



(12)发明专利

(10)授权公告号 CN 105900084 B

(45)授权公告日 2019.12.31

(21)申请号 201480069156.3

(22)申请日 2014.12.19

(65)同一申请的已公布的文献号
申请公布号 CN 105900084 A

(43)申请公布日 2016.08.24

(30)优先权数据
61/919,627 2013.12.20 US
62/036,502 2014.08.12 US
62/063,498 2014.10.14 US
14/576,006 2014.12.18 US

(85)PCT国际申请进入国家阶段日
2016.06.17

(86)PCT国际申请的申请数据
PCT/US2014/071606 2014.12.19

(87)PCT国际申请的公布数据
W02015/095762 EN 2015.06.25

(73)专利权人 高通股份有限公司
地址 美国加利福尼亚州

(72)发明人 高大山 钟辛
安纳桑帕玛纳翰·埃拉山尼帕莱·
坎哈达
史蒂文·道格拉斯·拉韦尔

(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

(51)Int.Cl.
G06F 16/732(2019.01)
G06F 16/78(2019.01)
G06F 16/783(2019.01)
G06K 9/00(2006.01)
G06T 7/20(2017.01)

(56)对比文件
CN 102870123 A,2013.01.09,
US 2010082585 A1,2010.04.01,
CN 102207966 A,2011.10.05,
CN 101571875 A,2009.11.04,
CN 103186790 A,2013.07.03,
CN 101276410 A,2008.10.01,
Gulrukh Ahanger,Dan Benson,etc.Video
query formulation.《PROCEEDINGS OF SPIE》
.1995,pages 280-291.
Gulrukh Ahanger,Dan Benson,etc.Video
query formulation.《PROCEEDINGS OF SPIE》
.1995,pages 280-291.

审查员 于亚男

权利要求书2页 说明书47页 附图61页

(54)发明名称

用于图像检索的系统、方法及设备

(57)摘要

一种图像检索的方法包含获得识别多个选定对象的信息且选择多个候选几何布置中的一者。此方法还包含通过至少一个处理器且响应于所述选择而识别多个数字图像中的描绘根据所述选定候选几何布置而布置的所述多个选定对象的至少一个数字图像。

方法 M500



1. 一种图像检索的方法,所述方法包括:
在处理器处获得识别多个选定对象的信息;
在所述处理器处接收多个候选几何布置中的一几何布置的选择,其中每个候选几何布置表示对象相对于参考位置的分布;及
通过至少一个处理器响应于所述选择而识别多个数字图像中的至少一个数字图像,所述至少一个数字图像描绘根据经选定的所述几何布置而布置的经识别的所述多个选定对象。
2. 根据权利要求1所述的图像检索的方法,其中所述所获得的信息包含与所述多个选定对象相关联的至少一个标记,且其中识别所述至少一个数字图像包含在与所述多个数字图像相关联的元数据内搜索所述至少一个标记。
3. 根据权利要求1所述的图像检索的方法,其中识别所述至少一个数字图像包括在与所述多个数字图像相关联的元数据内搜索与所述候选几何布置相关联的标记。
4. 根据权利要求1所述的图像检索的方法,其中所述至少一个数字图像包含来自第一视频文件的多个帧及来自第二视频文件的多个帧。
5. 根据权利要求1所述的图像检索的方法,其中使用语音辨识或触摸屏中的至少一者来执行获得识别所述多个选定对象的信息。
6. 根据权利要求1所述的方法,其中所述参考位置是特定位置空间中的坐标位置。
7. 一种存储指令的非暂时性计算机可读介质,所述指令在由处理器执行时致使所述处理器:
获得识别多个选定对象的信息;
接收多个候选几何布置中的一几何布置的选择,其中每个候选几何布置表示对象相对于参考位置的分布;及
响应于所述选择而识别多个数字图像中的至少一个数字图像,所述至少一个数字图像描绘根据经选定的所述几何布置而布置的经识别的所述多个选定对象。
8. 根据权利要求7所述的非暂时性计算机可读介质,其中所述所获得的信息包含与所述多个选定对象相关联的至少一个标记,且
其中所述介质存储有由所述处理器执行时致使所述处理器在与所述多个数字图像相关联的元数据内搜索所述至少一个标记的指令。
9. 根据权利要求7所述的非暂时性计算机可读介质,其中所述介质存储有由所述处理器执行时致使所述处理器在与所述多个数字图像相关联的元数据内搜索与所述候选几何布置相关联的标记的指令。
10. 根据权利要求7所述的非暂时性计算机可读介质,其中所述至少一个数字图像包含来自第一视频文件的多个帧及来自第二视频文件的多个帧。
11. 根据权利要求7所述的非暂时性计算机可读介质,其中使用语音辨识或触摸屏中的至少一者来获得识别所述多个选定对象的所述信息。
12. 一种用于图像检索的设备,所述设备包括:
用于获得识别多个选定对象的信息的装置;
用于接收多个候选几何布置中的一几何布置的选择的装置,其中每个候选几何布置表示对象相对于参考位置的分布;及

用于响应于所述选择而识别多个数字图像中的至少一个数字图像的装置,所述至少一个数字图像描绘根据经选定的所述几何布置而布置的经识别的所述多个选定对象。

13.根据权利要求12所述的设备,其中所述所获得的信息包含与所述多个选定对象相关联的至少一个标记,且

其中所述设备包含用于在与所述多个数字图像相关联的元数据内搜索所述至少一个标记的装置。

14.根据权利要求12所述的设备,其中所述设备包括用于在与所述多个数字图像相关联的元数据内搜索与所述候选几何布置相关联的标记的装置。

15.根据权利要求12所述的设备,其中所述至少一个数字图像包含来自第一视频文件的多个帧及来自第二视频文件的多个帧。

16.根据权利要求12所述的设备,其中所述用于获得识别所述多个选定对象的信息的装置经配置以使用语音辨识或触摸屏中的至少一者来获得所述信息。

17.一种用于图像检索的设备,所述设备包括:

鉴别器,其经配置以获得识别多个选定对象的信息并且接收多个候选几何布置中的一几何布置的选择,其中每个候选几何布置表示对象相对于参考位置的分布;及

搜索引擎,其经配置以响应于所述选择而识别多个数字图像中的至少一个数字图像,所述至少一个数字图像描绘根据经选定的所述几何布置而布置的经识别的所述多个选定对象。

18.根据权利要求17所述的设备,其中所述所获得的信息包含与所述多个选定对象相关联的至少一个标记,且

其中所述搜索引擎经配置以在与所述多个数字图像相关联的元数据内搜索所述至少一个标记。

19.根据权利要求17所述的设备,其中所述搜索引擎经配置以在与所述多个数字图像相关联的元数据内搜索与所述候选几何布置相关联的标记。

20.根据权利要求17所述的设备,其中所述至少一个数字图像包含来自第一视频文件的多个帧及来自第二视频文件的多个帧。

21.根据权利要求17所述的设备,其中所述鉴别器经配置以使用语音辨识或触摸屏中的至少一者来获得识别所述多个选定对象的所述信息。

用于图像检索的系统、方法及设备

[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日申请的第61/919,627号美国临时专利申请案、2014年8月12日申请的第62/036,502号美国临时专利申请案、2014年10月14日申请的第62/063,498号美国临时专利申请案及2014年12月18日申请的第14/576,006号美国非临时专利申请案的优先权,这些专利申请案的内容明确地以全文引用的方式并入本文中。

技术领域

[0003] 本发明大体上涉及图像检索。

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] 一种根据一般配置的图像检索方法包含:获得识别多个选定对象的信息及选择多个候选几何布置中的一者。此方法还包含通过至少一个处理器且响应于所述选择而识别多个数字图像中的描绘根据选定候选几何布置而布置的多个选定对象的至少一个数字图像。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。

[0006] 一种根据一般配置的用于图像检索的设备包含用于获得识别多个选定对象的信息的装置及用于选择多个候选几何布置中的一者的装置。此设备还包含用于识别多个数字图像中的描绘根据选定候选几何布置而布置的多个选定对象的至少一个数字图像的装置。

[0007] 一种根据另一一般配置的用于图像检索的设备包含鉴别器,其经配置以获得(A)识别多个选定对象的信息及(B)多个候选几何布置中的一者的选择。所述方法还包含搜索引擎,其经配置以识别多个数字图像中的描绘根据选定候选几何布置而布置的多个选定对象的至少一个数字图像。

[0008] 由所揭示的实施例中的至少一者提供的一个特定优点是使用移动装置上的图像跟踪及帧检索的改善的用户体验。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

- [0009] 图1是描绘电子装置的框图。
- [0010] 图2A是描绘对象及检测模块的特定说明性实施例的框图。
- [0011] 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。
- [0012] 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。
- [0013] 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。
- [0014] 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。
- [0015] 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。
- [0016] 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。
- [0017] 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。
- [0018] 图9是描绘平滑模块的特定说明性实施例的框图。
- [0019] 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。
- [0020] 图11A展示根据一般配置的产生元数据的方法M100的流程图。
- [0021] 图11B展示方法M100的实施方案M110的流程图。
- [0022] 图11C展示方法M100的实施方案M120的流程图。
- [0023] 图11D展示方法M110及M120的实施方案M130的流程图。
- [0024] 图12展示选择三个对象的用户动作的序列的实例。
- [0025] 图13展示选择三个对象的用户动作的序列的另一实例。
- [0026] 图14展示帧内的对象及对象的增强显示的实例。
- [0027] 图15A到C展示使用定向敏感装置来操纵选择点的实例。
- [0028] 图16展示对象的图像之间的距离与实际距离进行比较之间的差异的实例。
- [0029] 图17A展示将对象的质量中心投影到对象的底部界线上以获得对应确定位置的实例。
- [0030] 图17B展示将链接点投影到地平面中的确定位置的实例。
- [0031] 图18A及18B展示像素协调空间中的距离与地平面中的距离之间的对应关系。
- [0032] 图19A展示篮球场的俯视图,且图19B到D展示球场的位置空间的非均匀分割的若干实例。
- [0033] 图20是描绘使用对象跟踪的视频处理的另一特定实施例。
- [0034] 图21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。
- [0035] 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。
- [0036] 图23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。
- [0037] 图24描绘使用直角坐标编码对象位置的元数据的特定实施例。
- [0038] 图25描绘使用极坐标编码对象位置的元数据的特定实施例。
- [0039] 图26是用于编码元数据的方法的特定说明性实施例的流程图。
- [0040] 图27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28是使用对象跟踪描绘视频处理的另一特定实施例。

- [0042] 图29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30A展示根据一般配置的设备A100的框图。
- [0044] 图30B展示设备A100的实施方案A110的框图。
- [0045] 图30C展示设备A100的实施方案A120的框图。
- [0046] 图30D展示设备A110及A120的实施方案A130的框图。
- [0047] 图31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。
- [0048] 图32A展示根据一般配置的设备MF100的框图。
- [0049] 图32B展示设备MF100的实施方案MF110的框图。
- [0050] 图32C展示设备MF100的实施方案MF120的框图。
- [0051] 图32D展示设备MF110及MF120的实施方案MF130的框图。
- [0052] 图33A展示一组模板的实例,且图33B展示对模板的修改的实例。
- [0053] 图34A展示选定对象的不同阵型的实例。
- [0054] 图34B展示非唯一映射的实例。
- [0055] 图34C展示选定对象的唯一阵型。
- [0056] 图35展示一组决策度量的实例。
- [0057] 图36展示一组决策度量的另一实例。
- [0058] 图37A展示方法M200的实施方案M200的流程图。
- [0059] 图37B到D分别展示方法M110、M120及M130的实施方案M210、M220及M230的流程图。
- [0060] 图38A展示根据一般配置的用于产生元数据的设备A100的框图。
- [0061] 图38B展示设备A100的实施方案A110的框图。
- [0062] 图38C展示设备A100的实施方案A120的框图。
- [0063] 图38D展示设备A110及A120的实施方案A130的框图。
- [0064] 图38E展示设备A100的实施方案A200的框图。
- [0065] 图38F、38G及39A分别展示设备A110、A120及A130的实施方案A210、A220及A230的框图。
- [0066] 图39B展示设备A100的实施方案A300的框图。
- [0067] 图39C及39D展示设备A200及A230的实施方案A310及A330的框图。
- [0068] 图40A展示根据一般配置的用于产生元数据的设备MF100的框图。
- [0069] 图40B展示设备MF100的实施方案MF110的框图。
- [0070] 图40C展示设备MF100的实施方案MF120的框图。
- [0071] 图40E展示设备MF100的实施方案MF200的框图。
- [0072] 图40F、40G及41A分别展示设备MF110、MF120及MF130的实施方案MF210、MF220及MF230的框图。
- [0073] 图41B展示设备MF100的实施方案MF400的框图。
- [0074] 图41C展示设备MF100的实施方案MF300的框图。
- [0075] 图41D及41E展示设备MF200及MF230的实施方案MF310及MF330的框图。
- [0076] 图42展示描绘使用对象跟踪的视频处理的特定实施例。
- [0077] 图43展示编码对象位置的元数据的特定实施例。
- [0078] 图44展示说明视频处理方法的特定实施例的流程图。

- [0079] 图45A展示根据一般配置的方法M500的流程图。
- [0080] 图45B展示根据一般配置的设备A500的框图。
- [0081] 图45C展示根据一般配置的设备MF500的框图。

具体实施方式

[0082] 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。

[0083] 除非通过其上下文明确限制,否则在本文中术语“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”。

[0084] 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”。

[0085] 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的整数数量。

[0086] 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。

[0087] 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。

[0088] 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。

[0089] 在一个配置中,显示器为从物理触摸(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限界框。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。

[0090] 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。

[0091] 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。

[0092] 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和对对象检测器108,以为运动跟踪器106和对对象检测器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。

[0093] 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。

[0094] 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。

[0095] 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。

[0096] 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。

[0097] 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。

[0098] 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。

[0099] 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。

[0100] 电子装置102可包含视频处理模块137。视频处理模块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。

[0101] 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。

[0102] 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。

[0103] 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。

[0104] 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)。

[0105] 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。

[0106] 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。

[0107] 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。

[0108] 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。

[0109] 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。

[0110] 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,以确保目标对象为下一视频帧)。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。

[0111] 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。

[0112] 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(即,目标对象可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)的情况下,目标对象可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。

[0113] 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)。

[0114] 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,限界框)以形成经组合跟踪结果(例如,限界框)并计算检测和跟踪置信度值256。

[0115] 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对象检测。

[0116] 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。

[0117] 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。

[0118] 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。

[0119] 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。

[0120] 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。

[0121] 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)。

[0122] 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。

[0123] 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。

[0124] 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。

[0125] 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。

[0126] 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。

[0127] 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 $\hat{x}_{t+k} = x_{t+k}$ 。

[0128] 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。

[0129] 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。

[0130] 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子集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。

[0131] 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)。

[0132] 搜索范围可由用于特定位置的窗口大小子集表示,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,如果最近视频帧并未检测到目标对象(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而不限搜索范围。

[0133] 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。

[0134] 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。

[0135] 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。

[0136] 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:

$$[0137] \quad X_t = W * X'_t + (1 - W) * X_{t-1} \quad (1)$$

[0138] 其中X'_t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'_t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。

[0139] 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_{location}=0.8且W_{size}=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。

[0140] 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平滑权重分别设定成W_{location}=0.65且W_{size}=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。

[0141] 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)。

[0142] 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:

$$[0143] \quad x_k = F_k x_{k-1} + w_k \quad (2)$$

$$[0144] \quad z_k = H x_{k-1} + v_k \quad (3)$$

[0145] 其中 x_{k-1} 为时间 $k-1$ 处的先前状态, x_k 为由 $x_k = [x, y, \dot{x}, \dot{y}]$ 界定的当前状态,其中 (x, y) 为限界框中心位置, (\dot{x}, \dot{y}) 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:

[0146]

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

[0147]

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

[0148] 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, $w_k \sim N(0, Q)$):

[0149]

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

[0150] 其中 σ_1 为可调参数。类似地, w_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, $v_k \sim N(0, R)$):

[0151]

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

[0152] 其中 σ_2 为可调参数。

[0153] 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。

[0154] 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)。

[0155] 数字图像(例如,视频流的帧)描绘可包含一或多个对象的场景。通常,所述对象是物理及有形对象(例如,人)。在图12中,例如,场景中的对象包含三个人、星星及树。

[0156] 可能需要从一或多个数字图像(例如,视频流的帧)产生具有表示选定对象随时间在一或多个图像内的位置的信息的相关联的流或文件。在一个实例中,此类信息包含关注点(例如,选定对象的质量中心)的绝对位置,其可随时间改变。在另一实例中,此类信息包含选定对象相对于关注点(例如,选定对象的质量中心)的位置的位置,其可随时间改变。关注点的位置(例如,选定对象的位置或多个对象的质量中心的位置)可被编码为元数据(例如,一或多个相应的位置码簿中的一或多个索引,例如码字)。

[0157] 图11A展示根据第一配置的产生元数据的方法M100的流程图,所述方法包含任务T100及T300。任务T100使用至少一个数字图像选择多个对象,其中视频流相对于物理空间描绘对象。基于指示位置空间中的选定对象的位置的信息,任务T300产生识别选定对象的多个候选几何布置中的一者的元数据。任务T100及T300可例如由图1的电子装置102、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。

[0158] 任务T100使用至少一个数字图像选择多个对象,其中所述至少一个数字图像相对于物理空间描绘所述对象。可响应于用户的直接动作而执行所述选择。在此情况下,屏幕可用于向用户显示至少一个数字图像,且用户可输入指示以在显示器中出现的对象中选择特定所要的对象。图11B展示方法M100的实施方案M110的流程图,其包含显示至少一个数字图像的任务T50。如果在触摸屏上执行显示,那么用户可通过触摸屏幕以选择所要的对象内的点而指示选择。

[0159] 所述至少一个数字图像可为视频流的一或多个帧。视频流描述一连串帧,其中每一帧表示像素协调空间中的图像。视频流通常包含可用于恢复来自所述流的帧的其它信息(例如,每一帧的对应的帧起始代码或包及帧结束代码或包)。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(其可对可见及/或其它波长敏感)产生、从另一装置串流传输,或由解码器(例如,从存储在磁性或光学媒体上的信息)产生,且可处于压缩或未经压缩的形式。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相机捕获的图像的视频流(例如,微软Kinect)。此类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在触摸屏上显示。视频流可为实况视频流、延时视频流,或从存储装置检索(例如,经预先记录)。

[0160] 视频流是数字图像(“帧”)的时间序列。在一个实例中,以特定帧速率捕获图像的序列(例如,以供以同一帧速率呈现)。视频流可包含指示每一帧的开始及结束的信息。在一个实例中,视频流包含每一帧的第一像素之前的帧起始代码及每一帧的最后像素之后的帧

结束代码。在另一实例中,所述视频流包含载运所述视频流的帧的像素信息的一或多个数据信号,及指示视频流的每一帧的开始及结束的一或多个同步信号。取决于主要视频流的特定格式,主要视频流可包含额外信息(例如,标头、元数据)。主要视频流还可能在帧开始及结束代码之间包含除帧数据(即,像素值)之外的数据。此类其它数据可包含(例如)传感器配置细节及/或图像统计值。

[0161] 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11C展示方法M100的实施方案M120的流程图,其包含解析视频流以产生至少一个数字图像的任务T30。或者,另一过程可执行以解析所述视频流(例如,用于显示),使得方法M100(例如,在存储器中的缓冲器中)可在任何时间使用视频流的一或多个帧。图11D展示方法M110及M120的实施方案M130的流程图。

[0162] 此类缓冲器(例如,存储器缓冲器110)通常一次能够存储至少一个全帧。对主要视频流的解析可经配置以在多个帧缓冲器之间交替。例如,一个此类帧缓冲器可存储全帧以用于处理,而序列中的后面的帧被提取并存储到另一帧缓冲器。或者,任务T30或解析过程可经配置以将解析的帧存储到循环缓冲器,使得在提取新帧时更新缓冲器(且覆写较旧的帧)。

[0163] 解析所述主要视频流可包含多路分用操作以将主要视频流与一或多个相关联的音频流及/或元数据流分离。或者,此操作可在任务T30或解析过程的上游执行(例如,主要视频流可由多路分用器提供)。

[0164] 图12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13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12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显示的视野(例如,执行夹捏动作以加宽视野,且执行缩放动作以使视野变狭窄)。除操纵触摸屏以外,用于直接用户选择的其它可能的模态包含致动一或多个按钮或其它开关及/或手势辨识。作为用户进行此类直接动作的替代方案,任务T100可通过向来自至少一个数字图像(例如,视频流的一或多个帧)的信息应用一或多个预定准则而自动执行所述选择。此类准则可包含(例如)球员的球衣号码的辨识、面部辨识及/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以响应于任务T100中的对象的选择而跟踪对象。例如,可实施如下文所描述的位置确定任务T200以还跟踪视频流内的一或多个选定对象(例如,如上文参考对象跟踪与检测模块204及/或对象跟踪与检测模块804所描述)。或者,对象跟踪可在任务T100之前开始。在此情况下,可实施任务T100以用于从一组被跟踪对象当中进行选择。例如,用户可从被跟踪对象的显示当中进行选择。可以用于此类对象跟踪(例如,通过任务T200)的算法的实例包含卢卡斯-金出武雄方法(及变体,例如金出武雄-卢卡斯-托马西跟踪器)及平均移位跟踪。对象的跟踪可包含将对象从主要视频流的帧的其余部分分段,但还可在不存在此类分段的情况下执行对象跟踪。

[0169] 对象从数字图像(例如,视频流的帧)的其余部分的分段可用于支持对象相对于帧的其余部分的增强显示(例如,显出轮廓或勾勒出轮廓)。可响应于任务T100中的对象的选择而执行分段(例如,通过显示任务T50)。此分段可用于通过确认对象的选择而向用户提供视觉反馈。还可在选择开始之前执行分段。在此情况下,可从已经分段的一组对象当中进行任务T100中的选择(例如,其中分段的对象被增强的显示)。

[0170] 图14中的左图展示来自视频流的描绘包含对象的场景的帧的两个实例。可在任务T100中的选择之前或响应于此类选择而执行对象的增强显示(例如,通过显示任务T50)增强显示的一个实例是对象的显出轮廓,其可以与背景形成对比的色彩显示(例如,黑色或黄色,如图14的中心图中所展示)。此对比色可对于所有对象都相同或可在对象之间不同(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(其同样可以对比色完成)包含勾勒对象的轮廓并且在对象处、内或周围显示图形元素(例如,对象的顶点或质心处的三角形、对象周围的椭圆,如图14的右侧图中所展示等)。

[0171] 将视频流的显示用作用户对对象的直接选择的背景对于缓慢移动对象可为适当的。然而,如果所要的对象中的任一者在显示内快速移动,此布置可让用户受挫。在此类快速移动的情况下,可能需要替代地显示视频流的单个帧以作为用户对对象的直接选择的背景。如上文所描述,可从其中帧中的对象被分段及增强的显示进行选择。如果流显示选择及帧显示选择两者都支持,那么这些显示模式之间的选择可由用户及/或自动执行。对于自动选择,可使用例如光流、运动向量、帧之间的像素差等量度来量化流中的运动程度。在此情

况下,在量度的值高于阈值时可选择帧显示模式,且在量度的值低于阈值时可选择流显示模式。

[0172] 基于指示位置空间中的选定对象的位置的信息,任务T300产生识别选定对象的多个候选几何布置中的一者的元数据。位置空间可为(例如,视频流的)至少一个数字图像的像素协调空间,例如数字图像(例如,视频流的帧)的(x,y)笛卡尔坐标空间。通常将数字图像(例如,视频帧)的像素坐标空间的原点指定为图像的左上方或左下方隅角。选定对象的位置可例如为用户在其处触摸屏幕以选择对象的位置。或者,可由对象跟踪器(例如,如本文中所描述的对象跟踪与检测模块204或804)来提供选定对象的位置。

[0173] 几何布置是选定对象的空间中的布置。例如,几何布置可为由选定对象的位置(例如,作为多边形的顶点)描述的形状。可实施任务T300以通过将一或多个度量应用到对象位置而识别候选几何布置中的适当一者。在此情况下,任务T300可使用决策树根据度量值而选择候选者。所产生的元数据还可包含额外信息,例如以下各者中的任一者:将应用于候选者的比例因子及/或方面因素、候选者的定向方向,及指示候选者是否围绕轴翻转的镜旗标。

[0174] 可实施任务T300以确定多个候选几何布置中的哪一者最类似于由选定对象的位置描述的形状。图35展示可借以实施任务T300以使由三个位置A、B、C描述的形状与标记为“等距”、“近-远-远”及“近-近-远”的一组三个候选几何布置中的一者匹配的另一实例。

[0175] 在此实例中,点之间的三个距离(即,AB、BC及AC)经过计算且通过量值被排序以确定其中的最小值、中间值及最大值。将最小距离与是最大距离的三分之二的值进行比较。如果最小距离更大,那么任务T300选择候选“等距”,如由图35的左上方的方框指示。如果最小距离更小,那么将中间距离与是最小距离及最大距离的值的二分之一的值进行比较。如果中间距离更大,那么任务T300选择候选“近-远-远”,如由图35的底部处的方框指示。如果中间距离更小,那么任务T300选择候选“近-近-远”,如由图35的右上方的方框指示。

[0176] 图35还展示落到决策边界的三个情况,其可视需要由任务T300指派给邻近候选者中的任一者。在此情况下,识别适当的候选者的元数据可为具有三个可能值中的一者的参数,每一可能值指示三个候选几何布置中的对应一者。所述元数据还可指示将应用于候选者的定向(例如,作为旋转)。在一个实例中,所述定向被指示为从几何布置的中心到与最小距离相对的点的线的方向。在另一实例中,所述定向被指示为从几何布置的中心到对应于选定对象中的特定一者的点的线的方向。

[0177] 图37A展示包含任务T200的方法M200的实施方案M200的流程图,其获得指示位置空间中的选定对象的位置的信息。可实施任务T200以响应于任务T100中的选择而执行。位置空间可为(例如,视频流的)至少一个数字图像的像素坐标空间,例如数字图像(例如,视频流的帧)的(x,y)笛卡尔坐标空间。数字图像(例如,视频帧)的像素坐标空间的原点通常被指定为图像的左上方或左下方隅角。图37B到D分别展示包含任务T200的方法M110、M120及M130的实施方案M210、M220及M230的流程图。

[0178] 显示的像素坐标空间可与位置空间相同,或位置空间可包含显示的像素坐标空间。例如,显示的帧可包含视频流的对应帧的一部分(例如,响应于使视野变狭窄的缩放动作)。或者,显示的像素坐标空间可包含位置空间。例如,视频流可显示在小于显示的帧的显示的窗口中,其中其它内容显示在显示的另一窗口中。

[0179] 可实施任务T200以将选定对象的位置确定为像素坐标空间中的选定对象的图像的位置。例如,所确定的位置可为对象的质量中心。如果采用恒定密度,那么所述质量中心与质心相同,但密度中的差异还可由例如色彩及/或透明度中的差异指示。所确定的位置的其它实例包含对象的顶部或底部像素、对象的质量中心到对象的顶部或底部的垂直投影,或对象的特定特征的位置。图17A展示将对象的质量中心C10投影到对象的底部界线上以获得对应确定的位置DP10的实例。

[0180] 或者或另外,可实施任务T200以确定多个对象的阵型的参考位置。例如,可实施任务T200以将参考位置计算为选定对象的阵型的位置。在此情况下,所确定的参考位置可为阵型的质量中心(其中不同对象具有同一或不同密度)、对象中的顶部或底部像素、阵型的质量中心到阵型的顶部或底部的垂直投影,或对象中的特定一者的位置。在此情况下,任务T200可将选定对象的位置确定为相对于参考位置的位置(例如,使用参考位置作为原点)。

[0181] 选定对象可彼此物理地分开,使得他们不是共同对象的部分。例如,对象中的至少一者可能够独立于其它对象移动。在特定实例中,对象中的一或多者是不同的人、动物及/或车辆。

[0182] 视频流或显示的像素坐标空间可为二维或三维的。三维显示的实例包含立体(例如,左/右)及体积显示。在此位置空间中的对象的用户选择可包含例如使用手套、手势(例如,使定向敏感装置倾斜)或压敏触摸屏在三个维度上操纵选择点。

[0183] 或者,位置空间可为物理空间(例如,场景空间)。在一个此类实例中,所确定的参考位置是基于使用包含深度信息的视频流(例如,来自结构光成像器或其它深度相机(例如,微软Kinect)的视频流)选择的至少一个点。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在触摸屏上显示。所确定的参考位置可基于物理空间中的每一选定对象的图像的位置,或可为选定对象的阵型的位置,如以上实例中(例如,质量中心)。在此位置空间中的对象的用户选择可包含例如使用手套、手势(例如,使定向敏感装置倾斜)或压敏触摸屏在三个维度上操纵选择点。

[0184] 图15A到C展示使用定向敏感装置在所显示的场景的深度维度中操纵选择点的实例。此定向敏感装置可例如为具有指示装置相对于重力轴(例如,穿过装置及地球中心的轴)的定向的一或多个定向传感器的移动装置。所述一或多个定向传感器可包含例如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)。

[0185] 在图15A到C中的每一者中,左图指示定向敏感装置的屏幕上的同一数字图像的显示,且右图指示装置相对于重力轴g的对应定向(以穿过显示的垂直线的横截面展示,其以粗体指示)。这些图中的每一者中显示的场景包含前景(即,最靠近捕获装置)中的对象、中间地面中的对象,及背景(即,最远离捕获装置)中的对象。在这些实例中的每一者中,由显示的左边处的滑块指示倾斜程度。在装置垂直于地面时,如图15A中所展示,倾斜程度较低且选择前景对象(如中左图中的突出显示指示)。在装置与地面水平时,如图15C中所展示,倾斜程度较高且选择背景对象(如左图中的突出显示指示)。在装置在这些定向之间与地面成四十五度角时,如图15B中所展示,倾斜程度处于中点,且选择中间地面中的对象(如由左图中的突出显示指示)。

[0186] 在另一实例中,可通过来自对象中或上的一或多个位置传感器(例如,GPS传感器)的信息来指示物理空间中的对象的位置。例如,所述对象可包含一或多个此类传感器(例

如,球)或穿戴一或多个此类传感器(例如,人或动物)。可使用触摸屏装置中的位置及/或定向传感器及/或对象特征(例如,色彩、球衣号码)的辨识来执行显示的像素坐标空间中的用户选定位置与物理空间中的对象的感测的位置的联系。

[0187] 另外或替代地,可从像素坐标空间中的对象的图像的位置投影物理空间中的对象的位置。显示平面中的对象图像的位置之间的空间关系可不与物理空间中的对应对象的位置之间的空间关系充分相关。在显示平面(例如,捕获视频流期间的焦平面)垂直于地平面时,此类差异可最高。

[0188] 图16展示与场景空间中的实际对象之间的距离进行比较的2-D像素坐标空间中的对象的图像之间的距离之间的差异的实例。上图A展示数字图像(例如,视频流的帧),且下图A展示所描绘的场景的左边的三个人的质心之间的距离。这些距离指示最左边的人与他们彼此相比离其他两个人中的每一者更远。上图B展示从正上方观看的同一场景,且下图B展示中间的人与他们彼此相比离其它两个人实际上远得多,使得下图A中指示的距离关系相对于场景空间中的实际距离关系是不准确的。

[0189] 物理空间可为二维空间,例如场景空间中的地平面。在一个此类实例中,位置空间是场景空间中的地平面,且像素坐标空间中点被投影到地平面中的对应点。地平面可例如是体育场,例如图16的图C中展示的篮球场。可能需要使用此类位置空间来支持对象位置在具有不同视野及/或从不同视角捕获的同一或类似空间中的类似事件的视频流中的一致性。

[0190] 可通过将场景空间中的链接点(例如,由来自一或多个传感器及/或深度相机的数据指示)向下投影到地平面中(或在适当时将链接点向上投影到地平面中)而执行使数字图像的像素坐标空间中的点与地平面中的对应点相关联。图17B展示链接点LP10(其可由如上文所论述的对象上的GPS传感器及/或其它位置传感器指示)到地平面中的确定的位置DP20的此类投影的实例。

[0191] 或者,可通过将来自像素坐标空间的点投影到如图18A及18B中所展示的地平面而执行使数字图像的像素坐标空间中的点与地平面中的对应点相关联,其展示在正交于地平面的平面中穿过捕获装置的焦轴的横截面)。如果捕获装置相对于地平面的定向角 α 是已知的(例如,从如上文所描述的一或多个定向传感器),那么可直接执行此类投影。或者,可从视频帧中的参考信息(例如,地平面参考点(例如,体育场的边界线、地平线(例如由光中的差异指示)、给对象定界的水平特征(例如,指示表演舞台的前面及后面的线)、或至少给水平特征定界的后面)推断定向角 α 。如图18B中所展示,通过定向角 α 使地平面中的距离 r 及 s 之间的比率与焦平面中的对应距离 p 及 q 之间的比率相关。可假设视野角度 b 是已知的,且焦平面中的距离 p 及 q 之间的比率等于显示空间中成像的对应距离(即,在那些距离相对于彼此沿着显示的像素坐标空间中的垂直线出现时)。

[0192] 在一个此类实例中,实施任务T200以使用捕获装置的定向及/或图像参考信息将显示平面映射到地平面(例如,如上文所描述)。在此实施方案中,任务T200通过将显示平面中对象的质量中心向下投影到对象的底部(例如,如图17A中所展示)而确定显示平面中的每一选定对象的位置,且通过将此显示位置映射到地平面中的对应位置而确定位置空间中的对象位置。在此情况下,根据上文描述的实例中的任一者,可实施任务T200以将所确定的参考位置计算为地平面中的对象位置的平均值,或计算为地平面中的对象位置的阵型的位置。

置。

[0193] 可实施任务T300以使元数据与至少一个数字图像(例如,与视频流)相关联。此元数据还可包含关于被跟踪对象的其它信息,例如选定对象的标记(例如,人的姓名)、选定对象中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及/或从所接收的至少一个图像或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、捕获的时间及/或位置)。

[0194] 任务T300的此实施方案可将元数据编码为所要的格式,例如XML(可扩展标记语言)、KLV(密钥长度值)、*.srt(SubRip文件格式)及*.vtt(网络视频文本跟踪格式)中的任一者。在一些格式总,元数据中的至少一些可包含在数字图像或视频流内(例如,在帧数据的为嵌入数据保留的一部分内)。或者,可实施任务T300以将元数据打包为单独文件或视频流同步的元数据流。

