

# (12) 发明专利申请

(10) 申请公布号 CN 102262725 A

(43) 申请公布日 2011. 11. 30

(21) 申请号 201110132428. 8

(22) 申请日 2011. 05. 20

(30) 优先权数据

61/349, 907 2010. 05. 31 US

12/854, 187 2010. 08. 11 US

(71) 申请人 普莱姆森斯有限公司

地址 以色列特拉维夫市

(72) 发明人 A·格曼 T·亚尼尔 M·布兰德

E·顾恩德曼

(74) 专利代理机构 北京北翔知识产权代理有限

公司 11285

代理人 徐燕 杨勇

(51) Int. Cl.

G06K 9/00 (2006. 01)

G06T 7/20 (2006. 01)

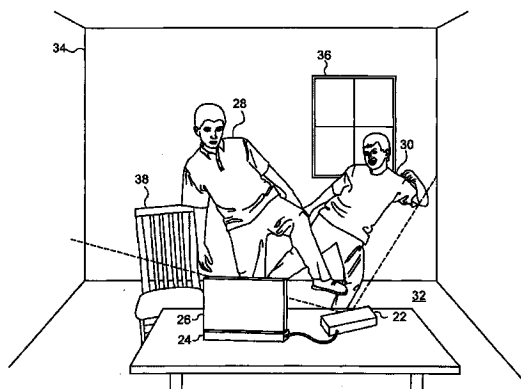
权利要求书 4 页 说明书 11 页 附图 7 页

(54) 发明名称

三维场景的分析

(57) 摘要

一种用于处理数据的方法和装置,包括接收包含人形形状的场景的深度图。所述深度图被处理,以识别所述场景中的三维(3D)连接分量,每个连接分量都包括一组相互邻近并具有相互邻近深度值的像素。将分离的第一连接分量和第二连接分量都识别为属于所述人形形状,并且产生包括所述第一连接分量和所述第二连接分量的所述人形形状表示。



1. 一种用于处理数据的方法,包括:

接收场景的暂时的深度图序列,所述场景包括静态背景和隐藏所述背景一部分的一个或多个移动的前景目标,所述深度图包括像素矩阵,所述像素的至少一些具有各自的像素深度值并对应于所述场景中的各自位置;

使用数字处理器,处理所述序列中的第一深度图,以识别没有被所述第一图中所述移动的前景目标所隐藏的所述静态背景的第一部分;

处理所述序列中至少一个第二深度图,以识别所述静态背景的至少一个第二部分,由于所述前景目标中的至少一个在所述序列上的运动,所述第二部分不同于所述第一部分;以及

通过至少组合所述静态背景的所述第一部分和所述第二部分,构造所述场景的背景模型。

2. 根据权利要求 1 所述的方法,还包括基于所述背景模型,通过将一或多个背景目标从所述深度图中移除来产生一或多个前景图。

3. 根据权利要求 2 所述的方法,其中所述移动的前景目标中的至少一个是人形体形,并且其中所述方法包括从前景图中提取所述人形体形,所述前景图从所述第二深度图连同与所述第一深度图中所述人形体形相关的信息中得出。

4. 根据权利要求 3 所述的方法,其中提取所述人形体形包括:检测属于正在触摸所述背景目标的其中之一的人形体形的区域,并将该区域识别为所述人形体形的一部分。

5. 根据权利要求 1 所述的方法,还包括识别所述深度图中的平面表面。

6. 根据权利要求 1 所述的方法,还包括识别所述深度图中的目标的阴影。

7. 根据权利要求 1 所述的方法,其中构造所述背景模型包括:响应于在所述深度图序列上所述背景中出现的变化,来更新所述背景模型。

8. 根据权利要求 7 所述的方法,其中所述深度图被成像组件捕获,并且其中所述方法包括评估所述序列中深度图之间所述背景模型变化的程度,并且当所述变化的程度超过预定阈值时,检测所述成像组件已经移动。

9. 一种用于处理数据的方法,包括:

接收场景的深度图,所述场景包含一个或多个关于表面定位的目标,所述深度图包括像素矩阵,所述像素的至少一些具有各自的像素深度值并对应于在所述场景中的各自位置;

使用数字处理器,处理所述深度图,以识别并产生所述表面的参数模型;以及

通过将由所述参数模型表示的所述表面从所述深度图中移除,分割所述场景中所述一个或多个目标。

10. 根据权利要求 9 所述的方法,其中所述目标中的至少一个是人形体形,并且其中所述分割所述一个或多个目标包括从所述深度图中提取所述人形体形。

11. 根据权利要求 9 所述的方法,其中所述表面是对应于地板的平面表面,在所述地板上定位有所述目标中的至少一个。

12. 根据权利要求 11 所述的方法,其中处理所述深度图包括计算所述深度图中所述像素的局部法线,并且根据所述局部法线聚集所述像素以识别所述平面表面。

13. 根据权利要求 12 所述的方法,其中所述深度图被成像组件捕获,并且其中聚集所

述局部法线包括：处理所述像素的至少一些的各自深度值以便于预测所述成像组件相对于每个像素的高度，以及聚集具有相同的所述成像组件预测高度的像素。

14. 一种用于处理数据的方法，包括：

接收包含人形形状的场景的深度图，所述深度图包括像素矩阵，所述像素的至少一些具有各自的像素深度值并对应于在所述场景中的各自位置；

使用数字处理器，处理所述深度图，以识别所述场景中的三维（3D）连接分量，每个连接分量包括一组相互邻近并具有相互邻近深度值的像素；

将分离的第一连接分量和第二连接分量都识别为属于所述人形形状；以及

产生包括所述第一连接分量和所述第二连接分量的所述人形形状表示。

15. 根据权利要求 14 所述的方法，其中处理所述深度图包括：基于所述深度图构造所述场景的背景模型，将所述背景模型从所述深度图中移除以产生前景图，以及识别所述前景图中的所述 3D 连接分量。

16. 根据权利要求 14 所述的方法，其中处理所述深度图包括：定位所述深度图的边缘和所述边缘之间的像素块，并且聚集相邻的三维像素块以识别所述 3D 连接分量。

17. 根据权利要求 14 所述的方法，其中接收所述深度图包括接收暂时的深度图序列，并且其中识别所述第一连接分量和所述第二连接分量包括在所述序列中的多个连续深度图上跟踪所述连接分量。

18. 根据权利要求 17 所述的方法，其中跟踪所述连接分量包括：为所述序列中第一深度图中的所述第一连接分量和所述第二连接分量分配公共标识符，并使用所述公共标识符来确定所述第一连接分量和所述第二连接分量属于随后的第二深度图中的所述人形形状，在所述随后的第二深度图中，所述第一连接分量和所述第二连接分量已经被分离。

19. 根据权利要求 18 所述的方法，其中分配所述公共标识符包括：用所述公共标识符标注所述深度图中所述第一连接分量和所述第二连接分量的像素。

20. 根据权利要求 19 所述的方法，其中所述第二深度图中的所述人形形状与另一个目标接触，由此导致所述第二连接分量包含其他目标，并且其中使用所述公共标识符包括：将属于所述人形形状的所述第二连接分量的被标注像素从属于其他目标的像素中分离，以产生出现在所述第二深度图中的所述人形形状表示。

21. 根据权利要求 17 所述的方法，其中跟踪所述连接分量包括：估计所述连接分量的运动速度，并且其中产生所述表示包括：使用被估计的速度保持所述人形形状的被遮挡部分的识别。

22. 根据权利要求 14 所述的方法，其中所述深度图中的所述人形形状被另一个目标部分地遮挡，并且其中所述第一连接分量和所述第二连接分量对应于被该遮挡目标所分离的所述人形形状的各部分，并且其中产生所述表示包括将所述第一连接分量和所述第二连接分量都包含在所述人形形状表示中。

23. 根据权利要求 14 所述的方法，其中所述第一连接分量和所述第二连接分量对应于所述人形形状的第一部分和第二部分，其中在所述深度图中所述第二部分部分地遮挡所述第一部分，并且其中产生所述表示包括将所述第一连接分量和所述第二连接分量都包含在所述人形形状表示中。

24. 一种用于处理数据的装置，包括：

成像组件,其被配置为产生场景的暂时的深度图序列,所述场景包括静态背景和隐藏所述背景一部分的一个或多个移动的前景目标,所述深度图包括像素矩阵,所述像素的至少一些具有各自的像素深度值并对应于所述场景中的各自位置;以及

处理器,其被配置为处理所述序列中的第一深度图,以识别没有被所述第一图中所述移动的前景目标所隐藏的所述静态背景的第一部分,并且处理所述序列中至少一个第二深度图,以识别所述静态背景的至少一个第二部分,由于所述前景目标中的至少一个在所述序列上的运动,所述第二部分不同于所述第一部分,以及通过至少组合所述静态背景的所述第一部分和所述第二部分,构造所述场景的背景模型。

25. 根据权利要求 24 所述的装置,其中所述处理器被配置为基于所述背景模型,通过将一或多个背景目标从所述深度图中移除来产生一或多个前景图。

26. 根据权利要求 25 所述的装置,其中所述移动的前景目标中的至少一个是人形体形,并且其中所述处理器被配置为从前景图中提取所述人形体形,所述前景图从所述第二深度图连同与所述第一深度图中所述人形体形相关的信息中得出。

27. 根据权利要求 26 所述的装置,其中所述处理器被配置为检测属于正在触摸所述背景目标的其中之一的人形体形的区域,并将该区域识别为所述人形体形的一部分。

28. 根据权利要求 24 所述的装置,其中所述处理器被配置为识别所述深度图中的平面地板。

29. 根据权利要求 24 所述的装置,其中所述处理器被配置为识别所述深度图中的目标的阴影。

30. 根据权利要求 24 所述的装置,其中所述处理器被配置为响应于在所述深度图序列上所述背景中出现的变化,来更新所述背景模型。

31. 根据权利要求 30 所述的装置,其中所述处理器被配置为评估所述序列中深度图之间所述背景模型变化的程度,并且当所述变化的程度超过预定阈值时,检测所述成像组件已经移动。

