



(21) 申请号 202080004716.2

专利权人 乔纳森·林

(22) 申请日 2020.02.27

(72) 发明人 乔纳森·林 基根·奥斯利

(65) 同一申请的已公布的文献号

申请公布号 CN 112955729 A

大卫·A·罗伯茨

(43) 申请公布日 2021.06.11

(74) 专利代理机构 北京柏杉松知识产权代理事

务所(普通合伙) 11413

(30) 优先权数据

专利代理师 罗小晨 刘继富

62/836,537 2019.04.19 US

(51) Int.Cl.

(85) PCT国际申请进入国家阶段日

G01N 15/14 (2024.01)

2021.02.24

G01N 1/18 (2006.01)

(86) PCT国际申请的申请数据

(56) 对比文件

PCT/US2020/020095 2020.02.27

US 2014336942 A1, 2014.11.13

(87) PCT国际申请的公布数据

US 2010070904 A1, 2010.03.18

W02020/214248 EN 2020.10.22

审查员 刘丹

(73) 专利权人 贝克顿·迪金森公司

地址 美国新泽西州

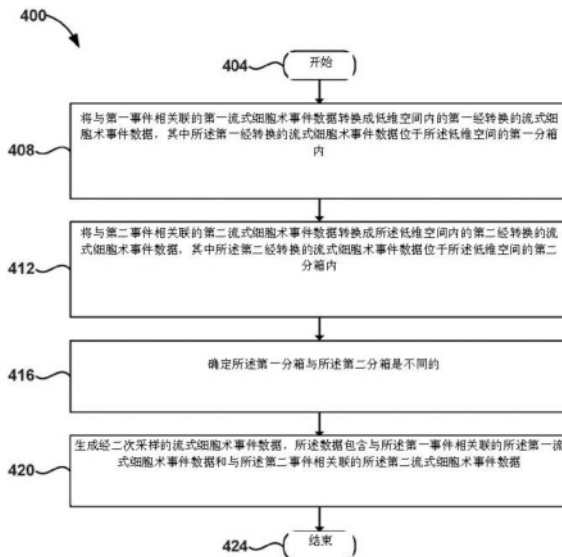
权利要求书8页 说明书21页 附图6页

(54) 发明名称

对流式细胞术事件数据进行二次采样

(57) 摘要

本发明包括用于对流式细胞术事件数据进行二次采样的系统、设备、计算机可读介质和方法。第一和第二流式细胞术事件数据可被转换成与多个分箱相关联的低维空间,且可被分配至第一分箱和第二分箱。本发明可以生成包含所述第一流式细胞术事件数据的经二次采样的流式细胞术事件数据。如果所述第一分箱与所述第二分箱有所不同,则所述经二次采样的流式细胞术事件数据可以包含所述第二流式细胞术事件数据。如果所述第一分箱与所述第二分箱是相同的,则所述经二次采样的流式细胞术事件数据可以不包含所述第二流式细胞术事件数据。



1. 一种用于对流式细胞术事件数据进行二次采样的方法,其包含:在控制器的控制下:

将高维空间内流式细胞术事件数据集中与第一多个事件中的第一事件相关联的第一流式细胞术事件数据转换成第一低维空间内与所述第一事件相关联的第一经转换的流式细胞术事件数据,其中所述第一事件与正二次采样需求相关联,其中所述第一低维空间与第一多个分箱相关联,且其中所述第一经转换的流式细胞术事件数据与所述第一多个分箱中的第一分箱相关联;

将所述高维空间内所述流式细胞术事件数据集中与所述第一多个事件中的第二事件相关联的第二流式细胞术事件数据转换成所述第一低维空间内与所述第二事件相关联的第二经转换的流式细胞术事件数据,其中所述第二事件与所述正二次采样需求相关联,且其中所述第二经转换的流式细胞术事件数据与所述第一多个分箱中的第二分箱相关联;

确定与所述第一经转换的流式细胞术事件数据相关联的所述第一分箱和与所述第二经转换的流式细胞术事件数据相关联的所述第二分箱是不同的;以及

生成所述流式细胞术事件数据的经二次采样的流式细胞术事件数据集,所述数据包含与所述第一事件相关联的所述第一流式细胞术事件数据和与所述第二事件相关联的所述第二流式细胞术事件数据。

2. 根据权利要求1所述的方法,其包含:接收包含所述第一流式细胞术事件数据和所述第二流式细胞术事件数据的流式细胞术事件数据。

3. 根据权利要求1-2中任一项所述的方法,其包含:

确定所述第一多个事件中所述第一事件的所述第一流式细胞术事件数据与所述正二次采样需求相关联;以及

确定所述第一多个事件中所述第二事件的所述第二流式细胞术事件数据与所述正二次采样需求相关联。

4. 根据权利要求1-2中任一项所述的方法,其包含:

确定所述第一经转换的流式细胞术事件数据与所述第一多个分箱中的所述第一分箱相关联;以及

确定所述第二经转换的流式细胞术事件数据与所述第一多个分箱中的所述第二分箱相关联。

5. 根据权利要求1-2中任一项所述的方法,其包含:

基于所述第一多个分箱中的所述第一分箱确定所述第一经转换的流式细胞术事件数据的第一描述符;以及

基于所述第一多个分箱中的所述第二分箱确定所述第二经转换的流式细胞术事件数据的第二描述符。

6. 根据权利要求5所述的方法,其中与所述第一分箱相关联的所述第一经转换的流式细胞术事件数据的所述第一描述符是所述第一多个分箱中所述第一分箱的第一分箱编号,且其中与所述第二分箱相关联的所述第二经转换的流式细胞术事件数据的所述第二描述符是所述第一多个分箱中所述第一分箱的第二分箱编号。

7. 根据权利要求1-2中任一项所述的方法,其中所述第一流式细胞术事件数据与第一稀有细胞相关联和/或所述第二流式细胞术事件数据与第二稀有细胞相关联,任选地,其中所述第一稀有细胞和所述第二稀有细胞是不同细胞类型的细胞。

8. 根据权利要求6所述的方法,其包含:

向内存数据结构中添加所述第一分箱、所述第一描述符和/或所述第一分箱编号;以及
向所述内存数据结构中添加所述第二分箱、所述第二描述符和/或所述第二分箱编号。

9. 根据权利要求8所述的方法,其包含:

将所述高维空间内所述流式细胞术事件数据集中与所述第一多个事件中的第三事件相关联的第三流式细胞术事件数据转换成所述第一低维空间内与所述第三事件相关联的第三经转换的流式细胞术事件数据,其中所述第三事件与所述正二次采样需求相关联,且其中所述第三经转换的流式细胞术事件数据与所述第一多个分箱中的第三分箱相关联;以及

确定与所述第三经转换的流式细胞术事件数据相关联的所述第三分箱是与所述第一经转换的流式细胞术事件数据相关联的所述第一分箱或与所述第二经转换的流式细胞术事件数据相关联的所述第二分箱,其中所述第三流式细胞术事件数据不在所述流式细胞术事件数据的所述经二次采样的流式细胞术事件数据中。

10. 根据权利要求9所述的方法,其包含:基于所述第一多个分箱中的所述第三分箱确定所述第三经转换的流式细胞术事件数据的第三描述符,其中与所述第三分箱相关联的所述第三经转换的流式细胞术事件数据的所述第三描述符是所述第一多个分箱中所述第三分箱的第三分箱编号。

11. 根据权利要求10所述的方法,其包含:确定所述第三分箱、所述第三描述符和/或所述第三分箱编号不在所述内存数据结构中。

12. 根据权利要求1-2中任一项所述的方法,其包含:

确定第四流式细胞术事件数据与负二次采样需求相关联,其中所述第四流式细胞术事件数据与所述第一多个事件中的第四事件相关联;以及

其中所述生成包含生成所述流式细胞术事件数据的所述经二次采样的流式细胞术事件数据集,所述事件数据包含与所述第四事件相关联的所述第四流式细胞术事件数据。

13. 根据权利要求12所述的方法,其包含:

接收定义多种目的细胞的多个门,其中所述第四流式细胞术事件数据与所述多种目的细胞中的目的细胞相关联。

14. 根据权利要求13所述的方法,其中所述第四流式细胞术事件数据与已分选细胞相关联。

15. 根据权利要求8所述的方法,其包含:

将所述高维空间内所述流式细胞术事件数据集中与第二多个事件中的第二事件相关联的第三流式细胞术事件数据转换成所述第一低维空间内与所述第二多个事件中的所述第二事件相关联的第三经转换的流式细胞术事件数据,其中所述第二多个事件中的所述第二事件与所述正二次采样需求相关联,其中所述第三经转换的流式细胞术事件数据与所述第一多个分箱中的第三分箱相关联,其中与所述第三经转换的流式细胞术事件数据相关联的所述第三分箱和与所述第一经转换的流式细胞术事件数据相关联的所述第一分箱是相同的,其中所述第三经转换的流式细胞术事件数据与所述第二多个事件中的所述第二事件相关联,且所述第一经转换的流式细胞术事件数据与所述第一多个事件中的所述第一事件相关联,

其中所述生成包含生成所述流式细胞术事件数据的所述经二次采样的流式细胞术事件数据集,所述事件数据包含与所述第二多个事件中的所述第二事件相关联的所述第三流式细胞术事件数据。

16. 根据权利要求15所述的方法,其包含:

确定所述第一多个事件中的最后事件与超过预定阈值的时间参数或事件数相关联;

重置所述内存数据结构;以及

向所述内存数据结构中添加与所述第三经转换的流式细胞术事件数据相关联的所述第三分箱,其中所述第三经转换的流式细胞术事件数据与所述第二多个事件中的所述第二事件相关联。

17. 根据权利要求16所述的方法,其包含:

接收二次采样参数水平;以及

基于所述二次采样参数水平确定所述预定阈值。

18. 根据权利要求1-2中任一项所述的方法,其中转换所述第一流式细胞术事件数据包含利用第一降维函数转换所述第一流式细胞术事件数据,和/或其中转换所述第二流式细胞术事件数据包含利用所述第一降维函数转换所述第二流式细胞术事件数据。

19. 根据权利要求18所述的方法,其中所述第一降维函数为线性降维函数。

20. 根据权利要求18所述的方法,其中所述第一降维函数为非线性降维函数。

21. 根据权利要求20所述的方法,其中所述非线性降维函数为t分布随机邻居嵌入。

22. 根据权利要求18所述的方法,其包含:接收所述第一降维函数或其标识。

23. 根据权利要求1-2中任一项所述的方法,

其中转换所述第一流式细胞术事件数据包含利用第二降维函数将所述第一流式细胞术事件数据转换成第二低维空间内与所述第一事件相关联的第一经转换的流式细胞术事件数据,其中所述第二低维空间与第二多个分箱相关联,且其中所述第二低维空间内的所述第一经转换的流式细胞术事件数据与所述第二多个分箱中的第一分箱相关联,和/或

其中转换所述第二流式细胞术事件数据包含利用所述第二降维函数将所述第二流式细胞术事件数据转换成所述第二低维空间内与所述第二事件相关联的第二经转换的流式细胞术事件数据,其中所述第二低维空间内的所述第二经转换的流式细胞术事件数据与所述第二多个分箱中的第二分箱相关联。

24. 根据权利要求23所述的方法,

其中所述第一多个分箱中的所述第一分箱与第一目的细胞类型相关联,

其中所述第二多个分箱中的所述第二分箱与第二目的细胞类型相关联,

其中所述第一多个分箱中的所述第二分箱与所述第一目的细胞类型无关联,其中所述第一多个分箱中的所述第二分箱与所述第二目的细胞类型无关联,

其中所述第二多个分箱中的所述第一分箱与所述第二目的细胞类型无关联,和/或其中所述第二多个分箱中的所述第一分箱与所述第一目的细胞类型无关联。

25. 根据权利要求24所述的方法,其中所述第一多个分箱中的所述第一分箱与所述第二多个分箱中的所述第一分箱的组合与第一目的细胞类型相关联,和/或其中所述第一多个分箱中的所述第二分箱与所述第二多个分箱中的所述第二分箱的组合与第二目的细胞类型相关联。

26. 根据权利要求25所述的方法, 其中所述第一多个分箱中的所述第一分箱与所述第二多个分箱中的所述第二分箱的组合与所述第一目的细胞类型和所述第二目的细胞类型无关联, 和/或其中所述第一多个分箱中的所述第二分箱与所述第二多个分箱中的所述第一分箱的组合与所述第一目的细胞类型和所述第二目的细胞类型无关联。

27. 根据权利要求1-2中任一项所述的方法, 其中所述第一多个分箱中的两个分箱具有相同尺寸。

28. 根据权利要求1-2中任一项所述的方法, 其中所述第一多个分箱中的每个分箱均具有相同尺寸。

29. 根据权利要求1-2中任一项所述的方法, 其中所述第一多个分箱中的两个分箱具有不同尺寸。

30. 根据权利要求1-2中任一项所述的方法, 其中所述第一多个分箱中的两个分箱包含数量相同的经转换的流式细胞术事件数据。

31. 根据权利要求1-2中任一项所述的方法, 其中所述第一多个分箱中的每个分箱均包含数量相同的经转换的流式细胞术事件数据。

32. 根据权利要求1-2中任一项所述的方法, 包含确定所述第一多个分箱中每个分箱的尺寸。

33. 根据权利要求32所述的方法, 包含基于多个门确定所述第一多个分箱中每个分箱的所述尺寸。

34. 根据权利要求32所述的方法, 包含基于与多种目的细胞相关联的所述经转换的流式细胞术事件数据确定所述第一多个分箱中每个分箱的所述尺寸。

35. 一种对流式细胞术事件数据进行二次采样的计算系统, 其包含: 被配置成存储可执行指令的非暂时性存储器; 以及

与所述非暂时性存储器进行通信的处理器, 所述处理器被所述可执行指令编程为:

将高维空间内流式细胞术事件数据集中与第一多个事件中的第一事件相关联的第一流式细胞术事件数据转换成第一低维空间内与所述第一事件相关联的第一经转换的流式细胞术事件数据, 其中所述第一事件与正二次采样需求相关联,

其中所述第一低维空间与第一多个分箱相关联, 且其中所述第一经转换的流式细胞术事件数据与所述第一多个分箱中的第一分箱相关联;

将所述高维空间内与所述第一多个事件中的第二事件相关联的第二流式细胞术事件数据转换成所述第一低维空间内所述流式细胞术事件数据集中与所述第二事件相关联的第二经转换的流式细胞术事件数据, 其中所述第二事件与所述正二次采样需求相关联, 且其中所述第二经转换的流式细胞术事件数据与所述第一多个分箱中的第二分箱相关联;

确定与所述第一经转换的流式细胞术事件数据相关联的所述第一分箱和与所述第二经转换的流式细胞术事件数据相关联的所述第二分箱是不同的; 以及

生成所述流式细胞术事件数据的经二次采样的流式细胞术事件数据集, 所述数据包含与所述第一事件相关联的所述第一流式细胞术事件数据和与所述第二事件相关联的所述第二流式细胞术事件数据。

