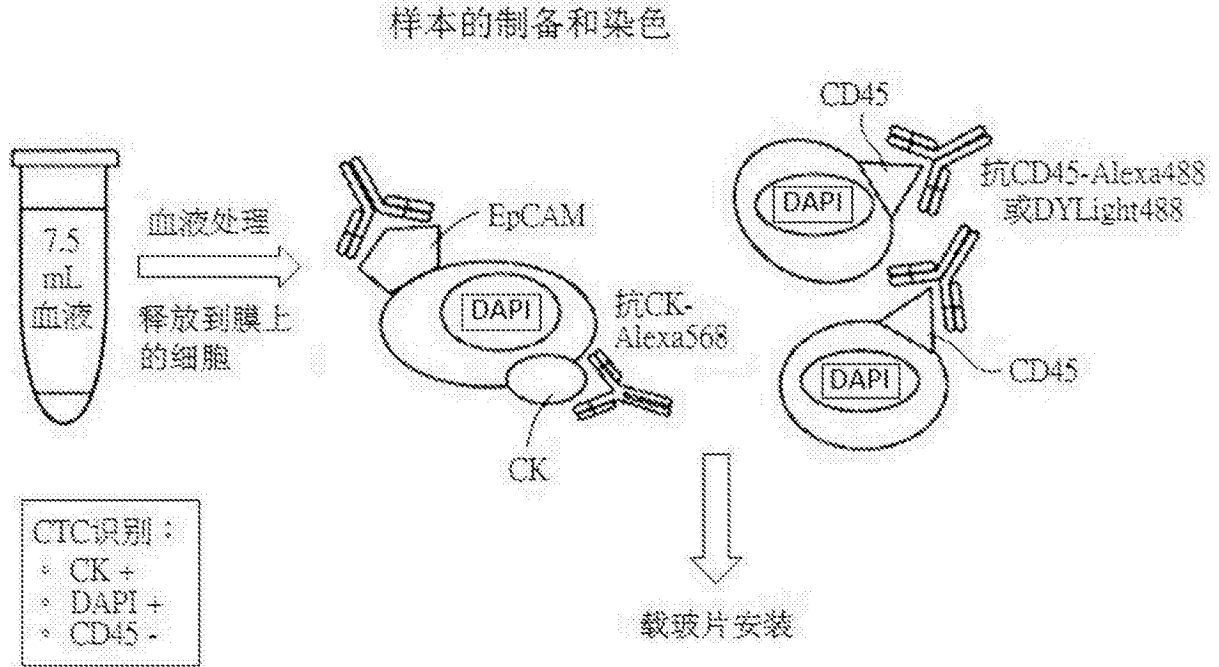


图2

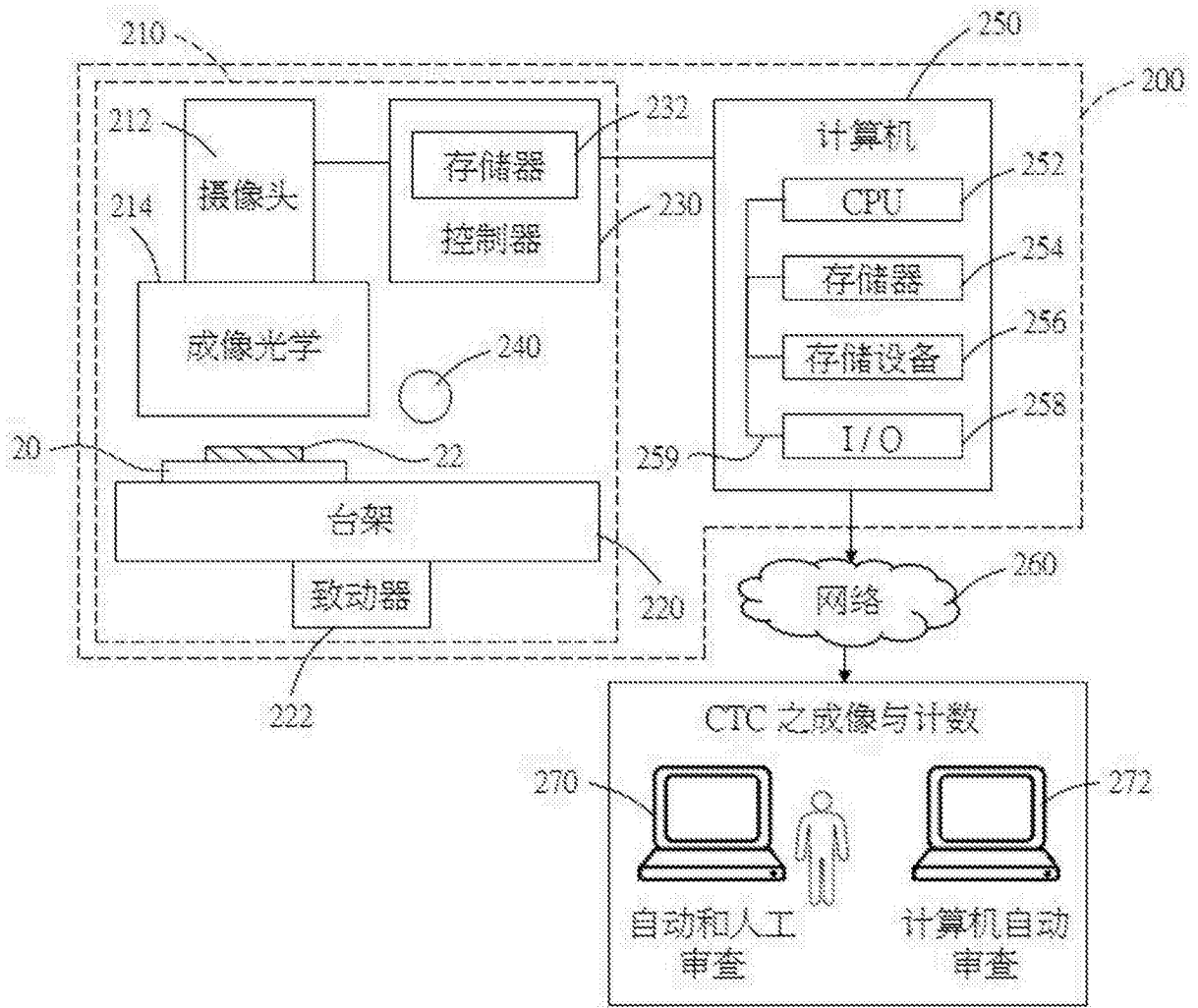