32. 一种用于处理数据的装置,包括:

成像组件,其被配置为生成场景的深度图,所述场景包含一个或多个关于表面定位的目标,所述深度图包括像素矩阵,所述像素的至少一些具有各自的像素深度值并对应于在所述场景中的各自位置;以及

处理器,其被配置为处理所述深度图,以识别并产生所述表面的参数模型,并通过将由所述参数模型表示的所述表面从所述深度图中移除,分割所述场景中所述一个或多个目标。

33. 根据权利要求 32 所述的装置,其中所述目标中的至少一个是人形体形,并且其中所述处理器被配置为从所述深度图中提取所述人形体形。

34. 根据权利要求 32 所述的装置,其中所述表面是对应于地板的平面表面,在所述地板上定位有所述目标中的至少一个。

35. 根据权利要求 34 所述的装置,其中所述处理器被配置为计算所述深度图中所述像素的局部法线,并且根据所述局部法线聚集所述像素以识别所述平面表面。

36. 根据权利要求 35 所述的装置,其中所述处理器被配置为处理所述像素的至少一些各自的深度值以便于预测所述成像组件相对于每个像素的高度,以及聚集具有相同的所述

成像组件预测高度的像素。

37. 一种用于处理数据的装置,包括:

成像组件,其被配置为生成包含人形形状的场景的深度图,所述深度图包括像素矩阵,所述像素的至少一些具有各自的像素深度值并对应于在所述场景中的各自位置;以及

处理器,其被配置为处理所述深度图,以识别所述场景中的三维(3D)连接分量,每个连接分量包括一组相互邻近并具有相互邻近深度值的像素,将分离的第一连接分量和第二连接分量都识别为属于所述人形形状,并且产生包括所述第一连接分量和所述第二连接分量的所述人形形状表示。

38. 根据权利要求 37 所述的装置,其中所述处理器被配置为基于所述深度图构造所述场景的背景模型,将所述背景模型从所述深度图中移除以产生前景图,以及识别所述前景图中的所述 3D 连接分量。

39. 根据权利要求 37 所述的装置,其中所述处理器被配置为定位所述深度图的边缘和所述边缘之间的像素块,并且聚集相邻的三维像素块以识别所述 3D 连接分量。

40. 根据权利要求 37 所述的装置,其中所述成像组件被配置为生成暂时的深度图序列,并且其中所述处理器被配置为在所述序列中的多个连续深度图上跟踪所述连接分量。

41. 根据权利要求 40 所述的装置,其中所述处理器被配置为向所述序列中第一深度图中的所述第一连接分量和所述第二连接分量分配公共标识符,并使用所述公共标识符来确定所述第一连接分量和所述第二连接分量属于随后的第二深度图中的所述人形形状,在所述随后的第二深度图中,所述第一连接分量和所述第二连接分量已经被分离。

42. 根据权利要求 41 所述的装置,其中所述处理器被配置为用所述公共标识符来标注所述深度图中所述第一连接分量和所述第二连接分量的像素。

43. 根据权利要求 42 所述的装置,其中所述第二深度图中的所述人形形状与另一个目标接触,由此导致所述第二连接分量包含其他目标,并且所述处理器被配置为将属于所述人形形状的所述第二连接分量的被标注像素从属于其他目标的像素中分离,以产生出现在所述第二深度图中的所述人形形状表示。

44. 根据权利要求 40 所述的装置,其中所述处理器被配置为估计所述连接分量的运动速度,并且使用被估计的速度保持所述人形形状的被遮挡部分的识别。

45. 根据权利要求 37 所述的装置,其中所述深度图中的所述人形形状被另一个目标部分地遮挡,并且其中所述第一连接分量和所述第二连接分量对应于被该遮挡目标所分离的所述人形形状的各部分,并且其中所述处理器被配置为将所述第一连接分量和所述第二连接分量都包含在所述人形形状表示中。

46. 根据权利要求 37 所述的装置,其中所述第一连接分量和所述第二连接分量对应于所述人形形状的第一部分和第二部分,其中在所述深度图中所述第二部分部分地遮挡所述第一部分,并且其中所述处理器被配置为将所述第一连接分量和所述第二连接分量都包含在所述人形形状表示中。

## 三维场景的分析

### 技术领域

[0001] 本发明总体涉及用于三维 (3D) 绘图 (mapping) 的方法和系统,并特别涉及 3D 图 (map) 数据的处理。

### 背景技术

[0002] 在本领域中已知许多用于产生深度图 (depth map) 的不同方法和系统。在本专利申请和权利要求中,术语“深度图”是指场景的二维像素矩阵表示,其中每个像素分别对应于场景中的一个位置并分别具有指示从某一参考位置到各自的场景位置的距离的像素值。(换句话说,深度图是图像的形式,其中像素值指示场景中目标的形貌信息,而不是亮度和/或颜色)。通过例如向目标投射激光斑纹图案并检测和处理该目标的图像,可以生成深度图,如在 PCT 国际申请 W02007/043036A1 中所述,该申请的公开内容以参引方式被纳入本文。

[0003] 深度图可以被处理,以分割和识别场景中的目标。深度图中人形形状(指的是结构类似于人的 3D 形状)的识别以及从一个场景到另一个场景这些形状的改变可以被用作控制计算机应用(application)的手段。例如,PCT 国际申请 W02007/132451 描述了一种计算机实现方法,其中深度图被分割,以便发现人形身体的轮廓,该申请的公开内容以参引方式被纳入本文。轮廓被处理,以识别身体的躯干以及一个或多个肢体。通过分析深度图中的至少一个被识别出的肢体的布置,产生一个输入以控制运行在计算机上的应用程序。

### 发明内容

[0004] 本发明的实施方案提供用于从深度图中提取信息的方法、装置和软件。

[0005] 因此,根据本发明的一个实施方案提供一种用于处理数据的方法,该方法包括接收场景的暂时的深度图序列,所述场景包括静态背景和隐藏所述背景一部分的一个或多个移动的前景目标,所述深度图包括像素矩阵,所述像素的至少一些具有各自的像素深度值并对应于所述场景中的各自位置。数字处理器处理所述序列中的第一深度图,以识别没有被所述第一图中所述移动的前景目标所隐藏的所述静态背景的第一部分。所述序列中至少一个第二深度图被处理,以识别所述静态背景的至少一个第二部分,由于所述前景目标中的至少一个在所述序列上的运动,所述第二部分不同于所述第一部分。所述处理器通过组合所述静态背景中的至少所述第一部分和所述第二部分,构造所述场景的背景模型。

[0006] 在一些实施方案中,所述方法包括基于所述背景模型,通过发现比所述背景模型更接近的像素来产生前景图。可以发现人形体形,其包括属于所述前景图的像素和被发现属于前一帧中人形体形的像素。

[0007] 在一个公开的实施方案中,所述方法包括识别所述深度图中的平面地板(planar floor)。所述目标中的一个的阴影可以在所述深度图中被识别,并且投射阴影的目标的深度可以被用于评估所述背景模型。

[0008] 通常,构造所述背景模型包括:响应于在所述深度图序列上所述背景中出现的变

化,来更新所述背景模型。在一个公开的实施方案中,所述深度图被成像组件所捕获,并且所述方法包括评估所述序列中深度图之间所述背景模型变化的程度,并且当所述变化的程度超过预定阈值时,检测所述成像组件已经移动。

[0009] 根据本发明的一个实施方案,还提供一种用于处理数据的方法,包括接收场景的深度图,所述场景包含一个或多个关于表面定位的目标。数字处理器处理所述深度图,以识别并产生所述表面的参数模型;以及通过将由所述参数模型表示的所述表面从所述深度图中移除,分割所述场景中所述一个或多个目标。

[0010] 在一个公开的实施方案中,所述表面是对应于地板的平面表面,在所述地板上定位有所述目标中的至少一个。处理所述深度图包括计算所述深度图中所述像素的局部法线(local normal),并且根据由所述局部法线所预测的平面来聚集所述像素,以识别所述平面表面。通常,所述深度图被成像组件所捕获,并且所述聚集操作包括处理所述像素的至少一些的各自深度值,以预测所述成像组件相对于每个像素的高度,以及聚集具有相同的所述成像组件预测高度的像素。

[0011] 附加地,根据本发明的一个实施方案提供一种用于处理数据的方法,包括接收包含人形形状的场景的深度图。数字处理器处理所述深度图,以识别所述场景中的三维连接分量(3D connected component)。每个连接分量包括一组相互邻近并具有相互邻近深度值的像素。所述处理器将分离的连接分量识别为属于所述人形形状,并产生包括一些连接分量的所述人形形状表示。

[0012] 在一个公开实施方案中,处理所述深度图包括定位所述深度图的边缘和所述边缘之间的像素块,并且聚集相邻的三维像素块以识别所述 3D 集群(cluster)。

[0013] 在一些实施方案中,接收所述深度图包括接收暂时的深度图序列,并且识别所述连接分量包括在所述序列中的多个连续深度图上跟踪所述连接分量。跟踪所述连接分量可以包括:为属于所述序列中第一深度图中的相同人形形状的连接分量分配公共标识符(common identifier),并使用所述公共标识符来确定所述连接分量属于随后的第二深度图中的所述人形形状。通常,分配所述公共标识符包括用所述公共标识符标注所述深度图中所述连接分量的像素。当所述第二深度图中的所述人形形状与另一个目标接触时,由此导致所述第二连接分量包含其他目标,可以使用所述公共标识符将属于所述人形形状的所述第二连接分量的被标注像素从属于其他目标的像素中分离,以产生出现在所述第二深度图中的所述人形形状表示。

[0014] 额外地或可替代地,跟踪所述连接分量包括估计所述连接分量的运动速度,以及产生所述表示包括使用被估计的速度保持所述人形形状的被遮挡部分的识别。

[0015] 在一个实施方案中,所述深度图中的所述人形形状被另一个目标部分地遮挡,并且所述连接分量对应于被该遮挡目标分离的所述人形形状的各部分,并且产生所述表示包括识别这种情况并将所述连接分量保留在所述人形形状表示中。