[0195] 可能需要实施方法M100以包含视频流上的任务T300的重复实例(例如,任务T200及T300的重复实例):对于视频流的每一帧,例如,或对于每一第n个帧(其中n为大于一的整数),或在指定事件之后(例如,跟踪任务对选定对象已移动的检测)。在此情况下,元数据文件或流内的元数据的项目可被加时戳以指示此项目与视频流的对应帧之间的所要的关联。在此情况下,任务T300还可包含将视频及元数据流多路复用(还可能与一或多个音频流一起)。

[0196] 对于其中含有将传输及/或存储的元数据的流的情况,可执行其它信息及/或格式化以符合特定流格式(例如,可将元数据及/或流打包为标头加上有效负载格式)。可实施任务T300以将含有元数据的流提供到可经配置以编码流的压缩器或其它编码器。可用于存储此流的存储文件格式的实例包含以下视频容器格式中的任一者:AVI、WMV、MOV、MKV、MPG及MP4。

[0197] 在一些应用中,选定对象的几何布置可为位置空间中的众多选定对象(例如,公共空间中的人群、大群动物、大群空中或水上交通工具或其它对象)的分布。任务T300可没有必要产生指示每一个别对象的位置的元数据,且选定对象的准确总数也可不必要的。

[0198] 在此情况下,可实施任务T300以确定多个候选几何布置中的哪一者最类似于位置空间中的选定对象的分布。例如,可能需要实施任务T300以确定最紧密近似相对于参考位置(例如,人群中的扬声器的位置)的选定对象的给定(例如,观测到的)分布的候选者。

[0199] 图36展示可借以实施任务T300以匹配相对于参考位置的众多对象的分布的一个实例,所述参考位置在图的右侧的十个候选者中的每一者的中心处指示。在此情况下,圆形位置空间被划分成十六个区,如图的左侧的图中所展示,且确定每一区中的对象的数目。

[0200] 图36的左边的四个大图说明可用于此选择的一组四个决策度量的一个实例:(A)内部区中的对象的数目与外部区中的对象的数目之间的比率,(B)最多人口的一半中的对象的数目与其它一半中的对象的数目之间的比率(或者,最多人口的一半中的对象的数目与对象的总数之间的比率),(C)最多人口的四分之一中的对象的数目与相反的四分之一中的对象的数目之间的比率,及(D)最多人口相反四分之一中的对象的数目与其它两个四分之一中的对象的数目之间的比率(或者,最多人口相反四分之一中的对象的数目与对象的总数之间的比率)。

[0201] 甚至在选定对象的数目较大时,可以非常低的计算复杂度计算位置空间中的对象

的特定分布的这四个度量的值。接着可使用决策树根据给定分布的这些度量的值在例如图36的右侧处的十个候选几何布置中进行选择。在此情况下,可实施任务T300以产生识别选定候选者并指示选定对象的总数的元数据。为进一步减小元数据的大小,可实施任务T300以产生指示选定对象的总数的近似值的元数据(例如,通过将总数除以指定量化因数且将结果舍入到最接近的整数)。

[0202] 还可实施任务T300以执行在几何布置中进行选择的额外实例中的任一者,如下文参考图33A、33B、34A到C及阵型、阵型数据及阵型码簿的相关联论述所描述。如本文中所描述的方法M100的实施方案可用于提供元数据(例如,在捕获视频流期间)以准许将帧编索引为描绘与选定对象中的特定几何布置的匹配以容易由通过搜索元数据的检索识别,因此避免检视整个视频流的需要。

[0203] 参考图20,展示描绘使用对象跟踪的视频处理的特定实施例。相对于图20中描绘的实施例而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。

[0204] 图20展示描绘在第一时间实例处捕获的场景的数字图像1700。例如,图像1700可对应于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,使用移动装置的相机)或从另一装置接收所述流。帧1700中描绘的场景包含第一对象1702、第二对象1704、第三对象1706及星星。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及第三对象1706中的至少一者可对应于经由移动装置控制的机器人。

[0205] 图20还展示移动装置的屏幕上的图像1700的显示1710。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T200)的用户选择机构1711。例如,用户可通过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以选择及启用对第一对象1702的跟踪而执行任务T100。响应于第一聚焦环1712的此类放置,移动装置可选择帧1700中的第一对象1702以进行跟踪。以类似方式,用户还可触摸屏幕上的用户选择机构1711且分别围绕第二对象1704及第三对象1706拖动第二聚焦环1714及第三聚焦环1716以选择并启用对第二对象1704及第三对象1706的跟踪。响应于第二聚焦环1714及第三聚焦环1716的此放置,移动装置可分别选择第二对象1704及第三对象1706以进行跟踪。

[0206] 图20还展示描绘在第二时间实例(例如,在第一时间实例之后)捕获的场景的图像1720。在图像1720中,每一对象1702到1706的位置已相对于描绘在第一时间实例处的场景的图像1700中的对应位置而改变。例如,在图像1720中,第一对象1702及第三对象1706已移动。虽然第二对象1704(例如,树)是静止对象,但第二对象1704可显得已移动。例如,移动装置(或其它捕获装置)的位置可移动,且继而可产生第二对象1704已移动的错觉。

[0207] 移动装置可监视选定对象1702到1706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,通过任务T200)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。

[0208] 图20还展示图像1720在移动装置的屏幕上的显示1730(例如,通过显示任务T50)。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。可基于图像1720中的第一对象1702的位置及图像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})。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。可基于图像1720中的第二对象1704的位置及图像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})。

[0209] 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每一指示(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图像(例如,一或多个视频流的帧)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)而将特定图像群集。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})而特定图像群集。例如,移动装置可确定由每一图像的每一指示(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。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。

[0210] 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(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。

[0211] 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。

[0212] 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多个者中的每一者的位置随时间的改变程度而改变所述频率。

[0213] 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(C_{01}),第二对象1704可在第二对象1704的位置处具有质量中心(C_{02})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。

[0214] 图20还展示描绘第三时间实例(例如,在第二时间实例之后)处捕获的场景的图像1740。在图像1740中,每一对象1702到1706的位置已相对于图像1720中在第二时间实例处描绘的对应位置而改变。

[0215] 图20还展示移动装置的屏幕上的图像1740的显示1750(例如,通过显示任务T50)。移动装置可在显示1750中显示几何形状的指示。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(L_{23})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(L_{13})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每一指示(L_{12} 、 L_{23} 、 L_{13})形成的三角形。

[0216] 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。

[0217] 还可能配置装置(例如,实施任务T200)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篮筐的二十英尺内及/或朝向目标/篮筐移动时进行指示,即使目标/篮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。

[0218] 图20中描绘的实施例可使得移动装置能够跟踪图像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21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20描述的技术。

[0219] 如本文中所描述的用于产生元数据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)。

[0220] 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21,展示用于基于集群而检索帧的移动装置的显示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。

[0221] 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21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。

[0222] 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)。第一群集可对应于屏幕1800上在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20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})可具有屏幕1730上的约150处的x坐标及屏幕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21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20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)。

[0223] 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)。第二群集可对应于屏幕1800上在选定对象之间的质量中心借以具有约100处的x坐标及约250处的y坐标的帧。参考图20中的第三帧,选定对象1702到1706的质量中心(C_{M3})可具有屏幕1750上的约100处的x坐标及屏幕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21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21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)。

[0224] 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产生的元数据之间,如上文所描述。

[0225] 图21还包含与视频流的帧的群集相关联的区的显示1802。例如,群集的显示1802可对应于图20中的场景的视频流。显示1802包含十个区(C1-C10)。在所说明的实施例中,区(C1-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-C10)可基于其它几何参数(例如,每一对象1702到1702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-C10)可具有相等大小(例如,每一区(C1-C10)可为“沃罗诺伊单元”)。

[0226] 每一区(C1-C10)可包含对应于与特定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20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-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。此集群信息可由任务T300产生,且作为元数据(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。

[0227] 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_{M1023-1026})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。

[0228] 参考图22,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。

[0229] 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20,用户可触摸屏幕1710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。

[0230] 在1904处,可跟踪每一对象之间的几何形状。例如,参考图20,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。

[0231] 在1906处,可在屏幕上显示几何形状的指示。例如,参考图20,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})。可基于图像1720中的第一对象1702的位置及图像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})。可基于图像1720中的第二对象1704的位置及图像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})。可基于图像1720中的第一对象1702的位置及图像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})。

[0232] 图22的方法1900可使得移动装置能够跟踪图像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(例如,通过任务T200)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。

[0233] 参考图23,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。

[0234] 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20,用户可触摸屏幕1710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。

[0235] 在2004处,可跟踪每一对象之间的几何形状。例如,参考图20,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可形成三个指示的三角形(L_{12} 、 L_{23} 、 L_{13})。

[0236] 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21,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21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)。

[0237] 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21,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21的屏幕1800上的第一群集(例如,群集1)而检索其中对象1702到1706与图20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。

[0238] 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21,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

[0239] 在2014年处,可基于特定几何形状而检索视频流的帧。例如,参考图21,移动装置可检索图20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。

[0240] 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21,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。

[0241] 图23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(例如,通过如本文中所描述的任务T300产生元数据),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。

[0242] 还可实施任务T300以基于所确定的参考位置而产生元数据,所述元数据识别划分位置空间的多个单独区中的一者,其中所述多个单独区包含不等大小的区。如上文所描述,位置空间可为二维或三维中的至少一个数字图像的像素坐标空间(例如,视频流的至少一个帧的像素坐标空间)或物理空间(例如,场景空间的地平面),且可延伸超过个别帧中表示的空间。此类元数据可被实施(例如)为一或多个索引(例如,进入一或多个量化码簿),所述一或多个索引指示包含参考位置的位置空间的区。

[0243] 划分方案可经配置以不均匀地划分位置空间(例如,以包含不等大小和/或形状的区域)。例如,一组观测到的数据点(例如,对象位置)可用于产生划分所述位置空间以使得每一单独区含有m个观测到的数据点的划分方案(例如,根据出现频率的划分)。或者或另外,划分方案可经配置以划分所述位置空间以使得高度关注的空间的区域被划分成比低度关注的区域更多的单独区。

[0244] 可根据与位置空间的相应的区域相关联的不同程度的关注而集中通过划分方案指示的各种区。图19B到D展示应用于篮球场的位置空间的非均匀划分的若干实例(例如,如图19A中所描绘)。此划分方案可将球场空间的划分编码为接近特定篮筐的较小区(例如,如图19B及19D中)或编码为接近任一篮筐的较小区(例如,如图19C中)。有可能凭经验指示位置空间的基于兴趣的非均匀划分(例如,在捕获流之前),使得可在数字图像或视频流的捕获期间执行方法M100。

[0245] 以类似方式,可根据位置空间的相应的区域内所映射的数据点(例如,对象的位置或多个对象的质量中心)的出现频率中的差异而集中通过划分方案指示的各种区。在此情

况下,可后验获得位置空间的非均匀划分(例如,作为对捕获的流训练的结果)。可实施所述划分以使得在每一单独区中找到相等数目的观测。

[0246] 在一个实例中,通过执行位置空间中的对象位置(例如,观测到的对象位置)的k均值集群为k个群集而获得划分方案。例如,可通过执行视频流的多个帧中观测到的对象位置(例如,选定对象中的一或多者的位置)的k均值集群为k个群集而获得划分方案。此集群将位置空间分割为沃罗诺伊单元,其中所得的k个均值是单元的产生器(例如,质心),且以此方式(例如,使用集群)获得的划分方案还可被称为量化码簿。群集可具有与另一群集不同数目的观测到的对象位置。为了获得初始条件(即,初始k个均值),可实施此划分方案产生任务以对观测到的位置执行分层集群方法(例如,分割或聚集集群)。可用于获得非均匀划分的类似性的量度的一个实例是位置空间中的观测到的对象位置之间的欧几里得距离。

[0247] 可由划分方案产生任务(例如,在存在离群值的情况下)使用的k均值集群的一个替代方案是k中心集群。另一替代方案是均值移位集群,其可用于将分布的模式分离到不同区中。任务T300可包含从多个划分方案(例如,多个码簿)当中进行选择,所述多个划分方案可具有不同长度及/或可对应于不同大小的区。例如,任务T300可使用一层级的划分方案,使得第一水平的划分方案指示多个单独区中的一者,并且还在多个此类第二水平方案当中指示第二水平的划分方案,其对应于所指示的区。此层级可用于将多个单独区中的一者与其它者不同地细分。在一个此类实例中,第一码簿将对应于观测到的位置的分布中的不同模式的不同量化区编索引,且第二码簿用于将这些模态区中的一或多者中的每一者划分为子区,使得任务T300可产生包含用于确定的位置的一个以上码字的元数据。

[0248] 可实施任务T300以通过例如选择其质心最接近所确定的参考位置的区而将所确定的参考位置映射到位置空间的对应区。用于索引选择的类似性的量度(例如,位置空间中的欧几里得距离)可与用于分割的类似性的量度相同或不同。对于其中通过量化码簿描述划分方案的情况,此类映射可包含选择到码簿(例如,指示选定区域的码字)中的索引。

[0249] 参考图24,展示用于使用直角坐标产生对象位置的元数据的方法M100的特定应用。图24中的实施例说明可用于逐帧地编码多个对象的确定的位置的码簿2102。所确定的位置可指示位置空间2106中的每一对象的位置。码簿2102在将位置空间划分2108为单独区(包含不等大小的区)的过程中将位置空间2106(例如,2132、2134、2136)中的位置映射到对应区(例如,2142、2144、2146)。

[0250] 码簿2102可用于产生多个帧或其它图像2104的元数据。此元数据可并入到视频流中及/或可单独地串流传输及/或存储在单独文件中。所确定的位置可包含帧中的对象的直角坐标数据(例如,x坐标及y坐标)。在特定实施例中,所确定的位置还可包含帧中的对象的三维坐标数据(例如,三维笛卡尔坐标系统的z坐标)。在所说明的实施例中,可通过将10,000个帧(或其它图像)上的三个对象的确定的位置集群而产生码簿2102。这些确定的位置还可能已经与对应的帧相关联以作为元数据。在特定实施例中,第一帧可对应于图20的第一图像(例如,显示1710),第二帧可对应于图20的第二图像(例如,显示1730),且第三帧可对应于图20的第三图像(例如,显示1750)。虽然在此实例中,使用10,000个帧的确定的位置以获得码簿2102,但在其它实施例中,码簿2102可基于更少(或额外)帧的确定的位置。或者或另外,码簿可基于如上文所描述的基于兴趣的划分方案(例如,如图19B到D中所展示),或不是通过将一组观测到的确定的位置集群而获得的另一划分方案。另外,所确定的位置可

包含更少(或额外)对象的位置信息。图24的以下描述是基于对应于第三图像的确定的位置。然而,类似技术可适用于其它确定的位置。

[0251] 第三图像中的所确定的位置可包含第一直角坐标数据2132、第二直角坐标数据2134及第三直角坐标数据2136。第一直角坐标数据2132可包含图20的第三图像中的第一对象1702的x坐标位置及第三图像中的第一对象1702的y坐标位置。第二直角坐标数据2134可包含第三图像中的第二对象1704的x坐标位置及第三图像中的第二对象1704的y坐标位置。第三直角坐标数据2136可包含第三图像中的第三对象1706的x坐标位置及第三图像中的第三对象1706的y坐标位置。

[0252] 在另一特定实施例中,直角坐标数据2132到2136中的每一对象1702到1706的x坐标及y坐标可相对于第三图像中的对象1702到1706的质量中心(C_{M3})。例如,可将质量中心(C_{M3})指定为原点,且每一对象1702到1706的位置(例如,x坐标及y坐标)可相对于所述原点。在此情况下,例如在图19A中所展示的篮球场或其它体育场的实例中,可能需要根据非均匀划分方案(例如图19B到D中展示的非均匀划分方案)量化选定对象(例如,运动选手)的此质量中心的位置。

[0253] 在另一特定实施例中,可将特定对象指定为原点,且其它对象的位置(例如,x坐标及y坐标)可相对于所述原点。所述特定对象(也被称为“主要关注的对象”)可包含以下实例中的任一者:球(例如,体育事件的视频中)、选定对手(例如,体育事件中)、犯罪受害者(例如,监控视频中)、新娘(例如,婚礼视频中)。主要关注的对象可随时间移动(例如,人),或可为在空间中具有固定位置的对象(例如,网、篮筐或体育视频中的其它目标)。作为非限制性实例,第一对象1702可被指定为原点,且第二对象1704及第三对象1706的位置可相对于所述原点。

[0254] 处理器可使用码簿2102根据将位置空间划分2108为单独区(包含不等大小的区)而编码直角坐标数据2132到2136以产生元数据。为了说明,处理器可将第一直角坐标数据2132编码(例如,量化)为第一经编码位置数据2142。例如,处理器可将第一直角坐标数据2132映射到对应值(例如,码簿2102的码字)且将第一直角坐标数据2132编码为值(例如,第一经编码位置数据2142)。另外,处理器可将第二直角坐标数据2134编码为第二经编码位置数据2144。例如,处理器可使用码簿2102将第二直角坐标数据2134映射到对应值且将第二直角坐标数据2134编码为值(例如,第二经编码位置数据2144)。以类似方式,处理器可将第三直角坐标数据2136编码为第三经编码位置数据2146。例如,处理器可使用码簿2102将第三直角坐标数据2136映射到对应值且将第三直角坐标数据2136编码为值(例如,第三经编码位置数据2146)。

[0255] 基于第一经编码位置数据2142,图20的第三图像中的第一对象1702(例如,对象1)的位置可在所产生的元数据中表示为网格2108中的对应位置(例如,指示的区的质心)。基于第二经编码位置数据2144,第三图像中的第二对象1704(例如,对象2)的位置可在产生的元数据中表示为网格2108中的对应位置。基于第三经编码位置数据2146,第三图像中的第三对象1706(例如,对象3)的位置可在产生的元数据中表示为网格2108中的对应位置。

[0256] 如上所述,所确定的位置可为多个选定对象的几何布置或“阵型”的参考位置。另外或在使用一或多个位置(或“参考”)码簿的替代方案中,可实施任务T300以使用一或多个阵型(或“相对”)码簿,其相对于参考位置(例如,对象的质量中心的位置)编码空间中的选

定对象的阵型(例如,它们的几何布置或描述此布置的形状)。在三个选定对象的非限制性实例中,可能需要将2-D空间中的选定对象的质量中心的位置存储为参考位置且产生元数据,所述元数据相对于参考位置而针对总共八个自由度指示2-D空间中的每一选定对象的位置。这些值可经量化为单个码字(例如,8-D空间中)。或者,可使用如上文所描述的一或多个位置码簿及对象阵型(例如,在6-D空间中)、使用一或多个阵型码簿(即,候选几何布置的编索引集合)将这些值单独地量化为参考位置(例如,在2-D空间中)。

[0257] 或者,到阵型或相对码簿中的索引可识别一组模板(即,候选几何布置)中的一者。在此情况下,可单独地量化对指示的模板的修改(例如,缩放、旋转或其它定向、纵横比等)。图33A展示三个模板的实例,且图33B展示对图33A中的最左边模板的修改的三个实例。

[0258] 阵型码簿可经配置以(例如)根据以下各者而不均匀地划分阵型空间(即,所有可能的阵型的空间的某一部分):可能的阵型之间的出现频率中的差异;可能的阵型之间的所关注水平的差异(例如,与预定阵型的类似性);可能的阵型与特定预定事件(例如,得分事件)或模式(例如,足球队的特定制阵型)的不同程度的关联;距空间参考最远的选定对象的距离;选定对象的选定子集彼此的距离;和/或选定对象中的任两者之间的最大距离。

[0259] 码字(例如,经编码位置数据2142到2146)可为固定宽度码字或可变宽度码字。在特定实施例中,处理器可基于对象的位置(例如,基于对象是在“高密度区域”中还是“低密度区域”中)而使用可变宽度码字。高密度区域可对应于图20中的图像的常常由对象1702到1706中的至少一者占据的区域。为减少元数据中的位的数目,对应于位置空间2106中的高密度区域的码字可比对应于低密度区域的码字具有更少的位(即,较短码字)(即,较不频繁被占据的网格位置具有较长码字)。或者或另外,非均匀划分方案2108可为在高密度区中具有更多子区(即,更多相异码字以提供较高分辨率)且在低密度区中具有更少子区(即,更少相异码字,从而提供较低分辨率)的分层方案。

[0260] 如果对象在高密度区域中,那么对象的码字可表示相对小的区以产生对象的更“精确”位置。例如,增加数目的码字可用于跟踪高密度区域中的对象1702到1706的位置以逐帧地获得更准确的跟踪。如果对象在低密度区域中,那么对象的码字可表示相对大的区。在特定实施例中,可基于历史数据的分析而确定高密度区域及低密度区域(例如,分析屏幕的哪些区域具有频繁被对象占据的历史)。在另一特定实施例中,高密度区域及低密度区域可为预先确定的。

[0261] 到阵型码簿中的码字可在n个选定对象到n个位置的唯一映射之间进行区分(例如,可包含用于每一此类映射的唯一码字)。例如,图34A展示各自可经编码为唯一码字的三个选定对象的四个不同阵型的实例(即,可由不同的相应的候选几何布置表示)。在这些实例中,选定对象的每一位置由唯一地识别所述对象的图标(例如,圆形、正方形或三角形)指示。或者,到阵型码簿中的码字可编码若干不同唯一映射。例如,图34B展示对应于图34A的阵型的非唯一映射的实例,其中任何选定对象的位置由x指示。在此情况下,图34C展示可由在图34B中最左边展示的单一个非唯一映射(即,候选几何布置)指示的选定对象的六个唯一阵型。

[0262] 码簿2102可为固定码簿或自适应码簿。自适应码簿可基于表示对象位置的历史数据而调整码字的长度(例如,可产生可变宽度码字)。例如,自适应码簿可基于历史数据而产生屏幕的频繁占据区域中的对象的较短码字且可产生屏幕的较不频繁占据区域中的对象

的较长码字。相比而言,固定码簿可不基于历史数据而改变,且可使用固定宽度码字或可变宽度码字。

[0263] 可能需要使用多个阵型码簿编码阵型数据,所述多个阵型码簿利用可能的阵型之间相对于相应的参考位置的出现频率或关注水平中的差异。例如,可依据参考位置的码字而选择不同大小的一组阵型码簿中的一者(例如,在参考位置处于球场中心时的较小码簿对在参考位置在目标附近时的较大码簿)。

[0264] 图24的实施例可使得用户能够基于通过任务T300使用码簿2102而产生的元数据来检索(例如,定位)特定图像,例如视频的特定部分。例如,位置空间2106可显示在屏幕上以作为用户交互网格,其使得用户能够选择不同对象的不同位置(例如,参数)。基于所述选择,移动装置的用户可容易定位视频的具有类似对象位置配置的所要部分(例如,帧),而不必导航(例如,播放、快进、快退等)穿过所述视频。例如,可实施方法M100以包含基于将产生的元数据与指定划分方案中的区中的一者的搜索请求进行比较的结果而检索一或多个图像(例如,从存储装置)的任务。在特定实施例中,位置空间2106的此显示可以与图21中的群集的屏幕1802大体上类似的方式使用。另外,产生的元数据2142到2146可被压缩并且从编码器传输到解码器。压缩元数据2142到2146且将其传输到解码器可使得解码器能够使用相对低量的数据(例如,经量化值)确定被跟踪对象的位置。例如,解码器可执行再现操作以基于码簿2102中的经量化值而解码与对象1702到1706相关联的视频。

[0265] 在一个实例中,所检索的帧包含选定运动选手之间的仅在特定播放或一组播放期间发生的选定阵型,或足球比赛或其它体育事件中的特定首发阵型。在另一实例中,所检索的帧包含在跑道上接近的特定赛车之间的选定阵型。通过此类选择性检索驱动的视频分析还可用于检测事件与特定阵型的出现的相关度(其先前可能未识别)。例如,潜在应用包含跟踪集会(例如,立法会议、街头示威、社交聚会)里的人且使选定人的接近与指定事件相关(例如,以识别决策者、滋事者、促进因素)。还可应用如本文中所描述的方法M100的实施方案以分析圈养(例如,实验室或动物园中)或野生的动物之间的社会互动。

[0266] 参考图25,展示用于使用极坐标产生对象位置的元数据的方法M100的另一特定应用。图25中的实施例说明可用于逐帧地编码多个对象的确定的位置的码簿2202。所确定的位置可指示极位置空间2206中的每一对象的位置。码簿2202在将位置空间2206划分2208为单独区(包含不等大小的区)的过程中将位置空间2206(例如,2232、2234、2236)中的位置映射到对应区(例如,2242、2244、2246)。

[0267] 码簿2202可用于产生多个帧或其它图像2204的元数据。此元数据可并入到视频流中及/或可单独地串流传输及/或存储在单独文件中。所确定的位置可包含帧中的对象的极坐标数据。在特定实施例中,所确定的位置还可包含帧中的对象的球或圆柱坐标数据(例如,3D坐标数据)。在所说明的实施例中,可通过将10,000个帧(或其它图像)上的三个对象的确定的位置集群而产生码簿2102。这些确定的位置还可能已经与对应的帧相关联以作为元数据。在特定实施例中,第一帧可对应于图20的第一图像(例如,显示1710),第二帧可对应于图20的第二图像(例如,显示1730),且第三帧可对应于图20的第三图像(例如,显示1750)。虽然在此实例中,使用10,000个帧的确定的位置以获得码簿2202,但在其它实施例中,码簿2202可基于更少(或额外)帧的确定的位置。或者或另外,码簿可基于如上文所描述的基于兴趣的划分方案(例如,如图19B到D中所展示),或不是通过将一组观测到的确定的

位置集群而获得的另一划分方案。另外,所确定的位置可包含更少(或额外)对象的位置信息。图25的以下描述是基于对应于第三图像确定的位置。然而,类似技术可适用于其它确定的位置。

[0268] 第三图像中的所确定的位置可包含极定向数据2230、第一极坐标数据2232、第二极坐标数据2234及第三极坐标数据2236。极定向数据2230可指示质量中心(C_{M3})的位置(例如,图20的第三图像中的质量中心(C_{M3})的x坐标(C_x)及第三图像中的质量中心(C_{M3})的y坐标(C_y))。极定向数据2230还可指示具有质量中心(C_{M3})的圆的半径(R)作为中心。可实施为比例因子的半径(R)可大到足以包含每一被跟踪对象。相对于极位置空间2206描绘说明性实例。在一个实例中,实施任务T300以将定向位置(例如,质量中心的位置)编码为指示包含不等大小的区的多个区中的一者的元数据,其划分在其内指示定向位置的对应位置空间,所述位置空间可不同于位置空间2206(例如,如图24中所展示的直角坐标空间)。

[0269] 第一极坐标数据2232可指示第一对象1702(对象1)的第一极坐标。第一极坐标可包含第一对象1702距质量中心(C_{M3})的第一距离(a)及从零度参考测得的第一对象1702的第一角(θ_a)(例如,与水平的角度移位)。第二极坐标数据2234可指示第二对象1704(对象2)的第二极坐标。所述第二极坐标可包含第二对象1704距质量中心(C_{M3})的第二距离(b)及从水平测得的第二对象1704的第二角度(θ_b)。第三极坐标数据2236可指示第三对象1706(对象3)的第三极坐标。所述第三极坐标可包含第三对象1706距质量中心(C_{M3})的第三距离(c)及从水平测得的第一对象1702的角度(θ_c)。

[0270] 在特定实施例中,可基于与质量中心(C_{M3})最远的对象的距离而将极定向数据2230的半径(R)归一化。例如,如果第一距离(a)大于第二距离(b)及第三距离(c),那么半径(R)可等于第一距离(a)。

[0271] 处理器可使用码簿2202根据将位置空间划分2108为多个区(包含不等大小的区)而编码极坐标数据2232到2236以产生元数据。为了说明,处理器可将第一极坐标数据2232编码(例如,量化)为第一经编码位置数据2242。例如,处理器可将第一极坐标数据2232(其是基于极定向数据2230)映射到对应值(例如,码簿2202的码字)且将第一极坐标数据2232编码为值(例如,第一经编码位置数据2242)。另外,处理器可将第二极坐标数据2234编码为第二经编码位置数据2244。例如,处理器可使用码簿2202将第二极坐标数据2234(其是基于极定向数据2230)映射到对应值且将第二极坐标数据2234编码为值(例如,第二经编码位置数据2244)。以类似方式,处理器可将第三极坐标数据2236编码为第三经编码位置数据2246。例如,处理器可使用码簿2202将第三极坐标数据2236映射到对应值且将第三极坐标数据2236编码为值(例如,第三经编码位置数据2246)。

[0272] 基于第一经编码位置数据2242,图20的第三图像中的第一对象1702(例如,对象1)的位置可在所产生的元数据中表示为极网格2208中的对应位置(例如,指示的区的质心)。基于第二经编码位置数据2244,第三图像中的第二对象1704(例如,对象2)的位置可在产生的元数据中表示为极网格2208中的对应位置。基于第三经编码位置数据2246,第三图像中的第三对象1706(例如,对象3)的位置可在产生的元数据中表示为极网格2208中的对应位置。例如,极网格2208可基于距网格原点的径向距离的范围(例如,基于距质量中心(C_{M3})的径向距离的范围)及与零度参考的角度移位的范围而被划分成包含不等大小的区的若干区。

[0273] 图25的实施例可使得用户能够基于通过任务T300使用码簿2202而产生的元数据来检索(例如,定位)特定图像(例如,视频的帧)。例如,极位置空间2206可显示在屏幕上以作为用户交互网格,其使得用户能够选择不同对象的不同位置(例如,参数)。基于所述选择,移动装置的用户可容易定位视频的具有类似对象位置配置的所要部分(例如,帧),而不必导航(例如,播放、快进、快退等)穿过所述视频。例如,可实施方法M100以包含基于将产生的元数据与指定划分方案中的区中的一者的搜索请求进行比较的结果而检索一或多个图像(例如,从存储装置)的任务。在特定实施例中,极位置空间2206的此显示可以与图21中的群集的屏幕1802大体上类似的方式使用。另外,产生的元数据2242到2246可被压缩并且从编码器传输到解码器。压缩元数据2242到2246且将其传输到解码器可使得解码器能够使用相对低量的数据(例如,经量化值)确定被跟踪对象的位置。例如,解码器可执行再现操作以基于码簿2202中的经量化值而解码与对象1702到1706相关联的视频。

[0274] 码字(例如,经编码位置数据2242到2246)可为固定宽度码字或可变宽度码字。在特定实施例中,处理器可基于对象的位置(例如,基于对象是在“高密度区域”中还是“低密度区域”中)而使用可变宽度码字。高密度区域可对应于图20中的图像的常常由对象1702到1706中的至少一者占据的区域。为减少元数据中的位的数目,对应于极位置空间2206中的高密度区域的码字可对应于低密度区域的码字具有更少的位(即,较短码字)(即,较不频繁被占据的网格位置具有较长码字)。或者或另外,非均匀划分方案2208可为在高密度区中具有更多子区(即,更多相异码字以提供较高分辨率)且在低密度区中具有更少子区(即,更少相异码字,从而提供较低分辨率)的分层方案。

[0275] 如果对象在高密度区域中,那么对象的码字可表示相对小的区以产生对象的更“精确”位置。例如,增加数目的码字可用于跟踪高密度区域中的对象1702到1706的位置以逐帧地获得更准确的跟踪。如果对象在低密度区域中,那么对象的码字可表示相对大的区。在特定实施例中,可基于历史数据的分析而确定高密度区域及低密度区域(例如,分析屏幕的哪些区域具有频繁被对象占据的历史)。在另一特定实施例中,高密度区域及低密度区域可为预先确定的。

[0276] 码簿2202可为固定码簿或自适应码簿。自适应码簿可基于表示对象位置的历史数据而调整码字的长度(例如,可产生可变宽度码字)。例如,自适应码簿可基于历史数据而产生屏幕的频繁占据区域中的对象的较短码字且可产生屏幕的较不频繁占据区域中的对象的较长码字。相比而言,固定码簿可不基于历史数据而改变,且可使用固定宽度码字或可变宽度码字。

[0277] 参考图26,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,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2300。

[0278] 方法2300可包含在2302处在移动装置处跟踪特定帧中的第一对象的第一位置及所述特定帧中的第二对象的第二位置。例如,参考图20,可在第三帧中跟踪第一对象1702的所述第一位置且可在所述第三帧中跟踪第二对象1704的所述第二位置。另外,可在第三帧中跟踪第三对象1706的第三位置。

[0279] 在2304处,可产生与所述第一位置相关联的第一坐标数据及与所述第二位置相关

联的第二坐标数据。例如,参考图24,处理器可产生与第三帧中的第一对象1702的位置相关联的第一直角坐标数据2132及与第三帧中的第二对象1704的位置相关联的第二直角坐标数据2134。另外,处理器可产生与第三帧中的第三对象1706的位置相关联的第三直角坐标数据2136。

[0280] 作为另一实例,参考图25,处理器可产生与第三帧中的第一对象1702的位置相关联的第一极坐标数据2232及与第三帧中的第二对象1704的位置相关联的第二极坐标数据2234。另外,处理器可产生与第三帧中的第三对象1706的位置相关联的第三极坐标数据2236。

[0281] 在2306处,可使用码簿编码第一坐标数据及第二坐标数据以产生元数据以作为第一经编码位置数据及第二经编码位置数据。例如,参考图24,处理器可使用码簿2102编码第一直角坐标数据2132以产生第一经编码位置数据2142(例如,码簿2102的码字),且处理器可使用码簿2102编码第二直角坐标数据2134以产生第二经编码位置数据2144。另外,处理器可使用码簿2102编码第三直角坐标数据2136以产生第三经编码位置数据2146。

[0282] 作为另一实例,参考图25,处理器可使用码簿2202编码第一极坐标数据2232以产生元数据以作为第一经编码位置数据2242(例如,码簿2202的码字),且处理器可使用码簿2202编码第二极坐标数据2234以产生元数据以作为第二经编码位置数据2244。另外,处理器可使用码簿2202编码第三极坐标数据2236以产生元数据以作为第三经编码位置数据2246。