图3

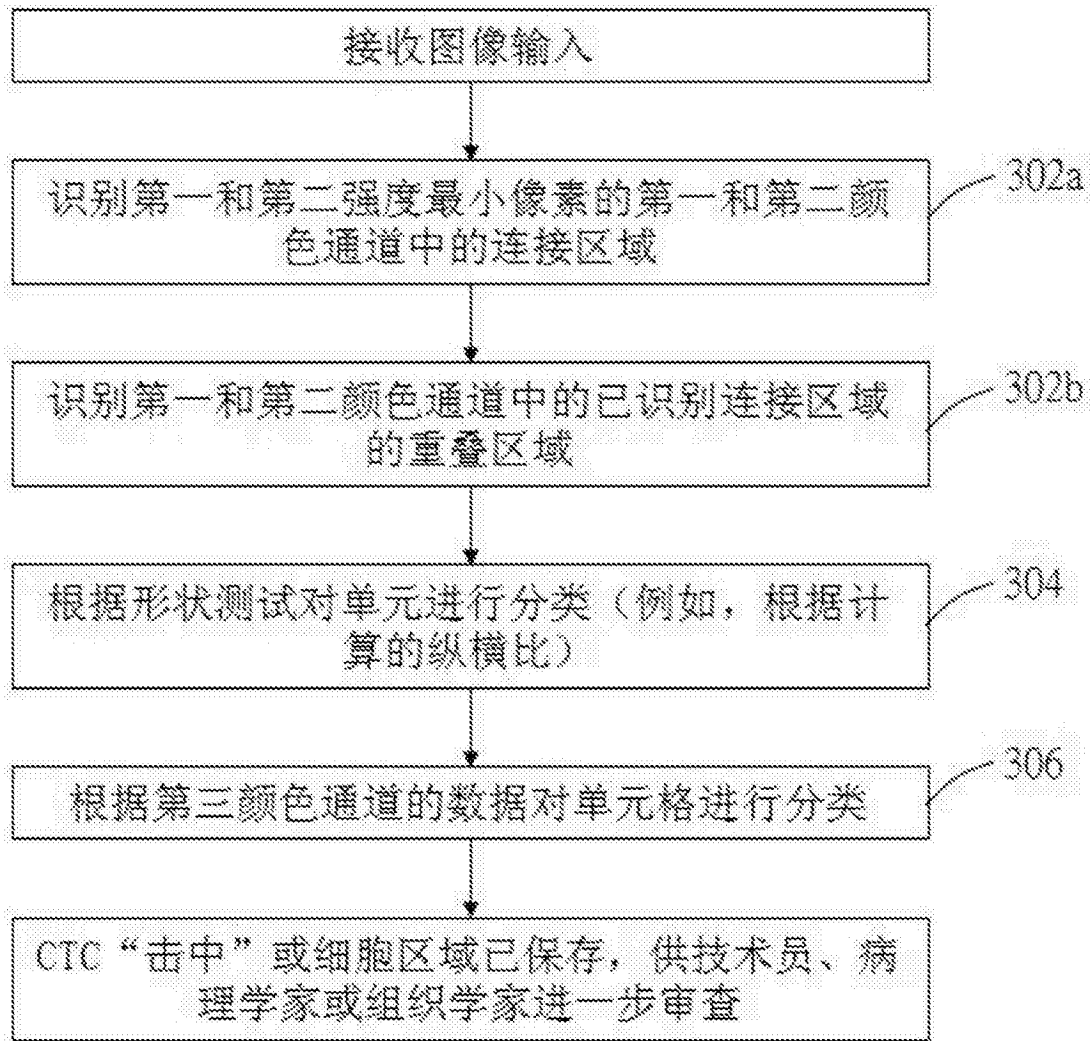


图4