[0016] 再者,根据本发明的一个实施方案提供一种用于处理数据的装置,包括成像组件,该成像组件被配置为产生场景的深度图,所述场景包含一个或多个关于表面定位的目标。处理器被配置为处理所述深度图,以识别并产生所述表面的参数模型,并通过将由所述参数模型表示的所述表面从所述深度图中移除来分割所述场景中所述一个或多个目标。

[0017] 根据本发明的一个实施方案,还提供一种用于处理数据的装置,包括成像组件,该

成像组件被配置为产生包含人形形状的场景的深度图。处理器被配置为处理深度图,以识别所述场景中的三维(3D)连接分量,每个连接分量包括一组相互邻近并具有相互邻近深度值的像素,将分离的连接分量都识别为属于所述人形形状,并且产生包括一些连接分量的所述人形形状的代表。

[0018] 根据本发明的一个实施方案,还提供一种计算机软件产品,包括有形的计算机可读介质,其中程序指令被储存在所述计算机可读介质中,当指令被计算机读取时,计算机接收场景的暂时的深度图序列,所述场景包含静态背景和隐藏所述背景一部分的一个或多个移动的前景目标,并且执行上述方法中的一个或多个。

[0019] 参照附图,从下文实施方案的详细描述中能够更完整地理解本发明。

### 附图说明

[0020] 图1是根据本发明的一个实施方案的用于3D绘图和场景分析的系统的示意性图示;

[0021] 图2是示意性示出根据本发明的一个实施方案的3D成像组件和相关计算机的元件的框图;

[0022] 图3是示意性示出根据本发明的一个实施方案的用于3D场景分析的方法的流程图;

[0023] 图4是示意性示出根据本发明的一个实施方案的用于识别3D场景中地板的流程图;

[0024] 图5是根据本发明的一个实施方案的深度图的示意性正面表示,所述深度图示出了对应于已经在深度图中被识别的目标的块;

[0025] 图6是根据本发明的一个实施方案的深度图的示意性俯视图,所述深度图示出了已经在深度图中被识别的块的集群;以及

[0026] 图7是根据本发明的一个实施方案的深度图的示意性正面表示,所述深度图示出了目标在深度图中的分割。

### 具体实施方案

#### [0027] 概述

[0028] 深度图提供有价值的信息,尤其是当深度图被连续呈现时。然而,基于深度图的游戏和其他应用仅因为从这些深度图中捕获、处理和提取高级信息的固有困难而发展缓慢。

[0029] 下文中描述的本发明的实施方案通过提供用于处理深度图,尤其是用于定位和跟踪深度图序列中人形形状的强健、有效的方法来解决这些困难。这些深度图通常表示包含静止背景和一个或多个移动的前景目标(隐藏了背景的一部分)的场景。在公开的实施方案中,数字处理器识别和建模背景,该背景包括通常在室内场景的背景中出现的地板和墙壁。处理器使用模型,以便于检测遮挡了背景模型的前景中的人形形状的各部分(并因此会被认为从此属于前景),并可以将它与不在前景中的形状的其他部分关联。为了识别和跟踪感兴趣的形状,尤其是人形形状,处理器分析移动的前景目标。

[0030] 在一些公开的实施方案中,处理器通过识别每个深度图中不被前景目标隐藏的静态背景的各部分来构造背景模型。由于前景目标从一个图移动到另一个图,所以背景中的



不同部分通常在不同图中显现。处理器通过组合序列中的不同图的静态背景的不同部分来创建背景模型。

[0031] 附加地,处理器可以特别地识别具有可以被参数建模的特定几何形状的表面,尤其是平坦表面,例如地板和墙壁,这些表面构成背景的一部分。为此,在一个实施方案中,通过考虑相邻像素处的深度值,处理器计算深度图中像素的局部法线。为了识别地板和其他可能的平面表面,处理器使其局部法线大致指向相同方向的像素聚集。

[0032] 通过识别深度图中 3D 连接分量,处理器从场景中提取人形形状(以及其他可能感兴趣的目标)。每个连接分量包括一组相互邻近并具有相互邻近深度值的像素,这意味着相邻像素的深度值不能偏离多于(小)预定的阈值量。为了计算人形形状的几何模型,处理器识别和跟踪这些连接分量,所述几何模型然后可以被进一步处理,以识别高级特征,例如身体姿势、移动和手势。

[0033] 尽管理想地,每个人形形状应该作为单个连接分量出现,但由于场景本身的各方面原因,常常会发生的是给定的人形形状会作为一组两个或更多个分离的连接分量出现。这个分离使得完全识别人形形状变得困难,更不要说对高级特征建模。例如,在一些深度图中,人形形状可以被另一个目标部分地遮挡,从而该形状被分成两个或更多个分离的连接分量。在其他情况中,人形形状的一个部分可以部分地遮挡深度图中另一个部分。然而,在下文描述的实施方案中,处理器为构成给定人形形状的所有连接分量分配一个公共标识符,并且在跟踪这些连接分量时使用该标识符,即使当这些连接分量变得分离或者彼此遮挡时也是如此。这些实施方案的方法使得处理器能够解决异常情况并产生人形形状连续表示(即使面对遮挡时)。

[0034] 系统描述

[0035] 图 1 是根据本发明的一个实施方案的 3D 用户界面系统 20 的示意性图示。用户界面基于 3D 成像组件 22,该成像组件 22 捕获包括用户 28 和用户 30 的身体的至少一部分的 3D 场景信息。组件 22 也可以捕获场景的彩色视频图像。组件 22 输出包含 3D 图数据(以及可能的彩色图像数据)的帧序列到计算机 24,计算机 24 从图数据中提取高级信息。该高级信息经由 API 被提供给在计算机 24 上运行的应用,计算机 24 相应地驱动显示器 26。在画出的实施例中,例如,用户 28 和用户 30 通过移动他们的肢体和改变身体姿势与在计算机 24 上运行的游戏软件互动。

[0036] 为了重新构造包含用户 28 和用户 30 的场景的深度图(也被称为“3D 图”),计算机 24 处理由组件 22 产生的数据。在一个实施方案中,组件 22 在场景上投射斑点图案并且捕获被投射图案的图像。接着,基于图案中斑点的横向位移,组件 22 或计算机 24 通过三角测量(triangulation)计算场景中各点(包括用户身体表面上的点)的 3D 坐标。该方法的优势是不需要用户持有或佩戴任何类型的信标、传感器或其他标记。该方法给出相对于离组件 22 特定距离处的预定参考平面的场景中各点的深度坐标。用于此类使用投射图案的、基于三角测量的 3D 绘图的方法和设备例如在 PCT 国际申请 W02007/043036、W02007/105205 和 W02008/120217 中描述,这些申请的公开内容以参引方式被纳入本文。或者,基于单个或多个本领域公知的摄像机或其他类型的传感器,系统 20 可以使用其他 3D 绘图方法。

[0037] 在本实施方案中,在用户移动他们身体的同时,系统 20 捕获和处理包含用户 28 和用户 30 的三维(3D)图序列。在组件 22 和/或计算机 24 上运行的软件处理 3D 图数据,以

提取对应于场景中用户的人形形状。为此（如将在下文详细描述），软件识别和建模静态场景背景，包括平面表面——例如地板 32 和墙壁 34，以及其他背景目标，例如在该实施例中的窗户 36。为了检测和跟踪用户 28 和用户 30 的人形形状以及其他前景目标（例如椅子 38），该软件将图像中的前景部分与所检测的人形形状的部分结合。

[0038] 为了提取每个身体的骨架，组件 22 或计算机 24 通常分析人形形状的几何特征，包括用户的手和关节的 3D 位置以及方向。为了识别手描绘的手势，也可以分析手在多个帧上的轨迹。然而，系统 20 的这些操作方面超出了本专利申请的范围。骨架和手势信息经 API 提供给运行在计算机 24 上的应用程序。作为对骨架和 / 或手势信息的响应，该程序例如可以移动和修改呈现在显示器 26 上的图像。

[0039] 计算机 24 通常包括通用计算机处理器，该处理器被用软件编程以执行下文描述的功能。所述软件例如可以经网络以电子形式被下载到处理器，或者可以作为替代地在有形的、非短暂介质——例如光、磁或电子存储介质——上提供。替代地或额外的，所描述的计算机的部分或全部功能可以在专用硬件——例如定制集成电路或者半定制集成电路或者可编程数字信号处理器 (DSP)——中实施。尽管在图 1 中以实施例方式将计算机 24 显示为成像组件 22 的分离单元，但是计算机的部分或全部处理功能可以由合适的微处理器和软件执行，或者由成像组件壳内或者与成像组件相关联的专用电路执行。

[0040] 作为另一个替代，至少部分的处理功能可以由合适的处理器执行，所述处理器与显示屏 26（例如在电视机中）或者任何其他合适类型的计算机化设备——例如游戏控制器或媒体播放器——集成。组件 22 的感应功能同样可以集成到计算机或其他被传感器输出控制的计算机化装置。

[0041] 图 2 是示意性示出根据本发明的一个实施方案的系统 20 中成像组件 22 和计算机 24 的元件的框图。成像组件 22 包括照射子组件 42，照射子组件 42 在感兴趣的场景上投射图案。深度成像子组件 44——例如合适配置的摄像机，捕获场景上图案的图像。通常，照射子组件 42 和成像子组件 44 在红外范围内运行，尽管也可以使用其他频谱范围。可选地，彩色摄像机 46 捕获场景的 2D 彩色图像，并且麦克风 48 也可捕获声音。

[0042] 处理器 50 从子组件 44 接收图像，并且将每个图像中的图案与存储器 52 中存储的参考图案对比。参考图案通常通过预先投射图案到距离组件 22 一已知距离的参考平面来捕获。一般来说，该平面垂直于子组件 44 的光轴。处理器 50 计算子组件 44 捕获的图像中图案的各部分相对于参考图案的局部位移，并将这些位移转化为深度坐标。该过程的细节例如在 PCT 国际申请 WO 2010/004542 中描述，该申请的公开内容以参引方式被纳入本文。作为替代地，如前文注意到的，组件 22 可以被配置为通过本领域公知的其他手段——例如立体成像或飞行时间 (time-of-flight) 测量——产生深度图。

