

(12) 发明专利申请

(10) 申请公布号 CN 102843523 A

(43) 申请公布日 2012. 12. 26

(21) 申请号 201210092921. 6

(22) 申请日 2012. 03. 29

(30) 优先权数据

1105218. 0 2011. 03. 29 GB

(71) 申请人 索尼公司

地址 日本东京都

(72) 发明人 罗伯特·马克·斯特凡·波特

(74) 专利代理机构 北京东方亿思知识产权代理

有限责任公司 11258

代理人 李晓冬

(51) Int. Cl.

H04N 5/262(2006. 01)

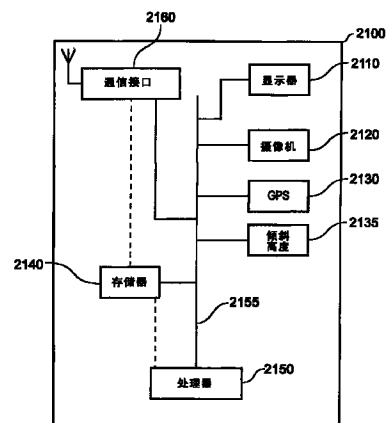
权利要求书 2 页 说明书 25 页 附图 22 页

(54) 发明名称

方法、装置和手持机

(57) 摘要

公开了一种方法、装置和手持机。一种在显示器上摄像机捕获的场景图像中的多个对象的方法,方法包括:接收 i) 表示要被应用到每个对象的不同评注的元数据,和 ii) 标识图像中的评注要被应用到的场景中每个对象的真实世界位置的位置信息;确定摄像机的焦距和应用于摄像机的倾斜;确定摄像机关于被捕获的场景的位置;和根据位置信息将评注应用到摄像机捕获的图像。



1. 一种在显示器上评注摄像机捕获的场景图像中的多个对象的方法,所述方法包括:
接收 i) 表示要被应用到每个对象的不同评注的元数据,和 ii) 标识所述图像中的评注要被应用到的所述场景中的每个对象的真实世界位置的位置信息;

确定所述摄像机的焦距和应用于所述摄像机的倾斜;

确定所述摄像机相对于被捕获的所述场景的位置;

根据所述位置信息将所述评注应用到所述摄像机捕获的所述图像;和

从与所述摄像机不同的具有所述场景的视场的图像捕获设备获得所述位置信息,其中其中根据所述图像捕获设备所捕获的所述场景的图像来确定所述位置信息。

2. 根据权利要求 1 所述的方法,还包括根据所接收的所述对象的真实世界位置信息、所述摄像机的位置信息、所述摄像机的焦距和应用于所述摄像机的倾斜,识别所述摄像机所捕获的所述图像中的至少一个对象;以及根据识别出的对象将所述评注应用于所述图像。

3. 根据权利要求 2 所述的方法,其中,识别所述对象的步骤包括检测所述图像中的对象。

4. 根据权利要求 2 或 3 所述的方法,还包括根据所述摄像机的存储的光学畸变特征来识别所述对象。

5. 根据权利要求 1 所述的方法,其中,所述位置信息是全球定位系统位置信息。

6. 根据权利要求 1 所述的方法,其中,所述一个对象是位于所述场景中的静态对象,或者是所述场景中的独特对象。

7. 根据权利要求 1 所述的方法,还包括存储所述元数据和根据所存储的元数据显示所述评注。

8. 一种装置,包括显示器和摄像机,所述显示器可操作为用于显示具有多个对象的场景图像,所述图像由所述摄像机捕获,所述装置还包括:

接收设备,可操作为接收 i) 表示要被应用到每个所述对象的不同评注的元数据,和 ii) 标识所述图像中的评注要被应用到的所述场景中的每个对象的真实世界位置的位置信息;

确定设备,被配置为:

确定所述摄像机的焦距和应用于所述摄像机的倾斜;和

确定所述摄像机相对于被捕获的场景的位置;

并且,所述装置包括控制器,该控制器可操作为根据所述位置信息将所述评注应用到所述摄像机捕获的所述图像;并且所述装置还包括获得设备,该获得设备可操作为从与所述摄像机不同的具有所述场景的视场的图像捕获设备获得所述位置信息,并且其中根据所述图像捕获设备所捕获的所述场景图像来确定所述位置信息。

9. 根据权利要求 8 所述的装置,其中,所述控制器还可操作为根据所接收的所述对象的真实世界位置信息、所述摄像机的位置信息、所述摄像机的焦距和应用于所述摄像机的倾斜,识别所述摄像机所捕获的所述图像中的至少一个对象;以及根据识别出的对象将所述评注应用于所述图像。

10. 根据权利要求 9 所述的装置,其中,所述控制器可操作为检测所述图像中的对象。

11. 根据权利要求 9 或 10 所述的装置,其中,所述控制器还可操作为根据所述摄像机的

存储的光学畸变特征来识别所述对象。

12. 根据权利要求 8 所述的装置,其中,所述位置信息是全球定位系统位置信息。

13. 根据权利要求 8 所述的装置,其中,所述一个对象是位于所述场景中的静态对象,或者是所述场景中独特的对象。

14. 根据权利要求 8 所述的装置,还包括存储设备,该存储设备可操作为存储所述元数据,并且所述控制器可操作为根据所存储的元数据显示所述评注。

15. 一种移动手持机,其包括用于连接到网络的收发器和根据权利要求 8 所述的装置。

16. 一种计算机程序,其包括当载入计算机时将所述计算机配置为执行根据权利要求 1 所述的方法的计算机可读指令。

17. 一种存储介质,其配置为在其中或其上存储有根据权利要求 16 所述的计算机程序。

方法、装置和手持机

技术领域

[0001] 本发明涉及一种方法、装置和手持机。

背景技术

[0002] 当观看事件或场景时,获得场景事件的进一步细节可能是有用的。这有时被称作增强现实。

[0003] 与增强现实相关的一个问题是在现实生活的图像上覆盖评注的速度和精确性。本发明实施例的目的是解决这些问题。

发明内容

[0004] 根据第一方面,提供了一种在显示器上评注摄像机捕获的场景图像中的多个对象的方法,所述方法包括:接收 i) 表示要被应用到每个对象的不同评注的元数据,和 ii) 标识图像中的评注要被应用到的场景中每个对象的真实世界位置的位置信息;确定摄像机的焦距和应用于摄像机的倾斜;确定摄像机相关于被捕获的场景的位置;和根据位置信息将评注应用到摄像机捕获的图像。

[0005] 方法还可包括从不同于摄像机的具有场景的视场的图像捕获设备获得位置信息,并且其中根据图像捕获设备所捕获的场景图像来确定位置信息。

[0006] 方法还可包括根据所接收的对象的真实世界位置信息、摄像机的位置信息、摄像机的焦距和应用于摄像机的倾斜来识别摄像机所捕获的图像中的至少一个对象;和根据所识别对象将评注应用于图像。

[0007] 识别对象的步骤包括检测图像中的对象。

[0008] 方法还可包括根据摄像机的存储的光学畸变特征来识别对象。

[0009] 位置信息可以是全球定位系统的位置信息。

[0010] 一个对象可以是位于场景中的静态对象,或者是场景中独特的对象。

[0011] 方法还可包括存储元数据和根据所存储元数据显示评注。

[0012] 根据另一个方面,提供了一种包括显示器和摄像机的装置,显示器操作为用于显示具有多个对象的场景图像,图像由摄像机捕获,所述装置还包括:

[0013] 接收设备,可操作为接收 i) 表示要被应用到每个对象的不同评注的元数据,和 ii) 标识图像中的评注要被应用到的场景中每个对象的真实世界位置的位置信息;

[0014] 确定设备,可操作为:确定摄像机的焦距和应用于摄像机的倾斜;和确定摄像机相关于被捕获的场景的位置;

[0015] 并且装置包括控制器,该控制器可操作为根据位置信息将评注应用到摄像机捕获的图像。

[0016] 装置还可包括获得设备,该获得设备可操作为从不同于摄像机的具有场景的视场的图像捕获设备获得位置信息,并且其中根据图像捕获设备所捕获的场景图像来确定位置信息。

[0017] 控制器还可操作为根据所接收的对象的真实世界位置信息、摄像机的位置信息、摄像机的焦距和应用于摄像机的倾斜来识别摄像机所捕获的图像中的至少一个对象；和根据所识别对象将评注应用于图像。

[0018] 控制器可操作为检测图像中的对象。

[0019] 控制器还可操作为根据摄像机的存储的光学畸变特征来识别对象。

[0020] 位置信息可以是全球定位系统的位置信息。

[0021] 一个对象可以是位于场景中的静态对象，或者是场景中独特的对象。

[0022] 装置还可包括存储设备，该存储设备可操作为存储元数据并且控制器可操作为根据所存储元数据显示评注。

[0023] 根据另一个方面，提供了一种移动手持机，其包括用于连接到网络的收发器和根据上面任何一个实施例所述的装置。

附图说明

[0024] 根据要结合附图来阅读的说明性实施例的以下详细描述，本发明的上面和其它目的、特征和优点将是明显的，在附图中：

[0025] 图 1 示出了根据本发明第一实施例的系统；

[0026] 图 2 示出了第一实施例的系统中的客户端设备；

[0027] 图 3 示出了根据本发明第二实施例的系统；

[0028] 图 4A 示出了本发明第一实施例的服务器；

[0029] 图 4B 示出了本发明第二实施例的服务器；

[0030] 图 5 示出了说明根据第一实施例或第二实施例的客户端设备向服务器的注册处理的流程图；

[0031] 图 6 示出了根据可应用于第一和第二实施例的本发明示例的对象跟踪方法的流程图；

[0032] 图 7A 示出了根据本发明第一和第二实施例的对象关键字的创建；

[0033] 图 7B 示出了根据本发明第一和第二实施例的对斜度的 3D 模型的方向表示的添加；

[0034] 图 8 示出了根据本发明第一和第二实施例的多个选手和他们相关联的边界框；

[0035] 图 9 示出了根据本发明第一和第二实施例的对象跟踪和遮挡检测方法的流程图；

[0036] 图 10A 和 10B 示出了根据本发明第一和第二实施例的对象跟踪和遮挡检测的某些示例；

[0037] 图 11 示出了根据本发明第一实施例的位于服务器内的重定格式设备；

[0038] 图 12 示出了根据本发明第二实施例的位于服务器内的重定格式设备；

[0039] 图 13 是根据本发明第一和第二实施例的用于确定摄像机的位置和摄像机视场内的对象之间的距离的系统的示意图；

[0040] 图 14 是根据本发明第一和第二实施例的用于确定摄像机和摄像机视场内的对象之间的距离的系统的示意图；

[0041] 图 15A 示出了根据本发明第一实施例的客户端设备；

[0042] 图 15B 示出了根据本发明第二实施例的客户端设备；

- [0043] 图 16A 示出了位于图 15A 的客户端设备中的客户端处理设备；
- [0044] 图 16B 示出了位于图 15B 的客户端设备中的客户端处理设备；
- [0045] 图 17 示出了根据本发明另一个实施例的联网系统；
- [0046] 图 18 示出了用于生成突出包的位于图 17 的联网系统中的根据第一或第二实施例的客户端设备；
- [0047] 图 19A 和 19B 示出了用于生成突出包的位于图 17 的联网系统中的根据第一或第二实施例的客户端设备；
- [0048] 图 20 示出了根据本发明另一个实施例的其中可在便携式设备上实现增强现实的体育场的平面图；
- [0049] 图 21 示出了根据图 20 的便携式设备的框图；
- [0050] 图 22 示出了当启动增强现实时图 20 和图 21 的便携式设备的显示器；并且
- [0051] 图 23 示出了说明本发明的增强现实实施例的流程图。

具体实施方式

[0052] 图 1 中示出了系统 100。在该系统 100 中，由摄像机装置 130 捕获场景图像。在实施例中，场景是体育赛事，例如足球比赛，尽管本发明不限于此。在该摄像机布置 130 中，三个高清摄像机位于平台上（未示出）。装置 130 使得能够生成缝合图像。装置 130 因此具有都捕获相同场景的不同部分的摄像机，所述不同部分具有每个摄像机之间的视场中的小重叠部分。三个图像都是高清图像，其当被缝合在一起时，产生超高清图像。摄像机布置 130 中的三个摄像机捕获的三个高清图像被馈送到执行图像编辑（例如颜色增强）的图像处理器 135。此外，图像处理器 135 从摄像机布置 130 中的摄像机接收与摄像机参数（例如焦距、缩放因子等）相关的元数据。增强的图像和元数据被馈送到将在随后参考图 4A 说明的第一实施例的服务器 110 或将在随后参考图 4B 说明的第二实施例的服务器 110’。

[0053] 在实施例中，在用户设备 200A-N 中执行实际的图像缝合。然而，为了减少用户设备 200A-N 内的计算消耗，在图像处理设备 135 连接到的服务器 110 内计算执行缝合所需的参数。服务器 110 可直接或经由网络（例如局域网、广域网或因特网）有线或无线地连接到图像处理器 135。计算参数和实际执行缝合的方法在 GB 2444566A 中描述。在 GB 2444566A 中还公开了适当类型的摄像机布置 130。这里合并了与参数计算、缝合方法和摄像机布置相关的 GB 2444566A 的内容。

[0054] 如 GB 2444566A 中所述，确定用于摄像机布置 130 中每个摄像机的摄像机参数。这些参数包括用于每个摄像机的焦距和相对侧转、斜度和旋转以及校正透镜畸变、桶形畸变等的参数并且在服务器 110 上被确定。此外，缝合图像所需的其它参数，例如色差校正参数、色度测量和曝光校正参数，也可在服务器 110 中计算。此外，如技术人员将会理解的，可能有图像缝合处理中所需的在服务器 110 中计算的其它值。这些值在 GB2444566A 中进行说明并且因此为了简洁将不在以下进行说明。服务器 110 中计算的这些值被发送到每个用户设备 200A-N，如将在随后所说明。

[0055] 除了在服务器 110 内计算的图像缝合参数，也出现其它计算。例如，出现对象检测和分割，其识别和提取可对其应用三维效果的图像中的对象。还在服务器 110 内确定标识图像内每个所检测对象的位置的位置信息。

[0056] 此外,在服务器 110 内生成深度图。深度图用所捕获场景中离开摄像机的相应距离来分配摄像机所捕获的图像中的每个像素。换言之,一旦为所捕获图像完成了深度图,可以确定对应于像素的场景中的点和捕获图像的摄像机之间的距离。还在服务器 110 内维护的是被周期性更新的背景模型。更新背景模型以使得以不同的速率来更新背景图像的不同部分。特定地,取决于图像的部分是否被检测为先前帧中的选手来更新背景模型。

[0057] 作为候选,服务器 110 可具有两个背景模型。在该情况下,在服务器 110 内维持长期背景模型和短期背景模型。长期背景模型定义了更长的时段(例如 5 分钟)内图像中的背景,而短期模型定义了更短的时段(例如 1 秒钟)内的背景。短期和长期背景模型的使用使得能够考虑到短期事件(例如照明变化)。

[0058] 在服务器 110 内计算的深度图被发送到每个用户设备 200A-N。在实施例中,摄像机布置 130 内的每个摄像机是固定的。这意味着深度图不会随着时间改变。然而,在触发以允许新的用户设备连接到服务器 110 时,用于每个摄像机的深度图被发送到每个用户设备 200A-N。例如,当新用户设备向服务器 110 注册时或者按时周期性地,可将深度图发送出去。如将会理解的,如果摄像机的视场移动了,将需要更频繁地重新计算深度图并且将其发送到用户设备 200A-N。然而,还设想,深度图被持续地发送到每个用户设备 200A-N。

[0059] 将在随后说明生成深度图和背景模型的方式。此外,将在随后说明执行对象检测和对象分割的方式。

[0060] 还连接到服务器 110 的是多个用户设备 200A-N。在实施例中,这些用户设备 200A-N 经由因特网 120 连接到服务器 110。然而,要理解本发明不限于此并且用户设备 200A-N 可经由任何类型的网络(例如局域网(LAN))连接到服务器 110,或者可有线连接到服务器 110 或无线连接到服务器 110。还连接到每个用户设备的是相应的显示器 205A-N。显示器 205A-N 可以是电视、或者监视器或能够显示可由用户感觉为三维图像的图像的任何种类的显示器。

[0061] 在本发明的实施例中,用户设备 200A-N 是 PlayStation®3 游戏终端。然而,本发明不限于此。实际上,用户设备可以是机顶盒、计算机或能够处理图像的任何其它类型的设备。

[0062] 还经由因特网 120 连接到服务器 110 和每个用户设备 200A-N 的是社区集线器 1700(有时称作为网络服务器)。将在随后说明社区集线器 1700 的结构和功能。

[0063] 图 2 中示出了用户设备 200A 的示意图。用户设备包含存储介质 220。在本发明的实施例中,存储介质 220 是硬盘驱动器,但是本发明不限于此。存储介质可以是光学介质、或半导体存储器等。

[0064] 连接到存储介质 220 的是中央处理器 250。在实施例中,中央处理器 250 是单元处理器。在实施例中,单元处理器是有利的,因为它特别适于复杂计算(例如图像处理)。

[0065] 还连接到中央处理器 250 的是无线附件接口 210,其可适合于连接到无线附件 210A 并且与其通信。在实施例中,无线附件 210A 是用户操作的设备,其可以是六轴控制器,尽管本发明不限于此。六轴控制器允许用户与用户设备 200A 交互并且对其进行控制。

[0066] 此外,图形处理器 230 连接到中央处理器 250。图形处理器 230 可操作为连接到显示器 205A 并且控制显示器 205A 显示立体像。

[0067] 如将会理解的,其它处理器(例如音频处理器 240)连接到中央处理器 250。

[0068] 参考图 3, 示出了系统 100 的不同实施例。该不同系统被称为 100' (其中相同的数字指代相同的部件) 并且被配置为经由长期演进 3GPP 网络提供内容。在该不同的实施例中, 服务器 110' 连接到服务网关 305 并且提供特别适于经由移动网络分发的内容。如技术人员将会理解的, 服务网关 305 向多个增强的节点 B 和从其路由用户数据。为了简洁, 图 3 中示出了单个增强的节点 B 310。增强的节点 B 310 与多个用户装置 315A-C 通信。

[0069] 图 4A 示出了服务器 110 的实施例。在图 4A 的实施例中, 图像处理 135 处理的图像被馈送到图像缝合设备 1101。如上所述, 图像缝合设备 1101 生成由被缝合到一起的三个单独所捕获图像所构成的超高清图片。这在 GB 2444566 中被描述并且因此将不在以下进行描述。

[0070] 所缝合图像被馈送到背景生成器 1102, 背景生成器 1102 从所缝合图像移除前景对象。换言之, 背景生成器 1102 生成仅包含所缝合图像背景的图像。将在随后说明背景生成器 1102 的结构和功能。另外, 所缝合图像被馈送到对象关键字产生设备 1103。其识别所缝合图像中的前景对象并且确定每个所识别对象的位置, 这将被说明。

[0071] 所生成的背景被馈送到重定格式设备 1104 和对象关键字产生设备 1103。重定格式设备 1104 将所生成的背景格式化为经由网络 120 传输更适合的格式, 这将在随后说明。

[0072] 来自对象关键字产生设备 1103 的输出被馈送到加法器 1105 和先进视频编码 (AVC) 编码器 1106。特定地, 对象关键字产生设备 1103 的一个输出可操作为控制与 AVC 编码器 1106 相关的量化器。AVC 编码器 1106 的输出产生包括来自摄像机布置 130 的缝合图像和所提取对象的合成流, 这将在随后说明。来自对象关键字产生设备 1103 的输出还包含与对象相关联的元数据。例如, 元数据可包括选手姓名、选手号码或选手生命信息。该元数据被馈送到连接到网络 120 的数据流产生设备 1108。