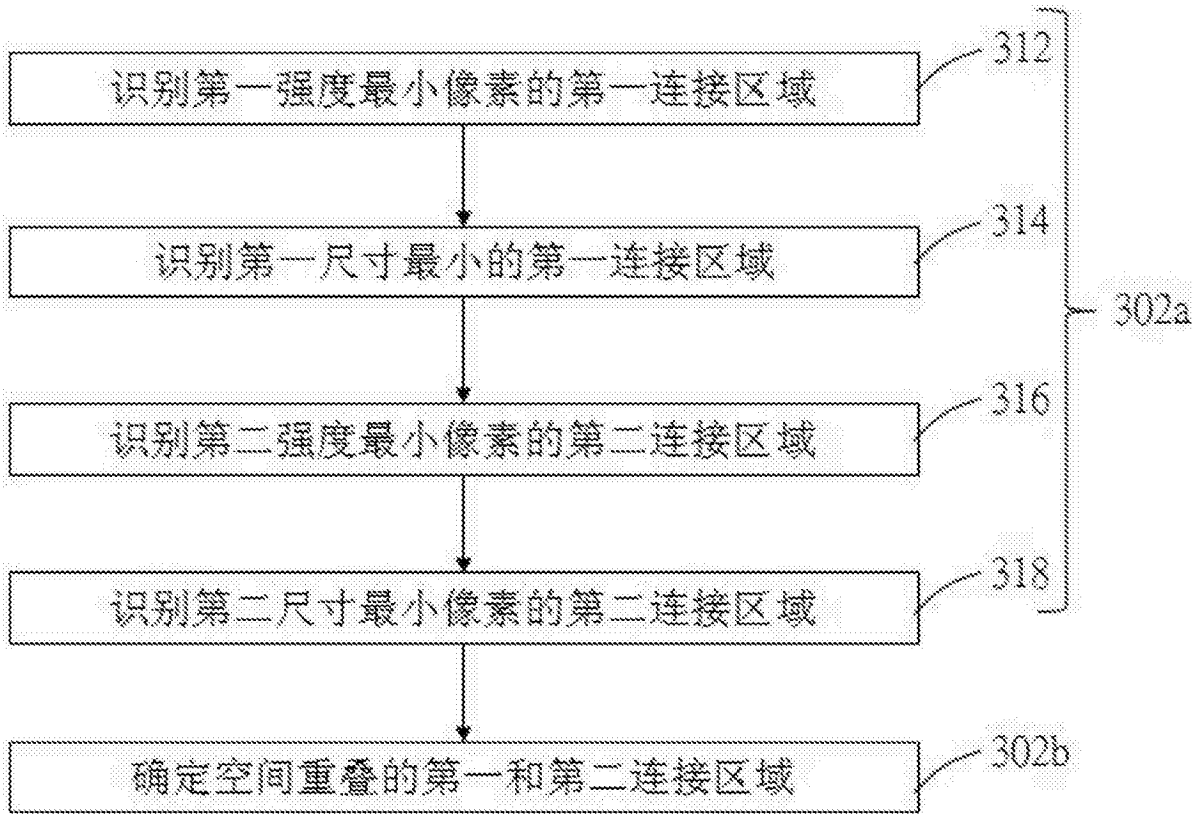


图5

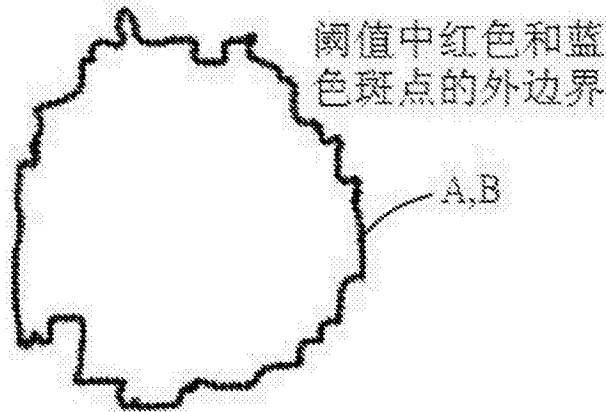