[0043] 处理器 50 经由通信链路 54——例如通用串行总线 (USB) 连接，将深度图输出到计算机 24 合适的接口 56。计算机包括带有存储器 60 和用户接口 62 的中央处理单元 (CPU) 58，该计算机驱动显示器 26 并且也可以包括其他部件。如上文注意到的，成像组件 22 可以替代地仅输出来自子组件 44 的原始图像，并且上文描述的深度图计算可以由 CPU 58 用软件执行，或者下文中提到的 CPU 58 所具有的功能的至少一部分可以由处理器 50 执行。用于从深度图中提取高级信息的软件可以运行在处理器 50、CPU 58 或者二者上；并且在本专利申请和权利要求中的所提到的 CPU 或处理器应该被理解为指的是这些元件中的一个或二者。

如下文将详细描述，CPU 58 通常运行一个或多个应用程序，所述应用程序基于经 API 由中间件提供的信息驱动用户接口 62。

#### [0044] 场景分析方法

[0045] 图 3 是示意性示出根据本发明的一个实施方案的用于 3D 场景分析的方法的流程图。为了清楚，参考上文描述的系统 20 的部件以及图 1 和图 2 示出的元件来描述该方法。然而，该方法的原理基本可以在任何合适的用于深度绘图和分析的系统中实施。同样为了描述的方便和清楚，假设该方法的步骤由 CPU 58 在软件中执行，但是替代地，这些步骤的一部分或全部可以由处理器 50 或另外的合适的数字处理器执行。

[0046] 在深度图输入步骤 70，计算机 24 从成像组件 22 接收暂时的深度图序列。深度图通常由成像组件在固定的帧速率下以视频图像的通用格式输出，带有光栅顺序的像素值。（因此在本文中的序列中的图也被称作图帧 (map frame)，或者简化为帧）。深度图中的像素值可以是实际的深度坐标值或图案视差 (disparity) 值（以像素为单位），所述像素值与实际的深度值有已知的单调关系。该方法在序列中每个图帧上迭代执行。

[0047] 在地板建模步骤 72，CPU 58 分析每个深度图以定位和建模地板 32。该步骤将在下文参考图 4 详细描述。地板位置（或者静态背景的其他部分）的明显变化可以给 CPU 提供摄像机已经移动的指示，例如由于被一个用户偶然或故意的接触。CPU 将已经被识别属于地板平面的像素从深度图中除去，通常通过将像素的深度值赋值为零来实现。

[0048] 在背景构造步骤 74，遍经图帧序列，CPU 58 逐渐建立场景背景模型。对于每一个像素，CPU 存储整个序列上的最大深度值（或者等价地，与参考图像比较的最大视差）。背景中的像素的深度值是该像素随时间的最大视差值。如果对于多于特定数量的连续帧，像素都没有可提供的深度值（不是因为阴影），则背景模型会在此像素处丢失其值。如果像素是由于阴影不具有深度值，并且投射该阴影的目标比当前背景深度值还要远，则像素会失去其当前背景值。如果背景模型中的像素不具有深度值且在当前帧中具有一个值，并且包含该像素的前景连接分量与背景相邻，则该像素的背景值会变成当前值。

[0049] 对于属于给定图中移动目标的像素，CPU 58 可以在这些像素被前景目标阻挡之前或者在前景目标已经移走之后为这些像素存储背景值。例如，在图 1 示出的场景中，窗 36 的区域被用户 30 的头部遮掩，但是当用户在随后的图帧中改变其姿势时，同一区域可以显现给成像组件 22。因此，遍经多个帧，作为在每个帧中显现的背景的不同部分的组合，CPU 58 累积地建立背景模型。

[0050] 如上文示出的，前景目标——例如用户 28 和用户 30 的身体——可以在其后面的目标上投射阴影，结果是成像组件 22 将不能提取阴影中的深度值。为了避免由于阴影而造成的深度图中的假象，CPU 58 将邻近一个目标的没有深度值的区域识别为属于该目标。阴影被允许包括只到特定最大阴影宽度的像素，而特定最大阴影宽度是由深度图中前景部分和背景部分之间的图案位移差所决定。几何上示出最大阴影宽度（以像素为单位）是前景目标和背景目标之间图案位移差的最大值的 1/8。CPU 58 可以计算阴影图像，指示图的阴影区，以在控制背景模型的更新时作为屏蔽 (mask) 使用。

[0051] 为了消除随后处理中可能增加噪声的假象，CPU 58 可以“转储 (dump)”背景模型。例如，如果 CPU 检测给定区域中的连接分量处于的深度接近在该区域中背景模型的深度，则它可以根据所述连接分量的值来修正背景模型的深度值。另一方面，如果 CPU 已经跟

踪前面帧的连接分量,并且识别它为人形形状的一部分(或者至少是人形形状一部分的候选),则保持分离,即使当连接分量接触背景时也是如此。遍经特定数量的帧,带有前景深度值的且没有被识别作为移动目标一部分的像素同样可以被附到背景中。

[0052] 在分量检测步骤 76, CPU 58 移走地板,并随后试图在前景像素和仍留在图中的之前被检测的像素之中发现可以对应于人形形状各部分的 3D 连接分量。CPU 可以将之前没有被识别作为人形形状各部分并且不是前景的像素“移除”到前景深度值的范围外(通过分配所述像素人为深度值),以使得这些像素不被认为是前景像素的分割,但是仍被认为是用于其它用途(将在下文描述)的场景分量。

[0053] CPU 分配标识符来标注它已经发现的每个 3D 连接分量,并计算连接分量的属性,包括:

[0054] 1. 像素数量;

[0055] 2. 包含所述连接分量的边界框;

[0056] 3. 质量中心;

[0057] 4. 面积;

[0058] 5. 前景像素的数量(假设连接分量的一部分具有背景深度值);

[0059] 6. 其他附近连接分量的标识符。作为替代或额外的,可以计算和存储其他属性。CPU 在跟踪人形形状从一帧到另一帧的移动中使用标识符和连接分量属性。因此,给定帧中的连接分量通常由前景像素和被发现属于前一帧中人形形状的像素组成。

[0060] 在集群检测步骤 78, CPU 58 处理深度图以识别新的 3D 集群。这些集群可以对应于没有在前面的帧中识别的人形身体,并且当系统开始运行时,其本身被用于在人形形状的初始识别中。步骤 78 在下文中会参考图 5 和图 6 做更详细地描述。

[0061] 在步骤 78, 识别的集群由具有近似相同深度值的相邻的像素块构成,所述块被聚集成具有相互邻近深度值的 3D 集群。跟踪在连续帧中出现的且具有不大于预定阈值(例如,10cm)的相对移动的集群。CPU 在每个候选集群周围建立边界框并跟踪它在多个连续图帧上的移动。CPU 将累积移动大于一些最小阈值(例如,20cm)的候选识别为移动目标。为此,移动被定义为集群的两个边缘的移动,即左侧和右侧都应大体在相同方向上移动。

[0062] 在分割步骤 80, CPU 58 处理在步骤 76 和步骤 78 发现的连接分量和集群,为的是分割和识别可以对应于人形身体各部分的分量。在分量集合步骤 82, 它将相邻的分量集合在一起,以重构人形身体形状。该步骤可以包括若干子操作。在简单的情况中,人形形状会作为单个 3D 连接分量出现,并且以此被标注。这种形状可以很容易被识别,例如通过执行校准过程,其中用户以规定的姿势面对成像组件 22 站立。所述过程帮助 CPU 58 在计算机 24 上运行游戏或其他应用的同时,随着用户随后的移动而跟踪用户身体各部分,

[0063] 在其他情况中,在粘合(gluing)子步骤 84, CPU 58 可能需要集合不同深度的两个或更多个连接分量,以构成单个人形身体形状。如上文注意到的,如果整个形状在前一个图帧中被识别,则公共标识符和元数据可以被用于保持各分量之间的联系。额外地或作为替代地,互相邻近的未标注的运动的连接分量可以被集合,并且被分配一个公共标识符(即使之前没有被标注)。

[0064] 另一方面,当用户接触前景目标(包括另一个用户)时,所述用户和目标可以暂时地作为单个 3D 连接分量出现在所述图帧的一个或多个中。在这种情况下, CPU 58 可以错

误地分配给它们一个公共标识符。随后,当用户从目标处移开时,在深度图中可以有二个带有相同标识符的独立的连接分量。在这种情况下,CPU 可以重设标识符并重建实际对应于用户身体的 3D 连接分量(或多个分量)。

[0065] 作为替代地,在这类情况中,CPU 58 可以使用在前一个帧中分配的标识符,以将属于用户身体的合并连接分量的被标注像素从属于与所述身体接触的其他目标的像素中分离。因此,尽管有接触,但 CPU 能够正确地分割身体。特别是当所述其他目标是另一个用户时,CPU 会为属于场景中二个(或更多个)不同人形形状中的每一个的像素分配不同的特定标识符。然后,已经被分配到一个帧的不同标识符可以被用在随后的帧中,以将合并连接分量的像素中的每一个分配到正确的身体。

[0066] 当在一个或多个深度图中一部分人形形状被另一个目标遮挡时,会出现一些具体的困难。在遮挡检测步骤 86,CPU 58 检测并解决这些异常。该步骤会在下文参考图 7 做更详细地描述。

[0067] 在总结步骤 82,在大部分情况中,所有的人形形状已经被识别。对于这些人形形状的每一个,CPU 58 分配一个公共标识符,以标注组成所述形状的所有 3D 连接分量。CPU 用合适的标识符标注属于这些连接分量的所有像素。(不属于任何人形形状的像素接受空的标识符。)结果是得到了分割的深度图,该深度图随后可以被进一步处理,以提取人形形状的骨架。同时,在图 3 方法的下一次迭代中,在处理 CPU 在步骤 70 接收的下一个深度图时,被标注的像素的矩阵被用作偏差(departure)点。