[0073] 重定格式设备 1104 的输出也馈送到加法器 1105。来自加法器 1105 的输出被馈送到 AVC 编码器 1106。来自 AVC 编码器 1106 的输出被馈送到数据流产生设备 1108。然后, 数据流产生设备 1108 将信号多路复用到一起。然后, 被多路复用的流被转换为数据分组并且经由因特网 120 传送到适当的用户设备。

[0074] 图 4B 示出了候选服务器 110'。在候选服务器 110' 中, 许多组件与相关于图 4A 所论述的相同。这些相同的组件具有相同的标号。然而, 该实施例中的背景生成器 1102' 没有到重定格式设备 1104' 的输出。代替地, 来自图像缝合设备 1101 的输出被馈送到背景生成器 1102' 和重定格式设备 1104'。

[0075] 此外, 在候选服务器 110' 中, 没有加法器。代替地, 来自重定格式设备 1104' 的输出被直接馈送到 AVC 编码器 1106'。此外, 在该实施例中的对象关键字产生设备 1103' 不产生如图 4A 的实施例中所产生的组合图像。

[0076] 用户注册

[0077] 在从服务器 110 向任何用户设备 200A-N 或从候选服务器 110' 向用户装置 315A-C 发送任何内容之前, 各个设备或装置需要向适当的服务器注册。以下涉及用户设备 200A 向服务器 110 的注册并且在图 5 中说明。应当注意, 用户装置将以相同方式向候选服务器 110' 注册。

[0078] 当用户开启用户设备 200A 时, 用户使用无线附件 210A 来选择他们希望在显示器 205A 上观看的特定事件。该事件可以是流行音乐会、体育赛事或任何种类的事件。在以下

示例中,事件是足球比赛。该选择是起始步骤 S50。

[0079] 为了观看事件,用户可需要支付一次性费用,或者事件可以是预订包的一部分。该费用或预订包可通过在观看事件之前,在用户设备 200A 中输入信用卡细节来购买。作为候选,可通过任何其它方法来购买事件,或者实际上事件可以是免费的。为了观看事件,用户将需要向服务器 110 注册。用户设备 200A 因此用作为相关于服务器 110 的客户端设备。该注册发生在步骤 S55 并且允许服务器 110 从用户设备 200A 获得必要的信息,例如 IP 地址等,以使得能够在服务器 110 和用户设备 200A 之间发生通信。此外,在该阶段可由服务器 110 收集其它信息,例如与要由用户观看的事件相关的信息,这允许发生于该用户的针对性广告。

[0080] 在注册后,在步骤 S510 中用户确认他们希望观看的事件并且确认支付细节。

[0081] 在步骤 S515 中,用户设备 200A 从服务器 110 和显示器 205A 接收初始化信息。来自显示器 205A 的初始化信息可包括与屏幕大小相关的信息。这可直接从显示器 205A 获得或者由用户输入。来自服务器 110 的初始化信息可包括深度图。可响应于来自用户设备 200A 的请求提供初始化信息或者可响应于注册从服务器 110 传送初始化信息。作为候选,初始化信息可周期性地被传送到连接到服务器 110 的每个用户设备 200A。这里应当注意,深度图仅需要被提供给用户设备 200A 一次,因为摄像机装置 130 是固定的。在摄像机装置 130 可移动的情况下,那么初始化信息将被更加定期地提供。初始化信息存储在用户设备 200A 内的存储介质 220 中。

[0082] 在步骤 S520 中,服务器 110 提供已从在图像缝合设备 1101 中缝合到一起的图像生成的背景的格式化的高清图像。用户设备 200A 的中央处理器 250 使用格式化的背景图像来生成超高清图像用于显示。另外,处理器 250 生成超高清图像的左版本和右版本和/或超高清图像的可变视场以显示超高清图像的 3D(或立体)表示或图像的视场。

[0083] 如这里所述,用户还可确定它们希望的事件的视场。将使用接口 210A 来选择该视场。GB 2444566A 中还描述了用户设备 200A 使用的允许选择适当视场的方法。

[0084] 另外,对于每个所捕获图像,服务器 110 分析图像以检测图像中的对象。该检测在对象关键字产生设备 1103 中执行,对象关键字产生设备 1103 的功能在下面论述。在检测图像中的对象之后,产生对象块。对象块包含前景对象。这将在随后说明。还产生的是标识在图像中所提取的对象位于哪里的位置数据。这也在随后论述。

[0085] 高清背景图像、图像内的所分割对象和位置数据被发送到用户设备 200A。

[0086] 在用户设备 200A 从服务器 110 接收前述信息之后,用户设备 200A 生成超高清图像。这是步骤 S325。另外,使用深度图、分离的对象块和图像中所检测对象的位置数据,用户设备 200A 将三维效果应用到超高清图像。此外,其它元数据被提供给用户设备 200A。为了改善用户体验,提供对象元数据,例如选手信息。此外,连同每一个对象块一起,可提供宏块数目。这标识了与每一个对象块相关的宏块数目。这减少了将对象块置于背景图像上的用户设备 200A 内的计算消耗。

[0087] 关于候选服务器 110',向用户装置 320A 提供类似信息。然而,在该实施例,提供了重定格式的所捕获和所缝合的图像(而不是服务器 110 的实施例的重定格式的背景图像)。另外,不提供对象块,因为在该实施例中没有另外的三维效果被应用到所检测对象。

[0088] 对象检测和跟踪

[0089] 现在将参考图 6、7 和 8 描述根据本发明示例的对象跟踪。特定地,以下的对象检测和跟踪指服务器 110。然而,相同的对象检测和跟踪技术用于候选服务器 110' 中。

[0090] 图 6 示出了根据本发明示例的对象跟踪方法的流程图。为了跟踪对象,根据被检测为在预定帧数目内基本静止的所接收视频的那部分来构建背景模型。在第一步骤 S60 中,从布置 130 内的一个摄像机接收的视频图像(其表示足球场)被处理以构建图像的背景模型。背景模型被构建以便创建帮助识别和跟踪个人选手的前景掩模。前景掩模将被用于生成随后说明的对象关键字。在步骤 S60,通过为每个像素确定像素的均值以及连续帧之间像素值的差异以便构建背景模型,形成背景模型。因此,在像素均值变化不大的连续帧中,这些像素可被识别为背景像素以便识别前景掩模。

[0091] 该背景/前景分割是图像处理领域中已知的处理并且本技术可利用在 2004 的 ICVGIP 会议中公布的 Manzanera 和 Richefeu 的标题为“A robust and Computationally Efficient Motion Detection Algorithm Based on $\Sigma - \Delta$ Background Estimation”的文档中所描述的算法。然而,本技术不应当被认为限于该已知技术并且用于相关于背景模型生成前景掩模以用于跟踪的其它技术也是已知的。

[0092] 将会理解,在摄像机的视场包含某些观众的情况下,观众未必被包括在背景模型中,因为他们将可能是走来走去的。这是不期望的,因为可能在执行对象跟踪时增加了单元处理器的处理负载,并且是不必要的,因为大多数体育广播员不太可能对跟踪观众中的人感兴趣。

[0093] 在本发明的示例中,可构建单个背景模型或者实际上可构建两个背景模型。在构建单个背景模型的情况下,取决于在先前帧中的该位置是否检测到了选手,以不同的速率更新背景的不同部分。例如,在先前帧中存在选手的时候,可不那么频繁地更新背景以使得选手不会变成背景图像的一部分。

[0094] 作为候选,在创建两个背景模型的情况下,一个模型可在比赛开始时构建并且甚至可在选手来到球场之前完成。这被称为长期背景模型。另外,另一个背景模型在比赛期间被周期性地重新计算以便考虑照明条件的任何变化,例如在比赛期间可变化的阴影。这是短期背景模型。在比赛开始时创建的背景模型和被周期性重新计算的背景模型都存储在服务器 110 中,存储在存储介质中(未示出)。对于以下的说明,使用单个背景模型。

[0095] 在步骤 S605 中,来自摄像机的进入图像去除背景模型以便识别差异区域。因此,从图像去除背景模型并且使用结果图像来生成用于每个选手的掩模。在步骤 S610 中,,相对于当已去除背景模型后的结果图像版本中的像素值创建阈值。通过首先确定视频图像的一系列帧内像素的均值来生成背景模型。根据每个像素的均值,可根据视频图像的帧来计算每个像素的差异。然后,像素的差异用于确定阈值,阈值对于视频图像的全部像素中的每个像素将是变化的。对于对应于部分图像的像素,在变化较大时,例如包括观众的部分,阈值可被设置为高值,而对应于球场的部分图像将具有较低的阈值,因为球场的颜色和 content 将一贯地相同(除去选手的存在)。因此,阈值将确定前景元素是否存在,并且因此可相应地识别前景掩模。在步骤 S615 中,基于与均值人类形状模型的相关性的形状概率被用于提取前景掩模内的形状。此外,从图像提取颜色特征以便创建颜色概率掩模,以便例如根据选手衬衣的颜色来识别选手。因此,每个队衬衣的颜色可被用于将选手相互区分。为此,服务器 110 取决于每个足球队的队伍装备的已知颜色来生成颜色模板。因此,需要每个队的

衬衣颜色,守门员和裁判员的衬衣的颜色。然而,将会理解,可使用其它适当的颜色模板和/或匹配处理的模板。上面说明的背景生成在背景生成器 1102 中执行。

[0096] 返回图 6,在步骤 S615 中,服务器 110 比较每个颜色模板的每个像素和与选手图像的衬衣区域对应的像素。然后,服务器 110 生成表示颜色模板的像素和所选择像素之间的相似性的概率值,以便基于色相饱和度值 (HSV) 色空间中离队伍和衬衣颜色模型的距离形成颜色概率。另外,形状概率被用于局部化选手,这基于和平均人类形状模型的相关性。此外,运动概率基于离由递归最小平方估计器使用开始位置、速率和加速度参数所预测的位置的距离。

[0097] 图 7A 中示出了对象关键字创建设备 1106 对对象关键字的创建。图 7A 示出了装置 130 中的一个摄像机所生成的足球场的摄像机视野 710。如已经说明的,球场形成背景模型的一部分,而选手 730、732、734、736、738、740 应当形成前景掩模的一部分并且都分别如上所述。选手边界框(可被称作为矩形轮廓)被示出为每个选手周围的虚线。

[0098] 迄今,相关于摄像机图像处理来执行步骤 S60、S605、S610 和 S615。设计前景掩模,在步骤 S620 中,在首先按到摄像机的接近度排序选手轨迹之后执行选手跟踪。因此,被识别为最接近摄像机的选手被首先处理以便从跟踪处理去除这些选手。在步骤 S630,更新选手位置以便最大化形状、颜色和运动概率。在步骤 S640 中,构建排除要由其它更接近的选手轨迹覆盖的已经已知的图像区域的遮挡掩模。这确保其它选手部分或完全遮挡的选手仅可被匹配到可视图像区域。遮挡掩模改善了跟踪可靠性,因为它减少了跟踪合并的发生率(由此在遮挡事件后两个跟踪跟随相同的选手)。这是当许多目标看起来相同时的特定问题,因为它们不能(容易地)按颜色区分。遮挡掩模允许像素被分配到附近的选手并且从另外的选手被排除,防止两个轨迹匹配到相同的像素组并且因此维持了他们的不同身份。

[0099] 然后,通过提取摄像机图像内提供的特征并且将它们映射到 3D 模型来遵循跟踪每个选手的处理,如图 7A 和 7B 中所示。因此,对于摄像机产生的 2D 图像内的相应位置,将 3D 位置分配到最大化形状、颜色和运动概率的选手。如将简短说明的,如果检测到遮挡事件,将修改选手从 2D 图像到 3D 模型的选择和映射。为了帮助从 2D 图像到 3D 模型的映射,在步骤 S625 中,初始化要跟踪的选手以便形状和颜色概率中的峰值被映射到对选手的最适当选择。应当强调,在步骤 S625 执行的跟踪初始化仅执行一次,通常在跟踪处理开始时。对于系统的良好跟踪初始化,选手应当被很好地分离。根据本技术在跟踪初始化之后跟踪选手中的任何错误被自动校正,其不需要手工干涉。

[0100] 为了实现根据 2D 图像位置在 3D 模型中的跟踪,通过使用射影矩阵 P 来实现变换。跟踪需要 2D 图像位置可涉及 3D 模型内的位置。该变换通过使用射影 (P) 矩阵来完成。2D 空间中的点等于 3D 空间中的线:

$$[0101] \quad \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} P_{00} & P_{01} & P_{02} & P_{03} \\ P_{10} & P_{11} & P_{12} & P_{13} \\ P_{20} & P_{21} & P_{22} & P_{23} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \\ w \end{bmatrix}$$

[0102] 2D 空间中的点等于 3D 空间中的线,因为作为离摄像机的距离的第三维度不是已知的并且因此将相应地呈现为跨越 3D 模型的线。对象(选手)的高度可被用于确定离摄像机的距离。通过选择位于已知地水准平面上固定高度(平均人类高度)处的沿着线的点

来获得 3D 空间中的点。在摄像机校准处理的匹配之前,针对每个摄像机射影矩阵 P 被推断地获得一次,其中球场的物理特征(例如球场 70 的角落 71a、71b、71c、71d 被用于确定摄像机参数,其因此可帮助将所识别的选手的 2D 位置映射到 3D 模型。这是已知的技术,使用确定的方法。根据物理参数,射影矩阵 P 合并摄像机的缩放级别、光心、3D 位置和 3D 旋转向量(当它被指向时)。

[0103] 步骤 S630 中执行的跟踪算法是可伸缩的并且可在一个或多个摄像机上操作,仅需要球场上的所有点从至少一个摄像机是可见的(以足够的分辨率)。

[0104] 除了颜色和形状匹配,步骤 S630 包括其中还包括被跟踪的选手的运动的处理以使用更大的概率来正确地识别每个选手。因此,可根据有关运动并且在方向上确定帧之间的选手的有关运动。因此,相对运动可被用于随后的帧以产生搜索区域以便识别特定选手。此外,如图 7B 中所示,足球场的 3D 模型可用线 730.1、732.1、734.1、736.1、738.1、740.1 来扩大,线 730.1、732.1、734.1、736.1、738.1、740.1 相对于选手位置的图形表示来定位以便反映足球场上选手运动的相对方向。

[0105] 在步骤 S640,一旦已在 3D 模型中识别了选手的相对位置,那么该位置被相应地投射回足球场的 2D 图像视图并且围绕根据 3D 模型中选手的位置所识别的选手来投射相对边界。此外,在步骤 S640,然后将选手周围的相对边界添加到用于该选手的遮挡掩模。

[0106] 图 7B 示出了足球场的虚拟模型 220 的平面图。在图 7B 中所示的示例中,选手 730、732 和 734(在球场的左手侧)已被服务器 110 识别为穿有颜色不同于选手 736、738 和 740(在球场的右手侧)的足球衬衣,因此表示他们在不同的队伍。以该方式区分选手使得在遮挡事件之后检测每个选手更容易,因为可通过他们衣服的颜色容易地对他们相互区分。

[0107] 返回参考图 6,在步骤 S630,使用已知技术(例如卡尔曼滤波)来跟踪每个选手的位置,尽管将会理解,可使用其它适当技术。该跟踪发生在摄像机视野 710 和虚拟模型 720 中。在本发明的示例中,由服务器 110 使用虚拟模型 720 中选手的位置来执行的速率预测被用于帮助跟踪摄像机视野 710 中的每个选手。

[0108] 重复步骤 S630 和 S640,直到所有选手已被处理为由判定框 S635 表示。因此,如果未处理全部选手,那么处理继续到步骤 S630,而如果处理完成,那么处理在步骤 S645 结束。

[0109] 如图 6 中所示,所示出的方法包括另外的步骤 S650,其在图像由多个摄像机产生时是必需的。照此,可对于来自每个摄像机的视频图像执行处理步骤 S60 到 S645。照此,将为每个选手提供来自每个摄像机的检测概率。因此,根据步骤 S650,根据来自每个摄像机的用于每个选手的概率和根据每个摄像机提供的概率中的最高者所估计的选手位置来估计每个选手的位置,以使得用于每个选手的具有最高概率的位置被识别为用于该选手的位置。该位置为上述的位置数据。

[0110] 如果已确定在跟踪足球场上的选手中出现了错误,那么在步骤 S655 中可重新初始化用于该选手的轨迹。在检测特定选手的概率对于特定轨迹相对较低并且因此重新初始化轨迹的情况下,产生了跟踪中的错误检测。

[0111] 执行图 6 中所示方法的结果是生成用于每个选手的路径数据,其提供了视频图像的每个帧中的选手位置,其表示了选手在比赛期间所采取的路径。该所计算位置是发送到用户设备 200A 的位置数据。因此,路径数据提供相关于时间的位置。

[0112] 如果一个选手遮挡了另一个选手的全部或部分,当从单个摄像机视野跟踪每个选手的位置时可出现问题,如图 8 中所示。

[0113] 图 8 示出了多个选手 810、820、830 和 840 以及与他们相关的边界框(如每个选手周围的虚线所示)。当选手 810 和 840 可被清楚地相互区分时,选手 820 遮挡了选手 830 的部分。这是所谓的遮挡事件。在一个选手的全部或部分遮挡了至少一个其他选手的全部或部分使得即使在考虑了其它因素(例如选手的相对运动和方向)之后对选手的跟踪仍变得模糊,可出现遮挡事件。然而,将会了解,可出现其中涉及两个或多个选手的遮挡事件。

[0114] 为了检测遮挡事件,服务器 110 检测与选手相关的全部或部分掩模是否出现在与和另一个选手相关的全部或部分掩模相同的图像区域中,如图 8 中所示。在遮挡事件中涉及的选手在相反的队伍并且因此具有不同颜色的衬衣的情况下,他们因此可被容易地区分和跟踪。然而,在遮挡事件之后,如果选手在相同的队伍,那么服务器 110 可能不能区分哪个选手是哪个,特别是因为遮挡事件之后他们的运动(其例如由碰撞造成)可能不能预测,并且因此可能不能正确地跟踪选手。结果,分配给每个选手的跟踪路径可被交换。

[0115] 为了解决所跟踪选手中的模糊,服务器 110 用遮挡事件中涉及的所有那些选手的身份来标记遮挡事件中涉及的全部选手。因此,在随后的时间,如果一个或多个选手变得可易于区分,服务器 110 使用该信息来将选手的身份重新分配到正确的选手,以便维持哪个选手是哪个的记录。该处理参考图 9 来被更详细地描述。

[0116] 图 9 示出了根据本发明示例的对象跟踪和遮挡检测的方法的流程图。

[0117] 在步骤 s900,服务器 110 执行对所捕获视频图像的图像处理以便提取一个或多个图像特征,如上参考图 6 所述。然后,将所提取的图像特征与从可能的对象示例提取的相应图像特征比较以便识别每个对象。在示例中,根据衬衣上的号码来识别选手。然后,服务器 110 生成识别每个对象的用于每个对象的对象标识。该标识作为元数据结合图像和位置信息来存储。作为候选,在示例中,每个对象(例如选手)经由操作者接口由操作者来标识。然后,服务器 110 使用从操作者接口输入的数据来生成对象标识数据。然而,技术人员将会理解,图像识别技术与操作者的标识相组合以便生成对象标识数据,或者可使用其它适当的对象标识方法,例如数字识别,其通过选手衬衣背面上的号码来识别选手。