36. 根据权利要求35所述的系统, 其中所述处理器被所述可执行指令编程为: 接收包含所述第一流式细胞术事件数据和所述第二流式细胞术事件数据的流式细胞术事件数据。

37. 根据权利要求35-36中任一项所述的系统,其中所述处理器被所述可执行指令编程为:

确定所述第一多个事件中所述第一事件的所述第一流式细胞术事件数据与所述正二次采样需求相关联;以及

确定所述第一多个事件中所述第二事件的所述第二流式细胞术事件数据与所述正二次采样需求相关联。

38. 根据权利要求35-36中任一项所述的系统,其中所述处理器被所述可执行指令编程为:

确定所述第一经转换的流式细胞术事件数据与所述第一多个分箱中的所述第一分箱相关联;以及

确定所述第二经转换的流式细胞术事件数据与所述第一多个分箱中的所述第二分箱相关联。

39. 根据权利要求35-36中任一项所述的系统,其中所述处理器被所述可执行指令编程为:

基于所述第一多个分箱中的所述第一分箱确定所述第一经转换的流式细胞术事件数据的第一描述符;以及

基于所述第一多个分箱中的所述第二分箱确定所述第二经转换的流式细胞术事件数据的第二描述符。

40. 根据权利要求39所述的系统,其中与所述第一分箱相关联的所述第一经转换的流式细胞术事件数据的所述第一描述符是所述第一多个分箱中所述第一分箱的第一分箱编号,且其中与所述第二分箱相关联的所述第二经转换的流式细胞术事件数据的所述第二描述符是所述第一多个分箱中所述第一分箱的第二分箱编号。

41. 根据权利要求35-36中任一项所述的系统,其中所述第一流式细胞术事件数据与第一稀有细胞相关联和/或所述第二流式细胞术事件数据与第二稀有细胞相关联,任选地,其中所述第一稀有细胞和所述第二稀有细胞是不同细胞类型的细胞。

42. 根据权利要求40所述的系统,其中所述处理器被所述可执行指令编程为:

向内存数据结构中添加所述第一分箱、所述第一描述符和/或所述第一分箱编号;以及向所述内存数据结构中添加所述第二分箱、所述第二描述符和/或所述第二分箱编号。

43. 根据权利要求42所述的系统,其中所述处理器被所述可执行指令编程为:

将所述高维空间内所述流式细胞术事件数据集中与所述第一多个事件中的第三事件相关联的第三流式细胞术事件数据转换成所述第一低维空间内与所述第三事件相关联的第三经转换的流式细胞术事件数据,其中所述第三事件与所述正二次采样需求相关联,且其中所述第三经转换的流式细胞术事件数据与所述第一多个分箱中的第三分箱相关联;以及

确定与所述第三经转换的流式细胞术事件数据相关联的所述第三分箱是与所述第一经转换的流式细胞术事件数据相关联的所述第一分箱或与所述第二经转换的流式细胞术事件数据相关联的所述第二分箱,

其中所述第三流式细胞术事件数据不在所述流式细胞术事件数据的所述经二次采样的流式细胞术事件数据中。

44. 根据权利要求43所述的系统,其中所述处理器被所述可执行指令编程为:基于所述第一多个分箱中的所述第三分箱确定所述第三经转换的流式细胞术事件数据的第三描述符,其中与所述第三分箱相关联的所述第三经转换的流式细胞术事件数据的所述第三描述符是所述第一多个分箱中所述第三分箱的第三分箱编号。

45. 根据权利要求44所述的系统,其中所述处理器被所述可执行指令编程为:确定所述第三分箱、所述第三描述符和/或所述第三分箱编号不在所述内存数据结构中。

46. 根据权利要求35-36中任一项所述的系统,其中所述处理器被所述可执行指令编程为:

确定第四流式细胞术事件数据与负二次采样需求相关联,其中所述第四流式细胞术事件数据与所述第一多个事件中的第四事件相关联,

其中为了生成所述经二次采样的流式细胞术事件数据集,所述处理器被所述可执行指令编程为:生成所述流式细胞术事件数据的所述经二次采样的流式细胞术事件数据集,所述事件数据包含与所述第四事件相关联的所述第四流式细胞术事件数据。

47. 根据权利要求46所述的系统,其中所述处理器被所述可执行指令编程为:

接收定义多种目的细胞的多个门,其中所述第四流式细胞术事件数据与所述多种目的细胞中的目的细胞相关联。

48. 根据权利要求47所述的系统,其中所述第四流式细胞术事件数据与已分选细胞相关联。

49. 根据权利要求42所述的系统,其中所述处理器被所述可执行指令编程为:

将所述高维空间内所述流式细胞术事件数据集中与第二多个事件中的第二事件相关联的第三流式细胞术事件数据转换成所述第一低维空间内与所述第二多个事件中的所述第二事件相关联的第三经转换的流式细胞术事件数据,

其中所述第二多个事件中的所述第二事件与所述正二次采样需求相关联,其中所述第三经转换的流式细胞术事件数据与所述第一多个分箱中的第三分箱相关联,其中与所述第三经转换的流式细胞术事件数据相关联的所述第三分箱和与所述第一经转换的流式细胞术事件数据相关联的所述第一分箱是相同的,其中所述第三经转换的流式细胞术事件数据与所述第二多个事件中的所述第二事件相关联,且所述第一经转换的流式细胞术事件数据与所述第一多个事件中的所述第一事件相关联,

其中为了生成所述经二次采样的流式细胞术事件数据集,所述处理器被所述可执行指令编程为:生成所述流式细胞术事件数据的所述经二次采样的流式细胞术事件数据集,所述事件数据包含与所述第二多个事件中的所述第二事件相关联的所述第三流式细胞术事件数据。

50. 根据权利要求49所述的系统,其中所述处理器被所述可执行指令编程为:

确定所述第一多个事件中的最后事件与超过预定阈值的时间参数或事件数相关联;
重置所述内存数据结构;以及

向所述内存数据结构中添加与所述第三经转换的流式细胞术事件数据相关联的所述第三分箱,其中所述第三经转换的流式细胞术事件数据与所述第二多个事件中的所述第二事件相关联。

51. 根据权利要求50所述的系统,其中所述处理器被所述可执行指令编程为:

接收二次采样参数水平;以及

基于所述二次采样参数水平确定所述预定阈值。

52. 根据权利要求35-36中任一项所述的系统,

其中为了转换所述第一流式细胞术事件数据,所述处理器被所述可执行指令编程为:利用第一降维函数转换所述第一流式细胞术事件数据,和/或

其中为了转换所述第二流式细胞术事件数据,所述处理器被所述可执行指令编程为:利用所述第一降维函数转换所述第二流式细胞术事件数据。

53. 根据权利要求52所述的系统,其中所述第一降维函数为线性降维函数。

54. 根据权利要求52所述的系统,其中所述第一降维函数为非线性降维函数。

55. 根据权利要求54所述的系统,其中所述非线性降维函数为t分布随机邻居嵌入。

56. 根据权利要求52所述的系统,其中所述处理器被所述可执行指令编程为:接收所述第一降维函数或其标识。

57. 根据权利要求35-36中任一项所述的系统,

其中为了转换所述第一流式细胞术事件数据,所述处理器被所述可执行指令编程为:利用第二降维函数将所述第一流式细胞术事件数据转换成第二低维空间内与所述第一事件相关联的第一经转换的流式细胞术事件数据,其中所述第二低维空间与第二多个分箱相关联,且其中所述第二低维空间内的所述第一经转换的流式细胞术事件数据与所述第二多个分箱中的第一分箱相关联,和/或

其中为了转换所述第二流式细胞术事件数据,所述处理器被所述可执行指令编程为:利用所述第二降维函数将所述第二流式细胞术事件数据转换成所述第二低维空间内与所述第二事件相关联的第二经转换的流式细胞术事件数据,其中所述第二低维空间内的所述第二经转换的流式细胞术事件数据与所述第二多个分箱中的第二分箱相关联。

58. 根据权利要求57所述的系统,

其中所述第一多个分箱中的所述第一分箱与第一目的细胞类型相关联,

其中所述第二多个分箱中的所述第二分箱与第二目的细胞类型相关联,

其中所述第一多个分箱中的所述第二分箱与所述第一目的细胞类型无关联,其中所述第一多个分箱中的所述第二分箱与所述第二目的细胞类型无关联,

其中所述第二多个分箱中的所述第一分箱与所述第二目的细胞类型无关联,和/或其中所述第二多个分箱中的所述第一分箱与所述第一目的细胞类型无关联。

59. 根据权利要求58所述的系统,其中所述第一多个分箱中的所述第一分箱与所述第二多个分箱中的所述第一分箱的组合与第一目的细胞类型相关联,和/或其中所述第一多个分箱中的所述第二分箱与所述第二多个分箱中的所述第二分箱的组合与第二目的细胞类型相关联。

60. 根据权利要求59所述的系统,其中所述第一多个分箱中的所述第一分箱与所述第二多个分箱中的所述第二分箱的组合与所述第一目的细胞类型和所述第二目的细胞类型无关联,和/或其中所述第一多个分箱中的所述第二分箱与所述第二多个分箱中的所述第一分箱的组合与所述第一目的细胞类型和所述第二目的细胞类型无关联。

61. 根据权利要求35-36中任一项所述的系统,其中所述第一多个分箱中的两个分箱具有相同尺寸。

62. 根据权利要求35-36中任一项所述的系统,其中所述第一多个分箱中的每个分箱均具有相同尺寸。

63. 根据权利要求35-36中任一项所述的系统,其中所述第一多个分箱中的两个分箱具有不同尺寸。

64. 根据权利要求35-36中任一项所述的系统,其中所述第一多个分箱中的两个分箱包含数量相同的经转换的流式细胞术事件数据。

65. 根据权利要求35-36中任一项所述的系统,其中所述第一多个分箱中的每个分箱均包含数量相同的经转换的流式细胞术事件数据。

66. 根据权利要求35-36中任一项所述的系统,其中所述处理器被所述可执行指令编程为:确定所述第一多个分箱中每个分箱的尺寸。

67. 根据权利要求66所述的系统,其中所述处理器被所述可执行指令编程为:基于多个门确定所述第一多个分箱中每个分箱的所述尺寸。

68. 根据权利要求66所述的系统,其中所述处理器被所述可执行指令编程为:基于与多种目的细胞相关联的所述经转换的流式细胞术事件数据确定所述第一多个分箱中每个分箱的所述尺寸。

69. 根据权利要求23所述的方法,其中所述第二降维函数是线性降维函数。

70. 根据权利要求23所述的方法,其中所述第二降维函数是非线性降维函数。

71. 根据权利要求70所述的方法,其中所述非线性降维函数是t分布随机邻居嵌入。

72. 根据权利要求57所述的系统,其中所述第二降维函数是线性降维函数。

73. 根据权利要求57所述的系统,其中所述第二降维函数是非线性降维函数。

74. 根据权利要求73所述的系统,其中所述非线性降维函数是t分布随机邻居嵌入。

对流式细胞术事件数据进行二次采样

技术领域

[0001] 本发明通常涉及自动颗粒评估技术领域,更具体而言,涉及样品分析和颗粒表征方法。

背景技术

[0002] 颗粒分析仪(例如,流式细胞仪)可基于光散射和荧光等电光测量值来表征颗粒。在流式细胞仪中,例如,流体悬浮液中的颗粒(例如分子、与分析物结合的微珠或个别细胞)通过检测区域,在所述区域内,所述颗粒暴露于通常来自一个或更多个激光器的激发光中,并且测定所述颗粒的光散射和荧光特性。颗粒或其组分通常用荧光染料标记以便进行检测。各种不同的颗粒或组分可以通过用光谱特性不同的荧光染料标记来同时加以检测。不同的细胞类型可以通过用荧光染料标记的抗体或其他荧光探针标记各种细胞蛋白或其他组分而得到/产生的光散射特性和荧光发射加以识别。利用多色流式细胞术进行细胞(或其他颗粒)分析所获得的数据是多维的,其中每个细胞对应于由所述被测参数定义的多维空间中的一个点。细胞群或颗粒群被识别为所述数据空间中的点的聚类。

发明内容

[0003] 本发明公开了用于对流式细胞术事件数据进行二次采样的系统、设备、计算机可读介质和方法。在一些实施例中,方法包含:在控制器的控制下:将高维空间内流式细胞术事件数据集中与第一多个事件中的第一事件相关联的第一流式细胞术事件数据转换成第一低维空间内与所述第一事件相关联的第一经转换的流式细胞术事件数据。所述第一事件可与正二次采样需求相关联。所述第一低维空间可与第一多个分箱相关联。

[0004] 所述第一经转换的流式细胞术事件数据可与所述第一多个分箱中的所述第一分箱相关联。所述方法可以包含:将所述高维空间内所述流式细胞术事件数据集中与第一多个事件中的第二事件相关联的第二流式细胞术事件数据转换成所述第一低维空间内与所述第二事件相关联的第二经转换的流式细胞术事件数据。所述第二事件可与所述正二次采样需求相关联。所述第二经转换的流式细胞术事件数据可与所述第一多个分箱中的第二分箱相关联。所述方法可以包含:确定与所述第一经转换的流式细胞术事件数据相关联的所述第一分箱和与所述第二经转换的流式细胞术事件数据相关联的所述第二分箱是不同的。所述方法可以包含:生成所述流式细胞术事件数据的经二次采样的流式细胞术事件数据集,所述数据集包含与所述第一事件相关联的所述第一流式细胞术事件数据和与所述第二事件相关联的所述第二流式细胞术事件数据。

[0005] 在一些实施例中,所述方法可以包含:接收包含所述第一流式细胞术事件数据和所述第二流式细胞术事件数据的流式细胞术事件数据。所述方法可以包含:确定所述第一多个事件中所述第一事件的所述第一流式细胞术事件数据与所述正二次采样需求相关联;和/或确定所述第一多个事件中所述第二事件的所述第二流式细胞术事件数据与所述正二次采样需求相关联。所述方法可以包含:确定所述第一经转换的流式细胞术事件数据与所

述第一多个分箱中的所述第一分箱相关联;和/或确定所述第二经转换的流式细胞术事件数据与所述第一多个分箱中的所述第二分箱相关联。所述方法可以包含:基于所述第一多个分箱中的所述第一分箱确定所述第一经转换的流式细胞术事件数据的第一描述符;和/或基于所述第一多个分箱中的所述第二分箱确定所述第二经转换的流式细胞术事件数据的第二描述符。与所述第一分箱相关联的所述第一经转换的流式细胞术事件数据的所述第一描述符可以是所述第一多个分箱中所述第一分箱的第一分箱编号;和/或与所述第二分箱相关联的所述第二经转换的流式细胞术事件数据的所述第二描述符可以是所述第一多个分箱中所述第一分箱的第二分箱编号。所述第一流式细胞术事件数据可与第一稀有细胞相关联和/或所述第二流式细胞术事件数据可与第二稀有细胞相关联。所述第一稀有细胞和所述第二稀有细胞可以是不同细胞类型的细胞。