[0283] 在2308处,可存储元数据(例如,第一经编码位置数据及第二经编码位置数据)以实现基于输入参数对特定帧的检索。例如,参考图24,处理器可将第一经编码位置数据2142及第二经编码位置数据2144存储在存储器中。移动装置的用户可输入参数(例如,第一对象1702及第二对象1704的所要的位置)。基于所述输入参数,移动装置可响应于对应于第一位置及第二位置的输入参数而检索第三帧以用于回放。

[0284] 作为另一实例,参考图25,处理器可将第一经编码位置数据2242及第二经编码位置数据2244存储在存储器中。移动装置的用户可输入参数(例如,第一对象1702及第二对象1704的所要的位置)。基于所述输入参数,移动装置可响应于对应于第一位置及第二位置的输入参数而检索第三帧以用于回放。

[0285] 在特定实施例中,方法2300还可包含跟踪多个帧上的场景的第一区域的第一密度。第一密度可基于第一对象1702或第二对象1704在第一区域中的次数。方法2300还可包含跟踪多个帧上的场景的第二区域的第二密度。所述第二密度可基于第一对象1702或第二对象1704在第二区域中的次数。方法2300可进一步包含将第一密度与第二密度进行比较且在第一密度大于第二密度的情况下将码簿2102中的第一量的元数据分配到第一区域,且将码簿2102中的第二量的元数据分配到第二区域。所述第一量的元数据可大于所述第二量的元数据。例如,码簿2102中的较大量码字可对应于第一区域(例如,较高分辨率区域)且较少量码字可对应于第二区域(例如,较低分辨率区域)以减少较不密集区域的码字的数目。

[0286] 图26的方法2300可使得用户能够检索(例如,定位)使用码簿2102编码的视频的特定部分(帧)。例如,网格2106可为显示在屏幕上的使得用户能够选择不同对象的不同位置(例如,参数)的用户交互式网格。基于所述选择,移动装置的用户可容易定位视频的具有类似对象位置配置的所要部分,而不必导航(例如,播放、快进、快退等)穿过所述视频。

[0287] 参考图27,展示说明使用对象跟踪的视频处理方法2600的特定实施例的流程图。在说明性实施例中,可使用机器人执行方法2600。

[0288] 方法2600可包含在2602处在机器人处跟踪多个对象之间的几何形状。例如,用户可将待跟踪的特定对象传送到机器人。在特定实施例中,所述机器人可包含显示接口且用户可经由使用所述显示接口的交互而与所述机器人通信。作为另一实例,机器人可基于环境因素而独立地确定将跟踪什么对象。例如,机器人可包含检测运动的传感器。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。

[0289] 在2604处,可基于几何形状执行处理功能。例如,机器人可基于几何形状而确定对象之间的关系。作为另一实例,机器人可基于对象的几何形状而确定环境(例如,场景)。

[0290] 参考图28,展示描绘使用对象跟踪的视频处理的特定实施例。相对于图28中描绘的实施例而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。

[0291] 图28展示描绘在第一时间实例处捕获的场景的数字图像2700。例如,图像2700可对应于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,使用移动装置的相机)或从另一装置接收所述流。帧2400中描绘的场景包含第一对象2702、第二对象2704、第三对象2706、星星及树。在所说明的实施例中,第一对象2702可对应于第一人、第二对象2704可对应于第二人,且第三对象2706可对应于第三人。在另一实施例中,对象2702到2706中的至少一者可对应于经由移动装置控制的机器人。

[0292] 图28还展示移动装置的屏幕上的帧2700的显示2710。移动装置的屏幕(例如,触摸屏)可对应于取景器。显示2710可包含使用户能够选择用于跟踪的对象(例如,通过任务T200)的用户选择机构2711。例如,用户可触摸屏幕2710上的用户选择机构2711且围绕第一对象2702拖拽第一聚焦环2712以实现第一对象2702的跟踪。响应于第一聚焦环2712的此放置,移动装置可选择用于跟踪的第一对象2712。以类似方式,用户还可触摸屏幕上的用户选择机构2711且分别围绕第二对象2704及第三对象2706拖拽第二聚焦环2714及第三聚焦环2716以实现第二对象2704及第三对象2706的跟踪。响应于第二聚焦环2714及第三聚焦环2716的此放置,移动装置可分别选择用于跟踪的第二对象2704及第三对象2706。

[0293] 图28还展示描绘在第二时间实例处捕获的场景的图像(例如,帧)2720。在图像2720中,每一对象2702到2706的位置已相对于在描绘在第一时间实例处的场景的图像2700中描绘的对应位置而改变。例如,第一对象2702已移动成更靠近星星,第二对象2704已移动成更靠近树,且第三对象2706已移动成更靠近图像的底部。

[0294] 图24还展示移动装置的屏幕上的图像2720的显示2730。可将显示2730分割(例如,通过显示任务T50)为第一窗口2732、第二窗口2734、第三窗口2736及第四窗口2738。响应于使用用户选择2711选择三个对象2702到2706,例如,移动装置可将屏幕2730分割为三个对应窗口2732到2736及第四窗口2738。

[0295] 第一窗口2732可显示包含第一对象2702的图像2720的第一部分。可基于图像2720中的第一对象2702的位置而确定图像2720的第一部分。例如,移动装置可跟踪第一对象2702的位置且可在第一窗口2732中显示聚焦在第一对象2702的位置上的视频流。第二窗口2734可显示包含第二对象2704的图像2720的第二部分。可基于图像2720中的第二对象2704

的位置而确定图像2720的第二部分。例如,移动装置可跟踪第二对象2704的位置且可在第二窗口2734中显示聚焦在第二对象2704的位置上的视频流。第三窗口2736可显示包含第三对象2706的图像2720的第三部分。可基于图像2720中的第三对象2706的位置而确定图像2720的第三部分。例如,移动装置可跟踪第三对象2704的位置且可在第三窗口2736中显示聚焦在第三对象2706的位置上的视频流。

[0296] 第四窗口2738可显示每一对象2702到2706之间的几何形状的指示(例如,质量中心(C_{M2}))。例如,第四窗口2738可包含网格,所述网格包含对应于图像2720中的区域的多个区段(表示为灰线)。例如,网格的左下部分可对应于图像2720的左下部分。以与相对于图21及22所描述的方式类似的方式,移动装置可确定图像2720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})。

[0297] 图28还展示描绘在第三时间实例(例如,在第二时间实例之后)处捕获的场景的图像2740。在图像2740中,每一对象2702到2706的位置已相对于描绘第二时间实例处的场景的图像2720中的对应位置而改变。例如,第一对象2702已移动成更靠近树,第二对象2704已移动成更靠近场景2740的底部,且第三对象2706已移动成更靠近星星。

[0298] 图28还使用由移动装置产生的窗口2732到2736在移动装置的屏幕上展示图像2740的显示2750。例如,第一窗口2732显示包含第一对象2702的图像2740的第一部分,第二窗口2734显示包含第二对象2704的图像2740的第二部分,且第三窗口2736显示包含第三对象2706的图像2740的第三部分。

[0299] 第四窗口2738可更新指示(例如,质量中心(C_{M2}))以反映第三时间实例处的图像2740。移动装置可逐帧地或以非连续时间间隔跟踪几何形状,且在第四窗口2738中显示经更新的几何形状的指示。例如,在第三帧处,第四窗口2738可显示第三时间实例处的图像2740中的对象2702到2706的质量中心(C_{M3})。第四窗口2738还可显示跟踪第二帧与第三帧之间的对象的质量中心(例如,通过任务T200指示)的路径(例如,虚线)。

[0300] 图28中描绘的实施例可使得移动装置能够在帧2700、2720、2740中跟踪对象2702到2706,且可在单独窗口2732到2736中显示对应于对象2702到2706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2738中查看对象2702到2706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28描述的技术。

[0301] 参考图29,展示说明使用对象跟踪的视频处理方法2800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2800。

[0302] 方法2800包含在2802处在移动装置处接收对场景中的第一对象、场景中的第二对象及场景中的第三对象的选择。例如,参考图28,用户可触摸屏幕2710上的用户选择机构2711且围绕第一对象2702拖动第一聚焦环2712,围绕第二对象2704拖动第二聚焦环2714,且围绕第三对象2706拖动第三聚焦环2716以分别启用对第一、第二和第三对象2702到2706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。

[0303] 在2804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28,可将显示2730分割成第一窗口2732、第二窗口2734、第三窗口2736及第四窗口。

[0304] 在2806处,可在第一窗口中显示包含第一对象的所显示场景的第一部分。例如,参

考图28,第一窗口2732可显示包含第一对象2702的图像2720的第一部分。可基于图像2720中的第一对象2702的位置而确定图像2720的第一部分。例如,移动装置可跟踪第一对象2702的位置,且可在第一窗口2732中显示聚焦在第一对象2702的位置上的视频流。

[0305] 在2808处,可在第二窗口中显示包含第二对象的所显示场景的第二部分。例如,参考图28,第二窗口2734可显示包含第二对象2704的图像2720的第二部分。可基于图像2720中的第二对象2704的位置而确定图像2720的第二部分。例如,移动装置可跟踪第二对象2704的位置,且可在第二窗口2734中显示聚焦在第二对象2704的位置上的视频流。

[0306] 在2810处,可在第三窗口中显示包含第三对象的所显示场景的第三部分。例如,参考图28,第三窗口2736可显示包含第三对象2706的图像2720的第三部分。可基于图像2720中的第三对象2706的位置而确定图像2720的第三部分。例如,移动装置可跟踪第三对象2706的位置,且可在第三窗口2736中显示聚焦在第三对象2706的位置上的视频流。

[0307] 在2812处,可在第四窗口中显示每一对象之间的几何形状的指示。例如,参考图28,第四窗口2738可显示每一对象2702到2706之间的质量中心(C_{M2})。第四窗口2738可更新质量中心(C_{M2})以反映第三时间实例处的场景。例如,在第三帧处,第四窗口2738可显示第三时间实例处的图像2740中的对象2702到2706的质量中心(C_{M3})。第四窗口2738还可显示跟踪第二帧与第三帧之间的对象的质量中心的路径(例如,虚线)。

[0308] 图29的方法2800可使得移动装置能够跟踪场景2700、2720、2740中的对象2702到2706且在单独窗口2732到2736中显示对应于对象2702到2706的个别化(例如,聚焦)视频流。方法2800还使得用户能够在第四窗口2738中查看对象2702到2706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方法2800。

[0309] 参考图42,展示描绘使用对象跟踪的视频处理的特定实施例。相对于图42中描绘的实施例而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。

[0310] 图42展示描绘第一时间实例处的场景的数字图像1100。例如,图像1100可对应于在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。图像1100中描绘的场景可包含第一关注区(ROI) 1102、第二ROI 1104及第三ROI 1106。每一ROI 1102到1106可对应于场景内的包含根据本文中描述的技术(例如,参考图1到10、对象跟踪与检测模块204及804及/或任务T200)而跟踪的一或多个对象的区。

[0311] 为了说明,第一ROI 1102可包含第一被跟踪对象群组(例如,第一被跟踪对象1110及第三被跟踪对象1114)。另外,第二ROI 1104可包含第二被跟踪对象群组(例如,第四被跟踪对象1120、第五被跟踪对象1122及第六被跟踪对象1124),且第三ROI 1106可包含第三被跟踪对象群组(例如,第七被跟踪对象1130、第八被跟踪对象1132,及第九被跟踪对象1134)。第二被跟踪对象1112可定位在ROI 1102到1104外部。

[0312] 可产生元数据(例如,通过任务T300)以识别场景1100中的每一被跟踪对象1110到1114、1120到1124、1130到1134的位置。在特定实施例中,所述元数据对应于直角坐标数据。例如,可产生第一元数据以识别第一被跟踪对象1110的位置(例如,x坐标位置及y坐标位置),可产生第二元数据以识别第二被跟踪对象1112的位置等。在其它实施例中,所述元数

据可对应于极坐标数据、球坐标数据、圆柱坐标数据,或三维笛卡尔坐标数据。

[0313] 每一ROI 1102到1106可由ROI 1102到1106中的被跟踪对象的质量中心界定。例如,第一ROI 1102可由第一被跟踪对象1110与第三被跟踪对象1114之间的第一质量中心(C_{M1})界定。第二ROI 1104可由第四被跟踪对象1120、第五被跟踪对象1122与第六被跟踪对象1124之间的第二质量中心(C_{M2})界定。第三ROI 1106可由第七被跟踪对象1130、第八被跟踪对象1132与第九被跟踪对象1134之间的第三质量中心(C_{M3})界定。

[0314] 如相对于图43更详细地描述,可使用码簿来编码场景1100中的每一被跟踪对象1110到1114、1120到1124、1130到1134的位置的元数据。定位在ROI 1102到1106中的一者内的被跟踪对象的元数据可以比定位在ROI 1102到1106外部的被跟踪对象(例如,第二被跟踪对象1112)的元数据更高的位速率编码。以较高的位速率编码定位在ROI 1102到1106中的一者内的被跟踪对象的元数据可使得能够以比其它被跟踪对象的位置更高水平的准确度编码ROI 1102到1106中的被跟踪对象的位置。

[0315] 参考图43,展示编码对象位置的元数据的特定实施例。图43中的实施例说明可用于逐帧地编码多个对象的位置的元数据的码簿1202。元数据可用于指示网格中的每一对象的位置。

[0316] 码簿1202可用于编码多个帧1204的元数据。所述元数据可包含帧中的对象的直角坐标数据(例如,x坐标及y坐标)、帧中的对象的极坐标、帧中的对象的球坐标、帧中的对象的圆柱坐标,或帧中的对象的三维笛卡尔坐标。在所说明的实施例中,码簿1202可编码10,000个帧上的三个对象的元数据。在特定实施例中,第三帧可对应于图43的第一帧(例如,场景1100)。虽然在码簿1202中描绘10,000个帧的元数据,但在其它实施例中,码簿1202可编码更少(或额外)帧的元数据。另外,元数据可包含更少(或额外)对象的位置信息。图43的以下描述是基于对应于第三帧的元数据。然而,类似技术可适用于其它元数据。

[0317] 第三帧中的元数据可包含第一坐标数据1232、第二坐标数据1234、第三坐标数据1236、第四坐标数据1237、...、及第九坐标数据1238。第一坐标数据1232可包含图42的图像1100中的第一被跟踪对象1110的坐标位置,第二坐标数据1234可包含图像1100中的第二被跟踪对象1112的坐标位置,第三坐标数据1236可包含图像1100中的第三被跟踪对象1114的坐标位置,第四坐标数据1237可包含图像1100中的第四被跟踪对象1120的坐标位置等。

[0318] 处理器可使用码簿1202编码坐标数据1232到1236。为了说明,处理器可使用第一位速率(例如,相对高的位速率)将第一坐标数据1232编码(例如,量化)为第一经编码位置数据1242。例如,处理器可将第一坐标数据1232映射到对应值(例如,码簿1202的码字)且可将第一坐标数据1232编码为所述值(例如,第一经编码位置数据1242)。处理器可使用第一位速率,因为第一被跟踪对象1110在ROI(例如,图42的第一ROI 1102)内部。

[0319] 另外,处理器可使用低于第一位速率的第二位速率将第二坐标数据1234编码为第二经编码位置数据1244。例如,处理器可使用码簿1202将第二坐标数据1234映射到对应值且可将第二坐标数据1234编码为所述值(例如,第二经编码位置数据1244)。处理器可使用第二位速率,因为第二被跟踪对象1112在图42的ROI 1102到1106中的每一者外部。

[0320] 处理器可使用大于第二位速率的第三位速率将第三坐标数据1236编码为第三经编码位置数据1246。例如,处理器可使用码簿1202将第三坐标数据1236映射到对应值且将第三直角坐标数据1236编码为所述值(例如,第三经编码位置数据1246)。处理器可使用第

三位速率,因为第三被跟踪对象1114在ROI(例如,图42的第一ROI 1102)内部。在特定实施例中,第三位速率可等于或大体上等于第一位速率。

[0321] 处理器可使用大于第二位速率的第四位速率将第四坐标数据1237编码为第四经编码位置数据1247。例如,处理器可使用码簿1202将第四坐标数据1237映射到对应值且可将第四坐标数据1237编码为所述值(例如,第四经编码位置数据1247)。处理器可使用第四位速率,因为第四被跟踪对象1120在ROI(例如,图42的第二ROI 1104)内部。在特定实施例中,第四位速率可大体上等于第一位速率。

[0322] 处理器可使用大于第二位速率的第九位速率将第九坐标数据1238编码为第九经编码位置数据1248。例如,处理器可使用码簿1202将第九坐标数据1238映射到对应值且可将第九坐标数据1238编码为所述值(例如,第九经编码位置数据1248)。处理器可使用第九位速率,因为第九被跟踪对象1134在ROI(例如,图42的第三ROI 1106)内部。在特定实施例中,第九位速率可大体上等于第一位速率。

[0323] 可针对图42中的其它被跟踪对象的坐标数据执行类似编码技术。可以比定位在ROI 1102到1106外部的被跟踪对象(例如,第二被跟踪对象1112)的元数据更高的位速率编码定位在ROI 1102到1106中的一者内部的被跟踪对象的元数据(例如,坐标数据)。以更高的位速率编码定位在ROI 1102到1106中的一者内部的被跟踪对象的元数据可使得能够以相对高水平的精度编码被跟踪对象的位置。

[0324] 图43的实施例可使得用户能够基于使用码簿1202编码的数据而检索(例如,定位)视频的特定部分。例如,位置空间可显示在屏幕上以作为用户交互网格,其使得用户能够选择不同对象的不同位置(例如,参数)。基于所述选择,移动装置的用户可容易定位视频的具有类似对象位置配置的所要部分(例如,帧),而不必导航(例如,播放、快进、快退等)穿过所述视频。另外,经编码位置数据1242到1248可被压缩并且从编码器传输到解码器。压缩经编码位置数据1242到1248且将其传输到解码器可使得解码器能够使用相对低量的数据(例如,经量化值)确定被跟踪对象的位置。

[0325] 参考图44,展示说明使用对象跟踪的视频处理方法1300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、图43的码簿1202或其任何组合执行方法1300。

[0326] 方法1300可包含在1302处在移动装置处跟踪(例如,通过任务T200)位于特定帧的关注区中的第一对象的第一位置。例如,参考图42,可跟踪第一被跟踪对象1110的第一位置。在1304处,可跟踪不位于特定帧的关注区中的第二对象的第二位置。例如,参考图42,可跟踪第二被跟踪对象1112的第二位置。

[0327] 在1306处,可产生与第一位置相关联的第一坐标数据及与第二位置相关联的第二坐标数据。例如,参考图43,处理器可产生与场景1100中的第一被跟踪对象1110的位置相关联的第一坐标数据1232及与场景1100中的第二被跟踪对象1112的位置相关联的第二坐标数据1234。

[0328] 在1308处,可使用码簿编码(例如,通过任务T300)第一坐标数据及第二坐标数据以产生第一经编码位置数据及第二经编码位置数据。例如,参考图43,处理器可使用所述1202编码第一坐标数据1232以产生第一经编码位置数据1242(例如,码簿1202的码字),且

处理器可使用码簿1202编码第二坐标数据1234以产生第二经编码位置数据1244。可以第一位速率(例如,高位速率)编码第一坐标数据1232且可以第二位速率(例如,低位速率)编码第二坐标数据1234。处理器可使用第一位速率编码第一坐标数据1232,因为第一被跟踪对象1110在ROI(例如,图42的第一ROI 1102)内部。处理器可使用第二位速率编码第二坐标数据1234,因为第二被跟踪对象1112在图42的ROI 1102到1106中的每一者外部。

[0329] 在1310处,可存储第一经编码位置数据及第二经编码位置数据以实现基于输入参数对特定帧的检索。例如,参考图43,处理器可将第一经编码位置数据1242及第二经编码位置数据1244存储在存储器中。移动装置的用户可输入参数(例如,第一被跟踪对象1110及第二被跟踪对象1112的所要的位置)。基于所述输入参数,移动装置可响应于对应于第一位置及第二位置的输入参数而检索第三帧以用于回放。

[0330] 在特定实施例中,方法1300可包含跟踪位于特定帧的所关注区中的第三对象的第三位置。例如,可跟踪第三被跟踪对象1114的第三位置。可产生与第三位置相关联的第三坐标数据。可使用码簿编码第三坐标数据以产生第三经编码位置数据。可使用大于第二位速率的第三位速率编码第三坐标数据。可存储所述第三经编码位置数据以实现基于输入参数对特定帧的检索。

[0331] 在特定实施例中,方法1300可包含跟踪位于特定帧的第二关注区中的第四对象的第四位置。例如,可跟踪第四被跟踪对象1120的第四位置。可产生与第四位置相关联的第四坐标数据。可使用码簿编码第四坐标数据以产生第四经编码位置数据。可使用大于第二位速率的第四位速率编码第四坐标数据。可存储第四经编码位置数据以实现基于输入参数对特定帧的检索。

[0332] 图44的方法1300可使得用户能够检索(例如,定位)使用码簿1202编码的视频的特定部分(帧)。另外,方法1300可使得能够以比ROI外部的对象更高的位速率编码ROI内部的被跟踪对象的元数据。以较高的位速率编码定位在ROI内部的被跟踪对象的元数据可使得能够以比其它被跟踪对象的位置更高水平的准确度编码ROI中的被跟踪对象的位置。

[0333] 图38A展示根据一般配置的用于产生元数据的设备A100的框图。设备A100包含经配置以使用至少一个数字图像选择多个对象的鉴别器100,其中所述至少一个数字图像相对于物理空间描绘所述对象(例如,如本文中参考任务T100所描述)。设备A100还包含选择器300,其经配置以基于指示位置空间中的所述选定对象的位置的信息而产生识别所述选定对象的多个候选几何布置中的一者的元数据(例如,如本文中参考任务T300所描述)。

[0334] 图38B展示设备A100的实施方案A110的框图。设备A110包含经配置以显示至少一个数字图像的显示模块50(例如,如本文中参考显示任务T50所描述)。

[0335] 图38C展示设备A100的实施方案A120的框图。设备A120包含经配置以解析视频流以产生至少一个数字图像的接口30(例如,如本文中参考解析任务T30所描述)。图38D展示设备A110及A120的实施方案A130的框图。

[0336] 图38E展示设备A100的实施方案A200的框图。设备A200包含经配置以确定位置空间中的所述多个对象的参考位置的对象跟踪器200(例如,如本文中参考任务T200所描述)。对象跟踪器200还可经配置以在视频流内跟踪所述多个对象(例如,如本文中参考任务T200所描述)。图38F、38G及39A分别展示包含对象跟踪器200的实例的设备A110、A120及A130的实施方案A210、A220及A230的框图。

[0337] 图39B展示设备A100的实施方案A300的框图。设备A200包含经配置以使由选择器300产生的元数据与视频流的对应帧相关联的视频流编辑器400(例如,如本文中参考任务T300所描述)。图39C及39D分别展示包含视频流编辑器400的实例的设备A200及A230的实施方案A310及A330的框图。

[0338] 如本文中所揭示的设备A100和/或A500的实施方案中的任一者可实施在移动装置内,例如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)。移动装置可包含经配置以显示至少一个数字图像的一或多个屏幕(例如,触摸屏)及/或经配置以捕获至少一个数字图像(例如,视频流)的一或多个成像单元(例如,相机或其它图像捕获装置)。

[0339] 参考图31,展示包含可操作以执行上文描述的视频处理技术的组件的无线装置2900的框图。装置2900包含耦合到存储器2932的处理器2910,例如数字信号处理器(DSP)。装置2900是可实施以包含如本文中所描述的设备A100和/或A500的实施方案中的任一者的元件的移动装置的一个实例。

[0340] 图31还展示耦合到处理器2910及显示器2928的显示控制器2926。在特定实施例中,所述显示器可对应于上文描述的取景器或屏幕。存储器2932可为包含可执行指令2956的有形非暂时性处理器可读存储媒体。指令2956可由例如处理器2910等处理器执行。例如处理器2910或显示控制器2926等处理器可执行指令2956以执行图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23的方法2300、图25的方法2500、图26的方法2600、图28的方法2800、方法M100、方法M110、方法M120、方法M130或其任何组合的至少一部分。

[0341] 译码器/解码器(CODEC)2934还可耦合到处理器2910。扬声器2936及麦克风2938可耦合到CODEC 2934。图31还指示无线控制器2940可耦合到处理器2910及天线2942。在特定实施例中,射频(RF)接口2980安置于无线控制器2940与天线2942之间。

[0342] 在特定实施例中,处理器2910、显示控制器2926、存储器2932、CODEC 2934及无线控制器2940包含在系统级封装或系统芯片装置2922中。在特定实施例中,输入装置2930(例如,如上文所描述的用于捕获场景的相机)及电力供应器2944耦合到系统芯片装置2922。此外,在特定实施例中,如图31中所说明,显示器2928、输入装置2930、扬声器2936、麦克风2938、天线2942、电力供应器2944及RF接口2980在系统芯片装置2922外部。例如,RF接口2980可为单独芯片。然而,显示器2928、输入装置2930、扬声器2936、麦克风2938、天线2942、电力供应器2944及RF接口2980中的每一者可耦合到系统芯片装置2922的组件,例如接口或控制器。

[0343] 上文描述使用位置相关元数据注释视频文件的众多实例(例如,参考任务T300)。此类注释可包含(例如)针对例如帧分类及集群以用于检索和或存档等应用而编码通过对象跟踪获得的位置。

[0344] 此类功能性可用于产生识别用户感兴趣的内容的一部分的元数据(例如,在俘获视频流时)。可能需要通过提供用于基于元数据对指定内容的快速检索的机制而充分利用对一或多个视频流的此基于兴趣的预滤波的结果。例如,可能需要支持对所要的内容的自

动提取(例如,从大量视频文件中),同时减少或消除手动检视的需要。

[0345] 图45A展示根据一般配置的图像检索的方法M500的流程图,其包含任务X100、X200及X300。任务X100获得识别多个选定对象的信息,且任务X200选择多个候选几何布置中的一者。响应于所述选择,任务X300识别多个数字图像中的描绘根据选定候选几何布置而布置的多个选定对象的至少一个数字图像。任务X100、X200及X300可例如由图1的电子装置102、处理器2910、无线装置2900或其任何组合执行。

[0346] 任务X100获得识别多个选定对象的信息。可实施任务X100以从用户获得所述信息(例如,经由触摸屏,例如触摸屏116)。在此情况下,用户可从触摸屏上显示的菜单选择对象中的一或多个者。例如,此菜单可为图标的显示,每一图标对应于特定一人(例如,用户的联系人列表中的人的缩略图像)。

[0347] 在另一实例中,用户可从触摸屏上的已经基于与一或多个视频文件相关联的元数据而产生的显示选择对象中的一或多个者。例如,屏幕可经配置以显示已经在视频文件内被跟踪的对象的图标的菜单(例如,在俘获期间)。在此情况下,与视频文件相关联的元数据可包含使识别被跟踪对象的标记与视频文件的一或多个帧中的每一者中的对象的对应位置相关联的信息。

[0348] 类似地,用户可从具有相关联的元数据的图像(例如,视频文件的帧)选择对象中的一或多个者。在此情况下,触摸屏接口可经配置以通过突出显示图像内的对象而显示所述图像,以指示所述对象的位置元数据在对应的视频文件中可用(即,所述对象已被跟踪)。在一个此类实例中,触摸屏接口可经配置以允许用户一次选择多个对象(例如,通过围绕对象拖拽套索,或通过使用夹捏及缩放动作一围绕对象产生方框)。

[0349] 作为此触摸屏输入的补充或替代,可实施任务X100以通过语音辨识从用户获得信息。用户可例如通过从屏幕上显示的可用的对象的列表(例如,如上文所描述的图标的菜单)读取对象的名称而选择所述对象中的一或多个者。或者,用户可简单地假设特定对象的位置元数据存在于待检索的至少一个视频文件中,且可通过说出对象的名称、选择与所述对象相关联的图标(例如,从联系人列表),或作为文本输入所述对象的名称而输入此选择。

[0350] 任务X200选择多个候选几何布置中的一者(例如,如本文中参看图35所描述)。可实施任务X200以基于从用户获得的信息(例如,经由触摸屏,例如触摸屏116)而执行此类选择。在一个实例中,用户指定描绘所要的布置的图像(例如,视频文件的帧)。在另一实例中,用户通过在此布置中放置图标(例如,通用图标或所要的对象的图标)而描述所要的布置。应注意,采用这两个实例中的任一者的触摸屏配置可用于在单个用户操作中将用户选择信息提供给任务X100及X200两者。例如,用户可使用被追踪对象的图标的显示来描述所要的布置且通过使用它们产生描述而指示选定对象。在另一实例中,用户可从此类布置的显示当中选择所要的候选几何布置。

[0351] 另外或替代地,可实施任务X200以通过语音辨识从用户获得信息。在此情况下,用户可使用描述符(例如,‘靠近’、‘附近’、‘远’、‘旁边’、‘离开’)指示所要的布置。例如,语音辨识接口可经配置以在以下形式中的任一或多个中辨识搜索请求:“找到紧靠着某人及凯特的汤姆”、“展示远离吉尔及山姆的汤姆”、“显示接近莎拉及远离吉尔的汤姆”及“搜索紧靠着莎拉并远离树的汤姆”。而且,此操作可用于在单个用户操作中将用户选择信息提供给任务X100及X200两者。

[0352] 在以上实例中的任一者中,可实施任务X200以通过例如将适当的决策度量(例如,如上文参看图35所描述)应用到由用户指示的布置而选择候选几何布置中的特定一者。在另一实例中,用户可从所述布置的显示当中选择所要的候选几何布置。

[0353] 响应于所述选择,任务X300识别多个数字图像中的描绘根据选定候选几何布置而布置的多个选定对象的至少一个数字图像。在一个实例中,所获得的信息包含多个选定对象中的每一者的与所述对象相关联的标记,且实施任务X300以搜索与多个数字图像(例如,视频文件)相关联的元数据内的标记。此标记可为对象的名称或在元数据内使用以指示对象的任何其它唯一或非唯一数据串。例如,人的标记可为与用户的联系人列表中的人的缩略图相关联的数据串。

[0354] 与多个数字图像相关联的元数据还可经配置以含有指示特定候选者的标记(例如,数据串)。在此情况下,可实施任务X300以在元数据内搜索与选定候选几何布置相关联的标记。

[0355] 或者,元数据可能缺乏指示任何候选几何布置的此数据串但仍可含有指示每一选定对象的位置的信息。在此情况下,可实施任务X200以基于此位置元数据而决定对象是否根据选定候选几何布置而布置。可应用任务X200的此实施方案(例如,如本文中所描述的一或多个决策度量(例如,参看图35))以执行此确定。在许多情况下由图像的元数据指示的位置之间的距离的比较可足以不用进一步考虑图像。

[0356] 可能需要使用如本文中所描述的方法M100的实施方案以基于选定对象的几何布置对视频文件或流(或数字图像的另一集合)进行滤波以用于信息检索。例如,可能需要使用此实施方案以识别其中在特定布置中找到人的图像。此类检索的潜在应用可包含指示选定对象(例如,人)的接近度以支持此类出现与其它事件的相关度。

[0357] 可能需要应用方法M100的实施方案以搜索多个所存档的视频文件(例如,响应于单个用户选择操作)。有可能多个视频文件的此群组中的每一者中的元数据的配置可在文件间不同。例如,与特定视频文件相关联的元数据可包含识别对象的元数据及指示位置的元数据,但可能缺乏两者之间的关联。在此情况下,可给视频文件加标签以指示在文件中展示特定人,且所述文件还可包含从跟踪视频流内的选定人获得的元数据,但被跟踪的人可能不是标签中识别的同一人。可能仍然需要将方法M100的实施方案的单个实例应用到此群组(例如,其中所有文件被识别为描绘选定对象的群组)。例如,一旦已通过识别具有被追踪对象的所要的布置的帧而结束此方法,可检视所述组所识别的帧以仅分离出其中被追踪对象也是选定对象的那些帧。

[0358] 如上文参考任务T200及T300所描述,与视频帧相关联的元数据可指示位置空间中的几何布置的参考位置。可实施方法M500以选择选定候选几何布置的参考位置(例如,基于选定候选几何布置的位置的用户的指示)。在此情况下,可实施任务X300以将其中根据候选几何布置相对于位置空间中的参考位置而布置的选定对象的图像与其中根据候选几何布置但不相对于位置空间中的参考位置而布置的选定对象的图像区分开。

[0359] 图45B展示根据一般配置的用于检索的设备A500的框图。设备A500包含鉴别器Z100,其经配置以获得(A)识别多个选定对象的信息及(B)多个候选几何布置中的一者的选择。鉴别器Z100可经配置以例如经由语音辨识接口及/或触摸屏接口而获得所述信息及/或选择。

[0360] 设备A500还包含搜索引擎Z300,其经配置以识别多个数字图像中的描绘根据选定候选几何布置而布置的多个选定对象的至少一个数字图像。在一个实例中,所获得的信息包含多个选定对象中的每一者的与对象相关联的标记,且实施搜索引擎Z300以搜索与多个数字图像相关联的元数据内的标记。另外或替代地,可实施搜索引擎Z300以在与多个数字图像相关联的元数据内搜索与选定候选几何布置相关联的标记。

[0361] 图45C展示根据另一一般配置的用于图像检索的设备MF500的框图。设备MF500包含用于获得识别多个选定对象的信息的装置Y100及用于选择多个候选几何布置中的一者的装置Y200。设备MF500还包含用于识别多个数字图像中的描绘根据选定候选几何布置而布置的多个选定对象的至少一个数字图像的装置Y300。在一个实例中,所获得的信息包含多个选定对象中的每一者的与对象相关联的标记,且实施装置Y300以搜索与多个数字图像相关联的元数据内的标记。另外或替代地,可实施装置Y300以在与多个数字图像相关联的元数据内搜索与选定候选几何布置相关联的标记。

[0362] 图40A展示根据一般配置的用于产生元数据的设备MF100的框图。设备MF100包含用于使用至少一个数字图像选择多个对象的装置F100,其中所述至少一个数字图像相对于物理空间描绘所述对象(例如,如本文中参考任务T100所描述)。设备MF100还包含用于基于指示位置空间中的所述选定对象的位置的信息而产生识别所述选定对象的多个候选几何布置中的一者的元数据的装置F300(例如,如本文中参考任务T300所描述)。