图6A

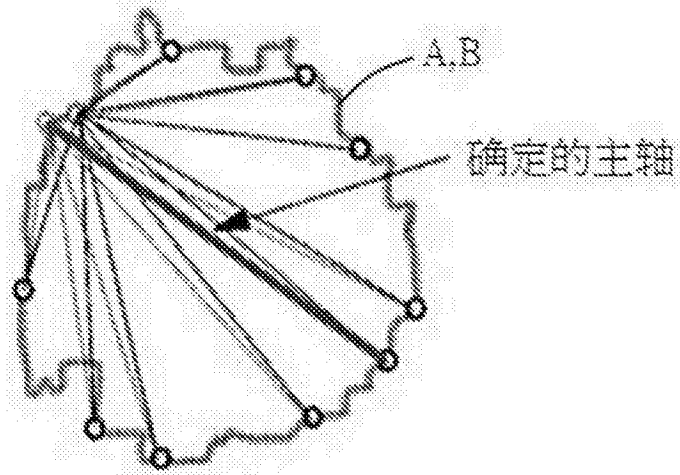


图6B

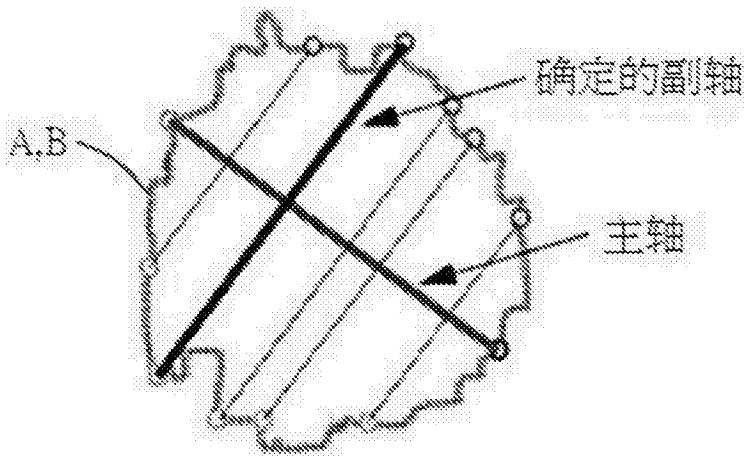


图6C

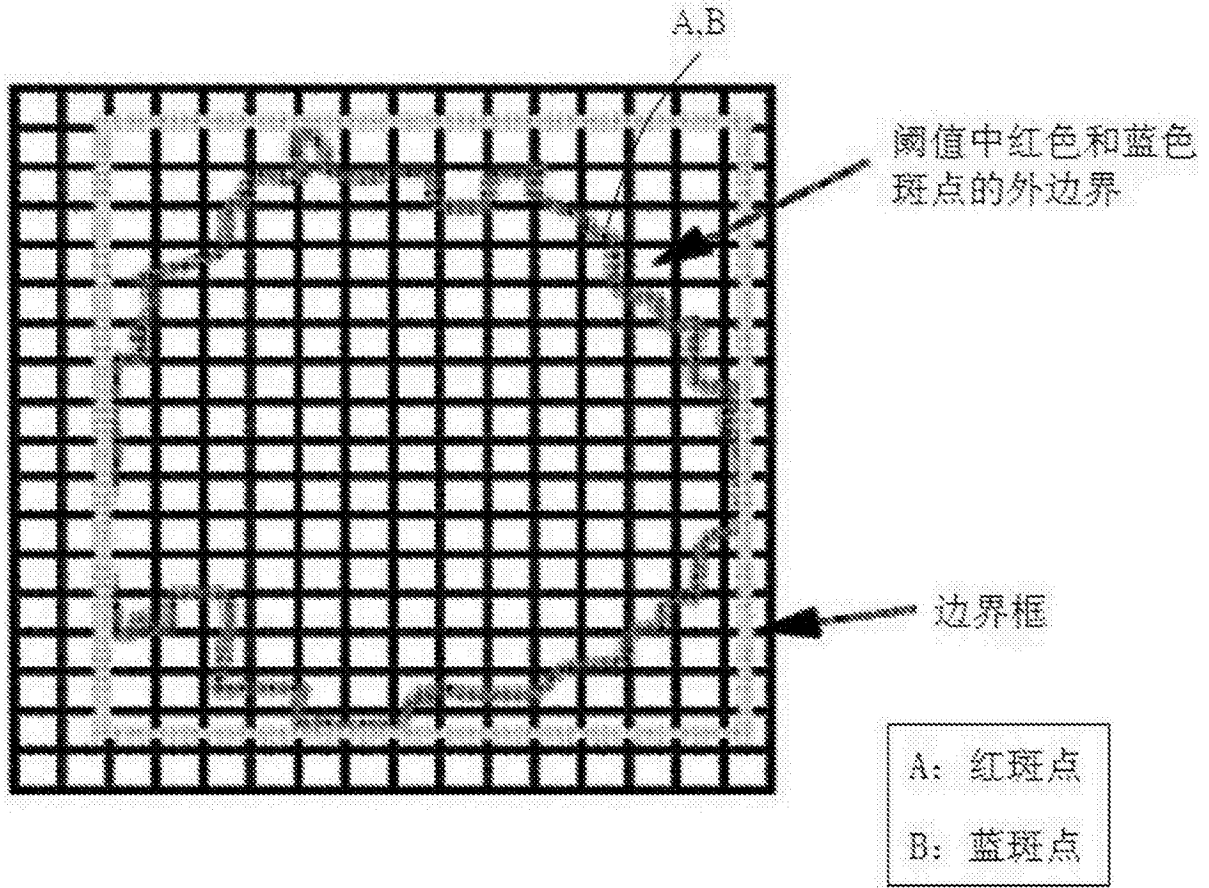