[0006] 所述方法可以包含:向内存数据结构中添加所述第一分箱、所述第一描述符和/或所述第一分箱编号;和/或向所述内存数据结构中添加所述第二分箱、所述第二描述符和/或所述第二分箱编号。

[0007] 在一些实施例中,所述方法包含:将所述高维空间内所述流式细胞术事件数据集中与第一多个事件中的第三事件相关联的第三流式细胞术事件数据转换成所述第一低维空间内与所述第三事件相关联的第三经转换的流式细胞术事件数据。所述第三事件可与所述正二次采样需求相关联。所述第三经转换的流式细胞术事件数据可与所述第一多个分箱中的所述第三分箱相关联。所述方法可以包含:确定与所述第三经转换的流式细胞术事件数据相关联的所述第三分箱是与所述第一经转换的流式细胞术事件数据相关联的所述第一分箱或与所述第二经转换的流式细胞术事件数据相关联的所述第二分箱。所述第三流式细胞术事件数据可能不在所述流式细胞术事件数据的所述经二次采样的流式细胞术事件数据中。所述方法可以包含:基于所述第一多个分箱中的所述第三分箱确定所述第三经转换的流式细胞术事件数据的第三描述符。与所述第三分箱相关联的所述第三经转换的流式细胞术事件数据的所述第三描述符可以是所述第一多个分箱中所述第三分箱的第三分箱编号。所述方法可以包含:确定所述第三分箱、所述第三描述符和/或所述第三分箱编号不在所述内存数据结构中。

[0008] 在一些实施例中,所述方法包含:确定第四流式细胞术事件数据(其与所述第一多个事件中的第四事件相关联)与负二次采样需求相关联。所述生成可以包含:生成所述流式细胞术事件数据的所述经二次采样的流式细胞术事件数据集,所述事件数据包含与所述第四事件相关联的所述第四流式细胞术事件数据。所述方法可以包含:接收定义多种目的细胞的多个门,其中所述第四流式细胞术事件数据与所述多种目的细胞中的目的细胞相关联。所述第四流式细胞术事件数据与已分选细胞相关联。

[0009] 在一些实施例中,所述方法包含:将所述高维空间内所述流式细胞术事件数据集中与第二多个事件中的第二事件相关联的第二流式细胞术事件数据转换成所述第一低维空间内与所述第二多个事件中的所述第二事件相关联的第二经转换的流式细胞术事件数据。

[0010] 所述第二多个事件中的所述第二事件可与所述正二次采样需求相关联。所述第二经转换的流式细胞术事件数据(其与所述第二多个事件中的所述第二事件相关联)可与所述第一多个分箱中的第二分箱相关联。与所述第二经转换的流式细胞术事件数据(其与所

述第二多个事件中的所述第二事件相关联) 相关联的所述第二分箱和与所述第一经转换的流式细胞术事件数据 (其与所述第一多个事件中的所述第一事件相关联) 相关联的所述第一分箱是相同的。所述生成可以包含: 生成所述流式细胞术事件数据的所述经二次采样的流式细胞术事件数据集, 所述事件数据包含与所述第二多个事件中的所述第二事件相关联的所述第二流式细胞术事件数据。所述方法可以包含: 确定所述第一多个事件中的最后事件与超过预定阈值的时间参数或事件数相关联。所述方法可以包含: 重置所述内存数据结构。所述方法可以包含: 向所述内存数据结构中添加与所述第二经转换的流式细胞术事件数据 (其与所述第二多个事件中的所述第二事件相关联) 相关联的所述第二分箱。在一些实施例中, 所述方法可以包含: 接收二次采样参数水平。所述方法可以包含: 基于所述二次采样参数水平确定所述预定阈值。

[0011] 在一些实施例中, 转换所述第一流式细胞术事件数据包含利用第一降维函数转换所述第一流式细胞术事件数据。转换所述第二流式细胞术事件数据可以包含利用所述第一降维函数转换所述第二流式细胞术事件数据。所述第一降维函数和/或所述第二降维函数可以是线性降维函数。所述第一降维函数和/或所述第二降维函数可以是非线性降维函数。所述非线性降维函数可以是t分布随机邻居嵌入 (t-SNE)。所述方法可以包含: 首先接收所述降维函数或其标识。

[0012] 在一些实施例中, 转换所述第一流式细胞术事件数据包含利用第二降维函数将所述第一流式细胞术事件数据转换成第二低维空间内与所述第一事件相关联的第一经转换的流式细胞术事件数据。所述第二低维空间可与第二多个分箱相关联。

[0013] 所述第二低维空间内的所述第一经转换的流式细胞术事件数据可与所述第二多个分箱中的第一分箱相关联。转换所述第二流式细胞术事件数据可以包含利用所述第二降维函数将所述第二流式细胞术事件数据转换成所述第二低维空间内与所述第二事件相关联的第二经转换的流式细胞术事件数据。所述第二低维空间内的所述第二经转换的流式细胞术事件数据可与所述第二多个分箱中的第二分箱相关联。所述第一多个分箱中的所述第一分箱可与第一目的细胞类型相关联。所述第二多个分箱中的所述第二分箱可与第二目的细胞类型相关联。所述第一多个分箱中的所述第二分箱与所述第一目的细胞类型无关联。所述第一多个分箱中的所述第二分箱与所述第二目的细胞类型无关联。所述第二多个分箱中的所述第一分箱与所述第二目的细胞类型无关联。所述第二多个分箱中的所述第一分箱与所述第一目的细胞类型无关联。所述第一多个分箱中的所述第一分箱与所述第二多个分箱中的所述第一分箱的组合与第一目的细胞类型相关联。所述第一多个分箱中的所述第二分箱与所述第二多个分箱中的所述第二分箱的组合与第二目的细胞类型相关联。所述第一多个分箱中的所述第一分箱与所述第二多个分箱中的所述第二分箱的组合与所述第一目的细胞类型和所述第二目的细胞类型无关联。所述第一多个分箱中的所述第二分箱与所述第二多个分箱中的所述第一分箱的组合与所述第一目的细胞类型和所述第二目的细胞类型无关联。

[0014] 在一些实施例中, 所述第一多个分箱中的两个分箱具有相同尺寸。所述第一多个分箱中的每个分箱可具有相同尺寸。所述第一多个分箱中的两个分箱可具有不同尺寸。所述第一多个分箱中的两个分箱可以包含数量 (大致) 相同的经转换的流式细胞术事件数据。所述第一多个分箱中的每个分箱可以包含数量大致相同的经转换的流式细胞术事件数据。

所述方法可以包含：确定所述第一多个分箱中每个分箱的尺寸。所述方法可以包含：基于多个门确定所述第一多个分箱中每个分箱的所述尺寸。所述方法可以包含：基于与多种目的细胞相关联的所述经转换的流式细胞术事件数据确定所述第一多个分箱中每个分箱的所述尺寸。

[0015] 本发明包括用于对流式细胞术事件数据进行二次采样的计算系统的实施例。在一些实施例中，所述计算系统可以包含：被配置成存储可执行指令的非暂时性存储器；以及与所述非暂时性存储器进行通信的处理器（例如，硬件处理器或虚拟处理器），所述处理器被所述可执行指令编程为：将高维空间内与第一多个事件中的第一事件相关联的第一流式细胞术事件数据转换成第一低维空间内流式细胞术事件数据集中与所述第一事件相关联的第一经转换的流式细胞术事件数据，其中所述第一事件与正二次采样需求相关联，其中所述第一低维空间与第一多个分箱相关联，且其中所述第一经转换的流式细胞术事件数据与所述第一多个分箱中的第一分箱相关联。所述处理器可被所述可执行指令编程为：将所述高维空间内所述流式细胞术事件数据集中与所述第一多个事件中的第二事件相关联的第二流式细胞术事件数据转换成所述第一低维空间内与所述第二事件相关联的第二经转换的流式细胞术事件数据，其中所述第二事件与所述正二次采样需求相关联，且其中所述第二经转换的流式细胞术事件数据与所述第一多个分箱中的第二分箱相关联。所述处理器可被所述可执行指令编程为：确定与所述第一经转换的流式细胞术事件数据相关联的所述第一分箱和与所述第二经转换的流式细胞术事件数据相关联的所述第二分箱是不同的。所述处理器可被所述可执行指令编程为：生成所述流式细胞术事件数据的经二次采样的流式细胞术事件数据集，所述数据包含与所述第一事件相关联的所述第一流式细胞术事件数据和与所述第二事件相关联的所述第二流式细胞术事件数据。

[0016] 在一些实施例中，所述处理器被所述可执行指令编程为：接收包含所述第一流式细胞术事件数据和所述第二流式细胞术事件数据的流式细胞术事件数据。所述处理器可被所述可执行指令编程为：确定所述第一多个事件中所述第一事件的所述第一流式细胞术事件数据与所述正二次采样需求相关联。所述处理器可被所述可执行指令编程为：确定所述第一多个事件中所述第二事件的所述第二流式细胞术事件数据与所述正二次采样需求相关联。所述处理器可被所述可执行指令编程为：

[0017] 确定所述第一经转换的流式细胞术事件数据与所述第一多个分箱中的所述第一分箱相关联。所述处理器可被所述可执行指令编程为：确定所述第二经转换的流式细胞术事件数据与所述第一多个分箱中的所述第二分箱相关联。

[0018] 在一些实施例中，所述处理器被所述可执行指令编程为：基于所述第一多个分箱中的所述第一分箱确定所述第一经转换的流式细胞术事件数据的第一描述符。所述处理器可被所述可执行指令编程为：基于所述第一多个分箱中的所述第二分箱确定所述第二经转换的流式细胞术事件数据的第二描述符。与所述第一分箱相关联的所述第一经转换的流式细胞术事件数据的所述第一描述符可以是所述第一多个分箱中所述第一分箱的第一分箱编号；和/或与所述第二分箱相关联的所述第二经转换的流式细胞术事件数据的所述第二描述符可以是所述第一多个分箱中所述第一分箱的第二分箱编号。所述第一流式细胞术事件数据与第一稀有细胞相关联和/或所述第二流式细胞术事件数据可与第二稀有细胞相关联。所述第一稀有细胞和所述第二稀有细胞可以是不同细胞类型的细胞。

[0019] 在一些实施例中,所述处理器被所述可执行指令编程为:向内存数据结构中添加所述第一分箱、所述第一描述符和/或所述第一分箱编号;和/或向所述内存数据结构中添加所述第二分箱、所述第二描述符和/或所述第二分箱编号。在一些实施例中,所述处理器被所述可执行指令编程为:将所述高维空间内所述流式细胞术事件数据集中与第一多个事件中的第三事件相关联的第三流式细胞术事件数据转换成所述第一低维空间内与所述第三事件相关联的第三经转换的流式细胞术事件数据。所述第三事件可与所述正二次采样需求相关联。所述第三经转换的流式细胞术事件数据可与所述第一多个分箱中的所述第三分箱相关联。所述处理器可被所述可执行指令编程为:确定与所述第三经转换的流式细胞术事件数据相关联的所述第三分箱是与所述第一经转换的流式细胞术事件数据相关联的所述第一分箱或与所述第二经转换的流式细胞术事件数据相关联的所述第二分箱。所述第三流式细胞术事件数据可能不在所述流式细胞术事件数据的所述经二次采样的流式细胞术事件数据中。所述处理器可被所述可执行指令编程为:

[0020] 基于所述第一多个分箱中的所述第三分箱确定所述第三经转换的流式细胞术事件数据的第三描述符。与所述第三分箱相关联的所述第三经转换的流式细胞术事件数据的所述第三描述符可以是所述第一多个分箱中所述第三分箱的第三分箱编号。所述处理器可被所述可执行指令编程为:确定所述第三分箱、所述第三描述符和/或所述第三分箱编号不在所述内存数据结构中。

[0021] 在一些实施例中,所述处理器被所述可执行指令编程为:确定第四流式细胞术事件数据(其与所述第一多个事件中的第四事件相关联)与负二次采样需求相关联。为了生成所述经二次采样的流式细胞术事件数据集,所述处理器可被所述可执行指令编程为:生成所述流式细胞术事件数据的所述经二次采样的流式细胞术事件数据集,所述事件数据包含与所述第四事件相关联的所述第四流式细胞术事件数据。所述处理器可被所述可执行指令编程为:接收定义多种目的细胞的多个门。所述第四流式细胞术事件数据可与所述多种目的细胞中的目的细胞相关联。所述第四流式细胞术事件数据与已分选细胞相关联。

[0022] 在一些实施例中,所述处理器被所述可执行指令编程为:将所述高维空间内所述流式细胞术事件数据集中与第二多个事件中的第二事件相关联的第二流式细胞术事件数据转换成所述第一低维空间内与所述第二多个事件中的所述第二事件相关联的第二经转换的流式细胞术事件数据。所述第二多个事件中的所述第二事件可与所述正二次采样需求相关联。所述第二经转换的流式细胞术事件数据(其与所述第二多个事件中的所述第二事件相关联)可与所述第一多个分箱中的第二分箱相关联。与所述第二经转换的流式细胞术事件数据(其与所述第二多个事件中的所述第二事件相关联)相关联的所述第二分箱和与所述第一经转换的流式细胞术事件数据(其与所述第一多个事件中的所述第一事件相关联)相关联的所述第一分箱是相同的。为了生成所述经二次采样的流式细胞术事件数据集,所述处理器可被所述可执行指令编程为:生成所述流式细胞术事件数据的所述经二次采样的流式细胞术事件数据集,所述事件数据包含与所述第二多个事件中的所述第二事件相关联的所述第二流式细胞术事件数据。所述处理器可被所述可执行指令编程为:

[0023] 确定所述第一多个事件中的最后事件与超过预定阈值的时间参数或事件数相关联。所述处理器可被所述可执行指令编程为:重置所述内存数据结构。所述处理器可被所述可执行指令编程为:向所述内存数据结构中添加与所述第二经转换的流式细胞术事件数据

(其与所述第二多个事件中的所述第二事件相关联)相关联的所述第二分箱。

[0024] 在一些实施例中,所述处理器被所述可执行指令编程为:接收二次采样参数水平。所述处理器可被所述可执行指令编程为:基于所述二次采样参数水平确定所述预定阈值。

[0025] 在一些实施例中,为了转换所述第一流式细胞术事件数据,所述处理器可被所述可执行指令编程为:利用所述第一降维函数转换所述第一流式细胞术事件数据;和/或为了转换所述第二流式细胞术事件数据,所述处理器可被所述可执行指令编程为:利用所述第一降维函数转换所述第二流式细胞术事件数据。所述第一降维函数和/或所述第二降维函数可以是线性降维函数。所述第一降维函数和/或所述第二降维函数可以是非线性降维函数。所述非线性降维函数可以是t分布随机邻居嵌入(t-SNE)。所述处理器可被所述可执行指令编程为:首先接收所述降维函数或其标识。