[0363] 图40B展示设备MF100的实施方案MF110的框图。设备MF110包含用于显示至少一个数字图像的装置F50(例如,如本文中参考显示任务T50所描述)。

[0364] 图40C展示设备MF100的实施方案MF120的框图。设备MF120包含用于解析视频流以产生至少一个数字图像的装置F30(例如,如本文中参考解析任务T30所描述)。图40D展示设备MF110及MF120的实施方案MF130的框图。

[0365] 图40E展示设备MF100的实施方案MF200的框图。设备MF200包含用于确定位置空间中的多个对象的参考位置的装置F200(例如,如本文中参考任务T200所描述)。图41B展示设备MF100的实施方案MF400的框图,其包含用于在视频流内跟踪所述多个对象的装置F210(例如,如本文中参考任务T200所描述)。图40F、40G及41A分别展示包含装置F200的实例的设备MF110、MF120及MF130的实施方案MF210、MF220及MF230的框图。还可实施此设备以包含装置F210以作为装置F200的补充或替代。

[0366] 图41C展示设备MF100的实施方案MF300的框图。设备MF200包含用于使由装置F300产生的元数据与视频流的对应帧相关联的装置F400(例如,如本文中参考任务T300所描述)。图41D及41E分别展示包含装置F400的实例的设备MF200及MF230的实施方案MF310及MF330的框图。

[0367] 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、MF120、MF130、MF200、MF210、MF220、MF230、MF300、MF310、MF330、MF400、MF500、A100、A110、A120、A130、A200、A210、A220、A230、A300、A310、A330及A500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。

