



(12)发明专利

(10)授权公告号 CN 105830009 B

(45)授权公告日 2019.05.14

(21)申请号 201480069055.6

(72)发明人 高大山 钟辛 A·A·坎哈达

(22)申请日 2014.12.19

S·比思 S·D·拉韦尔

(65)同一申请的已公布的文献号

(74)专利代理机构 北京律盟知识产权代理有限公司 11287
代理人 宋献涛

申请公布号 CN 105830009 A

(43)申请公布日 2016.08.03

(51)Int.Cl.

(30)优先权数据

G06F 3/0484(2013.01)

61/919,627 2013.12.20 US

G08B 13/19(2006.01)

14/575,945 2014.12.18 US

G06T 7/73(2017.01)

(85)PCT国际申请进入国家阶段日

G06T 7/254(2017.01)

2016.06.17

G06T 7/269(2017.01)

(86)PCT国际申请的申请数据

(56)对比文件

PCT/US2014/071567 2014.12.19

WO 2004/045215 A1,2004.05.27,

(87)PCT国际申请的公布数据

US 2004/0258152 A1,2004.12.23,

W02015/095743 EN 2015.06.25

US 2007/0126868 A1,2007.06.07,

(73)专利权人 高通股份有限公司

CN 101753999 A,2010.06.23,

地址 美国加利福尼亚州

审查员 吕晓光

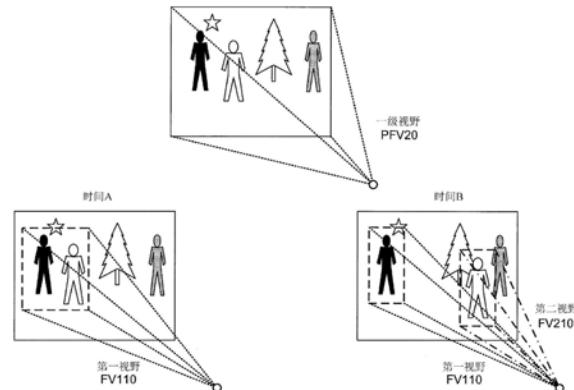
权利要求书4页 说明书47页 附图59页

(54)发明名称

用于视频处理的方法和设备

(57)摘要

本发明提供一种方法，其包含：选择主要视频流内的至少两个对象；从所述主要视频流产生包含所述选定对象中的第一者的第一视频流；以及从所述主要视频流产生包含所述选定对象中的第二者的第二视频流。所述主要视频流具有主要视野，且所述第一和第二视频流具有比所述主要视野更窄的相应第一和第二视野。所述第一视野包含所述主要视野的不在所述第二视野内的一部分，且所述第二视野包含所述主要视野的不在所述第一视野内的一部分。



1. 一种视频处理的方法,所述方法包括:

接收对具有一级视野的一级视频流内的第一对象的第一选择;

响应于接收到所述第一选择,从所述一级视频流产生包含所述第一对象且具有比所述一级视野更窄的第一视野的第一视频流;

接收对所述一级视频流内的第二对象的第二选择;以及

响应于接收到所述第二选择,从所述一级视频流产生包含所述第二对象且具有比所述一级视野更窄的第二视野的第二视频流,

其中所述第一视频流的产生在接收所述第二选择之前且独立于接收所述第二选择,

其中所述第二视频流的产生独立于接收所述第一选择,

其中在屏幕上显示所述第一视频流的同时在所述屏幕上显示所述第二视频流,且其中所述第一视野包含所述一级视野的不在所述第二视野内的一部分且所述第二视野包含所述一级视野的不在所述第一视野内的一部分。

2. 根据权利要求1所述的视频处理的方法,其进一步包括:

在所述一级视频流的第一多个帧或所述第一视频流的第二多个帧中的至少一者内检测所述第一对象的移动;以及

响应于检测到所述移动而改变所述第一视野的方向。

3. 根据权利要求2所述的视频处理的方法,其中改变所述第一视野的所述方向包括改变所述方向以跟随所述第一对象的所述移动。

4. 根据权利要求1所述的视频处理的方法,其中产生所述第一视频流包括从所述一级视频流的至少一个对应帧的一部分产生所述第一视频流的每一帧。

5. 根据权利要求1所述的视频处理的方法,其中对于所述第一视频流的每一帧,所述帧内所述第一对象的第一表示的质心的第一坐标位置与所述帧的中心坐标之间的第一距离小于所述一级视频流的对应帧内所述第一对象的第二表示的质心的第二坐标位置与所述对应帧的中心坐标之间的第二距离。

6. 根据权利要求1所述的视频处理的方法,其中产生所述第一视频流包括使所述第一视频流的第一多个帧中的每一者内的所述第一对象的大小归一化。

7. 根据权利要求1所述的视频处理的方法,其中产生所述第一视频流包括使所述第一视频流的第一多个帧中的每一者内的亮度设定、对比度设定或白平衡设定中的至少一者归一化。

8. 根据权利要求1所述的视频处理的方法,其进一步包括:

在所述屏幕上在第一窗口内显示所述第一视频流;以及

在所述屏幕上在第二窗口内显示所述第二视频流。

9. 根据权利要求8所述的视频处理的方法,其进一步包括:

确定在所述一级视频流的第一帧中所述第一对象到所述第一帧的第一帧边缘比所述第二对象到所述第一帧边缘更接近;

在确定所述第一对象更接近所述帧边缘之后,确定在所述一级视频流的第二帧中所述第二对象到所述第二帧的第二帧边缘比所述第一对象到所述第二帧边缘更接近,所述第二帧边缘对应于所述第一帧边缘;以及

响应于确定所述第二对象更接近所述第二帧边缘,在所述第一窗口内显示所述第二视

频流且在所述第二窗口内显示所述第一视频流。

10. 根据权利要求8所述的视频处理的方法,其进一步包括:

在所述在所述第一窗口内显示所述第一视频流之后,确定在所述一级视频流的帧、所述第一视频流的帧或所述第二视频流的帧中的至少一者中所述第二对象与所述第一对象重叠;以及

响应于确定所述第二对象与所述第一对象重叠,在所述第一窗口内显示所述第二视频流且在所述第二窗口内显示所述第一视频流。

11. 根据权利要求8所述的视频处理的方法,其进一步包括:

从所述一级视频流产生具有比所述一级视野更窄的二级视野的显示视频流;

在所述屏幕的第三窗口内显示所述显示视频流;

检测在所述显示视频流的帧中所述第一对象或所述第二对象中的至少一者靠近所述帧的边缘;以及

响应于所述检测,将所述显示视频流改变为具有三级视野,其中所述三级视野具有比所述二级视野宽的角度。

12. 根据权利要求1所述的视频处理的方法,其中所述第一对象是机器人对象,且其中所述方法进一步包括控制所述机器人对象的移动。

13. 一种用于视频处理的设备,所述设备包括:

用于接收对具有一级视野的一级视频流内的第一对象的第一选择的装置;

用于响应于接收到所述第一选择从所述一级视频流产生包含所述第一对象且具有比所述一级视野更窄的第一视野的第一视频流的装置;

用于接收对所述一级视频流内的第二对象的第二选择的装置;以及

用于响应于接收到所述第二选择从所述一级视频流产生包含所述第二对象且具有比所述一级视野更窄的第二视野的第二视频流的装置,

其中所述第一视频流的产生在接收所述第二选择之前且独立于接收所述第二选择,

其中所述第二视频流的产生独立于接收所述第一选择,

其中在屏幕上显示所述第一视频流的同时在所述屏幕上显示所述第二视频流,且其中在产生所述第一视频流和产生所述第二视频流之后,所述第一视野包含所述一级视野的不在所述第二视野内的一部分且所述第二视野包含所述一级视野的不在所述第一视野内的一部分。

14. 根据权利要求13所述的设备,其中所述用于产生所述第一视频流的装置经配置以使所述第一视频流的第一多个帧中的每一者内的所述第一对象的大小归一化。

15. 根据权利要求13所述的设备,其进一步包括:

用于在所述屏幕上在第一窗口内显示所述第一视频流且在第二窗口内显示所述第二视频流的装置;

用于确定在所述一级视频流的第一帧中所述第一对象到所述第一帧的第一帧边缘比所述第二对象到所述第一帧边缘更接近的装置;

用于在确定所述第一对象更接近所述第一帧边缘之后确定在所述一级视频流的第二帧中所述第二对象到所述第二帧的第二帧边缘比所述第一对象到所述第二帧边缘更接近的装置;以及

用于响应于确定所述第二对象更接近所述第二帧边缘而在所述第一窗口内显示所述第二视频流且在所述第二窗口内显示所述第一视频流的装置。

16. 根据权利要求13所述的设备，其进一步包括：

用于在所述屏幕上在第一窗口内显示所述第一视频流且在第二窗口内显示所述第二视频流的装置；

用于在所述第一窗口内显示所述第一视频流之后确定在所述一级视频流的帧、所述第一视频流的帧或所述第二视频流的帧中的至少一者中所述第二对象与所述第一对象重叠的装置；以及

用于响应于确定所述第二对象与所述第一对象重叠在所述第一窗口内显示所述第二视频流且在所述第二窗口内显示所述第一视频流的装置。

17. 根据权利要求13所述的设备，其进一步包括：

用于在所述屏幕上在第一窗口内显示所述第一视频流且在第二窗口内显示所述第二视频流的装置；

用于从所述一级视频流产生具有比所述一级视野更窄的二级视野的显示视频流的装置；

用于在所述屏幕的第三窗口内显示所述显示视频流的装置；

用于检测在所述显示视频流的帧中所述第一对象和所述第二对象中的至少一者靠近所述帧的边缘的装置；以及

用于响应于所述检测而将所述显示视频流改变为具有三级视野的装置，其中所述三级视野具有比所述二级视野宽的角度。

18. 一种用于视频处理的设备，所述设备包括：

鉴别器，其经配置以接收对具有一级视野的一级视频流内的第一对象的第一选择，且接收对所述一级视频流内的第二对象的第二选择；以及

视频流产生器，其经配置以响应于接收到所述第一选择，从所述一级视频流产生包含所述第一对象且具有比所述一级视野更窄的第一视野的第一视频流，且响应于接收到所述第二选择，从所述一级视频流产生包含所述第二对象且具有比所述一级视野更窄的第二视野的第二视频流，

其中所述第一视频流的产生在接收所述第二选择之前且独立于接收所述第二选择，

其中所述第二视频流的产生独立于接收所述第一选择，

其中在屏幕上显示所述第一视频流的同时在所述屏幕上显示所述第二视频流，且其中在产生所述第一视频流和产生所述第二视频流之后，所述第一视野包含所述一级视野的不在所述第二视野内的一部分且所述第二视野包含所述一级视野的不在所述第一视野内的一部分。

19. 根据权利要求18所述的设备，其进一步包括对象跟踪器，所述对象跟踪器经配置以在所述一级视频流的第一多个帧或所述第一视频流的第二多个帧中的至少一者内检测所述第一对象的移动，其中所述视频流产生器响应于检测到所述第一对象的所述移动而经配置以改变所述第一视野的方向以跟随所述第一对象的所述移动。

20. 根据权利要求18所述的设备，其中所述视频流产生器经配置以使所述第一视频流的第一多个帧中的每一者内的所述第一对象的大小归一化。

21. 根据权利要求18所述的设备,其进一步包括:

显示信号产生器,其经配置以在所述屏幕上在第一窗口内显示所述第一视频流且在第二窗口内显示所述第二视频流;以及

对象跟踪器,其经配置以:

确定在所述一级视频流的第一帧中所述第一对象到所述第一帧的第一帧边缘比所述第二对象到所述第一帧边缘更接近,以及

在确定所述第一对象更接近所述帧边缘之后,确定在所述一级视频流的第二帧中所述第二对象到所述第二帧的第二帧边缘比所述第一对象到所述第二帧边缘更接近,所述第二帧边缘对应于所述第一帧边缘,

其中所述显示信号产生器经配置以在所述第一窗口内显示所述第一视频流之后且响应于确定所述第二对象更接近所述第二帧边缘,在所述第一窗口内显示所述第二视频流且在所述第二窗口内显示所述第一视频流。

22. 根据权利要求18所述的设备,其进一步包括:

显示信号产生器,其经配置以在所述屏幕上在第一窗口内显示所述第一视频流且在第二窗口内显示所述第二视频流;以及

对象跟踪器,其经配置以确定在所述一级视频流的帧、所述第一视频流的帧或所述第二视频流的帧中的至少一者中所述第二对象与所述第一对象重叠,

其中所述显示信号产生器经配置以在所述第一窗口内显示所述第一视频流之后且响应于确定所述第二对象与所述第一对象重叠,在所述第一窗口内显示所述第二视频流且在所述第二窗口内显示所述第一视频流。

23. 根据权利要求18所述的设备,其中所述视频流产生器经配置以从所述一级视频流产生具有比所述一级视野更窄的二级视野的显示视频流,所述设备进一步包括:

显示信号产生器,其经配置以在所述屏幕上在第一窗口内显示所述第一视频流,在第二窗口内显示所述第二视频流,且在第三窗口内显示所述显示视频流;以及对象跟踪器,其经配置以检测在所述显示视频流的帧中至少所述第一对象或所述第二对象靠近所述帧的边缘,且

其中所述视频流产生器经配置以响应于所述检测而将所述显示视频流改变为具有三级视野,其中所述三级视野具有比所述二级视野宽的角度。

用于视频处理的方法和设备

[0001] 相关申请案的交叉参考

[0002] 本申请案主张共同拥有的2013年12月20日申请的美国临时专利申请案第61/919,627号及2014年12月18日申请的美国非临时专利申请案第14/575,945号的优先权，所述专利申请案的内容明确地以全文引用的方式结合在此。

技术领域

[0003] 本发明大体上涉及对象跟踪。

背景技术

[0004] 技术的进步已经产生了更小且更强大的计算装置。举例来说，当前存在多种便携式个人计算装置，包含无线计算装置，例如便携式无线电话、个人数字助理(PDA)和寻呼装置，其体积小，重量轻，且易于由用户携带。更确切地说，例如蜂窝式电话和因特网协议(IP)电话等便携式无线电话可经由无线网络传送语音和数据包。另外，许多此类无线电话包含并入其中的其它类型的装置。举例来说，无线电话还可包含数字静态相机、数码摄像机、数字记录器和音频文件播放器。而且，所述无线电话可处理可执行指令，其包含可用以接入因特网的软件应用，例如，网页浏览器应用。由此，这些无线电话可以包含很大的计算能力。

发明内容

[0005] 在一特定实施例中，一种方法包含在移动装置处接收对场景中的第一对象和场景中的第二对象的选择。所述场景对应于在所述移动装置的屏幕上捕获且显示的视频流。所述方法还包含将所述屏幕分割为至少第一窗口和第二窗口。所述方法进一步包含在第一窗口中显示所述场景的包含第一对象的第一部分。所述场景的第一部分是基于所述第一对象的位置而确定。所述方法还包含在第二窗口中显示所述场景的包含第二对象的第二部分。所述场景的第二部分是基于所述第二对象的位置而确定。

[0006] 在另一特定实施例中，一种移动装置包含存储器和耦合到所述存储器的处理器。所述处理器经配置以接收对场景中的第一对象和场景中的第二对象的选择。所述场景对应于在所述移动装置的屏幕上捕获且显示的视频流。所述处理器进一步经配置以将屏幕分割为至少第一窗口和第二窗口。所述处理器还经配置以在第一窗口中显示所述场景的包含第一对象的第一部分。所述场景的第一部分是基于所述第一对象的位置而确定。所述处理器进一步经配置以在第二窗口中显示所述场景的包含第二对象的第二部分。所述场景的第二部分是基于所述第二对象的位置而确定。

[0007] 在另一特定实施例中，一种方法包含在移动装置处接收对场景中的第一对象和场景中的第二对象的选择。所述场景对应于在所述移动装置的屏幕上捕获且显示的视频流。所述方法还包含将所述屏幕分割为第一窗口和第二窗口。所述方法进一步包含在第一窗口中初始地显示所述场景的包含第一对象的第一部分。所述场景的第一部分是基于所述第一对象的位置而确定。所述方法还包含在第二窗口中初始地显示所述场景的包含第二对象的

第二部分。所述场景的第二部分是基于所述第二对象的位置而确定。所述方法还包含基于当第一对象的一部分与第二对象的一部分重叠阈值量时控制在哪一窗口中显示所述第一对象和第二对象。

[0008] 在另一特定实施例中，一种移动装置包含存储器和耦合到所述存储器的处理器。所述处理器经配置以接收对场景中的第一对象和场景中的第二对象的选择。所述场景对应于在所述移动装置的屏幕上捕获且显示的视频流。所述处理器还经配置以将屏幕分割为第一窗口和第二窗口。所述处理器进一步经配置以在第一窗口中初始地显示所述场景的包含第一对象的第一部分。所述场景的第一部分是基于所述第一对象的位置而确定。所述处理器还经配置以在第二窗口中初始地显示所述场景的包含第二对象的第二部分。所述场景的第二部分是基于所述第二对象的位置而确定。所述处理器还经配置以基于当第一对象的一部分与第二对象的一部分重叠阈值量时控制在哪一窗口中显示所述第一对象和第二对象。

[0009] 在另一特定实施例中，一种方法包含在移动装置处接收对场景中的第一对象和场景中的第二对象的选择。所述场景对应于在所述移动装置的屏幕上捕获且显示的视频流。所述方法还包含将所述屏幕分割为第一窗口和第二窗口。如果在第二对象的选择之前选择第一对象，那么所述方法包含在第一窗口中显示所述场景的包含第一对象的第一部分且在第二窗口中显示所述场景的包含第二对象的第二部分。所述场景的第一部分是基于所述第一对象的位置而确定。所述场景的第二部分是基于所述第二对象的位置而确定。如果在第一对象的选择之前选择第二对象，那么所述方法包含在第二窗口中显示所述场景的第一部分且在第一窗口中显示所述场景的第二部分。

[0010] 在另一特定实施例中，一种移动装置包含存储器和耦合到所述存储器的处理器。所述处理器经配置以接收对场景中的第一对象和场景中的第二对象的选择。所述场景对应于在所述移动装置的屏幕上捕获且显示的视频流。所述处理器还经配置以将屏幕分割为第一窗口和第二窗口。如果在第二对象的选择之前选择第一对象，那么所述处理器经配置以在第一窗口中显示所述场景的包含第一对象的第一部分且在第二窗口中显示所述场景的包含第二对象的第二部分。所述场景的第一部分是基于所述第一对象的位置而确定。所述场景的第二部分是基于所述第二对象的位置而确定。如果在第一对象的选择之前选择第二对象，那么所述处理器经配置以在第二窗口中显示所述场景的第一部分且在第一窗口中显示所述场景的第二部分。

[0011] 在另一特定实施例中，一种方法包含在移动装置处接收对场景中的多个对象的选择。所述场景对应于在所述移动装置的屏幕上捕获且显示的视频流。所述方法还包含跟踪每一对象之间的几何形状。所述方法进一步包含在屏幕上显示所述几何形状的指示。

[0012] 在另一特定实施例中，一种移动装置包含存储器和耦合到所述存储器的处理器。所述处理器经配置以接收对场景中的多个对象的选择。所述场景对应于在所述移动装置的屏幕上捕获且显示的视频流。所述处理器还经配置以跟踪每一对象之间的几何形状。所述处理器还经配置以在屏幕上显示所述几何形状的指示。

[0013] 在另一特定实施例中，一种方法包含在移动装置处接收对场景中的多个对象的选择。所述场景对应于在所述移动装置的屏幕上捕获且显示的视频流。所述方法还包含跟踪每一对象之间的几何形状。所述方法进一步包含基于与所述几何形状相关联的至少一个参数将视频流的帧集群。

[0014] 在另一特定实施例中，一种移动装置包含存储器和耦合到所述存储器的处理器。所述处理器经配置以接收对场景中的多个对象的选择。所述场景对应于在所述移动装置的屏幕上捕获且显示的视频流。所述处理器还经配置以跟踪每一对象之间的几何形状。所述处理器进一步经配置以基于与所述几何形状相关联的至少一个参数将视频流的帧集群。

[0015] 在另一特定实施例中，一种方法包含在移动装置处接收对场景中的第一对象和场景中的第二对象的选择。所述场景对应于在所述移动装置的屏幕上捕获且显示的视频流。如果相对于屏幕的大小的第一对象的大小大于相对于屏幕的大小的第二对象的大小，那么所述方法包含在第一窗口中显示所述场景的包含第一对象的第一部分且在第二窗口中显示所述场景的包含第二对象的第二部分。所述场景的第一部分是基于所述第一对象的位置而确定。所述场景的第二部分是基于所述第二对象的位置而确定。所述第一窗口大于所述第二窗口。如果相对于屏幕的大小的第一对象的大小不大于相对于屏幕的大小的第二对象的大小，那么所述方法包含在第二窗口中显示所述场景的第一部分且在第一窗口中显示所述场景的第二部分。

[0016] 在另一特定实施例中，一种移动装置包含存储器和耦合到所述存储器的处理器。所述处理器经配置以接收对场景中的第一对象和场景中的第二对象的选择。所述场景对应于在所述移动装置的屏幕上捕获且显示的视频流。如果相对于屏幕的大小的第一对象的大小大于相对于屏幕的大小的第二对象的大小，那么所述处理器经配置以在第一窗口中显示所述场景的包含第一对象的第一部分且在第二窗口中显示所述场景的包含第二对象的第二部分。所述场景的第一部分是基于所述第一对象的位置而确定。所述场景的第二部分是基于所述第二对象的位置而确定。所述第一窗口大于所述第二窗口。如果相对于屏幕的大小的第一对象的大小不大于相对于屏幕的大小的第二对象的大小，那么所述处理器经配置以在第二窗口中显示所述场景的第一部分且在第一窗口中显示所述场景的第二部分。

[0017] 在另一特定实施例中，一种方法包含在机器人处跟踪多个对象之间的几何形状。所述方法还包含基于所述几何形状执行处理功能。

[0018] 在另一特定实施例中，一种方法包含在移动装置处接收对场景中的第一对象、场景中的第二对象和场景中的第三对象的选择。所述场景对应于在所述移动装置的屏幕上捕获且显示的视频流。所述方法还包含将屏幕分割为第一窗口、第二窗口、第三窗口和第四窗口。所述方法进一步包含在第一窗口中显示所述场景的包含第一对象的第一部分。所述场景的第一部分是基于所述第一对象的位置而确定。所述方法还包含在第二窗口中显示所述场景的包含第二对象的第二部分。所述场景的第二部分是基于所述第二对象的位置而确定。所述方法进一步包含在第三窗口中显示所述场景的包含第三对象的第三部分。所述场景的第三部分是基于所述第三对象的位置而确定。所述方法还包含在第四窗口中显示每一对象之间的几何形状的指示。

[0019] 在另一特定实施例中，一种移动装置包含存储器和耦合到所述存储器的处理器。所述处理器经配置以接收对场景中的第一对象、场景中的第二对象和场景中的第三对象的选择。所述场景对应于在所述移动装置的屏幕上捕获且显示的视频流。所述处理器还经配置以将屏幕分割为第一窗口、第二窗口、第三窗口和第四窗口。所述处理器进一步经配置以在第一窗口中显示所述场景的包含第一对象的第一部分。所述场景的第一部分是基于所述第一对象的位置而确定。所述处理器还经配置以在第二窗口中显示所述场景的包含第二对象的第二部分。

象的第二部分。所述场景的第二部分是基于所述第二对象的位置而确定。所述处理器进一步经配置以在第三窗口中显示所述场景的包含第三对象的第三部分。所述场景的第三部分是基于所述第三对象的位置而确定。所述处理器还经配置以在第四窗口中显示每一对象之间的几何形状的指示。

[0020] 在另一特定实施例中，一种方法包含在移动装置处接收视频流中的选定对象的特定几何形状的指示。所述方法还包含基于所述特定几何形状检索视频流的帧。所述方法进一步包含在移动装置的屏幕上显示检索的帧。

[0021] 在另一特定实施例中，一种移动装置包含存储器和耦合到所述存储器的处理器。所述处理器经配置以接收视频流中的选定对象的特定几何形状的指示。所述处理器还经配置以基于所述特定几何形状检索视频流的帧。所述处理器还经配置以在移动装置的屏幕上显示检索的帧。

[0022] 根据一般配置的视频处理的方法包含：选择具有一级视野的一级视频流内的至少两个对象；响应于所述选择，从所述一级视频流产生包含所述选定对象中的第一者且具有比所述一级视野更窄的第一视野的第一视频流；以及在所述选择之后，从所述一级视频流产生包含所述选定对象中的第二者且具有比所述一级视野更窄的第二视野的第二视频流。在此方法中，在所述产生第一视频流和所述产生第二视频流开始之后，所述第一视野包含一级视野的不在所述第二视野内的一部分，且所述第二视野包含一级视野的不在所述第一视野内的一部分。还揭示具有指令的计算机可读存储媒体（例如，非暂时性媒体），所述指令致使执行所述指令的机器执行此方法。

[0023] 根据一般配置的用于视频处理的设备包含：用于选择具有一级视野的一级视频流内的至少两个对象的装置；用于响应于所述选择从所述一级视频流产生包含所述选定对象中的第一者且具有比所述一级视野更窄的第一视野的第一视频流的装置；以及用于在所述选择之后从所述一级视频流产生包含所述选定对象中的第二者且具有比所述一级视野更窄的第二视野的第二视频流的装置。在此设备中，在所述产生第一视频流和所述产生第二视频流开始之后，所述第一视野包含一级视野的不在所述第二视野内的一部分，且所述第二视野包含一级视野的不在所述第一视野内的一部分。

[0024] 根据另一一般配置的用于视频处理的设备包含：鉴别器，其经配置以选择具有一级视野的一级视频流内的至少两个对象；以及视频流产生器，其经配置以响应于所述选择，从所述一级视频流产生包含所述选定对象中的第一者且具有比所述一级视野更窄的第一视野的第一视频流，以及在所述选择之后，从所述一级视频流产生包含所述选定对象中的第二者且具有比所述一级视野更窄的第二视野的第二视频流。在此设备中，在所述产生第一视频流和所述产生第二视频流开始之后，所述第一视野包含一级视野的不在所述第二视野内的一部分，且所述第二视野包含一级视野的不在所述第一视野内的一部分。

[0025] 由所揭示的实施例中的至少一者提供的一个特定优点是移动装置上的图像跟踪的改善的视觉显示。本发明的其它方面、优点和特征将在审阅全部申请案之后变得显而易见，所述全部申请案包含以下章节：附图说明、具体实施方式和权利要求书。

附图说明

[0026] 图1是描绘电子装置的框图。

- [0027] 图2A是描绘对象及检测模块的特定说明性实施例的框图。
- [0028] 图2B是描绘图2的对象及检测模块内的处理器实施组件的特定说明性实施例的框图。
- [0029] 图3是用于执行基于运动的跟踪及对象检测的方法的特定说明性实施例的流程图。
- [0030] 图4是用于执行基于运动的跟踪的方法的特定说明性实施例的流程图。
- [0031] 图5是用于基于前向-后向误差而估计基于运动的跟踪中的跟踪误差的方法的特定说明性实施例的流程图。
- [0032] 图6是用于执行对象检测的方法的特定说明性实施例的流程图。
- [0033] 图7是具有不同窗口大小的图像窗口的特定说明性实施例。
- [0034] 图8是描绘对象跟踪与检测模块的特定说明性实施例的框图。
- [0035] 图9是描绘平滑模块的特定说明性实施例的框图。
- [0036] 图10是用于平滑运动跟踪中的抖动的方法的特定说明性实施例的流程图。
- [0037] 图11是描绘使用对象跟踪的视频处理的特定实施例。
- [0038] 图12是使用对象跟踪的视频处理方法的特定说明性实施例的流程图。
- [0039] 图13是描绘使用对象跟踪的视频处理的另一特定实施例。
- [0040] 图14是使用对象跟踪的视频处理方法的特定说明性实施例的另一流程图。
- [0041] 图15是描绘使用对象跟踪的视频处理的另一特定实施例。
- [0042] 图16是使用对象跟踪的视频处理方法的特定说明性实施例的另一流程图。
- [0043] 图17是描绘使用对象跟踪的视频处理的另一特定实施例。
- [0044] 图18描绘用于基于集群检索帧的移动装置的屏幕的特定实施例。
- [0045] 图19是使用对象跟踪的视频处理方法的特定说明性实施例的另一流程图。
- [0046] 图20描绘使用对象跟踪的视频处理方法的特定说明性实施例的流程图。
- [0047] 图21是描绘使用对象跟踪的视频处理的另一特定实施例。
- [0048] 图22是使用对象跟踪的视频处理方法的特定说明性实施例的另一流程图。
- [0049] 图23是使用对象跟踪的视频处理方法的特定说明性实施例的另一流程图。
- [0050] 图24是描绘使用对象跟踪的视频处理的另一特定实施例。
- [0051] 图25是使用对象跟踪的视频处理方法的特定说明性实施例的另一流程图。
- [0052] 图26是包含可操作以执行视频处理技术的组件的无线装置的框图。
- [0053] 图27A到27E展示视野FV10以及对其的各种操作的结果。
- [0054] 图28A展示根据一般配置的视频处理的方法M100的流程图。
- [0055] 图28B展示一级视野PFV10的实例，且图28B和28C展示第一视野FV100和第二视野FV200的实例。
- [0056] 图29展示选择三个对象的用户动作的序列的实例。
- [0057] 图30展示选择三个对象的用户动作的序列的另一实例。
- [0058] 图31展示帧内的对象及对象的增强显示的实例。
- [0059] 图32A展示九个提取窗口的实例集合，且图32B展示一级视频流的帧的区与所述九个窗口中的每一者之间的对应。
- [0060] 图33A展示方法M100的实施方案M110的流程图。

- [0061] 图33B展示方法M100和M110的实施方案M120的流程图。
- [0062] 图33C展示方法M100的实施方案M200的流程图。
- [0063] 图33D展示方法M100的实施方案M300的流程图。
- [0064] 图34A到C展示方法M300的应用。
- [0065] 图35展示方法M100的应用的实例。
- [0066] 图36A到F展示显示窗口的默认布置的实例。
- [0067] 图37A展示方法M200和M300的实施方案M400的流程图。
- [0068] 图37B展示根据一般配置的设备A100的框图。
- [0069] 图37C展示设备A100的实施方案A110的框图。
- [0070] 图38A展示设备A100的实施方案A200的框图。
- [0071] 图38B展示设备A100的实施方案A300的框图。
- [0072] 图39A展示设备A200和A300的实施方案A400的框图。
- [0073] 图39B和39C展示设备A100的应用的框图。
- [0074] 图39D展示设备A300的应用的框图。
- [0075] 图39E展示根据一般配置的设备MF100的框图。
- [0076] 图40A展示设备MF100的实施方案MF110的框图。
- [0077] 图40B展示设备MF100的实施方案MF200的框图。
- [0078] 图40C展示设备MF100的实施方案MF300的框图。
- [0079] 图40D展示设备MF200和MF300的实施方案MF400的框图。
- [0080] 图41A和41B展示包含选定对象的一级视频流的视野的一部分的显示的两个实例。
- [0081] 图42展示显示视频流的帧的实例。
- [0082] 图43A和43C展示显示一级视频流的帧的窗口,且图43B展示指示选定对象的复合几何形状的图形的实例。

具体实施方式

- [0083] 无线电话或其它移动装置可使用相机捕获视频流及/或从另一装置及/或经由网络接收视频流。可需要用于跟踪视频流内的对象的新及/或提高的特征。
- [0084] 除非通过其上下文明确限制,否则在本文中使用术语“信号”来指示其一般含义中的任一者,包含如在电线、总线或其它传输媒体上表达的存储器位置的状态(或存储器位置的集合)。除非通过其上下文明确地限制,否则在本文中使用术语“产生”来指示其一般含义中的任一者,例如计算或以其它方式产生。除非通过其上下文明确限制,否则在本文中使用术语“计算”来指示其一般含义中的任一者,例如计算、评估、估计,和/或从多个值进行选择。除非明确地由其上下文限制,否则术语“获得”用于指示其普通含义中的任一者,例如计算、导出、接收(例如,从外部装置)及/或检索(例如,从存储元件阵列)。除非明确地由其上下文限制,否则术语“选择”用于指示其普通含义中的任一者,例如识别、指示、施加及/或使用一组两个或更多个中的至少一者及少于全部。在本发明描述及权利要求书中使用术语“包括”时,并不排除其它元件或操作。术语“基于”(如在“A是基于B”中)用于指示其一般含义中的任一者,包含以下情况(i)“从...导出”(例如,“B是A的前驱体”), (ii)“至少基于”(例如,“A至少基于B”)并且在特定上下文中在适当时,(iii)“等于”(例如,“A等于B”)。类似

地,使用术语“响应于”来指示其一般含义中的任一者,包含“至少响应于”。

[0085] 除非另有指示,否则对具有特定特征的设备的操作的任何揭示内容还明确地希望揭示具有类似特征的方法(且反之亦然),且对根据特定配置的设备的操作的任何揭示内容还明确地希望揭示根据类似配置的方法(且反之亦然)。术语“配置”可参考由其特定上下文指示的方法、设备和/或系统来使用。除非特定上下文另有指示,否则术语“方法”、“过程”、“程序”及“技术”通用地且可互换地使用。除非特定上下文另有指示,否则术语“设备”与“装置”也通用地且可互换地使用。术语“元件”和“模块”通常用以指示较大配置的一部分。除非通过其上下文明确地限制,否则术语“系统”在此用以指示其一般含义中的任一者,包含“相互作用以用于共同目的的一群组元件”。

[0086] 除非另有指示,否则使用术语“系列”来指示两个或更多个项目的序列。除非起初通过定冠词引入,否则用于修饰权利要求元素的序数术语(例如,“第一”、“第二”、“第三”等)本身不指示所述权利要求元素相对于另一权利要求元素的任何优先级或次序,而是仅区别所述权利要求元素与具有同一名称(如果没有序数术语)的另一权利要求元素。除非通过其上下文明确地限制,否则术语“多个”及“组”中的每一者在本文中用以指示大于1的整数量。

[0087] 参考图1,展示说明电子装置102的框图。电子装置102也可被称作无线通信装置、移动装置、移动台、订户台、客户端、客户端台、用户设备(UE)、远端台、接入终端、移动终端、终端、用户终端、订户单元等。电子装置的实例包含膝上型或桌上型计算机、蜂窝式电话、智能电话、无线调制解调器、电子阅读器、平板装置、游戏系统等。这些装置中的一些可根据一或多个行业标准来操作。

[0088] 例如智能电话或平板计算机的电子装置102可包含照相机。照相机可包含图像传感器114和将位于光学系统118的视野内的对象的图像聚焦于图像传感器114上的光学系统118(例如,透镜)。电子装置102还可包含照相机软件应用程序和显示屏。当执行照相机应用程序时,可由图像传感器114记录位于光学系统118的视野内的对象的图像。由图像传感器114所记录的图像可显示于显示屏上。可以相对高帧率快速连续显示这些图像,使得在任何给定时刻处,位于光学系统118的视野内的对象显示于显示屏上。尽管在所捕获的帧(例如,视频帧)方面描述实施例,但本文中所论述技术可用于任何数字图像上。因此,术语“帧”和“数字图像”可在本文中互换使用。

[0089] 照相机应用程序的用户接口120可准许跟踪显示于显示屏上的一或多个对象。电子装置102的用户可被准许选择待跟踪的对象。另外,所选择对象可用作稍后检测对象的参考。

[0090] 在一个配置中,显示屏为从物理触摸(例如,通过手指、触控笔或其它工具)接收输入的触摸屏116。触摸屏116可接收界定待跟踪目标对象的触摸输入。举例来说,如果电子装置102正捕获包含所关注动物的自然场景,那么(必要时)用户可围绕动物绘制指示希望跟踪或检测动物的界限框。可以任何合适方式选择目标对象。举例来说,面部辨识、行人辨识等可用于选择待跟踪、检测或所述两者的目标对象。在一个配置中,可跟踪多个对象。用户接口120可允许用户与对象跟踪与检测模块104互动(例如)以选择(即,界定)一或多个目标对象。触摸屏116可包含取景器131。取景器131可指触摸屏116显示视频流或实况馈送的部分。例如,取景器131可显示由电子装置102上的相机获得的视图。

[0091] 电子装置102可包含用于跟踪所选择对象和/或在视频帧中检测对象的对象跟踪与检测模块104。对象跟踪与检测模块104可包含用于跟踪一或多个对象的运动跟踪器106。运动跟踪器106可基于运动以用于逐帧跟踪图像(例如,视频帧)上的点的运动以估计目标对象在先前视频帧与当前视频帧之间的位置和/或位置改变。

[0092] 对象跟踪与检测模块还可包含用于检测视频帧中的对象的对象检测器108。对象检测器108可使用对象模型而非基于运动模型以通过比较当前视频帧的全部或一部分与所选择对象或(例如,视频帧序列中)所捕获的先前视频帧112的一部分来检测对象。对象检测器108可用于检测视频帧内的多个对象。

[0093] 对象跟踪与检测模块104还可包含存储器缓冲器110。存储器缓冲器110可存储一或多个所捕获帧和与所捕获视频帧相关联的数据。在一个实例中,存储器缓冲器110可存储先前所捕获视频帧112。对象跟踪与检测模块104可使用从存储器缓冲器110所提供的关于所捕获的先前视频帧112的数据来执行基于运动的跟踪和/或对象检测。可经由来自存储器缓冲器110的反馈将数据提供到运动跟踪器106或对象检测器108以便修整基于运动的跟踪和对象检测以更准确地跟踪和/或检测目标对象。举例来说,存储器缓冲器110可将位置和窗口大小数据提供到运动跟踪器106和对象检测器108,以为运动跟踪器106和对象检测器108提供当跟踪或检测对象时可用于更准确地精确指出对象的位置和大小的一或多个参数。

[0094] 如上文陈述,电子装置102可执行基于运动的跟踪。可使用多种方法执行基于运动的跟踪。在一个实例中,通过中值流方法执行跟踪,其中运动跟踪器106接受一对图像 I_t 、 I_{t+1} (例如,视频帧)和限界框 β_t 并输出限界框 β_{t+1} 。可在限界框 β_t 内的矩形网格上初始化点集合且可以跟踪所述点以在 I_t 与 I_{t+1} 之间产生稀疏运动流。可估计点预测质量并为每个点指派误差。可滤除最坏预测的一部分(例如,50%)同时将剩余预测用于估计整个限界框的移位。运动跟踪器106可对由电子装置102所捕获的每一视频帧执行基于运动的跟踪。以类似方法,可通过计算一或多个梯度(例如,x和y梯度)并使用一对帧之间的差来计算时间梯度且使用多个梯度值以准确跟踪当前视频帧内的目标对象来执行基于运动的跟踪。下文提供关于基于运动的跟踪的进一步细节。

[0095] 当执行基于运动的跟踪时,运动跟踪器106可基于运动的跟踪方法的所计算或所估计准确性来确定跟踪置信度值。在一些配置中,跟踪置信度值可为对应于目标对象落于当前视频帧内或视频帧的所界定窗口内的可能性或机率的介于0与1之间的实数。可将跟踪置信度值相比于跟踪阈值。如果跟踪置信度值大于跟踪阈值,那么发现目标对象处于当前视频帧内的可能性可较高。替代性地,如果跟踪置信度值小于或等于跟踪阈值,那么可能性可较低或不确定是否发现目标对象在当前视频帧内。可使用用于确定跟踪置信度值的各种方法。在一个配置中,通过计算当前视频帧中的所跟踪窗口(例如,跟踪贴片窗口)与来自先前所捕获视频帧的先前存储图像贴片之间的归一化互相关(NCC)来确定跟踪置信度值。下文提供关于确定跟踪置信度值的进一步细节。

[0096] 电子装置102也可执行对象检测。可使用多种方法执行对象检测。在一个配置中,使用滑动窗口方法执行对象检测,其中检视视频帧内的窗口的多个子集的内容以确定是否发现目标对象在当前视频帧中或当前视频帧的特定窗口或窗口子集内。可在视频帧中搜索所有可能窗口位置和大小的全部或子集。举例来说,每一窗口可对应于数据像素且对象检

测器108可使用数据像素执行一或多个计算以确定目标对象在特定窗口或子窗口内的置信水平(例如,二进制指标)。基于与一或多个窗口相关联的置信水平,可获得当前视频帧的检测器置信度值。另外,额外技术可用于增加对象检测的准确性或效率。下文解释这些技术中的一些。

[0097] 在一些配置中,运动跟踪器106和对象检测器108可依序而非并行操作。举例来说,电子装置102可对所选择对象(例如,目标对象)执行基于运动的跟踪并基于所跟踪参数依序执行所选择对象的对象检测。在一个配置中,电子装置102可对当前视频帧执行基于运动的跟踪。电子装置102可接着基于所跟踪参数对当前帧执行对象检测。在一个配置中,所跟踪参数可是基于置信度值与阈值之间的比较。举例来说,如果跟踪置信度值低于跟踪阈值,那么电子装置102可执行对象检测。替代性地,如果跟踪置信度值高于跟踪阈值,那么电子装置102可跳过当前视频帧的对象检测并基于当前视频帧的运动跟踪结果继续对下一视频帧执行基于运动的跟踪。换句话说,可仅当基于运动的跟踪并非极好(例如,跟踪置信度值低于跟踪阈值)时执行对象检测。当考虑是否执行和/或如何执行对象检测时可使用其它所跟踪参数。所跟踪参数的实例可包含目标对象区域、窗口位置、窗口大小、尺度层级、目标大小、跟踪和/或检测置信度值或可用于促进有效跟踪和/或检测目标对象的其它参数。

[0098] 依序执行基于运动的跟踪和基于所跟踪参数的对象检测可使得电子装置102能够在无需执行广泛计算的情况下跟踪和/或检测视频帧内的目标对象。具体来说,因为相比对象检测,基于运动的跟踪可较不计算密集型,所以电子装置102可跳过执行对象检测,其中基于运动的跟踪可用于准确跟踪当前视频帧内的目标对象。举例来说,如果电子装置102确定跟踪置信度值超出特定目标阈值,那么电子装置102可确定无需对当前视频帧进行对象检测以准确确定当前视频帧内的目标对象的位置或现状。另外,因为对象检测在许多状况下可是有益的,所以电子装置102可确定对象检测可用于更准确检测目标对象的状况或在基于运动的跟踪不充分基于与跟踪阈值的比较的状况下执行对象检测的状况。

[0099] 在一些配置中,基于运动的跟踪的结果和/或由存储器缓冲器110所提供的额外信息可用于窄化或修整执行对象检测的过程,而非跳过对当前视频帧的对象检测。举例来说,在无法使用基于运动的跟踪方法准确跟踪目标对象的情况下,电子装置102仍可估计或获得关于位置、窗口尺度或与目标对象相关联的其它所跟踪参数的信息,所述信息可在对象检测期间用以使用比并无经由基于运动的跟踪所提供参数的情况少的计算能力来更准确地检测对象。因此,即使在基于运动的跟踪并不提供超出跟踪阈值的跟踪置信度值的状况下,当随后执行对象检测时可使用基于运动的跟踪的结果。