[0026] 在一些实施例中,为了转换所述第一流式细胞术事件数据,所述处理器被所述可执行指令编程为:利用第二降维函数将所述第一流式细胞术事件数据转换成第二低维空间内与所述第一事件相关联的第一经转换的流式细胞术事件数据。所述第二低维空间可与所述第二多个分箱相关联。所述第二低维空间内的所述第一经转换的流式细胞术事件数据可与所述第二多个分箱中的第一分箱相关联。为了转换所述第二流式细胞术事件数据,所述处理器可被所述可执行指令编程为:利用所述第二降维函数将所述第二流式细胞术事件数据转换成所述第二低维空间内与所述第二事件相关联的第二经转换的流式细胞术事件数据。

[0027] 所述第二低维空间内的所述第二经转换的流式细胞术事件数据可与所述第二多个分箱中的第二分箱相关联。所述第一多个分箱中的所述第一分箱可与第一目的细胞类型相关联,所述第二多个分箱中的所述第二分箱可与第二目的细胞类型相关联,所述第一多个分箱中的所述第二分箱与所述第一目的细胞类型无关联,所述第一多个分箱中的所述第二分箱与所述第二目的细胞类型无关联,所述第二多个分箱中的所述第一分箱与所述第二目的细胞类型无关联,和/或所述第二多个分箱中的所述第一分箱与所述第一目的细胞类型无关联。所述第一多个分箱中的所述第一分箱与所述第二多个分箱中的所述第一分箱的组合可与第一目的细胞类型相关联,和/或所述第一多个分箱中的所述第二分箱与所述第二多个分箱中的所述第二分箱的组合可与第二目的细胞类型相关联。所述第一多个分箱中的所述第一分箱与所述第二多个分箱中的所述第二分箱的组合与所述第一目的细胞类型和所述第二目的细胞类型无关联,和/或所述第一多个分箱中的所述第二分箱与所述第二多个分箱中的所述第一分箱的组合与所述第一目的细胞类型和所述第二目的细胞类型无关联。

[0028] 在一些实施例中,所述第一多个分箱中的两个分箱具有相同尺寸。所述第一多个分箱中的每个分箱可具有相同尺寸。所述第一多个分箱中的两个分箱可具有不同尺寸。所述第一多个分箱中的两个分箱可以包含数量大致相同的经转换的流式细胞术事件数据。所述第一多个分箱中的每个分箱可以包含数量大致相同的经转换的流式细胞术事件数据。所述处理器可被所述可执行指令编程为:确定所述第一多个分箱中每个分箱的尺寸。所述处理器可被所述可执行指令编程为:基于多个门确定所述第一多个分箱中每个分箱的所述尺寸。所述处理器可被所述可执行指令编程为:基于与多种目的细胞相关联的所述经转换的流式细胞术事件数据确定所述第一多个分箱中每个分箱的所述尺寸。

附图说明

- [0029] 图1显示了用于分析和显示生物事件的分选控制系统的一个示例的功能框图。
- [0030] 图2A是根据本文所示的一个实施例的颗粒分选仪系统的示意图。
- [0031] 图2B是根据本文所示的一个实施例的另一颗粒分选仪系统的示意图。
- [0032] 图3显示了用于基于计算的样品分析和颗粒表征的颗粒分析系统的功能框图。
- [0033] 图4是显示对流式细胞术事件数据进行二次采样的示例性方法的流程图。
- [0034] 图5是被配置成实施对流式细胞术事件数据进行二次采样的方法的说明性计算系统。

具体实施方式

[0035] 在下文更为详细的描述中,参考了构成本文组成部分的附图。在附图中,除非上下文另外指出,否则类似符号通常标识类似组件。在详细描述、附图和权利要求中描述的说明性实施例并非旨在起限制作用。在不脱离本文所述主题的精神或范围的情况下,可使用其他实施例,且可做出其他变更。容易理解的是,本文通常所述的且在附图中示出的本发明的方面可以各种不同的配置来布置、替换、组合、分离和设计,所有这些都在本文中进行了明确的设想,且构成了本文公开内容的一部分。

[0036] 颗粒分析仪(例如,流式细胞仪和扫描式细胞仪)是一种分析工具,其可基于光散射和荧光等电光测量值来表征颗粒。在流式细胞仪中,例如,流体悬浮液中的颗粒(例如分子、与分析物结合的微珠或个别细胞)通过检测区域,在所述区域内,所述颗粒暴露于通常来自一个或更多个激光器的激发光中,并且测定所述颗粒的光散射和荧光特性。颗粒或其组分通常用荧光染料标记以便进行检测。各种不同的颗粒或组分可以通过用光谱特性不同的荧光染料标记来同时加以检测。

[0037] 在一些实施方案中,所述分析仪包括多个光电检测器,一个适用于待测定的每个散射参数,一个或更多个适用于待检测的每种不同染料。例如,一些实施例包括光谱结构,其中针对每种染料使用一个以上的传感器或检测器。所获得的数据包含针对所述光散射检测器和所述荧光发射中的每一项测得的所述信号。

[0038] 颗粒分析仪可以进一步包含用于记录所测得的数据并分析所述数据的装置。例如,可以使用与所述检测电子器件连接的计算机来进行数据存储和分析。例如,所述数据可以列表形式存储,其中每一行对应于一颗颗粒的数据,而所述列对应于每项被测特征。使用标准文件格式(例如,流式细胞术标准(“FCS”)文件格式)存储来自颗粒分析仪的数据有助于采用单独的程序和/或机器分析数据。采用当前分析方法时,所述数据通常以一维直方图或二维(2D)图的形式展示,以便于可视化,但也可采用其他方法来使多维数据可视化。

[0039] 例如,使用流式细胞仪测定的参数通常包括所述颗粒以狭角主要沿前向散射(被称为前向散射(FSC))的光;所述颗粒沿与所述激发激光器正交的方向散射(被称为侧向散射(SSC))的光;以及由一个或更多个检测器(用于在一定光谱波长范围内测量信号)中的荧光分子发射出的光,或者由主要在所述特定检测器或检测器阵列中检测到的荧光染料发射出的光。不同的细胞类型可以通过用荧光染料标记的抗体或其他荧光探针标记各种细胞蛋白或其他组分而得到/产生的光散射特性和荧光发射加以识别。

[0040] 流式细胞仪和扫描式细胞仪均可从诸如BD Biosciences(加利福尼亚州圣何塞)

等处购得。流式细胞术的相关描述请见以下出版物,例如,Landy等人(编),临床流式细胞术,纽约科学院年鉴,第677卷(1993年);Bauer等人(编),临床流式细胞术:原理与应用,Williams&Wilkins(1993年);Ormerod(编),流式细胞术:实用方法,牛津大学出版社(1994年);Jaroszeski等人(编),流式细胞术方案,Methods in Molecular Biology,第91期,Humana Press(1997年);以及Shapiro,实用流式细胞术,第4版,Wiley-Liss(2003年);这些出版物中的内容以引用方式并入本文。

[0041] 荧光成像显微镜检查的相关描述请见以下出版物,例如,Pawley(编),生物共聚焦显微镜检查手册,第2版,Plenum Press(1989年),其以引用方式并入本文。

[0042] 利用多色流式细胞术进行细胞(或其他颗粒)分析所获得的数据是多维的,其中每个细胞对应于由所述被测参数定义的多维空间中的一个点。细胞群或颗粒群被识别为所述数据空间中的点的聚类。可以通过在所述数据的一张或更多张二维图(被称为“散点图”或“点阵图”)中显示的某一群体周围绘制门来手动识别聚类和群体。或者,可以自动识别聚类,而且可以自动确定定义所述群体界限的门。用于自动设门的方法示例请见以下出版物,例如,编号如下的美国专利:4,845,653;5,627,040;5,739,000;5,795,727;5,962,238;6,014,904;

[0043] 6,944,338和8,990,047;这些专利中的内容以引用方式并入本文。

[0044] 流式细胞术是用于分析和分离生物颗粒(例如,细胞和组成分子)的有效方法,因此,它广泛应用于诊断和治疗中。所述方法利用流体介质线性分离颗粒,使所述颗粒可以排成一排,以通过检测装置。个别细胞可以根据其在流体介质中的位置以及是否存在可检测标记物来区分。因此,流式细胞仪可用于表征和生成生物颗粒群的诊断配置文件。

[0045] 生物颗粒的分离已通过使流式细胞仪具有分选或收集功能而实现。经检测具有一种或更多种所需特征的分离流中的颗粒通过机械或电气分离方式从所述样品流中单独分离出来。这种流式分选方法已用于分选不同类型的细胞,分离含有X和Y染色体的精子以进行动物育种,分选染色体以进行遗传分析,以及从复杂的生物种群中分离出特定的生物体。

[0046] 设门用于帮助理解可能由样品生成的大量数据并对其进行分类。鉴于一个给定样品所呈现出的大量数据,需要有效控制所述数据的图形显示。

[0047] 荧光激活颗粒分选或细胞分选是一种专用流式细胞术。

[0048] 荧光激活颗粒分选或细胞分选提供了一种将颗粒异质混合物分选至一个或更多个容器中的方法,所述方法基于各细胞的特定光散射和荧光特性实施,每次分选一种细胞。它记录来自个体细胞的荧光信号,并在物理上分离特定目的细胞。缩略词FACS是Becton, Dickinson and Company(新泽西州富兰克林湖)的商标且归Becton Dickinson所有,可用于指代执行荧光激活颗粒分选或细胞分选的设备。

[0049] 所述颗粒悬浮液置于狭窄、快速流动的液流中心附近。所述流的安置使得平均而言,在颗粒随机抵达(例如,泊松过程)所述检测区域时,相对于其直径,颗粒之间存在较大的间隔。振动机制可使流出的流体介质稳定破碎成单一液滴,所述液滴中包含先前在所述检测区域中表征的颗粒。通常可对所述系统进行调整,使液滴中存在一颗以上颗粒的概率偏低。如果将颗粒归类为待收集,则在一段时间内,可向所述流动池和流出流中施加电荷,以形成一滴或更多滴液滴并与所述流脱离。而后,这些带电液滴穿过静电偏转系统,所述系统依据向所述液滴施加的电荷,将液滴转移至目标容器中。

[0050] 一个样品可能包括(即使没有数百万也有)数千个细胞。可以对细胞进行分选以使样品纯化为目的细胞。所述分选过程通常可以识别出三种细胞:目的细胞、非目的细胞和无法识别的细胞。为了分选出具有高纯度的细胞(例如,高浓度的目的细胞),如果所需细胞与另一不合需求的细胞过于接近,则生成液滴的细胞分选仪可以电子方式中止所述分选,从而减少因无意间使含有目的颗粒的液滴内包括不合需求的颗粒而对已分选群体造成的污染。

[0051] 本发明公开了用于对流式细胞术事件数据进行二次采样的系统、设备、计算机可读介质和方法。在一些实施例中,方法包含:在控制器的控制下:将高维空间内与第一多个事件中的第一事件相关联的第一流式细胞术事件数据转换成第一低维空间内与所述第一事件相关联的第一经转换的流式细胞术事件数据。所述第一事件可与正二次采样需求相关联。

[0052] 所述第一低维空间可与第一多个分箱相关联。所述第一经转换的流式细胞术事件数据可与所述第一多个分箱中的所述第一分箱相关联。所述方法可以包含:将所述高维空间内与所述第一多个事件中的第二事件相关联的第二流式细胞术事件数据转换成所述第一低维空间内与所述第二事件相关联的第二经转换的流式细胞术事件数据。所述第二事件可与所述正二次采样需求相关联。所述第二经转换的流式细胞术事件数据可与所述第一多个分箱中的第二分箱相关联。所述方法可以包含:确定与所述第一经转换的流式细胞术事件数据相关联的所述第一分箱和与所述第二经转换的流式细胞术事件数据相关联的所述第二分箱是不同的。所述方法可以包含:生成所述流式细胞术事件数据的所述经二次采样的流式细胞术事件数据集,所述数据包含与所述第一事件相关联的所述第一流式细胞术事件数据和与所述第二事件相关联的所述第二流式细胞术事件数据。

[0053] 本发明包括用于对流式细胞术事件数据进行二次采样的计算系统的实施例。在一些实施例中,所述计算系统可以包含:被配置成存储可执行指令的非暂时性存储器;以及与所述非暂时性存储器进行通信的处理器(例如,硬件处理器或虚拟处理器),所述处理器被所述可执行指令编程为:将高维空间内与第一多个事件中的第一事件相关联的第一流式细胞术事件数据转换成第一低维空间内流式细胞术事件数据集中与所述第一事件相关联的第一经转换的流式细胞术事件数据,其中所述第一事件与正二次采样需求相关联,其中所述第一低维空间与第一多个分箱相关联,且其中所述第一经转换的流式细胞术事件数据与所述第一多个分箱中的第一分箱相关联。所述处理器可被所述可执行指令编程为:将所述高维空间内所述流式细胞术事件数据集中与所述第一多个事件中的第二事件相关联的第二流式细胞术事件数据转换成所述第一低维空间内与所述第二事件相关联的第二经转换的流式细胞术事件数据,其中所述第二事件与所述正二次采样需求相关联,且其中所述第二经转换的流式细胞术事件数据与所述第一多个分箱中的第二分箱相关联。

[0054] 所述处理器可被所述可执行指令编程为:确定与所述第一经转换的流式细胞术事件数据相关联的所述第一分箱和与所述第二经转换的流式细胞术事件数据相关联的所述第二分箱是不同的。所述处理器可被所述可执行指令编程为:生成所述流式细胞术事件数据的所述经二次采样的流式细胞术事件数据集,所述数据包含与所述第一事件相关联的所述第一流式细胞术事件数据和与所述第二事件相关联的所述第二流式细胞术事件数据。

[0055] 定义

[0056] 本文中使用的下文具体阐述的术语具有以下定义。除非本部分另有定义,否则本文中使用的所有术语的含义与本发明所属领域的普通技术人员通常理解的含义相同。

[0057] 本文中使用的“系统”、“仪器”、“装置”和“设备”通常涵盖硬件(例如,机械和电子硬件),而且在一些实施方案中,也涵盖相关软件(例如,用于图形控制的专用计算机程序)组件。

[0058] 本文中使用的“事件”或“事件数据”通常是指从单种颗粒(例如,细胞或合成颗粒)中测得的数据(例如,组合数据包)。通常情况下,从单种颗粒中测得的数据包括许多参数或特征,包括一个/种或更多个/种光散射参数或特征,以及至少一个/种利用从所述颗粒中检测到的荧光所得到的其他参数或特征,例如,所述荧光强度。因此,每个事件均可被表示为参数和特征测量值向量,其中每个被测参数或特征对应于数据空间的一个维度。在一些实施例中,从单种颗粒中测得的数据包括图像、电气数据、时间数据或声学数据。事件可与实验、测定或样品来源(可以通过测量数据来识别)相关联。

[0059] 本文中使用的颗粒“群体”或“子群体”(例如,细胞或其他颗粒)通常是指相对于一个或更多个被测参数具有特性(例如,光学、阻抗或时间特性)的一组颗粒,其使得被测参数数据在数据空间中形成聚类。因此,群体可被识别为数据中的聚类。相反,虽然通常也观测到了对应于噪点或本底的聚类,但每个数据聚类通常被解释为对应于特定类型的细胞或颗粒群。