[0368] 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。

[0369] 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。

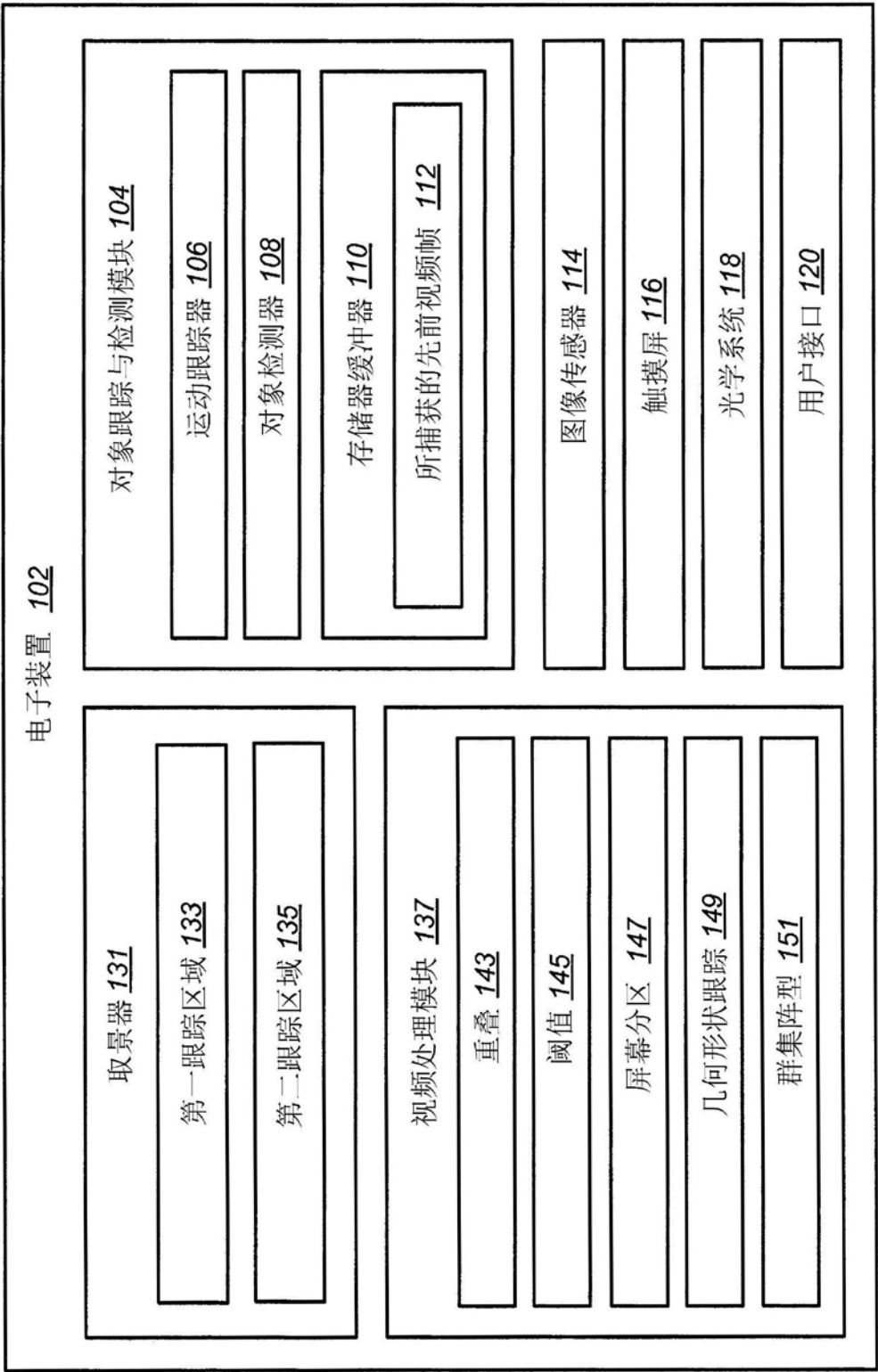


图1

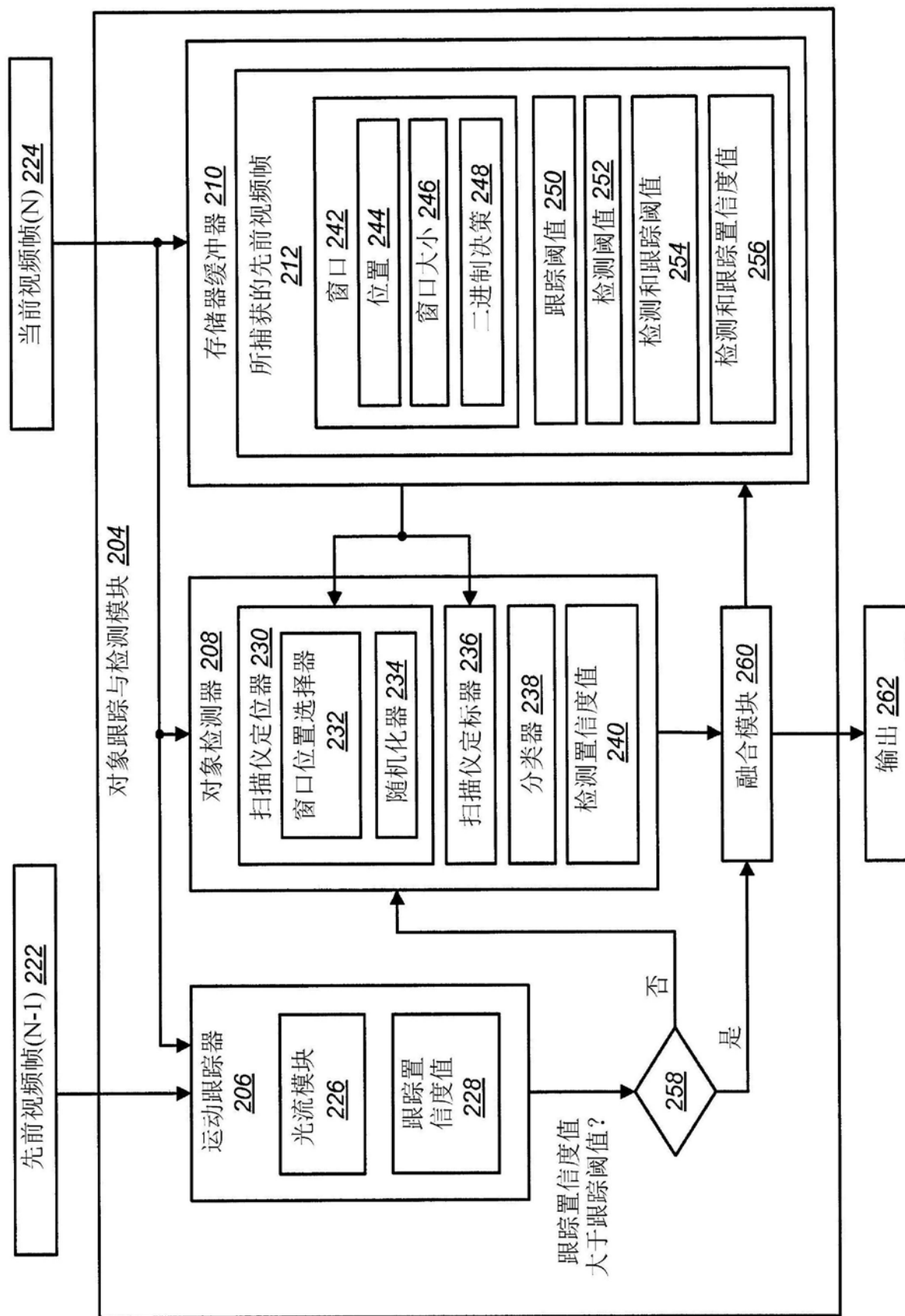


图2A

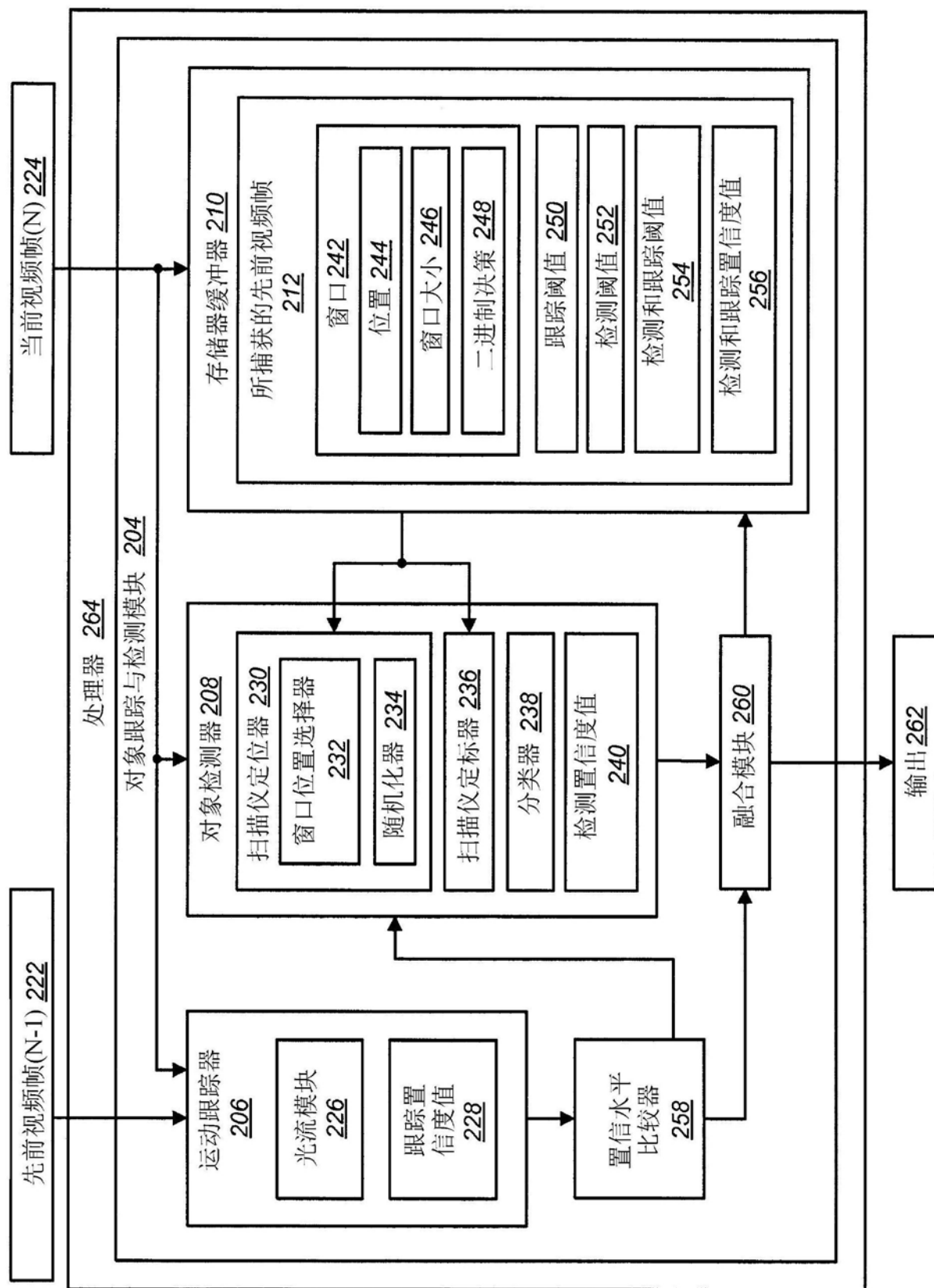


图2B

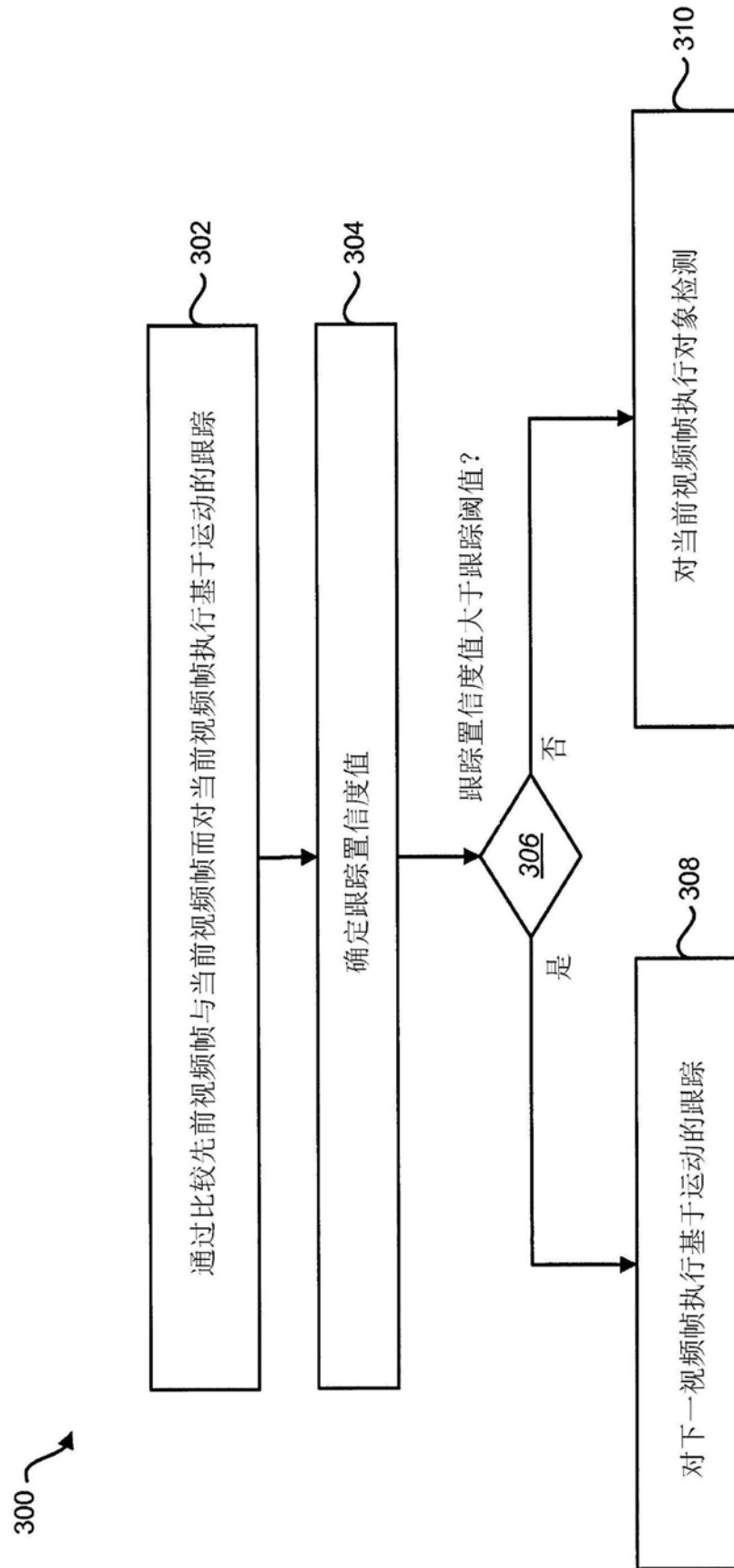


图3

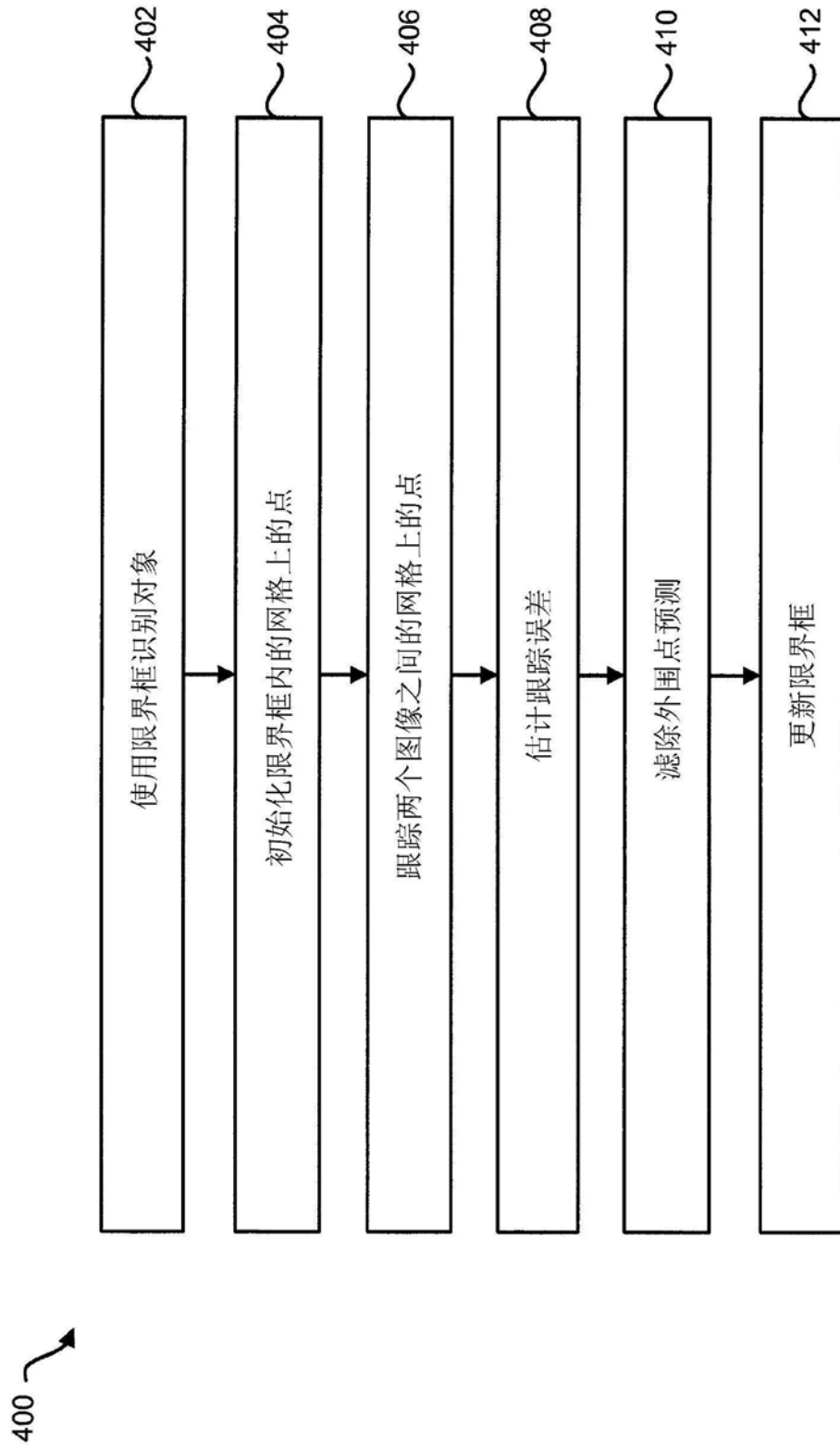


图4

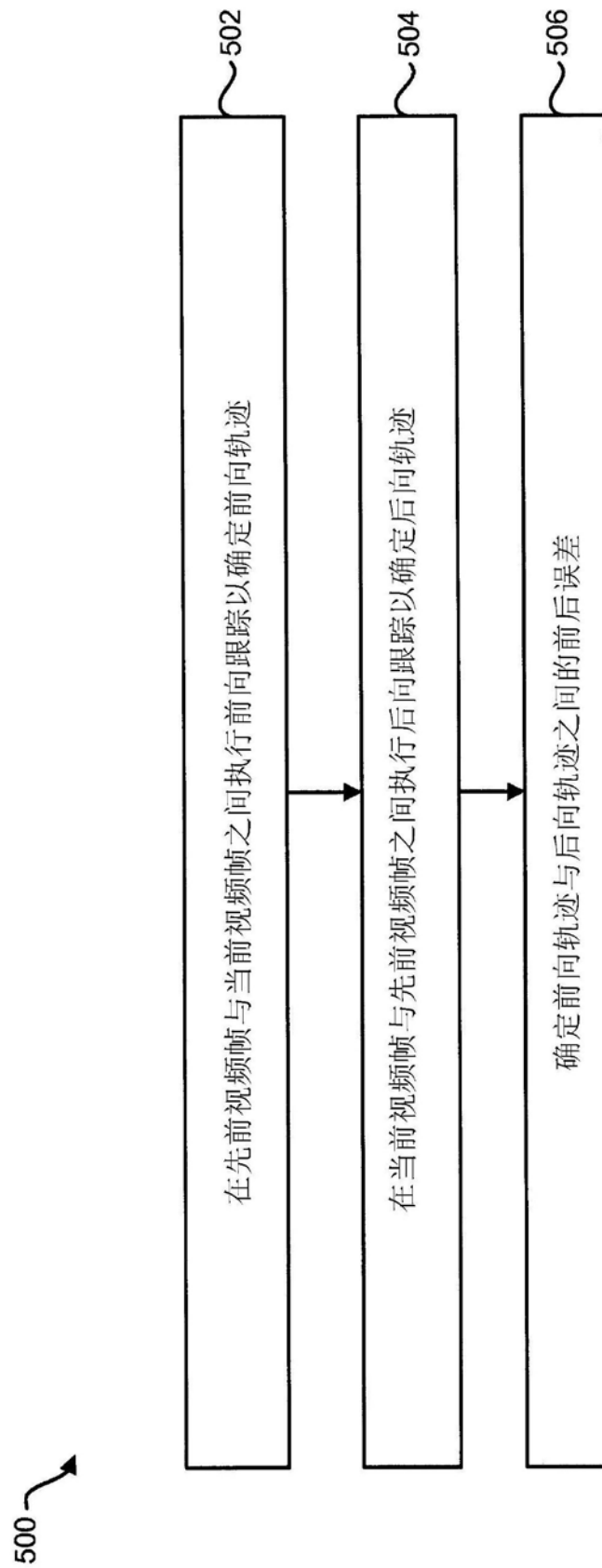


图5

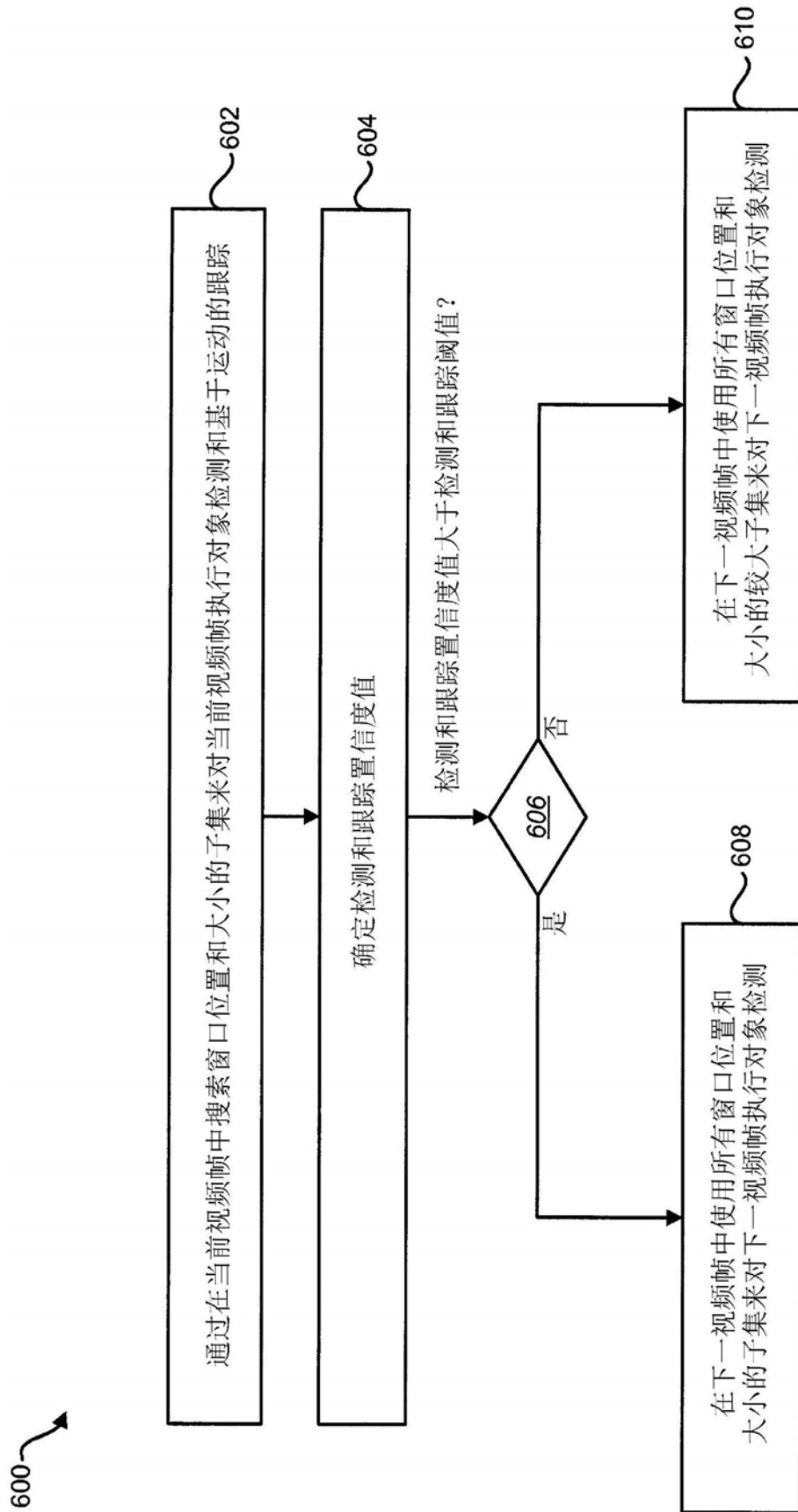


图6

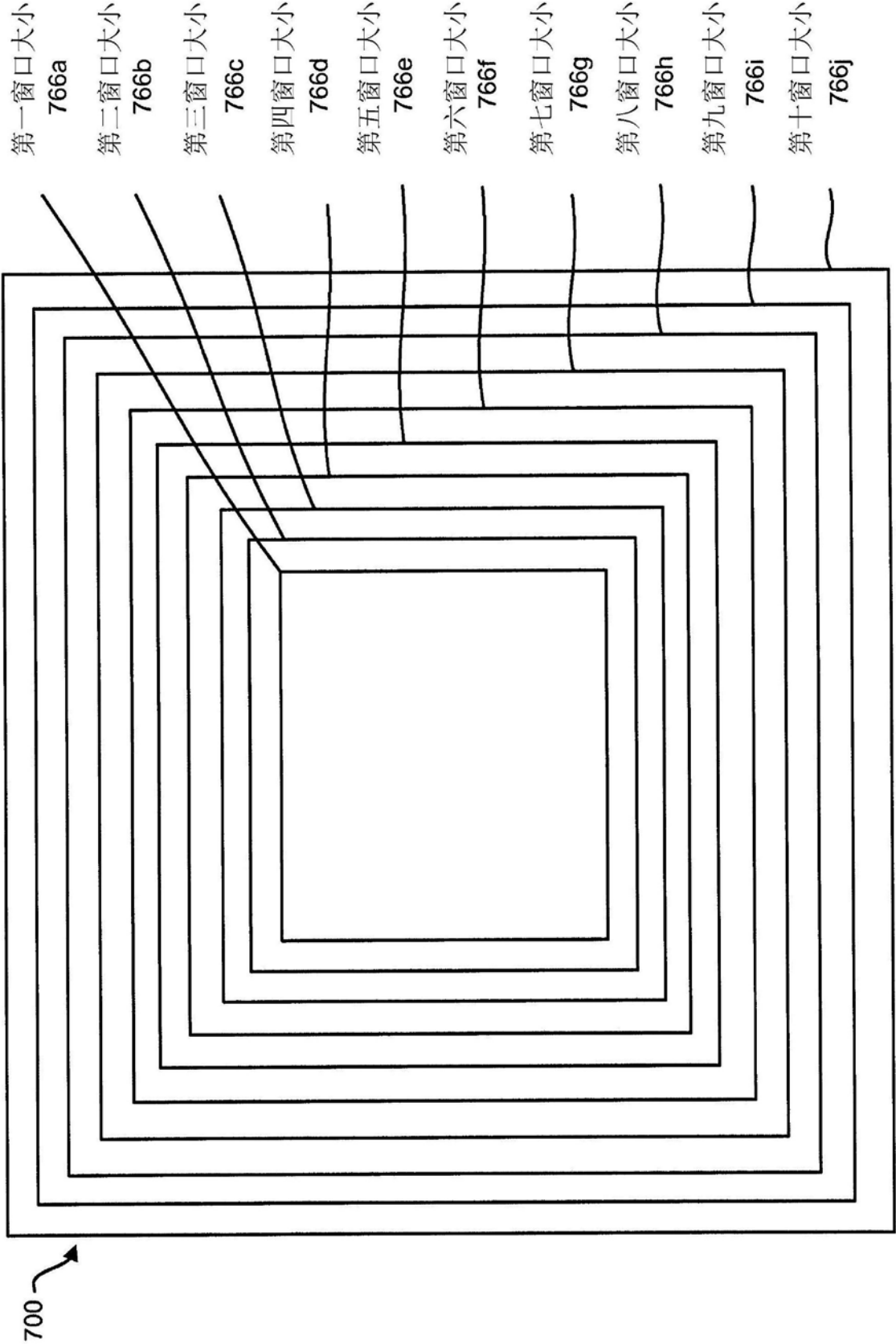


图7

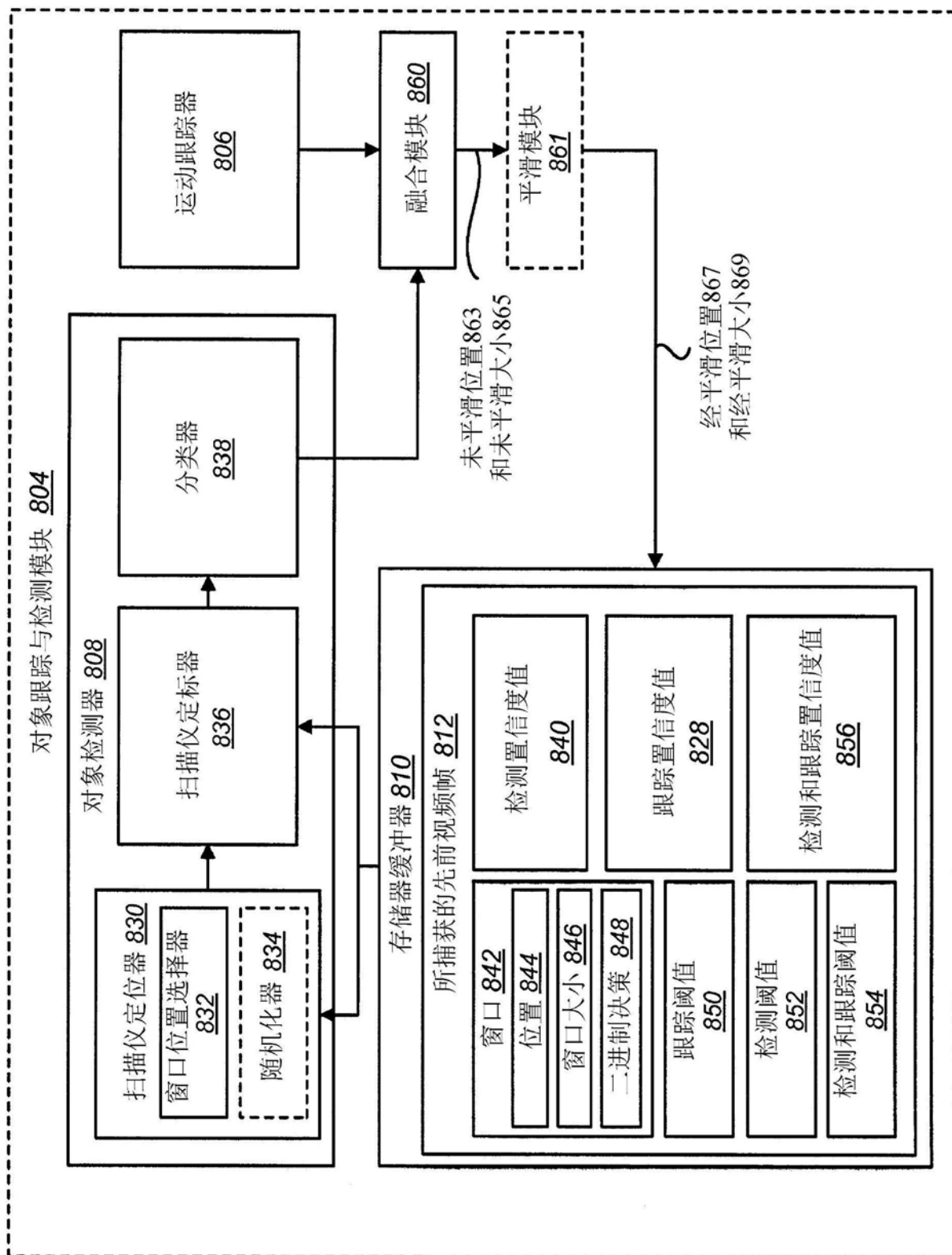


图8

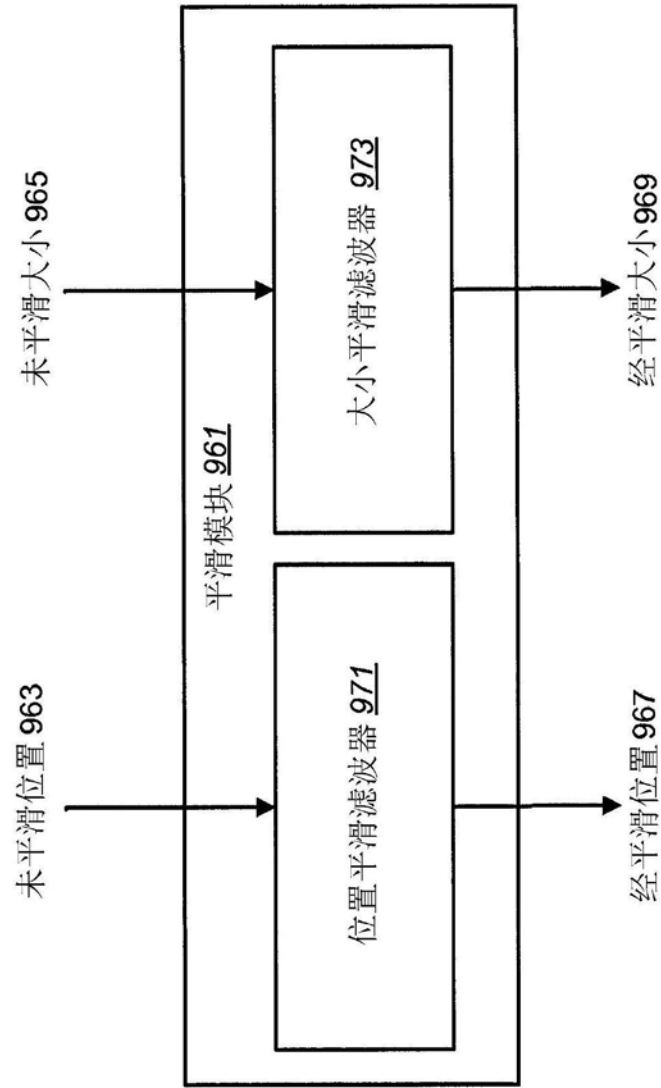


图9

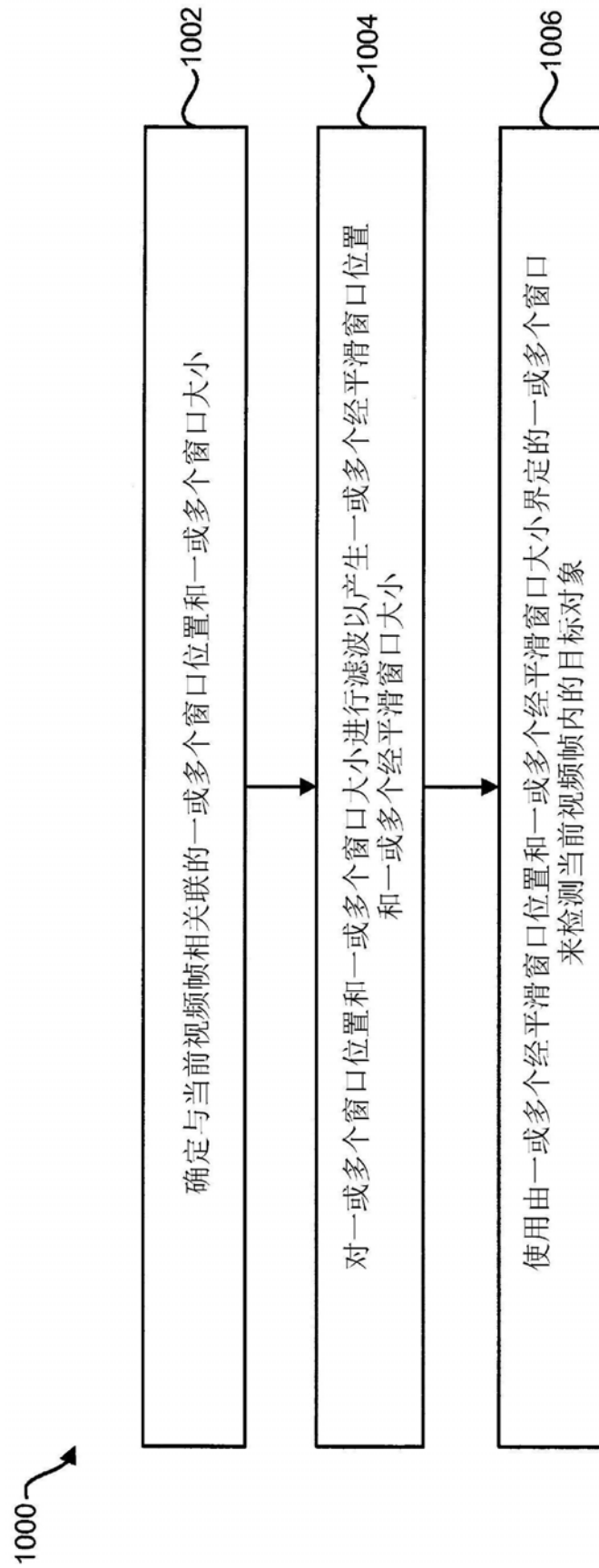


图10

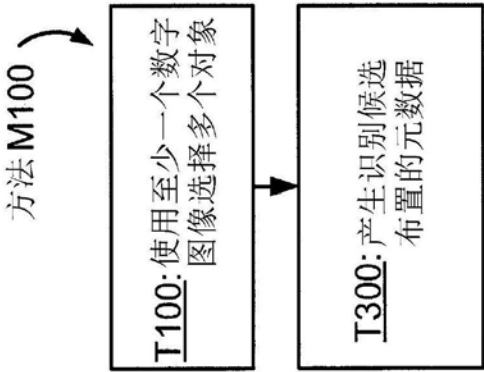


图11A

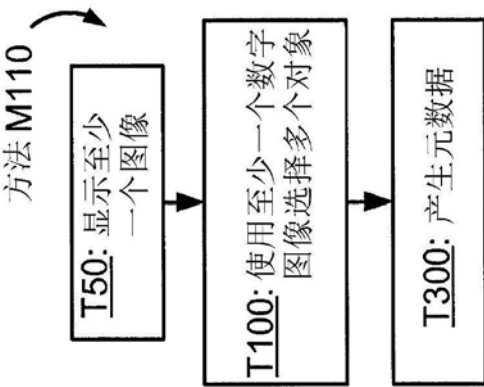


图11B

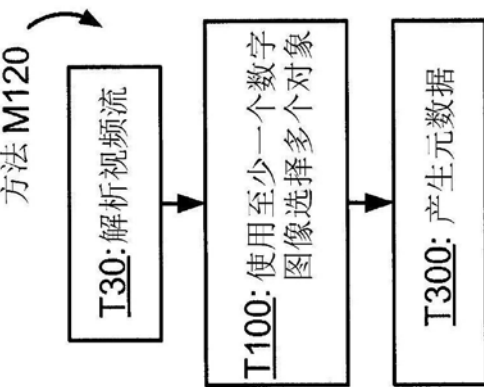


图11C

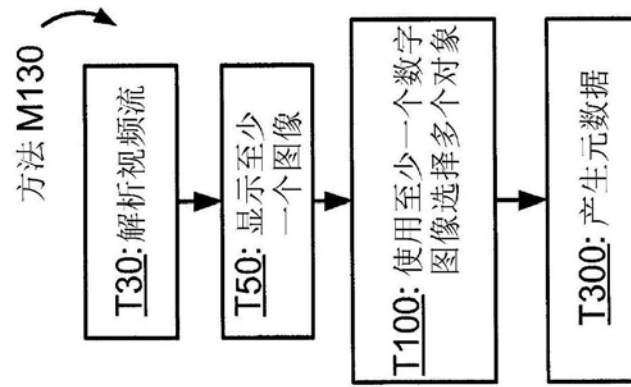


图11D

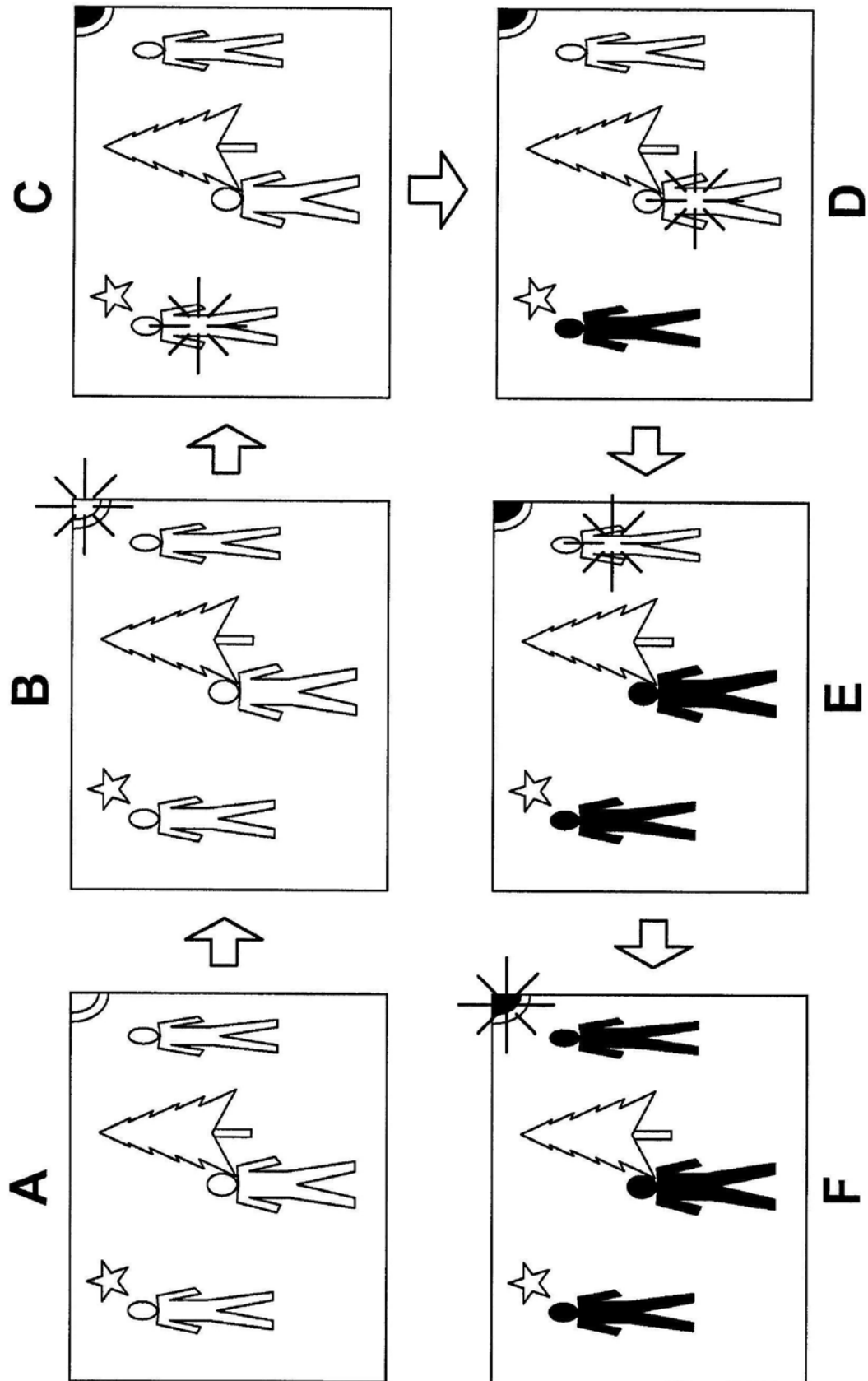


图12

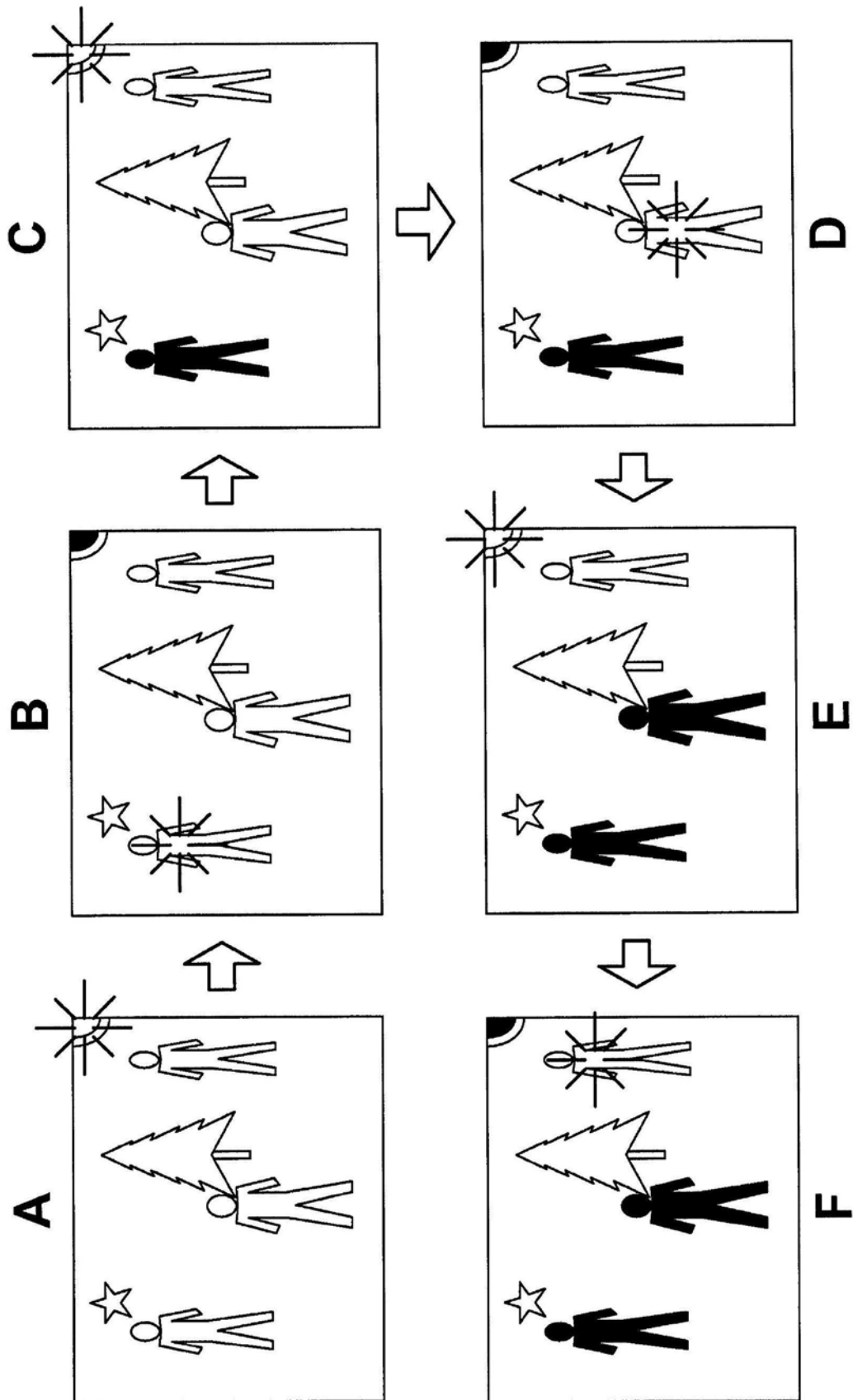


图13

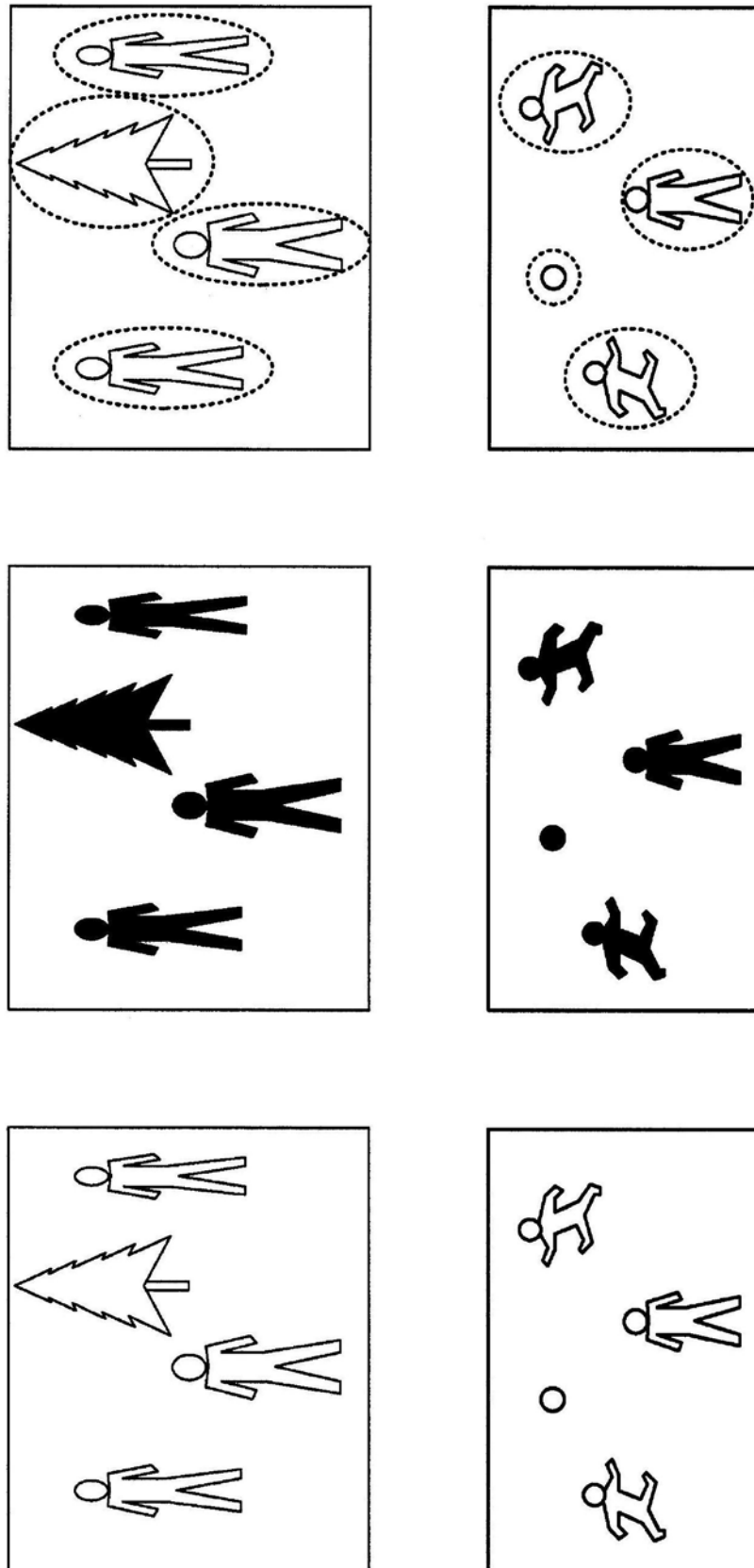


图14

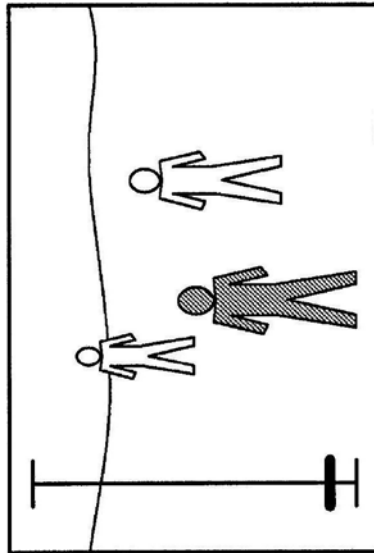
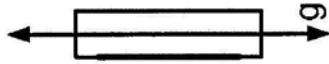