[0118] 在步骤 s905,服务器 110 依赖于在步骤 s900 提取的一个或多个图像特征来检测要被检测的任何对象(例如选手),如上面参考图 6 所述。如上所述,还使用虚拟模型 720 和摄像机视野 710 来跟踪每个选手。服务器 110 使用在跟踪处理期间生成的数据来生成和存储描述每个对象在所接收的视频图像内所采取的路径的对象路径数据。对象路径数据采用选手相关于时间的 x - y 坐标的采样形式。在本发明的示例中,路径数据具有格式 (t_i, x_i, y_i) ,其中 t_i 是采样时间,并且 x_i 和 y_i 是对象在采样时间 t_i 的 x 和 y 坐标。然而,将会理解,可使用其它适当的路径数据格式。

[0119] 在步骤 s915,服务器 110 记录用于每个对象的对象标识数据连同涉及每个对象在视频图像内所采取的路径的对象路径数据。所记录的数据存储在服务器 110 的硬盘驱动器(HDD)上或动态随机访问存储器(DRAM)中。这允许保存关于选手关联于每个所检测和跟踪的路径的记录。然后,所记录的数据可用于生成关于每个选手的以及在比赛期间他们在哪里的数据。例如,选手在球场特定区域中所花费的时间可根据相关记录中存储的数据来生成。该信息可在比赛期间或比赛结束时被发送到用户设备 200A,并且可被显示给用户(如

果他们期望的话)。在本发明的实施例中,所显示的所记录数据可包括选手覆盖的距离等。这将由用户设备 200A 的用户选择。此外,如果出于任何原因,选手和路径之间的关联变得模糊,例如如同可在遮挡事件之后发生的,那么该记录可被保存,直到解决了模糊性,如下所述。所记录的对象标识数据连同对象路径数据的示例在下面的表格 1 中示出。

[0120]

对象ID	t	x	y
A	t_1	x_1	y_1
A	t_2	x_2	y_2
A	t_3	x_3	y_3
\vdots	\vdots	\vdots	\vdots
A	t_i	x_i	y_i

[0121] 表格 1

[0122] 因此,用于每个对象的对象标识数据和用于该对象的对象路径数据之间的关联允许跟踪和识别每个对象。在上述示例中,可跟踪每个选手,因此使得广播员能够知道哪个选手是哪个,即使选手可能离得太远而不能由操作者或由服务器 110 执行的图像识别可视地识别。这允许广播员基于该关联来合并广播内容的观众可期望找到的另外特征和信息。在步骤 s920,服务器 110 检测是否已出现遮挡事件,如上参考图 6 所述。如果未检测到遮挡事件,那么处理返回到其中检测对象的步骤 s905。以该方式,每个对象可被单个地跟踪并且每个对象的路径唯一地关联于该对象的身份。

[0123] 然而,如果检测到遮挡事件,那么在步骤 s925,服务器 110 将用于遮挡事件中涉及的每个对象的对象标识数据与用于遮挡事件中涉及的每个对象的对象路径数据关联。例如,如果标记为 A 和 B 的两个对象分别关联于路径 P 和 Q,在检测到涉及对象 A 和 B 的遮挡事件之后,路径 P 将关联于 A 和 B,并且路径 Q 将关联于 A 和 B。然后记录在遮挡事件之后服务器 110 生成的关联,如上所述。这允许跟踪遮挡事件中涉及的对象(例如选手),而不必重新识别每个对象,即使关于哪个选手是哪个存在某些不确定性。因此,减少服务器 110 上的处理负载,因为仅遮挡事件中涉及的那些对象被模糊地识别,而遮挡事件中未涉及的对象仍可被识别。

[0124] 在步骤 s930,服务器 110 检查来察看是否已识别了遮挡事件中涉及的一个或多个对象,以使得可解决与所生成路径相关联的对象的身份。通过比较与该对象相关的一个或多个图像特征和从可能的对象示例提取的图像特征,由服务器 110 执行对至少一个对象的识别。如果未做出识别,那么处理进行到步骤 s905,用于每个对象的所生成路径数据关联于遮挡事件中涉及的所有那些对象。

[0125] 然而,如果检测到出现了遮挡事件中涉及的一个或多个对象的标识,那么在步骤 s935,更新所记录的路径数据以反映被确定地识别的对象的身份。在上面给出的示例中,将更新关联记录以使得 A 关联于路径 P,并且 B 关联于路径 Q。

[0126] 作为候选,由服务器 110 使用根据本发明示例的图像识别技术(如下所述)或通过两种技术的组合,可由操作者经由操作者接口执行对象的识别。然而,将会理解,可使用

适于区分或识别每个对象的任何其它识别技术。在图像识别的情况下,服务器 110 可生成表示要多可能地校正图像识别处理所进行的识别的可信度。在本发明的示例中,识别被确定是在可信度高于预定阈值的情况下。另外,操作者可向他们的识别分配可信度,并且如果该可信度超过预定阈值,那么检测识别。

[0127] 在本发明的示例中,生成表示何时更新了所记录路径数据的事件历史并且这也可被存储以便在明确的识别变为不正确的情况下用作为备份。例如,当操作者相信远离摄像机布置 130 的选手具有特定身份但是当选手更靠近摄像机(允许用户看到选手的更高分辨率的图像)时识别可变为不正确。在该情况下,他们可使用操作者接口来覆盖他们先前对选手的识别,以便服务器 110 可因此更新所记录的路径数据。在上面给出的示例中,识别事件历史可存储在服务器 110 的硬盘驱动器(HDD)上或者动态随机访问存储器(DRAM)中,该识别事件历史具有示出在确定的识别之前往往关联于 A 和 B 的路径 P 以及往往关联于 A 和 B 的路径 Q 的数据。

[0128] 识别事件历史也可包括在识别处理期间生成的可信度。如果随后的识别由具有比先前的确定识别更高的可信度的对象形成,那么随后识别的可信度可被用于验证或取消先前的验证。

[0129] 将会理解,在遮挡事件的检测之后,可在遮挡事件之后的任何时刻识别对象以便对遮挡事件中涉及的对象消除歧义。因此,在遮挡事件的检测之后,作为与步骤 s105 到 s125 同时运行的后台处理,服务器 110 可监测是否已出现了对象的确定识别。

[0130] 现在将参考图 10a 和 10b 描述根据本发明示例的对象跟踪和遮挡检测的某些示例。

[0131] 在图 10a 中所示的示例中,遮挡事件 1010 中涉及识别为 A 和 B 的两个对象。在遮挡事件之后,如箭头所表示的两个所检测对象路径关联于 A 和 B(AB)。一段时间后,对象 B 被确定地识别,如在较低路径上的 AB 所表示。然后,该识别被用于更新对象和路径之间的关联以使得在遮挡事件 1010 之后对象 A 关联于较高的路径,并且在遮挡事件 1010 之后对象 B 关联于较低的路径。

[0132] 在图 10b 中所示的示例中,在遮挡事件 420 中初始地涉及对象 A 和 B。然而,在可确定地识别对象 A 和 B 之前,在遮挡事件 1020 之后在较低路径上与 A 和 B 关联的对象在具有对象 C 的另一个遮挡事件 1030 中被涉及。因此,在遮挡事件 1030 之前,在遮挡事件 1020 之后较低路径上的对象是对象 A 还是对象 B 是不清楚的。因此,在遮挡事件 1030 之后,两个对象跟随的较高和较低路径关联于对象 A、B 和 C(ABC)。

[0133] 在随后的时间,在遮挡事件 1030 之后的较低路径上的对象被明确识别为是对象 B(ABC)。因此,可更新关联记录以使得遮挡事件 1030 之后的较高路径关联于对象 C。此外,该信息可用于更新关联记录以使得遮挡事件 1020 中涉及的两个对象可被消除歧义,因为当遮挡事件 1030 之后对象 B 被确定地识别为关联于较低路径时在遮挡事件 1030 中涉及的必然是对象 B。因此,可更新关联记录以使得在遮挡事件 1020 之后的较高路径关联于对象 A 并且遮挡事件 1020 之后的较低路径关联于对象 B。

[0134] 因此,本发明的示例允许对象关联于对象的所跟踪路径,即使在对象被明确识别之前可出现多个遮挡事件。此外,本发明的示例允许不同对象的身份被相互交叉引用以便允许每个路径关联于正确的对象。

[0135] 在某些示例中,识别对象开始位置的数据可被用于初始化和验证对象跟踪。用足球作为示例,选手可能在比赛场地上近似固定的位置开始比赛。每个选手可能位于距离比赛场地上特定坐标的阈值距离内。开始位置可依赖于队伍排列,例如 4-4-2(四个防守、四个中场、两个进攻)或 5-3-2,以及哪个队伍中场开球和哪个队伍防守中场开球。类似的位置可能由来自从地面进行的球门球的选手所采用。这些位置信息可被用于发起选手跟踪,例如通过比较位置数据和队伍表和排列信息。该位置信息也可被用于在出现遮挡事件时纠正路径信息。使用队伍排列信息是有利的,因为如果队伍排列的变化例如在替补或罚下之后变得明显了,这可由操作者在比赛过程期间重置。这将改善对象跟踪的精确性和可靠性。

[0136] 构建超高清图像内每个对象(或者在该示例中为选手)的位置。另外,构建在图 7A 中分别示出为框 730 到 740 的每个选手周围的框。每个框将包含选手的图像并且所以将被称作为“选手框”。当使用 AVC 编码器 1106' 编码图像时,选手框将形成图像内的一个或多个宏块。因为选手框对于用户以及对于用户设备上立体像的创建将是重要的,所以图像内的选手框的宏块地址由对象关键字生成器 1103' 生成。对象关键字生成器 1103' 向对象关键字生成器 1103' 内的量化控制提供宏块地址,这确保选手框被编码为相比于剩余图像更高的分辨率。这确保所编码图像经由其传送的网络带宽被最有效地利用。

[0137] 这里应当注意,在服务器 110 的对象关键字生成器 1103 中,除了生成的对象位置和宏块号码,从超高清图像提取选手块的内容。换言之,在对象关键字生成器 1103 中,从超高清图像提取单个选手。然而,在候选服务器 110' 的对象关键字生成器 1103' 中,仅生成位置和宏块号码,并且不提取选手框的内容。

[0138] 重定格式设备

[0139] 现在将参考图 11 描述服务器 110 的重定格式设备 1104。背景生成器所生成的超高清图像的背景被馈送到换算设备 1150。超高清图像的背景是 $6k \times 1k$ 的像素尺寸。换算设备 1150 将该比例减少为 3840×720 像素。如应当注意的,水平方向上的换算数量小于垂直方向的。换言之,水平方向上的数据减少小于垂直方向的数据减少。这在捕获像足球比赛这样的事件时特别有用,因为球在水平方向上移动并且选手的大部分移动是在水平方向上。因此,确保水平方向上的分辨率较高时重要的。然而,本发明不限于此并且如果存在垂直移动最重要的情况下捕获事件的图像的情况,那么垂直方向上的换算数量将少于水平方向的。

[0140] 换算的图像被馈送到帧分解器 1160。帧分解器 1160 在水平方向上平等地分解换算的背景图像。帧分解器 1160 被配置为产生 1920×1080 像素的两个帧。这是符合 108030P(1920) 帧 AVCHD 格式的。这两个帧被馈送到加法器 1105。

[0141] 如将在这里所述的,帧分解器 1160 在垂直方向上添加 360 个空白像素。然而,为了有效地利用带宽,该空白空间将具有其中插入的对象关键字生成器 1103 所提取的单独选手框。这意味着单独选手框可以有效方式经由因特网 120 被传送。在加法器 1105 中单独选手框被插入到两个图像内。这意味着被馈送到 AVC 编码器 1106 的来自加法器 1105 的输出包括合成图像(包括换算和分解的背景以及被插入到 360 个空白像素内的单独选手框)。

[0142] 参考图 12,描述了候选服务器 110' 的重定格式设备 1104'。在该情况下,超高清图像被馈送到换算器 1150', 换算器 1150' 被配置为将超高清图像换算为 2880×540 像素的图像。换算的图像被馈送到帧分解器 1160'。帧分解器 1160' 被配置为在水平方向上平等地

分解换算的图像并且形成 1440×1080 像素大小的一个图像并且因此符合 108030P(1440) 帧 AVCHD 格式。换言之,换算的图像的左侧形成所生成图像的上半部分并且换算的图像的右侧形成所生成图像的下半部分。该单个图像被馈送到 AVC 编码器 1106'。

[0143] AVC 编码

[0144] 现在将描述服务器 110 中的 AVC 编码器 1106 执行的 AVC 编码。如先前所述,对象关键字生成器 1104 生成选手框并且从超高清图像提取选手框的内容。选手框的内容在换算和分解的合成图像中的空白的 360 个像素中提供。与选手框的位置(即空白像素中每个选手框的位置)关联的宏块被馈送到 AVC 编码器 1106 中的量化器。特定地,控制合成图像中选手框的量化以使得 AVC 编码器 1106 相比于图像中的任何其它位置使用更多的位来编码选手框。这改善了选手框的质量,因为用户将注意观看选手框。

[0145] 使用 H. 264 编码对构成背景和选手框的两个合成图像进行 AVC 编码并且用近似 7Mbps 的比特率将其发送,尽管这可取决于网络性能而变化。

[0146] 在候选服务器 110' 中,由 AVC 编码器 1106' 执行 AVC 编码。如上所述,被馈送到 AVC 编码器 1106' 的重定格式的图像时 108030P(1440) 格式的超高清图像。不像服务器 110, 候选服务器 110' 中的对象关键字生成器 1103' 不提取选手框的内容。代替地,每个选手框的位置和与每个选手框相关联的宏块号码被用于控制 AVC 编码器 1106' 的量化。控制量化以确保用比图像的任何其它部分更多的位来编码选手框,以便确保清楚地再现选手。AVC 编码器 1106' 以大约 3Mbps 的比特率使用 H. 264 标准来编码图像,尽管这可取决于网络性能而变化。

[0147] 在任一种服务器中由编码器产生的所编码图像被馈送到数据流产生设备 1108。另外馈送到数据流产生设备 1108 的是与各个选手框相关联的宏块号码和所编码图像中每个选手框的位置。这被发送到客户端设备 200A 或用户装置作为元数据。

[0148] 深度图和位置数据生成

[0149] 现在将参考图 13 到 15 描述其中摄像机和摄像机捕获的图像内的对象之间的距离被用于确定偏移量的本发明实施例。这在位于服务器 110 和候选服务器 110' 中的深度图生成器 1107 中执行。

[0150] 图 13 是根据本发明实施例的用于确定摄像机的位置和摄像机视场内的对象之间的距离的系统的示意图。

[0151] 图 13 示出了布置为传送摄像机布置 130 中的摄像机(其捕获球场 70 的图像)的服务器 110。如上所述,服务器 110 可操作为分析摄像机捕获的图像以便跟踪球场 70 上的选手,并且确定他们在球场 70 上的位置。在某些实施例中,系统包括可操作为检测摄像机和摄像机视场内的对象之间的距离的距离检测器 1210。距离检测器 1210 和它的操作将随后在下面更详细地被描述。

[0152] 在某些实施例中,服务器 110 可使用跟踪数据和位置数据来确定摄像机的位置和球场上球员之间的距离。例如,服务器 110 可分析所捕获图像以便确定摄像机位置和选手 1201 之间的距离 1201a、摄像机位置和选手 1203 之间的距离 1203a、以及摄像机位置和选手 1205 之间的距离 1205a。

[0153] 换言之,本发明的实施例确定场景内对象和相关于摄像机定义的基准位置之间的距离。在参考图 13 所述的实施例中,基准位置位于摄像机的位置。

[0154] 另外,在某些实施例中,服务器 110 可操作为检测与场景内已知特征点对应的所捕获图像内的预定图像特征。例如,服务器 110 可使用已知技术分析所捕获图像以便检测对应于足球场特征(例如角落、中心点、罚球区等)的图像特征。基于所检测的已知特征点(图像特征)的所检测位置,服务器 110 然后可使用已知技术将球场 70 的三维模型映射到所捕获图像。因此,然后服务器 110 可分析所捕获图像以便取决于关于已映射到所捕获图像的 3D 模型的选手所检测位置来检测摄像机和选手之间的距离。

[0155] 在本发明的某些实施例中,服务器 110 可分析所捕获图像以便确定选手的脚与球场接触的位置。换言之,服务器 110 可确定对象(例如选手)与平面(例如球场 70)相符合的相交点。

[0156] 在对象被检测为在多个相交点与平面相符合时(例如选手的两只脚与球场 70 接触),那么服务器 110 可操作为检测哪个相交点最接近摄像机并且使用该距离用于生成偏移量。作为候选,当生成偏移量时用于该对象的全部所检测相交点的平均距离可被计算和使用。然而,将会理解,可选择其它适当的相交点,例如距离摄像机最远的相交点。

[0157] 然而,在某些情况下,如上所述确定摄像机位置和场景内对象之间距离的方法可造成三维图像的呈现失真。在图像由非常广角的摄像机捕获或通过许多高清摄像机捕获的图像缝合在一起而形成(例如本发明实施例中的情况)时该失真可能特别明显。

[0158] 例如,在球场 70 要被显示为选手和球在其上重叠的三维图像时,三维图像中的图像失真可能发生。在该情况下,角落 71b 和 71c 将呈现为相比于中心点 1214 在最接近摄像机 30 的边线上更加远离。边线因此可呈现为弯曲,即使在所捕获图像中边线为笔直的。

[0159] 当在相对小的显示器(例如计算机监视器)上观看三维图像时,该效果可尤其明显。如果在比较大的屏幕(例如电影院屏幕)上观看三维图像,那么该效果不那么明显,因为角落 71b 和 71c 更可能位于观众的外围视觉中。随后将在下面更详细地描述可将球场显示为三维图像的方式。

[0160] 解决该问题的可能方式将是对于图像的每个部分生成适当偏移量以便补偿失真。然而,这是计算密集型,并且依赖于多个物理参数,例如由于广角图像造成的失真程度、显示器大小等。

[0161] 因此,为了减少三维图像中的失真并且试图确保球场的在前面(即最接近摄像机的边线)以距离显示器的恒定深度呈现,特别是当要在相对小的显示器(例如计算机监视器或电视屏幕)上观看三维图像时,本发明的实施例确定对象和位于基准线上的基准位置之间的距离。基准线正交于摄像机的光轴并且通过摄像机的位置,并且基准位置位于基准线上对象位置线和基准线相交的点。对象位置线正交于基准线并且通过对象。这将在下面参考图 14 来描述。

[0162] 图 14 是根据本发明实施例的用于确定摄像机和摄像机视场内的对象之间的距离的系统的示意图。图 14 中所示的实施例基本上与上面参考图 9 所述的相同。然而,在图 14 中所示的实施例中,服务器 110 可操作为确定对象和虚线 1207 所表示的基准线之间的距离。

[0163] 如图 14 中所示,基准线 1207 正交于摄像机的光轴(即处于与光轴的正确角度)并且通过摄像机的位置。另外,图 14 示出了位于基准线 1207 上的基准位置 1401a、1403a 和 1405a。