[0060] 可以在所述维度的子集(例如,相对于被测参数的子集)中定义聚类,其对应于仅在被测参数或特征(提取自所述细胞或颗粒测量值)的子集中有所不同的群体。

[0061] 本文中使用的“门”通常是指识别目的数据子集的分类器边界。在细胞计量中,门可能与一组特定目的事件相关联。本文中使用的“设门”通常是指针对一个给定数据集,利用一个定义门对所述数据进行分类的过程,其中所述门可以是一个或更多个与布尔逻辑相结合的目的区域。

[0062] 下文进一步描述了各种实施例和系统的具体示例(可实现所述实施例和系统)。

[0063] 分选控制系统

[0064] 图1显示了用于分析和显示生物事件的分选控制系统(例如分析控制器100)的一个示例的功能框图。分析控制器100可被配置成实施用于控制生物事件图形显示的各种过程。

[0065] 颗粒分析仪或分选系统102可被配置成采集生物事件数据。例如,流式细胞仪可以生成流式细胞术事件数据。颗粒分析仪102可被配置成将生物事件数据提供给分析控制器100。可以在颗粒分析仪102和分析控制器100之间纳入一条数据通信通道。可以经由所述数据通信通道将所述生物事件数据提供给分析控制器100。

[0066] 分析控制器100可被配置成接收来自颗粒分析仪102的生物事件数据。接收自颗粒分析仪102的生物事件数据可以包括流式细胞术事件数据。分析控制器100可被配置成向显示设备106提供包括生物事件数据第一图的图形显示。例如,分析控制器100可进一步被配置成围绕由显示设备106所示的多个生物事件数据渲染作为门的目的区域,其覆盖在第一图上。在一些实施例中,所述门可以是在单参数直方图或双变量图上绘制的一个或更多个目的图形区域的逻辑组合。

[0067] 或者,分析控制器100可进一步被配置成以不同于门外生物事件数据中其他事件

的方式在门内的显示设备106上显示生物事件数据。例如,分析控制器100可被配置成使门内含有的生物事件数据的颜色与门外的生物事件数据的颜色不同。显示设备106可以显示器、平板计算机、智能电话或被配置成展示图形界面的其他电子设备的形式实现。

[0068] 分析控制器100可被配置成接收来自第一输入设备的识别所述门的门选择信号。例如,所述第一输入设备可以鼠标110的形式实现。鼠标110可以向分析控制器100发出门选择信号,以确定待在显示设备106上显示或经由显示设备106操纵的门(例如,当光标位于此处时,在所需门上或内单击)。在一些实施方案中,第一设备可以键盘108或用于向分析控制器100提供输入信号的其他装置(例如触摸屏、触控笔、光学检测器或语音识别系统)的形式实现。某些输入设备可能包括多种输入功能。在此等实施方案中,所述输入功能可以各自地被视为是一种输入设备。例如,如图1所示,鼠标110可以包括鼠标右键和鼠标左键,其均可产生触发事件。

[0069] 所述触发事件可以促使分析控制器100改变显示数据的方式(在显示设备106上实际显示部分数据),和/或提供输入以作进一步处理,例如选择目的群体进行颗粒分选。

[0070] 在一些实施例中,分析控制器100可被配置成检测鼠标110启动门选择的时间。分析控制器100可进一步被配置成自动修改图可视化,以便促进所述设门过程。所述修改可以基于由分析控制器100接收的生物事件数据的特定分布进行。

[0071] 分析控制器100可与存储设备104连接。存储设备104可被配置成接收和存储来自分析控制器100的生物事件数据。存储设备104还可被配置成接收和存储来自分析控制器100的流式细胞术事件数据。存储设备104还可被配置成允许分析控制器100检索生物事件数据,例如,流式细胞术事件数据。

[0072] 显示设备106可被配置成接收来自分析控制器100的显示数据。所述显示数据可以包含生物事件数据图和概述所述图部分的门。显示设备106可进一步被配置成根据接收自分析控制器100的输入以及接收自颗粒分析仪102、存储设备104、键盘108和/或鼠标110的输入来改变所展示的信息。

[0073] 在一些实施方案中,分析控制器100可以生成用户界面以接收示例事件,以便进行分选。例如,所述用户界面可以包括用于接收示例事件或示例图像的控件。可以在采集样品的事件数据之前或基于所述样品某一部分的初始事件集合来提供示例事件或图像或示例门。

[0074] 颗粒分选仪系统

[0075] 常用流式分选技术(被称为“静电细胞分选”)采用液滴分选,其中使含有线性分离颗粒的流或移动液柱破碎成液滴,且使含有目的颗粒的液滴带电并通过电场偏转至收集管中。液滴分选系统能够在通过直径小于100微米的喷嘴的流体介质中以100,000滴/秒的速率形成液滴。液滴分选通常要求液滴在与喷嘴头相距一定距离处与所述流脱离。所述距离通常距离喷嘴头约几毫米,而且对于不受干扰的流体介质,所述距离是稳定的,且可以通过以预先确定的频率以及使所述脱离保持恒定的振幅振荡喷嘴头来维持。例如,在一些实施例中,在给定频率下调整正弦波形电压脉冲的振幅,使所述脱离保持稳定和恒定。

[0076] 通常情况下,所述流中的线性分离颗粒的特征在于它们会通过位于流动池或比色皿内或喷嘴头正下方的观测点。一旦确定颗粒满足一项或更多项所需标准,便可以预测其到达液滴破碎点并以一滴液滴的形式与所述流脱离的时间。理想情况下,在含有选定颗粒

的液滴与所述流脱离之前,先对所述流体介质施加短暂的电荷,然后在液滴破碎后使其立即接地。

[0077] 待分选的液滴在与所述流体介质脱离时会保持带电状态,而所有其他液滴则不带电。所述带电液滴在电场作用下侧向偏离其他液滴的下行轨迹,并被收集在样品管中。所述不带电的液滴直接落入排放管。

[0078] 图2B是根据本文所示的一个实施例的颗粒分选仪系统200(例如,颗粒分析仪102)的示意图。在一些实施例中,颗粒分选仪系统200是细胞分选仪系统。如图2A所示,液滴形成传感器202(例如,压电振荡器)与流体导管201(其可与喷嘴203耦合,也可以包括或可以是喷嘴203)耦合。在流体导管201内,鞘液204以流体动力学的方式使样品流体206(包含颗粒209)聚集至移动液柱208(例如,流)中。在所述移动液柱208内,使颗粒209(例如,细胞)排成一排,以穿过由辐照源212(例如,激光器)辐照的受监测区域211(例如,激光流交点)。所述液滴形成传感器202振动使得移动液柱208破碎成多滴液滴210,其中一些含有颗粒209。

[0079] 在运行中,检测站214(例如,事件检测器)确定目的颗粒(或目的细胞)何时穿过受监测区域211。检测站214馈入定时电路228,转而馈入瞬时充电电路230。在液滴破碎点,在定时液滴延迟(Δt)通知后,可向所述移动液柱208中施加瞬时电荷,使目的液滴带电。所述目的液滴可以包括一颗/个或更多颗/个待分选颗粒或细胞。然后,可以通过启用偏转板(未示出)来分选所述带电液滴,使液滴偏转至诸如收集管或多孔或微孔样品板等容器中,其中孔或微孔可与特定目的液滴相关联。如图2A所示,所述液滴可被收集在排放容器238中。

[0080] 检测系统216(例如,液滴边界检测器)用于在目的颗粒通过受监测区域211时自动确定液滴驱动信号的相位。示例性液滴边界检测器如第7,679,039号美国专利所述,所述专利全文以引用方式并入本文。检测系统216允许仪器准确地计算每颗检测到的颗粒在液滴中的位置。检测系统216可以馈入振幅信号220和/或相位218信号,然后依次(通过放大器222)馈入振幅控制电路226和/或频率控制电路224。

[0081] 振幅控制电路226和/或频率控制电路224反过来控制液滴形成传感器202。振幅控制电路226和/或频率控制电路224可以包括在控制系统中。

[0082] 在一些实施方案中,分选电子器件(例如,检测系统216、检测站214、处理器240)可与被配置成存储所检测到的事件和基于其的分选决策的存储器耦合。所述分选决策可以包括在颗粒的事件数据中。在一些实施方案中,检测系统216和检测站214可以单个检测单元的形式实现或以通信方式耦合,使得事件测量值可以由检测系统216或检测站214中的一项收集并提供给所述非收集元件。

[0083] 图2B是根据本文所示的一个实施例的颗粒分选仪系统的示意图。图2B所示的颗粒分选仪系统200包括偏转板252和254。电荷通过倒钩中的流充电线施加。这产生了含有用于分析的颗粒210的液滴流210。所述颗粒可以用一个或更多个光源(例如,激光)照射,以产生光散射并生成荧光信息。所述颗粒信息通过诸如分选电子器件或其他检测系统(图2B中未示出)来进行分析。可以独立地控制偏转板252和254来吸引或排斥所述带电液滴,以将所述液滴引导至目的地收集容器中(例如,272、274、276或278中的一项)。如图2B所示,可以控制偏转板252和254,以沿着第一路径262朝向容器274或沿着第二路径268朝向容器278引导颗粒。如果所述颗粒并非目的颗粒(例如,在指定的分选范围内,不显示散射或照射信息),则偏转板可使所述颗粒沿着流路264继续流动。此类不带电液滴可以经由诸如抽吸器270进入

废弃物容器。

[0084] 可以包括所述分选电子器件,以开始收集测量值,接收颗粒的荧光信号,并确定如何调整所述偏转板以分选所述颗粒。图2B中所示的实施例的示例实施方案包括由Becton, Dickinson and Company(新泽西州富兰克林湖)市售的BD FACSAria™系列流式细胞仪。

[0085] 在一些实施例中,适用于颗粒分选仪系统200的一个或更多个所述组件可用于颗粒分析和表征,无论是否将所述颗粒物理分选至收集容器中。同样地,适用于颗粒分析系统300(图3)的一个或更多个下述组件也可用于颗粒分析和表征,无论是否将所述颗粒物理分选至收集容器中。例如,可以采用颗粒分选仪系统200或颗粒分析系统300中的一个或更多个所述组件,对颗粒进行分组或使其显示在树中,所述树包括至少三个本文所述的分组。

[0086] 图3显示了用于基于计算的样品分析和颗粒表征的颗粒分析系统的功能框图。在一些实施例中,颗粒分析系统300是流式系统。例如,图3所示的颗粒分析系统300可被配置成全部或部分执行本文所述的方法。颗粒分析系统300包括射流系统302。射流系统302可以包括样品管310和样品管内的移动液柱,或与所述样品管耦合,在所述样品管内,样品中的颗粒330(例如,细胞)沿共同样品路径320移动。

[0087] 颗粒分析系统300包括检测系统304,所述检测系统304被配置成在每颗颗粒沿所述共同样品路径穿过一个或更多个检测站时,收集来自每颗颗粒的信号。检测站308通常是指所述共同样品路径的受监测区域340。在一些实施方案中,检测可以包括在颗粒330穿过受监测区域340时,检测光或所述颗粒的一种或更多种其他特性。在图3中,显示了具有受监测区域340的检测站308。颗粒分析系统300的一些实施方案可以包括多个检测站。此外,某些检测站可能监测一个以上区域。

[0088] 为每个信号分配一个信号值,以构成每颗颗粒的数据点。如上所述,所述数据可以被称作事件数据。所述数据点可以是多维数据点,其包括针对颗粒的各所测特性的值。检测系统304被配置成在第一时间间隔内收集一系列所述数据点。

[0089] 颗粒分析系统300还包括控制系统306。控制系统306可以包括一个或更多个处理器、振幅控制电路226和/或频率控制电路224,具体如图2B所示。所示控制系统206在操作上可与射流系统302相关联。

[0090] 控制系统206可被配置成基于泊松分布以及在所述第一时间间隔内由检测系统304收集的数据点数量,针对所述第一时间间隔的至少一部分生成计算出的信号频率。控制系统306可进一步被配置成基于在所述第一时间间隔的所述部分内的数据点数量,生成实验信号频率。此外,控制系统306可将所述实验信号频率与计算出的信号频率或预先确定的信号频率进行比较。

[0091] 对流式细胞术事件数据进行二次采样

[0092] 本发明包括用于对数据集(例如,大型高维数据集)进行二次采样的系统、设备、计算机可读介质和方法,其能对稀有事件和群体进行加权,使得所述稀有事件和群体在所得子集中适当示出。在一些实施例中,在保存整个数据集不可取的情况下,保留所有群体(例如,包括稀有细胞和群体的所有群体)的数据集的子集可以在所述数据集或经二次采样的数据集的子集中示出。在一些实施例中,所述子集或经二次采样的数据集保留来自稀有子群体的代表性样品。在一些实施例中,可以在不舍弃稀有事件或目的事件(例如,对应于稀有细胞或目的细胞)的情况下对数据集进行二次采样。所述系统会自动检测稀有事件并予

以保存,同时更积极地舍弃常见事件。

[0093] 在一些实施例中,可以非随机地(例如,半随机地)对数据进行二次采样。可以选择所需的二次采样率,然后依次通过二次采样方法馈入所述数据。所述方法可以决定在单个事件基础上保存或舍弃事件(或与事件相关联的多维事件数据)。在不分析事件的整体分布的情况下即可舍弃事件的功能消除了保存和分析大量数据的需求。

[0094] 用户可以选择二次采样参数水平。所述二次采样参数水平可以确定所述算法“内存”的持续时间。

[0095] 用户可以选择一种或更多种转换或“指纹识别”函数。转换或指纹识别函数可以是某种方式,例如将所述数据从高维空间数据转换为低维空间数据的数学方程式。例如,转换或指纹识别函数可以是t分布随机邻居嵌入(t-SNE)。可以将事件转换为低维空间事件,所述空间被划分为分箱。所述分箱编号可以作为所述事件的描述符。

[0096] 在一些实施例中,分箱可以是一致的,也可以基于事件密度进行。分箱可以基于自动群体检测进行。分箱可以部分基于任意绘制的门(例如,用户绘制的门)进行。在一些实施例中,转换或指纹识别函数可以转换事件,使得类似事件具有相同的标识符。标识符可以小于用于生成所述标识符的数据。就计算方面而言,所述转换计算可能较为便宜。所述转换或函数的逆转可能存在也可能不存在。在一些实施例中,可以使用多种指纹识别函数。例如,可以使用不同的指纹识别函数定义不同的目标群体。作为另一示例,可以基于多种指纹识别函数的组合输出来定义目标群体。

[0097] 第三,用户可以描述不应被二次采样的事件。例如,目的区域周围的门可以自动绘制或由用户绘制。可能不会对目的区域周围的门内的事件进行二次采样。作为另一示例,可能不会对任何被分选的事件(例如,未被分选的细胞)进行二次采样。

[0098] 可以使用本文所揭示的二次采样方法对所述事件数据进行二次采样。例如,对于每个事件:

[0099] ii. 检查是否应对所述事件进行二次采样。如果答案为否,则保存所述事件。

[0100] iii. 如果应对所述事件进行二次采样,则使用所述指纹识别函数生成描述符。

[0101] 1. 将描述符与算法“内存”进行比较。以前是否见过此描述符?