[0068] 地板检测

[0069] 图 4 是示意性示出根据本发明的一个实施方案的在步骤 72 中使用的用于检测 3D 场景中地板的方法的详情的流程图。该地板在场景中被限定为平面表面,其被期望位于深度图的下部。场景中的任何平面都满足形式为  $AX+BY+CZ+D=0$  的方程,其中  $(X,Y,Z)$  是空间中坐标,并且  $A,B,C$  和  $D$  是常量。对于基于投射图案的视差值  $s$  的深度图(在每个像素  $(x,y)$  处计算),平面方程可以关于像素坐标和位移值  $(x,y,s)$  进行重铸,其中常量被适当调整。因此,图 4 的方法以及图 2 中示出的方法的其他方面可以在原始图数据上执行,甚至在计算实际深度值之前。在  $(x,y,s)$  域中的计算细节在上文提及的临时专利申请中描述。

[0070] 在地板发现步骤 90,对于每个连续的深度图,通过检查是否已经在之前的图帧中发现地板模型,CPU 58 开始发现地板的过程。如果是这样,在新模型计算步骤 92,CPU 为存在的地板计算新模型。为此,CPU 创建包含深度图中所有像素的地板图像,所述像素在存在的地板的平面(在预定公差内)上,或低于所述平面。(在步骤 74,地板图像随后可以从深度图中移除。)在模型标记步骤 94,CPU 计算新模型的置信度分数(confidence score),并且在模型评估步骤 96 使用所述分数,以决定接受或拒绝当前图帧的地板平面。所述置信度分数例如可以基于地板平面中像素的数量与低于该地板平面的像素数量的比较(并因此看得出是错误的)。在地板替换步骤 98,如果存在的地板平面是可接受的,则 CPU 可以基于地板图像调整平面参数。例如,CPU 可以对地板图像中的像素执行平面的最小二乘方拟合,以调整平面参数。

[0071] 在步骤 96,如果在地板图像中有太多的像素(例如,多于像素的  $1/6$ ) 低于存在的地板平面,则 CPU 拒绝新的地板模型。地板模型失效的原因可以是,例如由于用户 28 和用户 30 中一个有意或无意的接触,使得成像组件 22 已经移动。在摄像机运动检查步骤 100,

CPU 估计被移动的可能性。CPU 例如可以计算背景模型中深度值的变化速率。如果在变化速率中发现大的、突然的增加,则 CPU 得出摄像机已经被移动的结论。否则,地板模型被保留到下一帧。

[0072] 如果在步骤 100, CPU 发现摄像机已经移动,并且新的地板模型是无效的,则在地板无效步骤 102, CPU 将当前地板模型设置为错误,并随后继续发现新的地板。CPU 类似地在任何新图帧序列开始处开始新的地板搜索。

[0073] 紧随步骤 102,在穷尽搜索步骤 104, CPU 首先可以尝试穷尽性地地板搜索。在该步骤, CPU 识别深度图中(或者为了计算效率,可能在深度图的抽取版本(decimated version)中)所有水平像素,即,局部法线指向上的像素。局部法线是通过比较相邻像素的深度值来计算。基于透视规则,每个水平像素的空间坐标预测特定摄像机离地板的高度,假设该点在地板上。根据各自的摄像机高度预测值, CPU 聚集水平像素。随后, CPU 在那些包含多于一个预定最小区域的集群中选择位于空间中最低平面的集群作为地板基准。

[0074] 在搜索检查步骤 106, CPU 检查是否发现合适的集群。如果发现,如下文描述的,通过遍经集群中像素的线性回归, CPU 计算地板平面并且进入标记步骤 112。如果没有发现合适的集群,则在模型猜测步骤 108, CPU 采用粗略估计的替代过程。在该步骤中, CPU 随机地从深度图的下三分之一中选择三个像素,并计算空间中包含这三个对应点的平面。在模型检查步骤 110, CPU 检查该平面是否水平(法线指向上)。如果水平,则 CPU 接受该平面作为地板。否则,直到下一个帧才终止地板搜索。

[0075] 在模型标记步骤 112, CPU 58 评估在步骤 104 或步骤 108 发现的地板模型。该步骤以和上文描述的步骤 94 基本相同的方式进行。在新模型计算步骤 114,为了比较, CPU 计算新模型。该步骤以和上文描述的步骤 92 基本相同的方式执行。在新模型标记步骤 116, CPU 评估该新模型,在模型比较步骤 118, CPU 将评估结果与从步骤 112 得到的之前的模型分数比较,并且如果新模型分数更好则用新模型代替之前模型。CPU 可以重复步骤 114-118,直到其不能获得对模型的进一步改进,此时过程直到下一个帧才终止。

[0076] 尽管上文描述的方法特别涉及识别深度图中的地板(并在后续的分析中将地板移走),对该方法进行稍加改进,则可以使用它来发现(以及移除,如果期望)其他平面表面,例如墙壁、天花板、甚至桌面。再者,上文描述的方法可适合识别其他类型的能够参数表示的非平面表面,例如可以出现在深度图中的球面表面和抛物线表面。这些表面的参数可以从深度图中得出,例如,通过对深度图中点的合适的 3D 函数的线性最小二乘方拟合。可以用于此目的的拟合技术例如在 Flannery 等人的 Numerical Recipes in C: The Art of Scientific Computing, 第二版 (Cambridge University Press, 1992), 671-681 页中描述,该内容以参引方式被纳入本文。

[0077] 提取人形形状

[0078] 图 5 和图 6 分别示意性示出根据本发明的一个实施方案的深度图 120 的主视图和俯视图。图 5 示出深度图中对应于目标的块 122、124 和 126。图 6 示出了基于所述块在 XZ 平面上的投影的块的集群。这些图示出了连接分量检测步骤 74 的实施方式(图 3)。

[0079] 为了发现图 5 中示出的块, CPU 58 沿着 X(水平)方向区分深度图。为此,可被发现的没有深度值的像素被分配一个随机的、非常大的深度值。导数值大于特定阈值的像素被归类为边缘。为了避免边缘假象,在区分深度图之前,前景目标(如上所述)的阴影被分

配与投射该阴影的目标相同的深度值。随后,CPU 搜索被区分的深度图的行,以发现被之间没有边缘的左边正边缘和右边负边缘定界的像素行程 (run)。

[0080] CPU 创建包含这类行程的块 122、124、126,所述行程在多个相邻行发生,并且符合下列标准:

[0081] ●行程在至少最小数量(例如,五个)的连续行中发生。

[0082] ●行之间的深度值差不大于预定阈值。

[0083] ●行程在 X 方向至少重叠最小的重叠比例(例如,块中至少 80%的最短行程)。

[0084] CPU 计算每个块的几何特征,例如边界坐标、中间深度值、质量中心和面积。

[0085] 如图 6 中示出的,CPU 58 随后将块集成成 3D 集群。该图示出了深度图在 X-Z 平面的视图,其中深度(Z)从图的顶部增加到底部。三角形区域 128 表示成像组件 22 的观察范围。这里块 122、124 和 126 可以被看作处于各自不同的深度范围的集群。块 122 对应于中间深度处的人形形状的各部分,如在图 5 右侧看到的。块 124 对应于前景中的椅子,而块 126 是背景的一部分。

[0086] 如果块符合特定几何标准,则 CPU 将这些块集合在一个集群中。例如,CPU 可以要求块的各自边界框在 X-Z 平面至少被一些最小量重叠,并且每个块的质量中心被包含在集群中其他块的边界框内。如果块的面积和至少是矮小的人的面积,则集群被标记为属于人形形状的候选。

[0087] 如之前注意的,除了符合上述几何标准以外,如果经过多个图帧集群通常保持一些移动,则集群被认为是人形形状的候选。因此,如果给定的集群出现在和前一帧中被识别的人形形状相同的大体位置,则 CPU 保持该识别并且使用与前一帧相同的标识符标注该集群。

[0088] 否则,为了确定集群是否是人形形状(已经持续若干连续帧),计算机将集群中的块投射在地板平面上(在步骤 72 发现),并且计算投影的几何属性,例如面积、质量中心和边界框。深度图中所有落入边界框的连接分量被假设为属于相同的候选形状。如果所有这些连接分量的累积面积大于给定阈值,则 CPU 集合集群和其相关的连接分量,并为它们分配公共的标识符作为人形形状。CPU 计算和存储该形状的属性,并且在后续的图帧中跟踪该形状,如上文描述的。

[0089] 因此,再次参考图 5 和图 6,CPU 将发现块 122 的集群符合几何标准,并且将识别该集群属于计算机用户的人形形状。块 124 的集群太小,将被丢弃。尽管块 126 的集群大小足够,但是由于其是静态的因此最终也会被丢弃。

[0090] 图 7 是根据本发明的一个实施方案的深度图 130 的示意性正面表示,所述深度图 130 示出了深度图中目标的分割。该图基于一个人的实际深度图,例如可能在实际计算机游戏中发生的,该人迅速提起他的右腿,以进行空手道中的跳踢。人的身体作为 3D 连接分量 132 出现,3D 连接分量 132 在之前的图帧中出现而因此具有标识符,并被用计算的几何特性标注。新的连接分量 134——对应于腿——突然出现在该图中的大体比身体其他部位更短的距离处。连接分量 134 部分地遮挡身体分量 132 并且将该分量从另一个对应于人的手臂的连接分量 136 中切去。

[0091] 为了处理这类情况,当连接分量——例如分量 134——突然出现在图帧中(并因此没有预先分配的标识符)时,CPU 58 搜索附近其他的已经预先被识别并被标注为人形形

状的连接分量。如果新的连接分量在 X 和 Y(图 7 中的水平和垂直方向)中相邻,并且深度接近预先识别的连接分量(在该实施例中为分量 132),且在其之间没有深度不同的区域(除了无深度区域和遮挡目标),则 CPU 将连接分量联系在一起并且为二者分配公共的标识符。因此,在图 7 示出的实施例中,CPU 会将分量 134 和分量 136 附接到分量 132。CPU 会分配分量 132 的标识符给分量 134 和分量 136,并且会因此辨认并跟踪属于单个人形形状的这些分量。