图15A

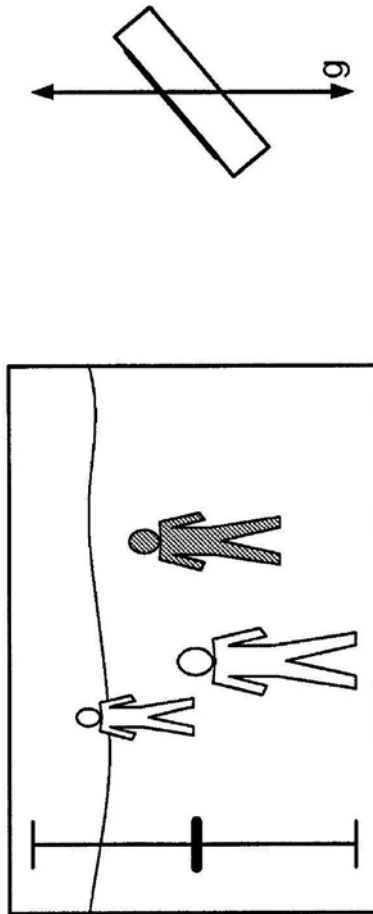


图15B

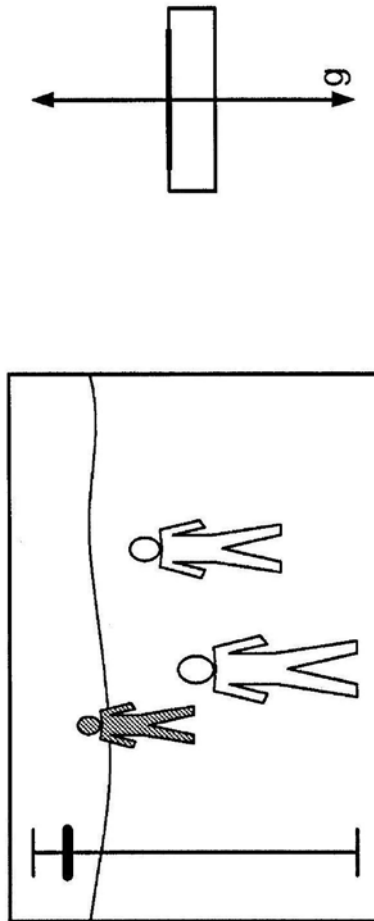


图15C

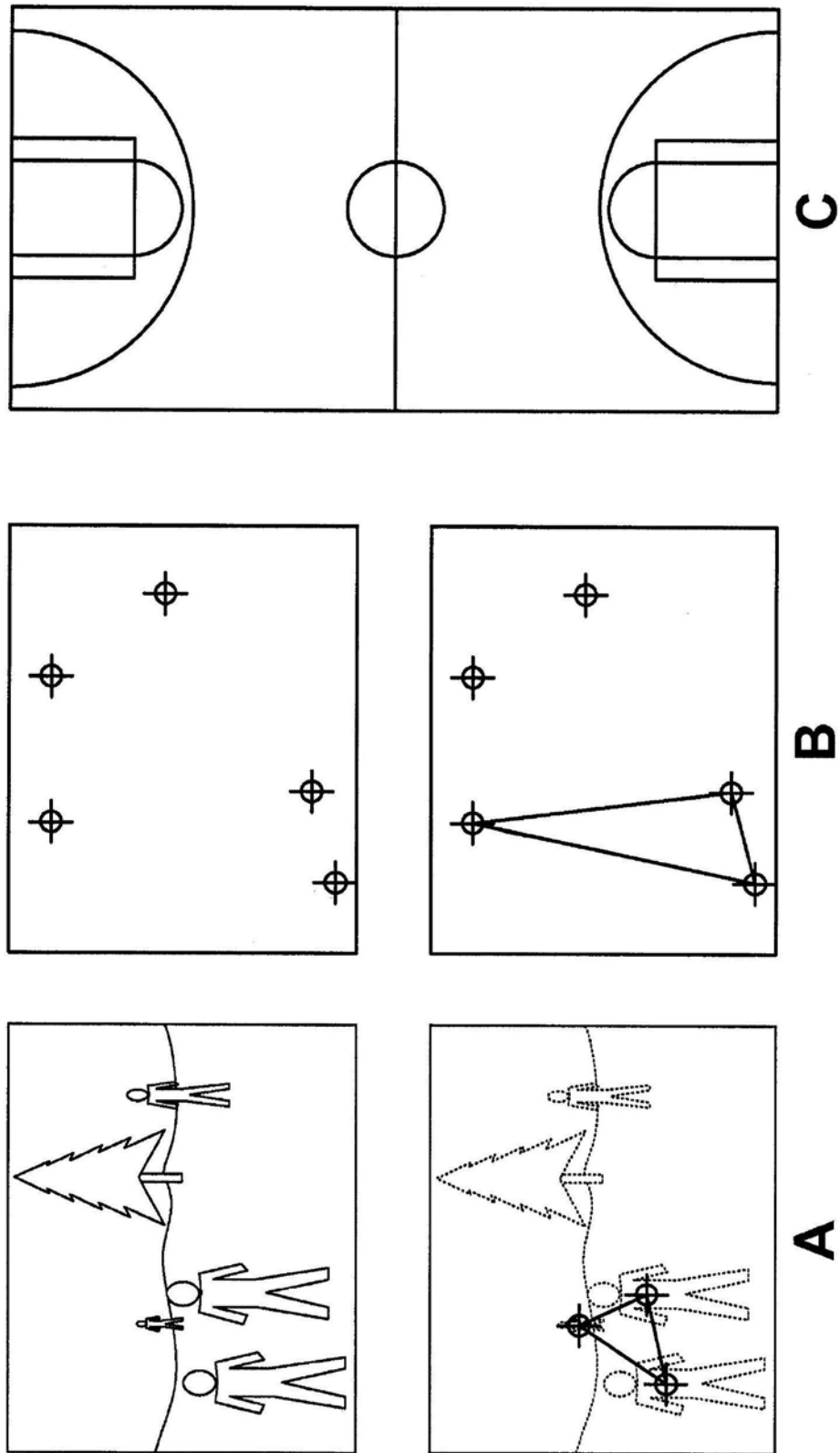


图16

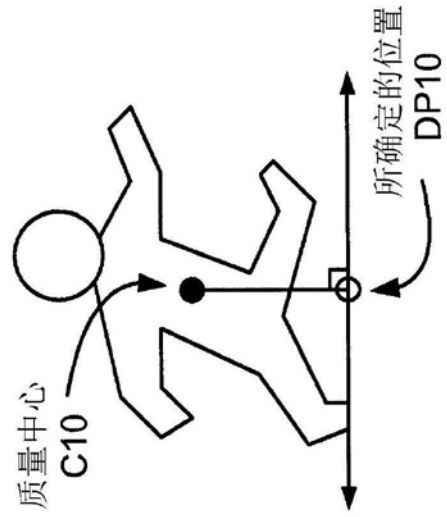


图17A

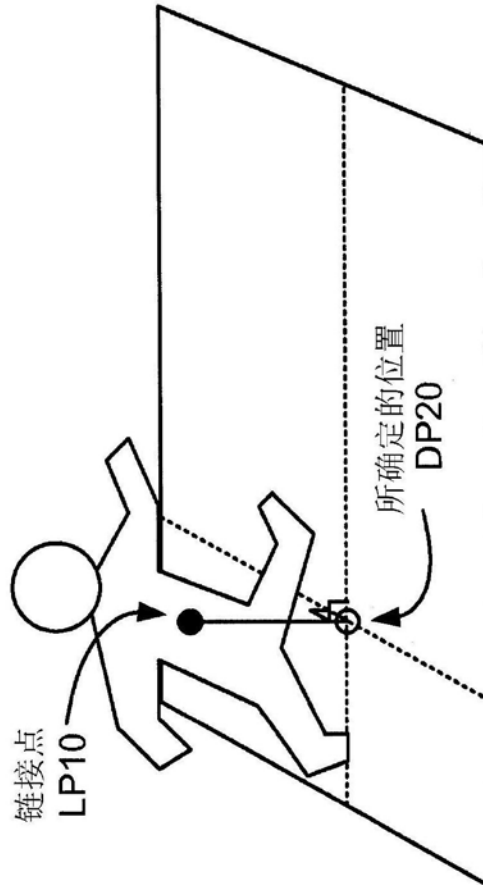


图17B

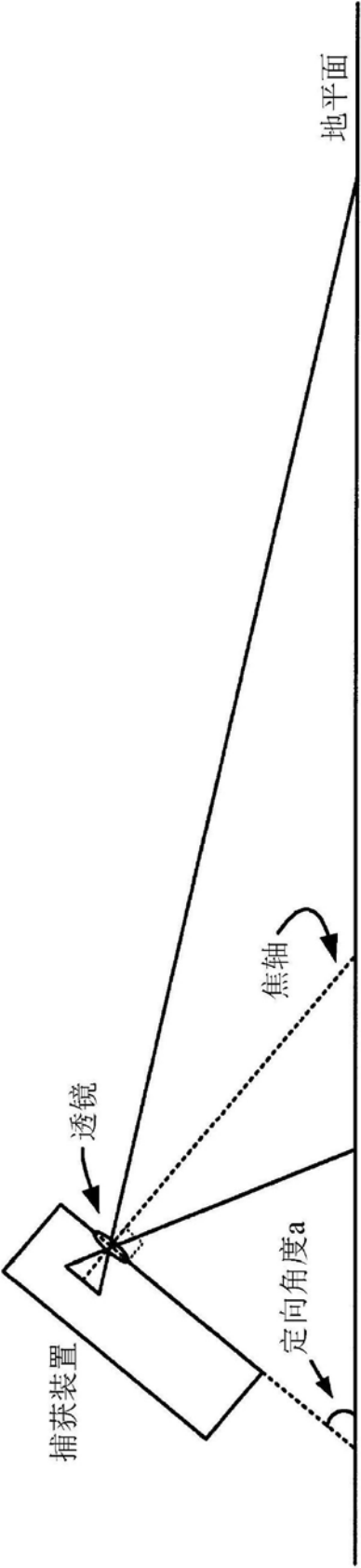


图18A

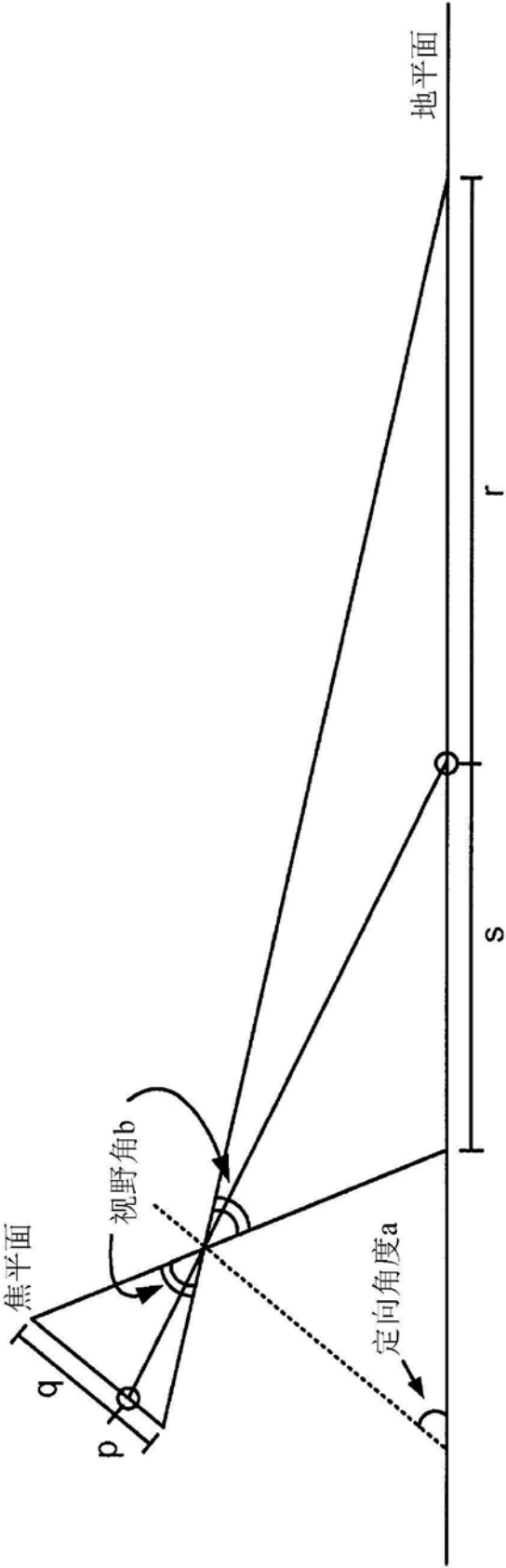


图18B

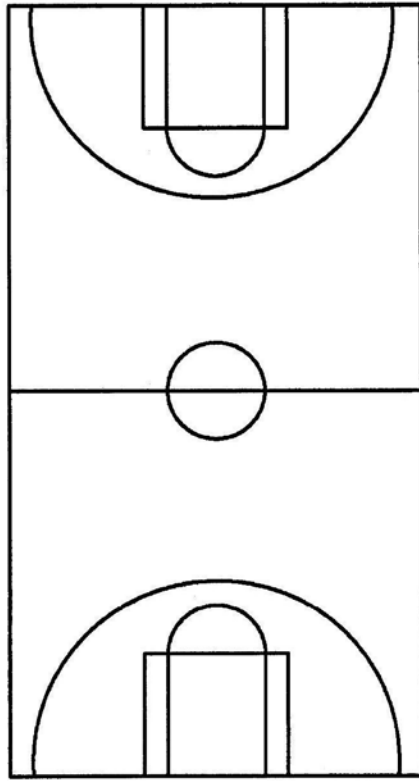


图19A

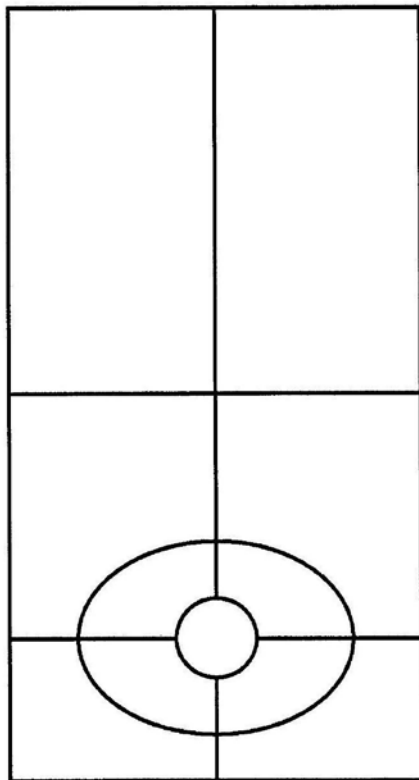


图19B

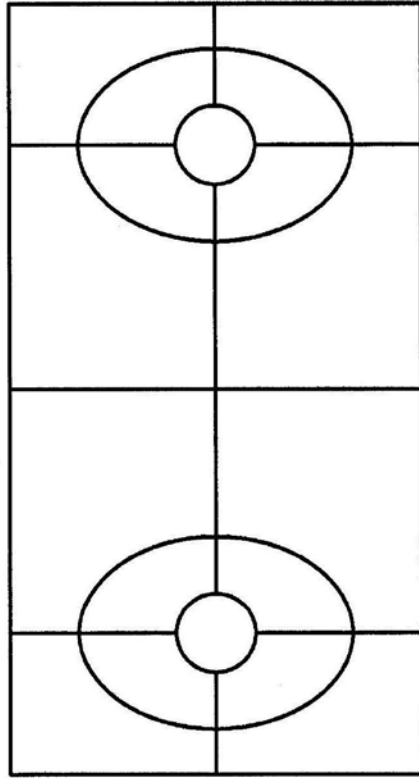


图19C

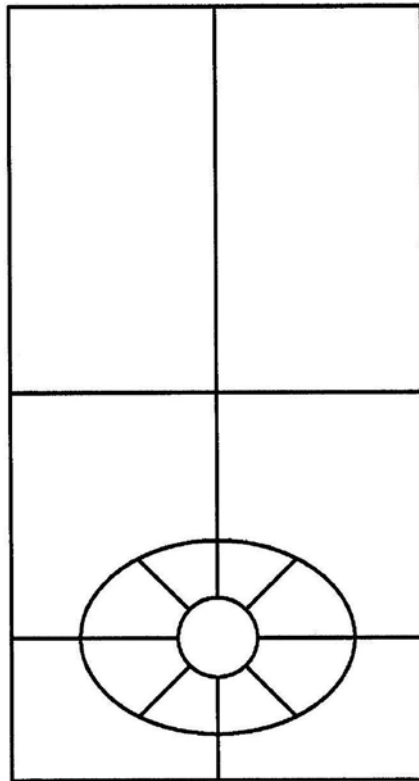


图19D

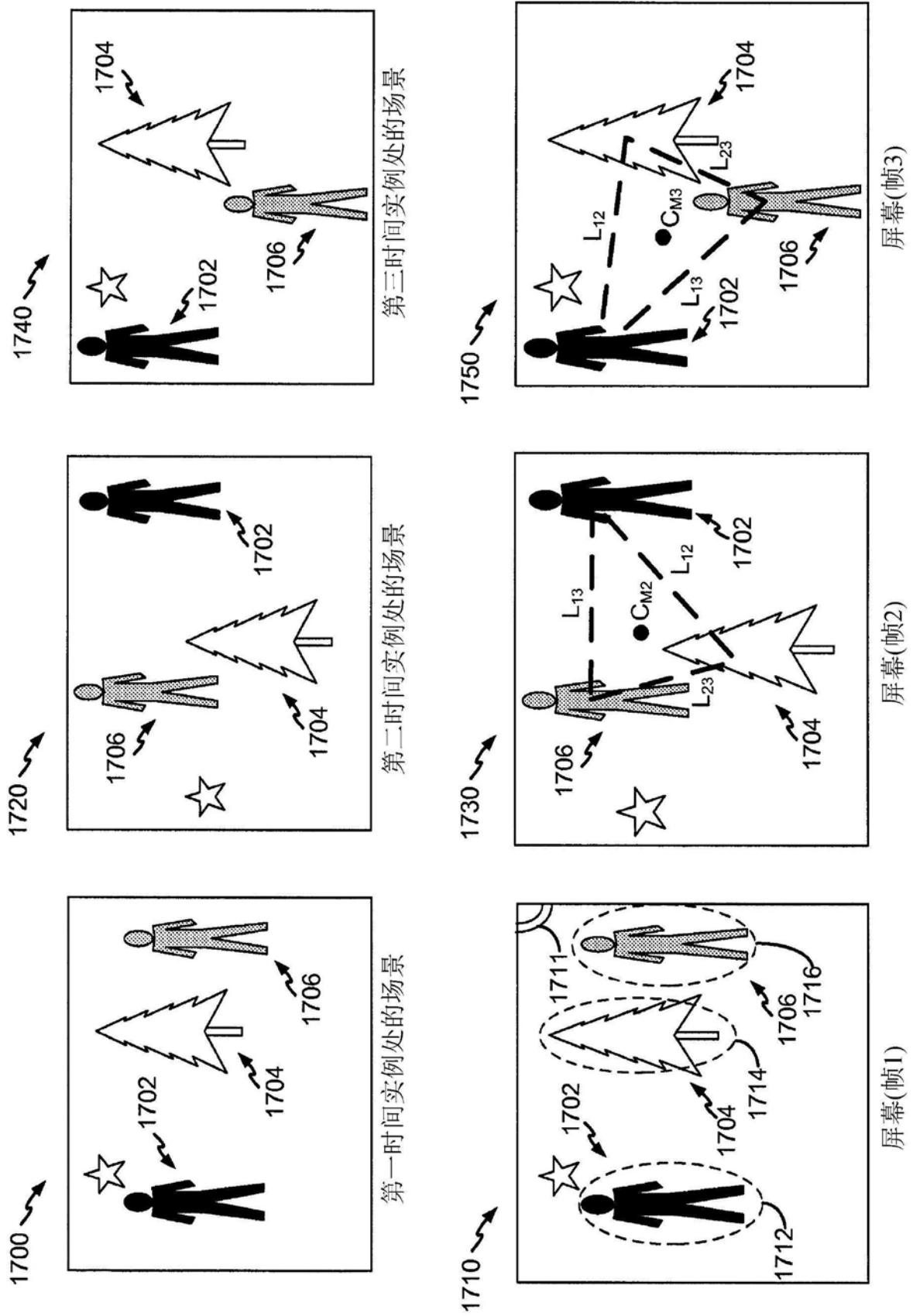


图20

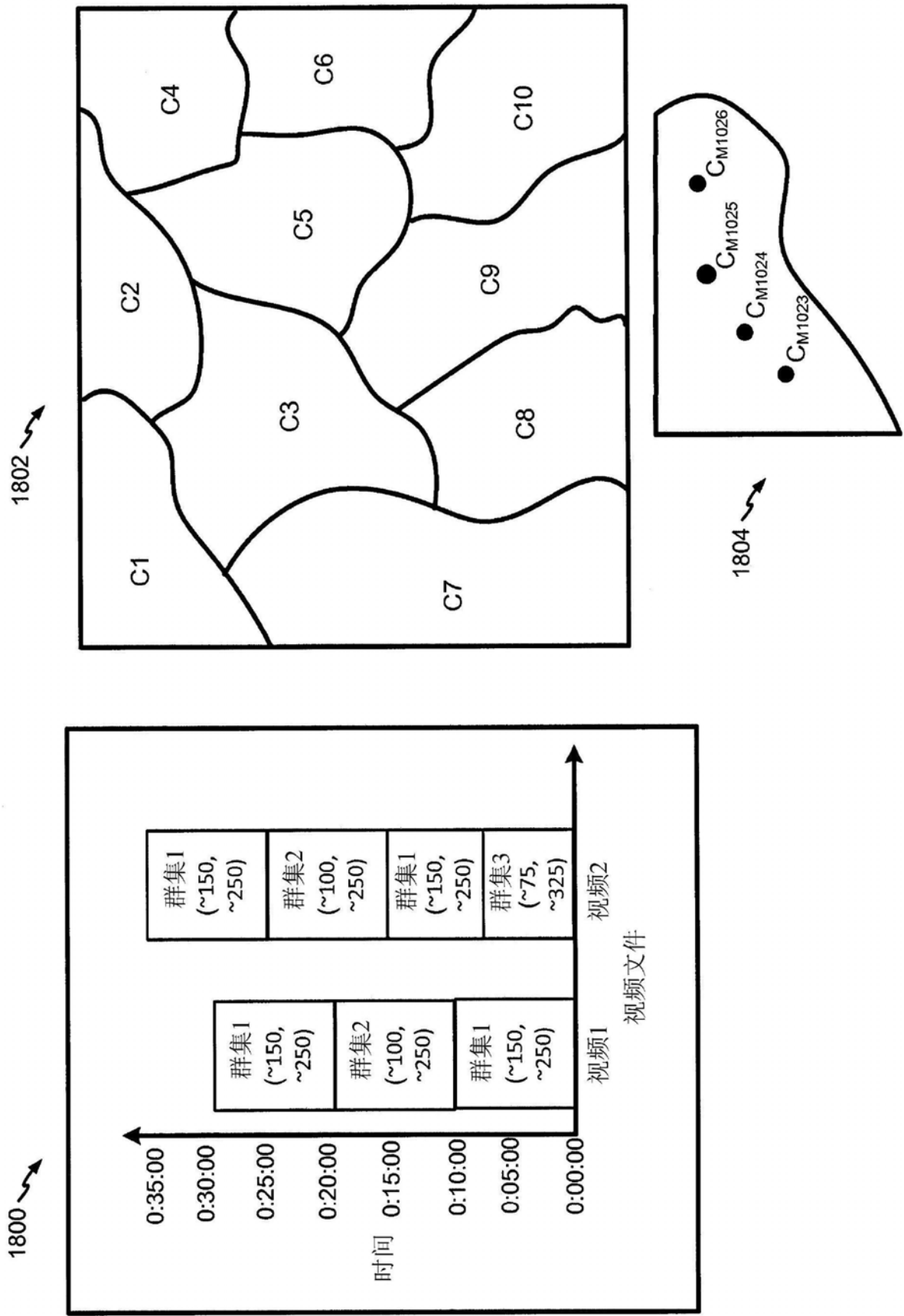


图21

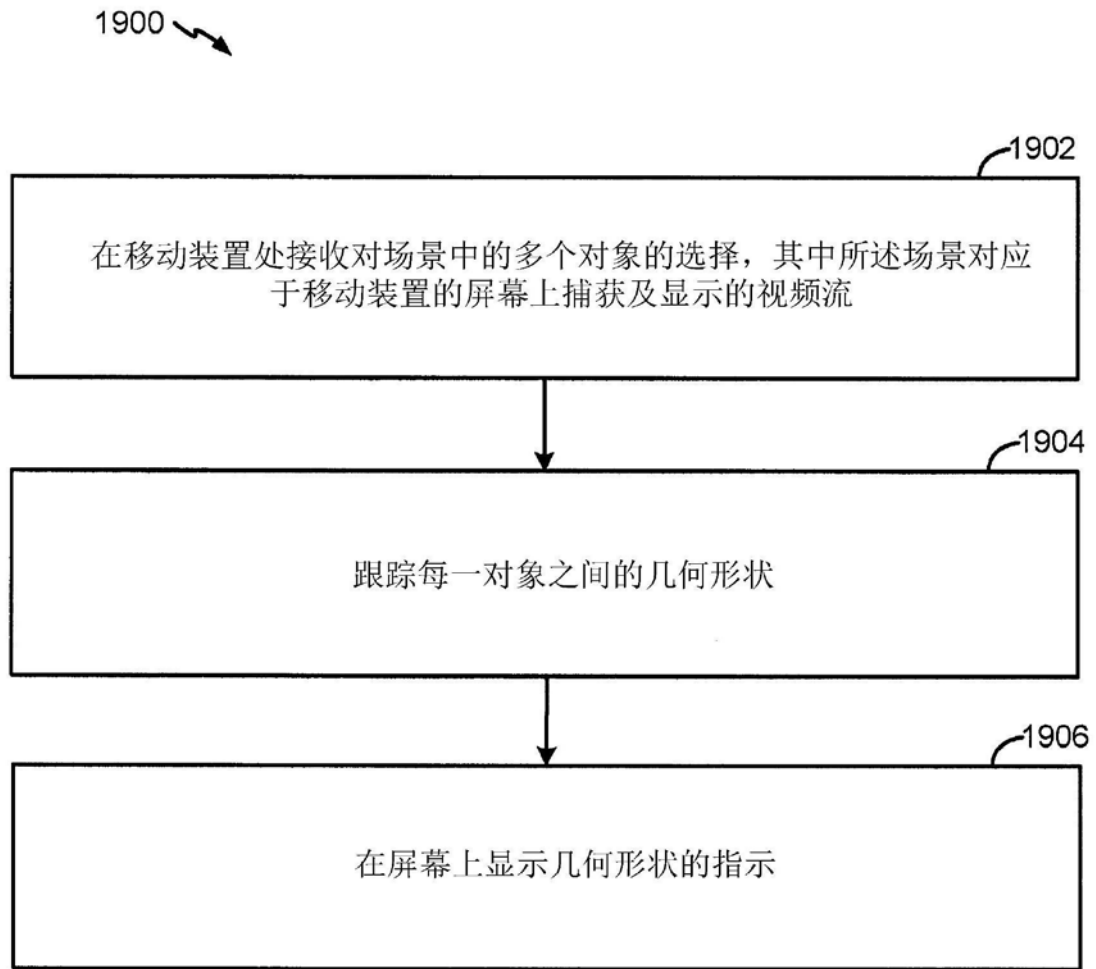


图22

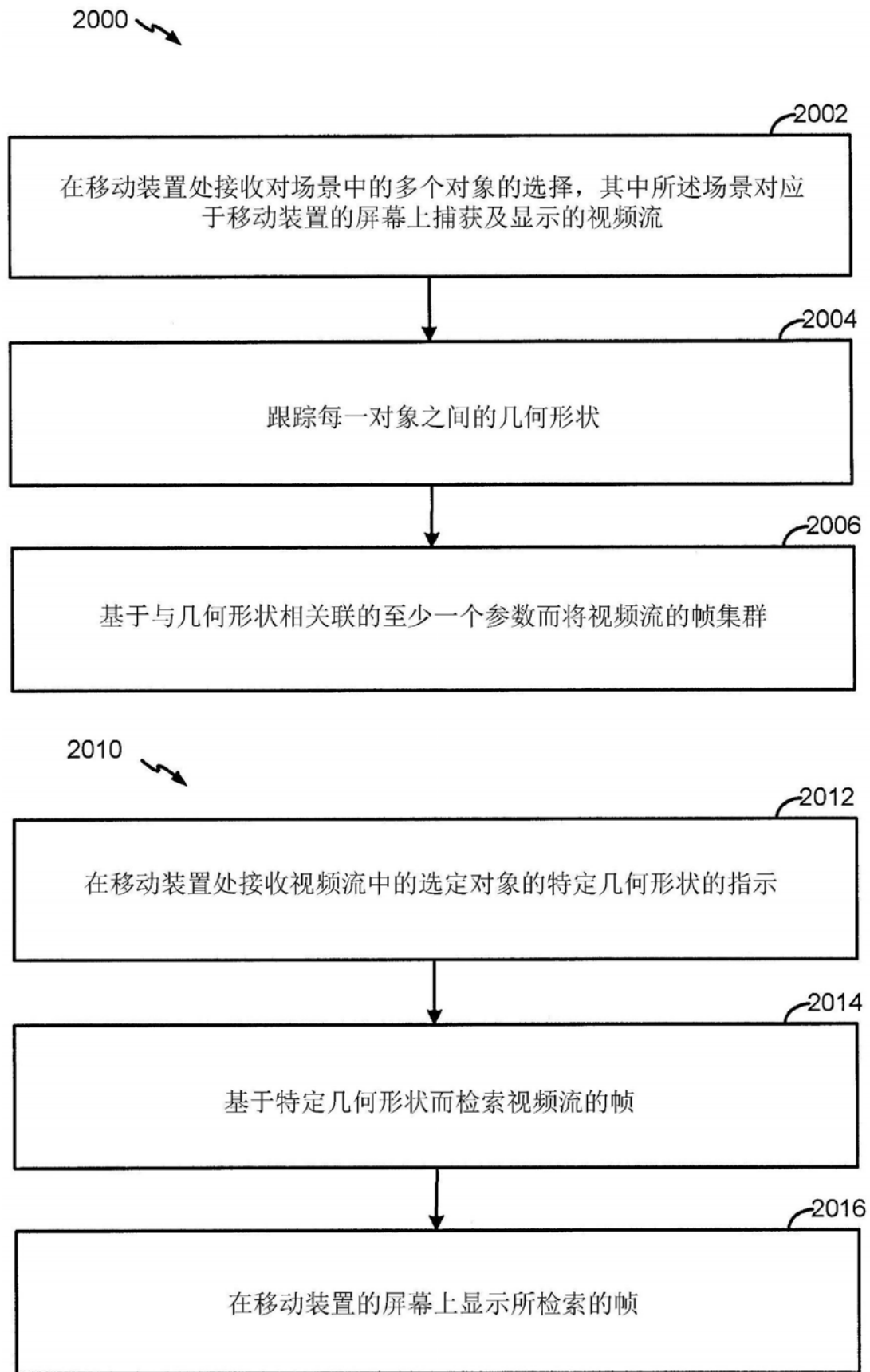


图23

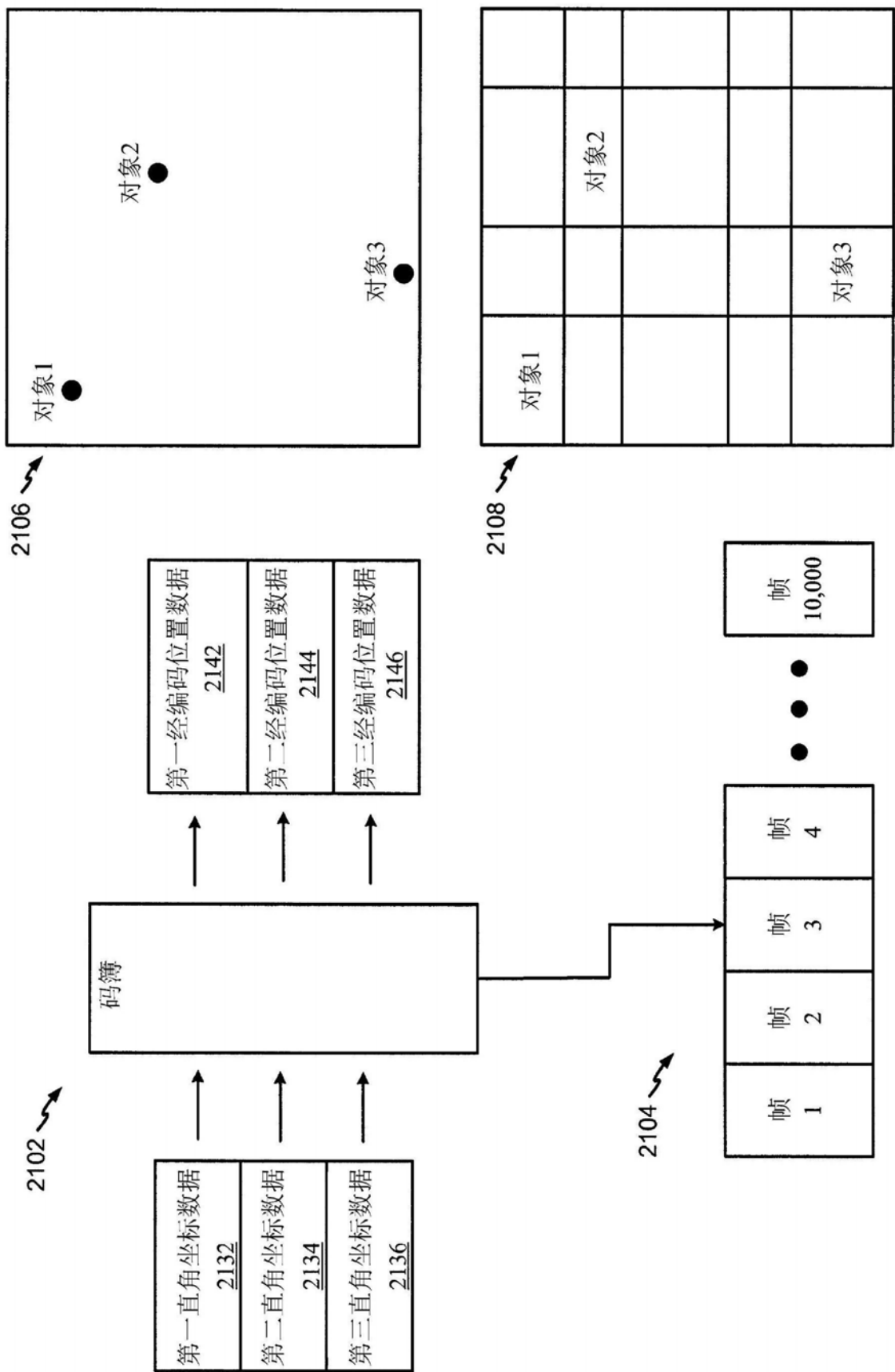


图24

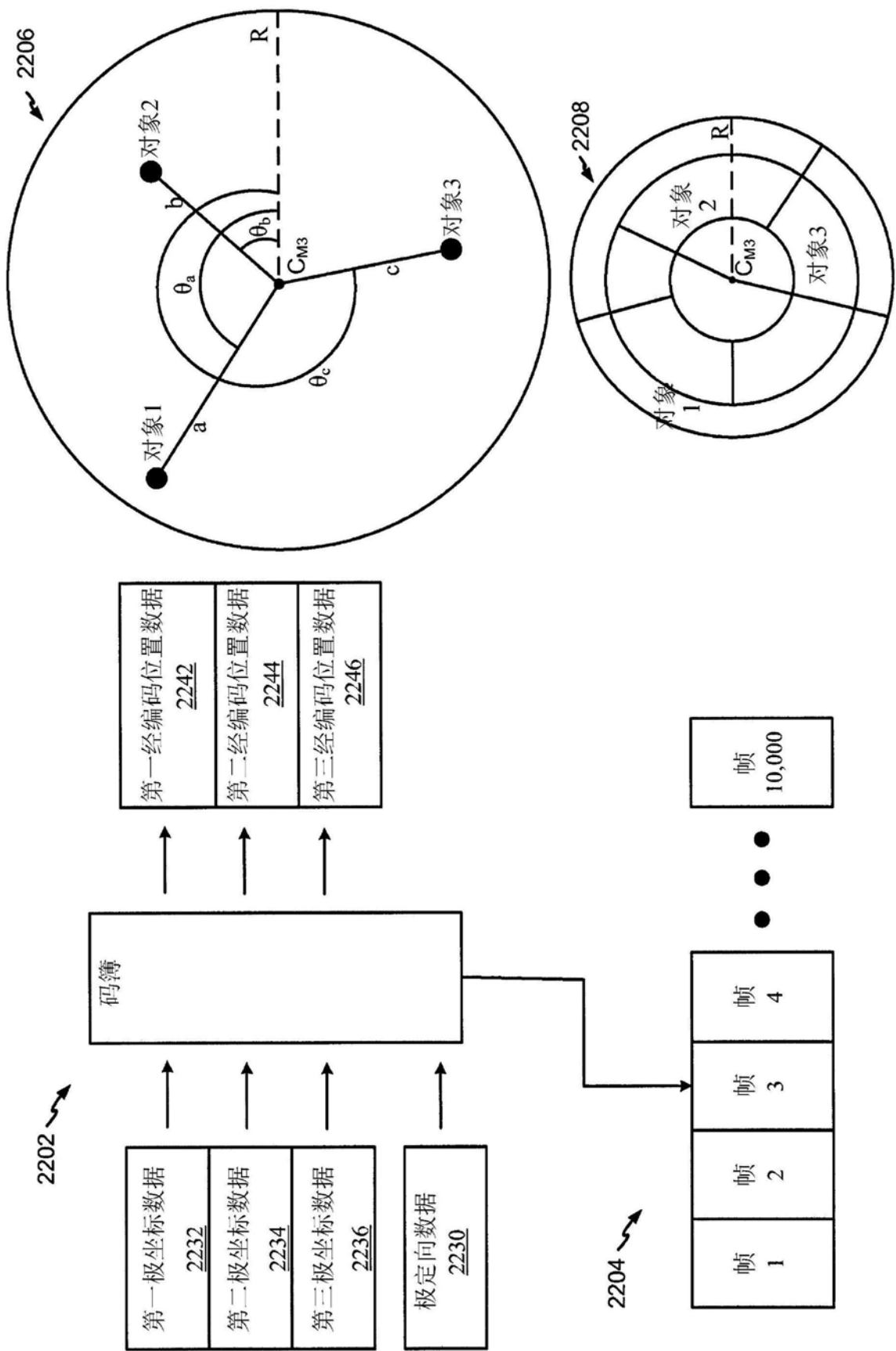


图25

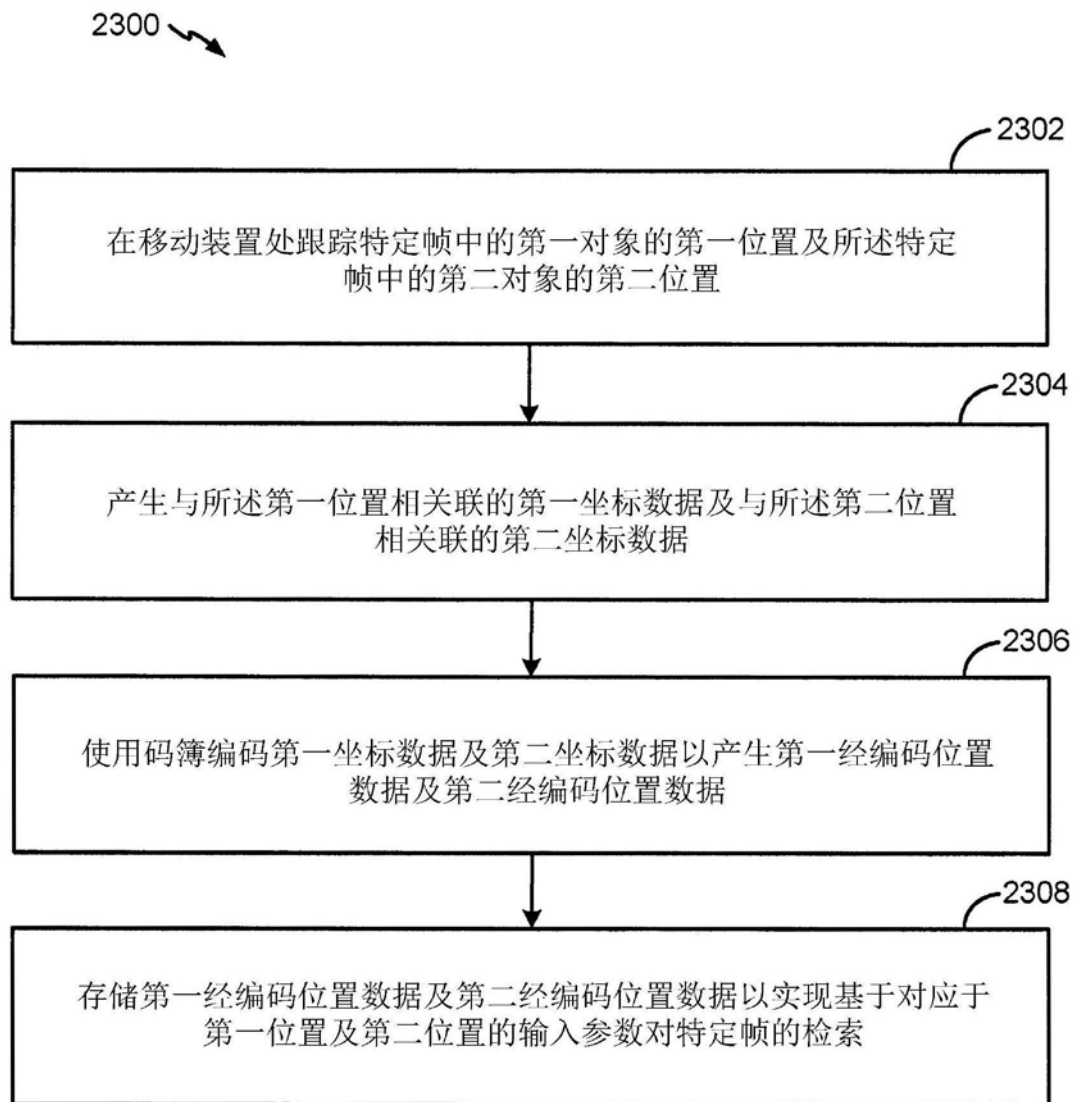


图26

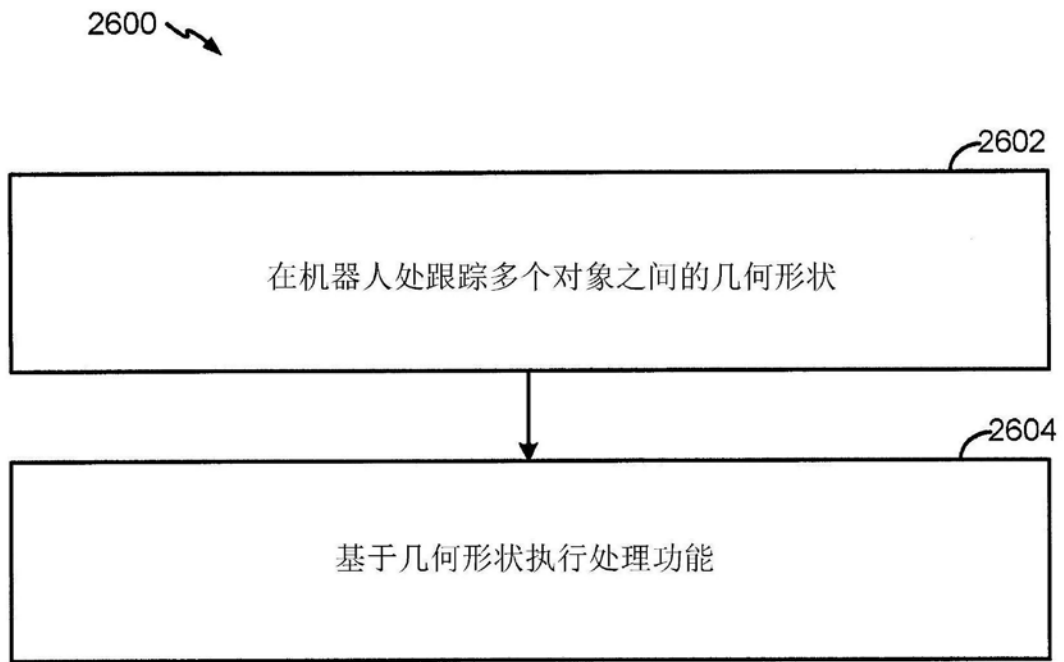


图27

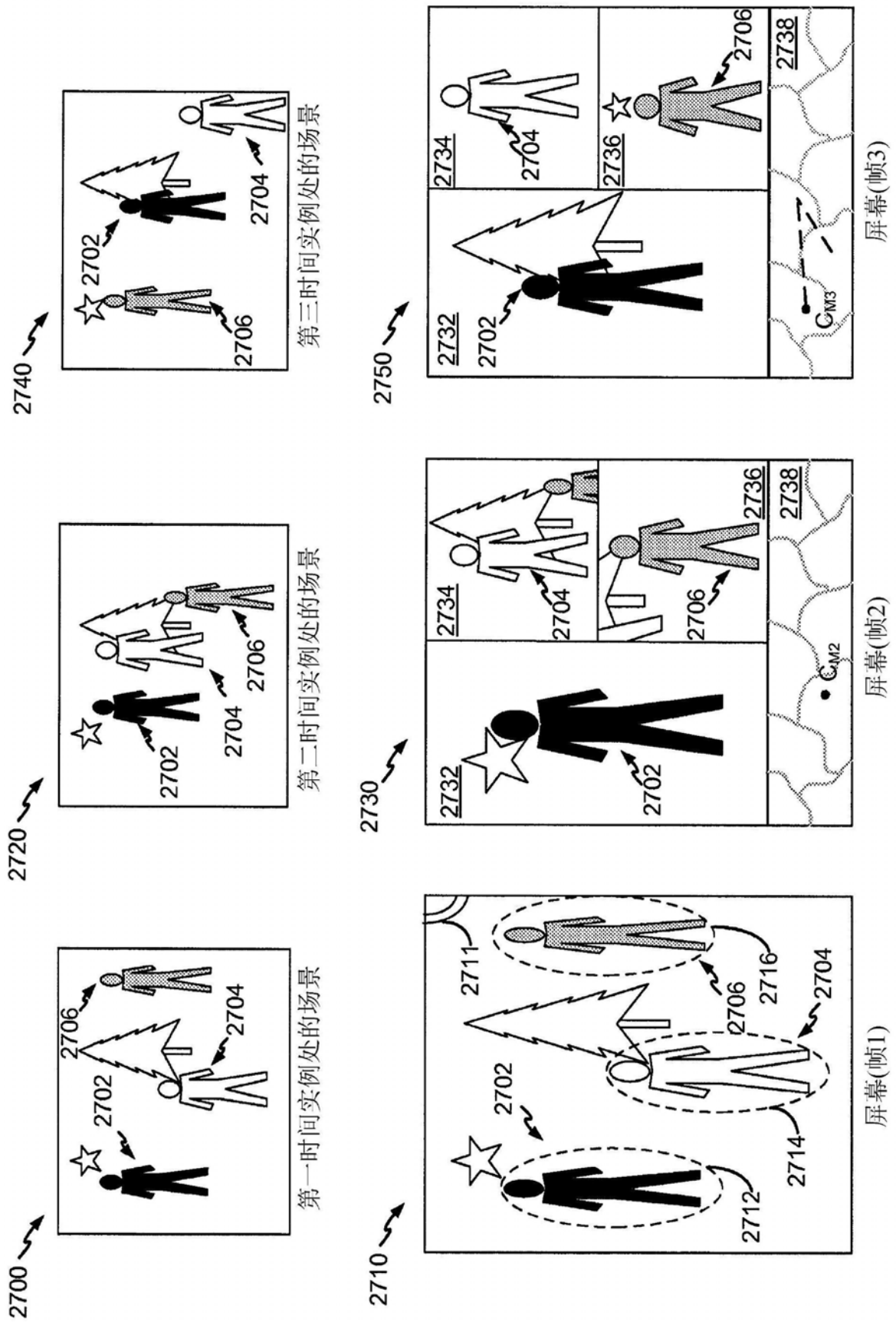


图28