[0102] a. 见过。舍弃事件

[0103] b. 未见过。保存事件并将描述符保存在内存中。

[0104] iv. 检查时间或事件数。如果所述时间和/或事件数超过基于用户的二次采样参数水平所生成的相应阈值,则重置算法存储器。

[0105] 本文所揭示的二次采样方法(一种非随机二次采样方法)可以是用于对大型数据集进行二次采样的补足性或补充性随机二次采样方法。对数据进行随机采样时,可以消除稀有群体。所述二次采样方法可以包括一些、多数或全部稀有群体。对于颗粒分析(例如流式细胞术分析),稀有事件可能极具价值。保留稀有群体可能是有用的,以便在分析经缩减的数据集时检测稀有群体。所述非随机二次采样方法可以有意偏离随机采样过程,使稀有群体更有可能在最终经二次采样的数据集中示出。

[0106] 在未降维的情况下对所述数据空间进行简单分割可能会使分箱因所谓的“维数灾难”而出现数据稀少的情况。所采用的降维变换或函数可以是关系保持嵌入,其允许在低维空间内进行分箱,且允许在二次采样之前对数据进行更为有效的分组。

[0107] 颗粒分析事件数据二次采样方法

[0108] 图4是显示对颗粒分选事件数据(例如,流式细胞术事件数据)进行二次采样的示例性方法400的流程图。方法400可以体现在存储于计算机可读介质(例如,计算系统的一个或多个磁盘驱动器)上的一组可执行程序指令中。例如,图5所示以及下文中更为详细地描述的计算系统500可以执行一组可执行程序指令,以实施方法400。在启用方法400时,所述可执行程序指令可被加载至诸如RAM的存储器中,并由计算系统500的一个或多个处理器执行。虽然就图5所示的计算系统500描述了方法400,但所述描述仅为了说明目的而提供,而非旨在进行限定。在一些实施例中,方法400或其部分可以由多个计算系统串行或并行执行。

[0109] 方法400在框404处开始后,方法400前进至框408,其中计算系统将高维空间内流式细胞术事件数据集中与第一多个事件中的第一事件相关联的第一流式细胞术事件数据转换成第一低维空间内与所述第一事件相关联的第一经转换的流式细胞术事件数据。所述第一事件可与正二次采样需求相关联。例如,对包含所述第一流式细胞术事件数据的流式细胞术事件数据进行二次采样时,所述经二次采样的流式细胞术事件数据可以不包括所述第一流式细胞术事件数据。所述第一低维空间可与第一多个分箱相关联。所述第一经转换的流式细胞术事件数据可与所述第一多个分箱中的所述第一分箱相关联。生成经二次采样的流式细胞术事件数据时,所述计算系统可以指示(例如,在数据结构中)应当包括所述第一流式细胞术事件数据。

[0110] 在一些实施例中,所述计算系统可以接收包含所述第一流式细胞术事件数据的流式细胞术事件数据。所述计算系统可以确定所述第一多个事件中所述第一事件的所述第一流式细胞术事件数据与所述正二次采样需求相关联。所述计算系统可以确定所述第一经转换的流式细胞术事件数据与所述第一多个分箱中的所述第一分箱相关联。

[0111] 所述处理器可被所述可执行指令编程为:基于所述第一多个分箱中的所述第一分箱确定所述第一经转换的流式细胞术事件数据的第一描述符。与所述第一分箱相关联的所述第一经转换的流式细胞术事件数据的所述第一描述符可以是所述第一多个分箱中所述第一分箱的第一分箱编号。所述计算系统可以:向内存数据结构中添加所述第一分箱、所述第一描述符和/或所述第一分箱编号。

[0112] 在一些实施例中,所述第一多个分箱中的两个分箱具有相同尺寸。所述第一多个分箱中的每个分箱可具有相同尺寸。所述第一多个分箱中的两个分箱可具有不同尺寸。所述第一多个分箱中的两个分箱可以包含数量大致相同的经转换的流式细胞术事件数据。所述第一多个分箱中的每个分箱可以包含数量大致相同的经转换的流式细胞术事件数据。所述计算系统可以:确定所述第一多个分箱中每个分箱的尺寸。所述处理器可被所述可执行指令编程为:基于多个门确定所述第一多个分箱中每个分箱的所述尺寸。所述计算系统可以:基于与多种目的细胞相关联的所述经转换的流式细胞术事件数据确定所述第一多个分箱中每个分箱的所述尺寸。

[0113] 在一些实施例中,为了转换所述第一流式细胞术事件数据,所述计算系统可以:利用第一降维函数转换所述第一流式细胞术事件数据。所述第一降维函数可以是线性降维函数。所述第一降维函数可以是非线性降维函数。所述非线性降维函数可以是t分布随机邻居嵌入(t-SNE)。所述计算系统可以:首先接收所述降维函数或其标识。

[0114] 方法400前进至框412,其中所述计算系统可以将所述高维空间内所述流式细胞术事件数据集中与第一多个事件中的第二事件相关联的第二流式细胞术事件数据转换成所述第一低维空间内与所述第二事件相关联的第二经转换的流式细胞术事件数据。所述第二事件可与所述正二次采样需求相关联。所述第二经转换的流式细胞术事件数据可与所述第一多个分箱中的第二分箱相关联。在一些实施例中,为了转换所述第二流式细胞术事件数据,所述计算系统可以:利用第二降维函数转换所述第一流式细胞术事件数据。所述第一降维函数和所述第二降维函数可以是相同的。

[0115] 在一些实施例中,所述计算系统可以接收包含所述第二流式细胞术事件数据的流式细胞术事件数据。所述计算系统可以确定所述第一多个事件中所述第二事件的所述第二流式细胞术事件数据与所述正二次采样需求相关联。所述计算系统可以确定所述第二经转换的流式细胞术事件数据与所述第一多个分箱中的所述第二分箱相关联。

[0116] 所述处理器可被所述可执行指令编程为:基于所述第一多个分箱中的所述第二分箱确定所述第二经转换的流式细胞术事件数据的第二描述符。与所述第二分箱相关联的所述第二经转换的流式细胞术事件数据的所述第二描述符可以是所述第一多个分箱中所述第一分箱的第二分箱编号。所述计算系统可以向所述内存数据结构中添加所述第二分箱、所述第二描述符和/或所述第二分箱编号。

[0117] 所述第一流式细胞术事件数据与第一稀有细胞相关联和/或所述第二流式细胞术事件数据可与第二稀有细胞相关联。所述第一稀有细胞和所述第二稀有细胞可以是不同细胞类型的细胞。

[0118] 方法400从框412前进至框416,其中所述计算系统可以确定与所述第一经转换的流式细胞术事件数据相关联的所述第一分箱和与所述第二经转换的流式细胞术事件数据相关联的所述第二分箱是不同的。生成经二次采样的流式细胞术事件数据时,所述计算系统可以指示(例如,在数据结构中)应当包括所述第二流式细胞术事件数据。

[0119] 在一些实施例中,为了转换所述第一流式细胞术事件数据,所述计算系统可以:利用第二降维函数将所述第一流式细胞术事件数据转换成第二低维空间内与所述第一事件相关联的第一经转换的流式细胞术事件数据。所述第二低维空间可与第二多个分箱相关联。所述第二低维空间内的所述第一经转换的流式细胞术事件数据可与所述第二多个分箱中的第一分箱相关联。为了转换所述第二流式细胞术事件数据,所述计算系统可以:利用所述第二降维函数将所述第二流式细胞术事件数据转换成所述第二低维空间内与所述第二事件相关联的第二经转换的流式细胞术事件数据。所述第二低维空间内的所述第二经转换的流式细胞术事件数据可与所述第二多个分箱中的第二分箱相关联。所述第一多个分箱中的所述第一分箱可与第一目的细胞类型相关联,所述第二多个分箱中的所述第二分箱可与第二目的细胞类型相关联,所述第一多个分箱中的所述第二分箱与所述第一目的细胞类型无关,所述第一多个分箱中的所述第二分箱与所述第二目的细胞类型无关,所述第二多个分箱中的所述第一分箱与所述第二目的细胞类型无关,和/或所述第二多个分箱中的所述第一分箱与所述第一目的细胞类型无关。

[0120] 所述第一多个分箱中的所述第一分箱与所述第二多个分箱中的所述第一分箱的第一组合可与第一目的细胞类型相关联,和/或所述第一多个分箱中的所述第二分箱与所述第二多个分箱中的所述第二分箱的第二组合可与第二目的细胞类型相关联。所述第一多

个分箱中的所述第一分箱与所述第二多个分箱中的所述第二分箱的第一组合与所述第一目的细胞类型和所述第二目的细胞类型无关联,和/或所述第一多个分箱中的所述第二分箱与所述第二多个分箱中的所述第一分箱的第二组合与所述第一目的细胞类型和所述第二目的细胞类型无关联。所述计算系统可以确定所述第一组合和所述第二组合是不同的。

[0121] 方法400前进至框420,其中所述计算系统可以生成所述流式细胞术事件数据的经二次采样的流式细胞术事件数据,所述数据包含与所述第一事件相关联的所述第一流式细胞术事件数据和与所述第二事件相关联的所述第二流式细胞术事件数据。

[0122] 在一些实施例中,所述计算系统可以:将所述高维空间内所述流式细胞术事件数据集中与第一多个事件中的第三事件相关联的第三流式细胞术事件数据转换成所述第一低维空间内与所述第三事件相关联的第三经转换的流式细胞术事件数据。所述第三事件可与所述正二次采样需求相关联。所述第三经转换的流式细胞术事件数据可与所述第一多个分箱中的所述第三分箱相关联。所述处理器可被所述可执行指令编程为:确定与所述第三经转换的流式细胞术事件数据相关联的所述第三分箱是与所述第一经转换的流式细胞术事件数据相关联的所述第一分箱或与所述第二经转换的流式细胞术事件数据相关联的所述第二分箱。所述第三流式细胞术事件数据可能不在所述流式细胞术事件数据的所述经二次采样的流式细胞术事件数据中。所述计算系统可以:基于所述第一多个分箱中的所述第三分箱确定所述第三经转换的流式细胞术事件数据的第三描述符。与所述第三分箱相关联的所述第三经转换的流式细胞术事件数据的所述第三描述符可以是所述第一多个分箱中所述第三分箱的第三分箱编号。所述计算系统可以:确定所述第三分箱、所述第三描述符和/或所述第三分箱编号不在所述内存数据结构中。

[0123] 在一些实施例中,所述计算系统包含:确定第四流式细胞术事件数据(其与所述第一多个事件中的第四事件相关联)与负二次采样需求相关联。所述计算系统可以:生成所述流式细胞术事件数据的所述经二次采样的流式细胞术事件数据集,所述事件数据包含与所述第四事件相关联的所述第四流式细胞术事件数据。所述计算系统可以:接收定义多种目的细胞的多个门。所述第四流式细胞术事件数据可与所述多种目的细胞中的目的细胞相关联。所述第四流式细胞术事件数据与已分选细胞相关联。

[0124] 在一些实施例中,所述计算系统可以:将所述高维空间内所述流式细胞术事件数据集中与第二多个事件中的第二事件相关联的第二流式细胞术事件数据转换成所述第一低维空间内与所述第二多个事件中的所述第二事件相关联的第二经转换的流式细胞术事件数据。所述第二多个事件中的所述第二事件可与所述正二次采样需求相关联。

[0125] 所述第二经转换的流式细胞术事件数据(其与所述第二多个事件中的所述第二事件相关联)可与所述第一多个分箱中的第二分箱相关联。与所述第二经转换的流式细胞术事件数据(其与所述第二多个事件中的所述第二事件相关联)相关联的所述第二分箱和与所述第一经转换的流式细胞术事件数据(其与所述第一多个事件中的所述第一事件相关联)相关联的所述第一分箱是相同的。所述计算系统可以:生成所述流式细胞术事件数据的所述经二次采样的流式细胞术事件数据集,所述事件数据包含与所述第二多个事件中的所述第二事件相关联的所述第二流式细胞术事件数据。

[0126] 所述计算系统可以:确定所述第一多个事件中的最后事件与超过预定阈值的时间参数或事件数相关联。所述计算系统可以:重置所述内存数据结构。所述处理器可被所述可

执行指令编程为:向所述内存数据结构中添加与所述第二经转换的流式细胞术事件数据(其与所述第二多个事件中的所述第二事件相关联)相关联的所述第二分箱。在一些实施例中,所述计算系统可以:接收二次采样参数水平。所述计算系统可以:基于所述二次采样参数水平确定所述预定阈值。

[0127] 方法400结束于框424处。

[0128] 执行环境

[0129] 图5描绘了示例计算设备500的总体架构,所述示例计算设备被配置成操作本文所揭示的代谢物、注释和基因整合系统。图5所示的计算设备500的所述总体架构包括计算机硬件和软件组件的布置。计算设备500可以包括相较于图5中示出的元件更多(或更少)的元件。但是,提供实现性公开内容时,无需示出所有此类常规元件。如图所示,计算设备500包括处理单元510、网络接口520、计算机可读介质驱动器530、输入/输出设备接口540、显示器550和输入设备560,所有这些都可以通过通信总线与另一设备进行通信。网络接口520可与一个或更多个网络或计算系统连接。处理单元510因此可以经由网络从其他计算系统或服务处接收信息和指令。

[0130] 处理单元510还可以与存储器570进行通信,且可以通过输入/输出设备接口540进一步为可选显示器550提供输出信息。输入/输出设备接口540还可以接受来自可选输入设备560的输入,例如键盘、鼠标、数位笔、麦克风、触摸屏、手势识别系统、语音识别系统、游戏手柄、加速度计、陀螺仪或其他输入设备。

[0131] 存储器570可以含有计算机程序指令(在一些实施例中被分组为模块或组件),所述指令由处理单元510执行以实现一个或更多个实施例。存储器570通常包括RAM、ROM和/或其他持久性、辅助性或非暂时性计算机可读介质。存储器570可以存储操作系统572,所述操作系统提供计算机程序指令以供处理单元510在计算设备500的常规管理和操作中使用。存储器570可以进一步包括用于实现本发明的方面的计算机程序指令和其他信息。

[0132] 例如,在一实施例中,存储器570包括用于对颗粒分析事件数据进行二次采样的二次采样模块574,例如,图4所述的二次采样方法400。另外,存储器570可以包括数据存储区590和/或一个或更多个存储所生成的流式细胞术事件数据集或经二次采样的流式细胞术事件数据集的其他数据存储区或与之通信。

[0133] 术语

[0134] 本文中使用的术语“确定”涵盖各种活动。例如,“确定”可以包括计算、计算机计算、处理、推导、调查、查找(例如,在表、数据库或另一数据结构中查找)、查明等。此外,“确定”可以包括接收(例如,接收信息)、访问(例如,访问存储器中的数据)等。此外,“确定”可以包括解析、选定、选择、建立等。