[0100] 电子装置102上的取景器131可包含第一跟踪区133及第二跟踪区135。第一跟踪区133及第二跟踪区135两者可由用户使用触摸屏116指定。例如,用户可在触摸屏116上将聚焦环拖动到第一跟踪区133及第二跟踪区135的所要位置。虽然不是必需的,跟踪区中的一者可静止。例如,第一跟踪区133可跟踪对象(例如,行走的人)且第二跟踪区135可覆盖静止的树。在一个配置中,第二跟踪区135可覆盖电子装置102上的整个触摸屏116。

[0101] 电子装置102可包含视频处理模块137。图片处理模块137可包含重叠143。重叠143可反映第一跟踪区133与第二跟踪区135之间的重叠的量。例如,如果第一跟踪区133及第二跟踪区135彼此完全不重叠,那么重叠143可为0%。同样,如果第一跟踪区133与第二跟踪区135完全重叠(或如果第二跟踪区135与第一跟踪区133完全重叠,其取决于哪个跟踪区更

大),那么重叠143可为100%。视频处理模块137可包含阈值145。可将重叠143与阈值145进行比较以确定是否应该执行视频处理,如相对于图13所描述。

[0102] 视频处理模块137还可包含屏幕分区147功能。例如,屏幕分区147可将取景器132分割为多个窗口以显示与第一跟踪区133及第二跟踪区135相关联的个别视频流,如相对于图11、13、15及21所描述。视频处理模块137还可包含几何形状跟踪149功能。例如,几何形状跟踪149功能可跟踪第一跟踪区133与第二跟踪区135之间的几何形状,如相对于图17及18所描述。所述几何形状可显示在取景器131上。视频处理模块137还可包含群集阵型151功能。群集阵型151功能可基于与几何形状相关联的至少一个参数(例如,质量中心)而产生群集。例如,每一群集可包含具有第一跟踪区133与第二跟踪区135之间的大体上类似的几何形状的视频帧,如相对于图17及18所描述。

[0103] 参看图2A,展示说明对象跟踪与检测模块204的框图。对象跟踪与检测模块204可实施于电子或无线装置内。对象跟踪与检测模块204可包含具有光流模块226和跟踪置信度值228的运动跟踪器206。对象跟踪与检测模块204还可包含具有扫描仪定位器230、扫描仪定标器236、分类器238和检测置信度值240的对象检测器208。存储器缓冲器210可存储可提供到运动跟踪器206和对象检测器208的与所捕获的先前视频帧212相关联的数据。对象跟踪与检测模块204、运动跟踪器206、对象检测器208和存储器缓冲器210可为上文结合图1所描述的对象跟踪与检测模块104、运动跟踪器106、对象检测器108和存储器缓冲器110的配置。

[0104] 运动跟踪器206可用于对当前视频帧(N)224执行基于运动的跟踪。举例来说,可(例如,由电子装置102)接收先前视频帧(N-1)222和当前视频帧(N)224。先前视频帧(N-1)222可在视频帧序列中在当前视频帧(N)224紧接着之前。可由对象跟踪与检测模块204获得并处理额外视频帧。可将先前视频帧(N-1)222提供到运动跟踪器206。另外,存储器缓冲器210可存储与先前视频帧(N-1)222(在本文中被称作所捕获的先前视频帧212)相关联的数据。在一些配置中,存储器缓冲器210可从电子装置102(例如,从照相机)直接获得关于先前视频帧(N-1)222的信息。存储器缓冲器210也可从融合模块260获得可指定对象在先前视频帧(N-1)222中的跟踪和/或检测位置的关于先前视频帧(N-1)222的跟踪结果。关于先前视频帧(N-1)222或其它先前所捕获视频帧的此信息可存储于存储器缓冲器210中。

[0105] 运动跟踪器206可随后接收视频帧序列中的当前视频帧(N)224。运动跟踪器206可比较当前视频帧(N)224与先前视频帧(N-1)222(例如,使用从存储器缓冲器210所提供的信息)。运动跟踪器206可使用光流模块226跟踪对象在当前视频帧(N)224上的运动。光流模块226可包含用于对当前视频帧(N)224上的对象执行基于运动的跟踪的硬件和/或软件。通过比较先前视频帧(N-1)222与当前视频帧(N)224,运动跟踪器206可确定与目标对象处于当前视频帧(N)224中的可能性相关联的跟踪置信度值228。在一个实例中,跟踪置信度值228为基于目标对象在当前视频帧(N)224内或当前视频帧(N)224内的窗口内的确定性百分比的实数(例如,介于0与1之间)。

[0106] 对象检测器208可用于检测当前视频帧(N)224上的对象。举例来说,对象检测器208可接收视频帧序列中的当前视频帧(N)224。对象检测器208可基于所跟踪参数对当前视频帧(N)224执行对象检测。所跟踪参数可包含对应于正准确跟踪目标对象的可能性的跟踪置信度值228。更确切地说,所跟踪参数可包含跟踪置信度值228与跟踪阈值250的比较。所

跟踪参数还可包含从存储器缓冲器210所提供的信息。当检测对象时可使用的所跟踪参数的一些实例包含区域、窗口位置、窗口大小或当执行对象检测时可由对象检测器208用作参数的其它信息。

[0107] 对象检测器208可包含扫描仪定位器230。扫描仪定位器230可包含窗口位置选择器232和随机化器234。窗口位置选择器232可在视频帧内选择多个窗口。举例来说，视频帧可包含多个窗口，每一窗口具有相关联位置和大小。在一个配置中，每一视频帧划分成多个（例如，大约10,000个）重叠窗口，每一窗口包含视频帧中的总像素的一部分。替代性地，可存在任何合适数目个窗口且其可并不重叠。扫描仪定位器230内的窗口位置选择器232可选择在其中尝试识别目标对象的窗口位置。随机化器234可随机选择具有变化大小和位置的窗口以用于检测对象。在一些配置中，随机化器234在视频帧内随机选择窗口。替代性地，随机化器234可基于一或多个因素较精确地选择窗口。举例来说，随机化器234可基于区域、大小或对象最可能定位的一般位置来限制窗口选择。可经由存储器缓冲器210获得此信息或可经由虽然并不足够准确以完全依赖但可提供当执行对象检测时有用的信息的基于运动的跟踪获得此信息。因此，虽然随机化器234可随机选择多个窗口以进行搜索，但可基于提供到对象检测器208的信息窄化窗口选择，且因此不完全随机。

[0108] 对象检测器208还可包含可用于绘制或选择具有一定大小的窗口的扫描仪定标器236。当检测对象或比较窗口选择与原始图像以检测图像是否在特定窗口内时，可由扫描仪定位器230使用窗口大小以窄化窗口大小。当界定对象时，扫描仪定标器236可最初选择具有某些大小或尺度层级的一或多个窗口，或替代性地基于从存储器缓冲器210所提供的信息绘制具有某些大小或尺度层级的一或多个窗口。

[0109] 分类器238可用于确定是否发现目标对象中的一些或全部在特定窗口中。在一些配置中，分类器238可产生每一窗口的二进制值以指示是否检测到目标对象在特定窗口或子窗口内。可针对由对象检测器208所搜索的每一窗口执行此分类（例如，二进制分类）。具体来说，对于检测到对象的每一窗口，分类器238可产生二进制1且对于未检测到对象的每一窗口，其可产生二进制0。基于所述数目或1和0的组合，对象检测器208可确定指示目标对象存在于当前视频帧（N）224内的可能性的检测置信度值240。在一些配置中，检测置信度值240为指示已准确检测对象的百分比或机率的介于0与1之间的实数。

[0110] 对象检测器208可根据多种所跟踪参数（包含区域、目标大小、窗口大小、尺度层级、窗口位置和一或多个置信度值）执行对象检测。一旦搜索到视频帧的窗口或窗口子集且对象检测器208获得每一所搜索窗口的二进制值，那么对象检测器208可确定窗口大小以及当前视频帧上具有最高置信度的位置或区域。此位置和窗口大小可用于后续跟踪和检测以更准确地跟踪和/或检测目标对象。

[0111] 可由对象检测器208使用各种技术来检测目标对象。在一个配置中，检测目标对象可包含在每一可能窗口位置和每一可能窗口大小处对窗口执行二进制分类。然而，搜索每一可能窗口为资源密集型操作。因此，在另一配置中，对象检测器可搜索窗口位置和大小的子集而非视频帧中的所有可能窗口。举例来说，对象检测器208可搜索所有可能窗口的1%。接着，如果检测不成功（例如，检测置信度值240低于检测阈值252），那么可在后续所捕获帧中搜索较高百分比的窗口位置（例如，2%）。所搜索窗口位置的百分比步长可是均匀的、不均匀的、慢的或快的，即，连续帧可具有1%、2%、3%、4%或1%、2%、4%、8%。在一个配置

中,响应于高检测置信度值,可将所搜索帧的百分比设定为极高(例如,80%、90%、100%) (即,以确保目标对象在下一视频帧中)。举例来说,响应于检测和跟踪置信度值超出检测和跟踪阈值256,所搜索帧的百分比可跳转到至少80%。替代性地,百分比可跳转到60%、70%、90%等。另外,可使用检测和跟踪阈值的任何合适值(例如,0.6、0.65、0.7、0.75、0.8、0.85等)。此外,可基于随机化器234(随机数产生器)随机确定所搜索窗口的百分比(例如,可在所捕获帧中搜索介于1%与15%之间的随机窗口百分比)。通过搜索所有窗口位置的子集,对象检测可使用电子装置102中的较少资源。

[0112] 本文中所描述的技术可针对每一位置搜索窗口大小的子集。每一窗口大小可在本文中被称作尺度层级,从而每一尺度层级对应于特定窗口大小。举例来说,可存在20个可能尺度层级。可在每一窗口位置处搜索尺度层级或窗口大小的子集而非搜索所有20个尺度层级。

[0113] 本文中所描述的技术也可使用来自存储器缓冲器210的反馈以修整所搜索窗口位置和大小。换句话说,成功检测和/或跟踪到目标对象的最后所捕获视频帧的位置和大小可用作用于搜索当前视频帧(N)224的开始点。举例来说,如果在最近视频帧中检测并跟踪到目标对象(即,最近所捕获视频帧的检测和跟踪置信度值256高于检测和跟踪阈值),那么扫描仪定位器可在与最近帧相关联的位置和大小处开始搜索当前所捕获帧。举例来说,在目标对象移动出光学系统的视野或消失于远处的情况下,目标对象很可能以相同于目标对象离开光学系统的视野或消失于远处时的大小重新出现。因此,当执行对象检测时可预测大小或大小范围以用于检测后续视频帧中的目标对象。

[0114] 所捕获视频帧(N)224中的所搜索窗口位置和窗口大小的搜索范围可限于类似于与最近视频帧(例如,先前视频帧(N-1)222)中的目标对象相关联的窗口位置和窗口大小的那些范围。如本文中所使用,术语“搜索范围”是指当检测和/或跟踪视频帧中的目标对象时可利用的候选窗口位置或候选窗口大小(或所述两者)的集合。举例来说,可基于目标对象在最近视频帧中的发现处从当前视频帧(N)224的一部分内选择所搜索窗口位置的子集(例如,象限中的一者或当前视频帧(N)224的半边)。换句话说,搜索空间可限于最后跟踪或检测到目标对象处附近。类似地,可基于在最近视频帧中发现目标对象的窗口大小限制针对每一窗口位置所搜索的帧大小。举例来说,如果使用具有尺度层级8的窗口在最近帧中检测到对象,那么对于当前视频帧(N)224,扫描仪定标器236可仅选择加上或减去3的窗口尺度层级8(即,尺度层级5到11)。此情况可进一步消除低机率搜索并增加对象检测效率。替代性地,如果最近(非当前)视频帧并未检测到目标对象(即,最近视频帧的检测和跟踪置信度值256低于检测和跟踪阈值),那么对象检测器208可扩展所搜索的搜索空间(窗口位置)(例如,更宽范围的图像或整个图像可经受搜索)。

[0115] 对象跟踪与检测模块204可包含用以合并多个窗口以形成单一窗口的融合模块260。最初存在两个置信度值:来自对象检测器208的检测置信度值240和来自运动跟踪器206的跟踪置信度值225。融合模块260可将两个置信度值(例如,选取较大者)组合成检测和跟踪置信度值256。检测和跟踪置信度值256可指示是否在视频帧上识别出目标对象。在一个配置中,检测和跟踪置信度值256可为介于0与1之间的实数,其中0指示在特定视频帧中识别出目标对象的最低可能置信度且1指示在特定视频帧中识别出目标对象的最高可能置信度。换句话说,检测和跟踪置信度值256可充当发现目标对象的总可能性指示。另外,检测

和跟踪置信度值256可为用于确定在下一视频帧中搜索的窗口位置、窗口大小或窗口百分比的参数。融合模块260可用于将关于当前视频帧(N)224的信息提供到存储器缓冲器210。在一个实例中,融合模块260可将关于所跟踪窗口242(例如,窗口位置244、窗口大小246等)以及检测和跟踪置信度值256的信息提供到存储器缓冲器210。融合模块260可使用来自运动跟踪器206和对象检测器208的跟踪结果(例如,界限框)以形成经组合跟踪结果(例如,界限框)并计算检测和跟踪置信度值256。

[0116] 存储器缓冲器210可存储与先前视频帧(N-1)222、当前视频帧(N)224或其它所捕获视频帧相关联的一或多个值。在一个配置中,存储器缓冲器210存储可包含对应于先前视频帧(N-1)222的信息的所捕获的先前视频帧212。所捕获的先前视频帧212可包含关于一或多个窗口242的信息,包含位置244、窗口大小246和每一窗口242的二进制决策248(例如,来自分类器238)。所捕获的先前视频帧212还可包含跟踪阈值250、检测阈值252以及检测和跟踪阈值254。可将跟踪阈值250提供到对象跟踪与检测模块204上的运动跟踪器206或电路系统(例如,置信水平比较器258)以确定跟踪置信水平是否大于跟踪阈值250。可将检测阈值252提供到对象跟踪与检测模块204上的对象检测器208或其它电路系统以确定检测置信度值240是否大于检测阈值252。检测和跟踪阈值254可为基于跟踪阈值250和检测阈值252的经组合值。可将检测和跟踪阈值254相比于检测和跟踪置信度值256以确定基于运动的跟踪和对象检测的经组合置信度值。阈值中的每一者可是基于目标对象位于视频帧内的可能性。对象跟踪与检测模块204可对当前视频帧(N)224执行基于运动的跟踪和/或检测直到获得特定检测和跟踪置信度值256为止。另外,可对多个视频帧序列中的每一视频帧执行基于运动的跟踪和对象检测。

[0117] 执行基于运动的跟踪和对象检测可包含依序执行基于运动的跟踪接着基于所跟踪参数执行对象检测。具体来说,本发明系统和方法可实施两级跟踪和检测方法。由于基于运动的跟踪是基于场景的相对运动而非如使用对象检测的实际对象识别,因此相比执行对象检测,基于运动的跟踪在电子装置中可是较不资源密集型的。因此,使用运动跟踪器206而非对象检测器208可是更有效的,其中可在无需也执行对象检测的情况下准确跟踪目标对象。

[0118] 因此,在缺乏运动跟踪器206的情况下,对象跟踪与检测模块204仅使用对象检测器208而非并行使用运动跟踪器206与对象检测器208(即,依序而非并行执行运动跟踪和对象检测(如果执行))。对于被执行跟踪的每一视频帧,运动跟踪器206可产生可为指示目标对象处于当前视频帧(N)224中的可能性的介于0与1之间的实数的跟踪置信度值228。

[0119] 在两级跟踪和检测方法的一个配置中,运动跟踪器206可首先对当前视频帧(N)224执行基于运动的跟踪。运动跟踪器206可基于基于运动的跟踪过程确定跟踪置信度值228。使用跟踪置信度值228和由存储器缓冲器210所提供的跟踪阈值250,对象跟踪与检测模块204内的电路系统(例如,置信水平比较器258)可确定跟踪置信度值228是否超出跟踪阈值250。如果跟踪置信度值228大于跟踪阈值250,那么对象跟踪与检测模块204可跳过执行对象检测并将跟踪结果提供到融合模块260以产生输出262。输出262可包含目标对象在当前视频帧(N)224内的指示。另外,输出262可包含关于目标对象的额外信息。

[0120] 如果跟踪置信度值228并不超出跟踪阈值250,那么对象检测器208可随后对当前视频帧(N)224执行对象检测。可对当前视频帧(N)224内的窗口的全部或子集执行对象检

测。对象检测器208也可基于基于运动的跟踪结果和/或从存储器缓冲器210所提供的信息来选择窗口、窗口大小或其它检测准则的子集。可基于提供到对象检测器208的一或多个所跟踪参数来使用或多或少稳定性过程执行对象检测。对象检测器208可确定检测置信度值240并比较检测置信度值240与检测阈值252。如果检测置信度值240高于检测阈值252，那么对象检测器208可将检测结果提供到融合模块260以产生输出262。输出可包含目标对象在当前视频帧(N)224内的指示和/或包含关于所检测对象的额外信息。

[0121] 替代性地，如果检测置信度值240小于或等于检测阈值252，那么对象检测器208可使用较稳定性方法再次执行对象检测(例如，搜索当前视频帧(N)224内的较大数目个窗口)。对象检测器208可重复对象检测过程直到获得令人满意的检测置信度值240为止。一旦获得令人满意的检测置信度值240而使得识别出当前视频帧内的目标对象，对象跟踪与检测模块204可用于对下一视频帧执行跟踪和检测。

[0122] 参考图2B，展示对象跟踪与检测模块204内的处理器264实施的组件的特定说明性实施例。如图2A中所展示，对象跟踪与检测模块204可由处理器264实施。不同处理器可用于实施不同组件(例如，一个处理器可实施运动跟踪器206、另一处理器可用于实施对象检测器208且又一处理器可用于实施存储器缓冲器210)。

[0123] 参考图3，展示用于执行基于运动的跟踪及对象检测的方法300的特定说明性实施例的流程图。可由电子装置102(例如，对象跟踪与检测模块104)实施方法300。电子装置102可通过比较先前视频帧(N-1)222与当前视频帧(N)224而对当前视频帧(N)224执行302基于运动的跟踪。可使用通过跟踪图像对之间的点的中值流方法来执行跟踪对象。也可使用基于运动的跟踪的其它方法。另外，可使用经由存储器缓冲器110所提供的关于所捕获的先前视频帧112的信息来对当前视频帧(N)224执行基于运动的跟踪。

[0124] 电子装置102可确定304跟踪置信度值228。跟踪置信度值228可指示已准确跟踪目标对象的可能性或确定性。电子装置102可确定306跟踪置信度值228是否大于跟踪阈值250。如果跟踪置信度值228大于跟踪阈值250，那么电子装置102可对下一视频帧执行308基于运动的跟踪。另外，电子装置102可基于基于运动的跟踪的结果来跳过对当前视频帧(N)224执行对象检测。换句话说，可仅当运动跟踪并非极好(即，如果跟踪置信度值228不超过跟踪阈值250)时对当前视频帧(N)224执行对象检测。然而，如果跟踪置信度值228不超过跟踪阈值250，那么电子装置102可对当前视频帧(N)224执行310对象检测。电子装置102可在基于运动的跟踪后依序执行对象检测。在一些配置中，可通过变化的稳定性多次执行对象检测以获得较高检测置信度值240。

[0125] 参考图4，展示用于执行基于运动的跟踪的方法400的特定说明性实施例的流程图。可由电子装置102(例如，对象跟踪与检测模块104)实施方法400。电子装置102可使用限界框识别402目标对象。可使用触摸屏116或选择所关注对象的其它输入方法手动地执行402对象识别。可以类似方式识别多个对象。另外，其它输入方法可用于识别待跟踪对象。在一个实例中，通过围绕目标对象绘制限界框来手动地识别对象。

[0126] 电子装置102可初始化404限界框内的网格上的点。网格上的点可遍及限界框均匀地间隔开。另外，可跟踪406两个图像(例如，先前视频帧(N-1)222与当前视频帧(N)224)之间的网格上的点。在一个实例中，由在图像之间产生稀疏运动流的卢卡斯-科纳德(Lucas-Kanade)跟踪器来跟踪点。电子装置102可估计408两个图像(例如，先前视频帧(N-1)222与

当前视频帧(N)224之间的跟踪误差。估计408跟踪误差可包含为所跟踪点中的每一点指派误差值。另外,可使用多种方法执行估计408跟踪误差,包含(例如)前后误差、归一化互相关(NCC)和平方差总和。所估计跟踪误差可用于获得跟踪置信度值228和最后确定目标对象处于当前视频帧(N)224中的可能性。在一个配置中,可通过计算当前视频帧(N)224与先前视频帧(N-1)222中的所跟踪窗口之间的归一化互相关(NCC)来获得跟踪置信度值228。也可使用额外技术估计跟踪误差,包含下文结合图5较详细描述的前后误差估计。另外,电子装置102可滤除410外围点预测。举例来说,电子装置可滤除最坏预测的50%。剩余预测可用于估计限界框的移位。

[0127] 电子装置102可更新412限界框。可执行更新412限界框使得经更新限界框变成用于下一视频帧的新限界框。可接着针对下一视频帧重复基于运动的跟踪过程,或如果跟踪置信度值228小于或等于跟踪阈值250,那么对于下一视频帧,可停止基于运动的跟踪过程直到可准确跟踪目标对象为止。在对当前视频帧(N)224的基于运动的跟踪并不提供令人满意结果的一些配置中,电子装置102可对当前视频帧(N)224执行对象检测以在定位目标对象时获得较高置信水平。在基于运动的跟踪无法产生令人满意结果的一些配置中(例如,当目标对象移动出视频帧范围时),可对任何后续视频帧执行对象检测直到检测到目标对象为止。

[0128] 参看图5,展示用于基于前向-后向误差而估计基于运动的跟踪中的跟踪误差的方法500的特定说明性实施例的流程图。可由电子装置102(例如,对象跟踪与检测模块104)实施方法500。在一些配置中,电子装置102可计算所跟踪窗口之间的归一化互相关(NCC)。归一化互相关(NCC)可用于确定跟踪置信度值228。电子装置102也可使用与归一化互相关(NCC)互补的各种跟踪误差估计技术(例如,前后误差、平方差总和)。在使用前后误差估计的实例中,电子装置102可在先前视频帧(N-1)222与当前视频帧(N)224之间执行502前向跟踪以确定前向轨迹。前向跟踪可包含跟踪向前k步长的图像。所得前向轨迹可等于 $(x_t, x_{t+1}, \dots, x_{t+k})$,其中 x_t 为时间点位置且k指示图像序列长度。电子装置102可在当前视频帧(N)224与先前视频帧(N-1)222之间执行504后向跟踪以确定后向轨迹。所得后向轨迹可等于 $(\hat{x}_t, \hat{x}_{t+1}, \dots, \hat{x}_{t+k})$,其中 $\hat{x}_{t+k} = x_{t+k}$ 。

[0129] 电子装置102可确定506前向轨迹与后向轨迹之间的前后误差。前后误差可界定为前向轨迹与后向轨迹之间的距离。另外,可界定各种距离以用于轨迹比较。在一个配置中,当确定前后误差时,可使用验证轨迹的初始点与结束点之间的欧几里得距离。在一个配置中,前后误差可用作可用于确定跟踪置信度值228的跟踪误差。

[0130] 参考图6,展示用于执行对象检测的方法600的特定说明性实施例的流程图。可由电子装置102(例如,对象跟踪与检测模块104)实施方法600。电子装置102可通过在当前视频帧(N)224中搜索窗口位置和大小的子集来对当前视频帧(N)224执行602对象检测和基于运动的跟踪。

[0131] 电子装置102可确定604检测和跟踪置信度值256。检测和跟踪置信度值256可提供发现目标对象在当前视频帧(N)224中或特定窗口内的置信水平。电子装置102也可确定606检测和置信度值256是否大于检测和跟踪阈值254。如果检测和置信度值256大于检测和跟踪阈值254,那么电子装置102可在下一视频帧中使用窗口和大小的子集(例如,同一子集)来对下一视频帧执行608对象检测。替代性地,如果检测和置信度值256小于检测和跟踪阈

值254,那么电子装置102可在下一视频帧中使用窗口位置和大小的较大子集来对下一视频帧执行610对象检测。在置信度值256小于检测和跟踪阈值254的一些配置中,电子装置102可使用整个搜索空间和/或下一视频帧的所有窗口来对下一视频帧执行610对象检测。

[0132] 参看图7,展示具有不同窗口大小766的图像窗口700的特定实施例。具体来说,图7说明十个可能窗口大小766a到766j的集合。每一窗口大小766可对应于尺度层级(例如,1到10)。尽管本文中展示为矩形,但搜索窗口可为任何形状(例如,正方形、矩形、圆形、椭圆形、自界定等)。此外,任何数目个窗口大小766或尺度层级可是可用的(例如,5个、15个、20个、30个等)。

[0133] 搜索范围可由用于特定位置的窗口大小的子集表示,例如当前视频帧(N)224中所搜索的窗口大小可限于类似于与最近帧中的目标对象相关联的窗口位置和窗口大小的那些窗口大小。举例来说,在并无反馈的情况下,对象检测器208可针对每一所选择窗口位置搜索所有十个窗口大小766a到766j。然而,如果在最近(非当前)视频帧中使用具有第五窗口大小766e的窗口检测对象,那么对于当前所捕获帧,扫描仪定标器236可仅选择加上或减去3的窗口大小5(即,窗口大小2到8)。换句话说,基于来自最近或先前视频帧(N-1)222的反馈,可并不搜索具有第一窗口大小766a、第九窗口大小766i和第十窗口大小766j的窗口。此情况可进一步消除低机率搜索并增加对象检测效率。换句话说,使用来自最近视频帧的反馈可有助于减少所执行计算。替代性地,如果最近视频帧并不检测目标对象(即,最近所捕获帧的检测和跟踪置信度值256小于检测和跟踪阈值254),那么对象检测器208可通过使用大小层级的子集并不限制搜索范围。

[0134] 参考图8,展示对象跟踪与检测模块804的特定说明性实施例。图8中所说明的对象跟踪与检测模块804可包含类似于图2中所说明的对象跟踪与检测模块204的模块并执行类似于其的功能性。具体来说,图8中所说明的对象检测器808、运动跟踪器806、扫描仪定位器830、窗口位置选择器832、随机化器834、扫描仪定标器836、分类器838、融合模块860、存储器缓冲器810、所捕获的先前视频帧812、窗口842、位置844、大小846、二进制决策848、跟踪阈值850、检测阈值852、检测和跟踪阈值854、检测置信度值840、跟踪置信度值828以及检测和跟踪置信度值856可对应于图2中所说明的对象检测器208、运动跟踪器206、扫描仪定位器230、窗口位置选择器232、随机化器234、扫描仪定标器236、分类器238、融合模块260、存储器缓冲器210、所捕获的先前视频帧212、窗口242、位置244、大小246、二进制决策248、跟踪阈值250、检测阈值252、检测和跟踪阈值254、检测置信度值240、跟踪置信度值228以及检测和跟踪置信度值256且具有类似于其的功能性。

[0135] 对象跟踪与检测模块804可包含用于减少归因于目标运动和跟踪误差的抖动影响的平滑模块861。换句话说,平滑模块861平滑跟踪结果,从而致使搜索窗口在位置(x,y)844和大小(宽度、高度)846两者中具有更平滑轨迹。平滑模块861可为简单移动平均(MA)滤波器或自动回归(AR)滤波器。位置844和大小846的平滑程度可不同。例如卡尔曼滤波器的预测性滤波器也可适于位置844平滑。因此,平滑模块861可接收未平滑位置863和未平滑大小865作为输入并输出经平滑位置867和经平滑大小869。

[0136] 参考图9,展示平滑模块961的特定说明性实施例。平滑模块961可用于减少归因于目标运动和跟踪误差的抖动影响(即,因此跟踪结果(限界框)在位置(x,y)和大小(宽度、高度)两者中具有更平滑轨迹)。在一个配置中,使用自动回归(AR)模型实施位置平滑滤波器

971和大小平滑滤波器973以接收未平滑位置963和未平滑大小965作为输入并输出经平滑位置967和经平滑大小969。

[0137] 在自动回归(AR)模型中,假定X为待平滑变数(位置抑或大小)。此外,使X'为由对象跟踪器输出的X的输出。在此配置中,可根据方程式(1)描述X在时间t处的经平滑滤波X_t:

$$[0138] X_t = W * X'_{t-1} + (1-W) * X_{t-1} \quad (1)$$

[0139] 其中X'为X在时间t处的跟踪器输出,X_{t-1}为X在时间t-1处的经平滑结果,且W(0<=W<=1)为控制平滑作用的平滑权重。举例来说,X'可为针对当前视频帧(N)224所选择的窗口位置或窗口大小且X_{t-1}可为用于先前视频帧(N-1)222的窗口位置或窗口大小。

[0140] 不同平滑权重W可用于位置平滑滤波器971和大小平滑滤波器973。举例来说,在一个实施方案中,W_{位置}=0.8且W_{大小}=0.4,使得存在对窗口位置的较少平滑作用但存在对窗口大小的较强平滑作用。此平滑权重选择将产生较少跟踪延迟较少抖动两者。

[0141] 当检测和跟踪置信度值856下降为低于一定阈值(例如,检测和跟踪阈值854)时,也可减少平滑权重选择。当潜在跟踪或检测误差较高时,此情况可带来较强滤波。举例来说,响应于低跟踪置信度(例如,检测和跟踪置信度值856低于检测和跟踪阈值854),可将用于位置和大小的平滑权重分别设定成W_{位置}=0.65且W_{大小}=0.2。换句话说,可降低权重中的一或两者,此情况可致使窗口位置和大小选择较大程度上依赖于先前视频帧的窗口位置和大小而非当前视频帧的窗口位置和大小。

[0142] 权重可是基于跟踪置信度值828或检测置信度值840而非检测和跟踪置信度值856。举例来说,响应于跟踪置信度值828下降为低于跟踪阈值850,可降低平滑权重W_{location}和W_{size}(即,响应于不佳运动跟踪可使用较强滤波)。替代性地,响应于检测置信度值840下降为低于检测阈值852,可降低平滑权重(即,响应于不佳对象检测可使用较强滤波)。

[0143] 在另一配置中,卡尔曼滤波可用于平滑窗口位置。在此配置中,可根据方程式(2)到(7)界定滤波:

$$[0144] x_k = F_k x_{k-1} + w_k \quad (2)$$

$$[0145] z_k = H x_{k-1} + v_k \quad (3)$$

[0146] 其中x_{k-1}为时间k-1处的先前状态,x_k为由x_k=[x, y, Ẃ, ẏ]界定的当前状态,其中(x, y)为限界框中心位置,(Ẃ, ẏ)为每一方向上的速度。此外,状态转变模型F_k和观察模型H可分别由方程式(4)到(5)界定:

[0147]

$$F_k = \begin{bmatrix} 1, 0, \Delta t, 0 \\ 0, 1, 0, \Delta t \\ 0, 0, 1, 0 \\ 0, 0, 0, 1 \end{bmatrix} \quad (4)$$

[0148]

$$H = \begin{bmatrix} 1, 0, 0, 0 \\ 0, 1, 0, 0 \end{bmatrix} \quad (5)$$

[0149] 其中Δt为可调参数。另外,w_k为假定取自具有根据方程式(6)的协方差Q的零均值多变量正态分布的过程噪声(即,w_k~N(0, Q)):

[0150]

$$Q = \begin{bmatrix} 1, 0, 0, 0 \\ 0, 1, 0, 0 \\ 0, 0, 1, 0 \\ 0, 0, 0, 1 \end{bmatrix} * \sigma_1^2 \quad (6)$$

[0151] 其中 σ_1 为可调参数。类似地, v_k 为假定为具有根据方程式(7)的协方差R的零均值高斯白噪声的观察噪声(即, $v_k \sim N(0, R)$)：

[0152]

$$R = \begin{bmatrix} 1, 0, 0, 0 \\ 0, 1, 0, 0 \\ 0, 0, 1, 0 \\ 0, 0, 0, 1 \end{bmatrix} * \sigma_2^2 \quad (7)$$

[0153] 其中 σ_2 为可调参数。

[0154] 参看图10,展示用于平滑运动跟踪结果中的抖动的方法1000的特定说明性实施例的流程图。可由电子装置102(例如,电子装置102中的对象跟踪与检测模块804)执行方法1000。电子装置102可确定1002与当前视频帧224相关联的一或多个窗口位置和一或多个窗口大小(例如,未平滑位置863和未平滑大小865)。电子装置102也可对一或多个窗口位置和一或多个窗口大小进行滤波1004以产生一或多个经平滑窗口位置867和一或多个经平滑窗口大小869。举例来说,此操作可包含使用移动平均滤波器、自动回归滤波器或卡尔曼滤波器。在一个配置中,响应于低跟踪置信度(例如,检测和跟踪置信度值856低于检测和跟踪阈值854),可减少用于位置和大小的平滑权重。替代性地,可基于检测置信度值840或跟踪置信度值828减少平滑权重。电子装置也可使用由一或多个经平滑窗口位置867和一或多个经平滑大小869所界定的一或多个窗口来检测1006当前视频帧224内的目标对象。

[0155] 在场景被成像的情形下,术语“对象”是指场景内的物理对象。在视频流的情形下,术语“对象”是指对象在视频流内的表示(例如,视频流的帧中的对象的图像)。如本文所使用的术语“移动装置”包含呈以下形状因数中的任一者的装置:可抓握的东西(例如,智能电话)、可驾驶的东西(例如,车辆或机器人)、可穿戴的东西(例如,服装或配饰),及可飞行的东西(例如,遥控飞机)。移动装置可包含一或多个屏幕(例如,触摸屏)及/或一或多个图像捕获装置(例如,相机)。

[0156] 可能需要从单个光学视野获得多个不同视野。此能力可用以通过一个相机获得多相机效果。举例来说,此能力可应用于支持通过仅一个相机同时放大场景的两个不同部分(可能以不同的相应变焦速率)。此能力也可以应用于支持在较大的一级视野内彼此独立地移动(例如,平移)的若干二级视野。

[0157] 如本文中所描述系统、方法及设备可实施以对输入视频流操作,所述输入视频流也被称为“一级视频流”。所述一级视频流描述一系列帧,其中每一帧表示像素坐标空间中的图像。一级视频流通常包含可用以从流恢复帧的其它信息(例如,用于每一帧的对应帧起始代码或包以及帧结束代码或包)。视频流还可包含嵌入数据(例如,元数据),其可与特定帧相关联。视频流可由相机或其它成像装置(可对可见波长和/或其它波长具有敏感性)产生,从另一装置串流,或由解码器产生(例如,从存储在磁性或光学媒体上的信息),且可呈压缩或未经压缩形式。一级视频流可包含深度信息,例如基于由结构光成像器或其它深度摄像机(例如,微软Kinect)捕获的图像的视频流。此视频流可通过例如将每一像素的深度

值映射到对应色彩而显示于触摸屏上。所述流可为即时的、经延迟,或从存储装置检索(例如,预记录)。

[0158] 一级视频流具有视野,其可表达为相对于视点(例如,相机或其它图像俘获装置的透镜)的视角。视野的角宽度可由于光学和/或数字变焦操作而随时间改变。图27A展示在正成像的场景的情形中的视野FV10的实例。在此图中,视野FV10由点线指示,视点由小圆指示,且场景的成像的部分由实心矩形指示。图27B展示对如图27A中所示的视野FV10执行缩小操作的结果,这造成其加宽。图27C展示对如图27A中所示的视野FV10执行放大操作的结果,这造成其变窄。在图27B和27C中,虚线矩形和实心矩形分别展示在变焦操作之前和之后的场景的成像的部分。在所有图27A、27B和27C中,视野FOV10是相对于同一视点。

[0159] 视频流描绘可包含一或多个对象的场景。通常,对象是物理且有形的对象(例如,人)。在图27A中,例如场景中的对象包含三个人、一个星和一棵树。然而,还预期对象可为虚拟的或另外为人工的,如绘制和/或计算机产生的内容流(例如,化身)。

[0160] 视野的方向可经界定为视野的中心射线的方向(例如,视锥的中心轴的方向)。此方向可例如由于图像捕获装置的移动和/或较大所捕获图像内的选定窗口的数字平移操作而随时间改变。图27E展示相对于图27D中的视野FV10的方向改变视野FV10的方向(如由点线指示)的结果。此改变可通过例如在视点处旋转图像捕获装置和/或将数字平移操作应用于所捕获图像内的选定窗口而实现。在图27D和27E两者中,视野FOV10是相对于同一视点。

[0161] 图28A展示根据配置的视频处理的方法M100的流程图,其包含任务T200、T300和T350。任务T200选择具有一级视野的一级视频流内的至少两个对象。响应于任务T200中的选择,任务T300从一级视频流产生包含所述选定对象中的第一者的第一视频流。所述第一视频流具有与一级视野相比更窄的第一视野。

[0162] 在任务T200中的选择之后,任务T350从一级视频流产生包含所述选定对象中的第二者的第二视频流。所述第二视频流具有与一级视野相比也更窄的第二视野。所述第一视野包含一级视野的不在所述第二视野内的一部分,且所述第二视野包含一级视野的不在所述第一视野内的一部分。任务T200、T300和T350可例如由图1的电子装置102、由图2B的处理器264实施的对象跟踪和检测模块204、由处理器实施的图8的对象跟踪和检测模块804、由处理器实施的图9的平滑模块961或其任何组合执行。

[0163] 任务T200选择所述一级视频流内的至少两个对象。所述选择可响应于用户的直接动作而执行。在此情况下,屏幕可用以向用户显示所述一级视频流,且用户可输入指示以选择所述显示中出现的对象当中的特定所需对象。如果在触摸屏上执行所述显示,那么用户可通过触摸所述屏幕而指示选择以选择所需对象内的点。

[0164] 图29展示选择三个对象的用户动作的序列的实例。图A展示在装置的触摸屏上显示的场景。所述装置可由用户操纵以起始选择模式。例如,触摸屏可经配置以在用户触摸适当的图标时进入选择模式。图A在帧的右上角中展示选择工具图标的实例。在图B中,用户触摸选择工具图标以起始所述选择模式。在图C中,响应于触摸动作而突出显示工具图标以指示装置处于选择模式中,且用户触摸第一对象以对其进行选择。在图D中,响应于触摸动作而突出显示第一对象(例如,如所展示显出轮廓或勾勒出轮廓)以指示对象被选定,且用户触摸第二对象以对其进行选择。在图E中,响应于触摸动作而突出显示第二对象以指示对象被选定,且用户触摸第三对象以对其进行选择。在图F中,响应于触摸动作而突出显示第三

对象以指示所述对象被选定，且用户再次触摸选择工具图标以终止所述选择模式。所述装置可经配置以在用户在选择模式中再次触摸选定对象的情况下解除对所述选定对象的选择。

[0165] 图30展示选择三个对象的用户动作的序列的另一实例。在此情况下，装置与图29中所展示不同地在选择模式中操作。在图A中，用户触摸选择工具图标以起始选择模式。在图B中，响应于触摸动作而突出显示工具图标以指示装置处于选择模式中，且用户触摸第一对象以对其进行选择。在图C中，响应于触摸动作而突出显示第一对象以指示所述对象被选定，且响应于所述选择，终止选择模式且从工具图标移除突出显示。用户再次触摸选择工具图标以起始选择模式，且在图D中，响应于触摸动作而突出显示工具图标以指示装置处于选择模式中，且用户触摸第二对象以对其进行选择。在图E中，响应于触摸动作而突出显示第二对象以指示所述对象被选定，且响应于所述选择，终止选择模式且从工具图标移除突出显示。用户再次触摸选择工具图标以起始选择模式，且在图F中，响应于触摸动作而突出显示工具图标以指示装置处于选择模式中，且用户触摸第三对象以对其进行选择。

[0166] 在另一实例中，触摸屏可经配置以响应于二动作选择而进入选择模式：第一动作选择菜单模式（例如，用户触摸屏幕的指定区域以显示菜单）且第二动作从菜单选择选择模式。在另一实例中，触摸屏可经配置以使得用户通过触摸并保持选择工具图标、将图标拖动到屏幕上的所要的点，且随后释放所述图标以选择所述点而选择对象。或者，用户可操纵触摸屏以通过选择包含所要的对象的至少一部分的图像的区域（例如，限界框或椭圆或套索）而指示选择。在此情况下，用户可个别地及/或作为群组来选择所要的对象。

[0167] 可能需要允许用户在选择模式期间操纵正显示的视野（例如，执行夹捏动作以缩窄视野，和执行变焦动作以加宽视野）。除触摸屏的操纵以外，用于直接用户选择的其它可能的模式包含致动一或多个按钮或其它开关和/或手势辨识。作为对用户的此直接动作的一个替代方案，任务T200可通过将一或多个预定准则应用于来自一级视频流的信息而自动执行所述选择。此准则可包含例如玩家的球衣号码的辨识、人脸识别和/或速度检测（例如，以识别快速移动对象，例如体育事件中的球）。

[0168] 可实施方法M100以响应于其在任务T200中的选择而跟踪对象。图33C展示方法M100的实施方案M200的流程图，其包含跟踪选定对象的任务T250（例如，如上文参考对象跟踪和检测模块204和/或对象跟踪和检测模块804所描述）。替代地，对象跟踪可在任务T200之前开始。在此情况下，可实施任务T200用于从被跟踪对象的集合之中的选择。举例来说，用户可从被跟踪对象的显示之中进行选择。可以用于此对象跟踪（例如，由任务T250）的算法的实例包含卢卡斯-科纳德方法（和变体，例如卢卡斯-科纳德-托马西跟踪器）和平均移位跟踪。对象的跟踪可包含将对象与一级视频流的帧的其余部分分段，但也可以在无此分段存在下执行对象跟踪。

[0169] 对象与一级视频流的帧的其余部分的分段可用以支持对象相对于帧的其余部分的增强显示（例如，廓形或轮廓）。分段可响应于任务T200中的对象的选择而执行（例如，由任务T250或由如下所述的方法M300的任务T400）。此分段可用以通过确认对象的选择而对用户提供视觉反馈。分段也可以在选择开始之前执行。在此情况下，任务T200中的选择可从已经分段的一组对象当中做出（例如，其中经分段对象增强的显示）。

[0170] 图31中的左侧面板展示描绘包含对象的场景的来自视频流的帧的两个实例。对象

的增强显示可在任务T200中的选择之前或响应于此选择而执行(例如,由任务T250或由如下所述的方法M300的任务T400)。增强显示的一个实例是对象的廓形,其可以与背景(例如,黑色或黄色,如图31的中心面板中所示)形成对比的色彩显示。此对比色彩可以针对全部对象相同或者可以在对象之间不同(例如,以便与每一对象的局部背景形成对比)。增强显示的其它实例(可同样以对比色彩完成)包含对象的轮廓以及在对象处、对象内或对象周围显示图形元素(例如,在对象的顶点或质心处的三角形,如图31的右侧面板中所示的对象周围的椭圆,等)。

[0171] 使用一级视频流的显示作为用于用户的对象的直接选择的情形可适合于缓慢移动对象。然而,如果任何所需对象在显示内快速移动,那么此布置可令用户沮丧。在此快速移动的情况下,可能需要改为显示一级视频流的单个帧作为用于用户的对象的直接选择的情形。如上文所描述,可从其中帧中的对象经分段且增强的显示做出选择。如果流显示选择和帧显示选择都受到支持,那么这些显示模式之间的选择可以由用户执行和/或自动执行。对于自动选择,流中的运动程度可使用例如光流、运动向量、帧之间的像素差等等量度来定量。在此情况下,当所述量度的值高于阈值时可选择帧显示模式,且当所述量度的值低于阈值时可选择流显示模式。