[0092] 在这类情况中,可以使用在结合连接分量到单个身体单元的其他考虑,包括各自边界框的交叉和他们质量中心之间的距离。

[0093] 在步骤 86(图 3),系统 20 异常处理的另一类型是一个人形形状被另一人形形状或其他前景目标遮挡。这类情况在图 1 中示出,其中用户 28 的左腿遮住了用户 30 的身体。这类情况的结果是当 CPU 58 分析画出的场景中的深度图中的 3D 连接分量时,用户 30 的右腿的下部分会与身体的其他部分分离。

[0094] CPU 58 使用上文描述的数据结构处理这类情况。假设在之前的图帧中用户 30 的身体被完全识别而没有遮挡,则 CPU 会将一个与用户 30 身体的其他部位相同的公共标识符与该用户 30 的右腿区域中的连接分量相关联。再者,CPU 会认出:对应于用户 30 右腿的连接分量和对应于用户 30 身体的连接分量位于相同的深度范围,而对应于用户 28 身体的连接分量——包括分离用户 30 身体各部分的左腿——则处在较小的深度。在这些情况下,CPU 会得出结论:对应于用户 30 的腿和身体的分离的连接分量实际上是同一人形形状的不同部分,并且会将各连接分量用一个公共标识符“粘合”在一起。

[0095] 一旦已经检测并解决了这类遮挡,对于每个人形形状,CPU 58 存储遮挡所述人形形状的其他人形形状(以及其他前景目标)的记录和所述人形形状遮挡的其他人形形状的记录。这些记录对于解决发生在随后帧中的遮挡是有用的。类似的技术可以用在检测 and 解决一个人触碰另一个人的情况。

[0096] 可选地,CPU 58 可以跟踪人形形状各部分从一帧到另一帧的运动,并由此估计这些身体部分运动的速度。CPU 可以使用该速度推断用户身体各部分的位置(即使当他们被其他目标遮挡时),并且因此可以保持对人形形状被遮挡的各部分在深度图中的恰当识别。

[0097] 如之前注意到的,尽管为了解释的清楚和方便,上文的实施方案是参考特定类型的深度绘图系统和特定类型的计算机应用来描述,但是本发明的原理可以类似地被应用到基本任何合适系统的任何类型的深度图中,而不管其要求如何。具体地,上文描述的 3D 连接分量分析和分割的方法可以用于从深度图中提取其他类型目标,并且不只是人形形状。作为结果的被分割的深度图可以直接显示给用户以用于用户观察,并且可以被用作基本任何类型的应用程序的输入,所述应用程序接受分割的深度图或其他图像作为其输入。

[0098] 因此,应理解,上述实施方案是通过实施例的方式来叙述,但本发明不限于上文所具体示出和描述的内容。相反,本发明的范围包括上述各种特征的组合和子组合及其变化和变型,这些对于阅读了前文描述的本领域技术人员是可以想到的,并且没有在现有技术中公开。