[0164] 例如,工作站可操作为确定基准位置 1401a 和选手 1201 之间的距离 1401。基准位置 1401a 位于基准线 1207 上用于选手 801 的对象基准线(由虚线 1401b 表示)与基准线 1207 相交的地方。类似地,基准位置 1403a 位于基准线 1207 上用于选手 1203 的对象基准线(由虚线 1403b 表示)与基准线 1207 相交的地方,并且基准位置 1405a 位于基准线 1207 上对象基准线(由虚线 1405b 表示)与基准线 1207 相交的地方。对象基准线 1401b、1403b 和 1405b 正交于基准线 1207 并且分别通过选手 1201、1203 和 1205。

[0165] 在某些实施例中,基准线 1207 平行于连接角落 71b 和 71c 的边线,以使得当以适当方式在显示器上一起观看球场的所捕获图像和球场的所修改图像时,连接角落 71b 和 71c 的边线上的所有点呈现为似乎距离显示器恒定的距离(深度)。这改善了三维图像呈现,而不必生成补偿任何失真的偏移量,所述失真可在使用广角摄像机或根据通过组合两个或多个摄像机捕获的图像形成的合成图像(如本发明实施例中的情况)捕获图像时出现。然而,将会理解,基准线需要平行于边线,并且可平行于场景内的任何其它适当特征,或者相关于场景内的任何其它适当特征来布置。

[0166] 为了图像生成成为使得当被观看时,它们呈现为三维的,服务器 110 可操作为检测所捕获图像内对象(例如选手)的位置。服务器 110 检测图像内对象的方式在上面参考图 6 进行了描述。该信息被馈送到用户设备 200A。用户设备 200A 然后通过按偏移量移动所捕获图像内对象的位置根据所捕获图像生成所修改图像,以使得当在显示器 205 上将所修改图像和所捕获图像作为一对图像一起观看时,对象呈现为位于距离显示器的预定距离处。这将在下面说明。

[0167] 为了产生正确的移动以便模拟三维效果,用户设备 200A 需要知道对象离摄像机的距离。这可使用深度图,或某些其它方法来实现。在本发明的某些实施例中,系统包括可经由网络与服务器 110 或用户设备 200A 通信的距离检测器 1210。距离检测器 1210 可连接到摄像机布置 130 内的摄像机,或者它可与摄像机布置分离。距离检测器可操作为生成表示摄像机和球场 70 上的对象(例如选手)之间的距离的距离数据。距离检测器 1210 可操作经由适当通信链路向服务器 110 发送距离数据,如图 13 中的虚线 1212 所表示。然后服务器 110 可操作为依赖于从距离检测器 1210 接收的距离数据来确定摄像机和对象之间的距离。换言之,距离检测器 1210 用作距离传感器。该距离传感器是领域中已知的并且可使用红外光、超声波、激光等来检测到对象的距离。用于每个对象的距离数据然后被馈送到用户设备 200A。

[0168] 在某些实施例中,距离检测器可操作为生成对于所捕获图像的每个像素表示摄像机和符合该像素的场景内的场景特征之间的相应距离的深度图数据。然后,从服务器 110 发送到用户设备 200A 的距离数据可包括距离图数据。

[0169] 为了实现该功能,距离检测器可包括发射红外光脉冲的红外光源。然后摄像机可按预定时间间隔(通常是纳秒级)检测从摄像机视场内的对象反射的红外光的强度以便生成表示对象离摄像机的距离的灰度图像。换言之,灰度图像可被认为是根据检测红外光从源到摄像机的飞行时间所生成的距离图。

[0170] 为了简化设计,摄像机可包括红外光源形式的距离检测器。该摄像机是领域中已知的,例如 3DV 系统制造的“Z-Cam”。然而,将会理解,可使用生成 3D 深度图的其它已知方法,例如红外模式失真检测。

[0171] 将会理解,可使用任何其它的适当距离检测器。例如,可使用具有垂直于摄像机光轴的光轴的摄像机来捕获球场图像。这些另外捕获的图像可由服务器 110 来分析以检测和跟踪选手位置和与来自摄像机的图像数据相关的合成数据,以便更精确地对选手位置做三角测量。

[0172] 在某些实施例中,服务器 110 可操作为使用距离检测器 1210 来检测和跟踪摄像机视场中的其它对象,例如足球,尽管将会理解,可检测任何其它适当对象。例如,一个或多个另外摄像机捕获的图像可由服务器 110 分析并且与来自跟踪系统的数据组合,以便跟踪足球。该数据被馈送到用户设备 200A 作为位置和深度信息,以使得用户设备 200A 可因此生成适当的左手和右手图像。

[0173] 服务器 110 可操作为检测与场景内的对象对应的所捕获图像内的对象像素。在上述实施例中,对象像素对应于用于生成所修改图像的选手掩模的那些像素,如下所述。选手掩模被馈送到用户设备 200A 以使得用户设备 200A 可生成所修改图像。

[0174] 然后,用户设备 200A 使用距离图数据中与选手掩模的像素相关联的距离数据来确定摄像机和选手之间的距离。为了简化三维显示,与选手掩模的像素对应的距离图数据中的距离值的平均值可被用于生成偏移量,如上所述。然而,将会理解,可使用从对应于对象的距离图数据选择距离值的任何其它适当方法。

[0175] 用户设备 200A 可操作为生成偏移量以对于深度图数据中的每个像素应用在左手图像和右手图像之间。因此,在应用了不一致性之后,当左手图像和右手图像作为一对图像在显示器上被一起观看时(如上所述),对象可具有改善的三维呈现,因为对象的表面维度可被更精确地再现,而不是将对象显示为好像它是距离显示器某距离处的二维图像。

[0176] 用户设备 200A 和用户装置 320A

[0177] 现在将参考图 15A 描述用户设备 200A 的实施例。用户设备 200A 包括经由因特网接收多路复用的数据流的分路器 1505。分路器 1505 连接到 AVC 解码器 1510、音频解码器 1515 和客户端处理设备 1500。分路器 1505 将多路复用的数据流多路分解为 AVC 流(其被馈送到 AVC 解码器 1510)、音频流(其被馈送到音频解码器 1515)和深度图数据、选手元数据(例如选手的姓名)、和任何其它元数据(其被馈送到客户端处理设备 1500)。用户还可使用将数据发送到客户端处理设备 1500 的控制器 1520 与用户设备 200A 交互。将参考图 16A 详细描述客户端处理设备 1500。

[0178] 将参考图 15B 描述用户装置 315A 的实施例。如将明显的,用户装置 315A 中的许多组件与相关于用户设备 200A 所描述的是相同的或者提供类似的功能。这些组件具有相同的标号并且将不再被进一步描述。然而,如根据图 15B 将明显的,代替图 15A 中的客户端处理设备 1500,提供用户装置处理设备 1500'。然而,应当注意,用户装置处理设备 1500' 接收类似于到客户端处理设备 1500 的数据并且将在图 15B 中描述用户装置处理设备 1500' 的功能。图 15B 中的用户控制 1520 可作为触摸屏或键盘等被集成到用户装置 315A 内。

[0179] 客户端处理设备 1500

[0180] 客户端处理设备 1500 包括生成要显示的左图像和右图像的图像处理单元 1600。图像处理单元 1600 从服务器 110 接收两个合成的背景图像。来自服务器 110 的两个合成背景图像还被馈送到选手框提取设备 1615。选手框提取设备 1615 从合成图像提取选手框。所提取的选手框被馈送到图像处理单元 1600。还从选手框提取设备 1615 馈送到图像处理

单元 1600 的是每个背景合成图像上每个选手框的位置和与选手框相关联的宏块号码。这使得图像处理单元 1600 能够在背景合成图像上的正确位置放置选手框,以便有效地重新创建超高清图像的两个合成图像。由图像处理单元 1600 将两个合成图像缝合在一起以形成超高清图像。

[0181] 包括选手框中每个选手的名称的选手元数据在数据控制器 1610 中被接收。还馈送到数据控制器 1610 的是来自用户控制器 1520 的信息和提供允许用户选择适当视场的摄像机布置的参数等的另外元数据,如 GB2444566A 中所述。数据控制器 1610 的输出时包含该信息的多路复用的数据流。数据控制器 1610 的多路复用的输出被馈送到虚拟摄像机生成器 1605。此外,虚拟摄像机生成器 1605 接收深度图。当从用户控制 1520 向虚拟摄像机生成器 1605 馈送信息时,虚拟摄像机生成器 1605 识别虚拟摄像机的边界。换言之,用户操纵用户控制器 1520 以确定超高清图像的哪个区域或段对他们是很重要的。虚拟摄像机生成器 1605 选择重要的超高清片段并且显示该区域。GB 2444566A 中描述了生成和显示区域的方法。

[0182] GB 2444566A 中的方法涉及生成单个图像。然而,在本发明的实施例中,所选择区域可被立体地显示。换言之,所选择区域应当被显示以使得它可以 3D 形式被观看。为了实现此,生成移动的所选择段,其具有的背景具有按照取决于深度图的数量移动的数量像素和水平移动的前景对象。因为用户所选择区域的屏幕上的位置是已知的,并且要在其上显示图像的屏幕的大小是已知的,使用所选择区域距离摄像机的相应距离(即深度图),确定前景对象之间的不一致性(即,用户定义的段中的前景对象和第二所选择的段之间的水平移动),如技术人员将会理解的。该不一致性确定了与屏幕上的前景对象相关联的明显深度。然后,用户选择的段显示在要由用户左眼观看的显示器上并且移动的所选择段显示在要由用户的右眼观看的显示器上。用户选择的段和移动的所选择段被立体地显示。此外,用户可控制移动量,这允许用户调节所选择段的左眼图像和右眼图像之间的移动量以便调节 3D 图像中场景的明显深度。

[0183] 用户装置处理设备 1500'

[0184] 现在将参考图 16B 描述用户装置处理设备 1500'。经由 LTE 网络发送的合成图像被馈送到用户装置图像处理器 1600'。另外提供到用户装置图像处理器 1600' 的是提供允许用户选择超高清图像的区域以用于显示的摄像机参数等的另外元数据。所需的元数据在 GB 244566A 中叙述并且允许用户选择超高清图像的区域用于观看。在 GB 244566A 中还描述了选择和显示区域的方法。

[0185] 用户装置处理设备 1500' 还具有输入到其的选手元数据(表示选手位于合成图像中的哪里)。在实施例中,该选手元数据是在合成图像中定义围绕选手的框的一组坐标。另外的选手元数据可包括每个选手的姓名和统计数据,例如年龄、先前俱乐部、队伍中的位置等。选手元数据和另外的选手元数据被馈送到用户装置数据控制器 1610' 内。还馈送到用户装置数据控制器 1610' 的是由用户控制设备 1520' 产生的用户生成的控制信息。这允许用户与用户装置交互以便改变超高清图像中所选择区域的位置以及其它交互性控制。

[0186] 用户装置数据控制器 1610' 的输出被馈送到虚拟摄像机处理设备 1605' 作为多路复用的数据流。还馈送到虚拟摄像机处理设备 1605' 的是深度图。虚拟摄像机处理设备 1605' 以与上面关于虚拟摄像机生成器 1605' 所述的相同方式生成用户选择的左图像段

和右图像段。这提供了立体图像用于 3D 显示。应当注意,虚拟摄像机处理设备 1605'稍微不同于虚拟摄像机生成器 1605 之处在于整个图像被视为背景,所以所选择区域中的每个图像像素按照取决于深度图的数量被移动,不管它是构成背景的一部分还是前景对象的一部分。每个像素按照所计算的不一致性所提供的数量(其根据深度图和显示器大小计算,如技术人员将会理解的)被水平移动。这允许在显示器上对场景进行 3D 观看。

[0187] 应当注意,在参考图 16A 和 16B 所述的两个实施例中,定义虚拟摄像机的缩放、摇动、倾斜和收敛的信息,以及定义屏幕上所选择区域的位置的细节,和任何其它用户定义的信息(例如水平移动的任何变化)将由用户设备 200A 或用户装置 315A 存储。另外存储的是唯一标识符,例如与特定片段(其中体验该观看)相关联的 UMID。该信息将作为元数据被存储,其包含比所显示的图像数据更少的数据,并且可存储在用户设备 200A 或用户装置 315A 之一上或者网络服务器 1700 上。该所存储元数据当结合合成图像一起被提供时,选手关键字(如果需要)和选手信息将使得用户能够在用户设备 200A 或用户装置 315A 之一上重新创建相同的体验。此外,如果提供给不同的用户,那么该所存储元数据将使得不同用户能够重新创建第一用户的体验。将参考图 17 到 19B 来说明描述使用所存储元数据的实施例。

[0188] 社区观看

[0189] 网络服务器 1700 连接到因特网并且在图 17 中示出。网络服务器 1700 可同样地连接到用户装置 315A 和用户设备 200A。实际上,在实施例中,一个用户可使用用户账户将他的或她的用户装置 315A 和他的或她的用户设备 200A 连接到网络服务器 1700。然而,为了简洁,现在描述用户设备 200A 的连接和使用。

[0190] 参考图 17,网络服务器 1700 包含可以为光或磁记录介质的存储介质 1705。存储介质 1705 连接到在存储介质 1705 上存储信息的数据库管理器 1710。数据库管理器 1710 还用于检索存储介质 1705 上存储的数据。数据库管理器 1710 连接到控制对数据库管理器 1710 的访问的网络处理器 1715。网络处理器 1715 连接到允许数据经由因特网 120 被传送的网络接口 1720。

[0191] 当用户设备 200A 连接到因特网 120 时,用户设备 200A 可连接到网络服务器 1700。当用户设备 200A 首次连接到网络服务器 1700 时,用户被要求登录到网络服务器 1700 上他的或她的账户或者创建新账户。如果用户选择登录到账户,用户被要求输入用户名和密码。这对网络服务器 1700 验证用户。在正确验证之后(其由网络处理器 1715 执行),用户可访问存储介质 1705 上存储的他的或她的账户细节。账户细节可提供与用户喜欢的足球队或用户喜欢的选手相关的信息。通过提供该信息,可以突出部分向用户提供最相关的片段,如将在随后说明。

[0192] 通常,用户可拥有用户设备和用户装置。如果这种情况的话,网络服务器 1700 将存储用户所拥有的装置的细节。网络服务器 1700 还将通过对用户设备的询问来确定是用户设备还是用户装置连接到网络服务器 1700。用户一旦登录了,可从他或她的账户添加或删除设备。

[0193] 与用户帐户相关联的选项之一是上载用户设备 200A 上存储的元数据,这将允许用户或不同用户重新创建用户的观看体验。该元数据可在观看比赛时或当在观看比赛前用户登录到网络服务器 1700 时由用户设备 200A 收集,元数据可存储在网络服务器 1700 内。

如果在用户设备 200A 上收集元数据,那么在用户连接到网络服务器 1700 时,用户可将元数据上载到网络服务器 1700。这可被自动或者在用户指示下完成。

[0194] 除了使得观众的体验能够被复制的元数据,另外的元数据可被传送到网络服务器 1700。另外元数据的生成和形成将参考图 18 来说明,图 18 示出了用户用来生成元数据和另外的元数据的图形用户接口。图 18 中所示的图形用户接口允许用户生成对比赛的评注。这些评注增强了观众的比赛体验。此外,因为仅存储重新创建比赛的元数据,而不是视频剪辑本身,所以减少了被存储来重新创建比赛的数据量。

[0195] 在用户设备 200A 的显示器 205A 上示出了图形用户接口。用户使用控制器 210A 与接口交互。显示器包含显示缝合的超高清图像的缝合图像显示区域 1835。在超高清图像内的是使得用户能够选择缝合图像的视场的虚拟视场。这显示在虚拟视场区域 1800 中。为了用户识别超高清图像的哪部分形成了虚拟视场,虚拟视场 1840 的轮廓示出在超高清图像上。

[0196] 虚拟视场区域 1800 下面的是标准视频控制按钮 1805,例如暂停、快进、倒回、停止和记录。该组视频控制按钮不是限制性的并且可包括控制显示器上的视频动作的任何类型的按钮。在虚拟视场区域 1800 右侧的是编辑按钮 1810。这些编辑按钮 1810 允许对视频的另外评注,例如对视频添加文本、画线或添加形状。当被添加到视频时,这些另外的评注形成了另外元数据的一部分。

[0197] 存在允许元数据标签被添加到视频特定帧的元数据标签输入区域 1815。这可包括帧内容的文本描述,例如处罚、铲球、任意球等。此外,为了使得能够容易地评注,常见标签(例如黄牌、进球和冲突)作为热键 1720 被提供。此外,提供自由的文本输入区域 1825。这允许添加用户希望的任何文本。该文本连同元数据标签输入也形成另外元数据的一部分。

[0198] 最后,提供了事件列表区域 1830。事件列表区域 1830 可按元数据标签自动地更新,或可由用户创建。作为候选,事件列表可使用元数据标签自动地生成,并且可由用户纠正或验证。因为随着比赛进行用户更新进球和警告等,可以自动地生成事件列表。实际上,当在元数据中提供选手位置信息时,如果用户在图像中识别哪个选手进球得分了,那么用户设备 200A 知道哪个选手进球得分了。此外,如果自动地跟踪球的位置,那么用户设备 200A 可以自动将得分者定义为在产生“进球”元数据之前接触球的最后选手。通过使用元数据标签自动更新事件列表,容易生成事件列表。此外,通过使用元数据和另外的元数据,用户设备 200A 或网络服务器 1700 内存储的数据量减少了,因为事件列表是“实时”生成的并且因此不需要被存储。

[0199] 此外对于将元数据上载到网络服务器 1700 上,用户也可访问和观看网络服务器 1700 的其他用户生成的突出节目。换言之,对于访问他们生成的突出包,用户也可访问另外用户生成的突出包。

[0200] 为了实现此,用户设备 200A 需要原始的比赛片段以及不同用户上传的元数据和另外的元数据。原始的比赛片段可从网络服务器 1700 提供或者使用点到点系统(其将增加提供比赛片段的速率)提供。元数据和另外的元数据将由网络服务器 1700 提供。

[0201] 参考图 19A 和 19B 说明查找和观看其他用户的观看体验的方法。

[0202] 参考图 19A,显示器 1900 具有文本搜索框 1905。这允许搜索网络服务器 1700 上存储的自由文本元数据和元数据标签。在图 19A 中所示的示例中,对于“NUFC 和 MUFC”之

间的突出片段执行搜索。如根据图 19A 将理解的,比赛数据 1910 以时间顺序被返回。换言之,最近的比赛位于接近列表顶部,比较旧的比赛位于接近屏幕的底部。

[0203] 对于搜索的结果,网络服务器 1700 可使用用户帐户中提供的信息(例如喜欢的足球队或喜欢的选手)来返回最相关的结果,而不需用户执行搜索。例如,如果用户是纽卡斯尔联队足球俱乐部的爱好者,那么将在家庭屏幕上放置最新的纽卡斯尔联队足球比赛。类似地,如果用户表示他们是法布雷加斯的爱好者,那么将在家庭屏幕上放置包括元数据标签“法布雷加斯”的最新剪辑。

[0204] 邻近比赛数据 1910 的是用户数据 1915。这示出了已载了用于比赛的突出包的每个用户的用户名。邻近用户数据 1915 的是用户评级数据 1920。这给出了观看用户数据 1915 所识别的用户创建的其它比赛的突出包的其他用户所给出的平均分数。如果用户点击“评论”超链接,那么用户的评论也是可访问的。为了帮助用户选择要选择哪个其他用户的突出包,最流行的用户处于列表顶部,并且最不流行的位于列表底部。

[0205] 邻近用户评级数据 1920 的是比赛评级数据评级 1925。这提供了对用于该比赛的特定突出包的用户反馈。该类型的信息是有用的,因为通常执行卓越突出包的用户可产生用于该比赛的特别不好的突出包。作为候选,通常产生中等突出包的用户可执行用于该比赛的特别好的突出包。