[0172] 响应于任务T200中的选择,任务T300从一级视频流产生包含所述选定对象中的第一者的第一视频流。所述第一视频流具有与一级视野相比更窄的第一视野。在任务T200中的选择之后,任务T350从一级视频流产生包含所述选定对象中的第二者的第二视频流。所述第二视频流具有与一级视野相比也更窄的第二视野。

[0173] 所述第一视野包含一级视野的不在所述第二视野内的一部分,且所述第二视野包含一级视野的不在所述第一视野内的一部分。图28B展示一级视野PFV10的实例,且图28B和28C分别展示第一视野FV100和第二视野FV200的实例,其比一级视野PFV10更窄且其中的每一者包含一级视野PFV10的不在另一视野内的一部分。一级视野内的所述第一视野的宽度和/或方向可随时间改变。同样,一级视野内的所述第二视野的宽度和/或方向可随时间改变。

[0174] 视频流是数字图像(“帧”)的时间序列。在一个实例中,以特定帧速率捕获所述图像序列(例如,以供在同一帧速率下呈现)。视频流可包含指示每一帧的开始和结束的信息。在一个实例中,视频流包含在每一帧的第一像素之前的帧起始代码和在每一帧的最后像素之后的帧结束代码。在另一实例中,视频流包含一或多个数据信号,其载运视频流的帧的像素信息,以及一或多个同步信号,其指示视频流的每一帧的开始和结束。取决于一级视频流的特定格式,一级视频流可包含额外信息(例如,标头、元数据)。一级视频流也可能包含除帧起始与结束代码之间的帧数据(即,像素值)外的数据。此其它数据可包含例如传感器配置细节和/或图像统计值。

[0175] 对视频流的解析包含提取视频流的帧(例如,使用帧开始及结束信息)且使它们可用于处理。例如,解析可包含将每一帧存储到对应指定位置,例如缓冲器。图33A展示包含任务T100的方法M100的实施方案M110的流程图,所述任务剖析一级视频流以产生一级视频流的一系列帧。替代地,另一过程可执行以剖析一级视频流(例如,用于显示),以使得一级视频流的一或多个帧在任何时间可用于方法M100(例如,在存储器中的缓冲器中)。

[0176] 此缓冲器(例如,存储器缓冲器110)通常能够一次存储至少一个全帧。一级视频流

的剖析可经配置以在多个帧缓冲器之间交替。举例来说,一个此帧缓冲器可存储全帧用于处理,同时序列中的随后帧被提取且存储到另一帧缓冲器。替代地,任务T100或剖析过程可经配置以将经剖析帧存储到循环缓冲器,以使得在提取新帧时所述缓冲器经更新(且旧帧被覆写)。

[0177] 剖析一级视频流可包含多路分用操作以将一级视频流与一或多个相关联音频流和/或元数据流分离。替代地,此操作可在任务T100或剖析过程的上游执行(例如,一级视频流可由多路分用器提供)。

[0178] 任务T300和T350中的每一者从一级视频流产生对应二级视频流。此任务可经实施以通过针对二级视频流的每一帧提取一级视频流的一或多个对应帧的所需部分,而产生二级视频流的帧。所需帧部分可直接从所述流(例如,从流缓冲器)或从经剖析帧(例如,如上文所描述存储在帧缓冲器或循环缓冲器中)提取。有可能实施方法M100以使得任务T300响应于第一对象的选择且在第二对象的选择之前产生对应二级视频流。图33B展示方法M110的此实施方案M120的流程图,其表明任务T300和T350可独立地响应任务T200的两个选择T200A和T200B。

[0179] 所需帧部分可由例如对应对象(例如,对象的质量中心)或此对象的限界框的位置以及对应视野的所需宽度和纵横比等参数确定。这些参数的值可例如用以确定所需帧部分的边界作为一级视频流的像素坐标空间内的像素坐标(例如,所需帧部分的左上方和右下方像素的坐标)。

[0180] 所需帧部分的宽度和/或纵横比的值可为默认值,可由用户或过程指示,和/或可基于对象的大小(例如,对象的限界框的大小)。举例来说,所需帧部分的宽度和/或纵横比可由包含二级视频流的显示中的所需窗口数目指示。明确地注意,二级视频流可以但不需要具有与一级视频流相同的纵横比,且每一二级视频流可以但不需要具有唯一纵横比。

[0181] 参考在其像素坐标空间中布置的一级视频流的帧,所需帧部分可视为提取窗口。在一个实例中,实施任务T300和/或T350以通过基于一或多个选定对象的当前位置选择且应用提取窗口的预定集合中的一者,而提取所需帧部分。图32A展示参考一级视频流的帧的九个此类窗口的实例集合,其每一者具有固定大小(即,四分之一帧)和帧内的固定位置。如此实例中所示,所述窗口可重叠,使得对象位置可在包含于所述窗口中的一个以上内的帧的一部分内发生(如由图32B的帧划分中的窗口对应符号指示)。在此情况下,可实施任务以基于历史(例如,以在也当前指示用于选择的情况下继续使用最近选择的窗口)和/或运动(例如,以选择中心与下一预期对象位置最紧密一致的窗口,如由所述帧内的对象的当前位置和一或多个先前位置指示和/或如由捕获装置的移动和/或一级视野的移动指示)而选择指示窗口中的一者。

[0182] 在另一实例中,实施任务T300和/或T350以选择提取窗口大小的预定集合中的一者(例如,基于对象大小,例如对象的限界框的大小)。在此情况下,可实施所述任务以在帧(例如,其中心与一或多个对象的质心对准)内的所需提取位置应用所述选定大小的提取窗口。在另一个实例中,实施任务T300和/或T350以从一或多个选定对象的大小(例如,从对应选定对象的大小)计算提取窗口的大小,且在所述对象的位置(例如,质心)处应用所述窗口。在另一进一步实例中,实施任务T300和/或T350以应用当前经平滑或未经平滑搜索窗口(例如,如本文中参考图7到10所描述)作为提取窗口。在任何情况下,可能需要使对应二级

视频流的多个帧中的每一者内的选定对象的大小归一化。举例来说,可能需要响应于一级视频流内的被跟踪对象的大小的改变(例如,随着对象接近或移动远离捕获装置和/或由于一级视野的宽度的改变)而变化所提取帧部分的(例如,提取窗口的)大小。

[0183] 可能需要根据一或多个呈现约束而选择或计算提取窗口大小和/或位置。举例来说,可能需要响应于检测到对象与所需帧部分的任何边缘之间的距离小于最小像素数目(例如,五个、十个或二十个像素)而增加窗口大小和/或改变窗口位置。替代地或另外,可能需要改变窗口大小以便维持对象的大小与所需帧部分之间的所需关系(例如,以如上文所描述使窗口内的对象的大小归一化)。举例来说,可能需要维持对象尺寸(例如高度和/或宽度)与所需帧部分的对应尺寸的比率不小于预定最小值(例如,百分之五十、六十、七十或八十)和/或不大于预定最大值(例如,百分之八十、九十或九十五)。此边缘接近检测和/或归一化也可以对一级视频流执行(例如,以维持限定选定对象的框的尺寸与其内显示一级视频流的帧的显示窗口的对应尺寸之间的所需关系,例如上文所提及的任何比率)。

[0184] 可能需要实施任务T300和/或任务T350以通过对对应提取的所需帧部分执行一或多个处理操作而产生二级视频流的一或多个帧。可包含调整大小、内插、归一化和/或显示相关调整的此处理可在提取所需帧部分时和/或在帧部分的提取完成之后执行。举例来说,产生二级视频流的帧可包含使所述帧的一或多个色彩和/或明度特性归一化,例如亮度、对比度和白平衡中的一或者者。另外或替代地,任务T300和/或任务T350可包含执行对来自一级视频流的像素信息的一或多个显示相关调整以产生第二视频流。这些调整的实例包含将像素信息变换到不同色彩空间中(例如,YUV到RGB)且执行 γ 调整。

[0185] 可实施任务T300和/或T350以产生与一级视频流具有帧到帧对应的二级视频流。在此情况下,所述任务可从一级视频流的对应帧产生二级视频流的每一帧。可替代地实施任务T300和/或T350以产生具有与一级视频流不同帧速率的二级视频流。在一个此类实例中,实施所述任务以产生具有比一级视频流低的帧速率的二级视频流(例如,通过从一级视频流的对应第n帧产生第一视频流的每一帧,其中n为大于一的整数,和/或在一级视频流的两个或更多个帧之间内插以获得一级视频流的帧速率的非整数分数)。在另一实例中,实施所述任务以产生具有比一级视频流高的帧速率的二级视频流(例如,通过在一级视频流的两个或更多个帧之间内插以产生二级视频流的帧)。

[0186] 可实施任务T300和/或T350以产生对应二级视频流用于传输和/或存储。举例来说,可能需要此任务添加帧起始/结束代码到产生的帧和/或产生指示帧之间的此划分的一或多个同步信号。可执行其它信息和/或格式化以遵守特定流格式(例如,可实施此任务以将像素数据包化为标头加有效负载格式)。可实施任务T300和/或T350以提供二级视频流到压缩器或其它编码器,其可经配置以通过产生例如运动向量和/或预测帧等结构而对所述流进行编码。可使用的存储文件格式的实例包含以下视频容器格式中的任一种:AVI、WMV、MOV、MKV、MPG和MP4。

[0187] 可实施任务T300和/或T350以使元数据与二级视频流关联。此元数据可基于关于被跟踪对象的信息,例如对象位置、用于选定对象的标签(例如,个人的名称)、选定对象之间的几何布置的识别等。另外或替代地,元数据可包含可从装置的其它元件和/或从一级视频流中的元数据获得的信息(例如,相机设定、相机定向、传感器特性、捕获的时间和/或位置)。

[0188] 任务T300和/或T350的此实施方案可将元数据编码为所需格式,例如XML(可扩展标记语言)、KLV(关键长度值)、*.srt(SubRip文件格式)和*.vtt(网络视频文本跟踪格式)当中的任一者。在一些格式中,元数据中的至少一些可包含于二级视频流内(例如,保留用于嵌入数据的帧数据的一部分内)。替代地,可实施流产生任务以封装元数据作为与二级数据流同步的元数据流。举例来说,此流内的元数据的项目可带时戳以指示此项目与二级视频流的对应帧之间的所需关联。

[0189] 也可以实施方法M100以组合一或多个二级视频流与其它流。举例来说,二级视频流可与一或多个其它流(例如,一或多个音频流和/或元数据流)多路复用。另外或替代地,二级视频流可与一或多个其它视频流(例如,其它二级视频流)组合以产生具有带多个窗口的帧的输出流,每一窗口对应于视频流中的不同一者。另外或替代地,可实施方法M100以包含使元数据(例如,以如上文所描述的任何此方式)与一级视频流关联的任务,例如如下参考图17到20所述指示选定对象和/或复合几何形状(例如,选定对象的几何布置)之间的几何形状的信息。

[0190] 替代地或除提供二级视频流用于传输或存储之外,可实施任务T300和/或T350以提供二级视频流用于本地显示。举例来说,此流产生任务可使帧数据可用于显示任务,其可实施以将帧数据存储到显示缓冲器。举例来说,可实施显示任务以将每一二级视频流的帧存储到显示缓冲器的对应于显示的对应窗口的区。

[0191] 图33D展示包含显示任务T400(例如,如上文所描述)的方法M100的实施方案M300的流程图,所述显示任务显示第一视频流和/或第二视频流。可实施任务T400以在显示的相应窗口内显示每一此类流,具有一级视频流的显示(例如,在另一窗口中)或不具有此显示。图37A展示其中显示任务T400也响应于来自跟踪任务T250的输入的方法M200和M300的实施方案M400的流程图。

[0192] 图34A到34C展示方法M300的实施方案(例如,方法M400)的应用的实例,其中图34A展示一级视频流PS10的帧PF10且图34C展示显示视频流DS10的对应帧DF10。图34B的左下方的关键图展示帧PF10中的每一选定对象与对应提取窗口之间的关联。图34B中的中央图展示界定存储在帧缓冲器FB10中的帧的对应所需帧部分FP10到FP40的提取窗口以及由任务T300和/或T350的对应实例产生的二级视频流SS10到SS40。

[0193] 在此实例中,实施任务T400以根据对应于所述流的窗口的在显示帧DF10中的所需布置将二级视频流中的每一者存储到显示缓冲器DB10的显示部分DP20到DP50中的对应一者。也在此实例中,实施任务T400以将一级视频流PS10的经调整大小版本RPS10存储到显示缓冲器DB10的显示部分DP10用于在对应窗口中显示。图34B的右下方的关键图展示显示缓冲器DB10内的每一显示部分DP10到DP50与显示帧DF10的对应窗口之间的关联。

[0194] 任务T300和T350可在每一相对对象的选择之后或在选择操作完成(例如,全部所需对象已经选择,当前不再需要选择)的指示之后即刻开始。对于其中选定对象中的两个极为接近的情况,二级视频流中的一者的产生可经延迟直到对象之间的分离(例如,如由跟踪任务T250指示)达到或超过阈值距离。

[0195] 图35展示方法M100的应用的此实例。面板A展示一级视频流的一级视野PFV20。面板B展示在时间A由任务T300的实例产生的第一视频流的第一视野FV110。在时间A,第一视野FV110包含第一和第二对象两者。面板C展示在时间A之后的时间B由任务T300的实例产生

的第一视野FV110以及由任务T350的实例产生的第二视频流的第二视野FV210。在时间B，第一视野FV110包含第一对象且不再包含第二对象，且第二视野FV210包含第二对象。

[0196] 图11展示方法M100的实施方案(例如，方法M110、M200、M300或M400)的应用的实例。相对于图11中所描绘的实施例描述的视频处理技术可例如由图1的电子装置102、由图2B的处理器264实施的对象跟踪和检测模块204、由处理器实施的图8的对象跟踪和检测模块804、由处理器实施的图9的平滑模块961或其任何组合执行。

[0197] 图11展示一级视频流的帧1100，其描绘在第一时间实例捕获的场景。举例来说，帧1100可对应于将在移动装置的屏幕上显示的视频流的帧。所述移动装置可经配置以捕获所述流(例如，以移动装置的相机)或从另一装置接收所述流。帧1100中所描绘的场景包含第一对象1102、第二对象1104、第三对象1106、星和树。在所说明的实施例中，第一对象1102可对应于第一人，第二对象1104可对应于第二人，且第三对象1106可对应于第三人。在另一实施例中，对象1102到1106中的至少一者可对应于经由移动装置控制的机器人。

[0198] 图11还展示移动装置的屏幕上的帧1100的显示1110。所述移动装置的屏幕(例如，触摸屏)可对应于取景器。显示1110还可包含用户选择机构1111，其使用户能够选择所显示的对象用于跟踪(例如，通过跟踪任务T250)。举例来说，用户可触摸屏幕上的用户选择机构1111，且随后围绕第一对象1102拖动第一聚焦环1112以启用对第一对象1102的跟踪。响应于第一聚焦环1112的此放置，所述移动装置可选择第一对象1102用于所描绘场景内的跟踪。以类似方式，用户还可触摸用户选择机构1111且随后围绕第二对象1104拖动第二聚焦环1114并且触摸用户选择机构1111且随后围绕第三对象1106拖动第三聚焦环1116，以分别启用对第二对象1104和第三对象1106的跟踪。响应于第二聚焦环1114和第三聚焦环1116的此放置，移动装置可分别选择第二对象1104和第三对象1106用于所描绘场景内的跟踪。

[0199] 图11还展示视频流的帧1120，其描绘在第二时间实例(例如，在第一时间实例之后)捕获的场景。在帧1120中，对象1102到1106的位置相对于描绘在第一时间实例的场景的帧1100中的对应位置已改变。举例来说，在帧1120中第一对象1102已移动更接近星，第二对象1104已移动更接近树，且第三对象1106已移动更接近帧的底部。

[0200] 图11还展示移动装置的屏幕上的帧1120的显示1130。显示1130可分割(例如，由显示任务T400)为两个或更多个窗口。在此实例中，显示1130分割成第一窗口1132、第二窗口1134和第三窗口1136。响应于例如通过用户选择机构1111对三个对象1102到1106的选择，移动装置可将显示1130分割为三个对应窗口1132到1136。

[0201] 所述装置可经配置以响应于用户的对象选择完成的指示而应用选定分割。举例来说，虽然可通过在选择对应对象时产生每一窗口来应用默认分割，但可能需要避免应用所述分割直到对象选择完成(例如，以使得显示区域在对象选择期间保持可用)。所述装置可通过例如闪烁聚焦环和/或选定对象来指示对象选择在进展中，直到指示选择完成。用于每一选定对象的聚焦环或其它选择指示可继续跟随或另外指示所述对象(例如，在对象在场景内移动时)直到指示选择完成。

[0202] 第一窗口1132可显示帧1120的包含第一对象1102的第一部分。帧1120的第一部分可基于帧1120中的第一对象1102的位置而确定。举例来说，移动装置可跟踪第一对象1102的位置且在第一窗口1132中显示聚焦于(例如，居中)第一对象1102的位置的视频流。第二窗口1134可显示帧1120的包含第二对象1104的第二部分。帧1120的第二部分可基于帧1120

中的第二对象1104的位置而确定。举例来说，移动装置可跟踪第二对象1104的位置且在第二窗口1134中显示聚焦于第二对象1104的位置上的视频流。第三窗口1136可显示帧1120的包含第三对象1106的第三部分。帧1120的第三部分可基于帧1120中的第三对象1106的位置而确定。举例来说，移动装置可跟踪第三对象1104的位置且在第三窗口1136中显示聚焦于第三对象1106的位置上的视频流。

[0203] 图36A和36B展示窗口1132到1136的其它可能的默认布置的实例。特定默认布置可由用户选择作为装置的显示配置选项。所述布置可为固定的或可由用户修改。在一个实例中，所述装置经配置以使得用户可通过拖动窗口侧面中的一者而将窗口移动到显示中的不同位置，可通过拖动窗口的不位于显示的边界上的拐角而调整窗口的大小，且可通过将指尖或触控笔保持在窗口内（例如，达两秒的时段）而删除窗口。在另一个实例中，所述装置经配置以使得用户可在窗口内双轻触以致使所述装置交换所述窗口的内容（例如，窗口内所显示的二级视频流）与所述显示中当前最大的窗口的内容。

[0204] 图11还展示视频流的帧1140，其描绘在第三时间实例（例如，在第二时间实例之后）的场景。在帧1140中，每一对象1102到1106的位置相对于描绘在第二时间实例的场景的帧1120中的对应位置已经改变。举例来说，在帧1140中第一对象1102已移动更接近树，第二对象已移动更接近所述帧的底部，且第三对象1106已移动更接近星。

[0205] 图11还展示使用由移动装置产生的窗口1132到1136在所述移动装置的屏幕上的帧1140的显示1150。举例来说，第一窗口1132显示帧1140的包含第一对象1102的第一部分，第二窗口1134显示帧1140的包含第二对象1104的第二部分，且第三窗口1136显示帧1140的包含第三对象1106的第三部分。

[0206] 在所述窗口中的每一者内，可能需要使相应选定对象的位置归一化。举例来说，可能需要显示窗口以使得相应选定对象的中心（例如，质量中心）与窗口的中心一致，且在对象在场景内移动时维持此关系。另外或替代地，在所述窗口中的每一者内，可能需要使相应选定对象的大小归一化。举例来说，可能需要显示窗口以使得相应选定对象的最大尺寸在所述窗口的对应尺寸的百分之五十与百分之七十五之间，且在对象朝向或远离相机移动时维持此关系。此显示可例如通过对窗口中正显示的部分调整大小（例如，对其应用数字变焦操作）而实现。另外或替代地，在所述窗口中的每一者内，可能需要使窗口的一或多个显示器方面归一化，其可包含亮度、对比度和白平衡中的任一者。

[0207] 在一特定实施例中，在显示1130和1150中可包含（例如，由显示任务T400）第四窗口（例如，如图36C中所示的窗口1138）。所述第四窗口可说明如相对于图17进一步描述的每一对象之间的几何形状。另外，或在替代方案中，所述第四窗口可显示包含帧1120和1140的视频流。举例来说，第一窗口1132、第二窗口1134和第三窗口1136可对应于帧1120和1140的画中画拍摄，且所述第四窗口可对应于整个帧1120和1140。在此情况下，可能需要控制相机的光学和/或数字变焦操作以使得全部选定对象都在第四窗口中正显示的流内保持可见。响应于对象选择的变焦的此组合也可以在无显示分割存在下执行。替代地或另外，可能需要在第四窗口中正显示的流内指示所述选定对象。举例来说，标记可上覆于每一对象上（例如，具有所述对象相应的对比色彩且上覆于所述对象的质心的形状）和/或可在对象周围布置晕圈（例如，椭圆）（所述晕圈内可能具有较低对比度或其它区别外观）。图36D和36E展示窗口1132到1138的可能布置的两个其它实例。

[0208] 在另一特定实施例中,可选择对单个对象(例如,仅第一对象1112)的跟踪用于显示。响应于选择对象用于显示,所述显示可分割成两个窗口(例如,如图36F中所示)。在此情况下,第一窗口可包含基于场景中的选定对象的位置而确定的场景的一部分。举例来说,移动装置可跟踪所述选定对象的位置且聚焦于所述对象的所述位置上的视频流可在第一窗口中显示(例如,由任务T400)。另外,第二窗口可包含整个场景。因此,所述第一窗口可对应于场景中的选定对象的聚焦画中画拍摄。

[0209] 图41A和41B展示包含选定对象的一级视频流的视野的一部分的显示的两个实例。此显示可在如上文所描述的窗口(例如,窗口1132、1134、1136或1138)中显示,其可以展示一级视频流的整个视野的窗口来展示或替代于所述窗口而展示。这些实例中的每一者包含图标(图41A中的虚线框,和图41B的顶部附近的有界的线),其相对于窗口的大小的尺寸指示所显示(变窄)视野与一级视频流的视野的关系。在这些特定实例中,变窄的视野在选定对象的复合几何形状(由开放的圆指示)的位置处居中。这些实例中的每一者还包含指示此位置的最近轨迹的点线。

[0210] 图42展示显示视频流(例如,如由显示任务T400的实施方案产生)的帧的实例。此帧包含沿着所述帧的底部的三个窗口,其显示对应于选定对象A、B和C的二级视频流(例如,如由任务T300和T350的实例产生)。左上方窗口显示一级视频流,其可经调整大小以配合所述窗口。此窗口还包含覆层(在此实例中,十字准线),其指示对应帧中的选定对象的复合几何形状的位置。右上方窗口显示包含选定对象的一级视频流的视野的一部分,如上文参考图41A到41B所描述。

[0211] 图43B展示指示选定对象的复合几何形状、此几何形状的位置以及此位置的最近轨迹的图形的实例。此图形可显示为单独窗口或上覆于一级视频流的显示上。图43A和43C展示显示包含相似覆层的一级视频流的帧的窗口。

[0212] 图11中所描绘的实施例可使移动装置能够跟踪帧1100、1120、1140中的对象1102到1106且可在单独窗口1132到1136中显示对应于对象1102到1106的个别化(例如,聚焦)视频流。所述移动装置可使用单个相机或使用多个相机执行相对于图11描述的技术。在分割正经配置(例如,由用户)且应用于视频流的同时,所述装置可继续记录未分割的视频流和/或各自包含所述窗口中的一或者者的一或多个视频流。另外或作为存储各自包含所述窗口中的一或者者的一或多个视频流的一个替代方案,可能需要装置存储元数据(嵌入视频流中和/或作为单独文件),从所述元数据,可以从记录的未分割的流重新产生特定分割。

[0213] 参考图12,展示说明使用对象跟踪的视频处理方法1200的特定实施例的流程图。在说明性实施例中,方法1200可使用图1的电子装置102、由图2B的处理器264实施的对象跟踪和检测模块204、由处理器实施的图8的对象跟踪和检测模块804、由处理器实施的图9的平滑模块961或其任何组合执行。

[0214] 方法1200包含在1202在移动装置处接收对所显示场景中的第一对象和所显示场景中的第二对象的选择。举例来说,参考图11,用户可触摸屏幕上的用户选择机构1111,且随后在第一对象1102周围拖动第一聚焦环1112,且触摸屏幕上的用户选择机构1111,且随后在第二对象1104周围拖动第二聚焦环1114,以分别启用对第一对象1102和第二对象1104的跟踪。所显示场景可对应于在移动装置的屏幕上捕获和显示的视频流。

[0215] 在1204,所述显示可分割成至少第一窗口和第二窗口。举例来说,参考图11,显示

1130可分割成第一窗口1132、第二窗口1134和第三窗口1136。举例来说，响应于以用户选择机构1111选择三个对象1102到1106，所述移动装置可将显示1130分割为三个对应窗口1132到1136。

[0216] 在1206，可在第一窗口中显示所述场景的包含第一对象的第一部分。举例来说，参考图11，第一窗口1132可显示帧1120的包含第一对象1102的第一部分。帧1120的第一部分可基于帧1120中的第一对象1102的位置而确定。举例来说，所述移动装置可跟踪第一对象1102的位置，且聚焦于第一对象1102的所述位置上的视频流可在第一窗口1132中显示。

[0217] 在1208，可在第二窗口中显示所述场景的包含第二对象的第二部分。举例来说，参考图11，第二窗口1134可显示帧1120的包含第二对象1104的第二部分。帧1120的第二部分可基于帧1120中的第二对象1104的位置而确定。举例来说，所述移动装置可跟踪第二对象1104的位置，且聚焦于第二对象1104的所述位置上的视频流可在第二窗口1134中显示。

[0218] 图12的方法1200可使所述移动装置能够跟踪场景1100、1120、1140中的对象1102到1106，且可在单独窗口1132到1136中显示对应于对象1102到1106的个别化(例如，聚焦)视频流。所述移动装置可使用单个相机或使用多个相机执行方法1200。

[0219] 图13展示方法M100的实施方案(例如，方法M110、M200、M300或M400)的应用的实例。相对于图13中所描绘的实施例描述的视频处理技术可例如由图1的电子装置102、由图2B的处理器264实施的对象跟踪和检测模块204、由处理器实施的图8的对象跟踪和检测模块804、由处理器实施的图9的平滑模块961或其任何组合执行。

[0220] 图13展示一级视频流的帧1300，其描绘在第一时间实例捕获的场景。举例来说，帧1300可对应于将在移动装置的屏幕上显示的视频流的帧。所述移动装置可经配置以捕获所述流(例如，以移动装置的相机)或从另一装置接收所述流。帧1300中所描绘的场景包含第一对象1302、第二对象1304、第三对象1306、星和树。在所说明的实施例中，第一对象1302可对应于第一人，第二对象1304可对应于第二人，且第三对象1306可对应于第三人。在另一实施例中，对象1302到1306中的至少一者可对应于经由移动装置控制的机器人。

[0221] 图13还展示移动装置的屏幕上的帧1300的显示1310。所述移动装置的屏幕(例如，触摸屏)可对应于取景器。显示1310可包含用户选择机构1311，其使用户能够选择对象用于跟踪(例如，由跟踪任务T250)。举例来说，用户可触摸屏幕上的用户选择机构1311，且随后围绕第一对象1302拖动第一聚焦环1312以启用对第一对象1302的跟踪。响应于第一聚焦环1312的此放置，所述移动装置可选择第一对象1312用于所描绘场景内的跟踪。以类似方式，用户还可触摸屏幕上的用户选择机构1311且在第二对象1304和第三对象1306周围拖动第二聚焦环1314和第三聚焦环1316以分别启用对第二对象1304和第三对象1306的跟踪。响应于第二聚焦环1314和第三聚焦环1316的此放置，所述移动装置可分别选择第二对象1304和第三对象1306用于跟踪。

[0222] 图13还展示视频流的帧1320，其描绘在第二时间实例捕获的场景。在帧1320中，每一对象1302到1306的位置相对于描绘在第一时间实例的场景的帧1300中的对应位置已经改变。举例来说，在帧1320中第一对象1302已移动更接近第二对象1304，第二对象1304已移动更接近第一对象1302，且第三对象1306已移动更接近帧的底部。

[0223] 图13还展示移动装置的屏幕上的帧1320的显示1330。显示1330可分割(例如，由显示任务T400)为两个或更多个窗口。在此实例中，显示1330分割成第一窗口1332、第二窗口

1334和第三窗口1336。举例来说,响应于以用户选择机构1311选择三个对象1302到1306,所述移动装置可将显示1330分割为三个对应窗口1332到1336。

[0224] 所述移动装置可经配置以响应于用户的对象选择完成的指示而应用选定分割。举例来说,虽然可通过在选择对应对象时产生每一窗口来应用默认分割,但可能需要避免应用所述分割直到对象选择完成(例如,以使得显示区域在对象选择期间保持可用)。所述装置可通过例如闪烁聚焦环和/或选定对象来指示对象选择在进展中,直到指示选择完成。用于每一选定对象的聚焦环或其它选择指示可继续跟随或另外指示所述对象(例如,在对象在场景内移动时)直到指示选择完成。

[0225] 第一窗口1332可显示帧1320的包含第一对象1302的第一部分。帧1320的第一部分可基于帧1320中的第一对象1302的位置而确定。举例来说,所述移动装置可跟踪第一对象1302的位置且在第一窗口1332中显示聚焦于(例如,居中)第一对象1302的位置上的视频流。第二窗口1334可显示帧1320的包含第二对象1304的第二部分。帧1320的第二部分可基于帧1320中的第二对象1304的位置而确定。举例来说,所述移动装置可跟踪第二对象1304的位置且在第二窗口1334中显示聚焦于第二对象1304的所述位置上的视频流。第三窗口1336可显示帧1320的包含第三对象1306的第三部分。帧1320的第三部分可基于帧1320中的第三对象1306的位置而确定。举例来说,所述移动装置可跟踪第三对象1304的位置且在第三窗口1336中显示聚焦于第三对象1306的所述位置上的视频流。

[0226] 图13还展示帧1340,其描绘在第三时间实例(例如,在第二时间实例之后)的场景。在帧1340中,每一对象1302到1306的位置相对于描绘在第二时间实例的场景的帧1320中的对应位置已经改变。举例来说,在帧1340中第一对象1302和第二对象1304已彼此交叉,且第三对象1306已移动更接近所述帧的右下方部分。

[0227] 图13还展示使用由移动装置产生的窗口1332到1336在所述移动装置的屏幕上的帧1340的显示1350。将了解,帧1340的第一部分在第二窗口1334中显示且帧1340的第二部分在第一窗口1332中显示。举例来说,如果第一对象1302的一部分与第二对象1304的一部分重叠阈值量(如例如由跟踪任务T250指示),那么第一窗口1332可显示聚焦于第二对象1304上的视频流且第二窗口1334可显示聚焦于第一对象1302上的视频流。第三窗口1336显示帧1340的包含第三对象1306的第三部分。

[0228] 在一特定实施例中,所述阈值量可对应于第一对象1302与第二对象1304的完全重叠。举例来说,当与第一对象1302的位置相关联的每一x坐标(或y坐标)具有小于与第二对象1304的位置相关联的每一x坐标(或y坐标)的值时,第一对象1302可在当与第一对象1302的位置相关联的每一x坐标(或y坐标)具有大于与第二对象1304的位置相关联的每一x坐标(或y坐标)的值时的帧处与第二对象1304完全重叠(例如,交叉)。

[0229] 替代地,当与第一对象1302的位置相关联的每一x坐标(或y坐标)具有大于与第二对象1304的位置相关联的每一x坐标(或y坐标)的值时,第一对象1302可在当与第一对象1302的位置相关联的每一x坐标(或y坐标)具有小于与第二对象1304的位置相关联的每一x坐标(或y坐标)的值时的帧处与第二对象1304完全重叠(例如,交叉)。

[0230] 在一个实例中,当在一尺寸(例如,高度或宽度)中两个选定对象的限界框的中心之间的距离小于在所述尺寸中所述两个框的大小之和的二分之一时可指示重叠。在此情况下,跟踪任务可假定最接近于帧的底部的对象与另一对象重叠。在另一实例中,执行色彩直

方图以使相应色彩值范围与每一选定对象关联。在此情况下,当两个对象接近且(A)具有在一个对象范围内的色彩值的像素与(B)具有在另一对象范围内的色彩值的像素之间的空间边界被发现在所述对象中的一者的预期区域内时可指示重叠。

[0231] 在另一特定实施例中,所述阈值量可对应于第一对象1302与第二对象1304的部分重叠(例如,重叠对象的预期区域的百分之十、二十、二十五、三十、四十或五十)。举例来说,在其中与第一对象1302的位置相关联的x坐标(或y坐标)具有与第二对象1304的位置相关联的x坐标(或y坐标)相同的值的帧中,第一对象1302可部分地与第二对象1304重叠。

[0232] 图13中所描绘的实施例可使移动装置能够跟踪帧1300、1320、1340中的对象1302到1306,且可在单独窗口1332到1336中显示对应于对象1302到1306的个别化(例如,聚焦)视频流。图13中所描绘的实施例还可使每一窗口1332到1336中的视频流能够描绘每一对象1302到1306相对于彼此的相对位置。所述移动装置可使用单个相机或使用多个相机执行相对于图13描述的技术。

[0233] 可能需要实施显示任务T400以如上文所描述响应于检测到选定对象之间的遮挡而修改显示分割。举例来说,当特定对象彼此交叉(例如,如由跟踪任务指示)时,二级视频流与对应窗口1332到1336之间的关联可经更新(例如,切换)以反射所述交叉。可能需要配置所述装置以通过例如闪烁所述切换对象和/或窗口达短暂周期(例如,一、二或三秒)来唤起对切换的注意。

[0234] 可能需要应用惯性机构(例如,在跟踪任务T250和/或显示任务T400内)以在一些时间中禁止保持极为接近的对象之间的过度频繁切换。在一个此类实例中,应用延滞周期以使得两个对象之间的切换在延滞周期(例如,二、三、四或五秒)内将不会逆转。在另一实例中,将滞后功能应用于所述阈值量以使得两个对象之间的切换在延滞周期内将不会逆转,除非在所述延滞周期内发生第二较高阈值量的重叠。

[0235] 参考图14,展示说明使用对象跟踪的视频处理方法1400的特定实施例的流程图。在说明性实施例中,方法1400可使用图1的电子装置102、由图2B的处理器264实施的对象跟踪和检测模块204、由处理器实施的图8的对象跟踪和检测模块804、由处理器实施的图9的平滑模块961或其任何组合执行。

[0236] 方法1400包含在1402在移动装置处接收对所显示场景中的第一对象和所显示场景中的第二对象的选择。举例来说,参考图13,用户可触摸屏幕上的用户选择机构1311且围绕第一对象1302拖动第一聚焦环1312和围绕第二对象1304拖动第二聚焦环1314以分别启用对第一对象1302和第二对象1304的跟踪。所显示场景可对应于移动装置的屏幕上捕获且显示的视频流的帧(例如,帧1300)。

[0237] 在1404,可将屏幕分割成至少第一窗口和第二窗口。举例来说,参考图13,显示1330可分割成第一窗口1332、第二窗口1334和第三窗口1336。举例来说,响应于以用户选择机构1311选择三个对象1302到1306,所述移动装置可将显示1330分割为三个对应窗口1332到1336。

[0238] 在1406,可初始地在第一窗口中显示所述场景的包含第一对象的第一部分。举例来说,参考图13,第一窗口1332可初始地显示帧1320的包含第一对象1302的第一部分。帧1320的第一部分可基于帧1320中的第一对象1302的位置而确定。举例来说,所述移动装置可跟踪第一对象1302的位置,且聚焦于第一对象1302的所述位置上的视频流可在第一窗口

1332中显示。

[0239] 在1408,可初始地在第二窗口中显示所述场景的包含第二对象的第二部分。举例来说,参考图13,第二窗口1334可初始地显示帧1320的包含第二对象1304的第二部分。帧1320的第二部分可基于帧1320中的第二对象1304的位置而确定。举例来说,所述移动装置可跟踪第二对象1304的位置,且聚焦于第二对象1304的所述位置上的视频流可在第二窗口1334中显示。

[0240] 其中显示第一对象和第二对象的窗口可基于第一对象的一部分何时与第二对象的一部分重叠阈值量来控制,1410。举例来说,参考图13,帧1340的第一部分在显示1350的第二窗口1334中显示,且帧1340的第二部分在显示1350的第一窗口1332中显示。举例来说,如果第一对象1302的一部分与第二对象1304的一部分重叠阈值量,那么第一窗口1332可显示聚焦于第二对象1304上的视频流且第二窗口1334可显示聚焦于第一对象1302上的视频流。第三窗口1336显示帧1340的包含第三对象1306的第三部分。

[0241] 图14的方法1400可使移动装置能够跟踪帧1300、1320、1340中的对象1302到1306,且可在单独窗口1332到1336中显示对应于对象1302到1306的个别化(例如,聚焦)视频流。方法1400还可使每一窗口1332到1336中的视频流能够描绘每一对象1302到1306相对于彼此的相对位置。举例来说,当特定对象彼此交叉时,窗口1332到1336中的视频流可经更新(例如,切换)以反映所述交叉。所述移动装置可使用单个相机或使用多个相机执行方法1400。

[0242] 图15展示方法M100的实施方案(例如,方法M110、M200、M300或M400)的应用的实例。相对于图15中所描绘的实施例描述的视频处理技术可由图1的电子装置102、由图2B的处理器264实施的对象跟踪和检测模块204、由处理器实施的图8的对象跟踪和检测模块804、由处理器实施的图9的平滑模块961或其任何组合执行。

[0243] 图15展示一级视频流的帧1500,其描绘在第一时间实例捕获的场景。举例来说,帧1500可对应于将在移动装置的屏幕上显示的视频流的帧。所述移动装置可经配置以捕获所述流(例如,以移动装置的相机)或从另一装置接收所述流。帧1500中所描绘的场景包含第一对象1502、第二对象1504、第三对象1506、星和树。在所说明的实施例中,第一对象1502可对应于第一人,第二对象1504可对应于第二人,且第三对象1506可对应于第三人。在另一实施例中,对象1502到1506中的至少一者可对应于经由移动装置控制的机器人。

[0244] 图15还展示移动装置的屏幕上的帧1500的显示1510。所述移动装置的屏幕(例如,触摸屏)可对应于取景器。显示1510可包含用户选择机构1511,其使用户能够选择对象用于跟踪(例如,由跟踪任务T250)。举例来说,用户可触摸屏幕上的用户选择机构1511且围绕第一对象1502拖动第一聚焦环1512以启用对第一对象1502的跟踪。响应于第一聚焦环1512的此放置,所述移动装置可选择第一对象1502用于跟踪。以类似方式,用户还可触摸屏幕上的用户选择机构1511且围绕第二对象1504和第三对象1506拖动第二聚焦环1514和第三聚焦环1516以分别启用对第二对象1504和第三对象1506的跟踪。响应于第二聚焦环1514和第三聚焦环1516的此放置,所述移动装置可分别选择第二对象1504和第三对象1506用于跟踪。

[0245] 图15还展示视频流的帧1520,其描绘在第二时间实例(例如,在第一时间实例之后)捕获的场景。在帧1520中,每一对象1502到1506的位置相对于描绘在第一时间实例的场景的帧1500中的对应位置已经改变。举例来说,在帧1520中第一对象1502已移动更接近星,

第二对象1504已移动更接近树，且第三对象已移动更接近帧的底部。

[0246] 图15还展示在第二对象1504之前选择第一对象1502且在第三对象1506之前选择第二对象1504的情况下移动装置的屏幕上的帧1520的显示1530。举例来说，显示1530可分割(例如，由显示任务T400)为第一窗口1532、第二窗口1534和第三窗口1506。在每一窗口1532到1536中显示的视频流可取决于对象1502到1506的选择次序。

[0247] 如果在其它对象1504、1506的选择之前选择第一对象1502，那么第一窗口1532可显示帧1520的包含第一对象1502的第一部分。帧1520的第一部分可基于帧1520中的第一对象1502的位置而确定。举例来说，所述移动装置可跟踪第一对象1502的位置，且聚焦于第一对象1502的所述位置上的视频流可在第一窗口1532中显示。如果在第三对象1506的选择之前选择第二对象1504，那么第二窗口1534可显示帧1520的包含第二对象1504的第二部分。帧1520的第二部分可基于帧1520中的第二对象1504的位置而确定。举例来说，所述移动装置可跟踪第二对象1504的位置，且聚焦于第二对象1504的所述位置上的视频流可在第二窗口1534中显示。第三窗口1536可显示帧1520的包含第三对象1506的第三部分。帧1520的第三部分可基于帧1520中的第三对象1506的位置而确定。举例来说，所述移动装置可跟踪第三对象1504的位置，且聚焦于第三对象1506的所述位置上的视频流可在第三窗口1506中显示。

[0248] 图15还展示在第三对象1506之前选择第二对象1502且在第一对象1502之前选择第三对象1506的情况下移动装置的屏幕上的帧1520的显示1530。举例来说，显示1530可分割为第一窗口1532、第二窗口1534和第三窗口1506。在每一窗口1532到1536中显示的视频流可取决于对象1502到1506的选择次序。

[0249] 如果在其它对象1502、1506的选择之前选择第二对象1504，那么第一窗口1532可显示帧1520的包含第二对象1504的第二部分。如果在第一对象1502的选择之前选择第三对象1506，那么第二窗口1534可显示帧1520的包含第三对象1506的第三部分。第三窗口1536可显示帧1520的包含第一对象1502的第一部分。

[0250] 图15中所描绘的实施例可使移动装置能够跟踪帧1500、1520中的对象1502到1506且在单独窗口1532到1536中显示对应于对象1502到1506的个别化(例如，聚焦)视频流。在特定窗口1532到1536中显示的视频流可取决于对象1502到1506的选择次序。因此，用户可基于重要性(或其它选择准则)确定在特定窗口1532到1536中将放置哪一对象1502到1506。所述移动装置可使用单个相机或使用多个相机执行相对于图15描述的技术。

[0251] 参考图16，展示说明使用对象跟踪的视频处理方法1600的特定实施例的流程图。在说明性实施例中，方法1600可使用图1的电子装置102、由图2B的处理器264实施的对象跟踪和检测模块204、由处理器实施的图8的对象跟踪和检测模块804、由处理器实施的图9的平滑模块961或其任何组合执行。

[0252] 方法1600包含在1602在移动装置处接收对所显示场景中的第一对象和所显示场景中的第二对象的选择。举例来说，参考图15，用户可触摸屏幕上的用户选择机构1511且围绕第一对象1502拖动第一聚焦环1512和围绕第二对象1504拖动第二聚焦环1514以分别启用对第一对象1502和第二对象1504的跟踪。所显示场景可对应于在移动装置的屏幕上捕获和显示的视频流。

[0253] 在1604可将所述显示分割成第一窗口和第二窗口。举例来说，参看图15，显示1530

可分割成第一窗口1532、第二窗口1534和第三窗口1536。举例来说，响应于以用户选择机构1511选择三个对象1502到1506，所述移动装置可将显示1530分割为三个对应窗口1532到1536。