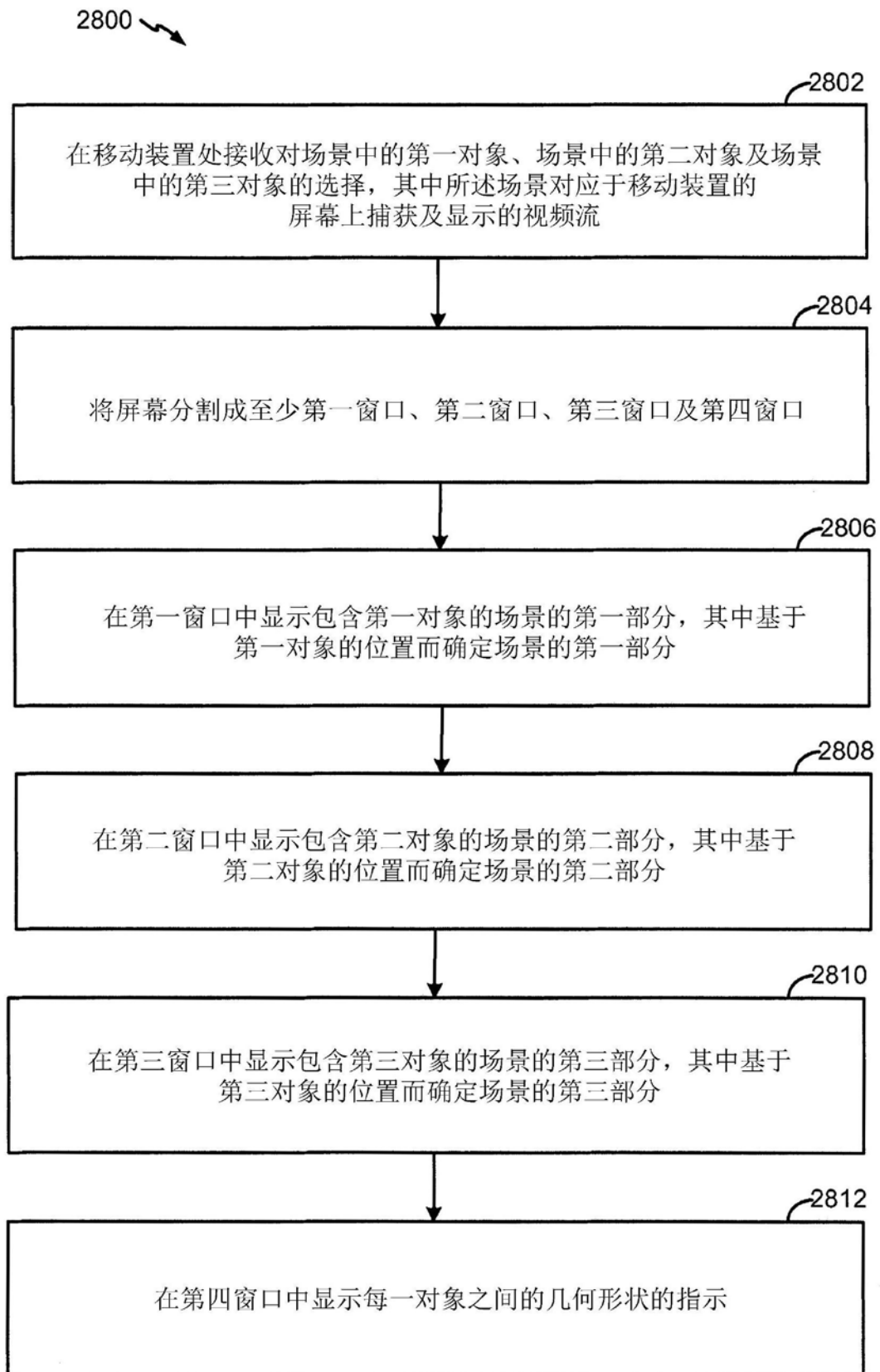


图29

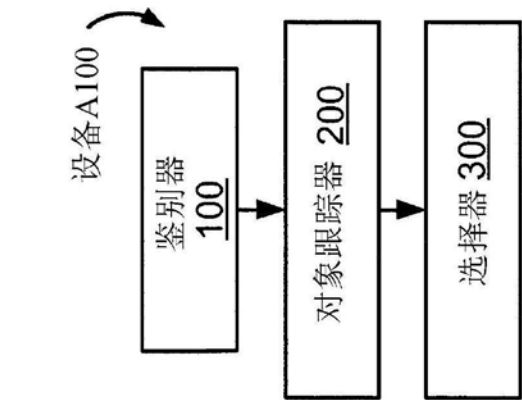


图30A

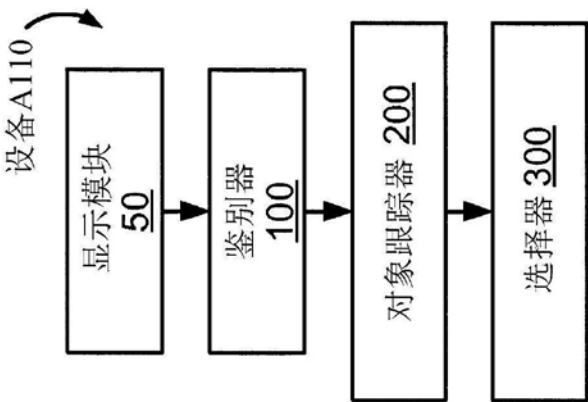


图30B

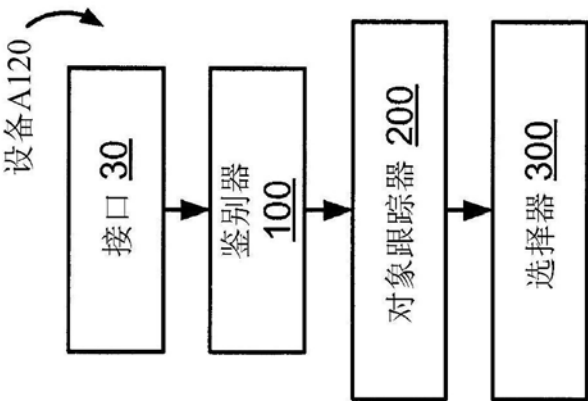


图30C

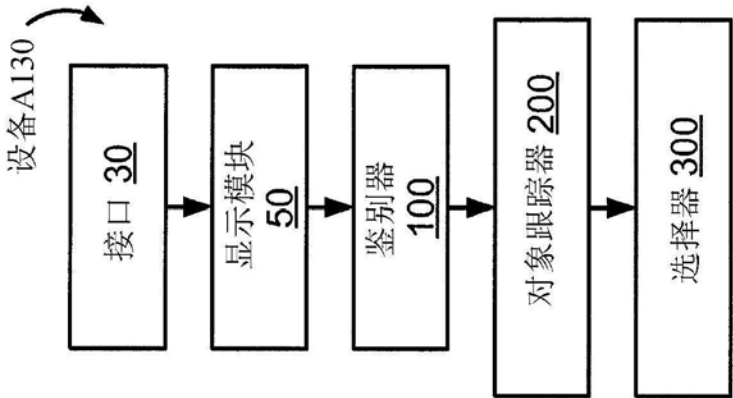


图30D

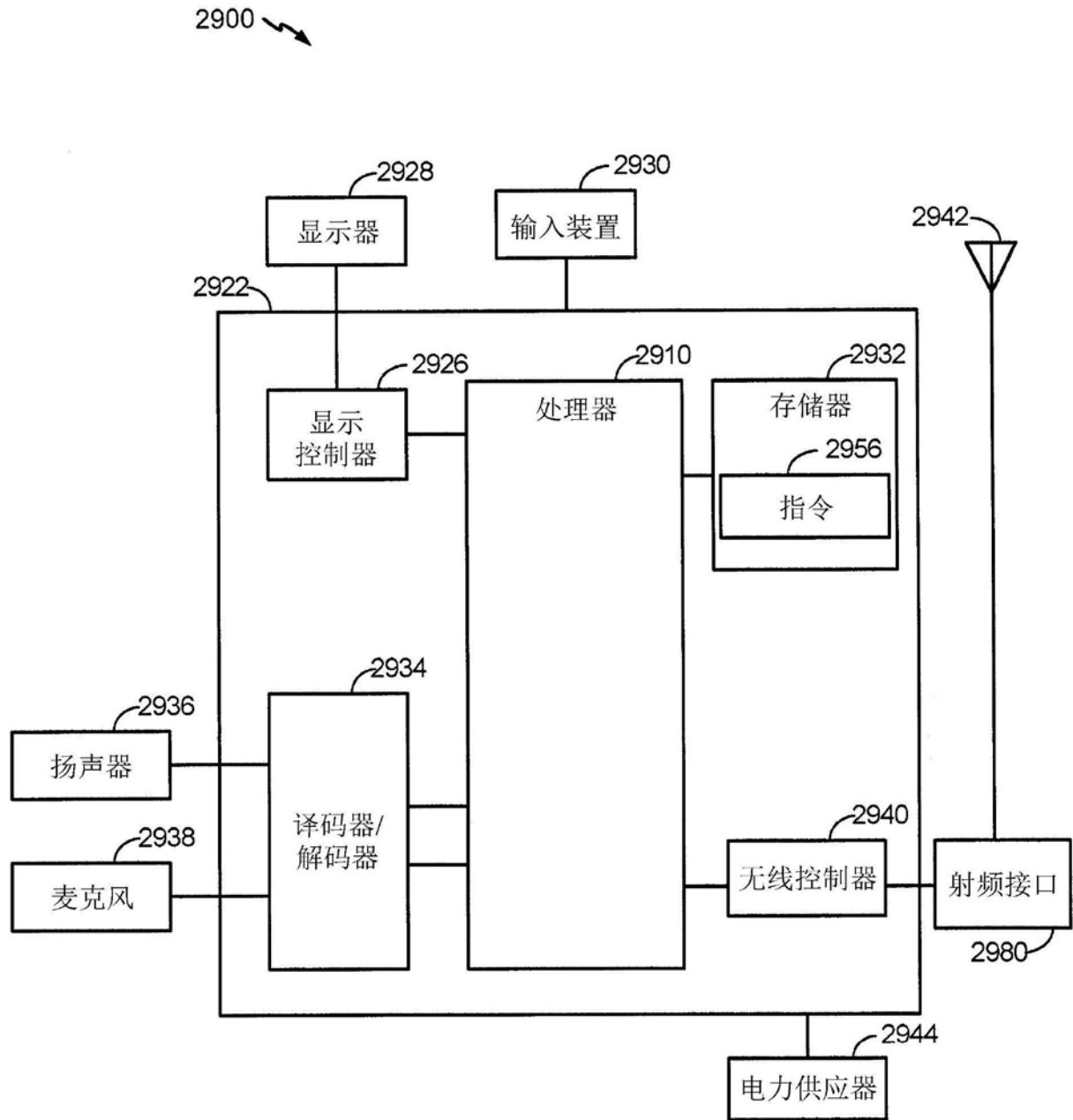


图31

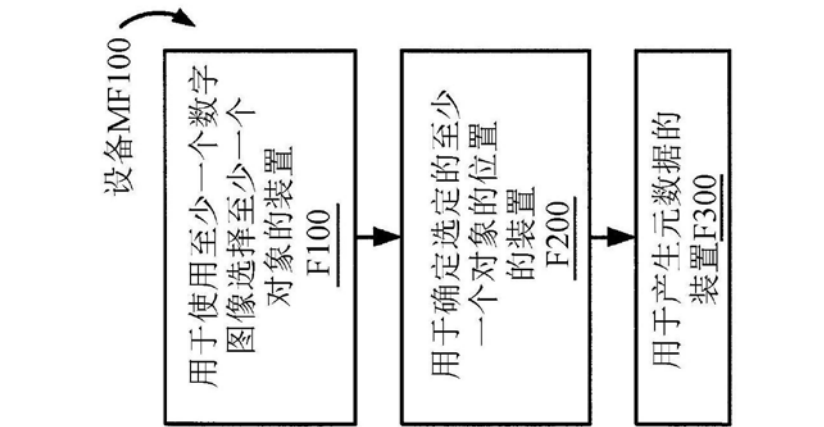


图32A

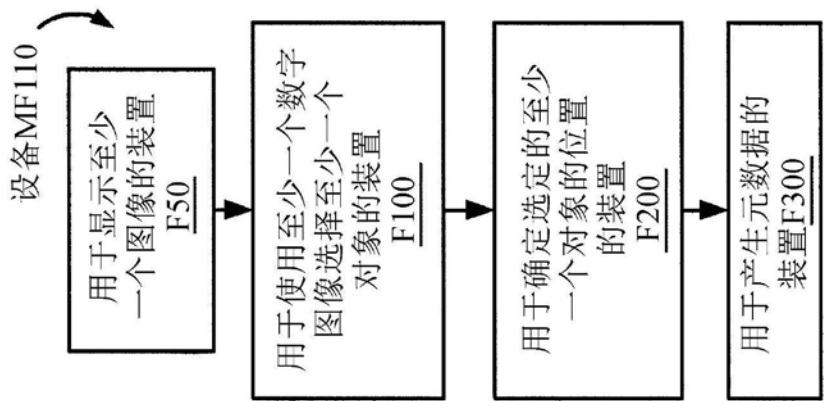


图32B

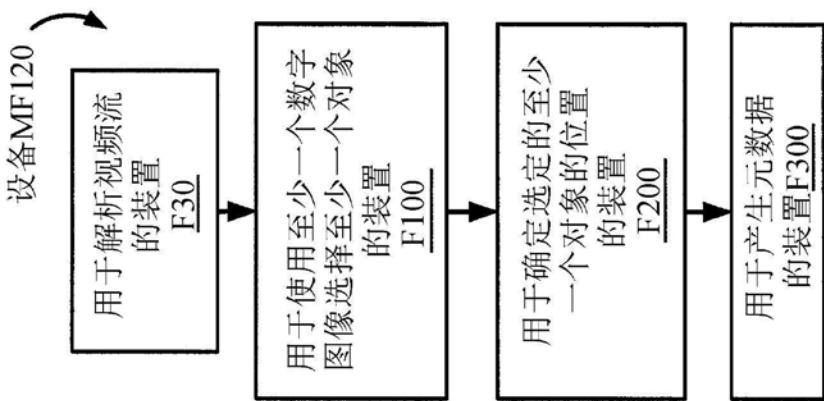


图32C

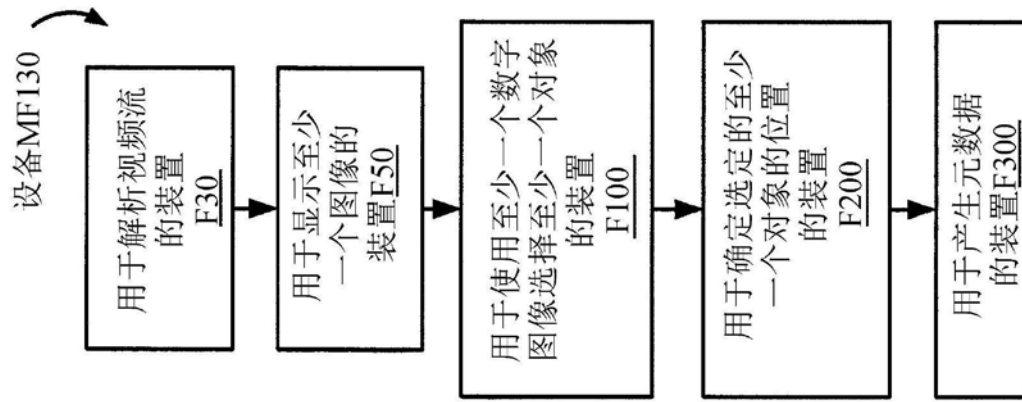


图32D

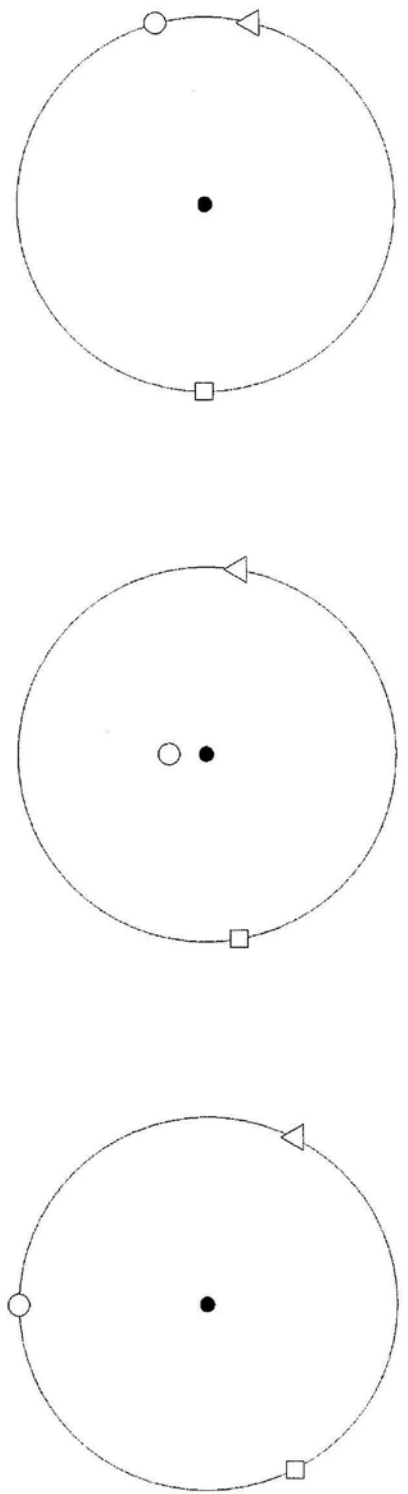


图33A

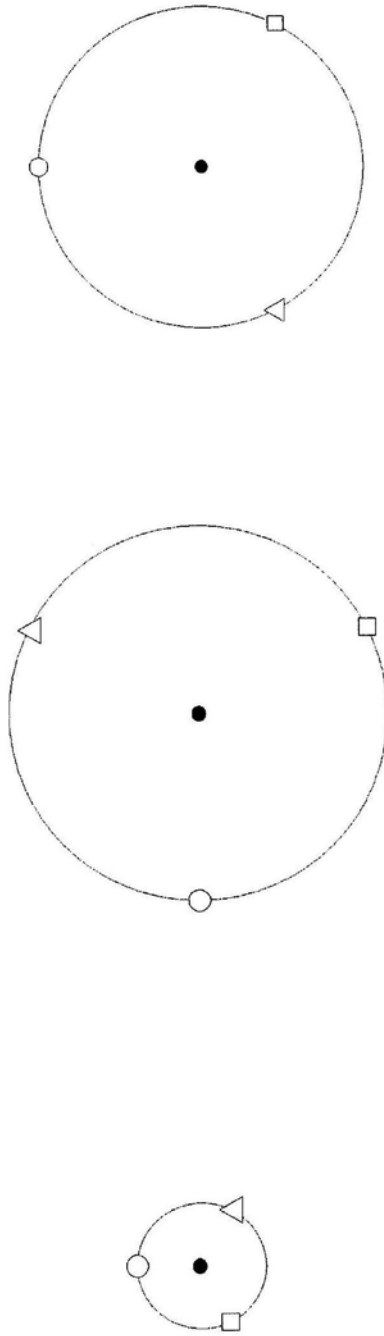


图33B

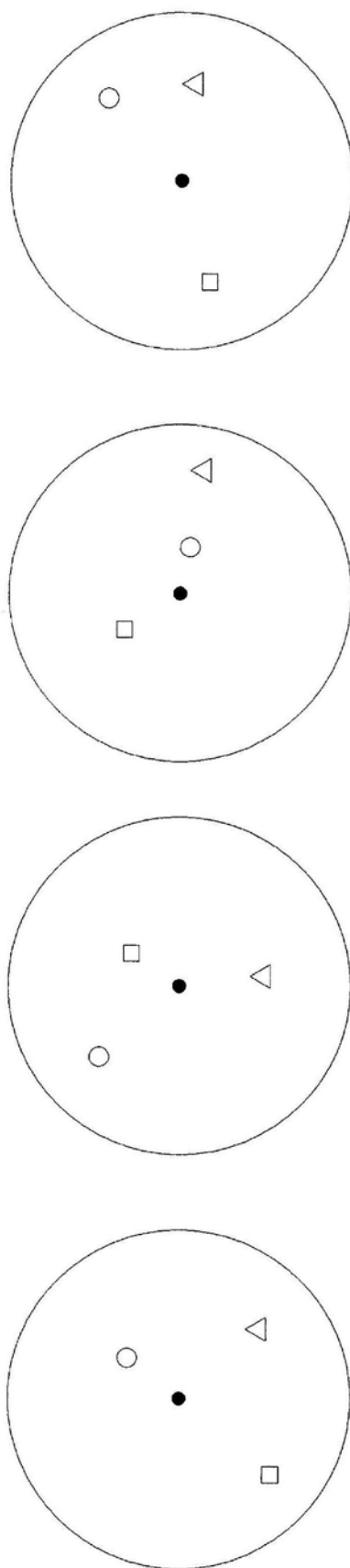


图34A

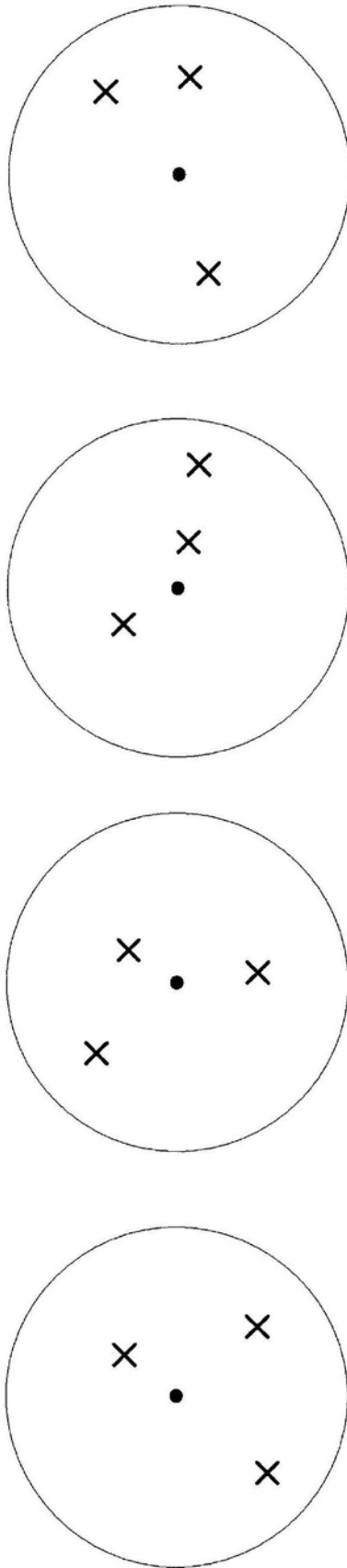


图34B

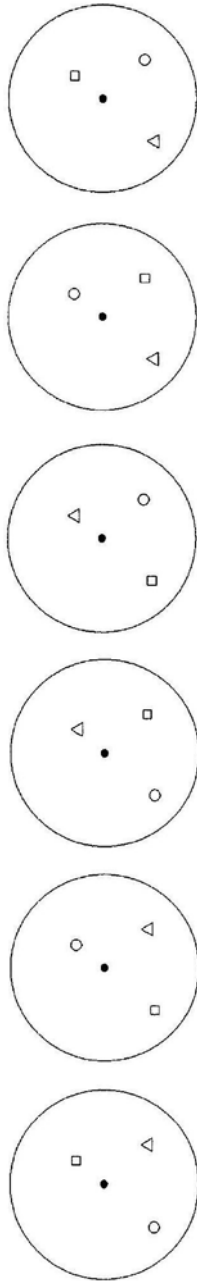


图34C

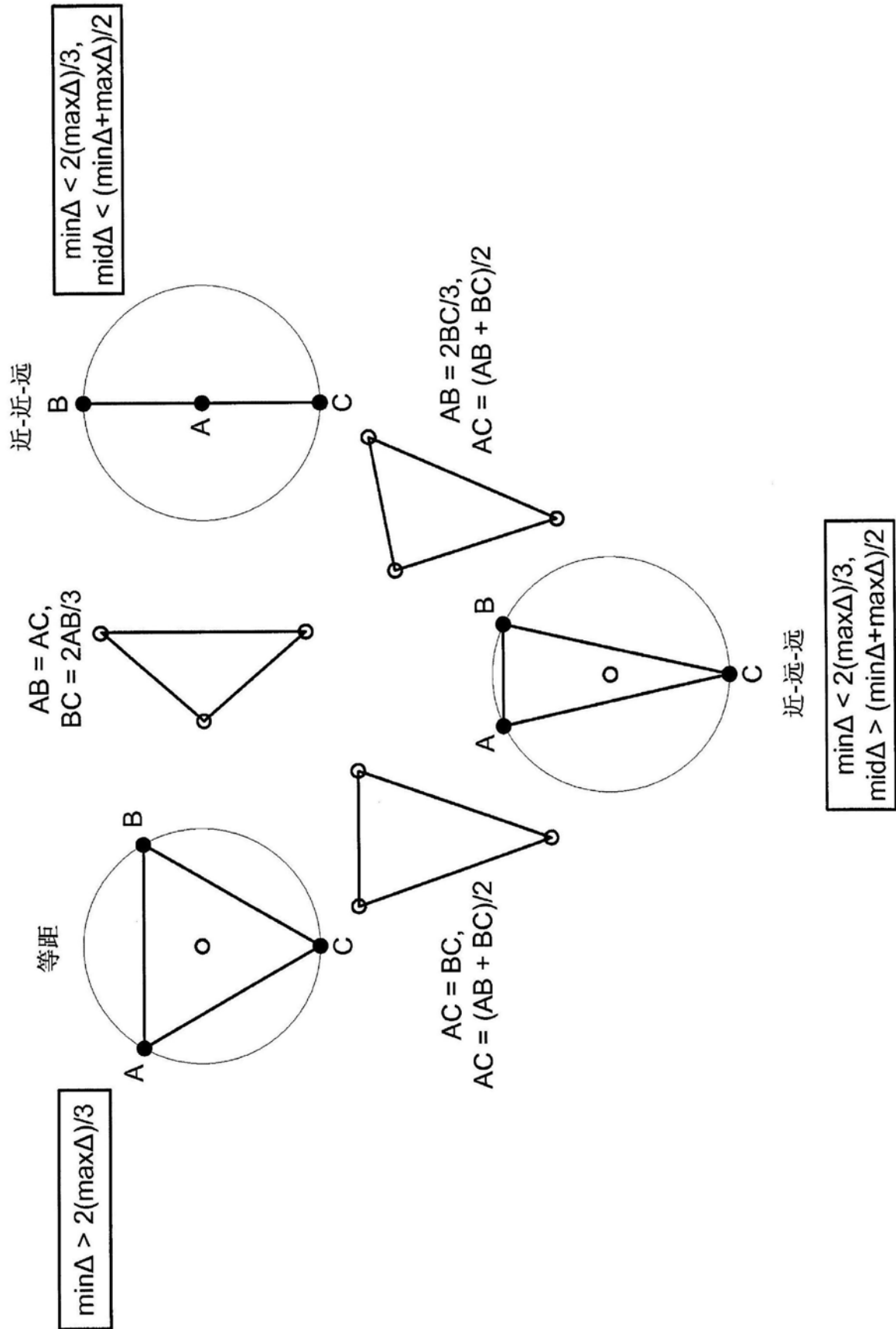


图35

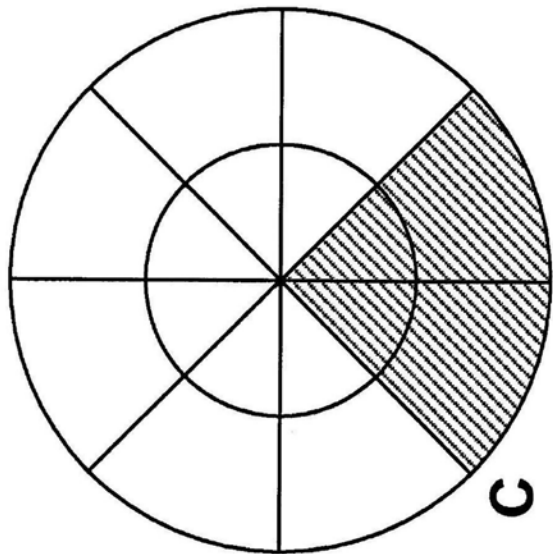
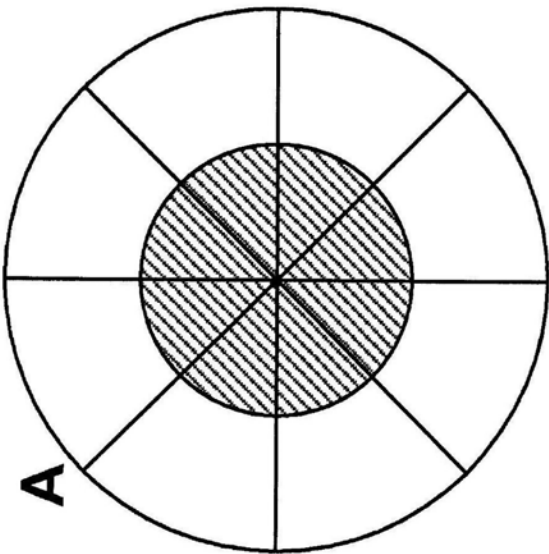
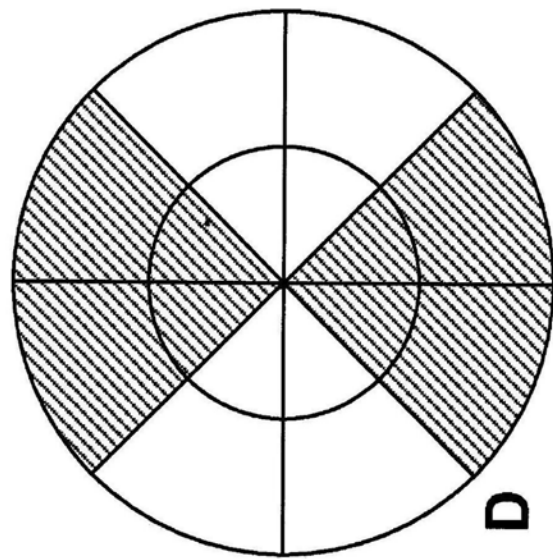
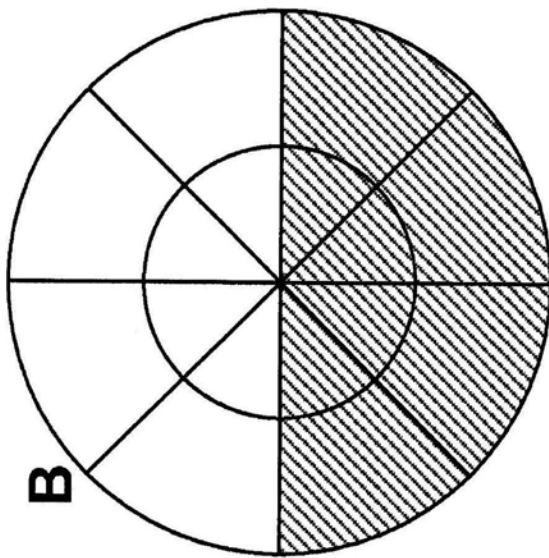
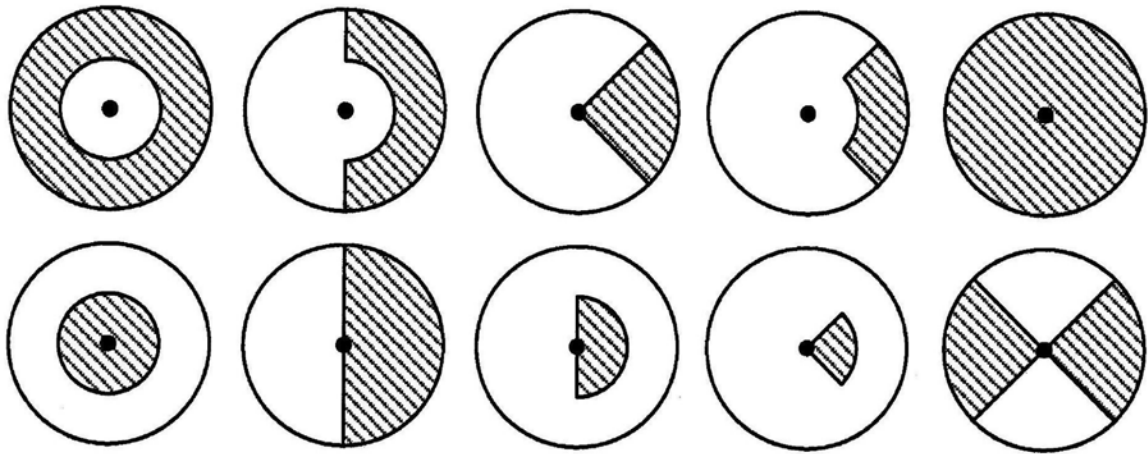