图7

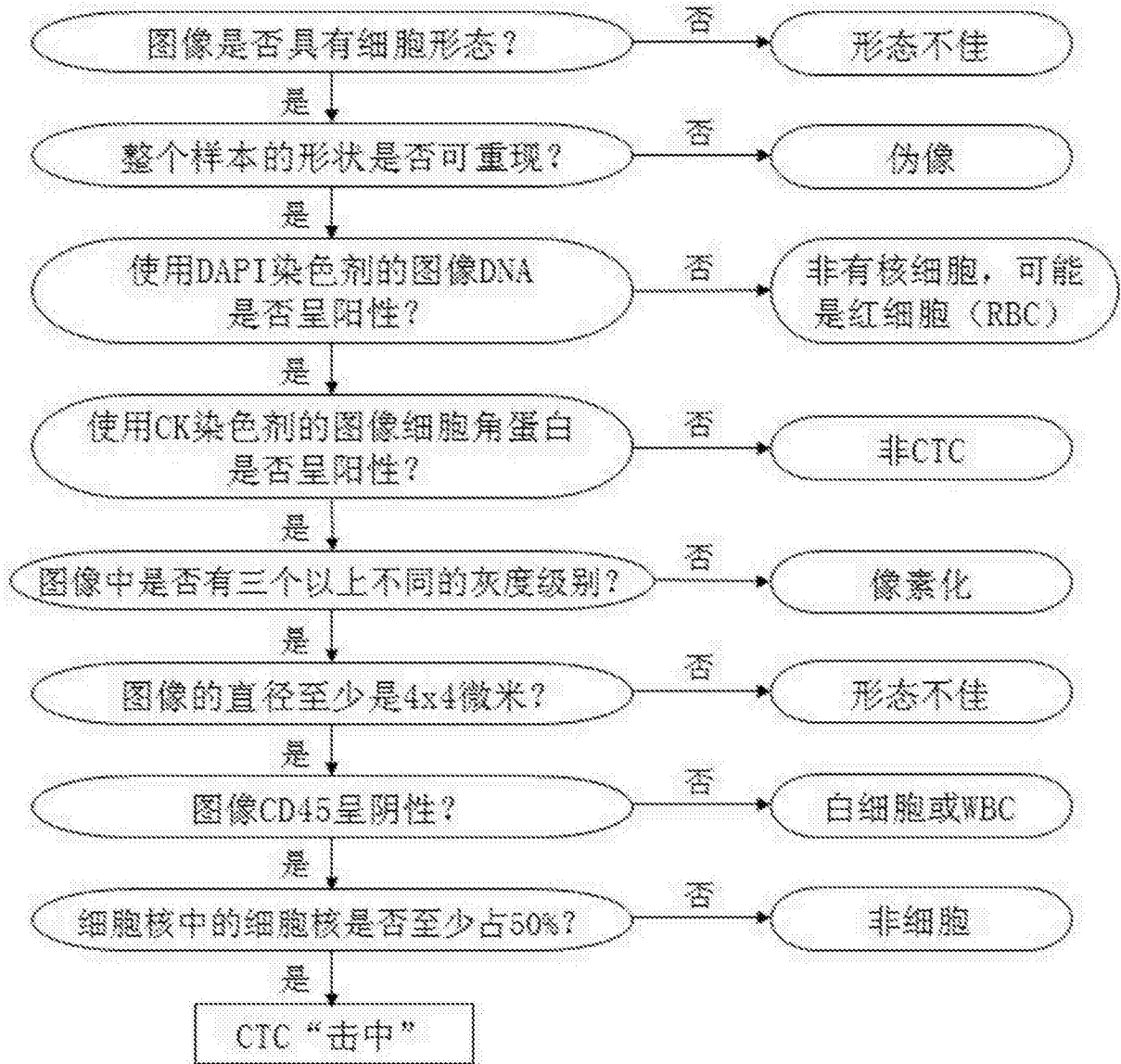


图8