[0254] 在1606，如果在第二对象的选择之前选择第一对象，那么可在第一窗口中显示所述场景的包含第一对象的第一部分。举例来说，参看图15，如果在其它对象1504、1506的选择之前选择第一对象1502，那么第一窗口1532可显示帧1520的包含第一对象1502的第一部分。帧1520的第一部分可基于帧1520中的第一对象1502的位置而确定。举例来说，所述移动装置可跟踪第一对象1502的位置，且聚焦于第一对象1502的所述位置上的视频流可在第一窗口1532中显示。

[0255] 在1608，如果在第二对象的选择之前选择第一对象，那么可在第二窗口中显示所述场景的包含第二对象的第二部分。举例来说，参看图15，第二窗口1534可显示帧1520的包含第二对象1504的第二部分。帧1520的第二部分可基于帧1520中的第二对象1504的位置而确定。举例来说，所述移动装置可跟踪第二对象1504的位置，且聚焦于第二对象1504的所述位置上的视频流可在第二窗口1534中显示。

[0256] 在1610，如果在第一对象的选择之前选择第二对象，那么可在第二窗口中显示所述场景的第一部分。另外，在1612，如果在第一对象的选择之前选择第二对象，那么可在第一窗口中显示所述场景的第二部分。

[0257] 图16的方法1600可使移动装置能够跟踪场景1500、1520中的对象1502到1506且可在单独窗口1532到1536中显示对应于对象1502到1506的个别化(例如，聚焦)视频流。在特定窗口1532到1536中显示的视频流可取决于对象1502到1506的选择次序。因此，用户可基于重要性(或其它选择准则)确定在特定窗口1532到1536中将放置哪一对象1502到1506。所述移动装置可使用单个相机或使用多个相机执行方法1600。

[0258] 图17展示方法M100的实施方案(例如，方法M110、M200、M300或M400)的应用的实例。相对于图17中所描绘的实施例描述的视频处理技术可由图1的电子装置102、由图2B的处理器264实施的对象跟踪和检测模块204、由处理器实施的图8的对象跟踪和检测模块804、由处理器实施的图9的平滑模块961或其任何组合执行。

[0259] 图17展示一级视频流的帧1700，其描绘在第一时间实例捕获的场景。举例来说，帧1700可对应于将在移动装置的屏幕上显示的视频流。所述移动装置可经配置以捕获所述流(例如，以移动装置的相机)或从另一装置接收所述流。帧1700中所描绘的场景包含第一对象1702、第二对象1704、第三对象1706和星。在所说明的实施例中，第一对象1702可对应于第一人，第二对象1704可对应于树，且第三对象1706可对应于第二人。在另一实施例中，第一对象1702和第三对象1706中的至少一者可对应于经由移动装置控制的机器人。

[0260] 图17还展示移动装置的屏幕上的帧1700的显示1710。所述移动装置的屏幕(例如，触摸屏)可对应于取景器。显示1710可包含使用户能够选择用于跟踪的对象(例如，通过任务T250)的用户选择机构1711。举例来说，用户可触摸屏幕上的用户选择机构1711且围绕第一对象1702拖动第一聚焦环1712以启用对第一对象1702的跟踪。响应于第一聚焦环1712的此放置，所述移动装置可选择帧1700中的第一对象1712用于跟踪。以类似方式，用户还可触摸屏幕上的用户选择机构1711且围绕第二对象1704和第三对象1706拖动第二聚焦环1714和第三聚焦环1716以分别启用对第二对象1704和第三对象1706的跟踪。响应于第二聚焦环

1714和第三聚焦环1716的此放置,所述移动装置可分别选择第二对象1704和第三对象1706用于跟踪。

[0261] 图17还展示视频流的帧1720,其描绘在第二时间实例(例如,在第一时间实例之后)捕获的场景。在帧1720中,每一对象1702到1706的位置相对于描绘在第一时间实例的场景的帧1700中的对应位置已经改变。举例来说,在帧1720中第一对象1702和第三对象1704已经移动。虽然第二对象1704(例如,树)是静止对象,但第二对象1704可表现为已经移动。举例来说,所述移动装置(或其它捕获装置)的位置可移动,这又可产生第二对象1704已移动的错觉。

[0262] 移动装置可监视选定对象之间的一或多个空间关系(也被称为“几何形状”)。例如,移动装置可跟踪并测量第一对象1702与第二对象1704之间的第一距离、第二对象1704与第三对象1706之间的第二距离,及第三对象1706与第一对象1702之间的第三距离。在特定实施例中,跟踪每一选定对象1702到1706之间的几何形状(例如,通过跟踪任务T250)可在触发事件之后起始。作为非限制性实例,跟踪每一选定对象1702到1706之间的几何形状可在第一对象1702越过第二对象1704之后起始。在其它实施例中,跟踪选定对象1702到1706之间的几何形状可在用户输入之后起始。作为非限制性实例,跟踪每一选定对象1702到1706之间的几何形状可在用户对对象1702到1706中的两者或更多者的选择之后起始。作为另一实例,跟踪每一选定对象1702到1706之间的几何形状可在用户选择起始跟踪模式之后起始。

[0263] 图17还展示帧1720在移动装置的屏幕上的显示1730(例如,通过显示任务T400)。移动装置可包含显示1730上的几何形状的指示。例如,第一对象1702与第二对象1704之间的第一距离的第一指示(L_{12})可包含在显示1730中。可基于帧1720中的第一对象1702的位置及帧1720中的第二对象1704的位置而确定所述第一距离。例如,移动装置可跟踪对象1702的位置及第二对象1704的位置以产生第一指示(L_{12})。第二对象1704与第三对象1706之间的第二距离的第二指示(L_{23})也可包含在显示1730中。可基于帧1720中的第二对象1704的位置及帧1720中的第三对象1706的位置而确定所述第二距离。例如,移动装置可跟踪第二对象1704的位置及第三对象1706的位置以产生所述第二指示(L_{23})。第一对象1702与第三对象1706之间的第三距离的第三指示(L_{13})也可包含在显示1730中。可基于帧1720中的第一对象1702的位置及帧1720中的第三对象1706的位置而确定第三距离。例如,移动装置可跟踪第一对象1702的位置及第二对象1704的位置以产生所述第三指示(L_{13})。

[0264] 移动装置还可跟踪对象1702到1706的复合几何形状(例如,选定对象的几何布置)。例如,在所说明的实施例中,对象1702到1706的几何布置可表征为由三个指示(L_{12} 、 L_{23} 、 L_{13})形成的三角形。每一对象之间的几何形状的至少一个参数或复合几何形状的至少一个参数可用于将视频流的帧群集。作为非限制性实例,可基于两个特定对象1702到1706之间的距离(例如,第一距离、第二距离及/或第三距离)而将特定帧群集。作为另一非限制性实例,可基于多个对象1702到1706的质量中心(C_{M2})而特定帧群集。例如,移动装置可确定由每一帧的三个指示(L_{12} 、 L_{23} 、 L_{13})形成的三角形的质量中心(C_{M2})。在特定实施例中,质量中心(C_{M2})可经计算、编索引并且在屏幕1730上显示。在另一实施例中,质量中心(C_{M2})可仅被计算及编索引。其中选定对象的几何布置具有大体上类似的质量中心的帧可被群集在一起且编索引到存储器中。在此类跟踪期间,可能需要控制显示平移操作,使得显示的中心与所

述质量中心一致。另外或替代地,可能需要控制相机的光学及/或数码变焦操作,使得所有选定对象在显示内保持可见。

[0265] 在特定实施例中,移动装置可跟踪每一对象1702到1706之间的几何形状及/或每一帧的复合几何形状(例如,逐帧地跟踪几何形状)。在其它实施例中,移动装置可以非连续间隔跟踪每一对象1702到1706之间的几何形状及/或复合几何形状(例如,以非连续帧跟踪几何形状)。以非连续帧跟踪几何形状可减少移动装置处消耗的电力的量(例如,延长电池使用时间)并且减少用于将与几何形状相关联的信息编索引及/或将信息集群的存储器量。

[0266] 例如,在特定实施例中,移动装置可跟踪每个帧处的每一对象1702到1706(例如,三个对象)。在跟踪期间,每一对象1702到1706可由与对象1702到1706的位置及对象1702到1706的大小相关联的至少三个值表示。例如,一个值可对应于对象1702到1706的x坐标,一个值可对应于对象1702到1706的y坐标,且另一值可对应于对象1702到1706的大小(例如,限界框)。每一值可对应于32位(例如,4字节)数字。因此,在跟踪期间每一帧至少288位的数据(3对象*3值/对象*32位/值)可被收集及记录到存储器中。假设移动装置每秒捕获30帧,为在一小时(例如,3600秒)内逐帧地跟踪三个对象1702到1706并编索引,至少3,888,000字节的数据被收集及记录到存储器中。

[0267] 然而,以非连续帧跟踪几何形状可放宽存储器要求。作为非限制性实例,在一小时内以每15帧的间隔对每一对象1702到1706进行跟踪及编索引可产生存储器空间中的93.33%节约。以非连续帧跟踪几何形状可在存在与对象1702到1706相关联的相对小量的移动的场景中特别有益。移动装置可基于用户输入(例如,用户偏好)而确定对帧进行跟踪及编索引的频率,及/或可基于(例如)用户指示及/或选定对象中的一或者者中的每一者的位置随时间的改变程度而改变所述频率。

[0268] 每一对象1702到1706还可具有对应于屏幕1730上的对象1702到1706的位置的个别质量中心。例如,第一对象1702可在第一对象1702的位置处一质量中心(C₀₁),第二对象1704可在第二对象1704的位置处具有质量中心(C₀₁)等。每一对象1702到1706的质量中心还可为移动装置跟踪的几何参数。

[0269] 图17还展示描绘第三时间实例(例如,在第二时间实例之后)处捕获的场景的视频流的帧1740。在帧1740中,每一对象1702到1706的位置已相对于描绘场景的帧1720中在第二时间实例处描绘的对应位置而改变。

[0270] 图17还展示移动装置的屏幕上的帧1740的显示1750(例如,通过显示任务T400)。移动装置可在显示1750中包含几何形状的指示。例如,第一对象1702与第二对象1704之间的第一距离的第一指示(L₁₂)可包含在显示1750中,第二对象1704与第三对象1706之间的第二距离的第二指示(L₂₃)可包含在显示1750中,且第一对象1702与第三对象1706之间的第三距离的第三指示(L₁₃)可包含在显示1750中。另外,移动装置还可跟踪对象1702到1706的复合几何形状。例如,在所说明的实施例中,对象1702到1706的复合几何形状可对应于由三个指示(L₁₂、L₂₃、L₁₃)形成的三角形。

[0271] 在另一实施例中,移动装置可跟踪每一对象1702到1706的移动且产生显示每一对象1702到1706随时间的移动的一连串曲线图。移动装置可在每个帧处产生每一对象1702到1706的曲线图,或可在非连续帧处产生每一对象1702到1706的曲线图以放宽存储器要求并提高电力节约,如上文所描述。

[0272] 还可能配置装置(例如,实施任务T250)以跟踪一或多个选定对象与当前在显示上不可见的对象(例如,当前不在由视频流捕获的场景内)的空间关系。例如,可跟踪移动对象与当前在屏幕外的固定(参考)对象的空间关系。在特定实例中,装置可经配置以每当选定运动选手在目标或筐的二十英尺内及/或朝向目标/筐移动时进行指示,即使目标/筐当前不在屏幕上也如此。在此情况下,装置可经配置以产生大于当前显示的场景的图。对于相机的位置是已知(例如,由位置传感器(例如GPS传感器)指示)的情况,相机的一或多个定向传感器可用于确定当前视图如何与更大的场景图相关。所述一或多个定向传感器可包含一或多个惯性传感器(例如,陀螺仪及/或加速度计),及/或一或多个磁场传感器(例如,磁力计),以指示相机相对于重力轴(例如,穿过装置及地球中心的轴)的定向。另外或替代地,可将固定参考对象的可见特征(例如,体育场或球场上的标记,或表演舞台的现场设备)用作参考点。移动的相机也可完成此类映射,但参考对象的外观可随着观看的角度改变而改变。

[0273] 图17中描绘的实施例可使得移动装置能够跟踪场景1700、1720、1740中的对象1702到1706的几何形状,并且将具有大体上类似的几何形状的帧群集以用于检索,如下文相对于图18所解释。移动装置可使用单个相机或使用多个相机执行相对于图17描述的技术。

[0274] 如本文中所描述的用于显示分割和/或对象跟踪的用途及应用可包含以下各者中的任一者:体育活动(例如,英式足球、篮球、足球、曲棍球)、社交事件(例如,婚礼、聚会、舞会)、艺术表演(例如,演出、校园演出、音乐会、二重奏),及安全或监控监视。移动装置还可为头戴式显示器,其中通过使用(例如)操纵杆、手势辨识相机系统或配备有触摸及/或定向传感器(例如,如上文所描述的一或多个定向传感器)的手套来检测用户命令而执行用户选择。视频流可为实况(例如,例如经由短程连接(例如蓝牙),或经由网络(例如,到因特网的Wi-Fi连接由装置捕获或从另一捕获装置串流传输)或经记录的(同样,由装置或由另一装置记录)。

[0275] 可能需要根据流内的选定对象的相对位置及/或之间的空间关系而对视频流的帧进行分类。参考图18,展示用于基于集群而检索帧的移动装置的屏幕上的显示1800的特定实施例。显示1800可包含具有水平轴线及垂直轴线的图表。水平轴线可对应于由移动装置捕获的不同视频文件。例如,在说明性实施例中,移动装置捕获第一视频文件及第二视频文件。垂直轴线可对应于视频文件的时间索引。

[0276] 每一视频文件可包含基于与视频文件中的选定对象的几何形状相关联的至少一个参数的帧的群集。例如,视频文件中的每一群集可包含在选定对象之间具有大体上类似的几何形状的帧的群组。图18中描绘的群集可基于对象之间的质量中心;然而,在其它实施例中,群集可基于一或多个不同参数(例如,形状、距离、空间特性、色彩方案等)。所述参数可由用户选择。

[0277] 第一视频的开始处的帧直到第一视频的约10:00标记可被分组为第一群集(例如,群集1)。第一群集可对应于在像素坐标空间中的选定对象之间的质量中心借以具有约150处的x坐标及约250处的y坐标的帧。参考图17中的第二帧,选定对象1702到1706的质量中心(C_{M2}) (例如,三角形的质量中心(C_{M2}))可具有显示1730上的约150处的x坐标及显示1730上的约250处的y坐标。因此,第二帧(及具有对象1702到1706之间的大体上类似的质量中心的其

它帧)可放置在图18的第一群集中。因此,如果用户选择第一群集,那么移动装置可显示其中选定对象1702到1706与图17中的第二帧具有大体上类似的配置的对应的视频流(例如,视频的部分)。

[0278] 从第一视频的约10:00标记直到第一视频的约20:00标记的帧可被分组为第二群集(例如,群集2)。第二群集可对应于在像素坐标空间中的选定对象之间的质量中心借以具有约200处的x坐标及约250处的y坐标的帧。参考图3中的第三帧,选定对象1702到1706的质量中心(C_{M3})可具有显示1750上的约200处的x坐标及显示1750上的约250处的y坐标。因此,第三帧(及具有对象1702到1706之间的大体上类似的质量中心的其它帧)可放置在图18的第二群集中。因此,如果用户选择第二群集,那么移动装置可显示其中选定对象1702到1706与图18中的第三帧具有大体上类似的配置的对应的视频流(例如,视频的部分)。

[0279] 额外群集(例如,第三群集)可对应于具有不同质量中心(例如,175处的x坐标及325处的y坐标)的帧。用户可通过选择具有与视频的特定部分相关联的配置(例如,质量中心)的群集而检索(例如,定位)视频的特定部分。因此,通过将帧编索引到基于选定对象1702到1706的几何参数(例如,每一及/或所有选定对象的距离、质量中心)中,移动装置的用户可容易定位视频的所要的部分而不必导航(例如,播放、快进、快退等)穿过所述视频。将帧编索引到群集中还可准许所述查看哪一配置(例如,几何形状)出现最频繁、最不频繁等。所述几何参数的所述编索引及/或值可包含在由任务T300和/或T350存储的元数据中,如上文所描述。

[0280] 图18还包含与视频流的帧的群集相关联的区的显示1802。例如,区的显示1802可对应于图17中的场景的视频流。显示1802包含十个区(C1到C10)。在所说明的实施例中,区(C1到C10)是基于特定帧处的对象1702到1706的质量中心。在其它实施例中,区(C1到C10)可基于可用以将帧编索引到群集中的其它几何参数(例如,每一对象1702到1706的质量中心、形状、距离、空间特性、色彩方案等)。在特定实施例中,每一区(C1到C10)可具有相等大小(例如,每一区(C1到C10)可为“沃罗诺伊单元”)。

[0281] 每一区(C1到C10)可包含对应于与对应群集相关联的质量中心的x坐标及y坐标。作为非限制性实例,对应于第一区(C1)的质量中心可具有x坐标75及y坐标580。另外,对应于第二区(C2)的质量中心可具有x坐标215及y坐标580。图17中的场景的视频流中的帧可通过对象1702到1706的质量中心编索引。每一帧可基于质量中心放置在对应于区(C1到C10)中的一者的群集中。例如,具有最接近第一区(C1)的质量中心的质量中心的帧可放置在对应第一群集内。此集群信息可作为元数据(例如,同步到视频流的元数据流中,如本文中所描述)与对应帧相关联。在此情况下,可容易通过搜索元数据而识别在特定群集内编索引的帧以进行检索,因此避免检视整个视频流的需要。

[0282] 移动装置可响应于对特定群集的用户选择而显示具有特定几何形状(例如,质量中心)的视频帧。例如,如果用户选择(例如,触摸)第一区(C1),那么移动装置可显示其中对象1702到1706的质量中心在第一(C1)区内的视频帧的群集。在特定实施例中,用户可触摸第一区(C1)处的屏幕且在超过阈值(例如,三秒)的时间周期内保持第一区(C1)。在保持第一区三秒之后,第一区(C1)可在显示1802上放大。例如,显示1802可描绘如1804中所展示的第一区(C1)。在此模式中,第一区1804可基于特定帧(例如,帧1023到帧1026)的质量中心($C_{M1023到1026}$)说明所述特定帧。在特定实施例中,用户可选择特定帧(例如,帧1024)且移动装

置可在帧1024处起始视频流的回放。

[0283] 参考图19,展示说明使用对象跟踪的视频处理方法1900的特定实施例的流程图。在说明性实施例中,可使用图1的电子装置102、由图2B的处理器264实施的对象跟踪与检测模块204、由处理器实施的图8的对象跟踪与检测模块804、由处理器实施的图9的平滑模块961或其任何组合执行方法1900。

[0284] 方法1900包含在1902处在移动装置处接收对场景中的多个对象的选择。例如,参考图17,用户可触摸屏幕上的用户选择机构1711且围绕第一对象1702拖动第一聚焦环1712,围绕第二对象1704拖动第二聚焦环1714,且围绕第三对象1706拖动第三聚焦环1716以分别启用对第一、第二和第三对象1702到1706的跟踪。所显示的场景可对应于移动装置的屏幕上捕获及显示的视频流。移动装置可经配置以捕获所述流(例如,使用移动装置的相机)或从另一装置接收所述流。

[0285] 在1904处,可跟踪每一对象之间的几何形状。例如,参考图17,移动装置可跟踪每一选定对象1702到1706之间的几何形状。例如,移动装置可跟踪并且测量第一对象1702与第二对象1704之间的第一距离、第二对象1704与第三对象1706之间的第二距离,及第三对象1706与第一对象1702之间的第三距离。

[0286] 在1906处,可在屏幕上显示几何形状的指示。例如,参考图17,可在显示1730中包含第一对象1702同样第二对象1704之间的第一距离的第一指示(L_{12})。可基于帧1720中的第一对象1702的位置及帧1720中的第二对象1704的位置而确定第一距离。例如,移动装置可跟踪第一对象1702的位置及第二对象1704的位置以产生第一指示(L_{12})。还可在显示1730中包含第二对象1704与第三对象1706之间的第二距离的第二指示(L_{23})。可基于帧1720中的第二对象1704的位置及帧1720中的第三对象1706的位置而确定第二距离。例如,移动装置可跟踪第二对象1704的位置及第三对象1706的位置以产生第二指示(L_{23})。还可在显示1730中包含第一对象1702与第三对象1706之间的第三距离的第三指示(L_{13})。可基于帧1720中的第一对象1702的位置及帧1720中的第三对象1706的位置而确定第三距离。例如,移动装置可跟踪第一对象1702的位置及第二对象1704的位置以产生所述第三指示(L_{13})。

[0287] 图19的方法1900可使得移动装置能够跟踪帧1700、1720、1740中的对象1702到1706的几何形状且将具有大体上类似的几何形状的帧群集以用于检索。移动装置可使用单个相机或使用多个相机执行方法1900。在另一实例中,所述方法可包含在检测到一或多个对象之间的指定空间关系时,例如在第一选定对象与第二选定对象之间的距离变得小于(或者,大于)指定阈值距离时,指示警报条件。

[0288] 参考图20,展示说明使用对象跟踪的视频处理方法2000、2010的特定实施例。在说明性实施例中,可使用图1的电子装置102、由图2B的处理器264实施的对象跟踪与检测模块204、由处理器实施的图8的对象跟踪与检测模块804、由处理器实施的图9的平滑模块961或其任何组合执行方法2000、2010。

[0289] 方法2000包含在2002处在移动装置处接收对场景中的多个对象的选择。例如,参考图17,用户可触摸屏幕上的用户选择机构1711且围绕第一对象1702拖动第一聚焦环1712,围绕第二对象1704拖动第二聚焦环1714,且围绕第三对象1706拖动第三聚焦环1716以分别启用对第一、第二和第三对象1702到1706的跟踪。所显示的场景1700可对应于移动装置的屏幕上捕获及显示的视频流。移动装置可经配置以捕获所述流(例如,使用移动装置

的相机)或从另一装置接收所述流。

[0290] 在2004处,可跟踪每一对象之间的几何形状。例如,参考图17,移动装置可跟踪每一选定对象1702到1706之间的几何形状。例如,移动装置可跟踪并测量第一对象1702与第二对象1704之间的第一距离、第二对象1704与第三对象1706之间的第二距离,及第三对象1706与第一对象1702之间的第三距离。另外,移动装置还可跟踪对象1702到1706的复合几何形状。例如,在所说明的实施例中,对象1702到1706的复合几何形状可对应于由三个指示(L_{12}, L_{23}, L_{13})形成的三角形。

[0291] 在2006处,可基于与几何形状相关联的至少一个参数而将视频流的帧集群。例如,参考图18,视频文件中的每一群集可包含在选定对象1706到1708之间具有大体上类似的几何形状的帧群组。图18中描绘的群集可基于对象1706到1708之间的质量中心(C_{M2}, C_{M3}) (例如,三角形的质量中心)。

[0292] 在特定实施例中,方法2000可包含在对象处于特定几何形状时检索帧。例如,参考图18,移动装置可基于帧中的对象1702到1706的质量中心(C_{M2}, C_{M3})而将帧编索引。可容易通过选择与特定质量中心相关联的群集而实现具有特定质量中心(例如,特定几何形状)的帧。例如,移动装置可通过选择图18的显示1800中所示的第一群集(例如,群集1)而检索其中对象1702到1706与图17的第二帧中的对象1702到1706具有大体上类似的几何形状的帧。例如,响应于用户选择第一群集,移动装置可检索对象1702到1706的质量中心借以具有约150处的x坐标及约250处的y坐标的帧。

[0293] 方法2010可包含在2012处在移动装置处接收视频流中的选定对象的特定几何形状的指示。作为非限制性实例,参考图18,移动装置可接收显示其中对象1702到1704的质量中心最接近第一区(C1)的质量的中心的指示(例如,用户触摸显示1802中展示的第一区(C1)处的屏幕)的视频流。

[0294] 在2014处,可基于特定几何形状而检索视频流的帧。例如,参考图18,移动装置可检索图17中的场景中的视频流的其中对象1702到1704的质量中心在对应于第一区(C1)的区域中的帧。

[0295] 在2016处,可在移动装置的屏幕上显示所检索的帧。例如,参考图18,移动装置可显示(例如,播放)视频流的其中对象1702到1704的质量中心在对应于第一区(C1)的区域中的帧。在特定实施例中,移动装置可以顺序次序播放所述视频流。

[0296] 图20的方法2000、2010可使得用户能够通过选择具有与视频的特定部分相关联的配置(例如,质量中心)的群集而检索(例如,定位)视频的所述特定部分。因此,通过将帧编索引到基于选定对象1702到1706的几何参数的群集中,移动装置的用户可容易定位视频的要的部分,而不必导航(例如,播放、快进、快退等)穿过所述视频。

[0297] 图21展示方法M100的实施方案(例如,方法M110、M200、M300或M400)的应用的实例。相对于图21中所描绘的实施例描述的视频处理技术可由图1的电子装置102、由图2B的处理器264实施的对象跟踪和检测模块204、由处理器实施的图8的对象跟踪和检测模块804、由处理器实施的图9的平滑模块961或其任何组合执行。

[0298] 图21展示一级视频流的帧2100,其描绘在第一时间实例捕获的场景。举例来说,帧2100可对应于将在移动装置的屏幕上显示的视频流的帧。所述移动装置可经配置以捕获所述流(例如,以移动装置的相机)或从另一装置接收所述流。帧1100中所描绘的场景包含第

一对象2102、第二对象2104、第三对象2106、星和树。在所说明的实施例中，第一对象2102可对应于第一人，第二对象2104可对应于第二人，且第三对象2106可对应于第三人。在另一实施例中，对象2102到2106中的至少一者可对应于经由移动装置控制的机器人。

[0299] 图21还展示移动装置的屏幕上的帧2100的显示2110。所述移动装置的屏幕(例如，触摸屏)可对应于取景器。显示2110可包含用户选择机构2111，其使用户能够选择对象用于跟踪(例如，由跟踪任务T250)。举例来说，用户可触摸屏幕上的用户选择机构2111且围绕第一对象2102拖动第一聚焦环2112以启用对第一对象2102的跟踪。响应于第一聚焦环2112的此放置，所述移动装置可选择第一对象2112用于跟踪。以类似方式，用户还可触摸屏幕上的用户选择机构2111且围绕第二对象2104和第三对象2106拖动第二聚焦环2114和第三聚焦环2116以分别启用对第二对象2104和第三对象2106的跟踪。响应于第二聚焦环2114和第三聚焦环2116的此放置，所述移动装置可分别选择第二对象2104和第三对象2106用于跟踪。

[0300] 图21还展示视频流的帧2120，其描绘在第二时间实例(例如，在第一时间实例之后)捕获的场景。在帧2120中，特定对象2102、2106的大小相对于描绘在第一时间实例的场景的帧2100中的对应大小已经改变。举例来说，第一对象2102已移动更接近相机透镜(例如，第一对象2102的大小表现为相对大)，且第三对象2104已移动更远地相机透镜(例如，第三对象2106的大小表现为相对小)。第二对象2104的大小表现为大于第三对象2106的大小且小于第一对象2102的大小。

[0301] 图21还展示移动装置的屏幕上的帧2120的显示2130。显示2130可分割(例如，通过显示任务T400)为第一窗口2132、第二窗口2134和第三窗口2136。举例来说，响应于以用户选择机构2111选择三个对象2102到2106，所述移动装置可将显示2130分割为三个窗口2132到2136。第一窗口2132可大于第二窗口2134，且第二窗口2134可大于第三窗口2136。帧2120在每一窗口2132到2136中显示的部分可基于相对于显示2130的大小的对象2102到2106的大小。

[0302] 如果相对于显示2130的大小的第一对象2102的大小大于相对于显示2130的大小的其它选定对象2104、2106的大小，那么第一窗口2132(例如，最大窗口)可显示帧2120的包含第一对象2102的第一部分。帧2120的第一部分可基于帧2120中的第一对象2102的位置而确定。举例来说，所述移动装置可跟踪第一对象2102的位置，且聚焦于第一对象2102的所述位置上的视频流可在第一窗口2132中显示。

[0303] 如果相对于显示2130的大小的第二对象2104的大小大于相对于显示2130的大小的第三对象2106的大小，那么第二窗口2134可显示帧2120的包含第二对象2104的第二部分。帧2120的第二部分可基于帧2120中的第二对象2104的位置而确定。举例来说，所述移动装置可跟踪第二对象2104的位置，且聚焦于第二对象2104的所述位置上的视频流可在第二窗口2134中显示。

[0304] 第三窗口2136(例如，最小窗口)可显示帧2120的包含第三对象2106(例如，相对于显示2130的大小的最小对象)的第三部分。帧2120的第三部分可基于帧2120中的第三对象2106的位置而确定。举例来说，所述移动装置可跟踪第三对象2104的位置，且聚焦于第三对象2106的所述位置上的视频流可在第三窗口2136中显示。

[0305] 图21还展示视频流的帧2140，其描绘在第三时间实例(例如，在第二时间实例之后)捕获的场景。在帧2140中，对象2102到2106的大小相对于描绘在第二时间实例的场景的

帧2120中所描绘的对应大小已经改变。举例来说,第一对象2102已移动更远离相机透镜(例如,第一对象2102的大小表现为相对小),第二对象已移动更接近相机透镜(例如,第一对象2102的大小表现为相对大),且第三对象2104也已移动更接近相机透镜(例如,第三对象2106的大小表现为大于第一对象2102的大小且小于第二对象2104的大小)。

[0306] 图21还展示移动装置的屏幕上的帧2140的显示2150,其包含由所述移动装置产生的窗口2132到2136。举例来说,第一窗口2132(例如,最大窗口)显示帧2140的包含第二对象2104的第二部分,第二窗口2134显示帧2140的包含第三对象2106的第三部分,且第三窗口2136(例如,最小窗口)显示帧2140的包含第一对象2102的第一部分。

[0307] 图21中所描绘的实施例可使移动装置能够跟踪帧2100、2120、2140中的对象2102到2106,且可基于对象2102到2106的相对大小在单独窗口2132到2136中显示对应于对象2102到2106的个别化(例如,聚焦)视频流。举例来说,聚焦于表现为大于其它对象2102到2106(例如,表现为占据屏幕的最大部分)的对象2102到2106上的视频流可在最大窗口(例如,第一窗口2132)中显示。在相似实施方案中,可实施显示任务T400以响应于对应于每一二级视频流的选定对象在一级视频流内的相对大小(例如,对象的界限框的相对大小)而将二级视频流指派于显示的对应窗口,如跟踪任务T250指示。所述移动装置可使用单个相机或使用多个相机执行相对于图21描述的技术。

[0308] 参考图22,展示说明使用对象跟踪的视频处理方法2200的特定实施例的流程图。在说明性实施例中,方法2200可使用图1的电子装置102、由图2B的处理器264实施的对象跟踪和检测模块204、由处理器实施的图8的对象跟踪和检测模块804、由处理器实施的图9的平滑模块961或其任何组合执行。

[0309] 方法2200包含在2202在移动装置处接收对场景中的第一对象和场景中的第二对象的选择。举例来说,参考图21,用户可触摸屏幕上的用户选择机构2111且围绕第一对象2102拖动第一聚焦环2112和围绕第二对象2104拖动第二聚焦环2114以分别启用对第一对象2102和第二对象2104的跟踪。所显示场景可对应于在移动装置的屏幕上捕获和显示的视频流。

[0310] 在2204,如果相对于所述显示的大小的第一对象的大小大于相对于所述显示的大小的第二对象的大小,那么可在第一窗口中显示所显示场景的包含第一对象的第一部分。举例来说,参考图21,如果相对于显示2130的大小的第一对象2102的大小大于相对于显示2130的大小的其它选定对象2104、2106的大小,那么第一窗口2132(例如,最大窗口)可显示帧2120的包含第一对象2102的第一部分。帧2120的第一部分可基于帧2120中的第一对象2102的位置而确定。举例来说,所述移动装置可跟踪第一对象2102的位置,且聚焦于第一对象2102的所述位置上的视频流可在第一窗口2132中显示。

[0311] 在2206,响应于相对于所述显示的大小的第一对象的大小大于相对于所述显示的大小的第二对象的大小的确定,可在第二窗口中显示所显示场景的包含第二对象的第二部分。举例来说,参考图22,如果相对于显示2130的大小的第二对象2104的大小大于相对于显示2130的大小的第三对象2106的大小,那么第二窗口2134可显示帧2120的包含第二对象2104的第二部分。帧2120的第二部分可基于帧2120中的第二对象2104的位置而确定。举例来说,所述移动装置可跟踪第二对象2104的位置,且聚焦于第二对象2104的所述位置上的视频流可在第二窗口2134中显示。第一窗口2132可大于第二窗口2134。

[0312] 在2208,如果相对于所述显示的大小的第一对象的大小不大于相对于所述显示的大小的第二对象的大小,那么可在第二窗口中显示所显示场景的第一部分。在2210,如果相对于所述显示的大小的第一对象的大小不大于相对于所述显示的大小的第二对象的大小,那么可在第一窗口中显示所述帧的第二部分。

[0313] 图22的方法2200可使移动装置能够跟踪帧2100、2120、2140中的对象2102到2106,且可基于对象2102到2106的相对大小在单独窗口2132到2136中显示对应于对象2102到2106的个别化(例如,聚焦)视频流。举例来说,聚焦于表现为大于其它对象2102到2106(例如,表现为占据屏幕的最大部分)的对象2102到2106上的视频流可在最大窗口(例如,第一窗口2132)中显示。所述移动装置可使用单个相机或使用多个相机执行方法2200。

[0314] 参考图23,展示说明使用对象跟踪的视频处理方法2300的特定实施例的流程图。在说明性实施例中,方法2300可使用机器人执行。

[0315] 方法2300可包含在2302在机器人处跟踪多个对象之间的几何形状。举例来说,用户可向机器人传达待跟踪的特定对象。在一特定实施例中,所述机器人可包含显示接口,且用户可以经由与所述显示接口的交互向机器人通信。作为另一实例,所述机器人可独立地基于环境因素而确定将跟踪何种对象。举例来说,所述机器人可包含用以检测运动的传感器。基于所检测到的运动,机器人可选择跟踪与所述运动相关联的对象。

[0316] 在2304,可基于几何形状执行处理功能。举例来说,所述机器人可基于几何形状确定对象之间的关系。作为另一实例,所述机器人可基于对象的几何形状确定设定(例如,场景)。

[0317] 图24展示方法M100的实施方案(例如,方法M110、M200、M300或M400)的应用的实例。相对于图24中所描绘的实施例描述的视频处理技术可例如由图1的电子装置102、由图2B的处理器264实施的对象跟踪和检测模块204、由处理器实施的图8的对象跟踪和检测模块804、由处理器实施的图9的平滑模块961或其任何组合执行。

[0318] 图24展示一级视频流的帧2400,其描绘在第一时间实例捕获的场景。举例来说,帧2400可对应于将在移动装置的屏幕上显示的视频流的帧。所述移动装置可经配置以捕获所述流(例如,以移动装置的相机)或从另一装置接收所述流。帧2400中所描绘的场景包含第一对象2402、第二对象2404、第三对象2406、星和树。在所说明的实施例中,第一对象2402可对应于第一人,第二对象2404可对应于第二人,且第三对象2406可对应于第三人。在另一实施例中,对象2402到2406中的至少一者可对应于经由移动装置控制的机器人。

[0319] 图24还展示移动装置的屏幕上的帧2400的显示2410。所述移动装置的屏幕(例如,触摸屏)可对应于取景器。显示2410可包含用户选择机构2411,其使用户能够选择对象用于跟踪(例如,由跟踪任务T250)。举例来说,用户可触摸屏幕上的用户选择机构2411且围绕第一对象2402拖动第一聚焦环2412以启用对第一对象2402的跟踪。响应于第一聚焦环2412的此放置,所述移动装置可选择第一对象2402用于跟踪。以类似方式,用户还可触摸屏幕上的用户选择机构2411且围绕第二对象2404和第三对象2406拖动第二聚焦环2414和第三聚焦环2416以分别启用对第二对象2404和第三对象2406的跟踪。响应于第二聚焦环2414和第三聚焦环2416的此放置,所述移动装置可分别选择第二对象2404和第三对象2406用于跟踪。

[0320] 图24还展示视频流的帧2420,其描绘在第二时间实例(例如,在第一时间实例之后)捕获的场景。在帧2420中,每一对象2402到2406的位置相对于描绘在第一时间实例的场

景的帧2400中的对应位置已经改变。举例来说,第一对象2402已移动更接近星,第二对象2404已移动更接近树,且第三对象2406已移动更接近帧的底部。

[0321] 图24还展示移动装置的屏幕上的帧2420的显示2430。显示2430可分割(例如,由显示任务T400)为两个或更多个窗口。在此实例中,显示2430分割成第一窗口2432、第二窗口2434、第三窗口2436和第四窗口2438。响应于例如以用户选择机构2411对三个对象2402到2406的选择所述移动装置可将显示2430分割为三个对应窗口2432到2436和第四窗口2438。

[0322] 第一窗口2432可显示帧2420的包含第一对象2402的第一部分。帧2420的第一部分可基于帧2420中的第一对象2402的位置而确定。举例来说,所述移动装置可跟踪第一对象2402的位置,且聚焦于第一对象2402的所述位置上的视频流可在第一窗口2432中显示。第二窗口2434可显示帧2420的包含第二对象2404的第二部分。帧2420的第二部分可基于帧2420中的第二对象2404的位置而确定。举例来说,所述移动装置可跟踪第二对象2404的位置,且聚焦于第二对象2404的所述位置上的视频流可在第二窗口2434中显示。第三窗口2436可显示帧2420的包含第三对象2406的第三部分。帧2420的第三部分可基于帧2420中的第三对象2406的位置而确定。举例来说,所述移动装置可跟踪第三对象2404的位置,且聚焦于第三对象2406的位置上的视频流可在第三窗口2436中显示。

[0323] 第四窗口2438可显示每一对象2402到2406之间的几何形状的指示(例如,质量中心(C_{M2}))。例如,第四窗口2438可包含网格,所述网格包含对应于帧2420中的区域的多个区段(表示为灰度级线)。例如,网格的左下区段可对应于帧2420的左下区段。以与相对于图18及19所描述的方式类似的方式,移动装置可确定帧2420中的对象的质量中心(C_{M2})。可在网格的对应区段处显示质量中心(C_{M2})。

[0324] 图24还展示帧2440,其描绘在第三时间实例(例如,在第二时间实例之后)的场景。在帧2440中,每一对象2402到2406的位置相对于描绘在第二时间实例的场景的帧2420中的对应位置已经改变。举例来说,第一对象2402已移动更接近树,第二对象已移动更接近场景2440的底部,且第三对象2406已移动更接近星。

[0325] 图24还展示使用由移动装置产生的窗口2432到2436在所述移动装置的屏幕上的帧2440的显示2450。举例来说,第一窗口2432显示帧2440的包含第一对象2402的第一部分,第二窗口2434显示帧2440的包含第二对象2404的第二部分,且第三窗口2436显示帧2440的包含第三对象2406的第三部分。

[0326] 第四窗口2438可更新指示(例如,质量中心(C_{M2}))以反映在第三时间实例的帧2440。移动装置可以逐帧为基础或在非连续时间间隔跟踪几何形状,且在第四窗口2438中显示经更新几何形状的指示。举例来说,在第三帧,第四窗口2438可显示在第三时间实例的帧2440中的对象2402到2406的质量中心(C_{M3})。第四窗口2438还可显示在第二帧与第三帧之间跟踪对象的质量中心(例如,如由跟踪任务T250指示)的路径(例如,点线)。

[0327] 图24中描绘的实施例可使得移动装置能够在帧2400、2420、2440中跟踪对象2402到2406,且可在单独窗口2432到2436中显示对应于对象2402到2406的个别化(例如,聚焦)视频流。所述实施例还使得用户能够在第四窗口2438中查看对象2402到2406的几何形状。移动装置可使用单个相机或使用多个相机执行相对于图24描述的技术。

[0328] 参考图25,展示说明使用对象跟踪的视频处理方法2500的特定实施例的流程图。在说明性实施例中,方法2500可使用图1的电子装置102、由图2B的处理器264实施的对象跟

踪和检测模块204、由处理器实施的图8的对象跟踪和检测模块804、由处理器实施的图9的平滑模块961或其任何组合执行。

[0329] 方法2500包含在2502在移动装置处接收对所显示场景中的第一对象、所显示场景中的第二对象和所显示场景中的第三对象的选择。举例来说，参考图24，用户可触摸屏幕上的用户选择机构2411且围绕第一对象2402拖动第一聚焦环2412，围绕第二对象2404拖动第二聚焦环2414，且围绕第三对象2406拖动第三聚焦环2416，以分别启用对第一、第二和第三对象2402到2406的跟踪。所显示场景可对应于在移动装置的屏幕上捕获和显示的视频流。

[0330] 在2504处，可将显示分割成第一窗口、第二窗口、第三窗口及第四窗口。例如，参考图24，可将显示2430分割成第一窗口2432、第二窗口2434、第三窗口2436及第四窗口。

[0331] 在2506，可在第一窗口中显示所显示场景的包含第一对象的第一部分。举例来说，参考图24，第一窗口2432可显示帧2420的包含第一对象2402的第一部分。帧2420的第一部分可基于帧2420中的第一对象2402的位置而确定。举例来说，所述移动装置可跟踪第一对象2402的位置，且聚焦于第一对象2402的所述位置上的视频流可在第一窗口2432中显示。

[0332] 在2508，可在第二窗口中显示所显示场景的包含第二对象的第二部分。举例来说，参考图24，第二窗口2434可显示帧2420的包含第二对象2404的第二部分。帧2420的第二部分可基于帧2420中的第二对象2404的位置而确定。举例来说，所述移动装置可跟踪第二对象2404的位置，且聚焦于第二对象2404的所述位置上的视频流可在第二窗口2434中显示。

[0333] 在2510，可在第三窗口中显示所显示场景的包含第三对象的第三部分。举例来说，参考图24，第三窗口2436可显示帧2420的包含第三对象2406的第三部分。帧2420的第三部分可基于帧2420中的第三对象2406的位置而确定。举例来说，所述移动装置可跟踪第三对象2406的位置，且聚焦于第三对象2406的所述位置上的视频流可在第三窗口2436中显示。

[0334] 在2412，可在第四窗口中显示每一对象之间的几何形状的指示。举例来说，参考图24，第四窗口2438可显示每一对象2402到2406之间的质量中心(C_{M2})。第四窗口2438可更新所述质量中心(C_{M2})以反映在第三时间实例的场景2440。举例来说，在第三帧，第四窗口2438可显示在第三时间实例的帧2440中的对象2402到2406的质量中心(C_{M3})。第四窗口2438还可显示在第二帧与第三帧之间跟踪对象的质量中心的路径(例如，点线)。

[0335] 图25的方法2500可使移动装置能够跟踪场景2400、2420、2440中的对象2402到2406且可在单独窗口2432到2436中显示对应于对象2402到2406的个别化(例如，聚焦)视频流。方法2500也使用户能够在单独窗口(例如，第四窗口2438)中查看对象2402到2406的几何形状。所述移动装置可使用单个相机或使用多个相机执行方法2500。