[0135] 本文中使用的术语“提供”涵盖各种活动。例如,“提供”可以包括将值存储在便于后续检索的存储设备的位置、将值通过至少一种有线或无线通信介质直接发送给接收者、发送或存储对值的引用等。

[0136] “提供”还可以包括通过硬件进行编码、解码、加密、解密、确认、验证等。

[0137] 本文中使用的术语“选择性地”或“选择性”可以涵盖各种活动。例如,“选择性”过程可以包括从多个选项中确定一个选项。“选择性”过程可以包括以下一项或更多项:动态确定的输入、预先配置的输入或用户为确定而发起的输入。在一些实施方案中,可以包括n

输入切换以提供选择性功能,其中n是用于进行选择的输入数量。

[0138] 本文中使用的术语“消息”涵盖用于传达(例如,发送或接收)信息的各种格式。消息可以包括机器可读的信息聚合,例如,XML文档、固定字段消息、逗号分隔消息等。在一些实施方案中,消息可以包括用于传输一种或更多种信息表示的信号。虽然以单数形式叙述,但应当理解,消息可以多个部分组成、发送、存储、接收等。

[0139] 本文中使用的“用户界面”(也被称为交互式用户界面、图形用户界面或UI)可以指基于网络的包括数据字段、按钮或用于接收输入信号或提供电子信息或向用户提供信息以响应于任何接收到的输入信号的其他交互控件。可以采用诸如超文本标记语言(HTML)、JAVASCRIPT™、FLASH™、JAVA™、.NET™、WINDOWSOS™、

[0140] macOS™、Web服务和丰富站点摘要(RSS)等技术来全部或部分实现UI。在一些实施方案中,可以将UI包括在被配置成根据所述的一个或更多个方面进行通信(例如,发送或接收数据)的独立客户端(例如,胖客户端、胖客户机)中。

[0141] 本文中使用的“数据存储区”可以体现在可以访问某一设备或由所述设备访问的硬盘驱动器、固态存储器和/或任何其他类型的非暂时性计算机可读存储介质中,所述设备包括存取设备、服务器或其他所述的计算设备等。如本领域所公知,在不脱离本发明的范围的情况下,数据存储区还可以或替代地在多个本地和/或远程存储设备上实现分布或分区。在其他实施例中,数据存储区可以包括或体现在数据存储网络服务中。

[0142] 本领域技术人员应当理解,可以使用各种不同技术和技巧中的任何一种来表示信息、消息和信号。例如,在上述整个说明书中可能提及的数据、指令、命令、信息、信号、位、符号和芯片可以用电压、电流、电磁波、磁场或粒子、光场或光学粒子或其任何组合表示。

[0143] 本领域技术人员应当进一步理解,结合本文所揭示的实施例描述的各种说明性逻辑块、模块、电路和算法步骤可以电子硬件、计算机软件或两者的组合的形式实现。为了清楚阐明硬件和软件的这种可互换性,上文已大体上就其功能描述了各种说明性组件、块、模块、电路和步骤。这种功能以硬件还是软件的形式实现取决于特定应用以及对于整个系统的设计约束。技术人员可以针对各特定应用以各种方式执行所述功能,但是这种执行决策不应被解释为导致脱离本发明的范围。

[0144] 本文所述的技术可以硬件、软件、固件或其任何组合的形式来实施。所述技术可以在各种设备中的任何一种中实施,例如,专门编程的事件处理计算机、无线通信设备或集成电路器件。被描述为模块或组件的任何功能都可以在集成逻辑设备中一同执行,或在离散但可互操作的逻辑设备中分别执行。如果以软件形式实施,则所述技术可以至少部分通过计算机可读数据存储介质来实现,所述计算机可读数据存储介质包括指令,执行所述指令时,即执行上述一种或更多种方法。所述计算机可读数据存储介质可以构成计算机程序产品的一部分,其可以包括包装材料。所述计算机可读介质可以包含存储器或数据存储介质,例如,随机存取存储器(RAM)(例如同步动态随机存储器(SDRAM))、只读存储器(ROM)、非易失性随机存取存储器(NVRAM)、电可擦可编程只读存储器(EEPROM)、闪速存储器、磁性或光学数据存储介质等。所述计算机可读介质可以是非暂时性存储介质。

[0145] 此外或另外,所述技术可以至少部分通过计算机可读通信介质来实现,所述计算机可读通信介质以指令或数据结构的形式携带或传递程序代码,而且可以通过计算设备来访问、读取和/或执行,例如传播的信号或波。

[0146] 所述程序代码可以由专门编程的分选策略处理器执行,所述分选策略处理器可以包括一个或更多个处理器,例如,一个或更多个数字信号处理器(DSP)、可配置微处理器、专用集成电路(ASIC)、现场可编程门阵列(FPGA)或其他等效的集成或离散逻辑电路系统。所述图形处理器可以被专门配置成执行本发明中所描述的任何技术。计算设备的组合(例如,DSP和微处理器的组合、多个微处理器、与DSP内核结合的一个或更多个微处理器或至少部分数据连接中的任何其他此等配置)可以执行一项或更多项所述功能。在一些方面,本文所述的功能可以在被配置成用于编码和解码的专用软件模块或硬件模块内提供,或者将其并入专用分选控制卡中。

[0147] 本领域技术人员应当理解,一般而言,本文使用的术语,特别是所附权利要求书中(例如,在所附权利要求书的主体部分中)使用的术语,通常应理解为“开放”术语(例如,术语“包括”应解释为“包括但不限于”,术语“具有”应解释为“至少具有”等)。本领域技术人员还应理解,如果意在所引入的权利要求中标明具体数目,则这种意图将在该权利要求中明确指出,而在无这种明确标明的情况下,则视为不存在这种意图。例如,为帮助理解,所附权利要求可能使用了引导短语“至少一个”和“一个或更多个”来引入权利要求中的特征。然而,这种短语的使用不应被解释为暗示着由不定冠词“一”或“一个”引入的权利要求特征将含有该特征的任意特定权利要求限制为仅含有一个该特征的实施例,即便是该权利要求既包括引导短语“一个或更多个”或“至少一个”又包括不定冠词如“一”或“一个”(例如,“一”和/或“一个”应当被解释为意指“至少一个”或“一个或更多个”);在使用定冠词来引入权利要求中的特征时,同样如此。

[0148] 另外,即使明确指出了所引入权利要求特征的具体数目,本领域技术人员应当认识到,这种列举应解释为意指至少是所列数目(例如,不存在其他修饰语的短语“两个特征”意指至少两个该特征,或者两个或更多该特征)。另外,在使用类似于“A、B和C等中至少一个”这样的表述的情况下,一般来说应该按照本领域技术人员通常理解该表述的含义来予以解释(例如,“具有A、B和C中至少一个的系统”应包括但不限于单独具有A、单独具有B、单独具有C、具有A和B、具有A和C、具有B和C和/或具有A、B和C的系统等)。在使用类似于“A、B或C等中至少一个”这样的表述的情况下,一般来说应该按照本领域技术人员通常理解该表述的含义来予以解释(例如,“具有A、B或C中至少一个的系统”应包括但不限于单独具有A、单独具有B、单独具有C、具有A和B、具有A和C、具有B和C和/或具有A、B和C的系统等)。本领域技术人员还应当理解,实质上任意表示两个或更多可选项目的转折连词和/或短语,无论是在说明书、权利要求书还是附图中,都应被理解为给出了包括这些项目之一、这些项目任一方或两个项目的可能性。例如,短语“A或B”应当被理解为包括“A”或“B”或“A和B”的可能性。

[0149] 另外,在以马库什组描述本发明的特征或方面的情况下,本领域技术人员应当认识到,本发明因此也是以该马库什组中的任意单独成员或成员子组来描述的。

[0150] 本领域技术人员应当理解,出于任意和所有目的,例如为了提供书面说明,本文所揭示的所有范围也包含任意及全部可能的子范围及其子范围的组合。任意列出的范围可以被容易地看作充分描述且实现了将该范围至少进行二等分、三等分、四等分、五等分、十等分等。作为非限制性示例,在此所讨论的每一范围可以容易地分成下三分之一、中三分之一和上三分之一等。本领域技术人员应当理解,所有诸如“直至”、“至少”、“大于”、“小于”之类的语言包括所列数字,并且指代了随后可以如上所述被分成子范围的范围。最后,本领域技

术人员应当理解,范围包括每一单独数字。因此,例如具有1-3个单元的组是指具有1、2或3个单元的组。类似地,具有1-5个单元的组是指具有1、2、3、4或5个单元的组,以此类推。

[0151] 本文所揭示的方法包含用于实现所述方法的一个或更多个步骤或动作。在不脱离权利要求书的范围的情况下,所述方法步骤和/或动作可以彼此互换。换言之,除非指定了步骤或动作的特定顺序,否则可以在不脱离权利要求书的范围的情况下,修改特定步骤和/或动作的顺序和/或用途。

[0152] 虽然已在本文中揭示了各方面和实施例,但是其他方面和实施例对于本领域技术人员而言是显而易见的。本文所揭示的各方面和实施例是为了说明目的而提供,而非旨在进行限定,其真实范围和精神由以下权利要求书指示。