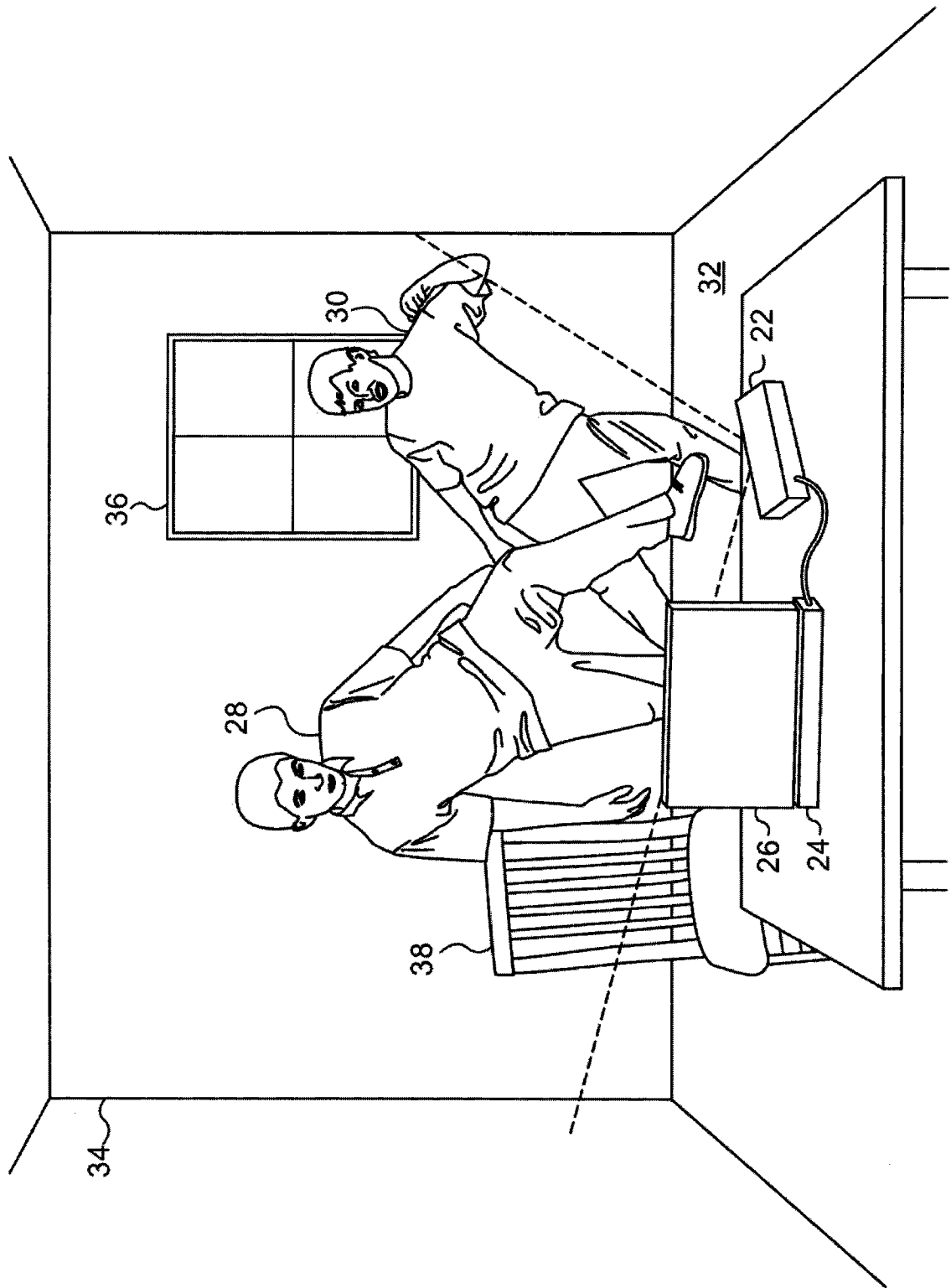


图 1

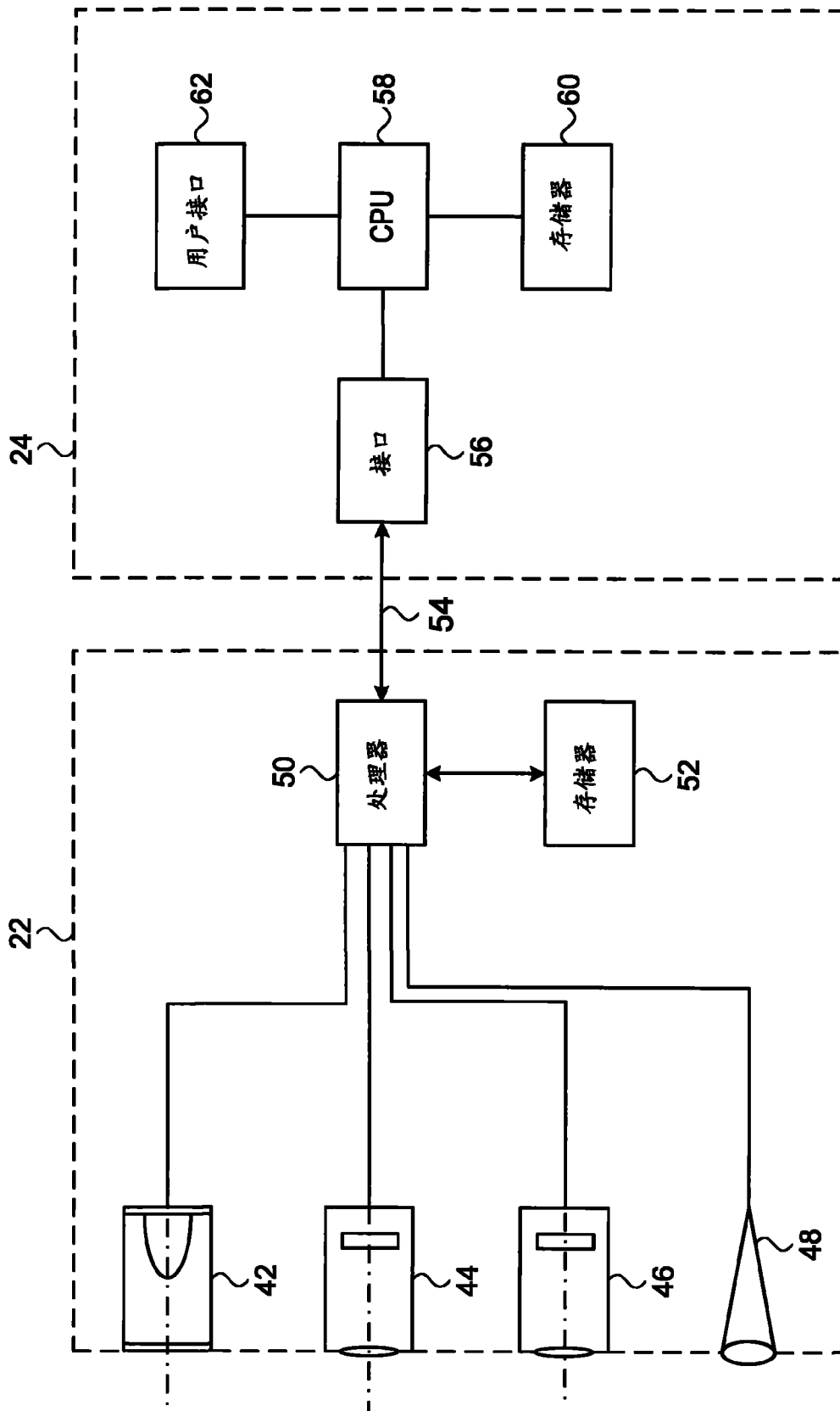


图 2

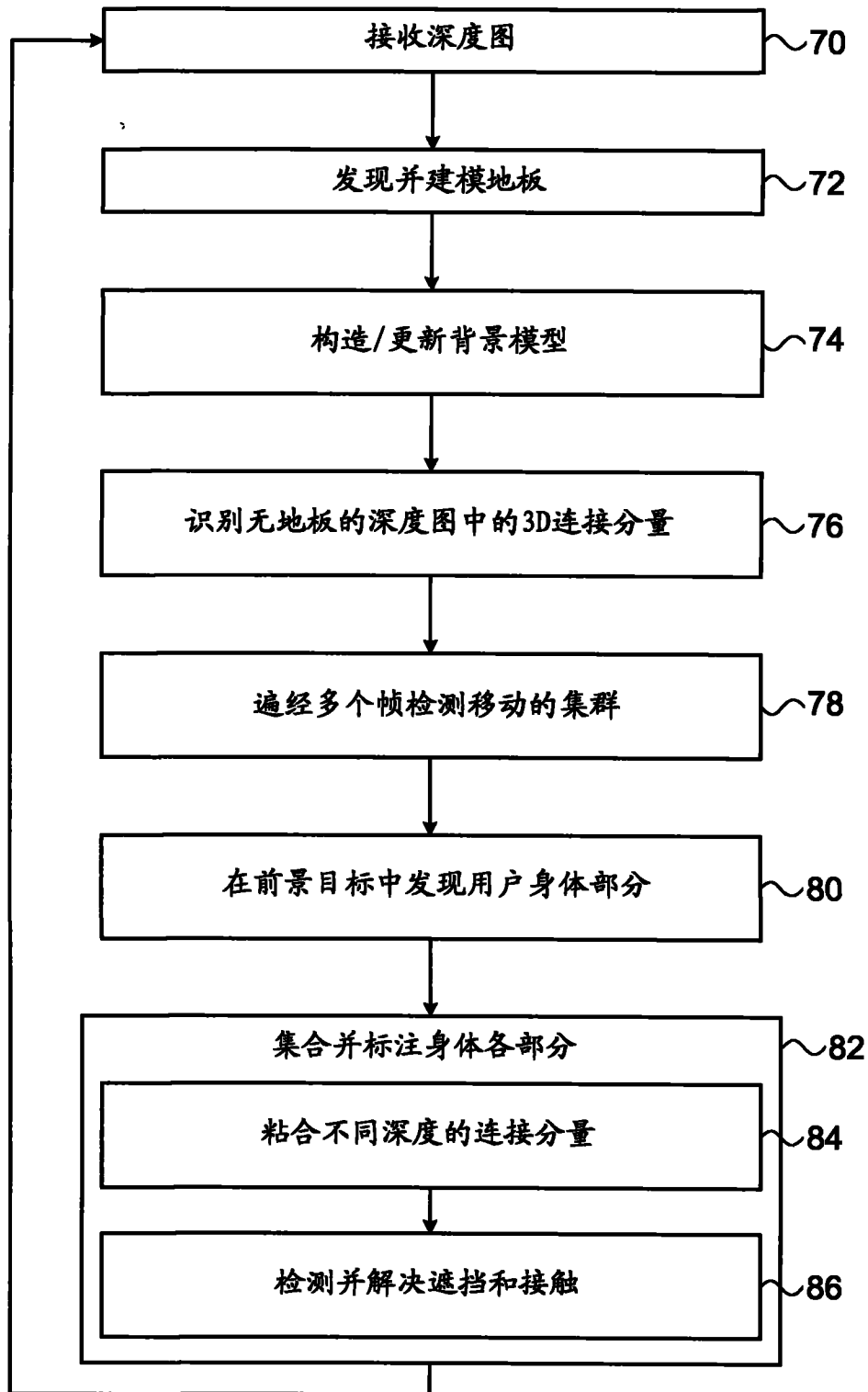


图 3

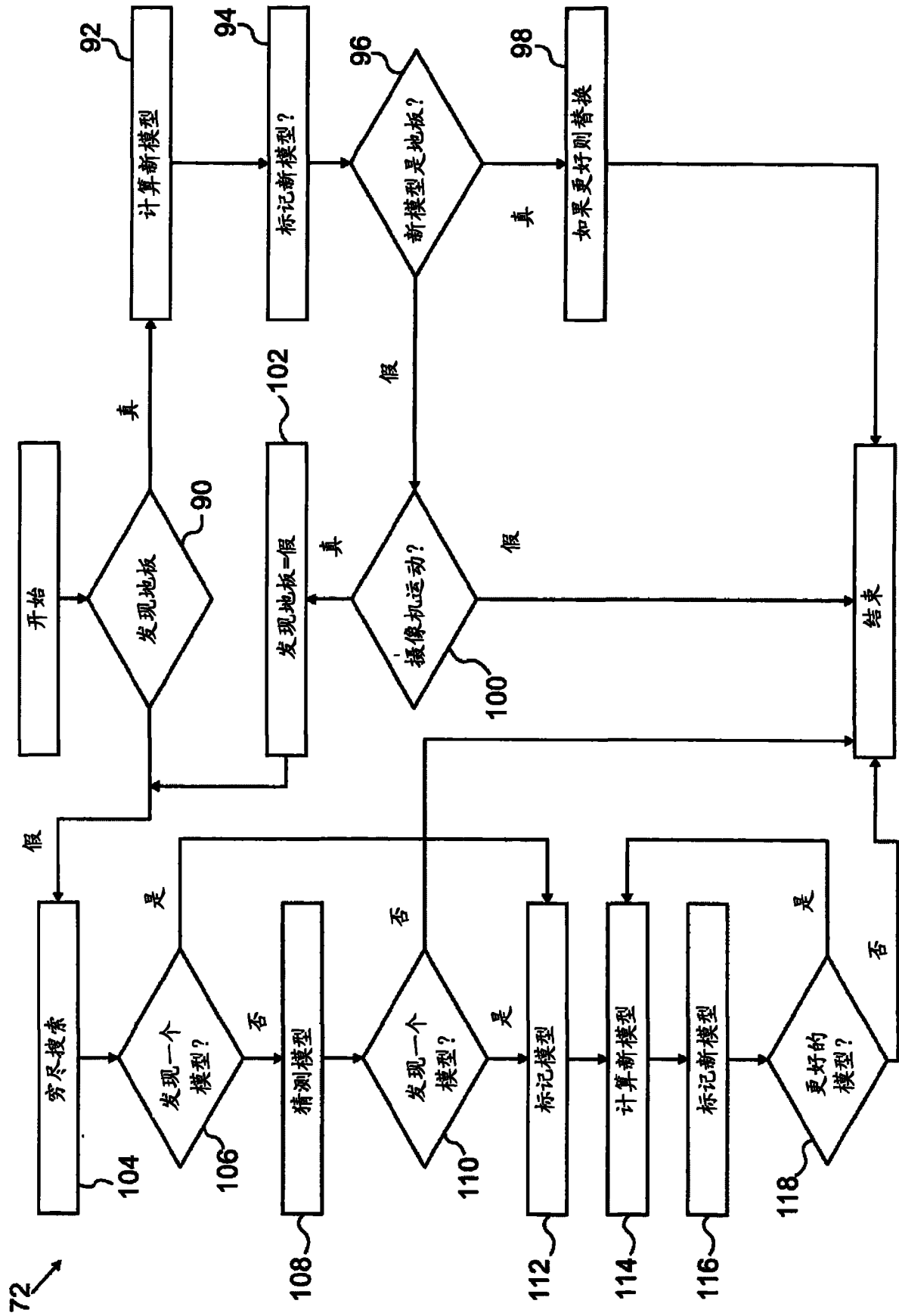