[0336] 图37B展示根据一般配置的设备A100的框图。设备A100包含鉴别器200，其经配置以选择具有一级视野的一级视频流PS10内的至少两个对象(例如，如本文中参考选择任务T200所描述)。设备A100还包含视频流产生器200，其经配置以响应于所述选择从一级视频流产生包含所述选定对象中的第一者且具有比一级视野更窄的第一视野的第一视频流VS10，且在所述选择之后从一级视频流产生包含所述选定对象中的第二者且具有比一级视野更窄的第二视野的第二视频流VS20(例如，如本文中参考流产生任务T300和T350所描述)。在所述产生第一视频流和所述产生第二视频流开始之后，所述第一视野包含一级视野的不在所述第二视野内的一部分，且所述第二视野包含一级视野的不在所述第一视野内的一部分。

[0337] 图37C展示设备A100的实施方案A110的框图。设备A110包含接口110,其经配置以剖析一级视频流PS10以产生一级视频流PS10的一系列帧(例如,如本文中参考剖析任务T110所描述)。接口110可包含经布置以存储一或多个经剖析帧的缓冲器,例如如本文中所描述的帧缓冲器或循环缓冲器(例如,帧缓冲器FB10,存储器缓冲器110)。图38A展示设备A100的实施方案A200的框图。设备A200包含对象跟踪器250,其经配置以跟踪选定对象(例如,如本文中参考跟踪任务T250、对象跟踪和检测模块204和/或对象跟踪和检测模块804所描述)。举例来说,对象跟踪器250可经配置以检测在显示视频流的帧中第一和第二对象当中的至少一个靠近所述帧的边缘;确定在一级视频流的多个帧中的第一者中第一对象比第二对象更接近所述帧的第一边缘;和/或确定在所述在第一窗口内显示第一视频流之后,在(A)一级视频流的帧、(B)第一视频流的帧和(C)第二视频流的帧中的至少一者中第二对象与第一对象重叠。

[0338] 图38B展示设备A100的实施方案A300的框图。设备A300包含显示信号产生器400,其经配置以基于第一和第二视频流中的至少一者产生显示信号(例如,如本文中参考显示任务T400所描述)。举例来说,显示信号产生器400可经配置以产生包含第一窗口内的第一视频流和第二窗口内的第二视频流的显示信号(例如,到屏幕)。另外或替代地,显示信号产生器400可经配置以产生展示几何形状和/或群集区的指示的显示信号(例如,如本文中参考图17、18和24所描述)。显示信号产生器400可包含如本文中所描述的显示缓冲器(例如,显示缓冲器DB10)。图39A展示设备A200和A300的实施方案A400的框图。

[0339] 本文所揭示的设备A100的实施方案中的任一者可实施于移动装置内,例如以下形状因数的任一者的装置:可抓握的东西(例如,智能电话)、可驾驶的东西(例如,车辆或机器人)、可穿戴的东西(例如,服装或配饰),及可飞行的东西(例如,遥控飞机)。移动装置可包含经配置以显示流DS10的一或多个屏幕SC10(例如,触摸屏)和/或经配置以产生一级视频流PS10的一或多个成像单元IU10(例如,相机或其它图像捕获装置)。图39B和39C展示设备A100的此些应用的实例的框图,且图39D展示设备A300的此应用的框图。

[0340] 参考图26,展示包含可操作以执行上述视频处理技术的组件的无线装置2600的框图。装置2600包含耦合到存储器2632的处理器2610,例如数字信号处理器(DSP)。装置2600是可实施以包含如本文中所描述的设备A100的实施方案中的任一者的元件的移动装置的一个实例。

[0341] 图26还展示耦合到处理器2610和显示器2628的显示控制器2626(例如,经配置以执行如本文中所描述的显示任务T400)。在一特定实施例中,显示器可对应于上述取景器或屏幕。存储器2632可为包含可执行指令2656的有形的非暂时性处理器可读存储媒体。指令2656可由处理器执行,例如处理器2610。例如处理器2610或显示控制器2626等处理器可执行指令2656以执行以下方法的至少一部分:图3的方法300、图4的方法400、图5的方法500、图6的方法600、图10的方法1000、图12的方法1200、图14的方法1400、图16的方法1600、图19的方法1900、图20的方法2000、2010、图22的方法2200、图25的方法2500、方法M100、方法M110、方法M120、方法M200、方法M300、方法M400或其任何组合。

[0342] 译码器/解码器(编解码器)2634也可耦合到处理器2610。扬声器2636和麦克风2638可耦合到编解码器2634。图26还指示无线控制器2640可耦合到处理器2610和天线2642。在一特定实施例中,射频(RF)接口2680安置于无线控制器2640与天线2642之间。

[0343] 在一特定实施例中,处理器2610、显示控制器2626、存储器2632、编解码器2634和无线控制器2640包含在系统级封装或芯片上系统装置2622中。在一特定实施例中,输入装置2630(例如,如上文所描述用以捕获场景的相机)和电力供应器2644耦合到芯片上系统装置2622。此外,在一特定实施例中,如图26中所说明,显示器2628、输入装置2630、扬声器2636、麦克风2638、天线2642、电力供应器2644和RF接口2680在芯片上系统装置2622的外部。举例来说,RF接口2680可为单独芯片。然而,显示器2628、输入装置2630、扬声器2636、麦克风2638、天线2642、电力供应器2644和RF接口2680中的每一者可耦合到芯片上系统装置2622的组件,例如接口或控制器。

[0344] 图39E展示根据一般配置的设备MF100的框图。设备MF100包含用于选择具有一级视野的一级视频流内的至少两个对象的装置F200(例如,如本文中参考任务T200所描述)。设备MF100还包含用于响应于所述选择从所述一级视频流产生包含选定对象中的第一者且具有比一级视野更窄的第一视野的第一视频流的装置F300(例如,如本文中参考任务T300所描述)。设备MF100还包含用于在所述选择之后从一级视频流产生包含选定对象中的第二者且具有比一级视野更窄的第二视野的第二视频流的装置F350(例如,如本文中参考任务T350所描述)。在产生第一视频流和所述产生第二视频流开始之后,所述第一视野包含一级视野的不在所述第二视野内的一部分,且所述第二视野包含一级视野的不在所述第一视野内的一部分。

[0345] 图40A展示设备MF100的实施方案MF110的框图。设备MF110包含用于剖析一级视频流以产生一级视频流的一系列帧的装置F100(例如,如本文中参考剖析任务T100所描述)。

[0346] 图40B展示设备MF100的实施方案MF200的框图。设备MF200包含用于跟踪对象的装置F250(例如,如本文中参考跟踪任务T250所描述)。举例来说,装置F250可实施为用于检测在显示视频流的帧中第一和第二对象当中的至少一个靠近所述帧的边缘的装置;用于确定在一级视频流的多个帧中的第一者中第一对象比第二对象更接近所述帧的第一边缘的装置;和/或用于确定在所述在第一窗口内显示第一视频流之后,在(A)一级视频流的帧、(B)第一视频流的帧和(C)第二视频流的帧中的至少一者中第二对象与第一对象重叠的装置。

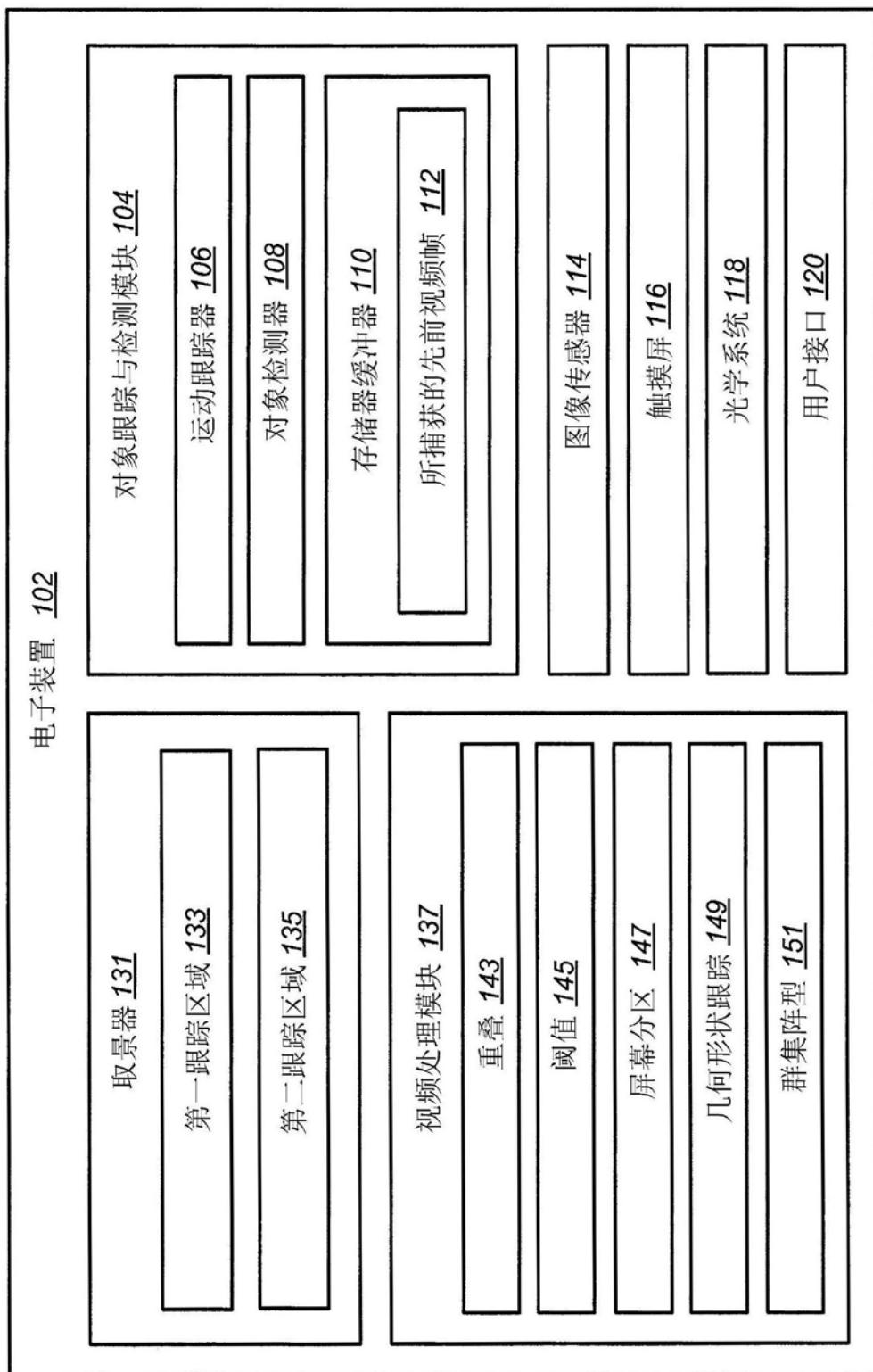
[0347] 图40C展示设备MF100的实施方案MF300的框图。设备MF300包含用于显示第一和第二视频流中的至少一者的装置F400(例如,如本文中参考显示任务T400所描述)。举例来说,装置F400可实施为用于在屏幕上显示第一窗口内的第一视频流和第二窗口内的第二视频流的装置。另外或替代地,装置F400可实施为用于产生展示几何形状和/或群集区的指示的显示信号的装置(例如,如本文中参考图17、18和24所描述)。图40D展示设备MF200和MF300的实施方案MF400的框图。

[0348] 所属领域的技术人员将进一步了解,结合本文中所揭示的实施例所描述的各种说明性逻辑块、配置、模块、电路和算法步骤可实施为电子硬件、由处理器执行的计算机软件,或两者的组合。例如,本文中揭示的设备(例如,设备MF100、MF110、MF200、MF300、MF400、A100、A110、A200、A300及A400中的任一者)的各种实施方案的一或多个元件可整体或部分地实施为经布置以在逻辑元件的一或多个固定或可编程阵列(例如,处理器2610)(例如微处理器、嵌入式处理器、IP核心、数字信号处理器、FPGA(现场可编程门阵列)、ASSP(专用标准产品)及ASIC(专用集成电路))上执行的一或多个指令集。上文已大体在其功能性方面描述各种说明性组件、块、配置、模块、电路和步骤。此功能性是实施为硬件还是处理器可执行

指令取决于特定应用和强加于整个系统的设计约束。熟练的技术人员可针对每一特定应用以不同方式实施所描述的功能性,但此类实施决策不应被解释为引起偏离本发明的范围。

[0349] 结合本文所揭示的实施例而描述的方法或算法的步骤可直接体现在硬件、由处理器执行的软件模块或所述两者的组合中。软件模块可驻留在随机存取存储器 (RAM)、快闪存储器、只读存储器 (ROM)、可编程只读存储器 (PROM)、可擦除可编程只读存储器 (EPROM)、电可擦除可编程只读存储器 (EEPROM)、寄存器、硬盘、可移除式磁盘、压缩光盘只读存储器 (CD-ROM) 或此项技术中已知的任何其它形式的非瞬时存储媒体中。示范性存储媒体耦合到处理器,使得处理器可从存储媒体读取信息并且将信息写入到存储媒体。在替代方案中,存储媒体可集成到处理器。处理器及存储媒体可以驻留在专用集成电路 (ASIC) 中。ASIC 可以驻留在计算装置或用户终端中。在替代方案中,处理器及存储媒体可以作为离散组件驻留在计算装置或用户终端中。

[0350] 提供对所揭示实施例的先前描述以使所属领域的技术人员能够制作或使用所揭示的实施例。所属领域的技术人员将容易对这些实施例的各种修改显而易见,并且可将本文界定的原理应用到其它实施例而不偏离本发明的范围。因此,本发明并不既定限于本文展示的实施例,而应被赋予与如由所附权利要求书界定的原理及新颖特征一致的可能最宽范围。