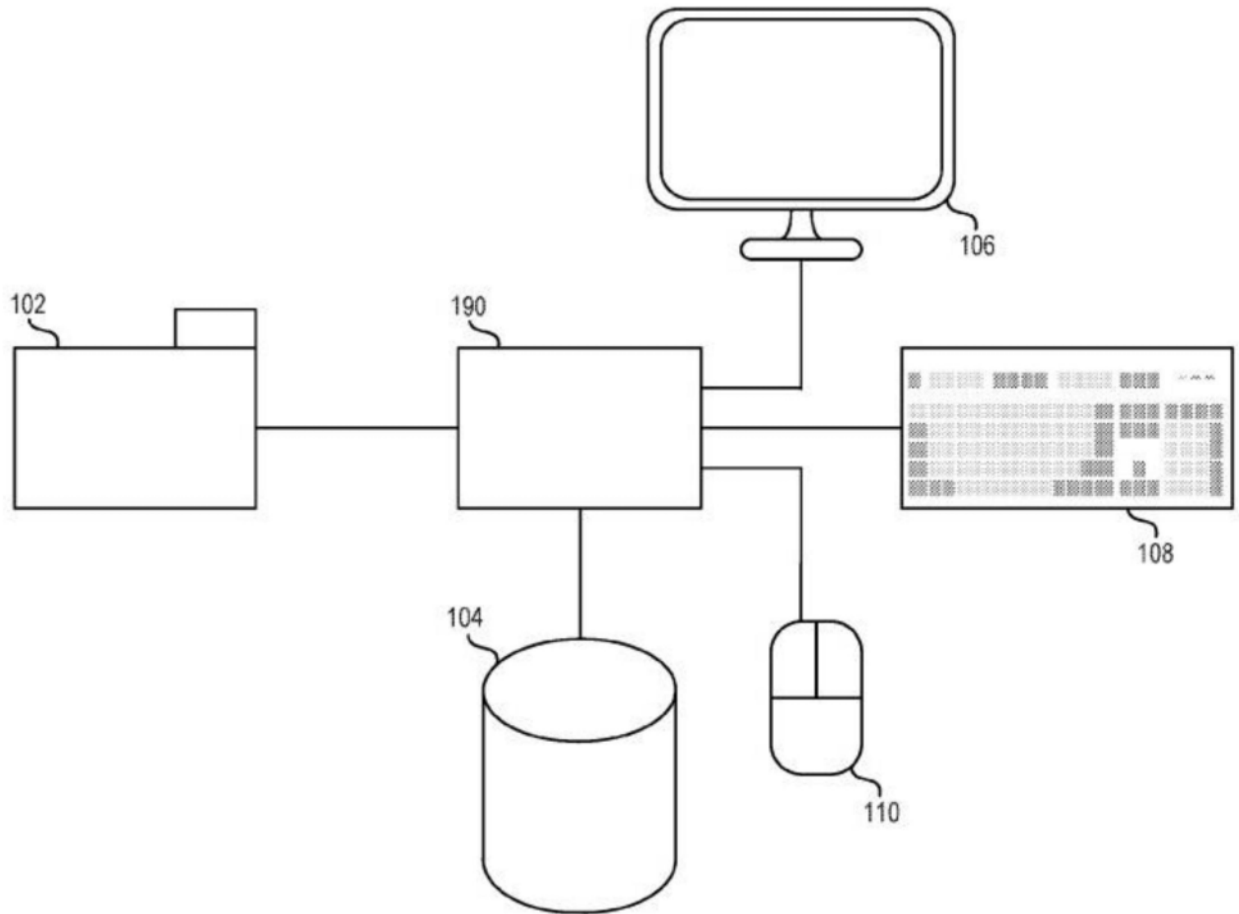


图1

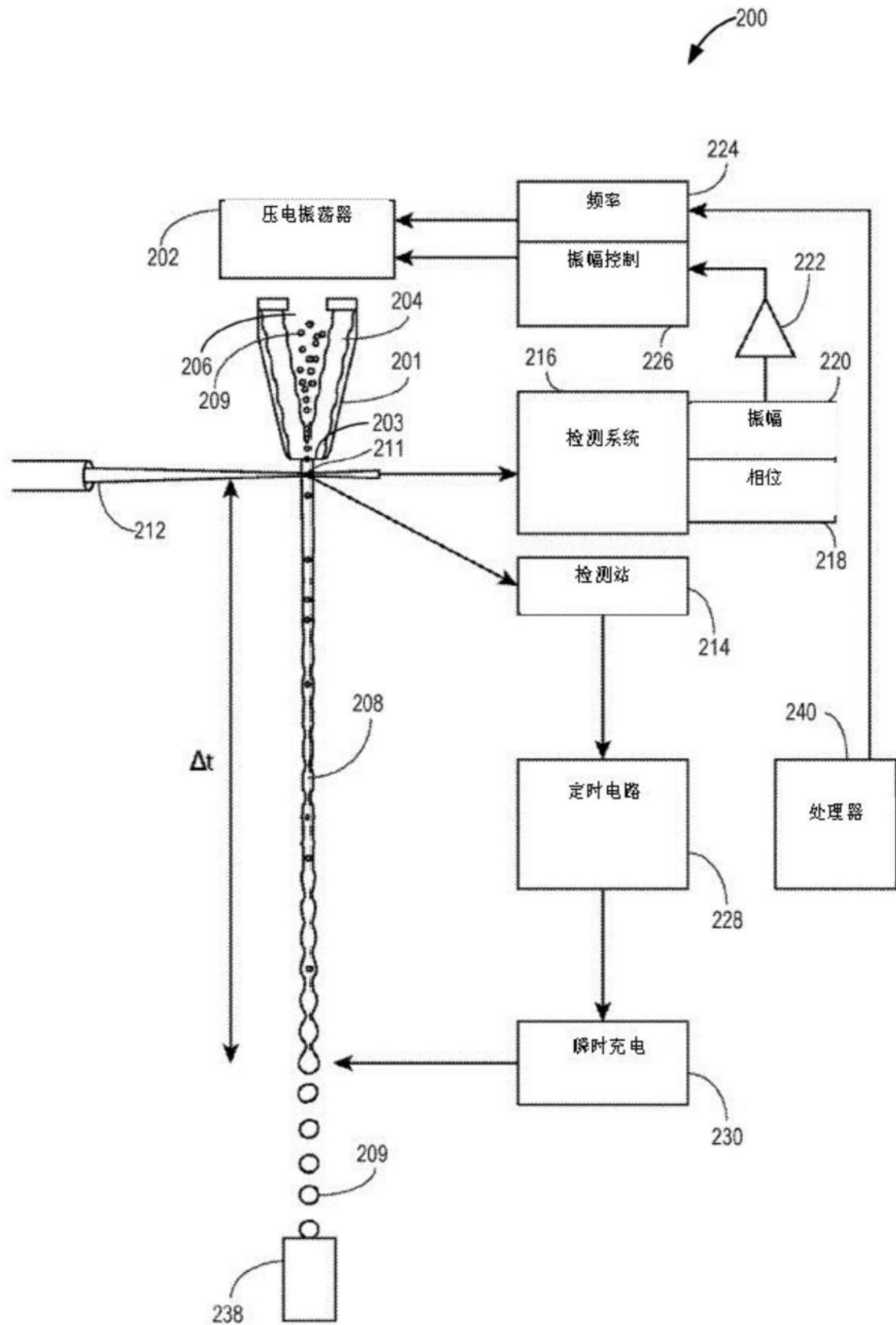


图2A

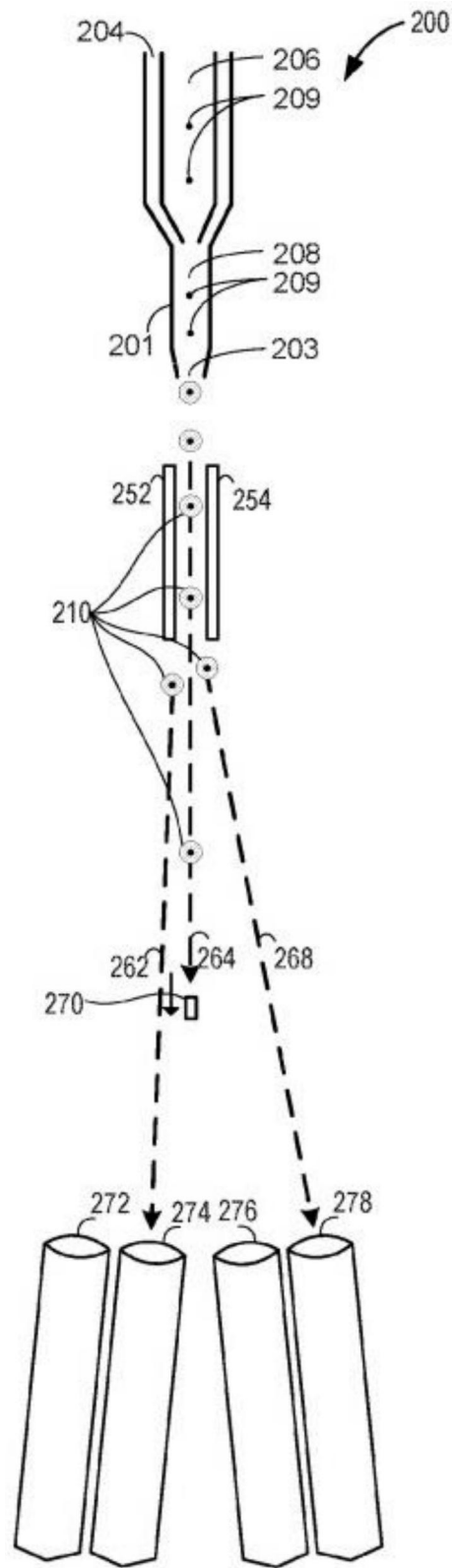


图2B

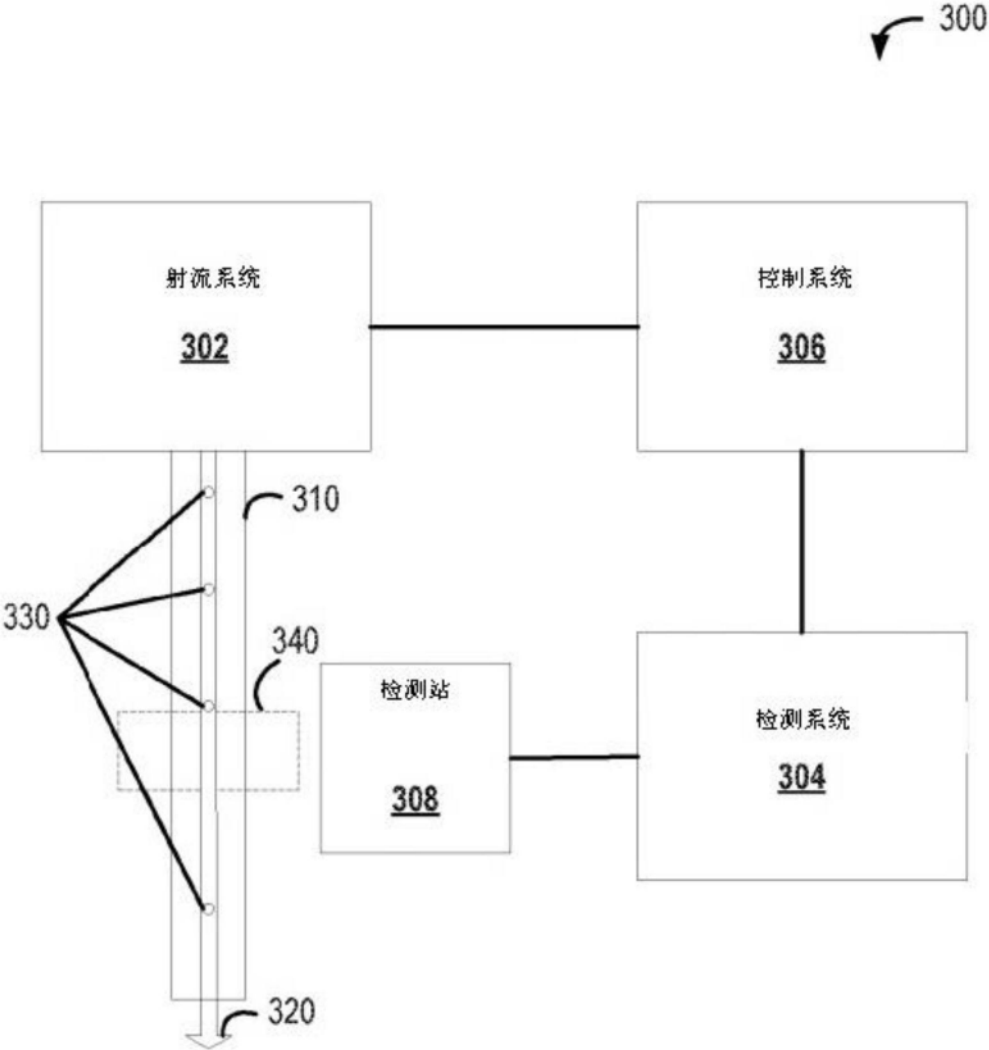


图3

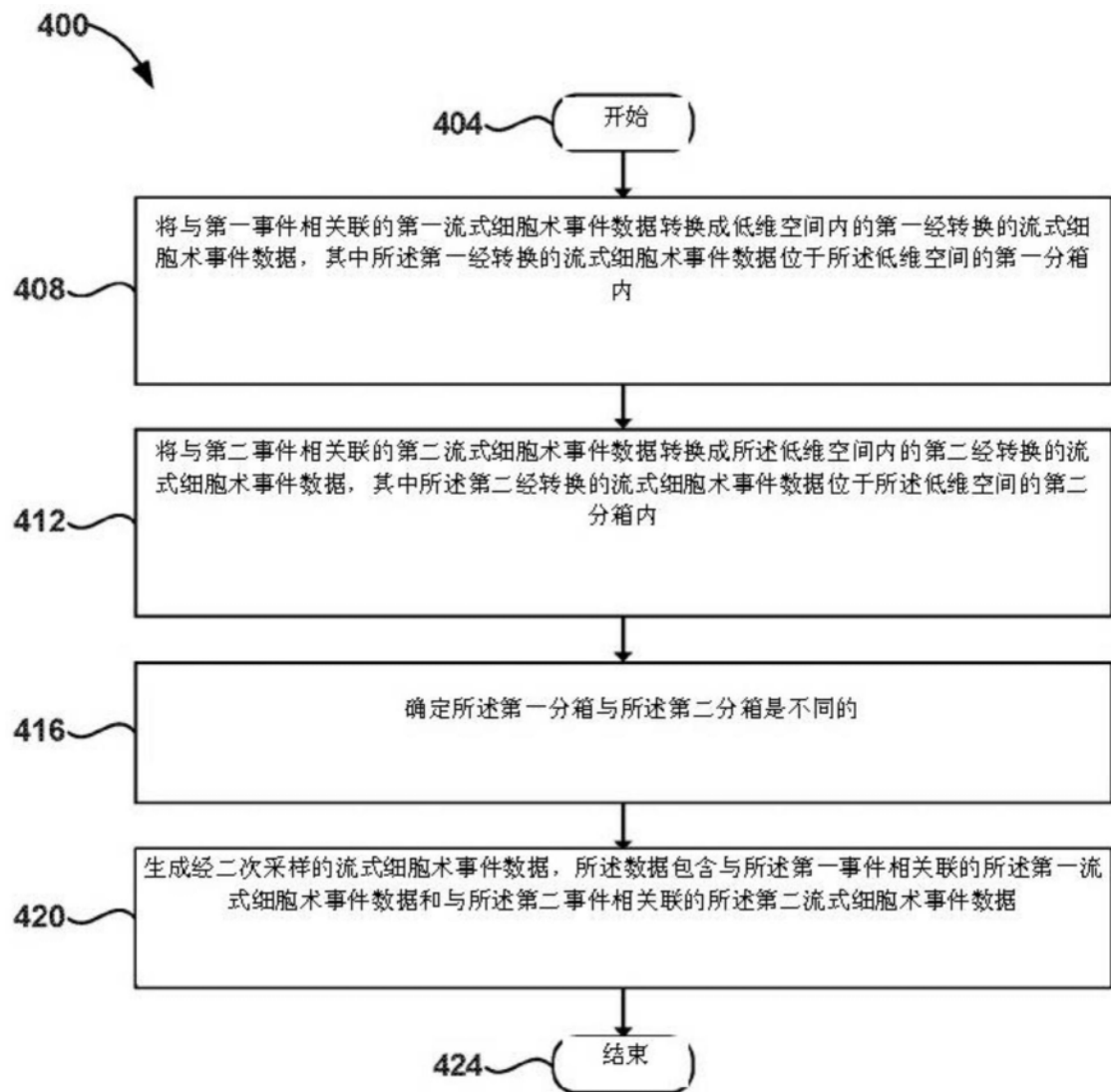


图4

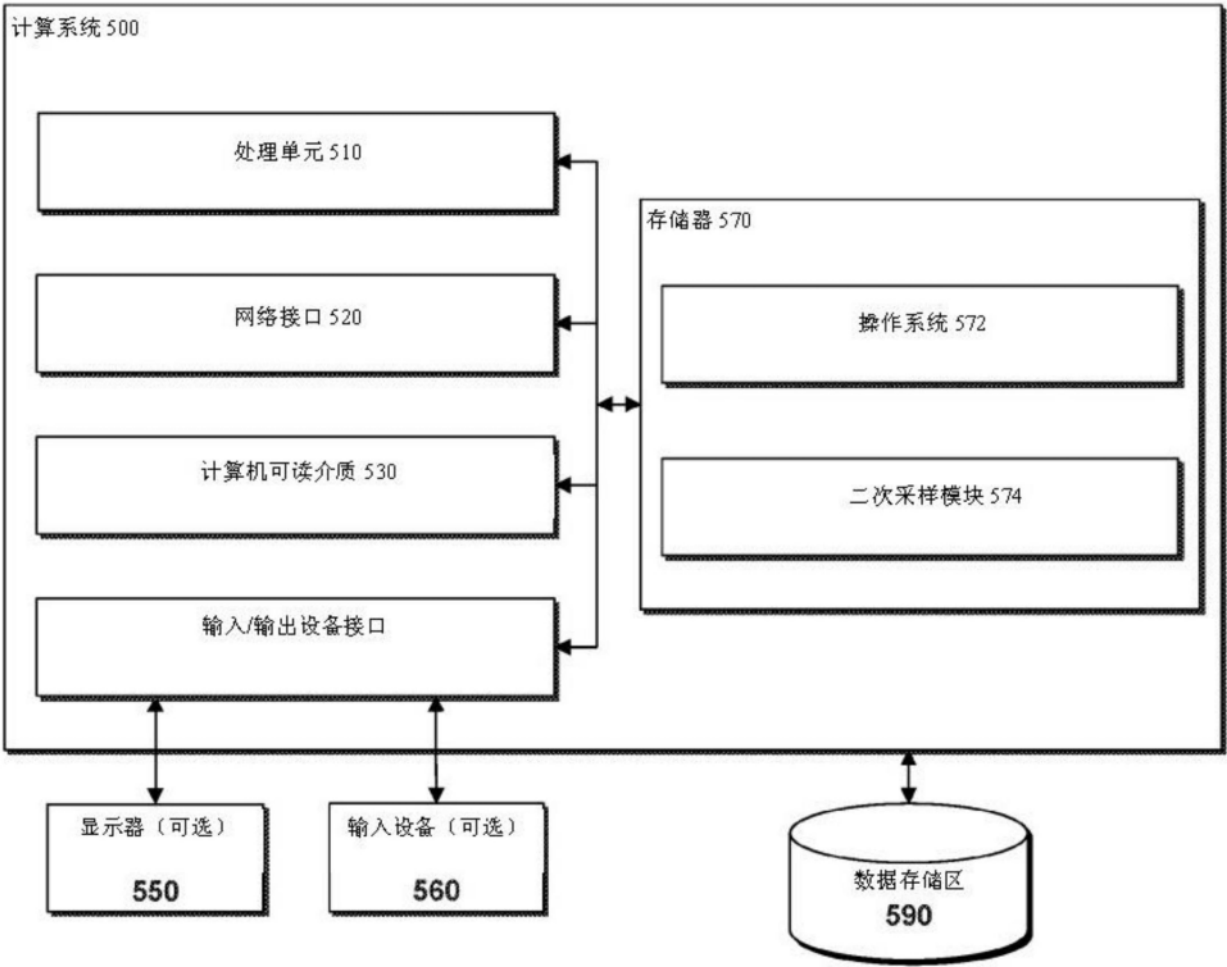


图5