[0206] 为了提供用户灵活性,可取决于用户偏好来改变每一列数据的排序。

[0207] 在用户选择了特定突出包之后,初始比赛被下载和本地存储在用户设备 200A 内。另外下载的(从网络服务器 1700)是用于显示产生突出包的其他用户所体验的视场的元数据和其他用户生成的任何其它元数据。因为元数据比它表示的数据更小,所以与元数据相关联的下载速度和存储需求相比于下载突出剪辑更小。

[0208] 参考图 19B,屏幕 1900 具有示出创建突出包的其他用户所体验的视场的视场区域 1930。这根据元数据和初始片段创建。事件列表区域 1935 也在显示器 1900 上。该列表对应于图 18 中的事件列表 1830。根据另外的元数据创建评注观看区域 1940。这显示了要显示给用户的最后的帧和其他用户添加的评注。例如,如果其他用户用标记突出了特定冲突,这将被放置在评注观看区域 1940 中。提供标准的视频控制按钮组 1945,例如加速或减慢视场 1930 中显示的视频。位于视频控制按钮 1945 邻近的下一个事件按钮 1950 允许用户略过下一个事件。下一个事件是用户特别感兴趣的一个片段。用户可通过下一个事件选择按钮 1955 来选择特别感兴趣的下一个事件。在该实施例中,下一个事件包括下一个进球、下一个任意球、下一个黄牌或红牌或者下一个角球。用户可通过围绕适当的下一个事件符号的框容易地看到选择了哪个事件。在实施例中,下一个事件突出框 1960 围绕下一个进球。

[0209] 用户还能够改进另一个用户的特定突出包以便例如改善虚拟摄像机定位,编辑突出包的持续时间或添加另外的评注。这可由用户在创建可编辑的突出包时许可。此外,关于特定突出包的另外评注可由其他用户添加。这使得另外的用户能够对特定突出包进行评论。例如,用户可添加识别也许由突出包的创建者遗漏的特定内容特征的评论。所以在足球比赛的上下文中,不同的用户可识别可能未由其他用户所注意的球场上选手的定位。这可导致用户组(都在观看相同的突出包)之间的实时通信。

[0210] 可以对于具有 1920×1080 像素分辨率的显示器上示出的视频输入突出包的作者所应用的评注。然而,其他用户可在具有更小显示器的便携式手持设备上观看所评注的视

频。例如,手持设备可以是具有 320×240 像素分辨率的显示器的设备。此外,便携式设备上的其他用户可对于在较大显示器上创建的突出包应用另外的评注。在实施例中,为了解决此,元数据可连同突出包一起存储以表示在其上创建评注的显示器的大小。因此,显示器上评注的像素位置可被换算或调节为确保当在不同大小的显示器上再现评注时,评注被放置在显示器的正确区域上。

[0211] 作为示例,如果在具有 1920×1080 像素的显示器上生成突出包并且具有 240×90 像素大小的评注被输入到具有 (430, 210) 的左上像素位置的突出包上的帧内,那么生成定义评注、评注的大小和像素位置、其上生成评注的显示器的大小的元数据。这和包一起存储。

[0212] 当另一个用户希望在便携式设备上观看突出包时,检索描述评注的元数据。便携式设备知道评注的大小和像素位置以及在其上创建评注的显示器的大小。因此,便携式设备换算评注以使得评注的大小对于显示器是正确的。特定地,便携式设备上评注的大小是 40×20 像素。当为便携式设备显示器换算时的评注位置将是像素 (71.6, 46.6)。为了选择正确的像素位置,将在像素位置 (72, 47) 处放置评注。这是简单的对最接近像素的向上舍入。然而,可想象当换算导致小数像素位置时进行像素选择的其它方法。

[0213] 如果便携式设备的用户创建了具有在像素位置 (140, 103) 处具有 38×28 的像素大小的另外评注,那么创建描述评注和在其上创建该评注的显示器的大小的元数据。

[0214] 因此,如果初始作者再次观看包,那么便携式设备的用户创建的评注将被向上换算为在像素位置 (840, 463.5) 处具有 228×126 大小的评注。此外,为了正确地在初始作者的显示器上显示评注,评注将被放置在像素位置 (840, 464) 处。

[0215] 最后,用户可以使用框 1970 来评定特定突出包的质量。用户选择适当标记(在该情况下为 5 中的一个),并且点击框 1970。然后该值被传送到网络服务器 1700,在那它与其他用户和该特定突出包相关联地存储。

[0216] 通过向网络服务器 1700 发送元数据和另外的元数据而不是视频剪辑,减少了经由网络发送的数据量。实际上,当经由不同方法向用户提供初始视频片段时,可进一步减少网络服务器 1700 处理的数据量。例如,用户可使用点到点系统或通过邮件等在记录介质上接收初始视频片段。

[0217] 创建突出包的用户,或观看突出包的用户可为此付费。费用可以每次观看时支付,或者可以是按月或按年的预订服务。

[0218] 尽管前面参考用户设备 200A 进行描述,但是可同样适用用户装置 315A。

[0219] 客户端设备上的增强现实

[0220] 图 20 示出了其中举行足球比赛的体育场 2000 的平面图。足球场 2020 位于体育场 2000 内并且比赛正由摄影机系统 2010 拍摄。摄像机系统 2010 包括摄像机布置 130、图像处理设备 135 和服务器 110。摄像机系统包括全球定位系统 (GPS) 传感器(未示出)、高度传感器和倾斜传感器。GPS 系统提供摄像机系统 2010 的坐标位置,高度传感器提供对摄像机系统高度的识别并且倾斜传感器提供应用到摄像机系统 2010 的倾斜量的表示。GPS 系统、高度和倾斜传感器是已知的并且所以不在以下进行描述。

[0221] 在球场上的是第一选手 2040、第二选手 2050、第三选手 2055、第四选手 2060、第五选手 2065、第六选手 2070 和第七选手 2075。还提供了由选手 2040 控制的球 2045。摄像机

系统 2010 正捕获足球比赛,如前面实施例中所述。

[0222] 位于人群中的是正在通过他的移动电话机 2100 观看比赛的观众 2030,在实施例中移动电话机 2100 是 Sony Ericsson Mobilem Communications 制造的 Xperia X10 电话机。将参考图 21 描述移动电话机 2100。移动电话机 2100 包括可使用 3G 或 LTE 网络标准经由蜂窝网络通信的通信接口 2160。实际上,通信接口 2160 可能使用任何网络标准(例如 WiFi 或蓝牙等)进行通信。还提供了存储器 2140。在存储器上存储了数据。存储器例如可以是固态存储器。存储器还存储计算机可读指令并且因此存储器 2140 是存储计算机程序的存储介质。另外,存储器 2140 存储其它类型的数据,例如元数据,或用户专有的数据,以及与移动电话机 2100 中的摄像机 2120 的光学畸变相关的数据。移动电话机 2100 配备有向用户显示信息的显示器 2110。

[0223] 摄像机 2120 被布置为捕获可存储在存储器 2140 或可直接显示到显示器 2110(存储或不存储在存储器 2140 中)的图像。还提供了为移动电话机 2100 提供全球唯一位置的 GPS 传感器 2130。此外,还提供了表示应用到移动电话机 2100 的倾斜和电话机 2100 的高度的倾斜和高度传感器 2155。另外,用于观看场景的摄像机 2120 的焦距由电话机 2100 确定。

[0224] 还提供的是控制每一个前述组件并且被布置为在其上运行计算机软件的处理器的处理器 2150。在本实施例中处理器 2150 的示例是 Qualcomm® 制造的 SnapDragon Processor。处理器 2150 使用数据总线 2155 连接到每个组件。

[0225] 图 22 示出了如用户 2030 所见的移动电话机 2100。用户 2030 正在持有移动电话机 2100 以使得他可容易地看到显示器 2110。用户在比赛时正指向移动电话机 2100 的摄像机 2120。显示器 2110 在移动电话机 2100 上示出了摄像机 2120 所捕获的比赛的实况图像。这在图 22 中示出,其中第一到第七选手中的每一个示出在球场 2020 上。另外,位于选手 2040 到 2075 中每一个上面的是每个选手的姓名。每个选手的姓名由处理器 2150 置于显示器 2110 上。每个选手的姓名从摄像机系统 2010 中生成的选手元数据提供。这将在随后参考图 23 来说明。除了每个选手上面的姓名,在显示器 2110 上提供示出比赛时间的时钟 2220,并且还显示当前比赛分数 2225。

[0226] 在实施例中,显示器 2110 是触摸屏,其允许用户 2030 通过按压显示器 2110 来向移动电话机 2100 发出命令。为了提供增强的用户能力,位于每个选手上面的姓名可由用户 2030 触摸来显示选手传记。选手传记可在比赛前存储在存储器 2140 中。作为候选或者另外地,通过按下选手上面的姓名可提供与选手相关的实时比赛统计数据。换言之,实时比赛统计数据提供选手得分的进球数量,选手完成的过人数量,并且因为摄像机系统 2010 使用选手跟踪,还提供选手奔跑的距离量。该信息可响应于用户触摸姓名而提供到电话机 2100。作为候选,该数据可经由网络被持续地更新并且存储在存储器 2140 中,以使得当用户触摸姓名时从存储器 2140 检索信息。这相比于经由网络请求信息更快。该信息由摄像机系统生成,如上参考图 9 所述。

[0227] 参考图 23,描述了在显示器 2110 上放置选手姓名的方法。移动电话机 2100 向摄像机系统 2010 注册。在注册处理期间,完成识别移动电话机 2100 的用户是否有资格访问信息的验证处理。例如,交换支付信息。这在步骤 S2310 中示出。

[0228] 如上所述,摄像机系统 2010 捕获比赛图像并且根据该捕获的图像,检测图像中每

个选手的位置并且确定选手的真实世界位置。为了实现此,摄像机系统 2010 使用图 14 中描述的技术来识别所检测的对象在球场上的位置。重要的是要注意,使用该技术的球场上选手的位置确定了选手相对于摄像机系统 2010 的位置。因此,当向摄像机系统 2010 提供它的 GPS 位置时,摄像机系统 2010 确定每个选手的 GPS 位置(或真实世界位置)。另外,因为每个选手的身份已知,所以还生成与选手相关联的元数据,例如选手姓名。这是步骤 S2320。

[0229] 真实世界的位置信息和元数据被发送到移动电话机 2100。这是步骤 S2330。应当注意,所检测的图像(例如足球),或者裁判员或助理裁判员也可被传送到移动电话机 2100。

[0230] 移动电话机 2100 接收与每个所检测的选手相关联的真实世界位置信息和所检测的球。移动电话机 2100 从 GPS 传感器检索识别移动电话机 2100 位置的 GPS 值。这是步骤 S2340。

[0231] 此外,从位于移动电话机 2100 内的高度和倾斜传感器检索高度和倾斜值。另外,确定电话机 2100 中摄像机 2120 的焦距。这是步骤 S2350。

[0232] 使用电话机 2100 的 GPS 位置、倾斜角和焦距,电话机 2100 确定使用摄像机 2120 捕获的球场区域。换言之,电话机 2100 确定摄像机所看到的真实世界位置的边界。这通过摄像机系统 2010 提供球场上基准点的真实世界位置而进一步便利。为了实现此,这些基准点被用于计算真实世界位置和球场平面的角度。使用电话机的 GPS 位置和它的倾斜角,计算表示真实世界中电话机的镜头指向的方向的三维向量。使用已知技术,可因此计算该向量把球场平面一分为二的真实世界点。该真实世界点是摄像机视场的中心。为了确定视场的范围,必须首先计算水平和垂直视场的角度。这些使用已知技术根据传感器大小和镜头的焦距来计算。

[0233] 作为示例,使用例如以下的公式:

[0234] $FOV(\text{horizontal}) = 2 * \arctan(\text{SensorWidth} / (\text{FocalLength} * 2))$

[0235] $FOV(\text{vertical}) = 2 * \arctan(\text{SensorHeight} / (\text{FocalLength} * 2))$

[0236] 然后,这些角度被用于旋转表示电话机镜头指向的方向的向量,以使得它通过摄像机图像的角落之一。此外,使用已知技术,计算该向量把球场平面一分为二的真实世界点。该真实世界点是摄像机视场的角落。该技术然后被重复用于摄像机视场的全部四个角落以便确定摄像机所看见的真实世界位置的边界。当移动电话机 2100 提供球场上选手的真实世界位置以及球场上的真实世界关键点时,电话机 2100 确定最可能在摄像机 2120 观看的图像中的什么位置看到选手和关键点。然后它在图像内的这些位置定位评注。

[0237] 在候选实施例中,为了增加评注放置的精确性,移动电话机 2100 然后对所捕获图像执行图像检测以便检测图像内的任何对象。这是步骤 S2360。因为移动电话机 2100 知道摄像机所看到的真实世界位置的边界,所以电话机 2100 识别图像内所检测的每个对象的真实世界位置。因此,通过比较电话机 2100 所捕获的每个对象的真实世界位置和摄像机系统 2010 所捕获的每个对象的真实世界位置,可以确定移动电话机 2100 捕获的图像内的哪个对象对应于哪个所检测的选手。摄像机系统 2010 提供的评注(作为元数据被提供)被应用到图像内的正确对象。这是步骤 S2370。这里应当注意,为了改善评注处理的精确性,考虑了移动电话机 2100 中摄像机的光学畸变。例如,如果摄像机内的光学畸变使得光线向左弯曲 5 个像素通过镜头,那么所检测对象的真实世界位置将不同于摄像机所捕获的。因此,

可对所捕获图像内的所检测位置应用校正以便纠正该错误。光学畸变存储在存储器 2140 中,并且当制造电话机时生成。然后处理结束(步骤 S2380)。

[0238] 使用该信息,与移动电话机摄像机的当前焦距组合,移动电话机可确定体育场的哪部分将出现在它的视场中并且因此计算摄像机系统所检测的任何选手应当出现在它屏幕上的什么位置。

[0239] 在实施例中,可使用块匹配技术等来执行移动电话机 2100 所捕获的图像中的对象检测。这可改善评注被放置在移动电话机 2100 的显示器上的精确性。

[0240] 摄像机系统可向移动电话机 2100 发送对象的表示(例如每个选手的裁剪)。移动电话机 2100 所检测的对象可与从摄像机系统 2010 接收的相比较。该改善了检测技术的质量。

[0241] 为了减少执行该对象比较所需的处理器功率,实施例中的移动电话机 2100 比较来自摄像机系统的已知基准位置和它的视场内的相应基准位置。例如,从摄像机系统 2010 接收的任何球场标记可与移动电话机 2100 捕获的图像中的任何所检测球场标记比较。比较球场标记是有用的,因为它们场景中是静态的,并且因此标记的位置将保持不变。如果不匹配,或者匹配概率低于比如 98% 的阈值,那么从摄像机系统 2010 接收的所检测球与移动电话机 2100 所检测的其它对象比较。因为用户可能专注于球,最希望的是移动电话机 2100 捕获的任何图像将包括球。此外,因为球是图像中的独特对象,检测该对象将是更容易的并且因此减少了移动电话机 2100 内的处理能力。

[0242] 如果没有球的匹配或者匹配的概率低于阈值,那么移动电话机 2100 所检测的对象再次与从摄像机系统 2010 发送的其它对象比较。当实现了确定的比较时,移动电话机 2100 所检测的对象位置与通过变换所计算的位置比较。这确立了校正值。然后校正值被应用于每个变换的位置值。该校正的变换位置值标识向其提供元数据(例如选手姓名)的选手的位置。移动电话机 2100 将姓名应用到最接近校正的变换位置值的所检测对象。特定地,移动电话机 2100 在所检测对象上面插入姓名。这改善了放置评注的精确性。为了提供增强的用户体验,比赛时间和比赛分数被应用到显示器的特定区域,例如在显示器的角落中。这些区域通常不是用户的焦点,所以将不会模糊动作。

[0243] 可设想增强现实实施例将是在移动电话机 2100 上运行的计算机程序。例如,实施例可以是所谓的“应用”。为了帮助用户,当初始化应用时,移动电话机 2100 将自动启动 GPS 传感器和高度和倾斜传感器。此外,如在比赛期间期望的,用户可能希望不与移动电话机 2100 交互。通常,为了节省电池能量,显示器将在一段不活动周期之后关闭。然而,将是不方便的。因此,应用将禁止自动关闭显示器。

[0244] 尽管前面描述了根据所捕获对象确定球场上不同对象的位置,但是本发明不限于此。例如,每个选手可以携带使用 GPS 系统提供球场上选手位置的设备。此外,类似的设备可放置在球中。这将减少系统的计算消耗,因为该信息将被自动提供而不需要计算位置。

[0245] 尽管这里参考附图详细描述了本发明的说明性实施例,应当理解,本发明不限于那些精确的实施例,并且可由本领域技术人员在其中实现各种改变和修改而不会偏离如所附权利要求所定义的本发明的范围和精神。