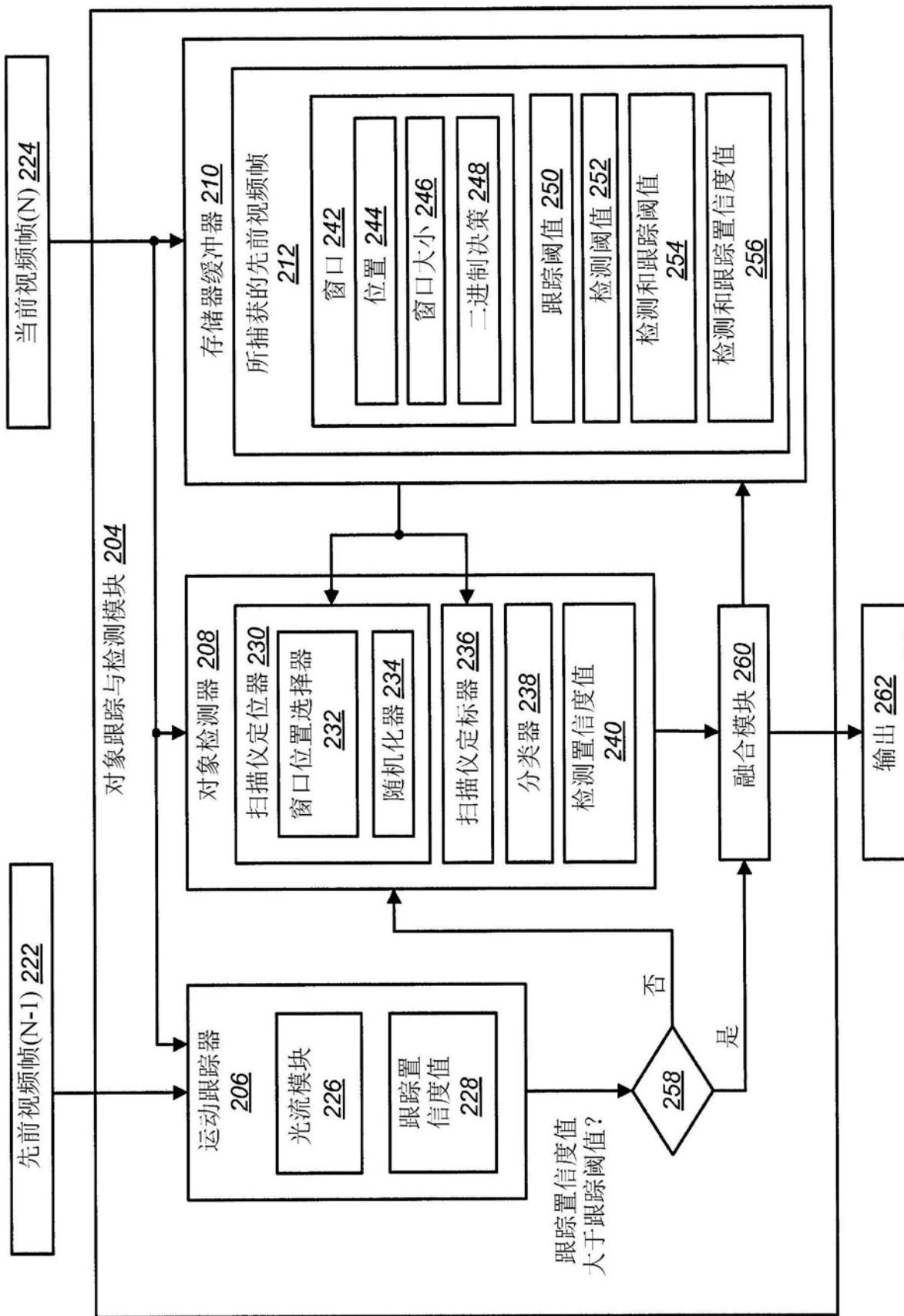


图2A

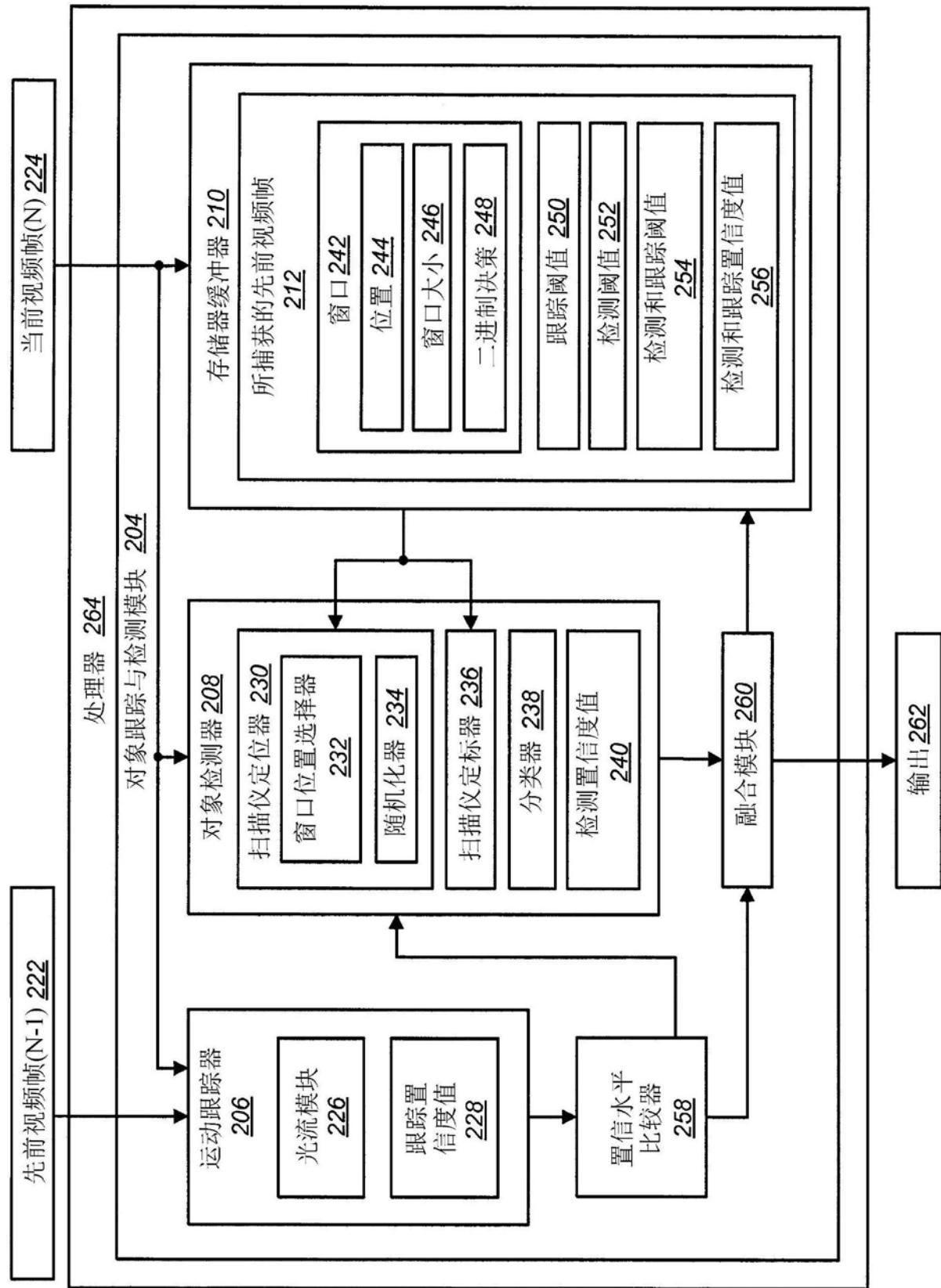


图2B

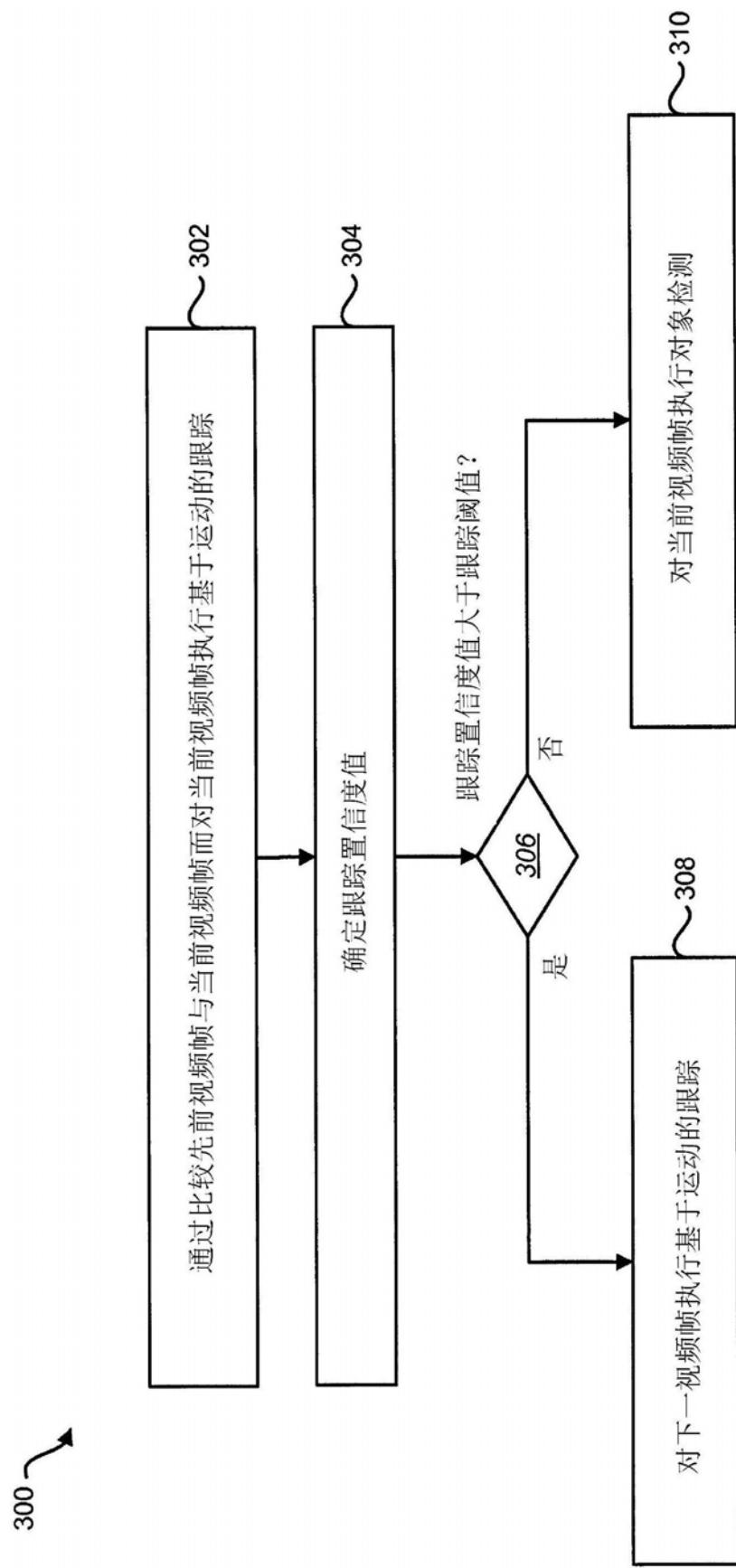


图3

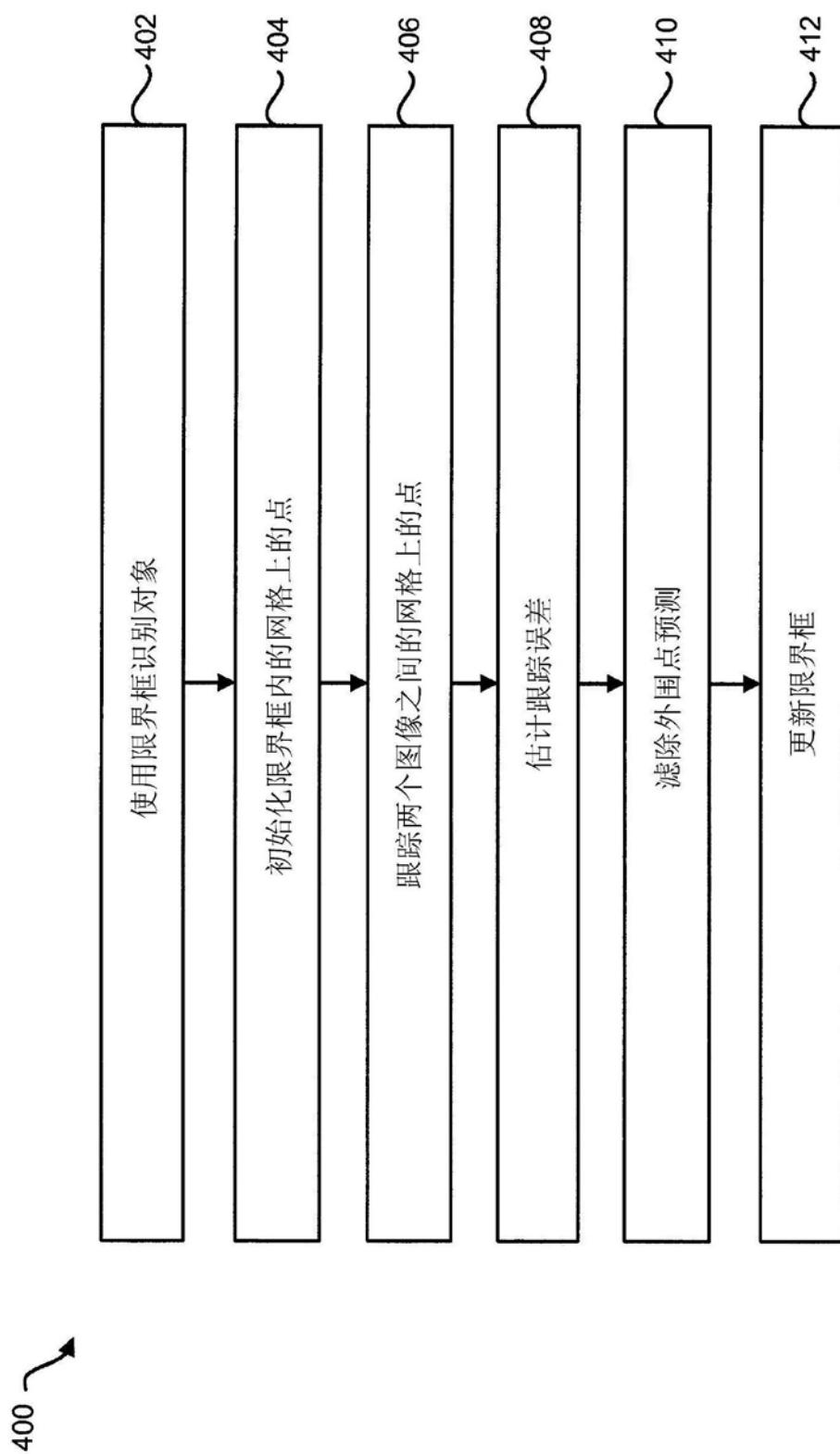


图4

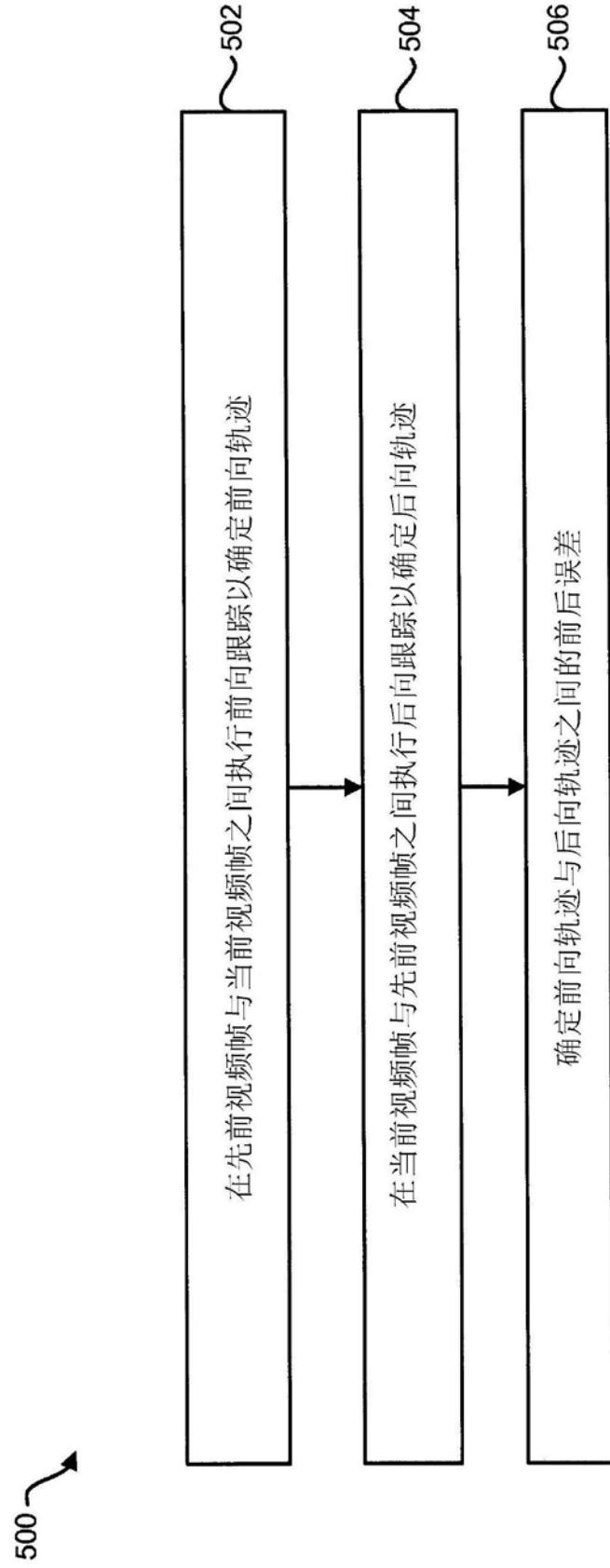


图5

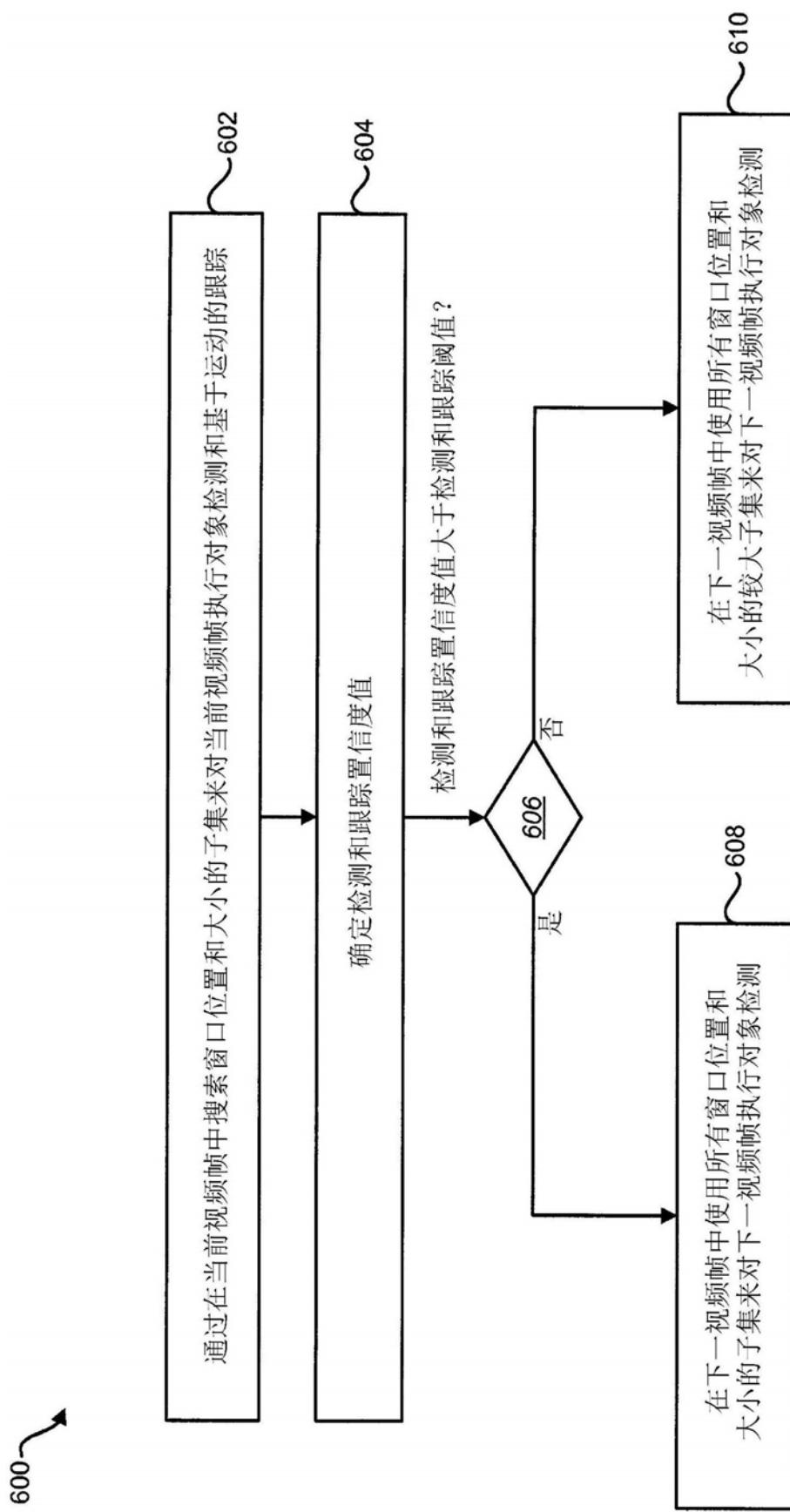


图6

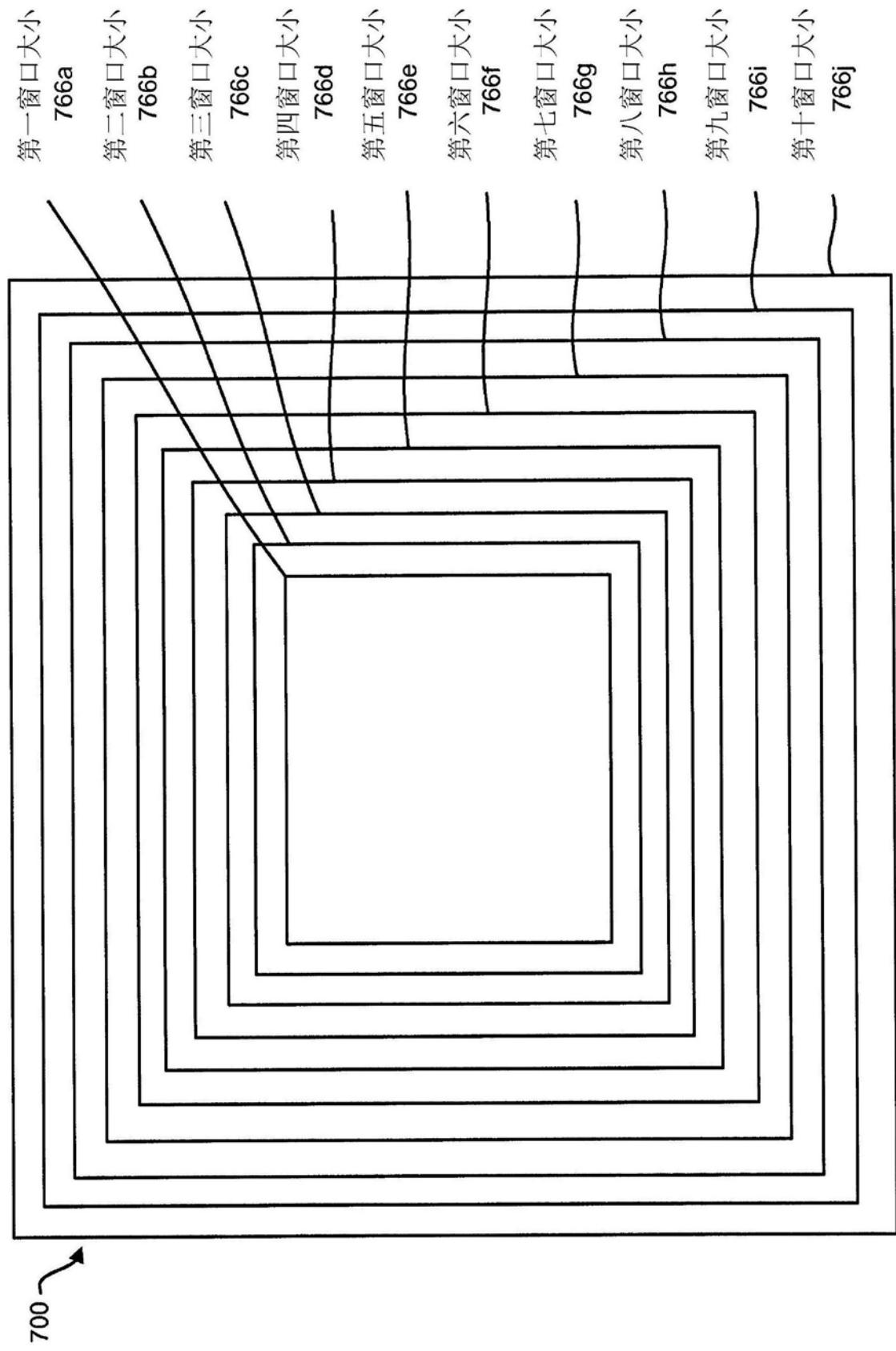


图7

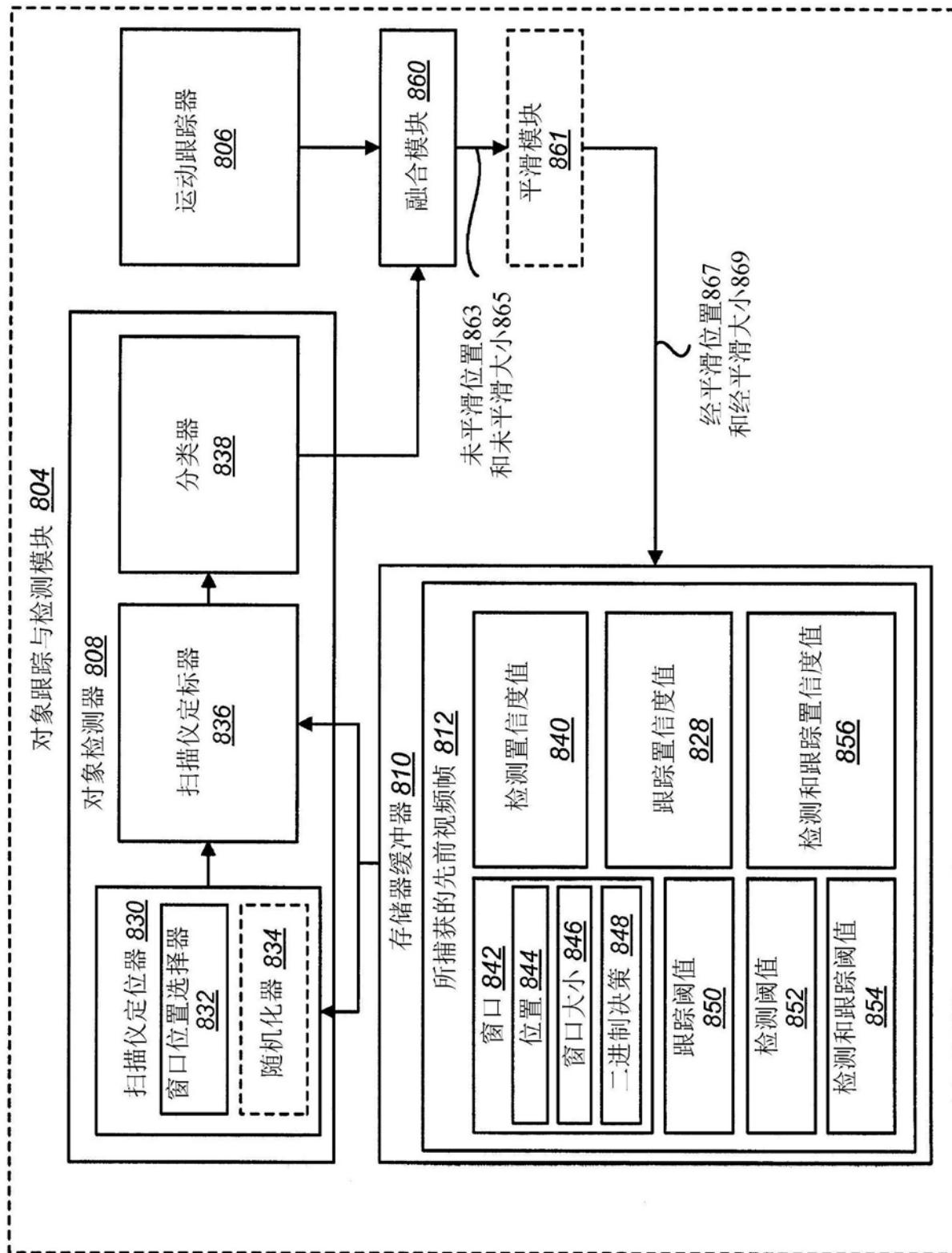


图8

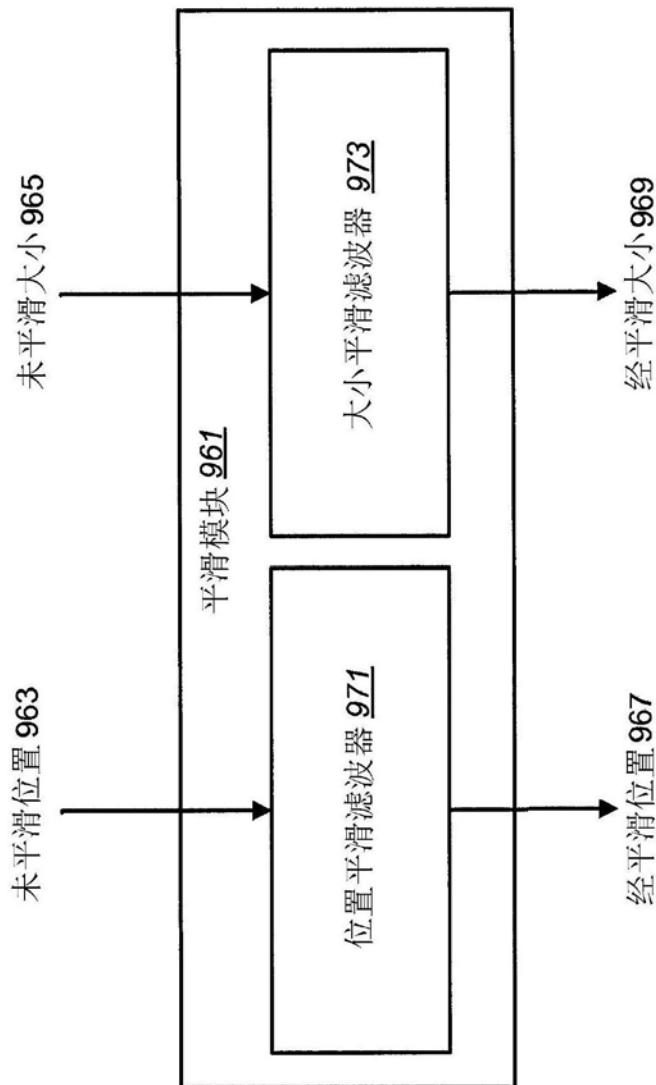


图9

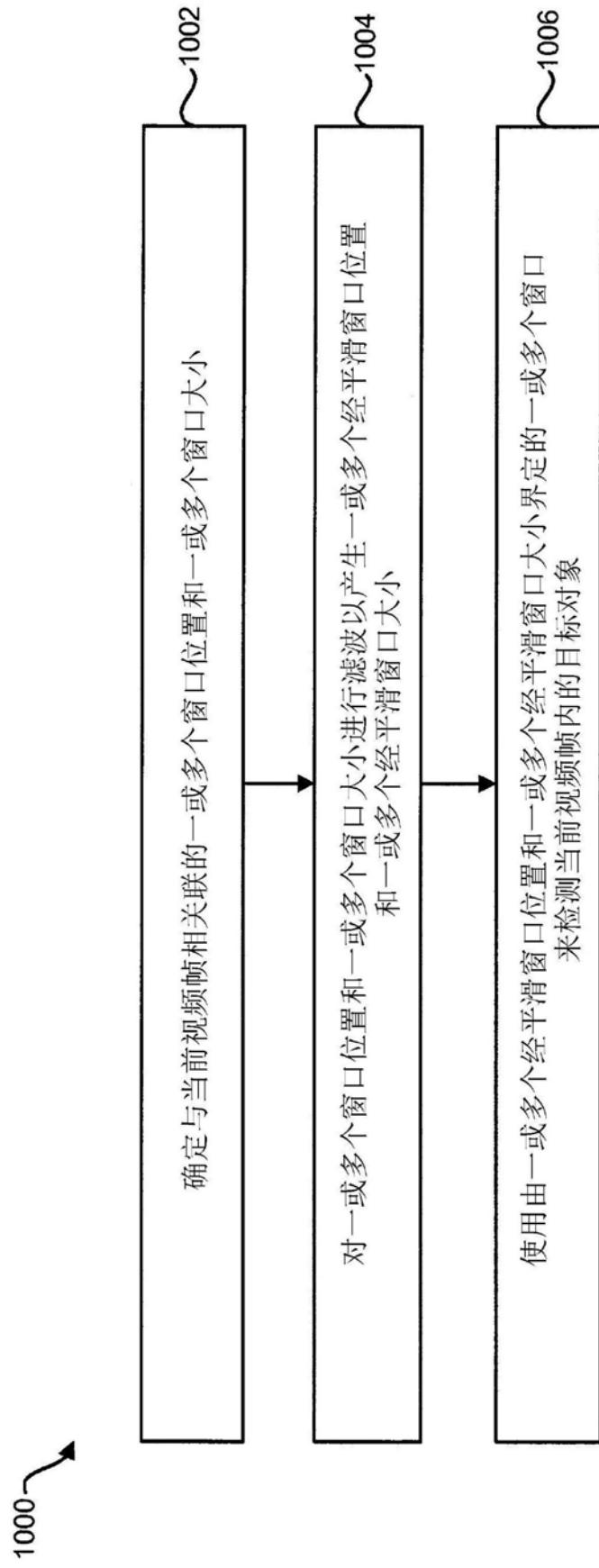


图10

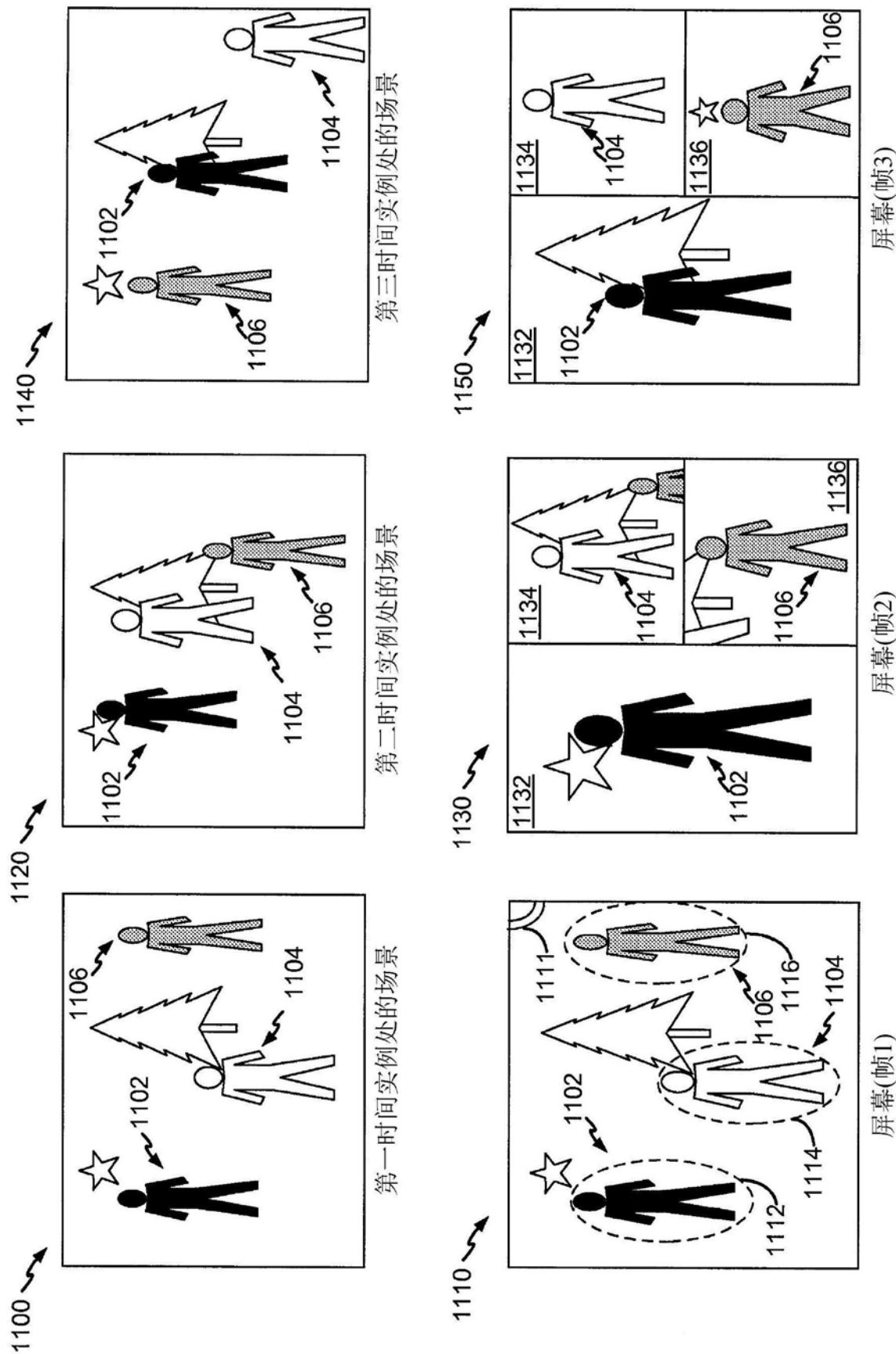


图11

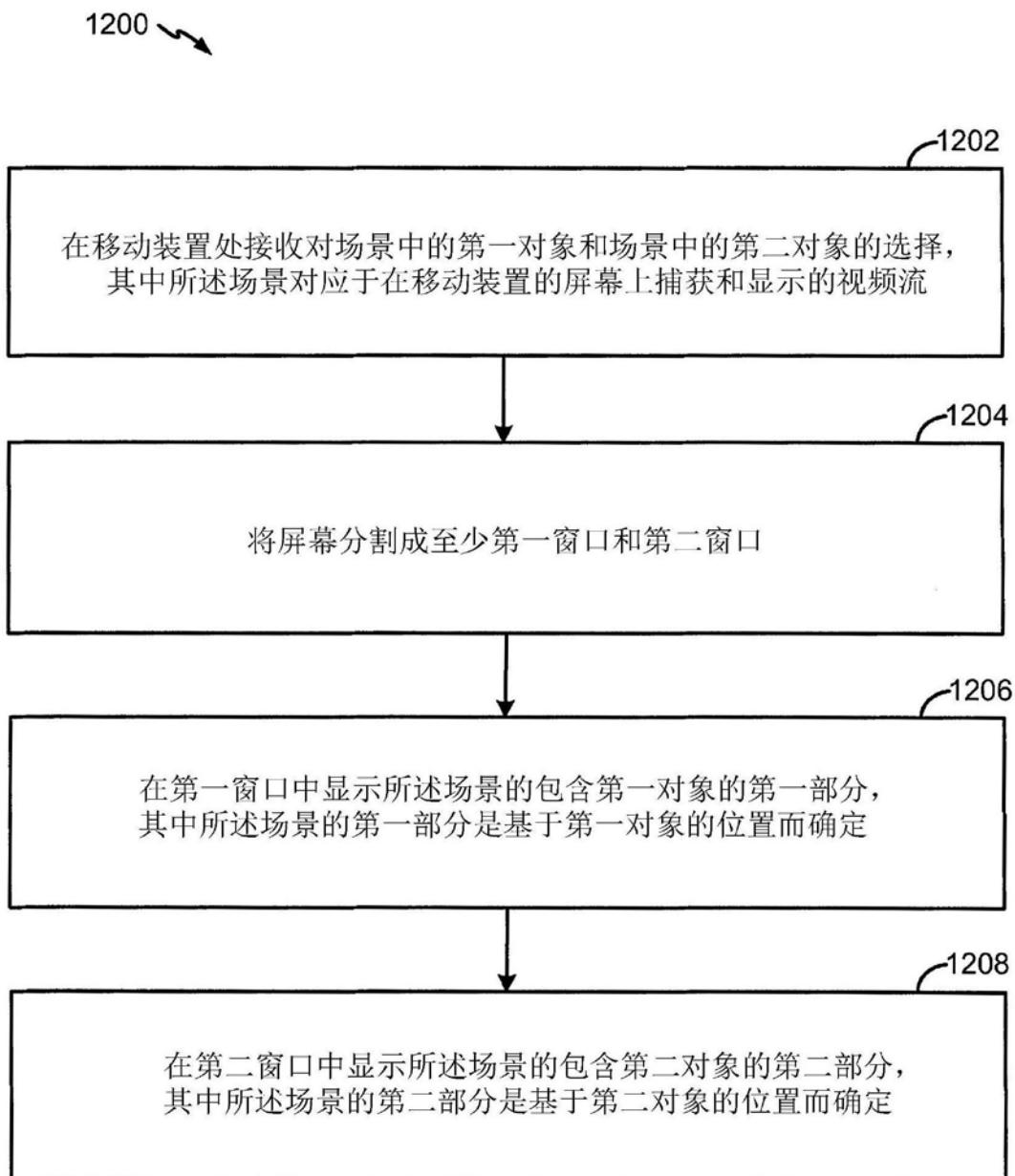


图12

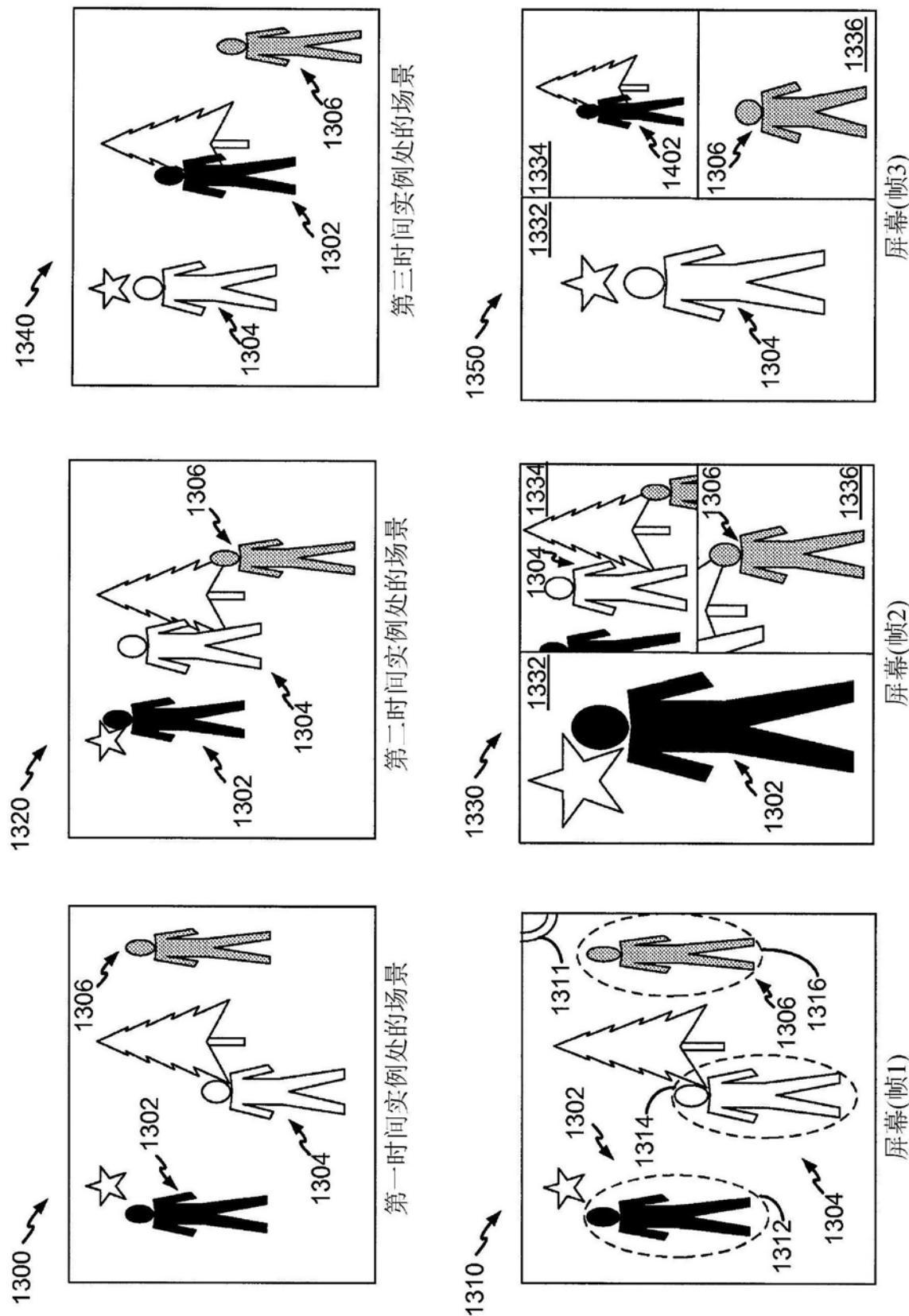


图13

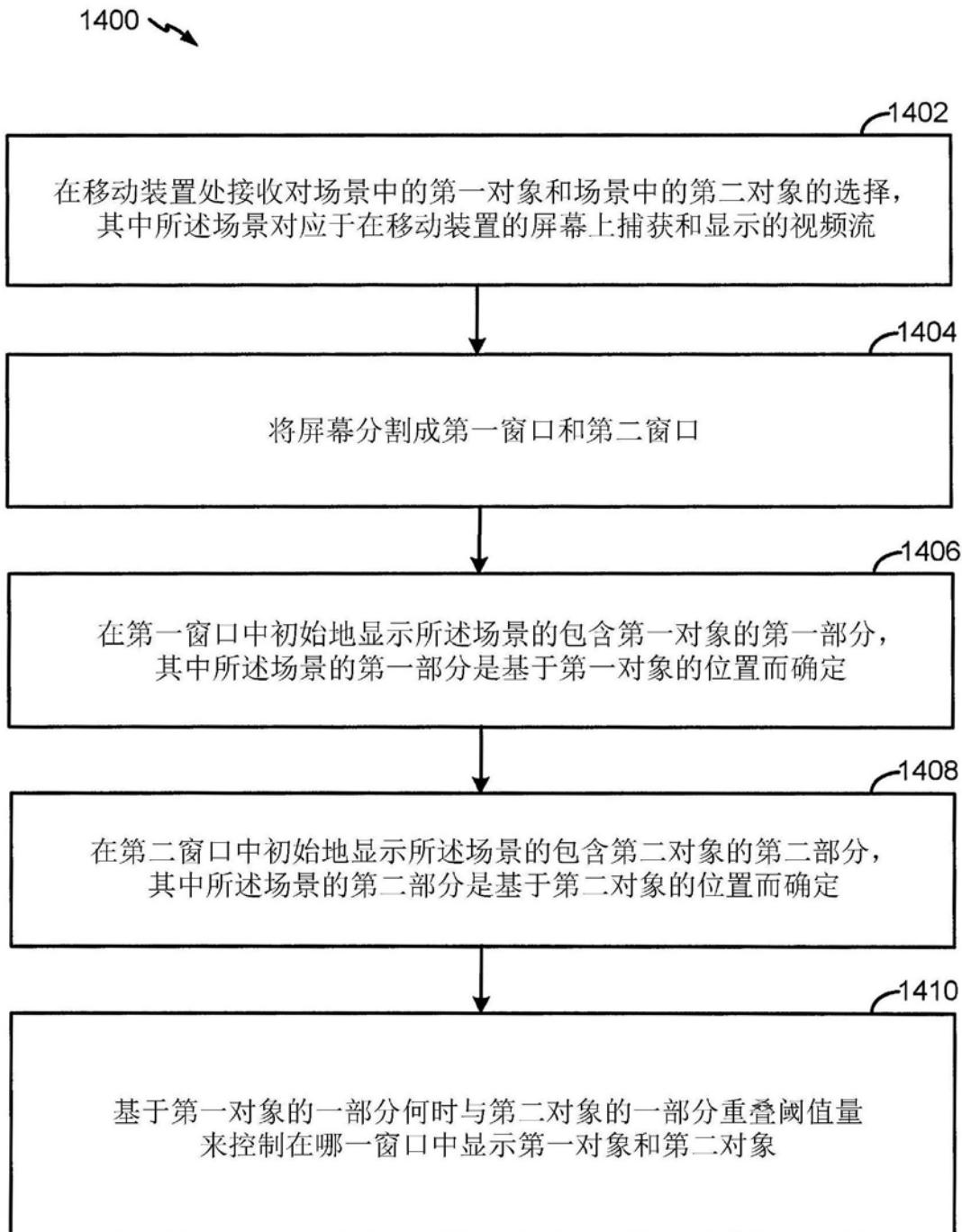


图14

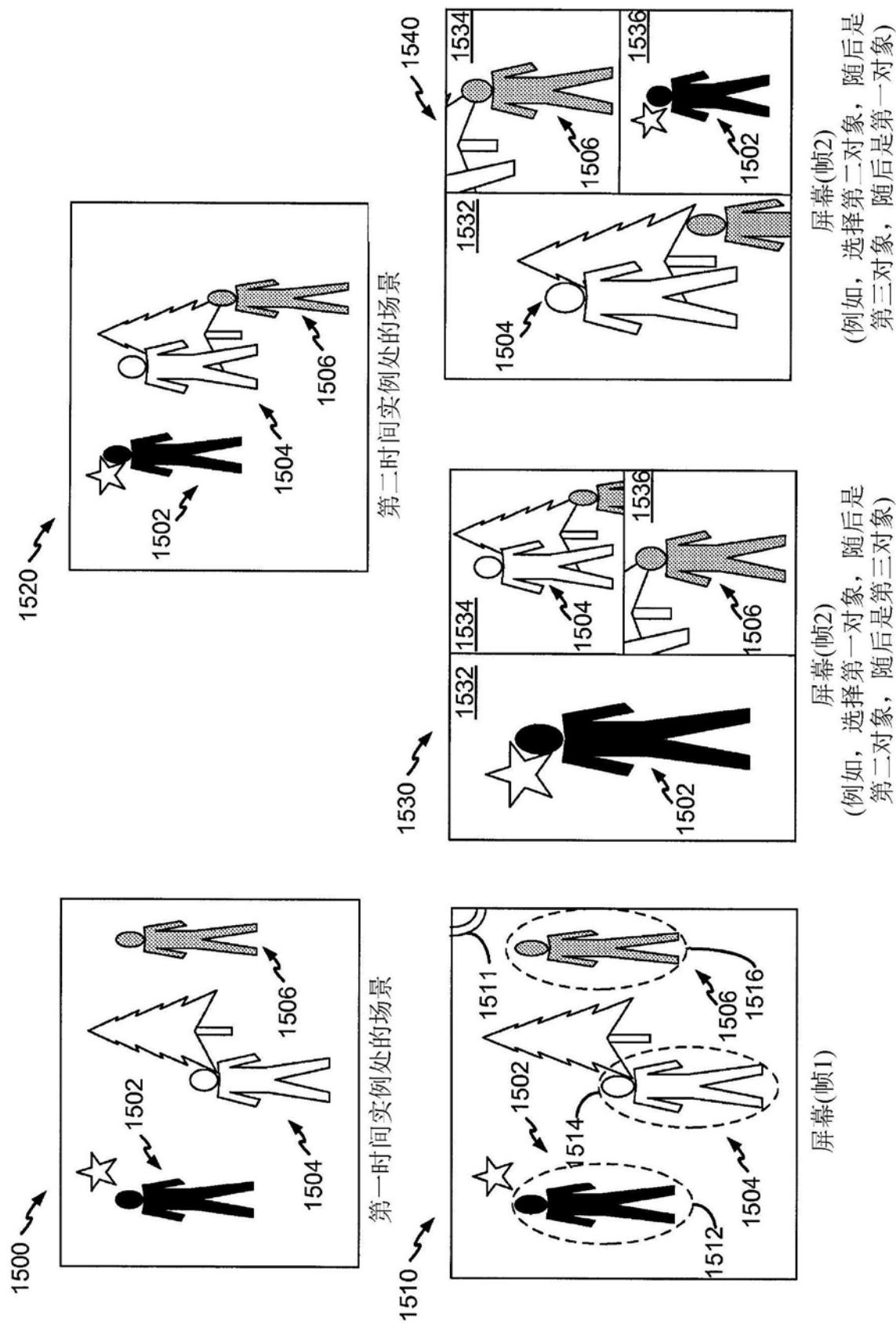


图15

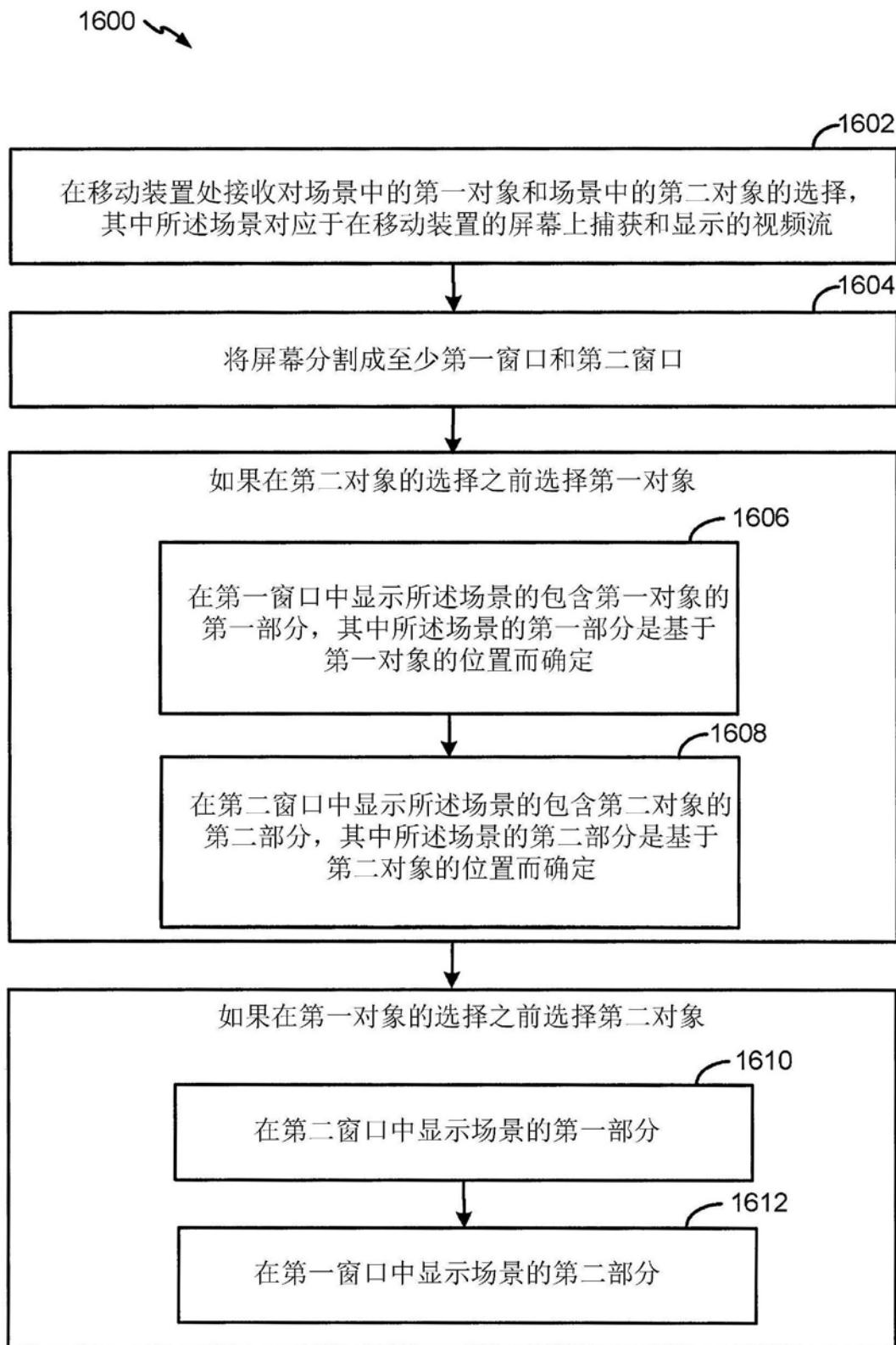


图16

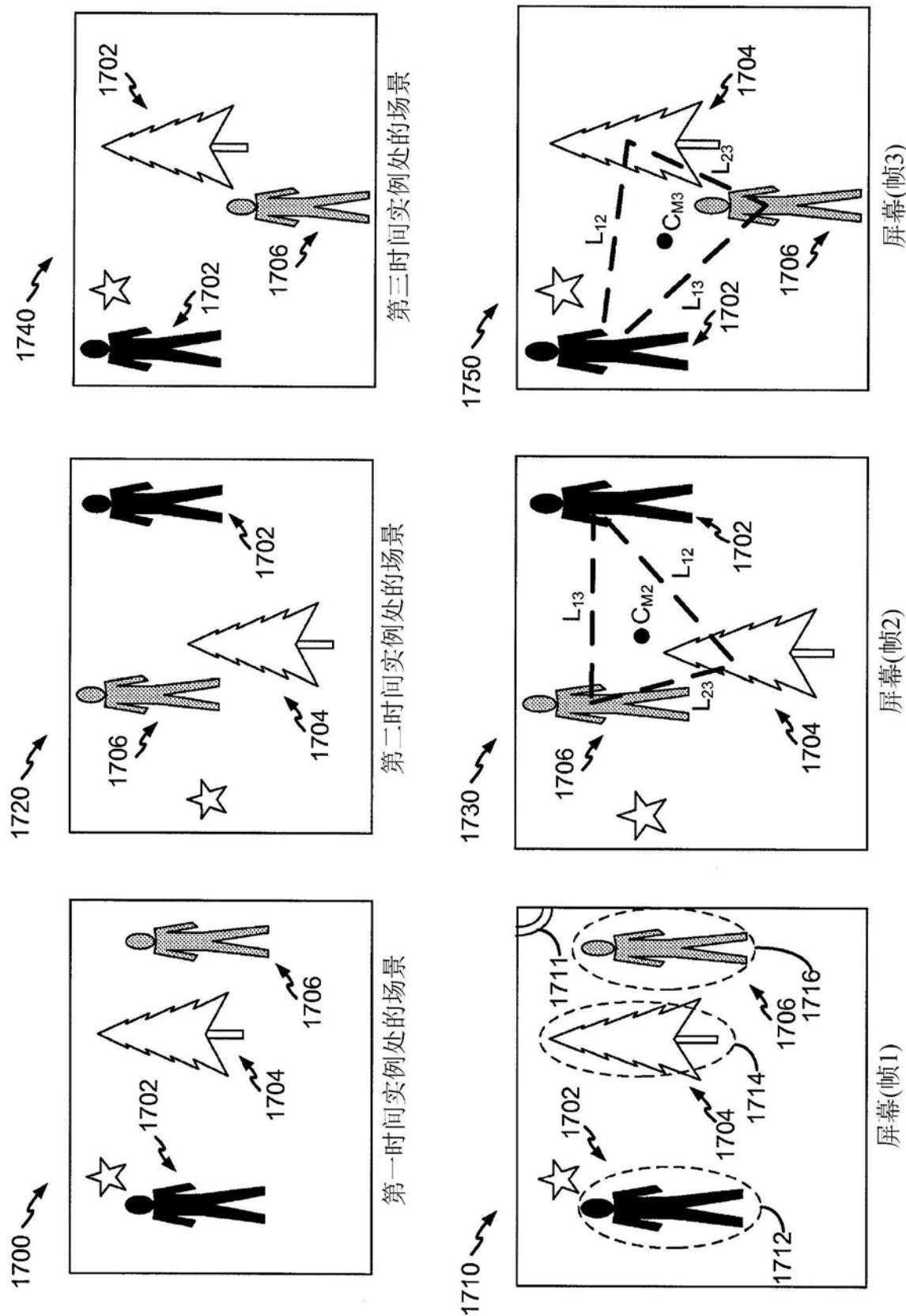


图17

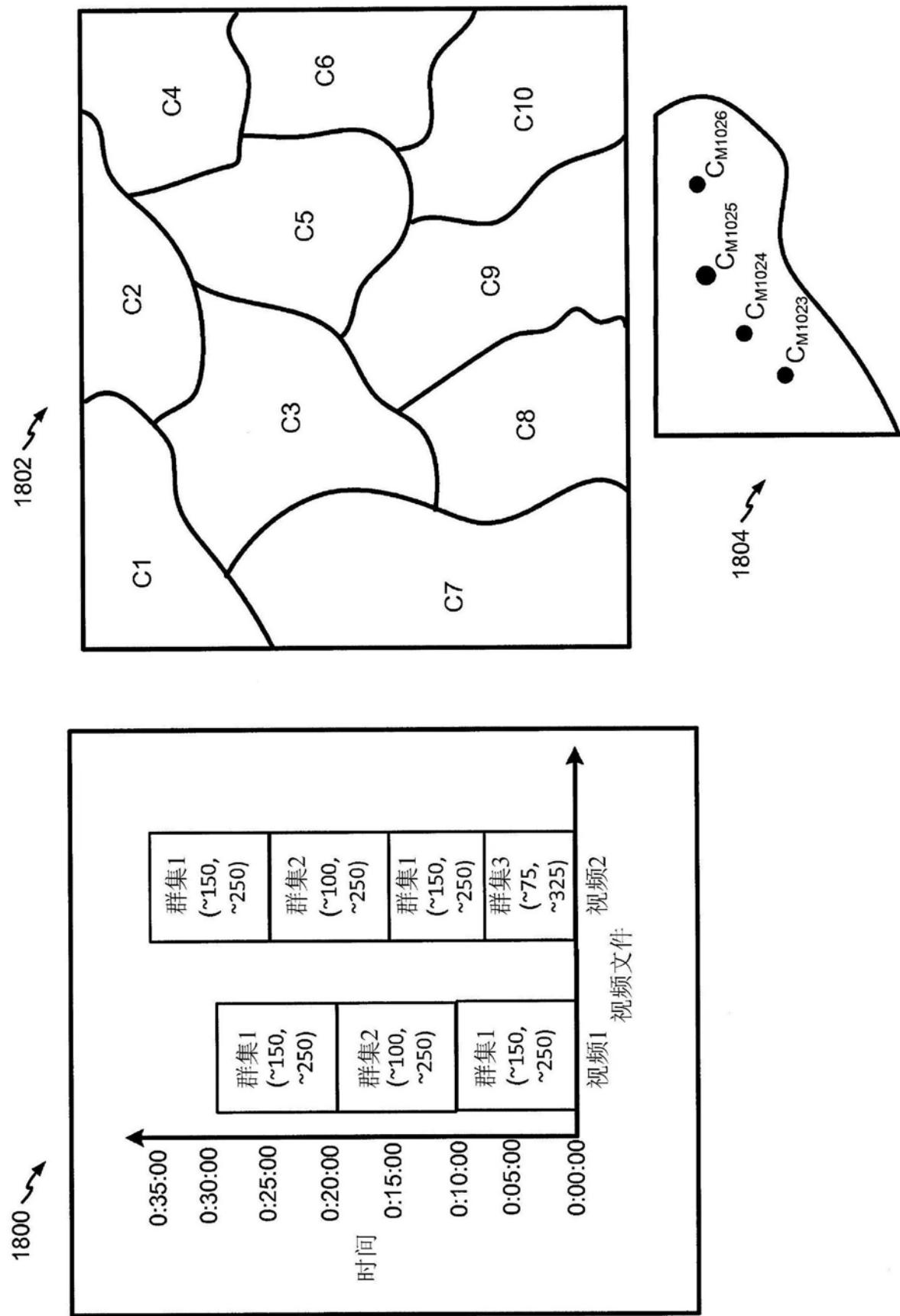


图18

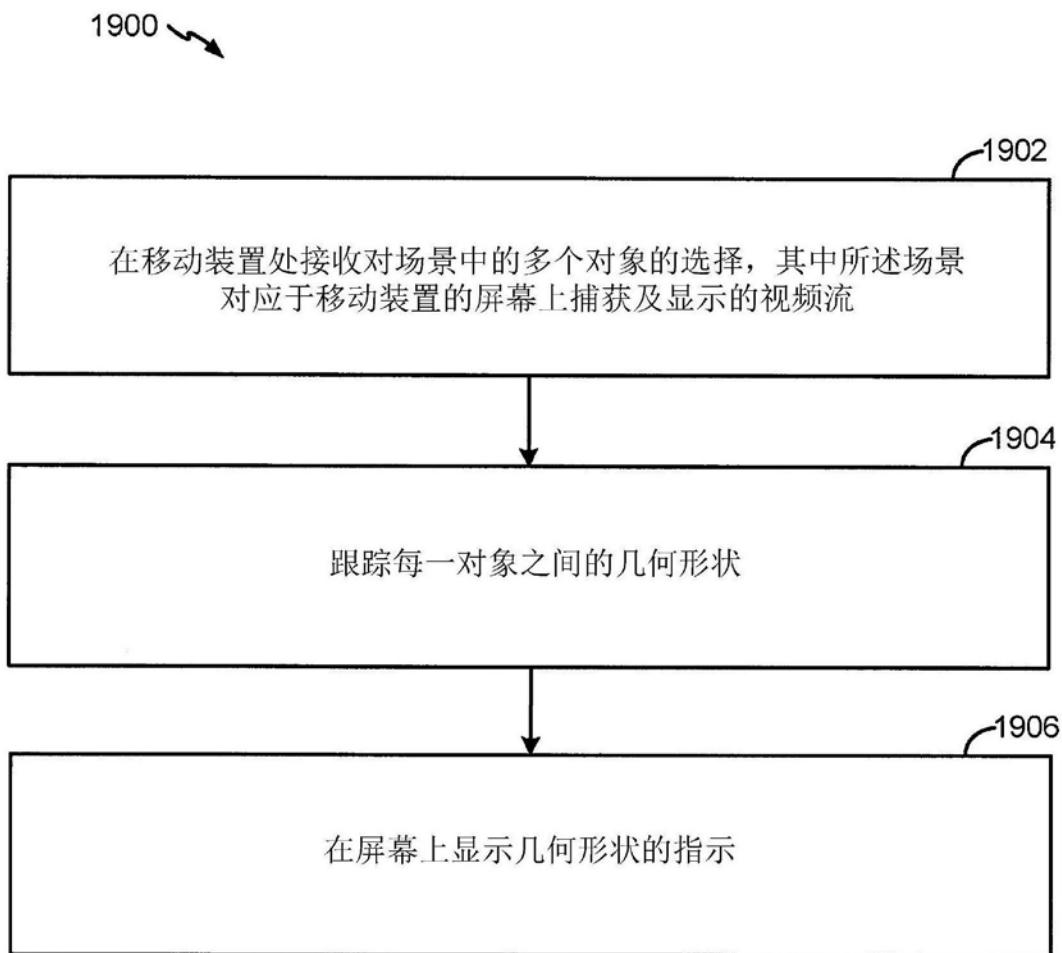


图19

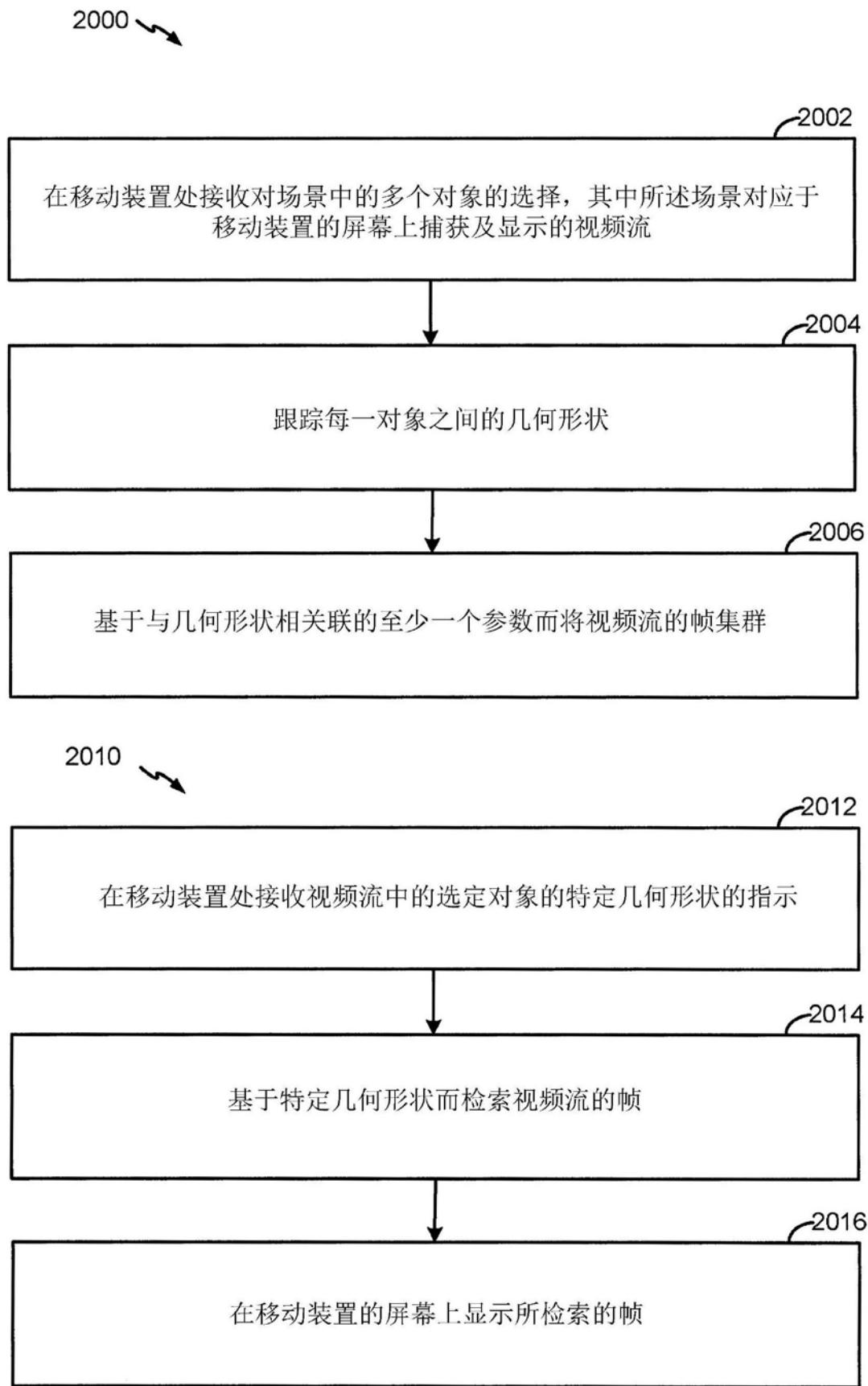


图20

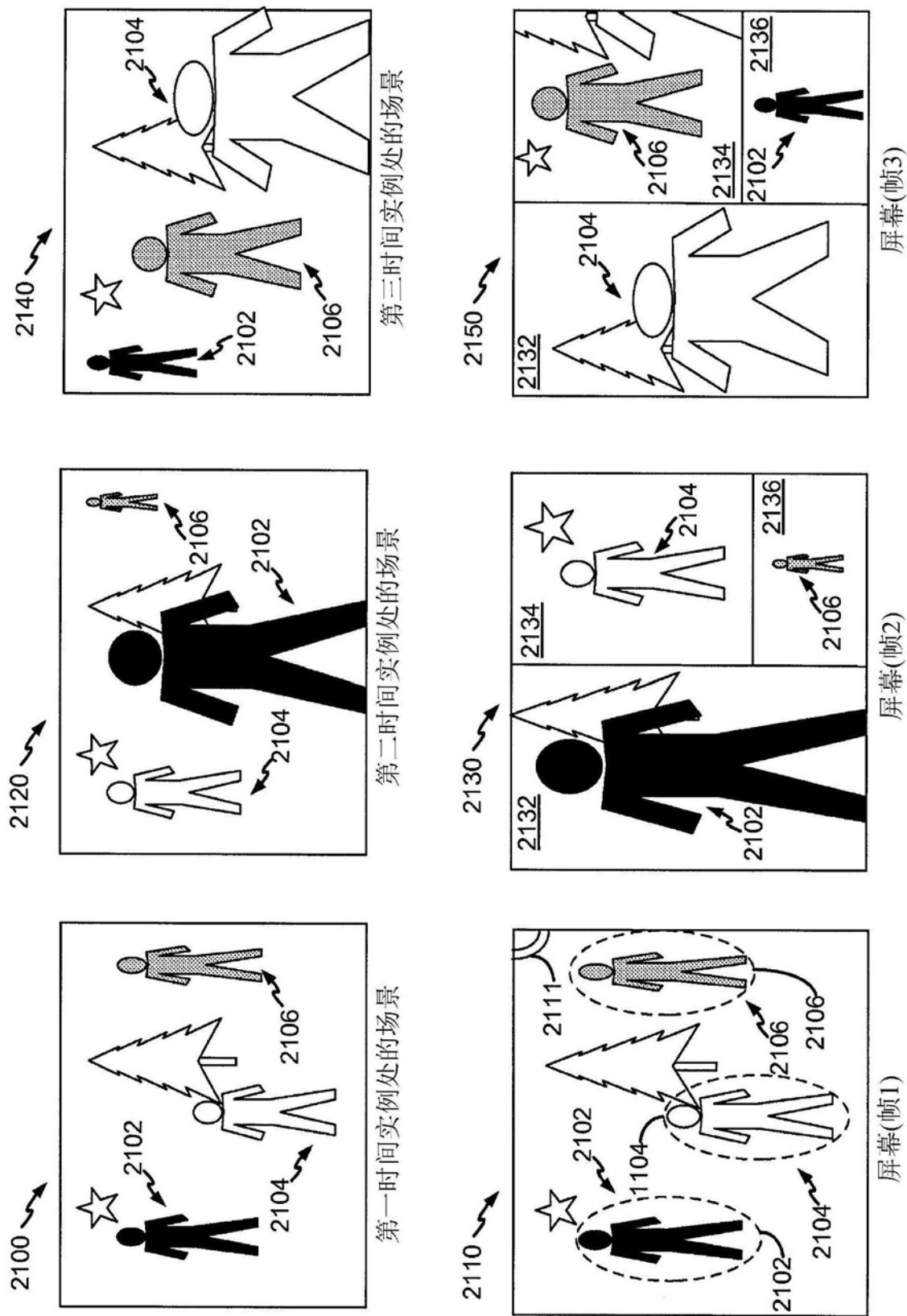


图21

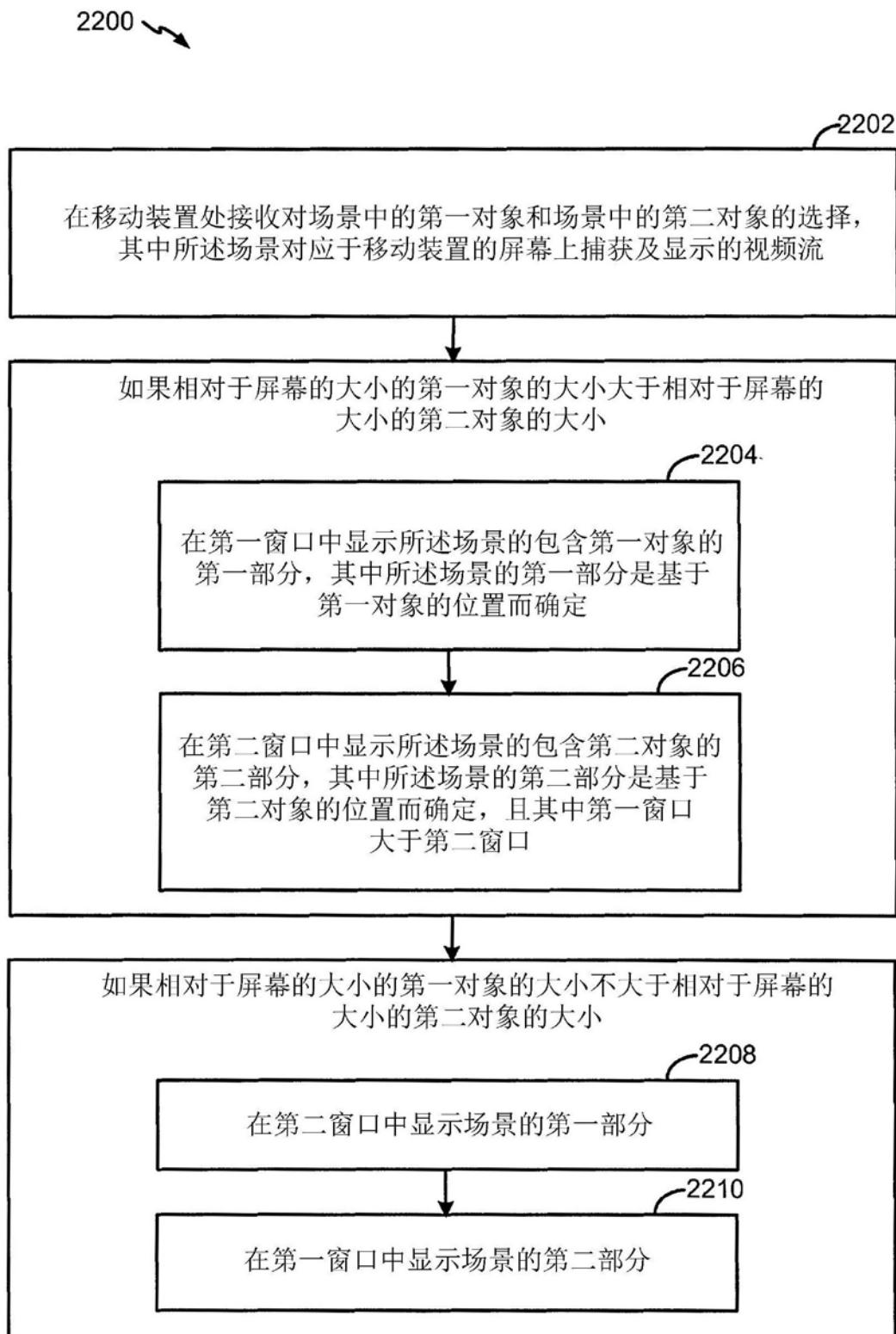


图22

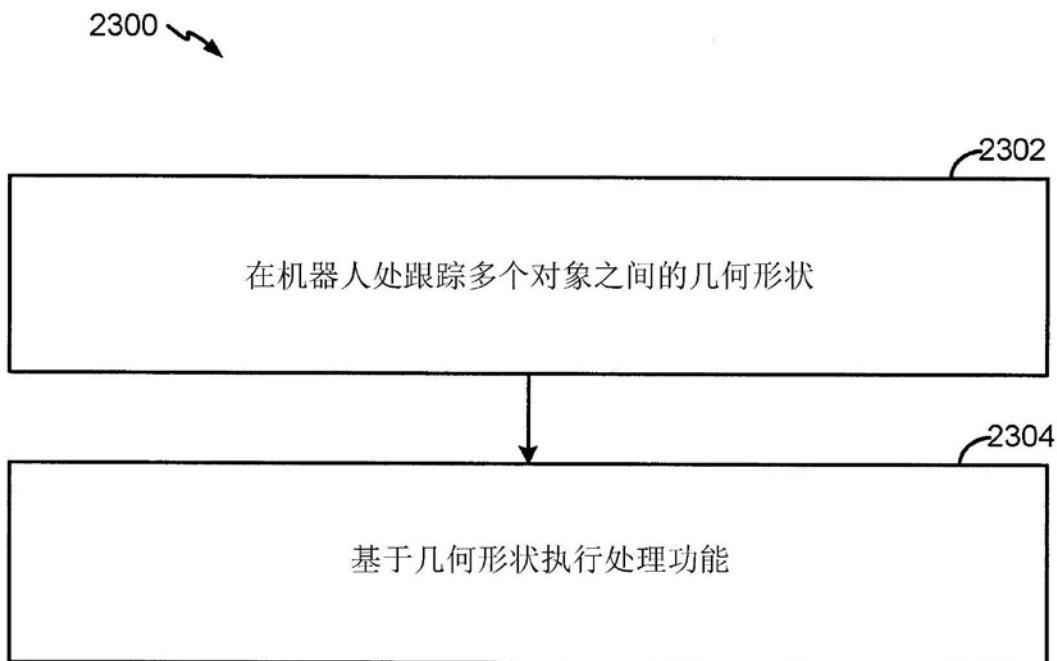


图23

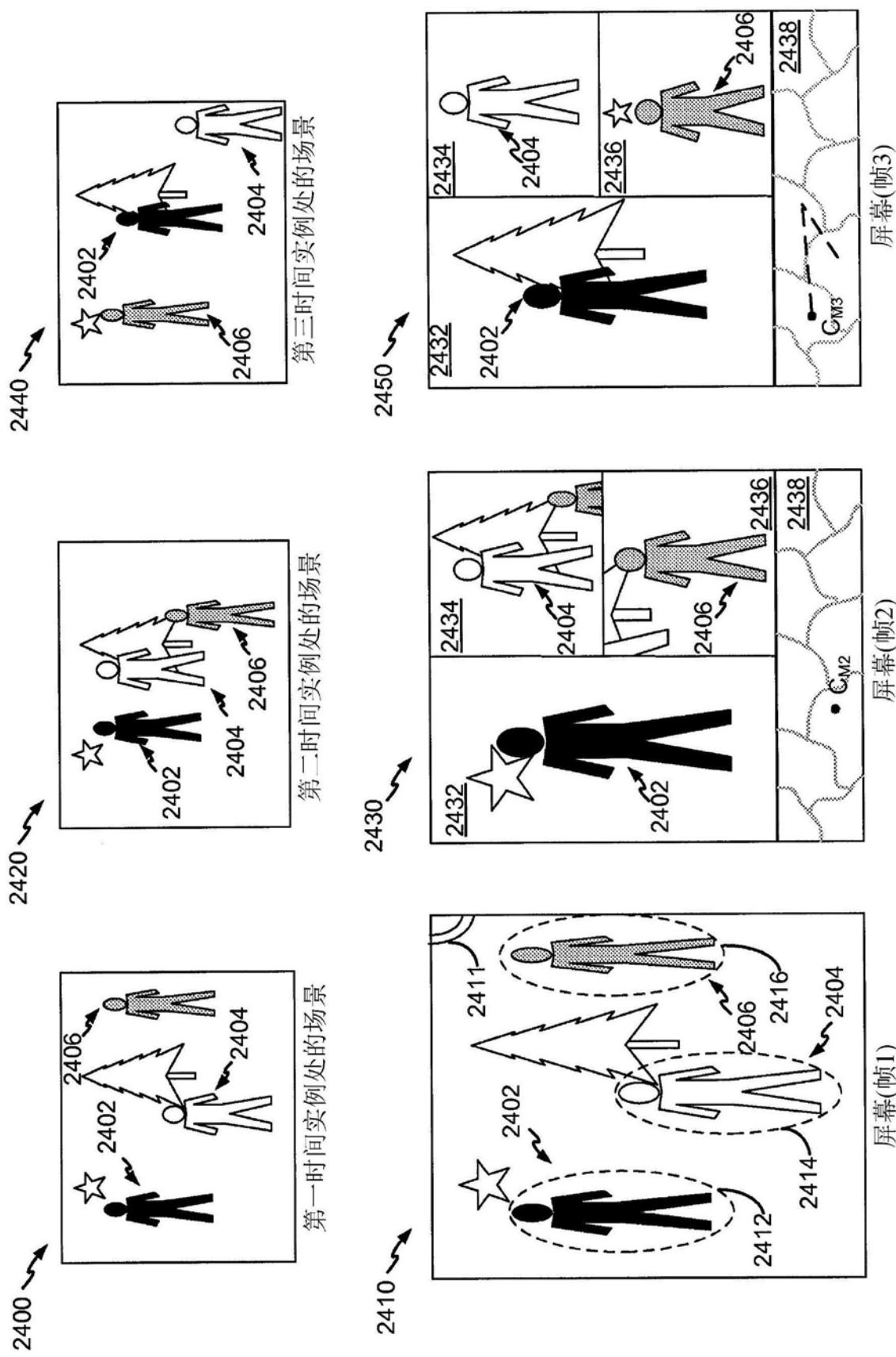


图24

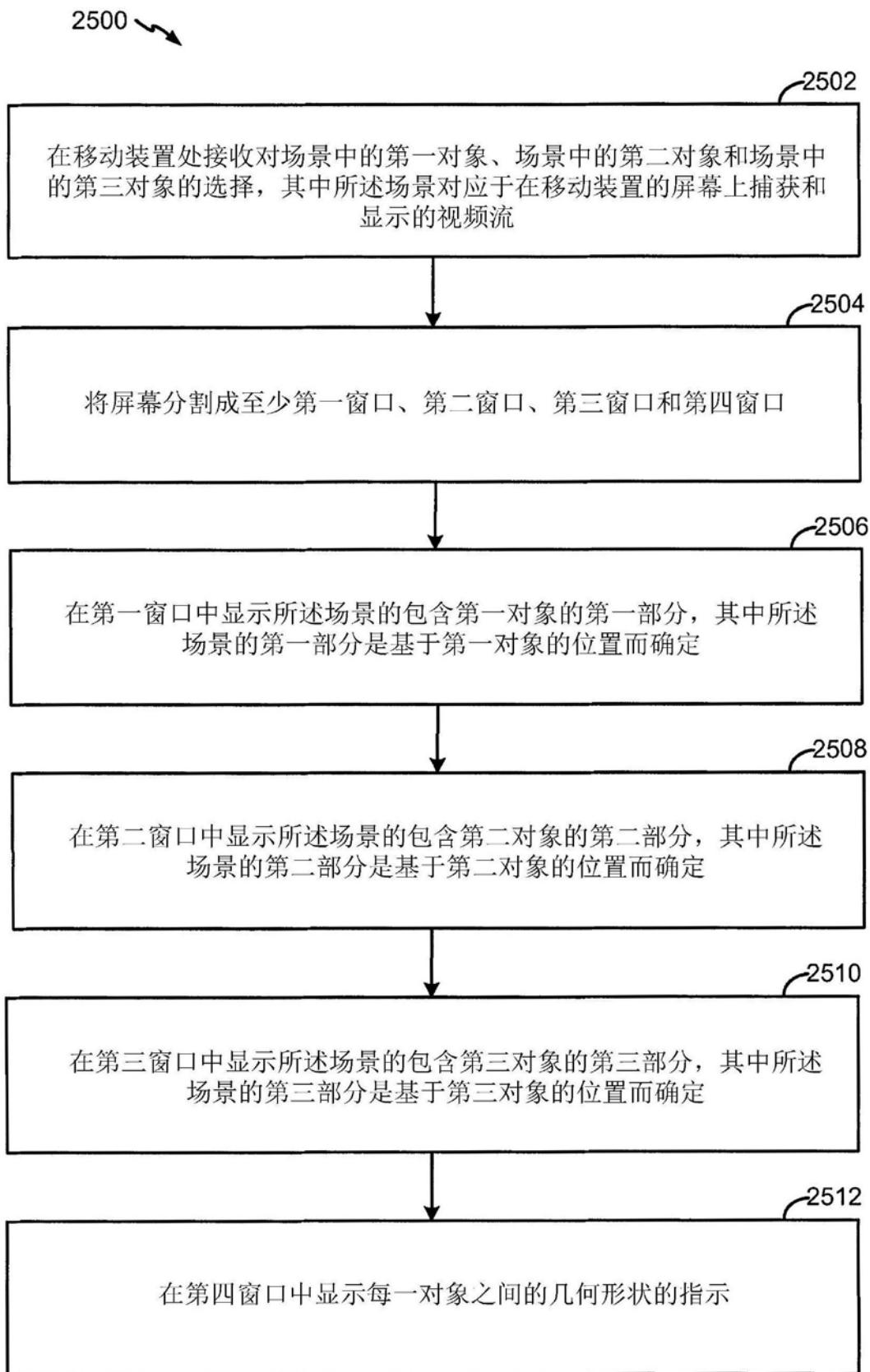


图25

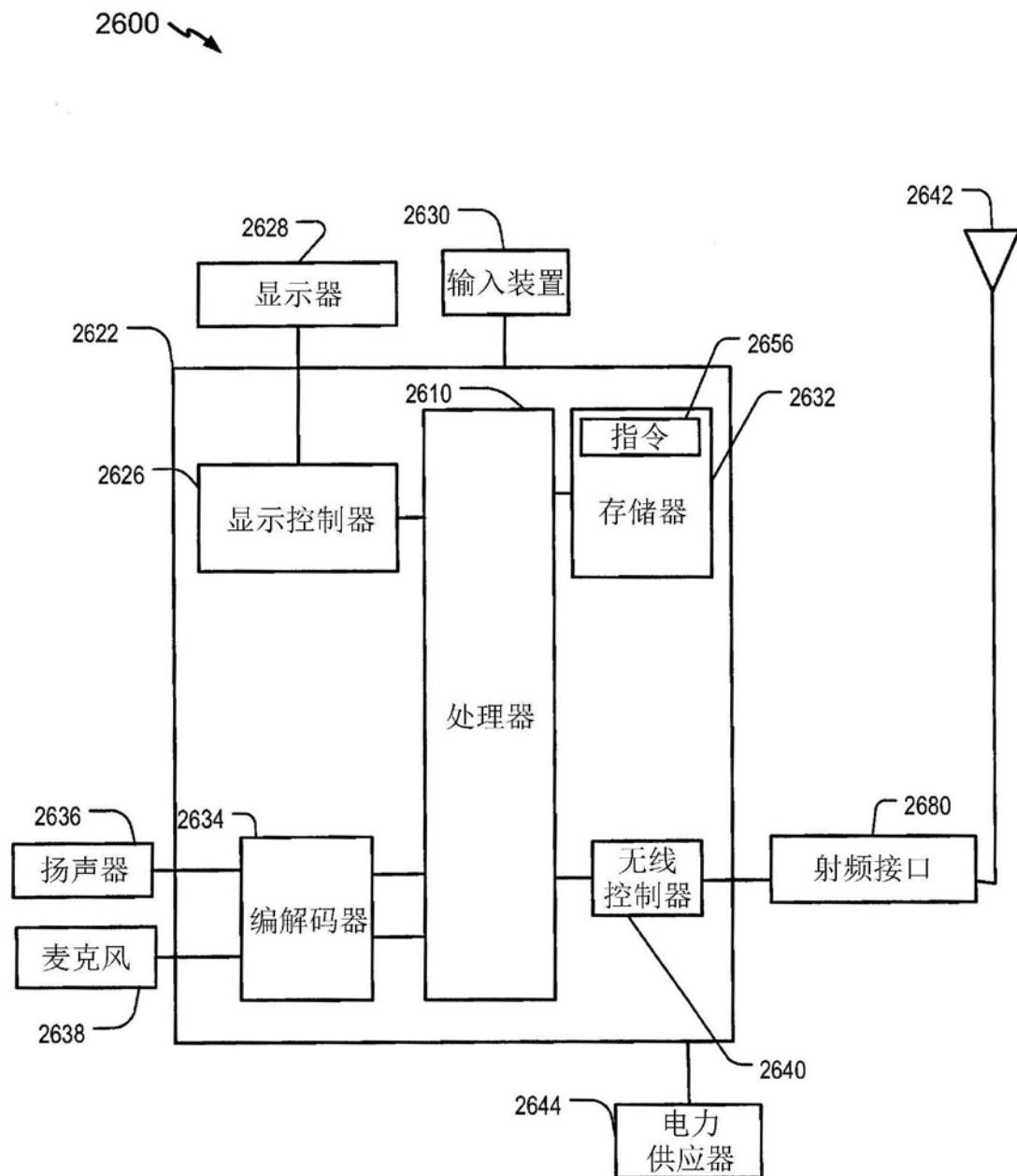


图26