图 4

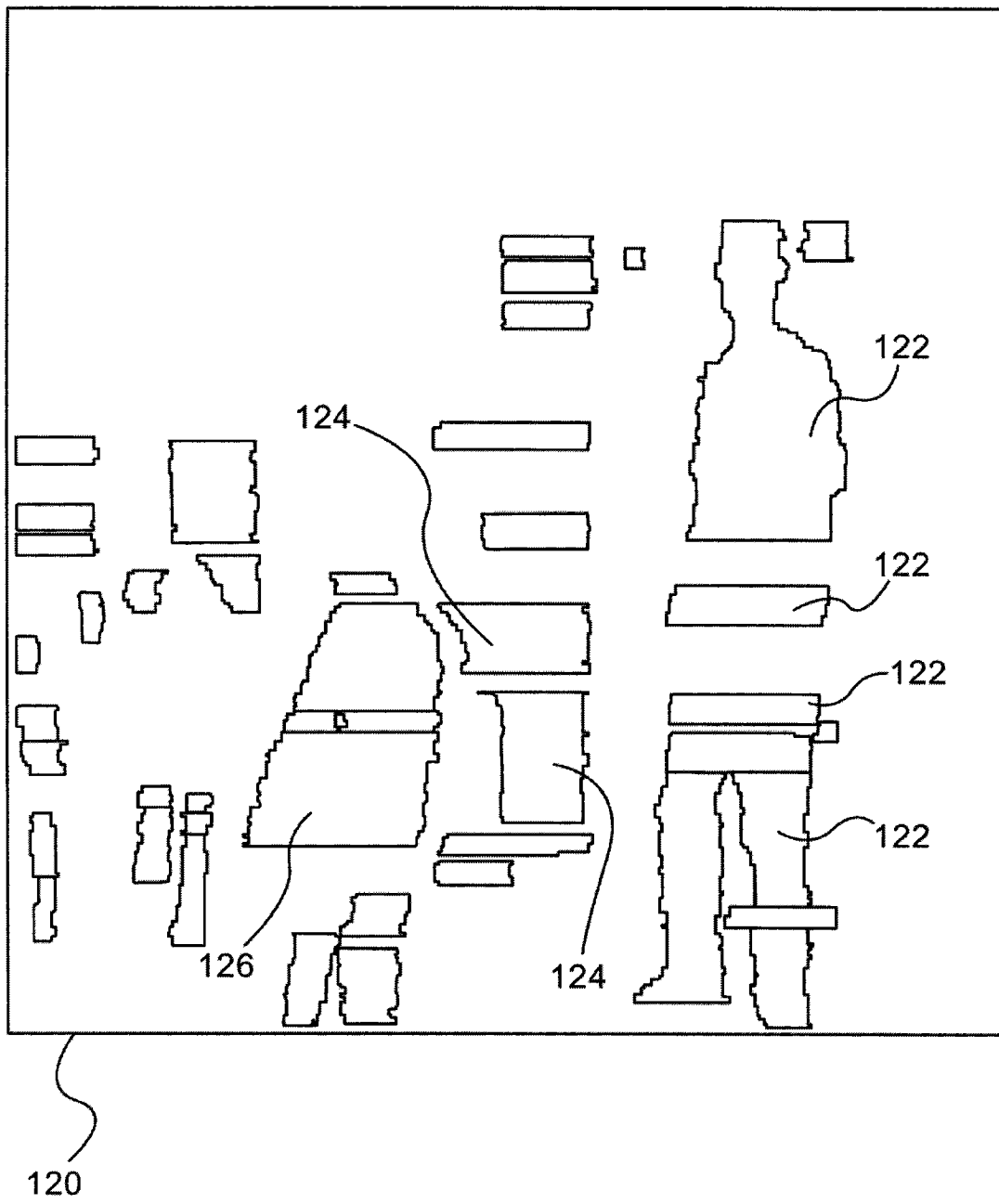


图 5

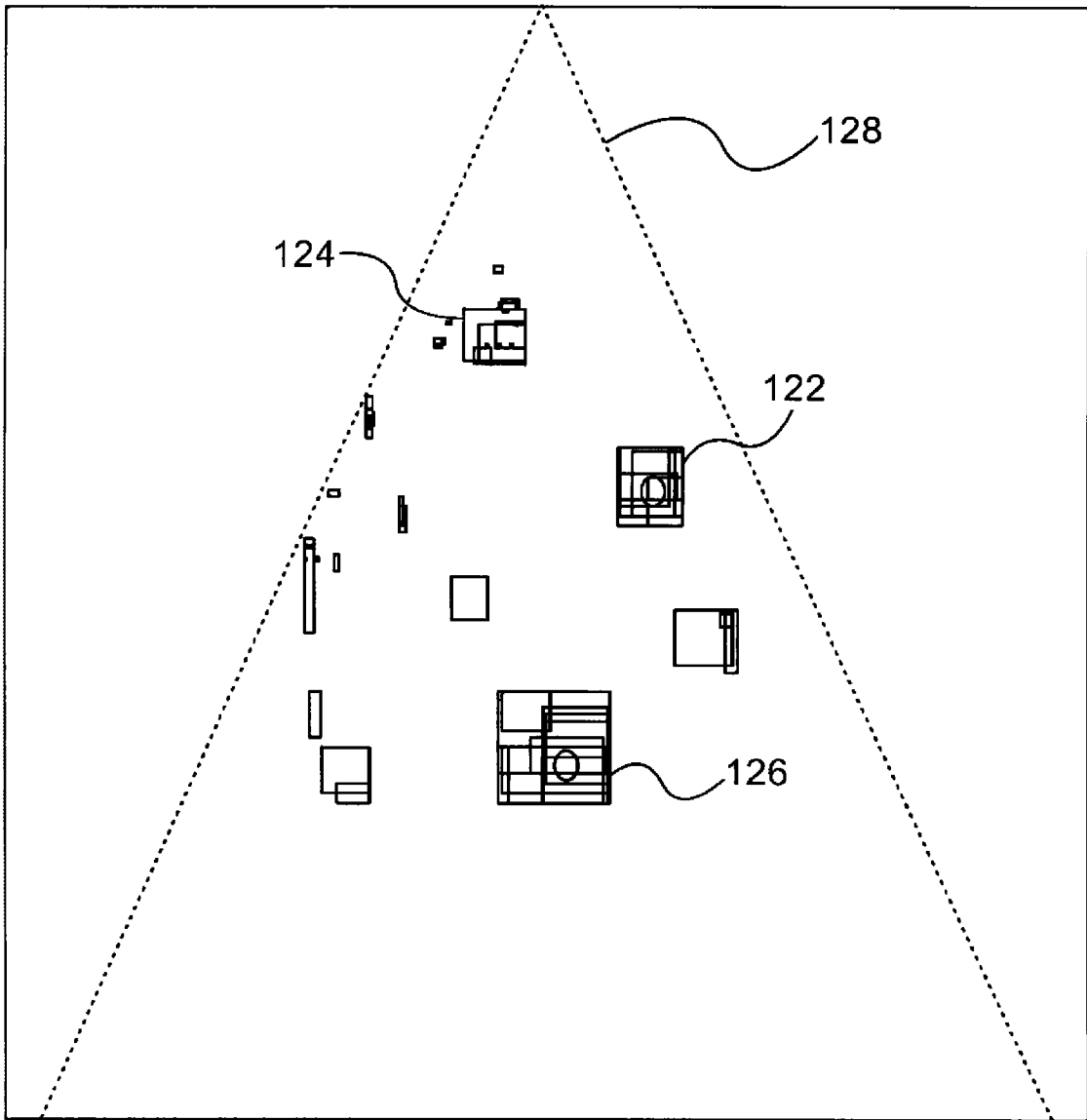


图 6

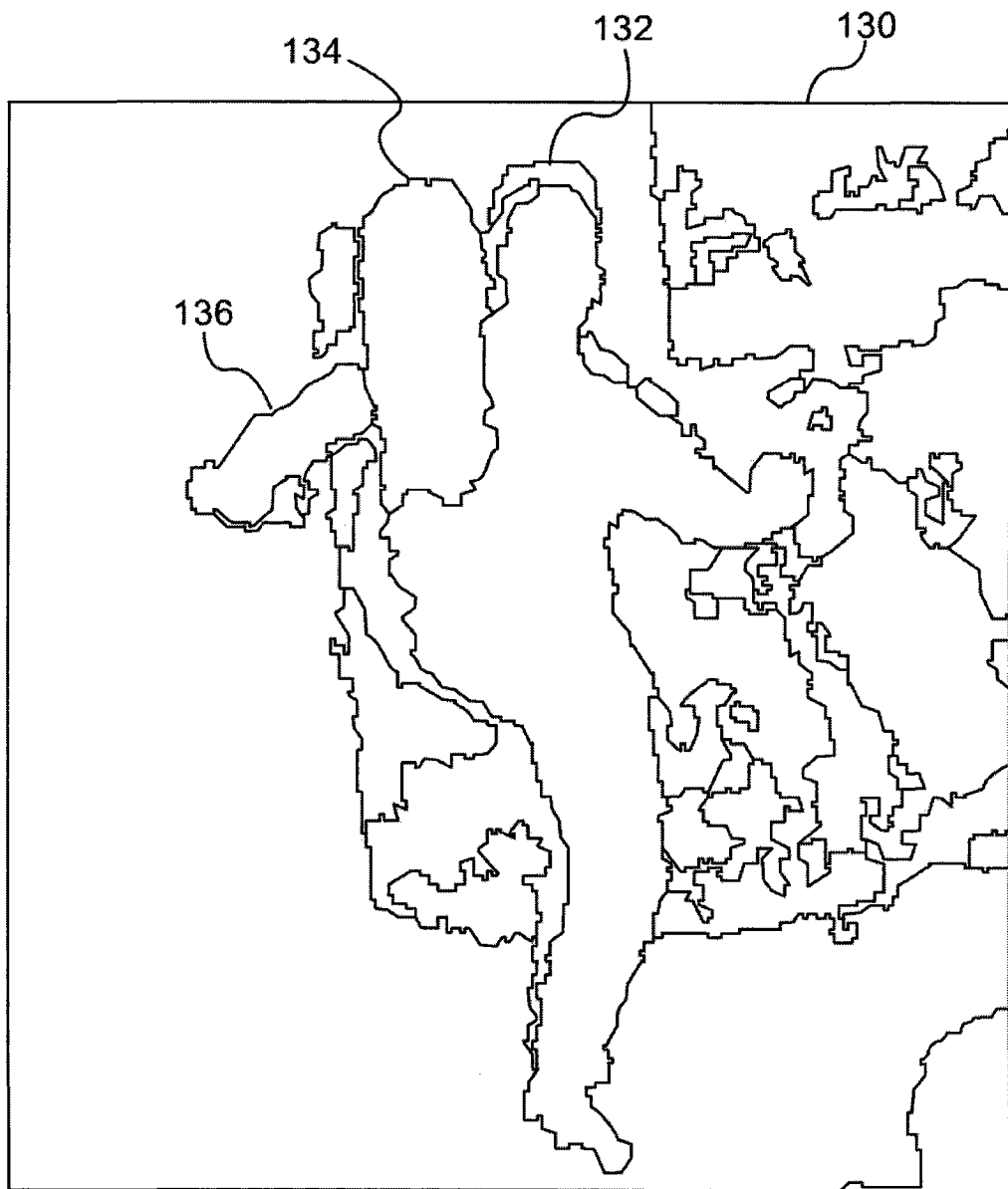


图 7