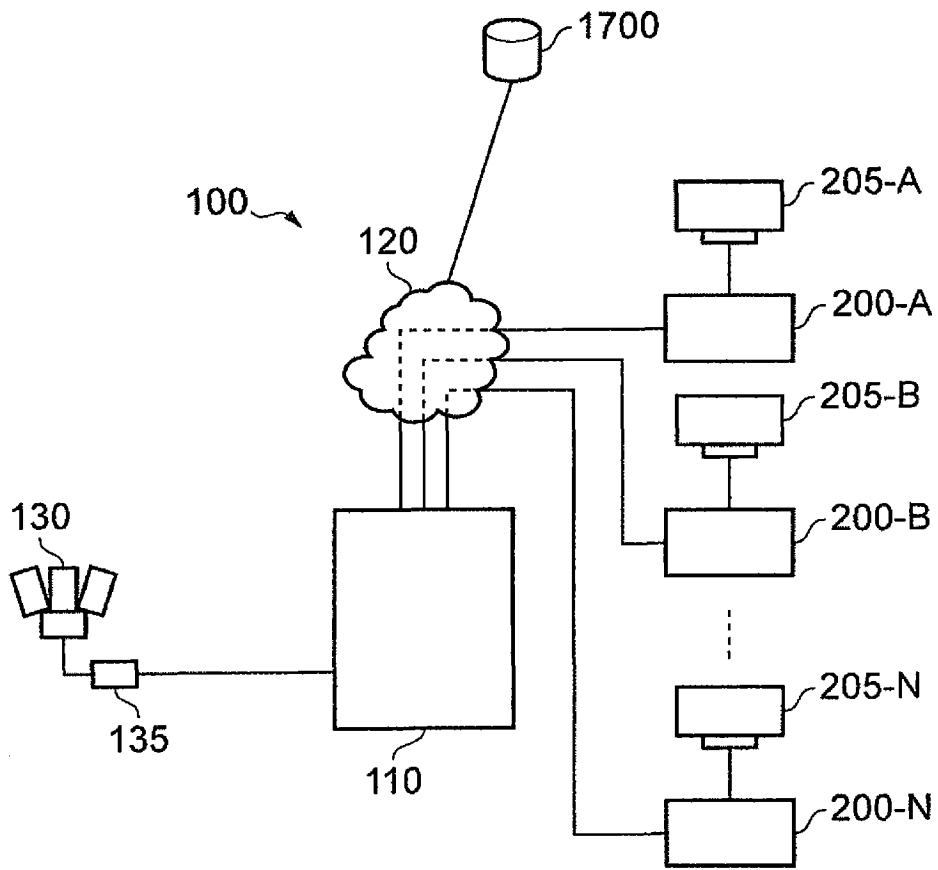


图 1

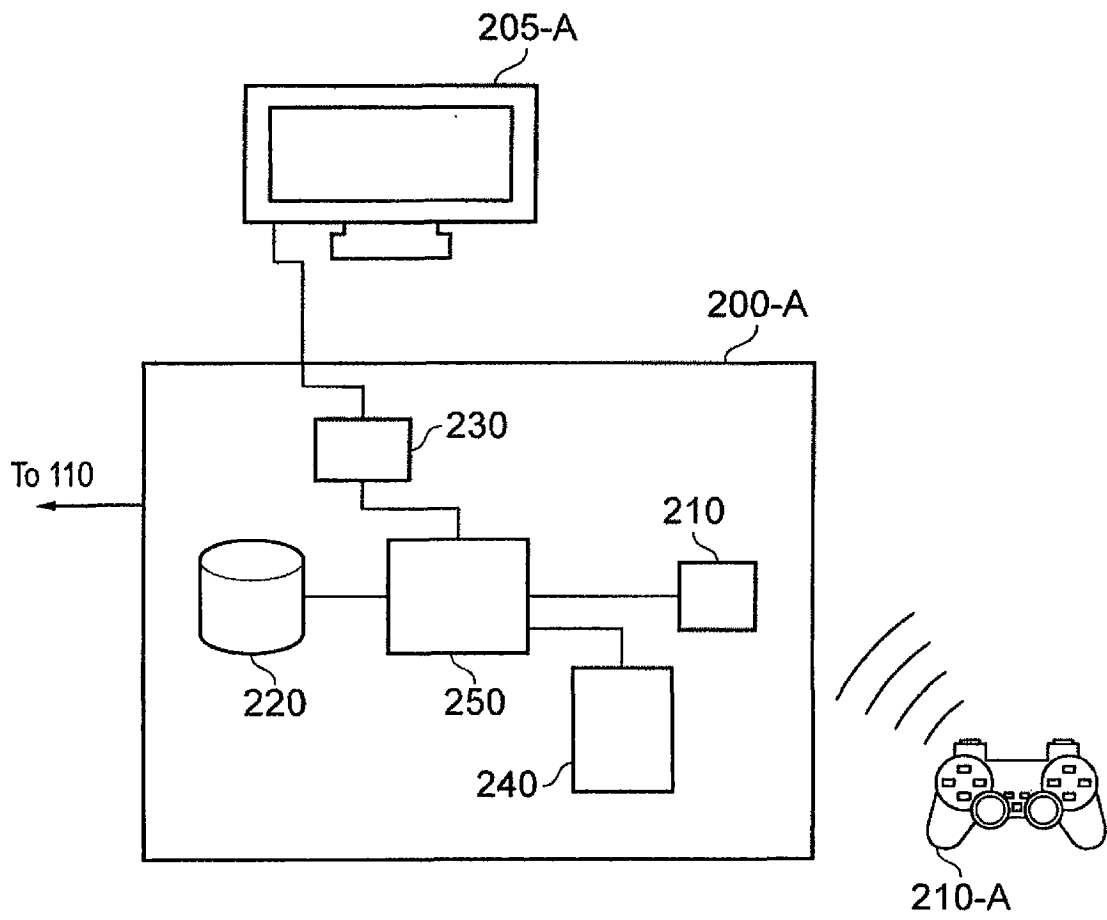


图 2

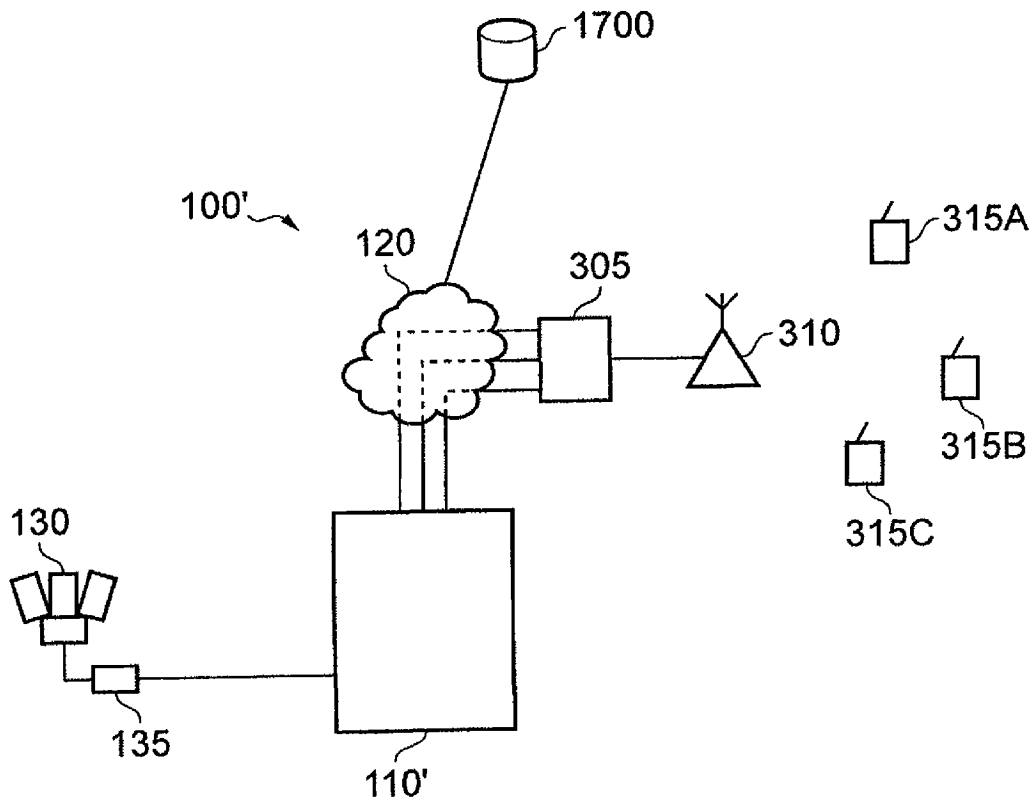


图 3

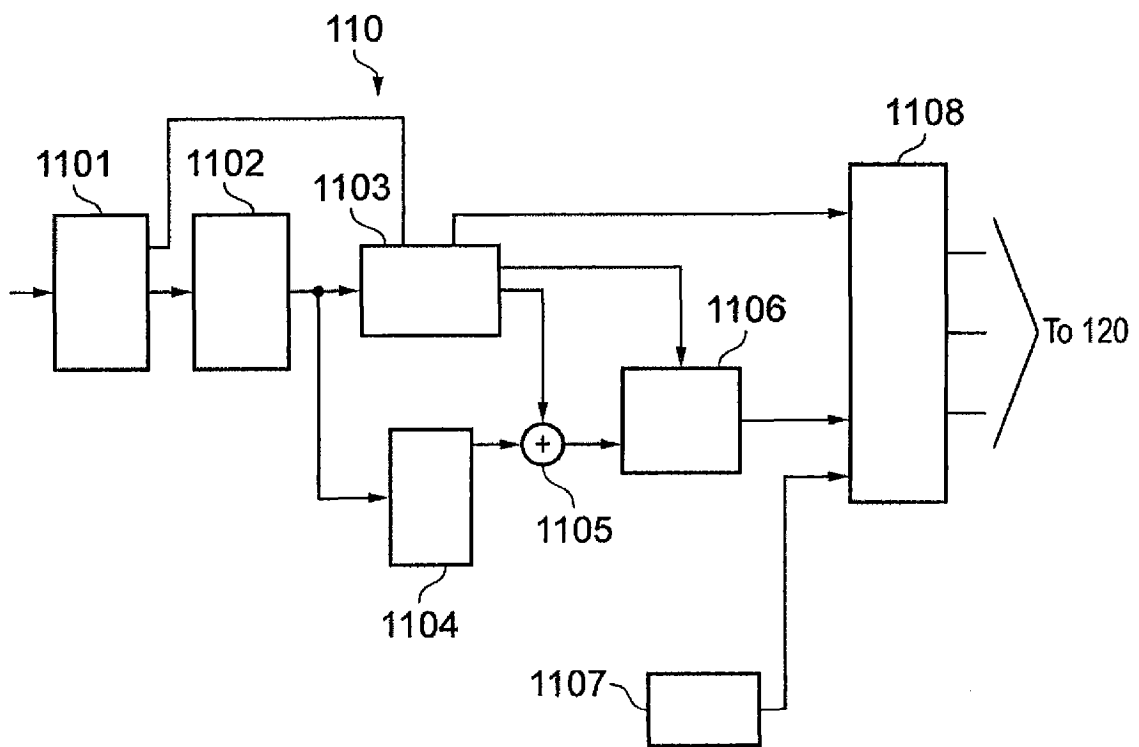


图 4A

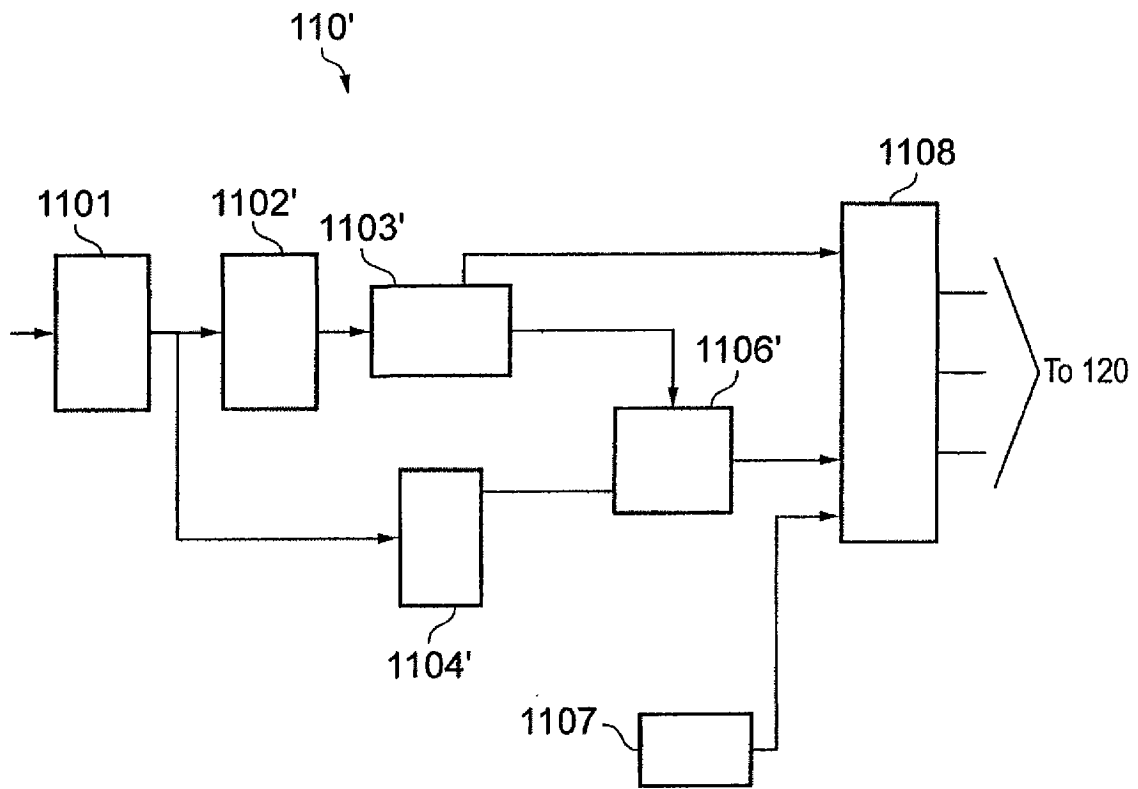


图 4B

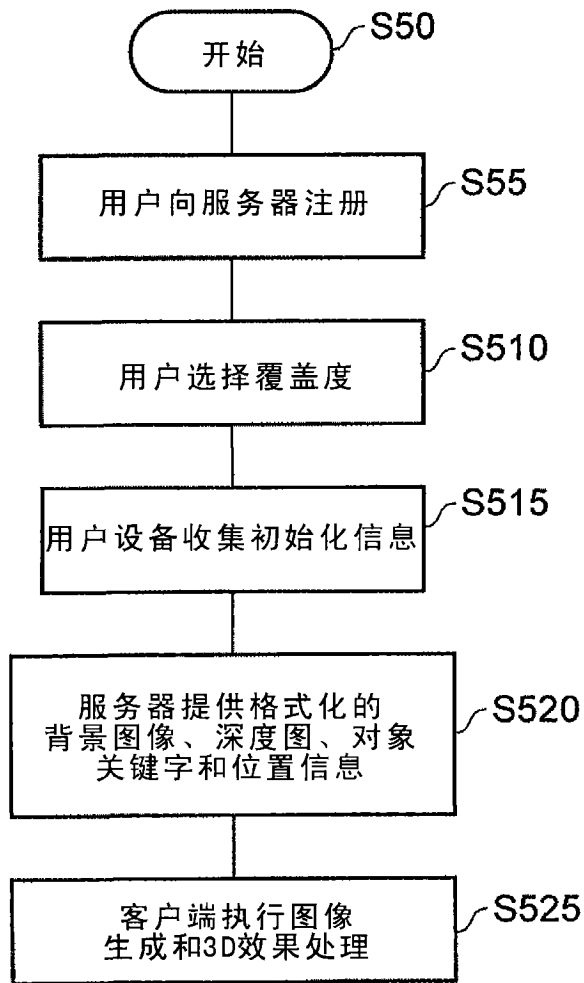


图 5

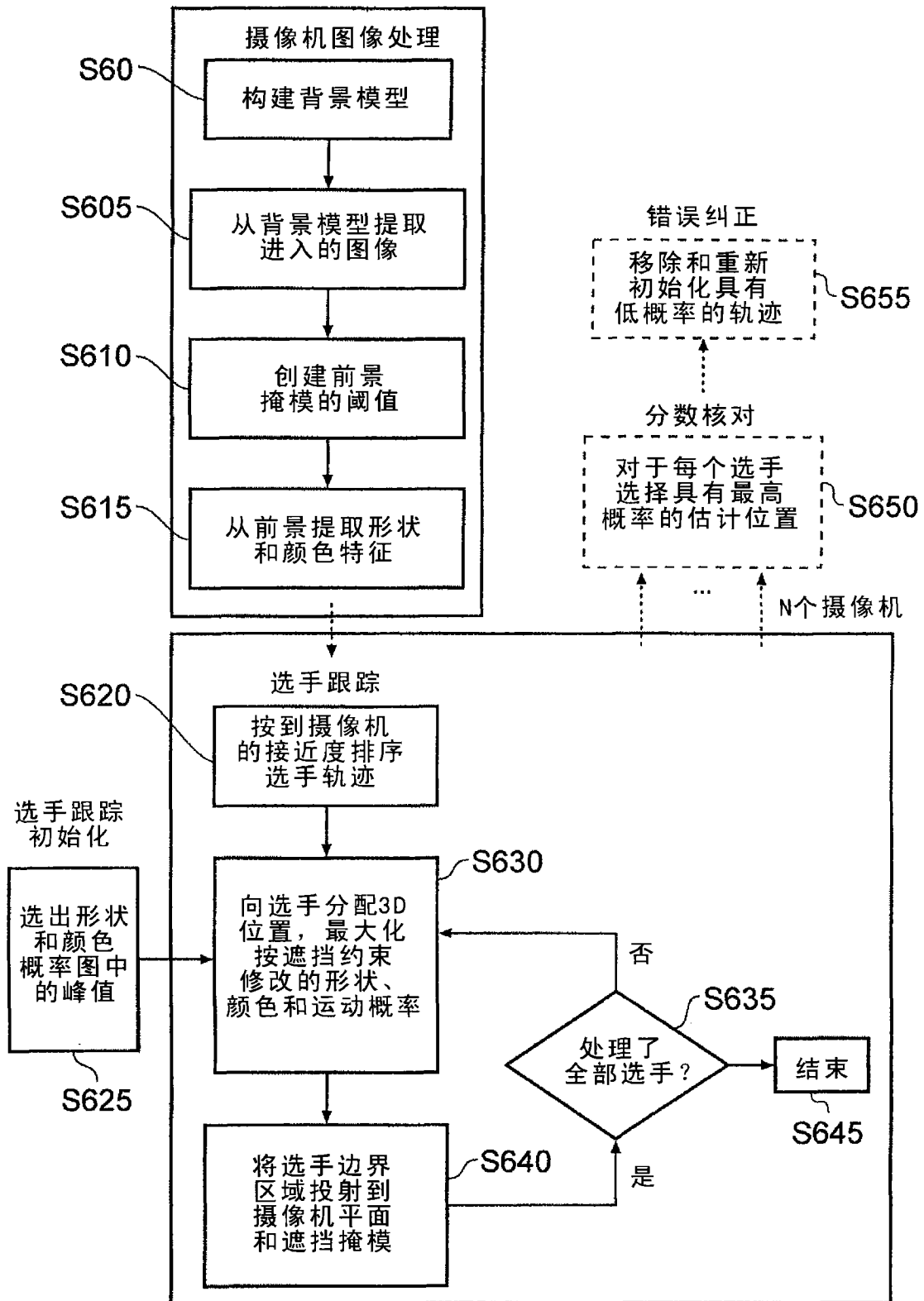


图 6

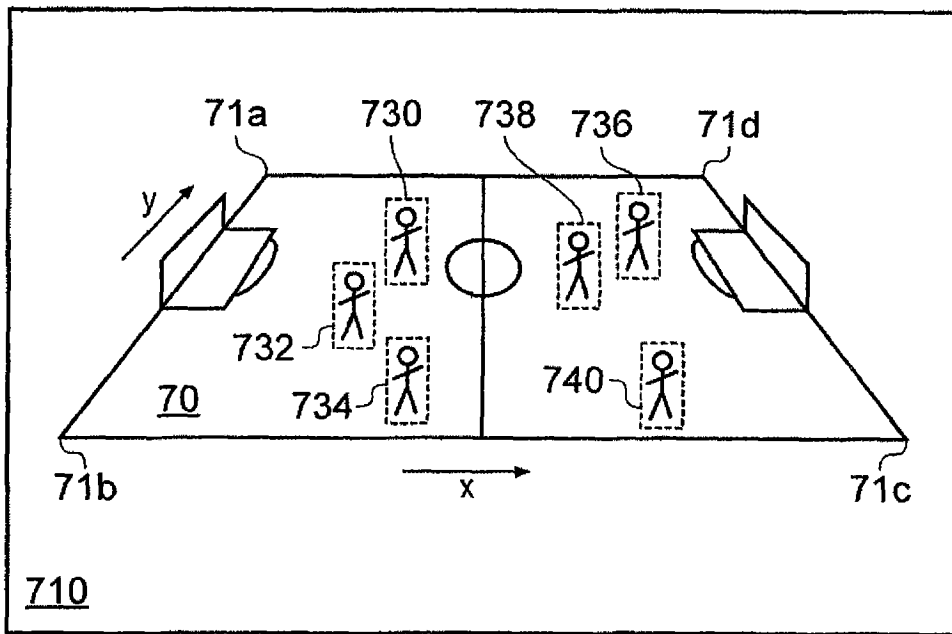


图 7A

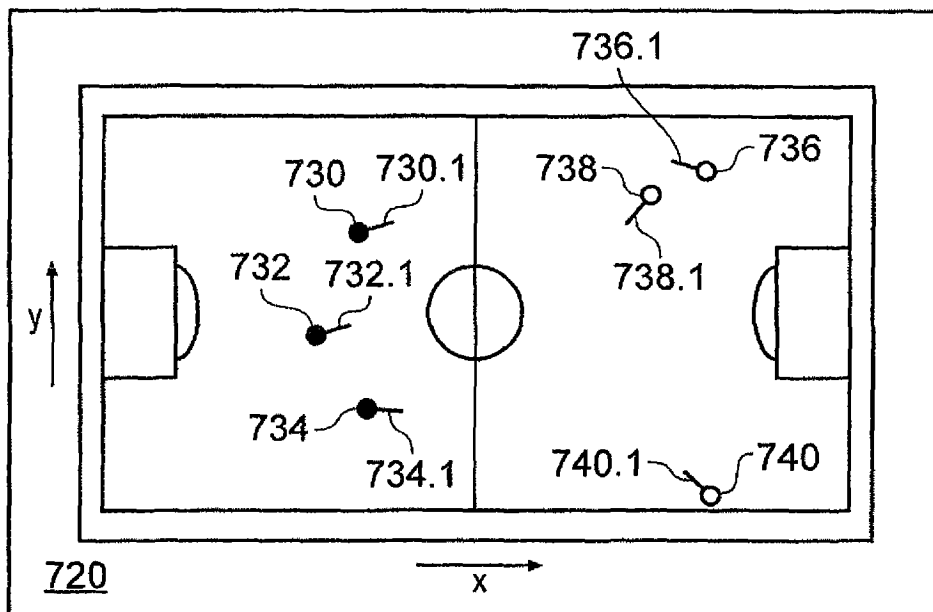