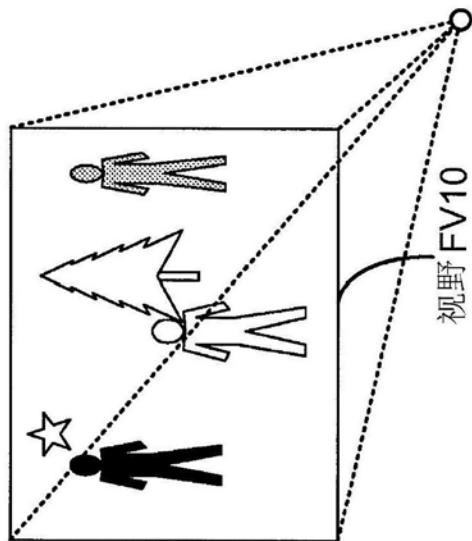


图27A

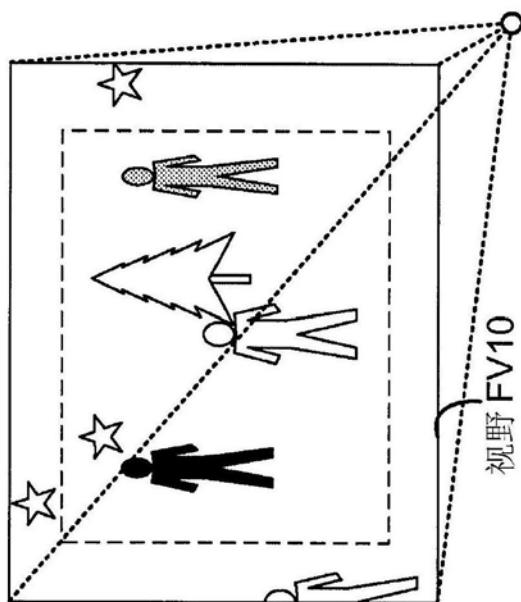


图27B

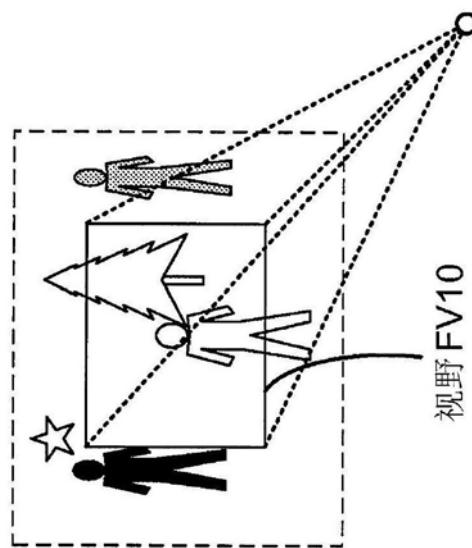


图27C

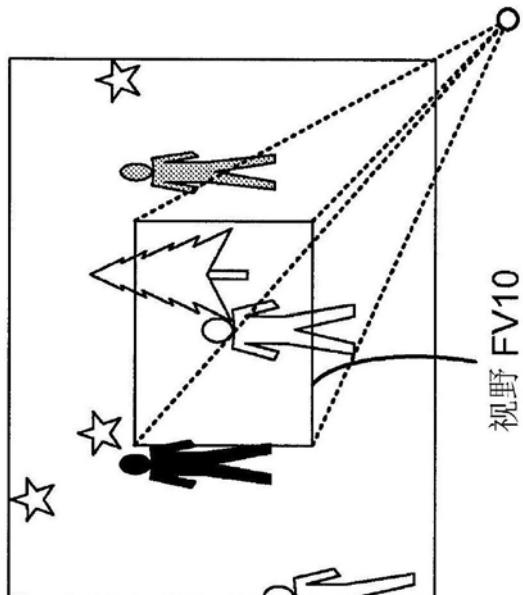


图27D

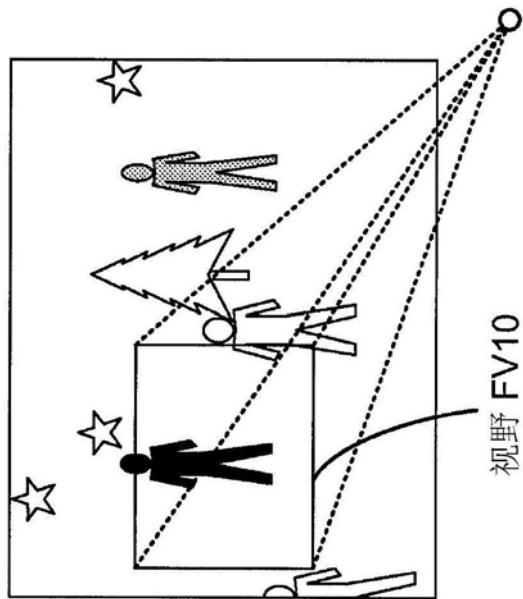


图27E

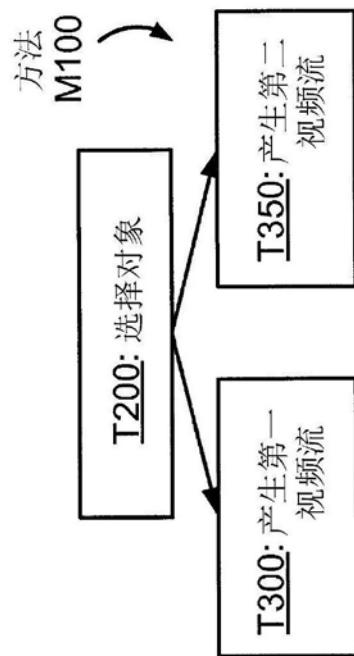


图28A

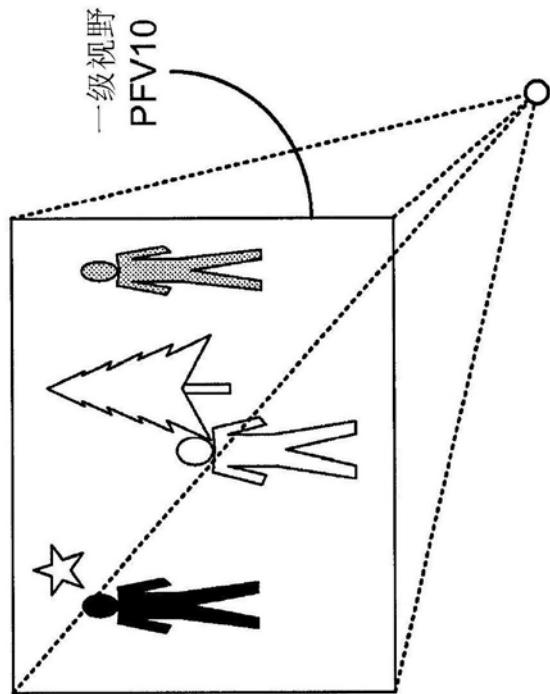


图28B

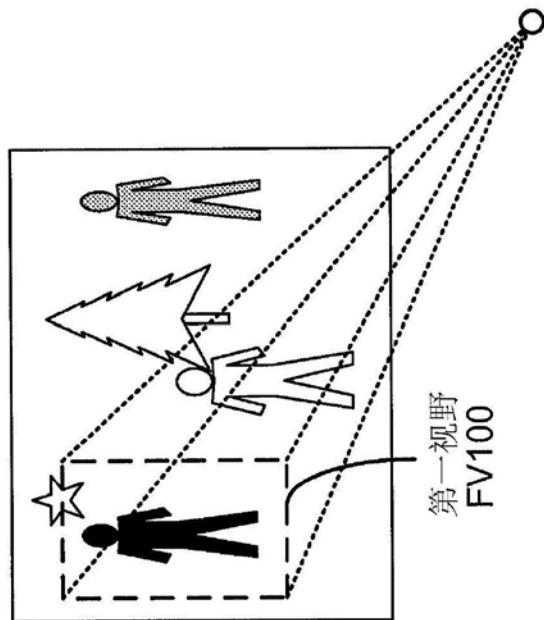


图28C

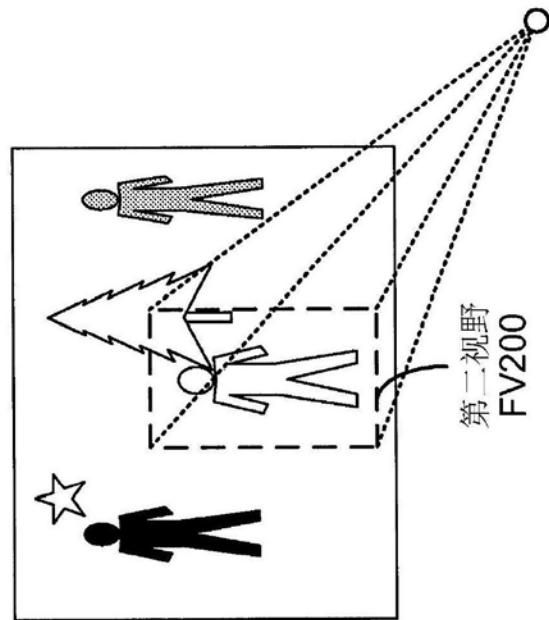


图28D

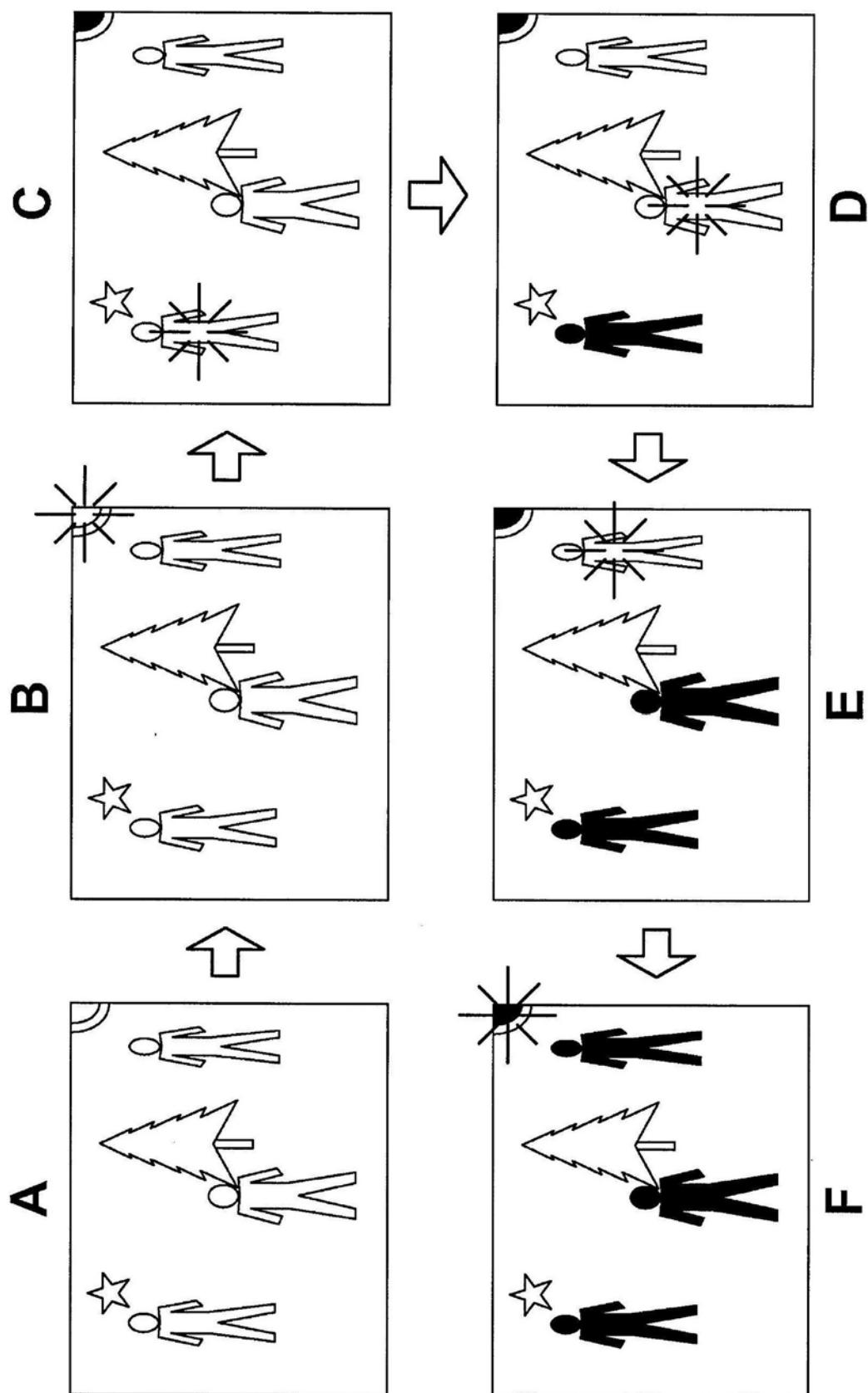


图29

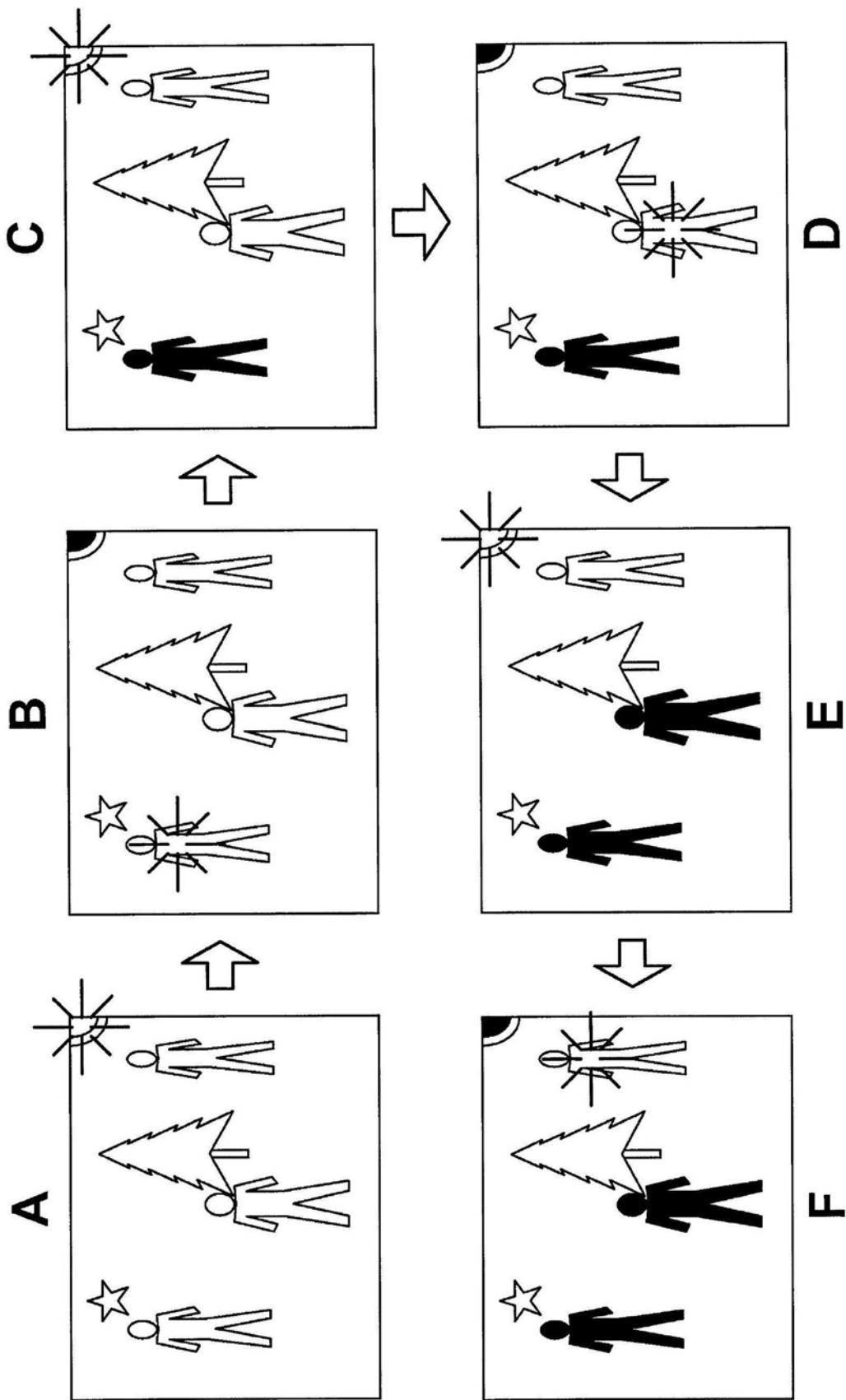


图30

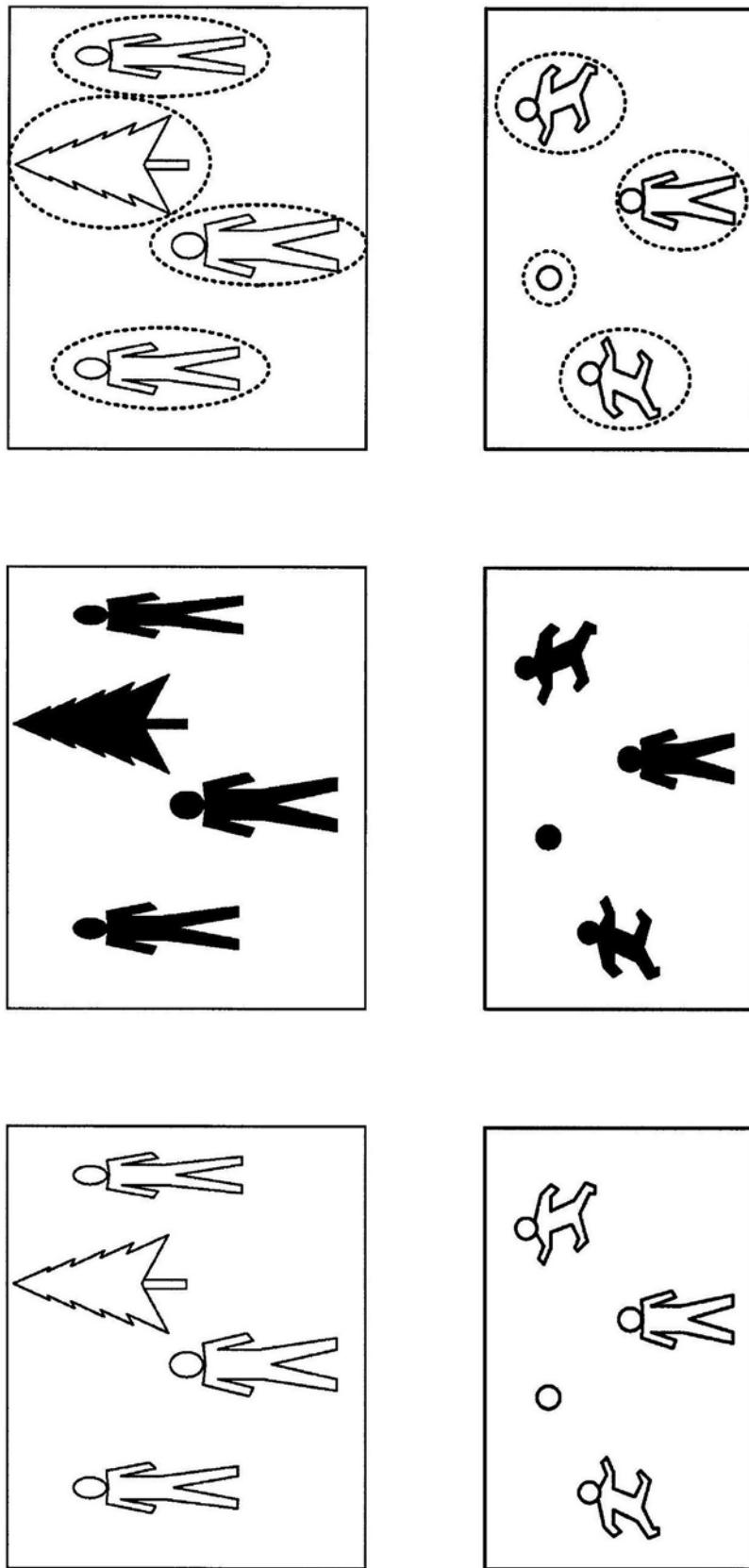


图31

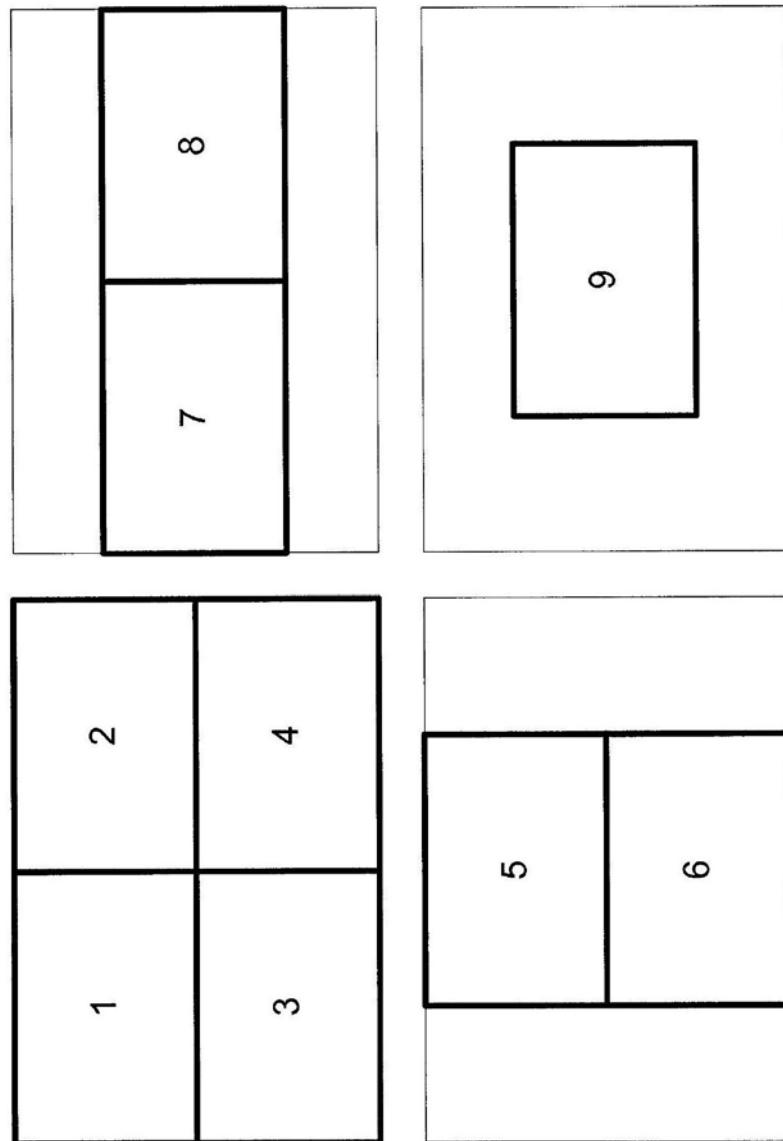


图32A

1	1,5	2,5	2
1,7	1,5,7,9	2,5,8,9	2,8
3,7	3,6,7,9	4,6,8,9	4,8
3	3,6	4,6	4

图32B

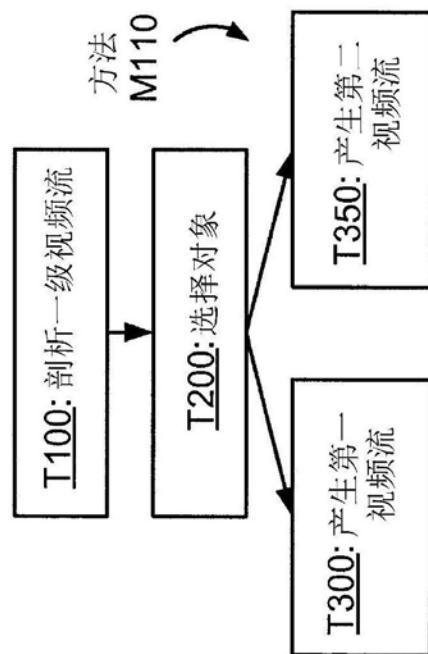


图33A

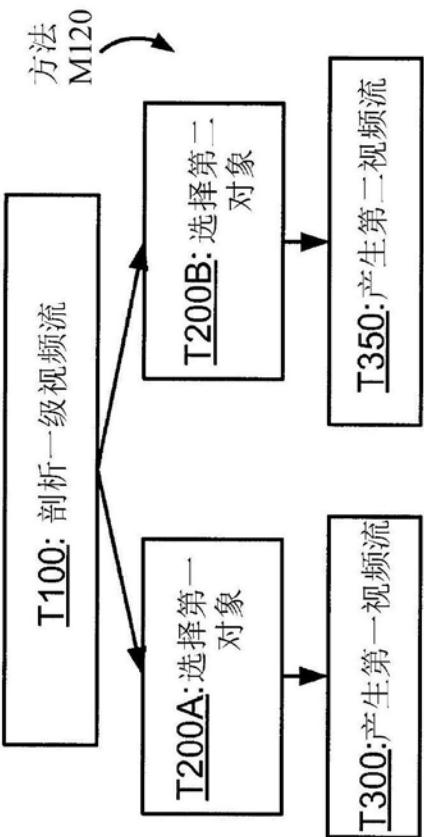


图33B

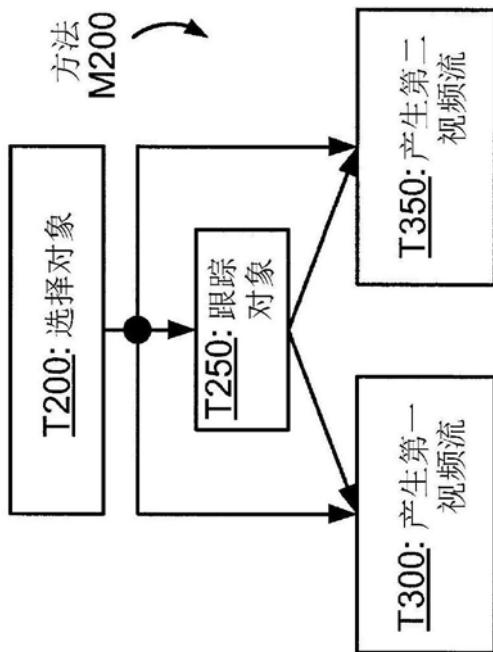


图33C

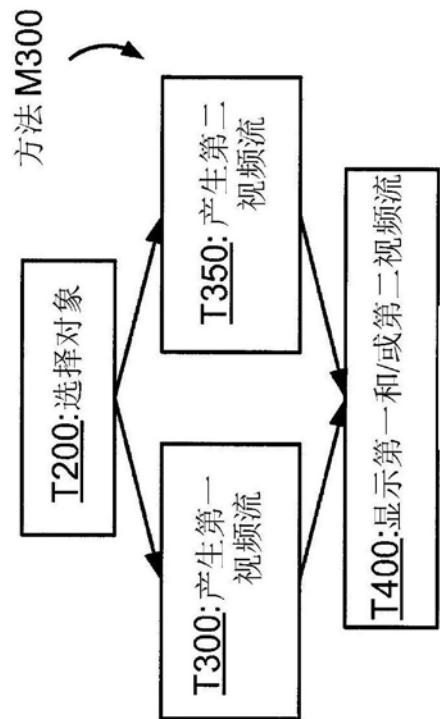


图33D

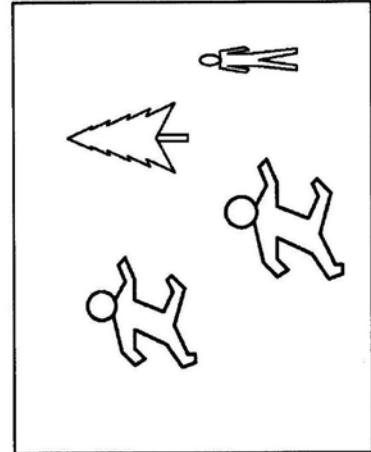


图34A

一级视频流PS10
的帧PF10

显示视频流DS10
的帧DF10

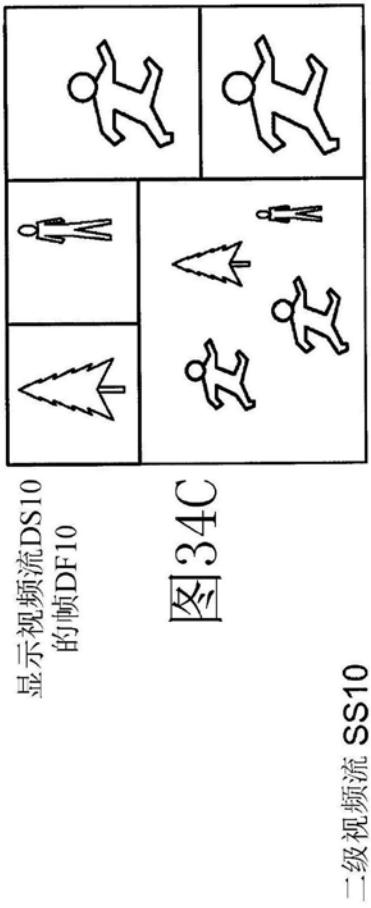


图34C

显示视频流DS10
的帧DF10

二级视频流 SS10

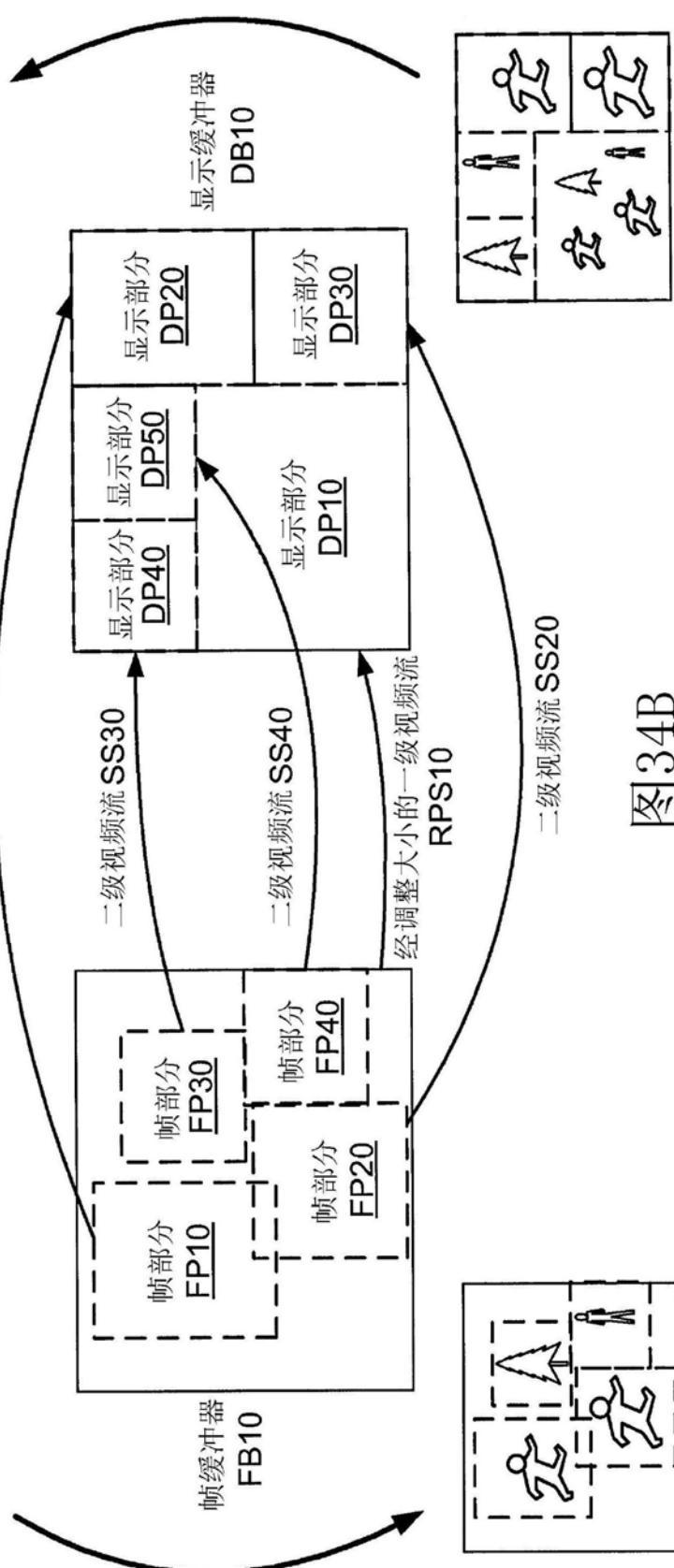
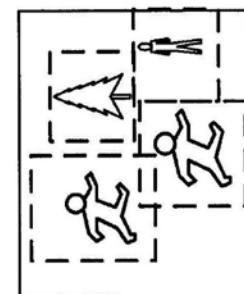


图34B

二级视频流 SS20



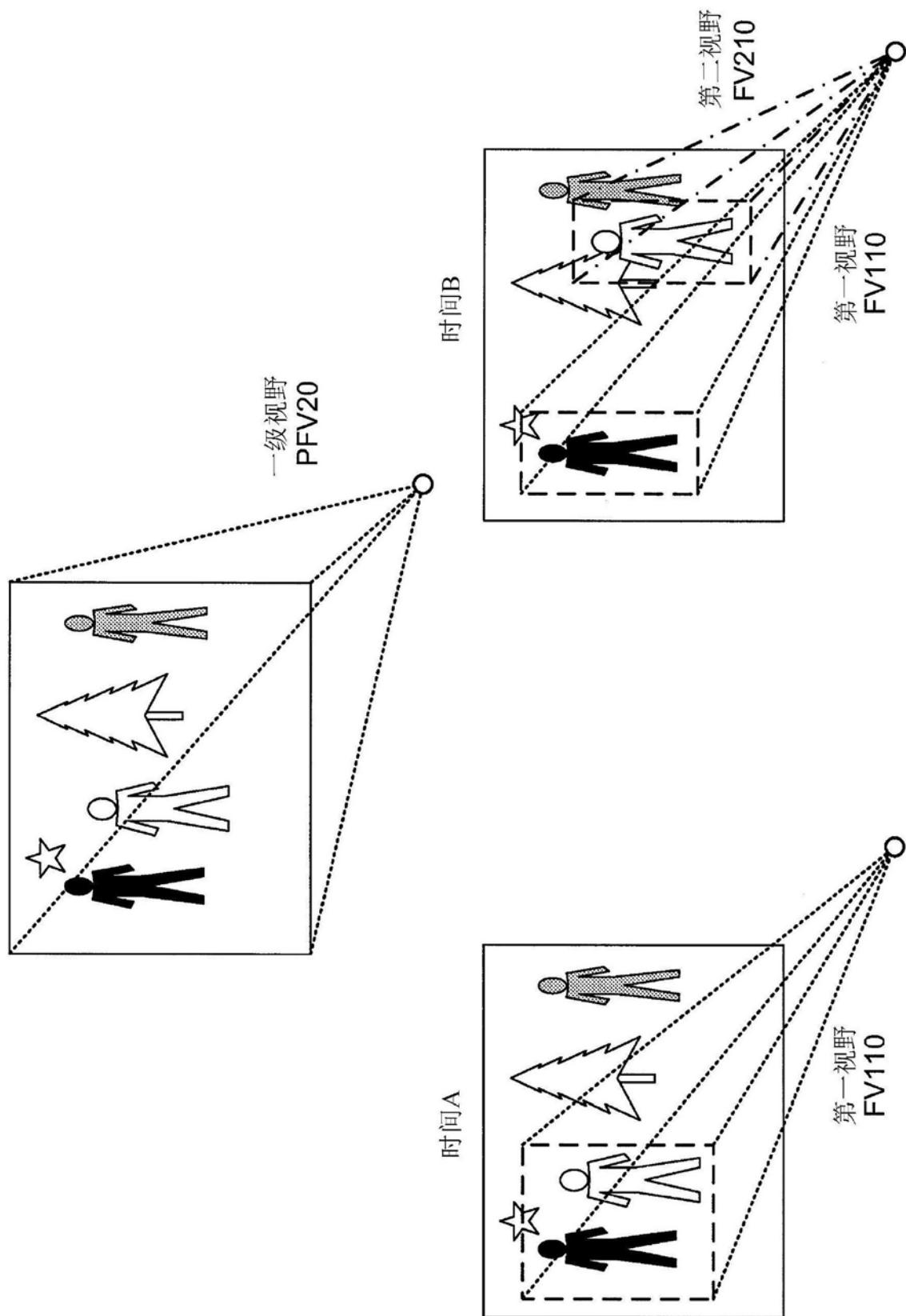


图35

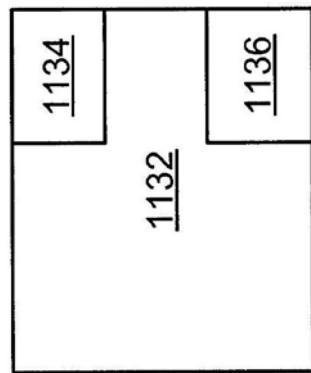