图36

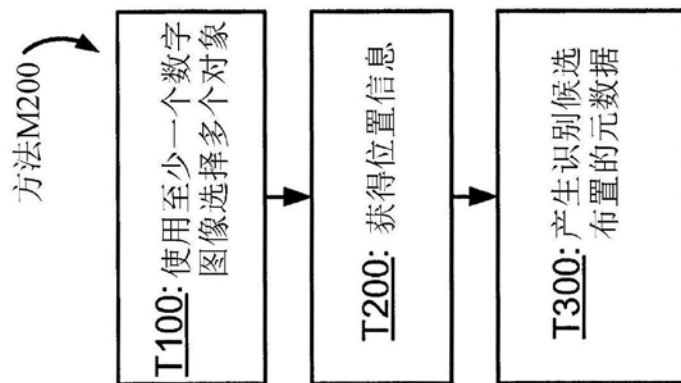


图37A

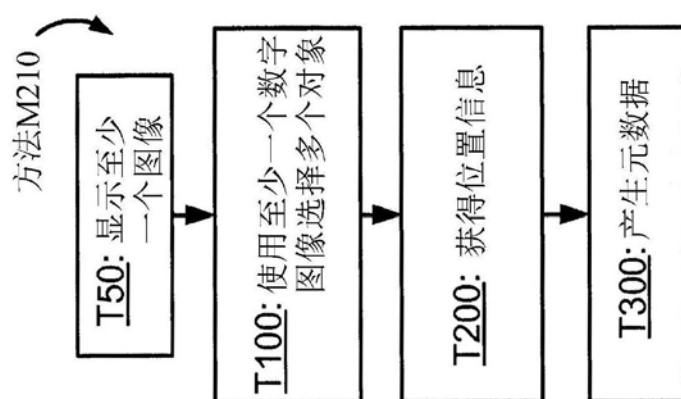


图37B

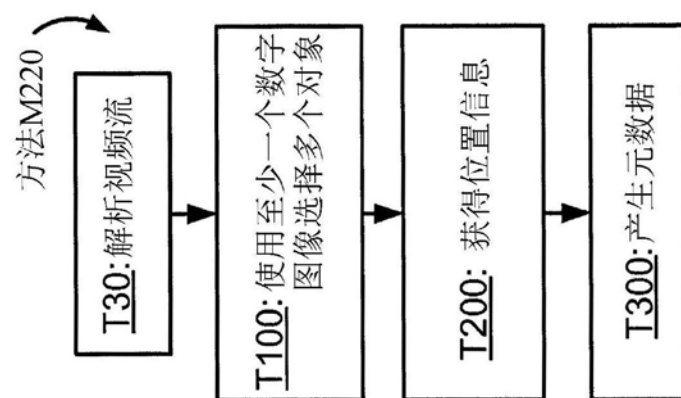


图37C

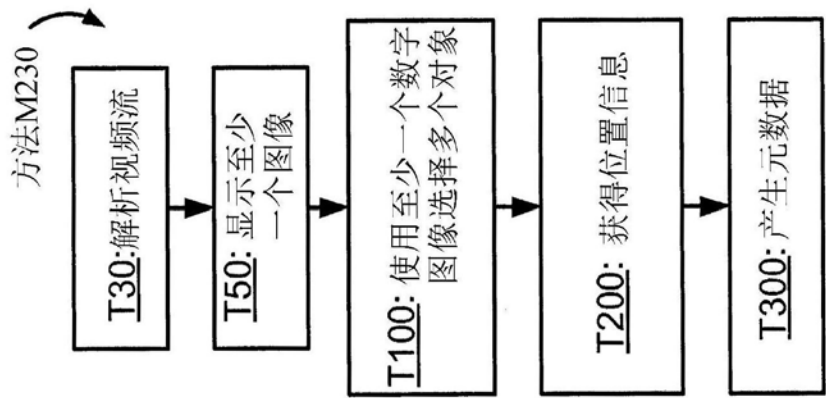


图37D

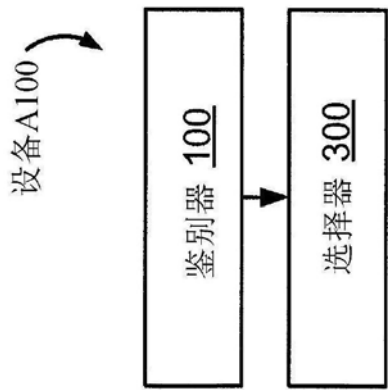


图38A

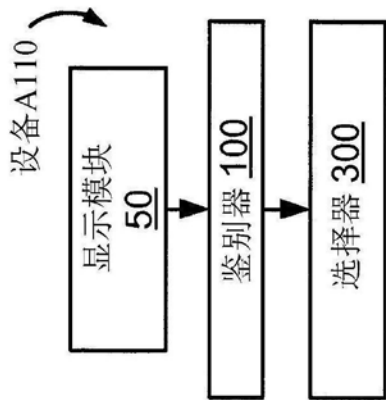


图38B

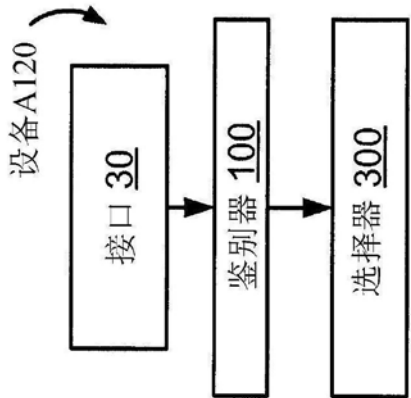


图38C

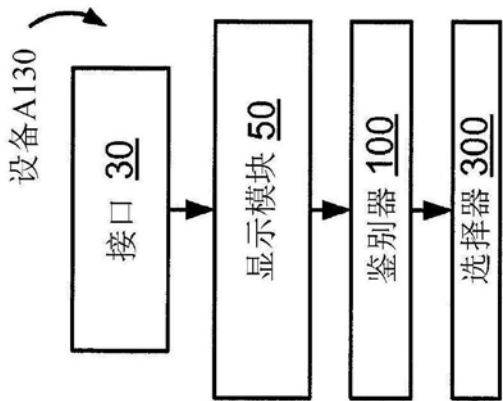


图38D

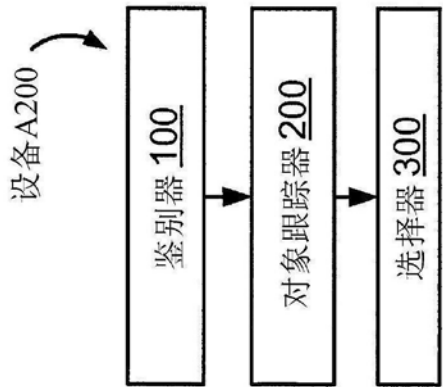


图38E

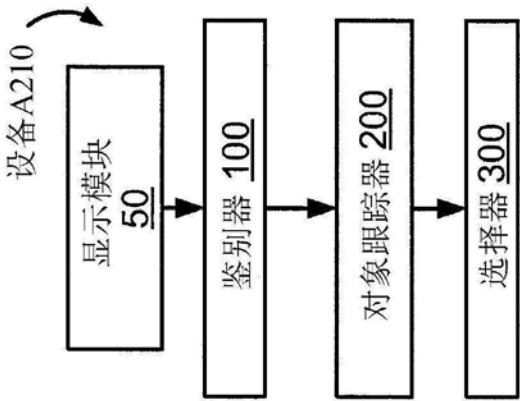


图38F

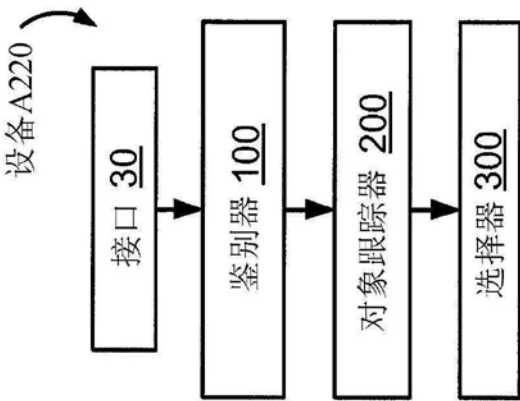


图38G

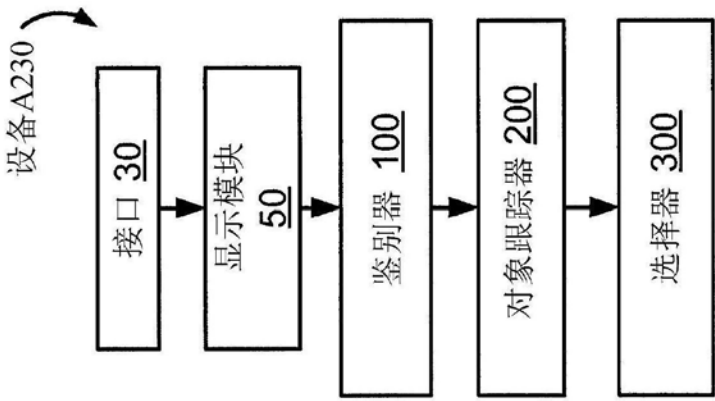


图39A

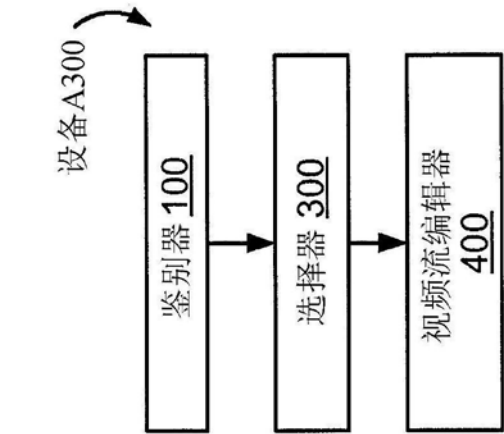


图39B

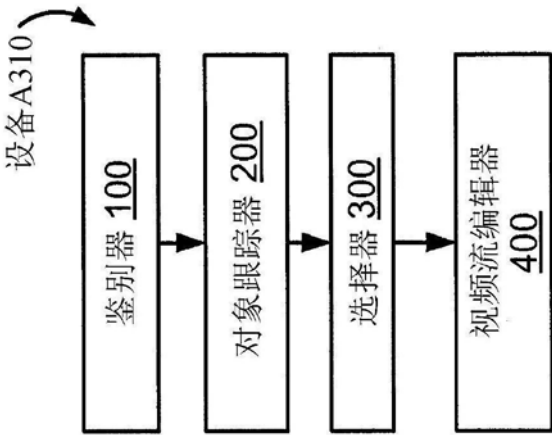


图39C

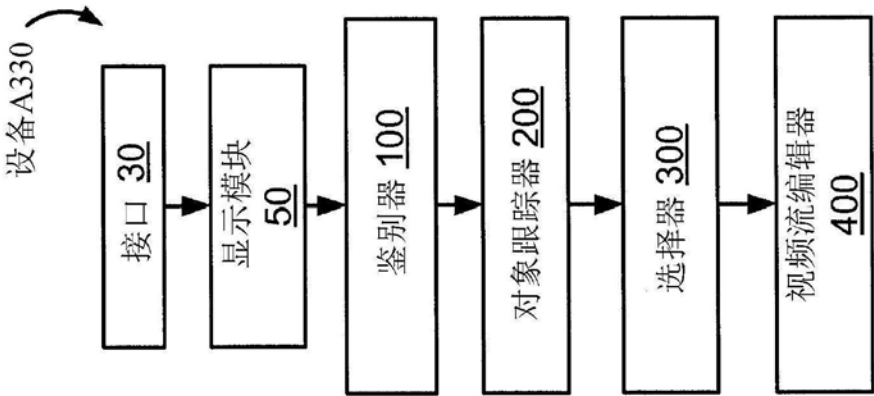


图39D

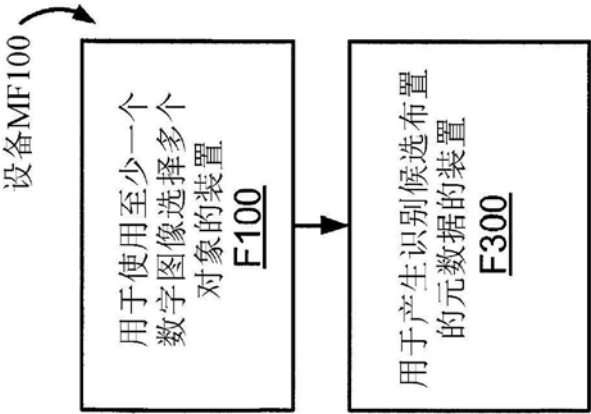


图40A

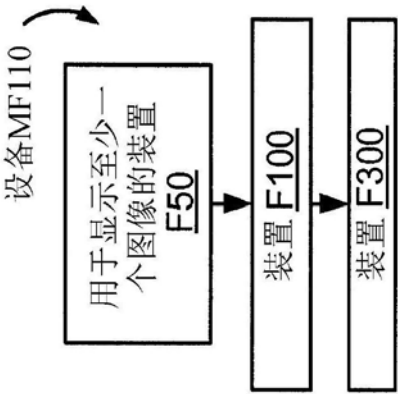


图40B

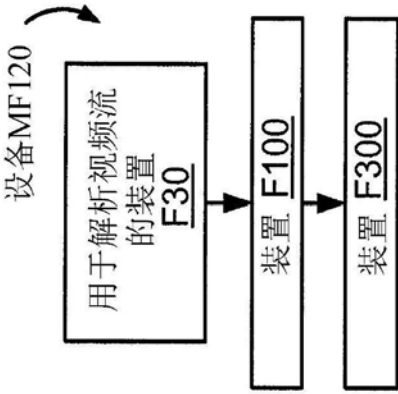


图40C

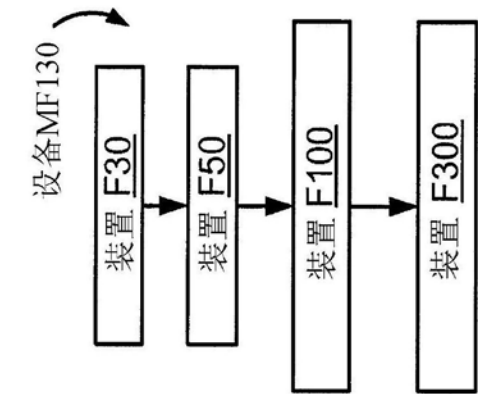


图40D

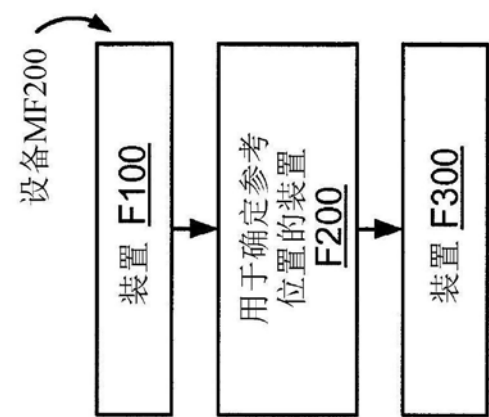


图40E

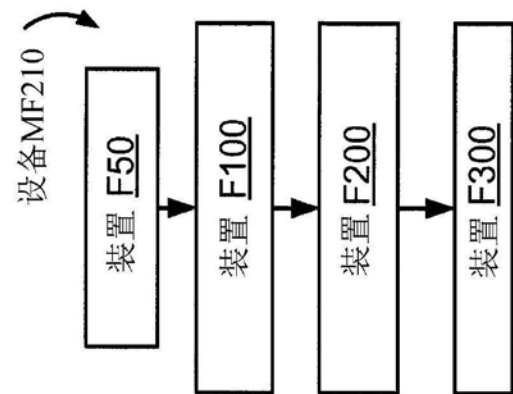


图40F

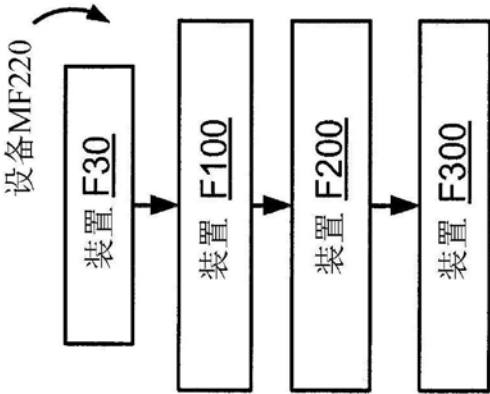


图40G

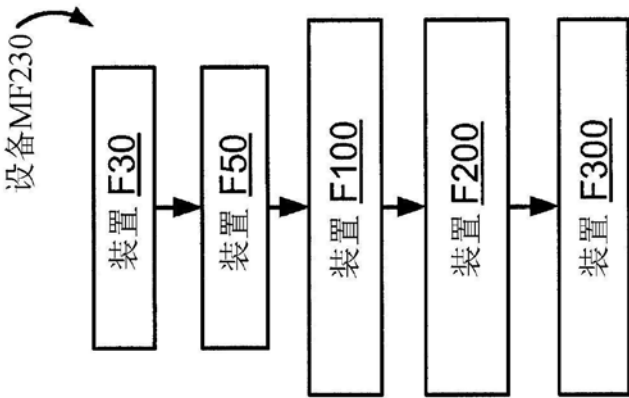


图41A

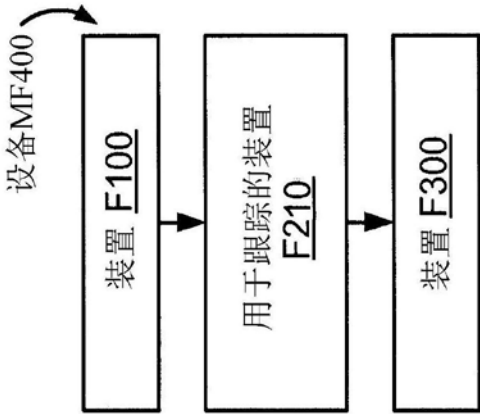


图41B

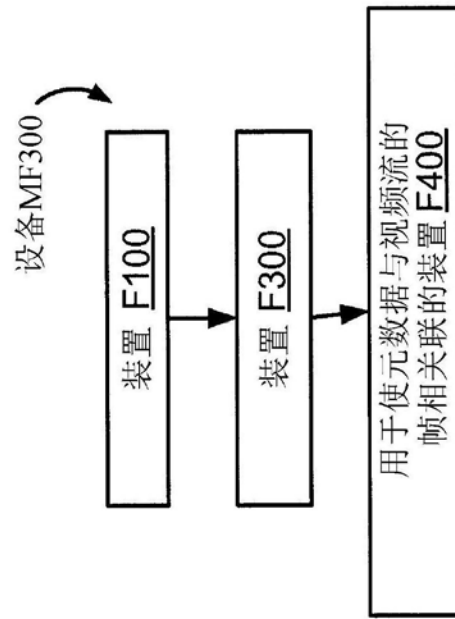


图41C

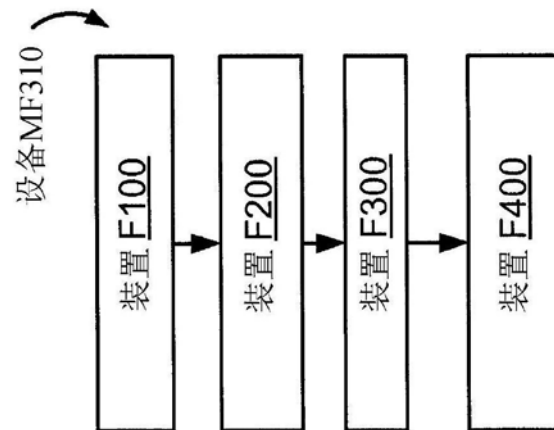


图41D

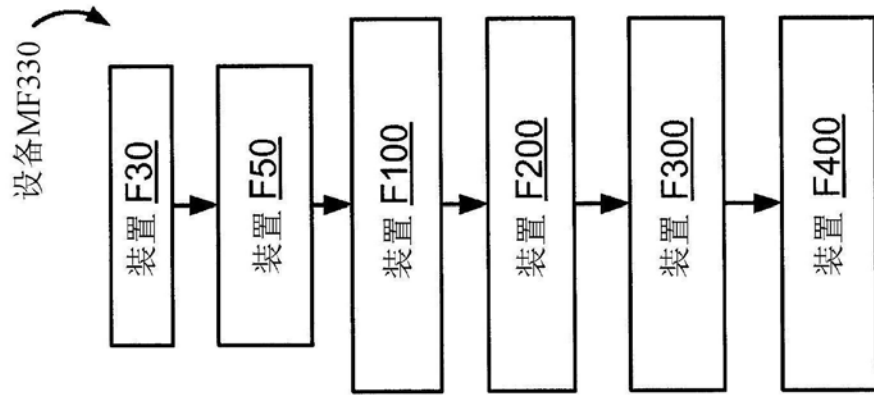


图41E

1100 ↗

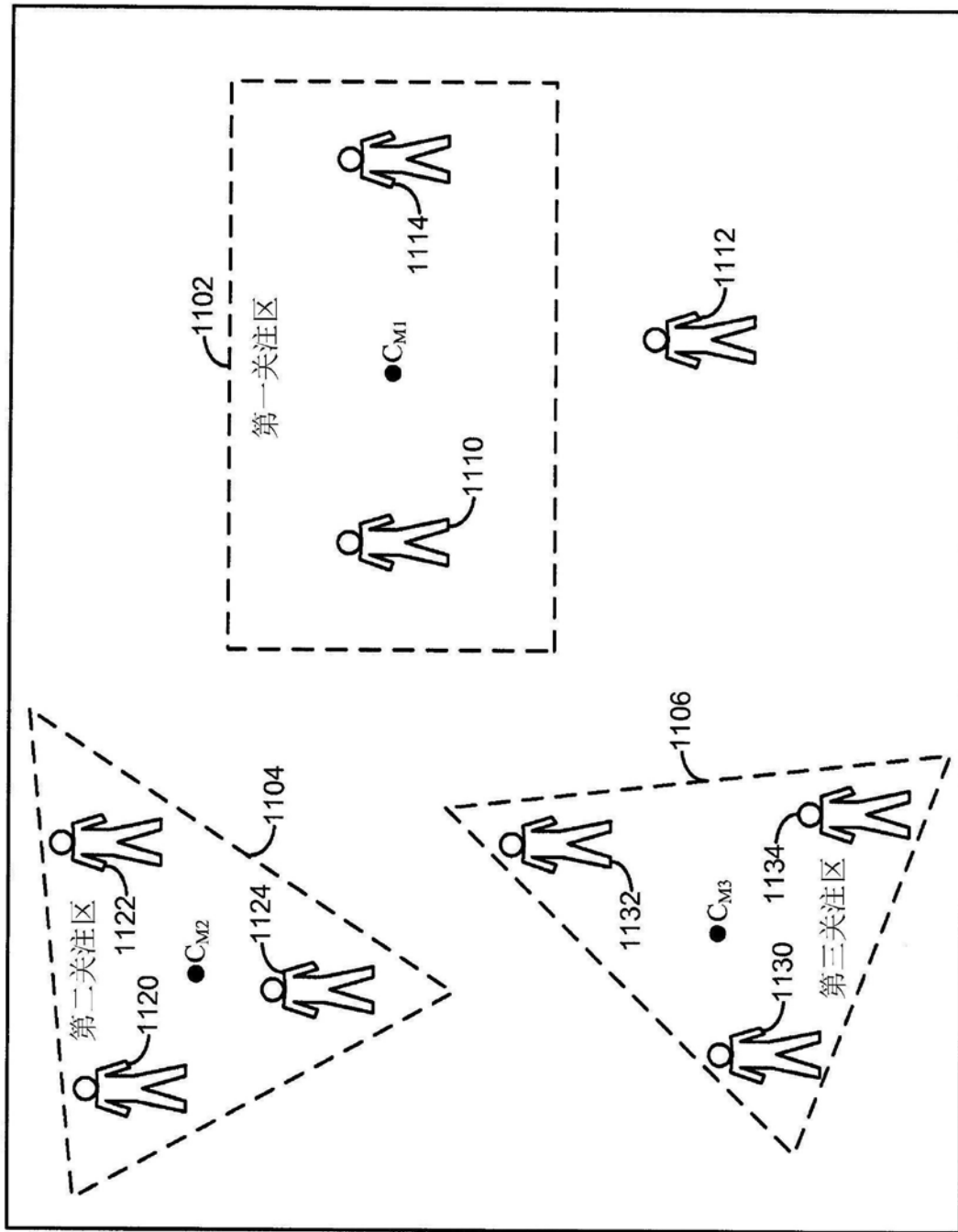


图42

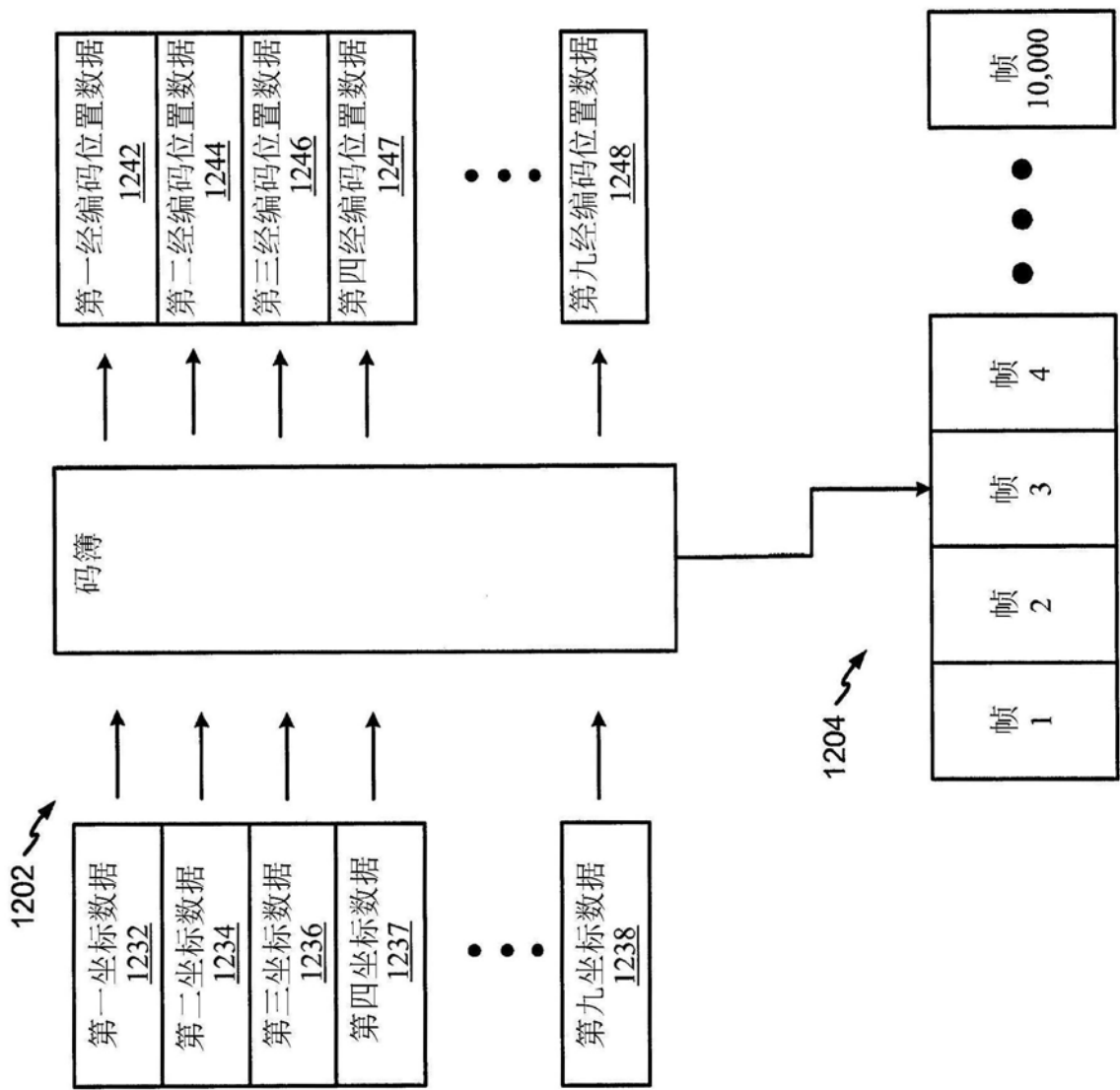


图43

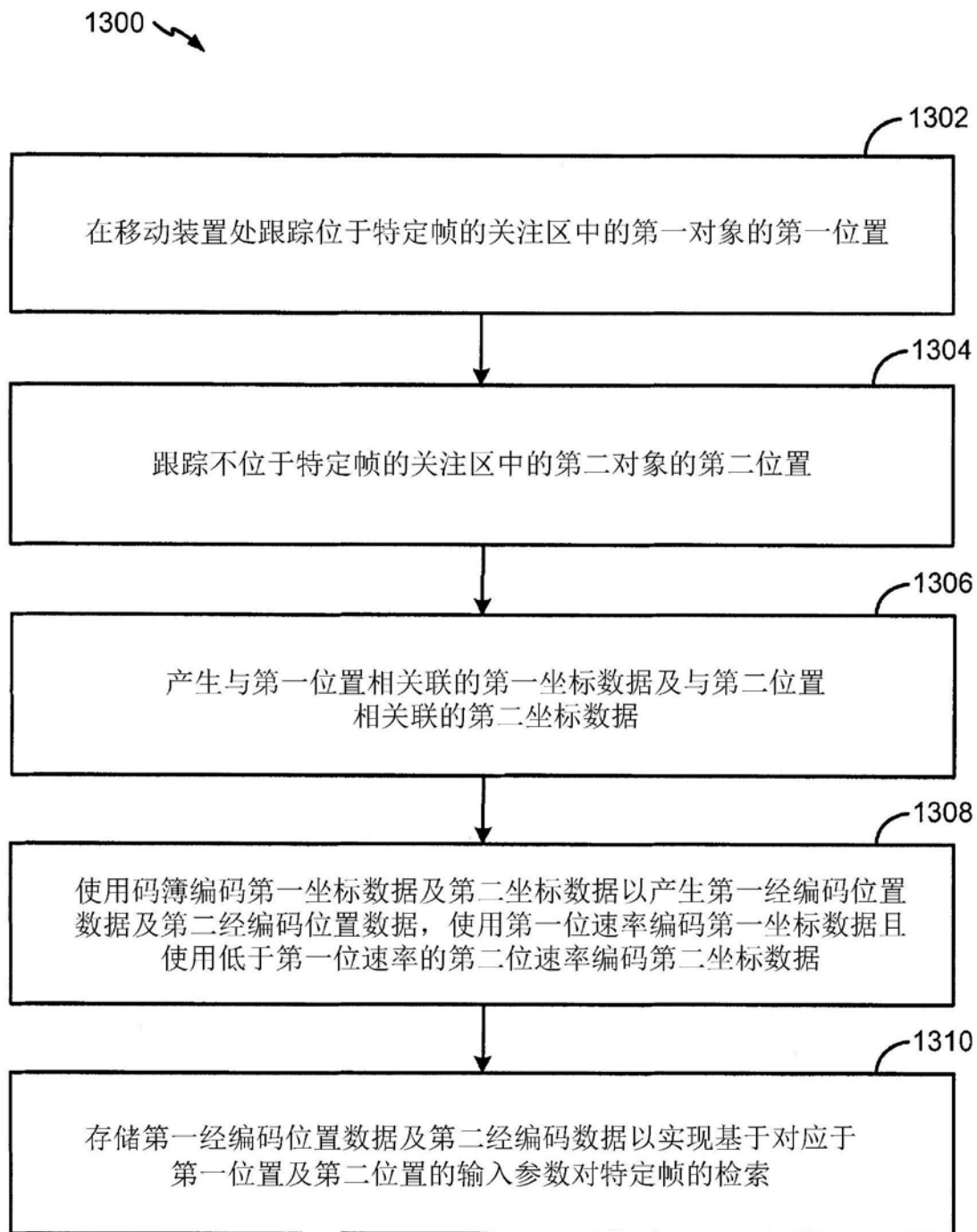


图44

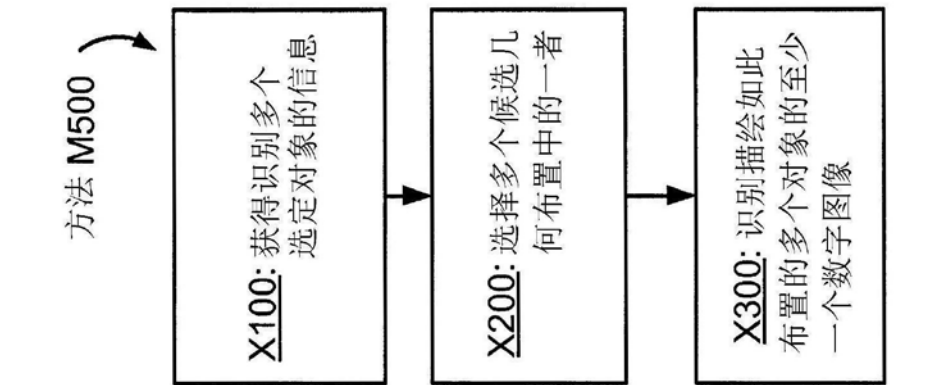


图45A

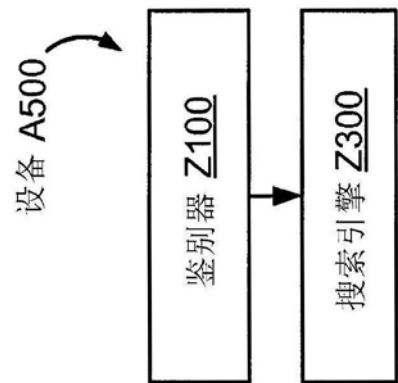


图45B

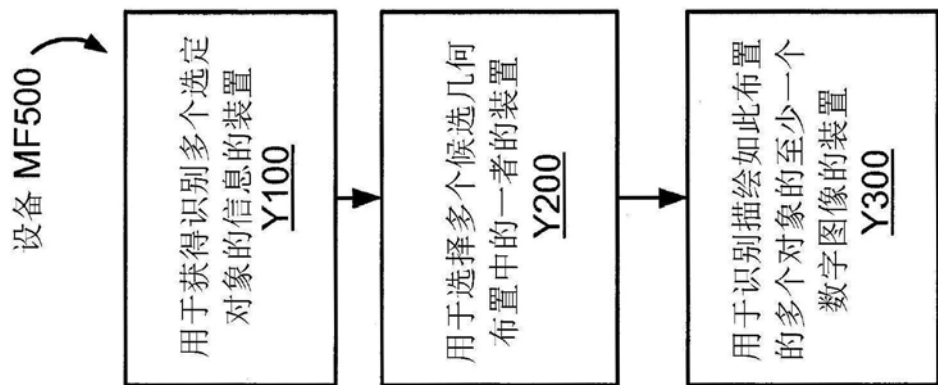


图45C