图 7B

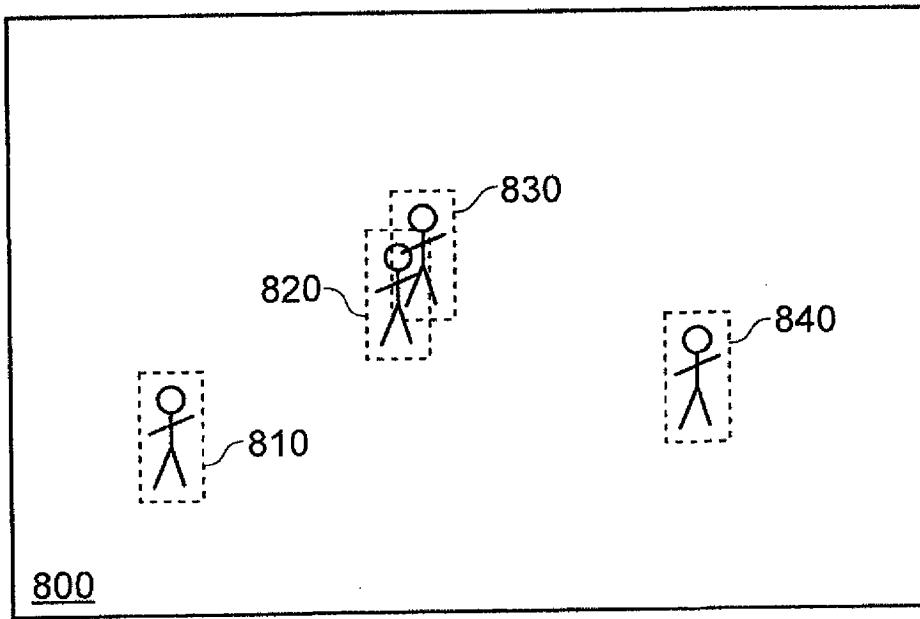


图 8

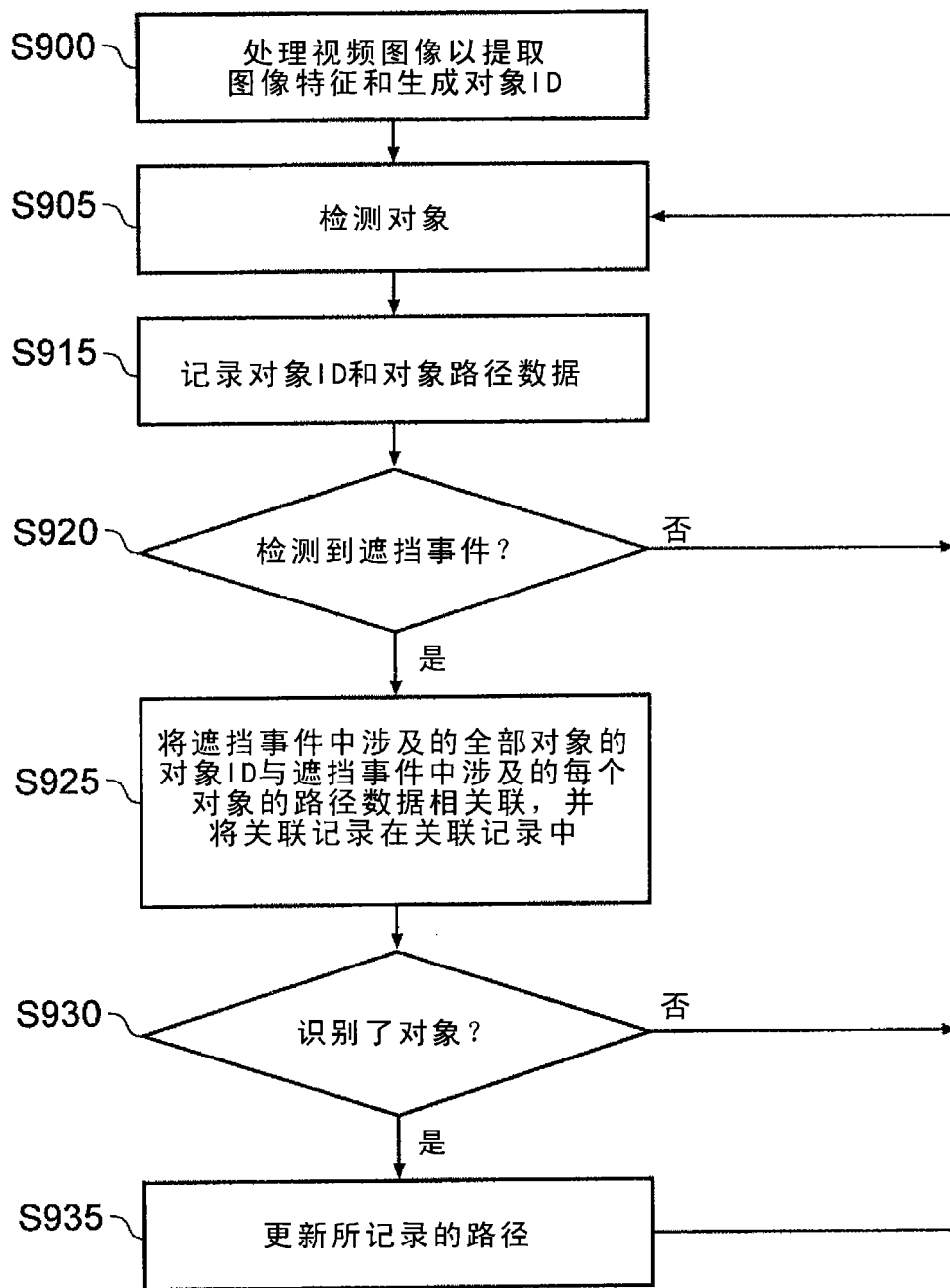


图 9

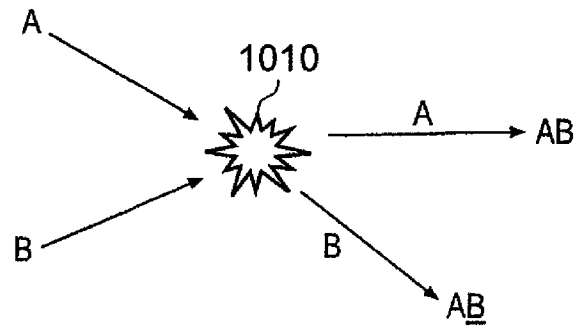


图 10A

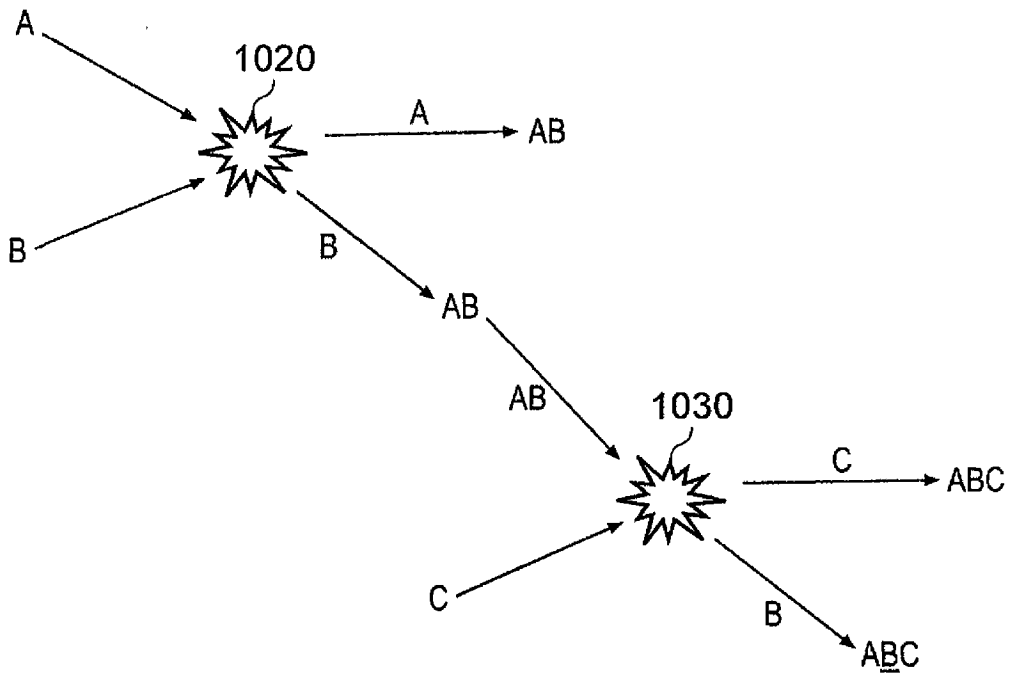


图 10B

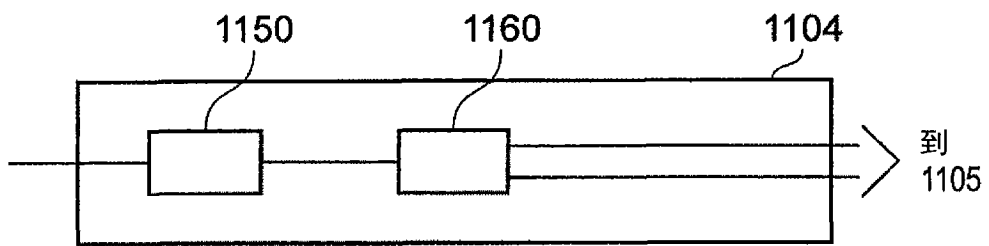


图 11

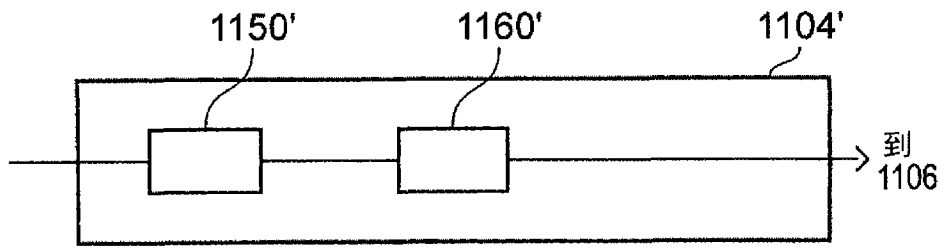


图 12

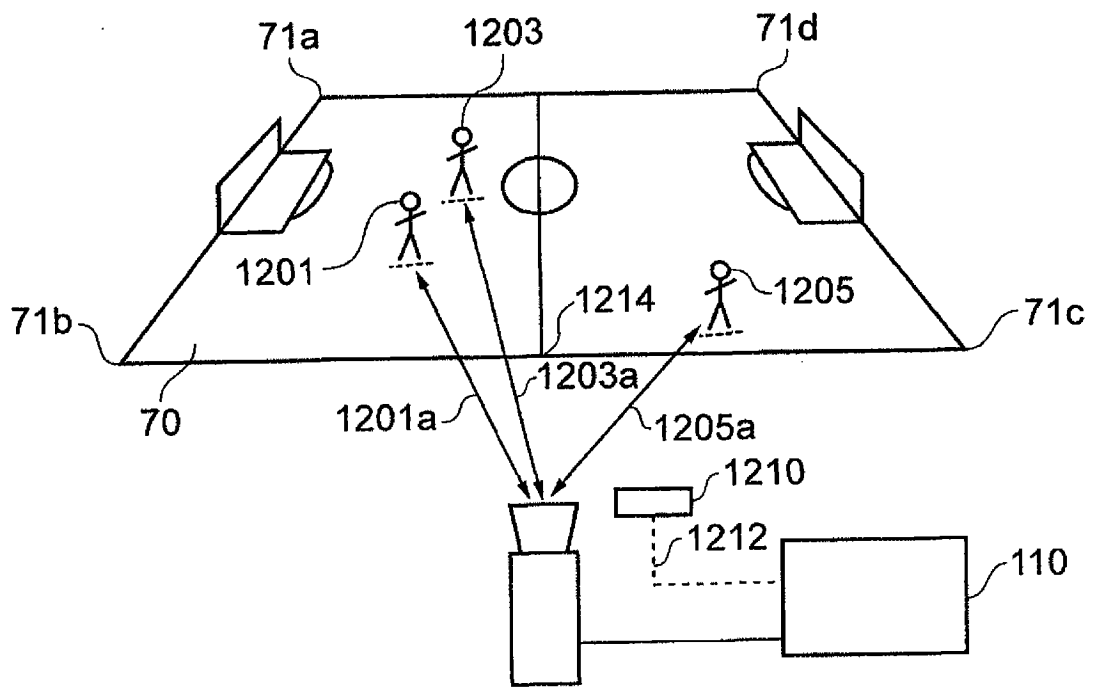


图 13

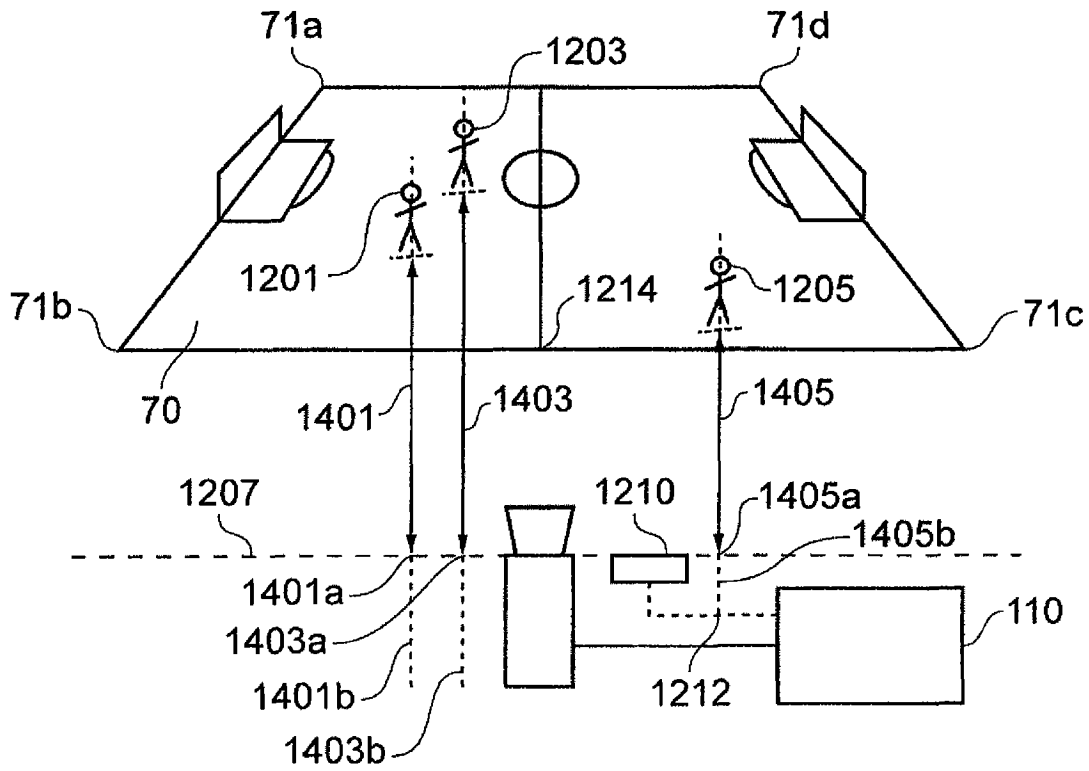