图36A

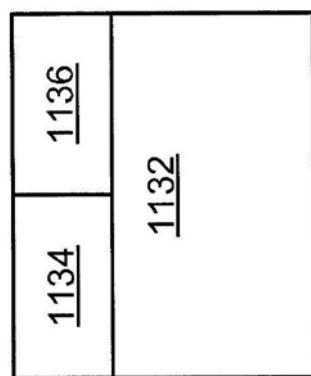


图36B

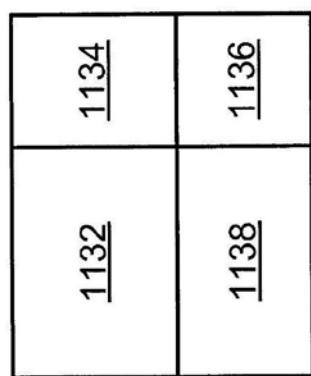


图36C

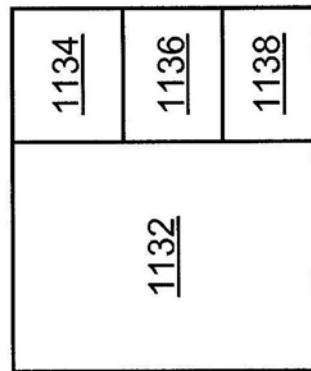


图36D

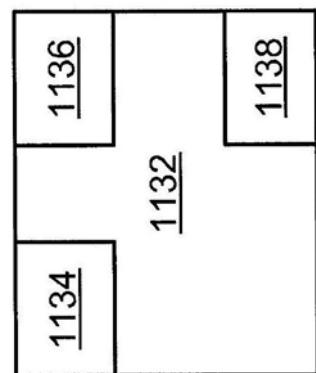


图36E

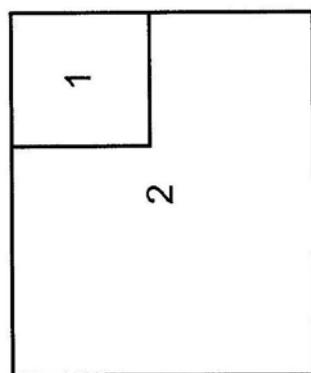


图36F

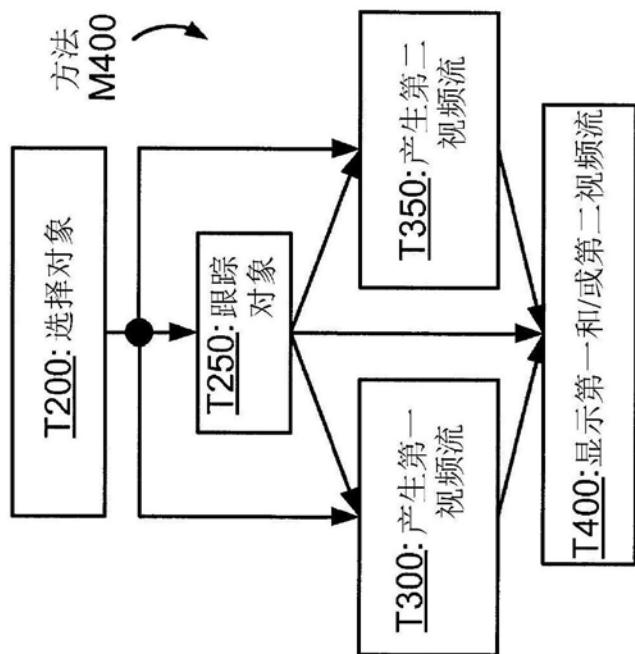


图37A

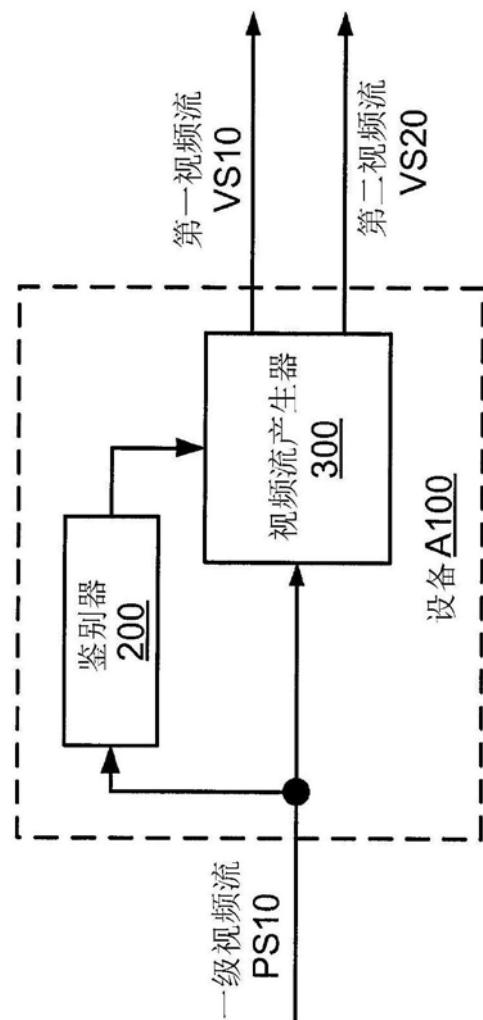


图37B

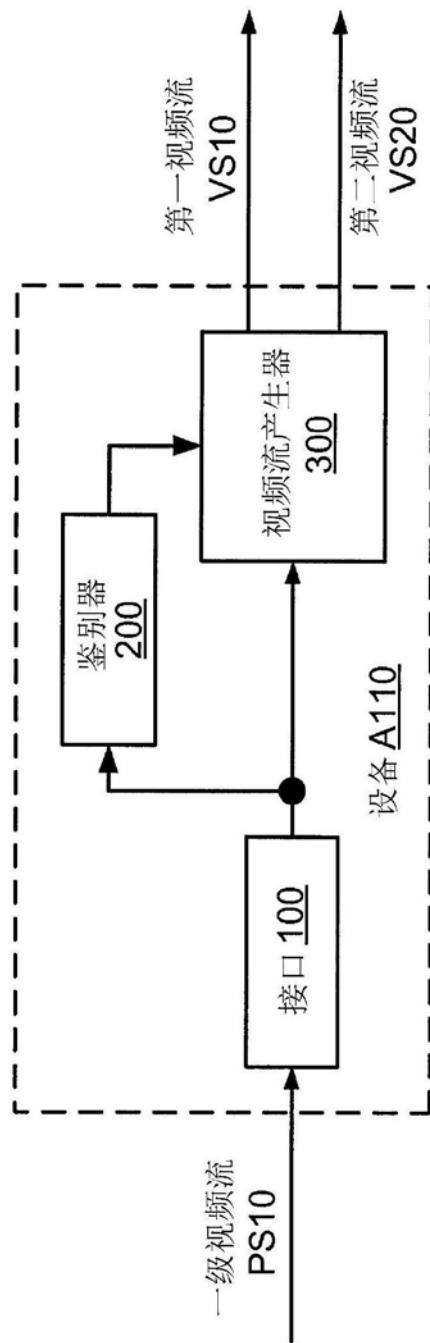


图37C

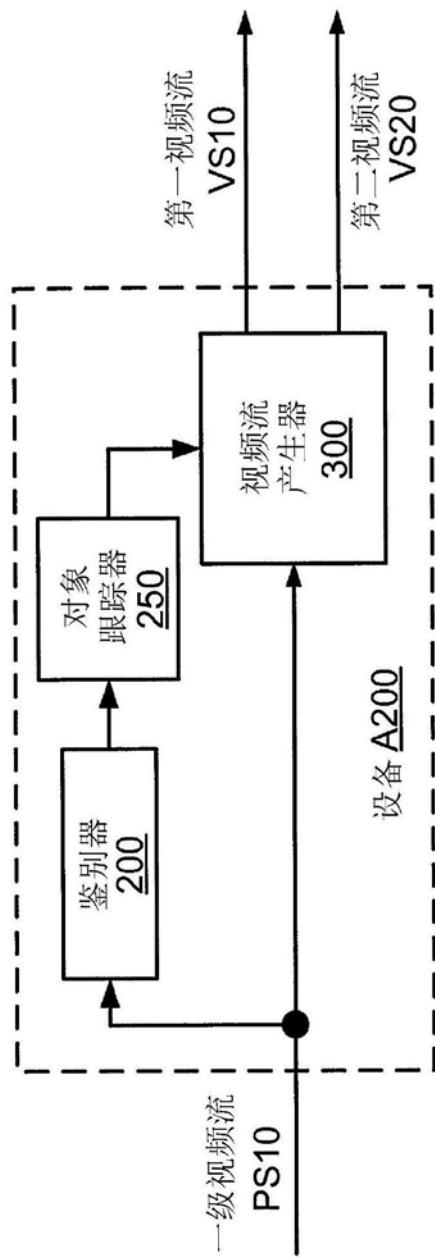


图38A

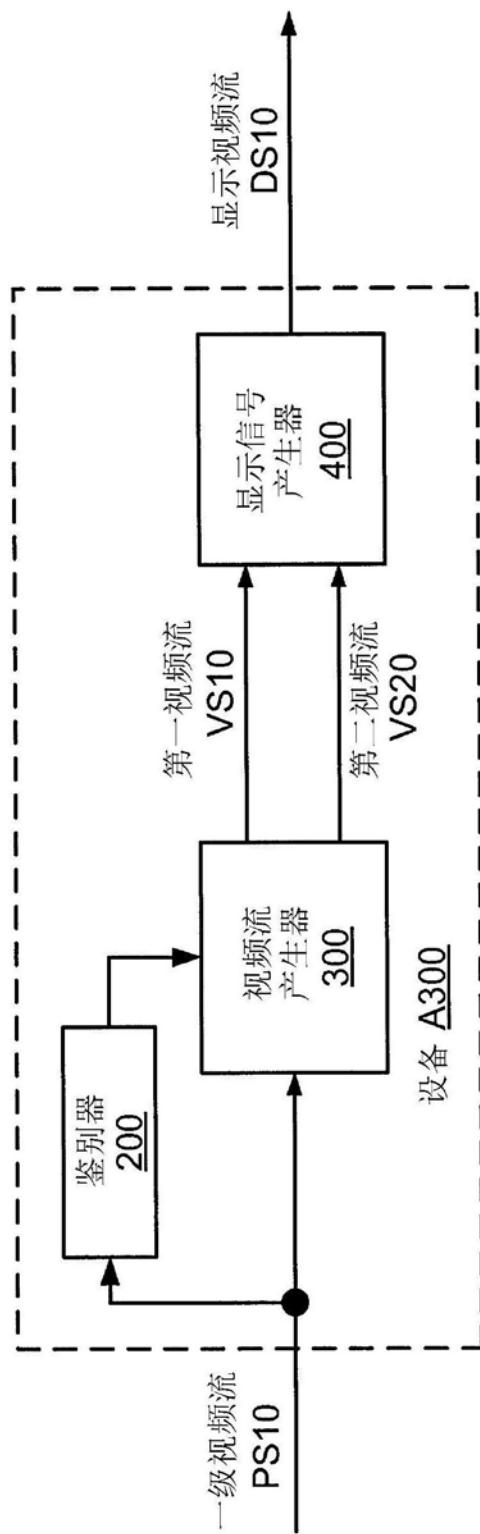


图 38B

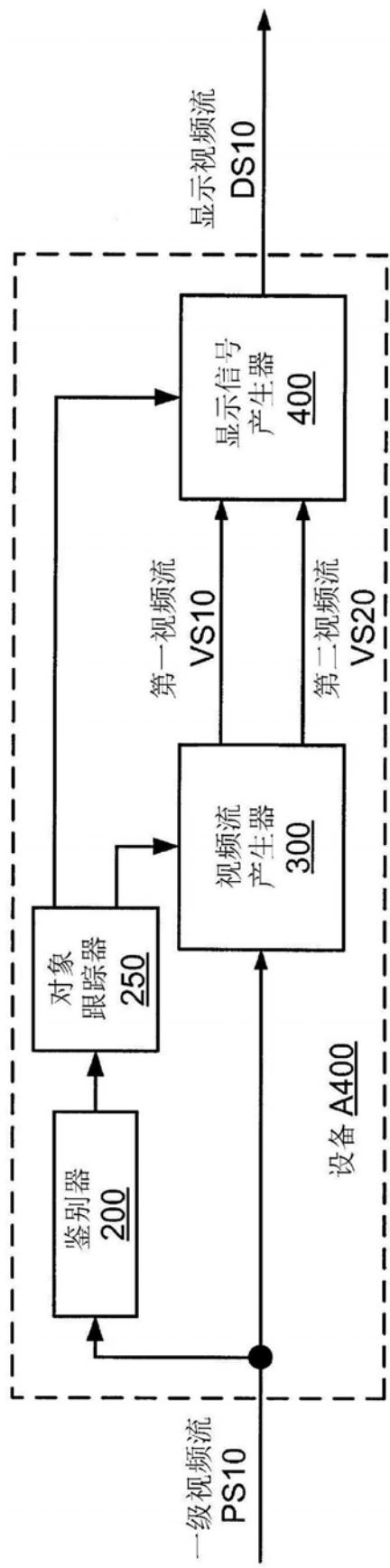


图39A

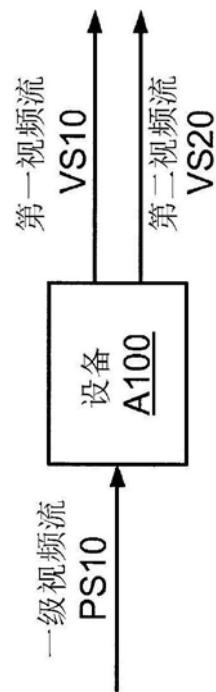


图39B

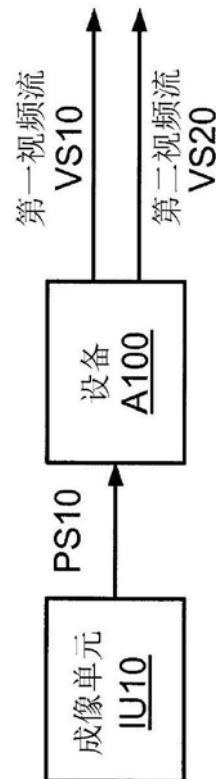


图39C

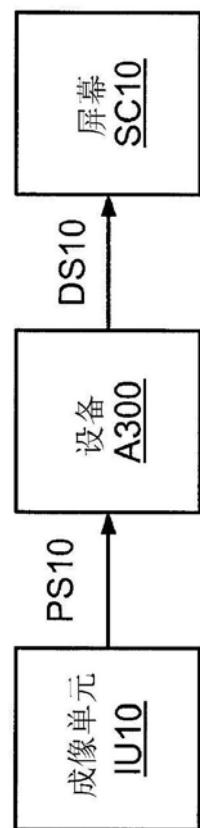


图39D

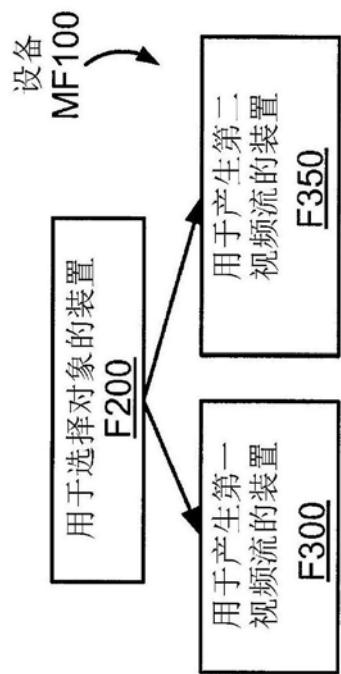


图39E

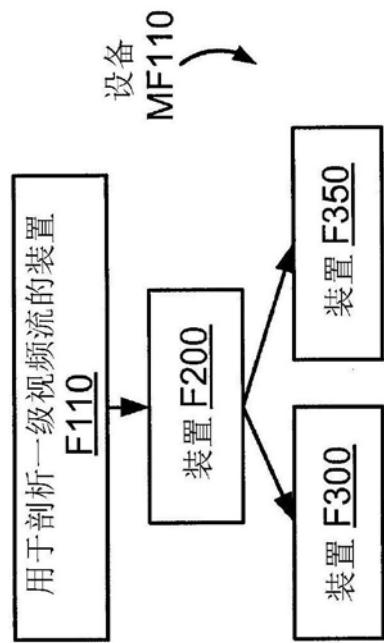


图40A

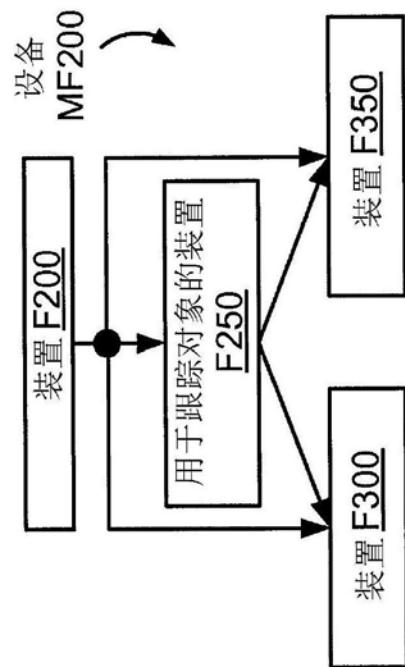


图40B

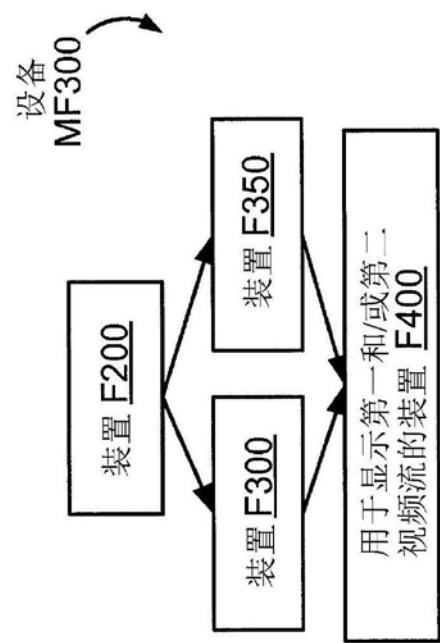


图40C

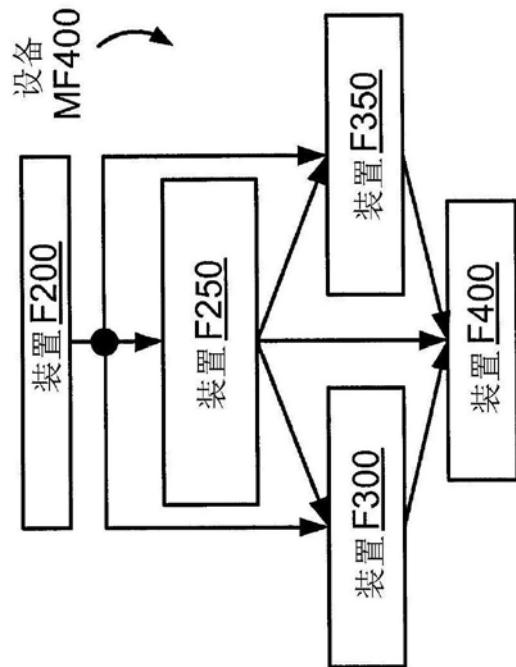


图40D

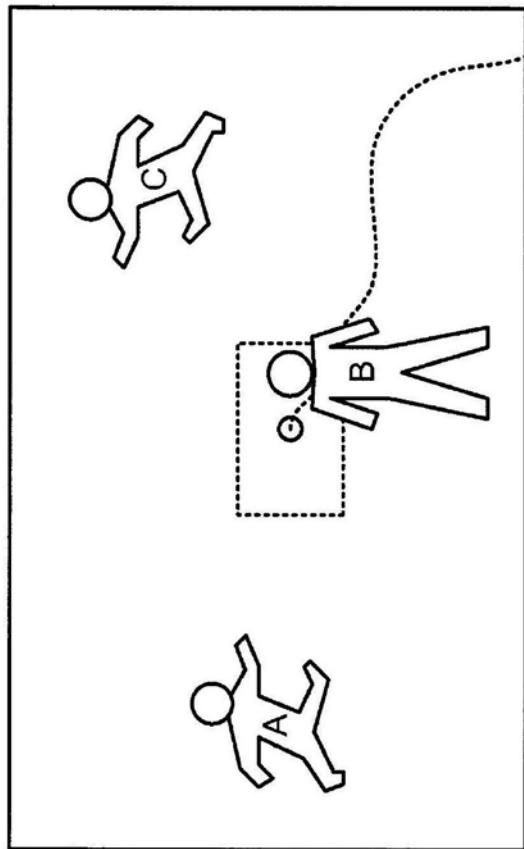


图41A

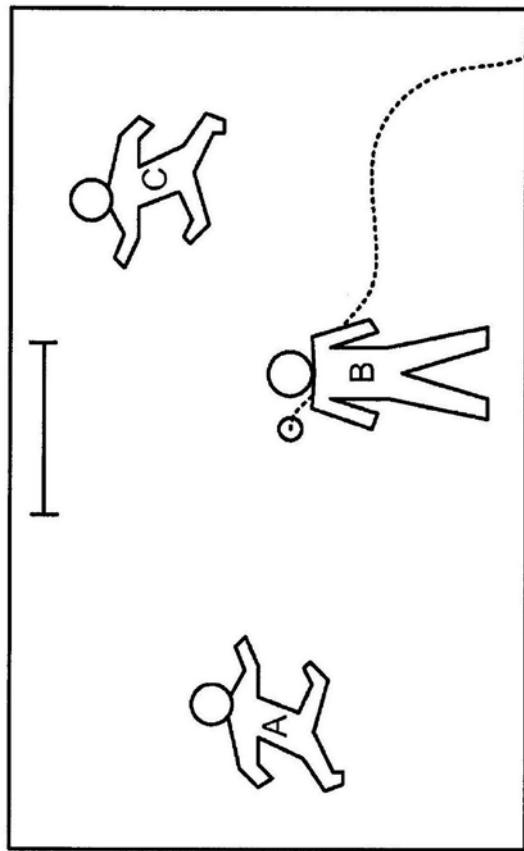


图41B

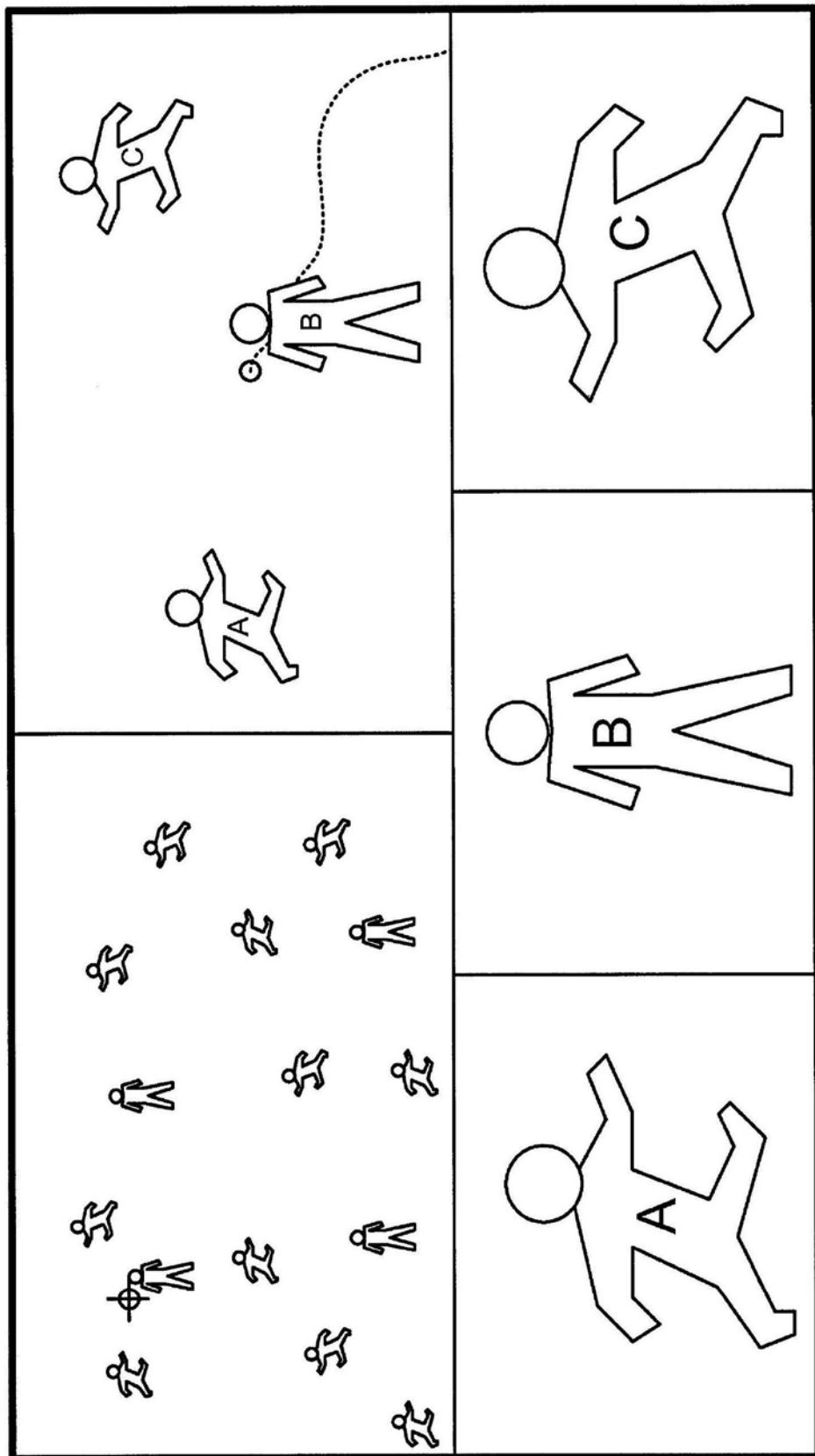


图42

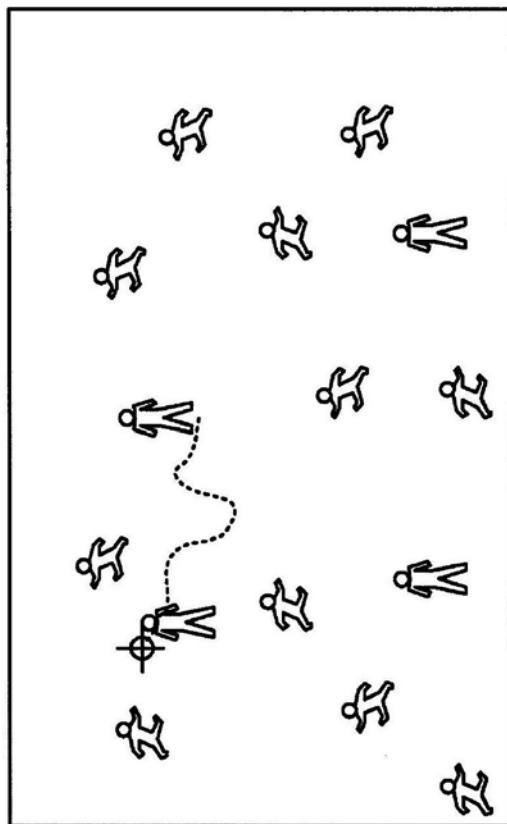


图43A

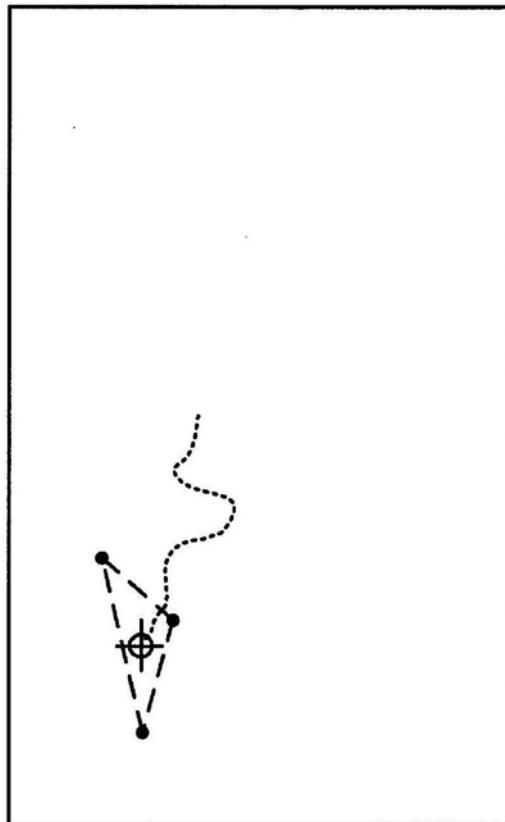


图43B

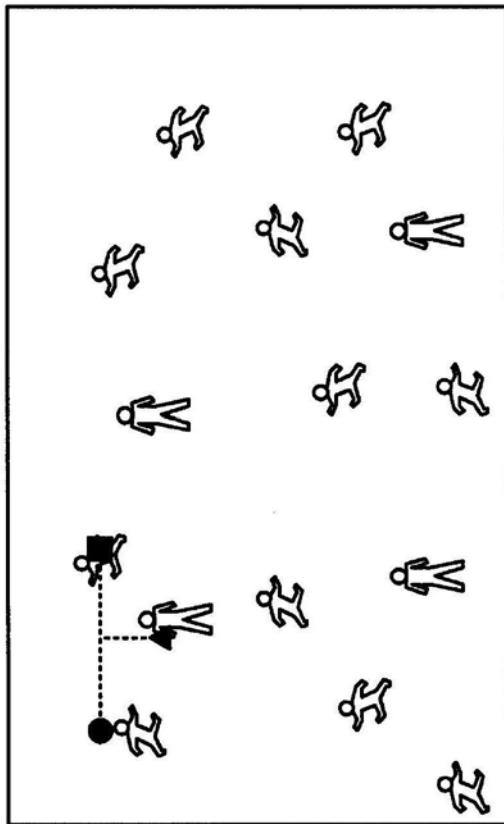


图43C