图 14

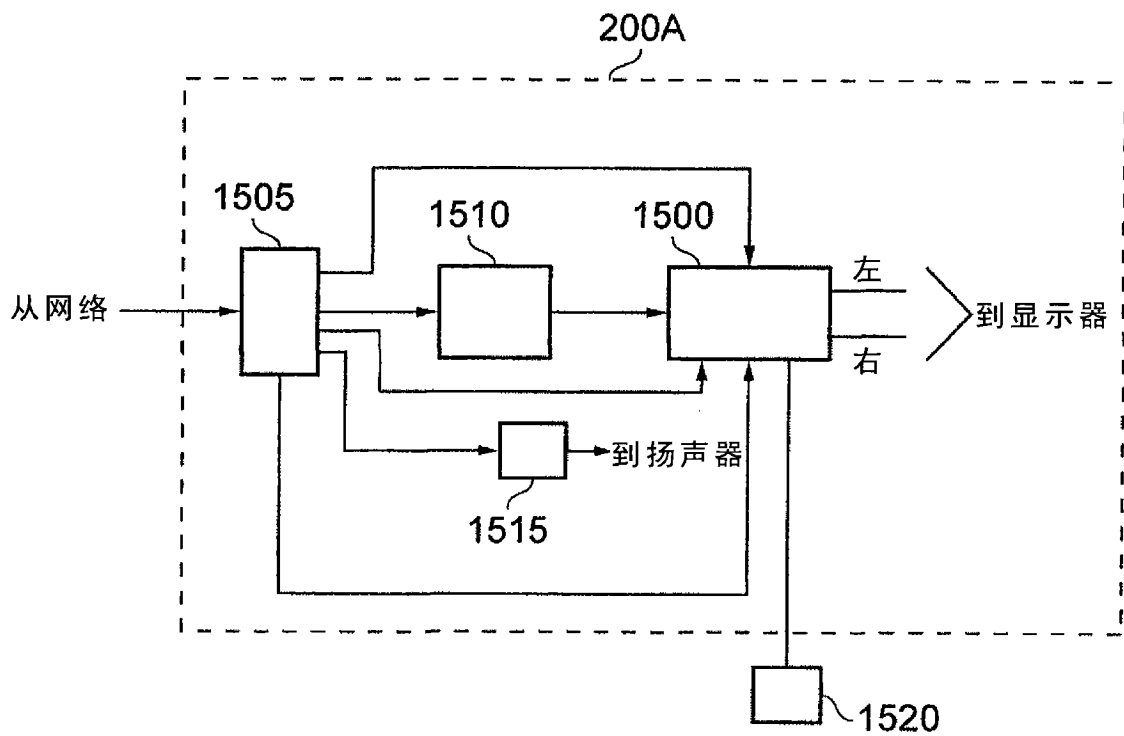


图 15A

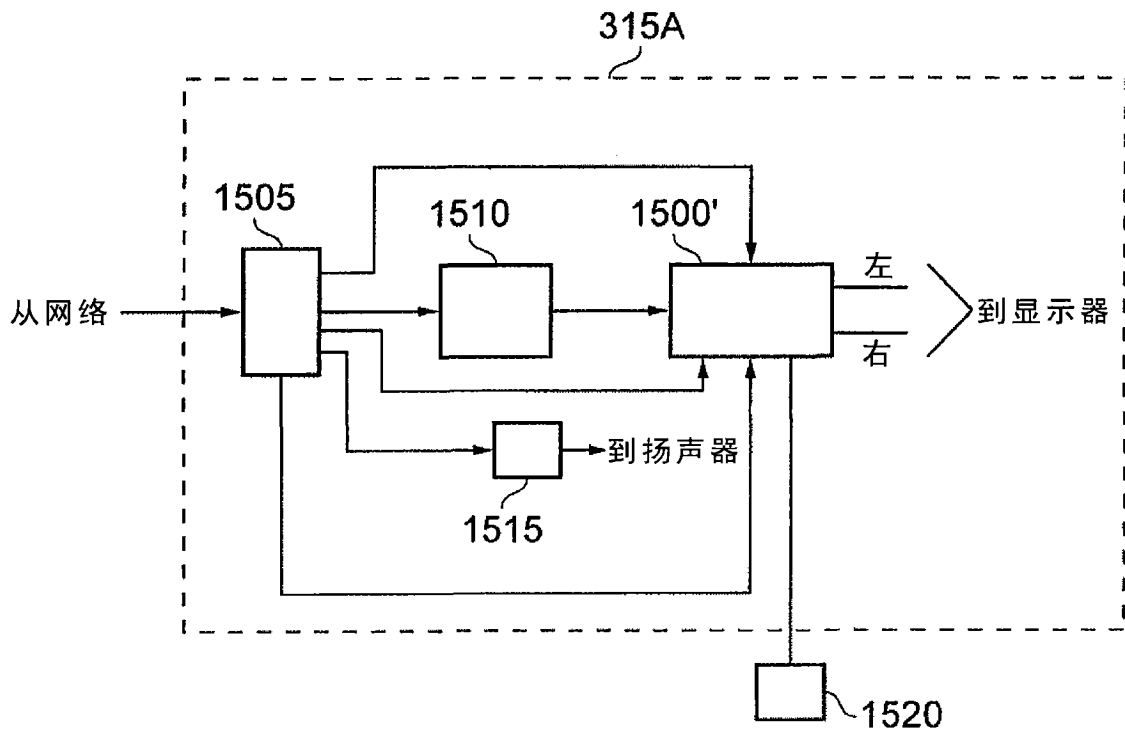


图 15B

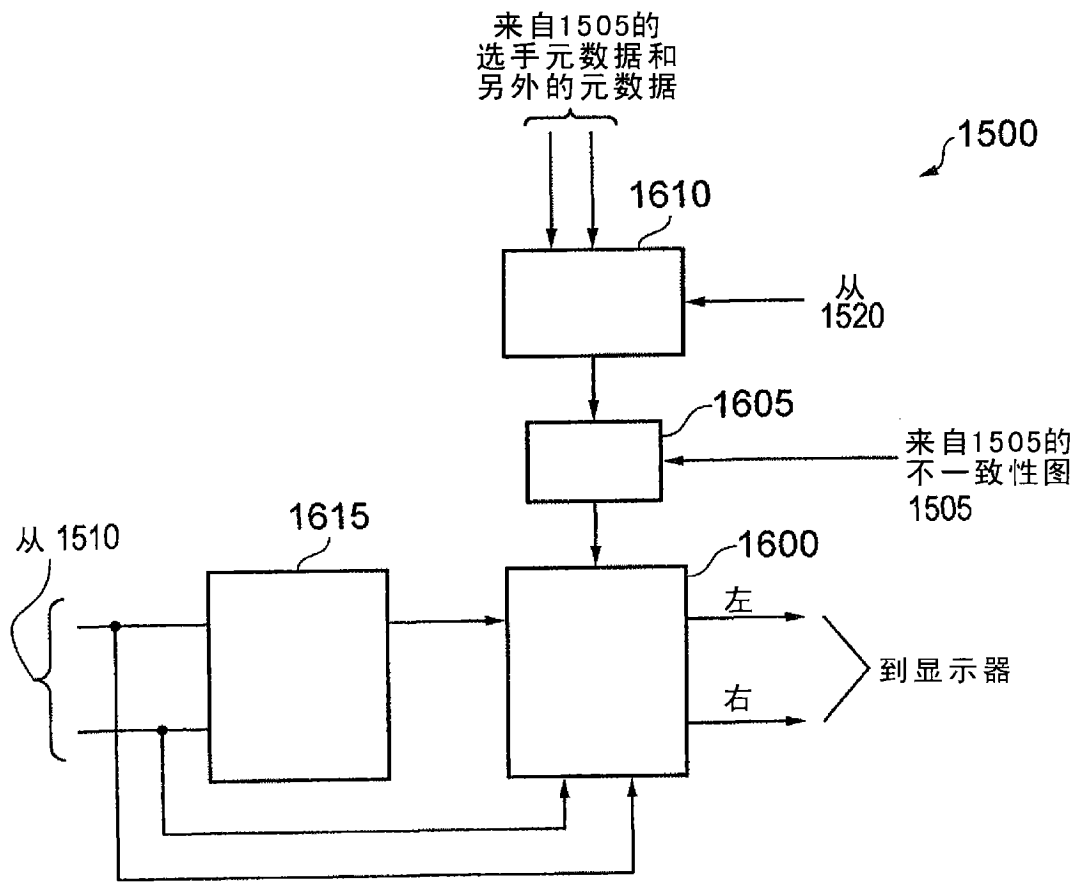


图 16A

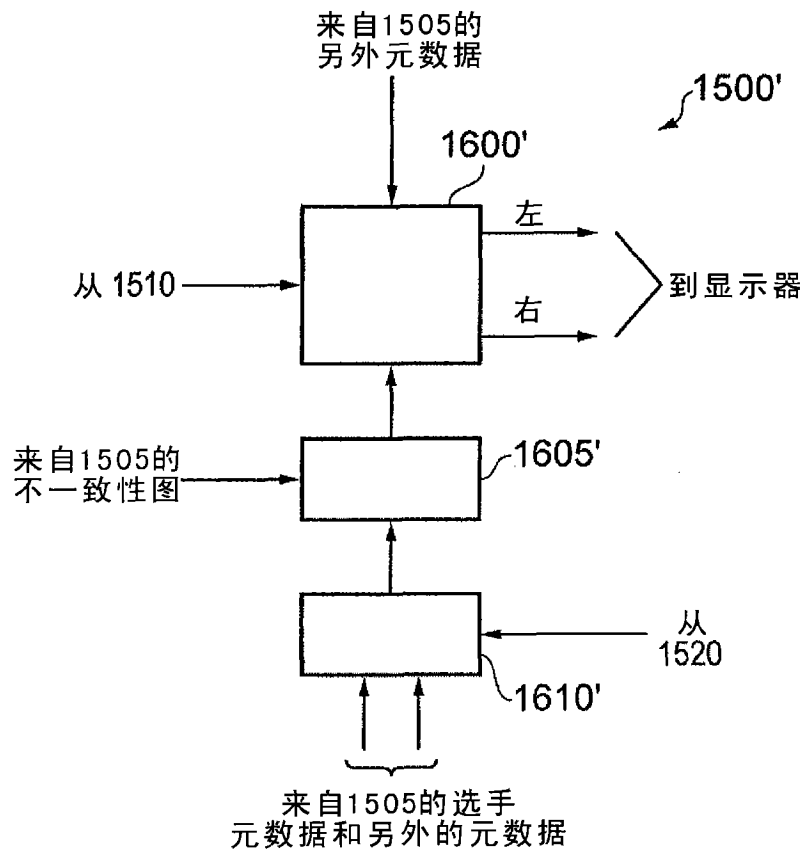


图 16B

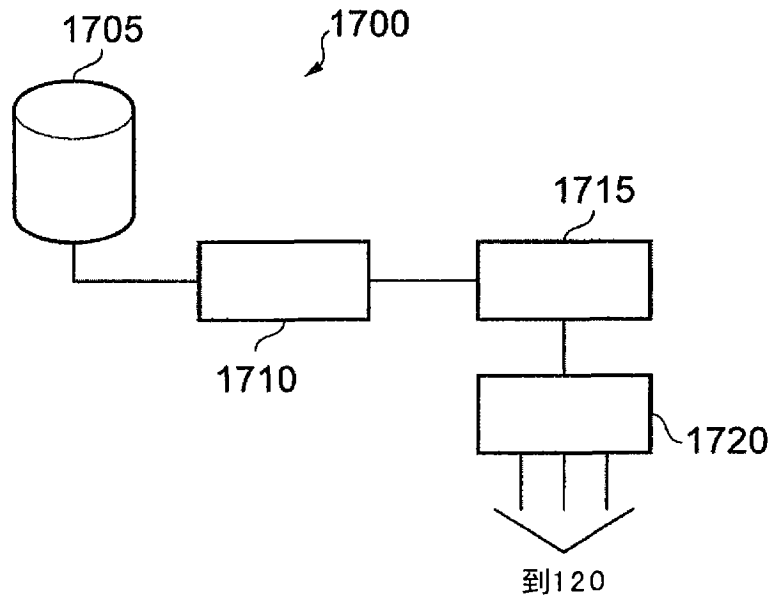


图 17

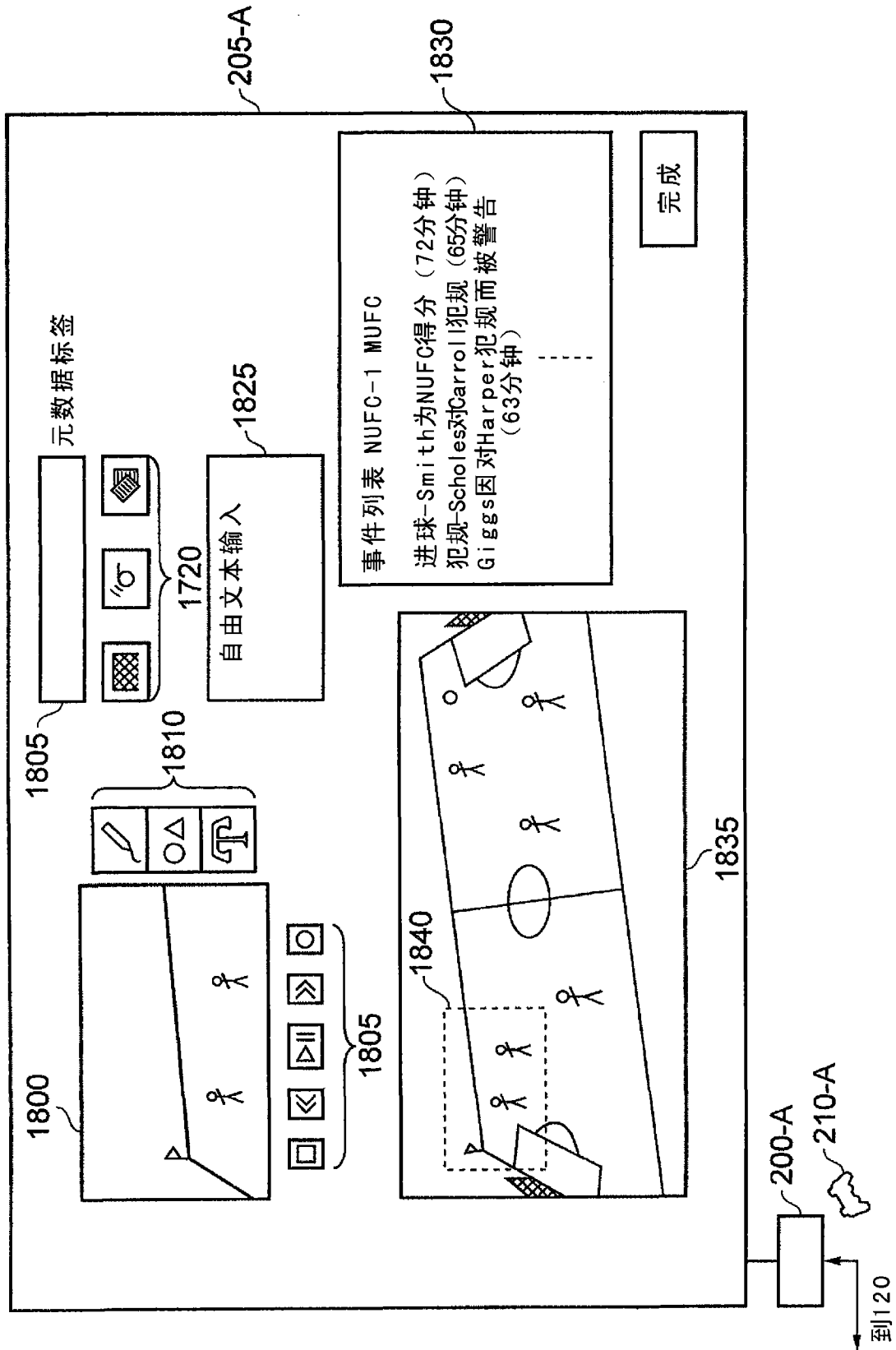


图 18

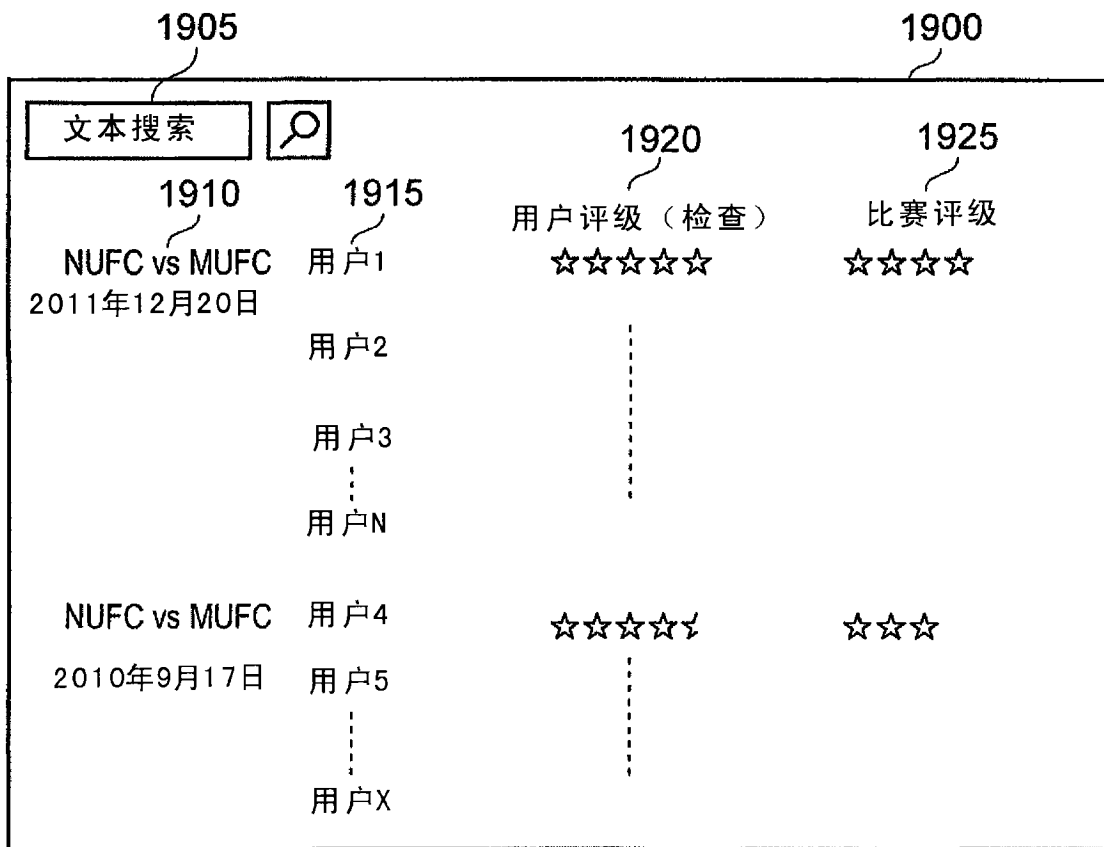


图 19A

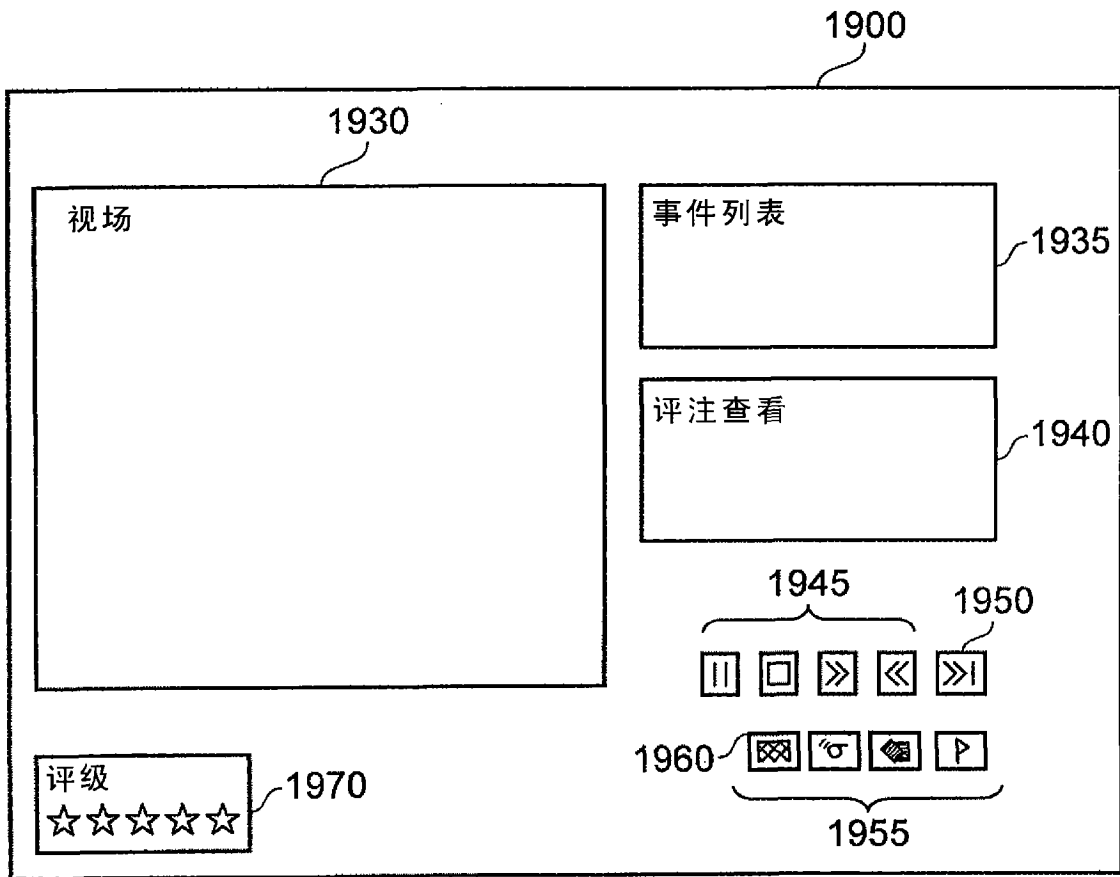


图 19B

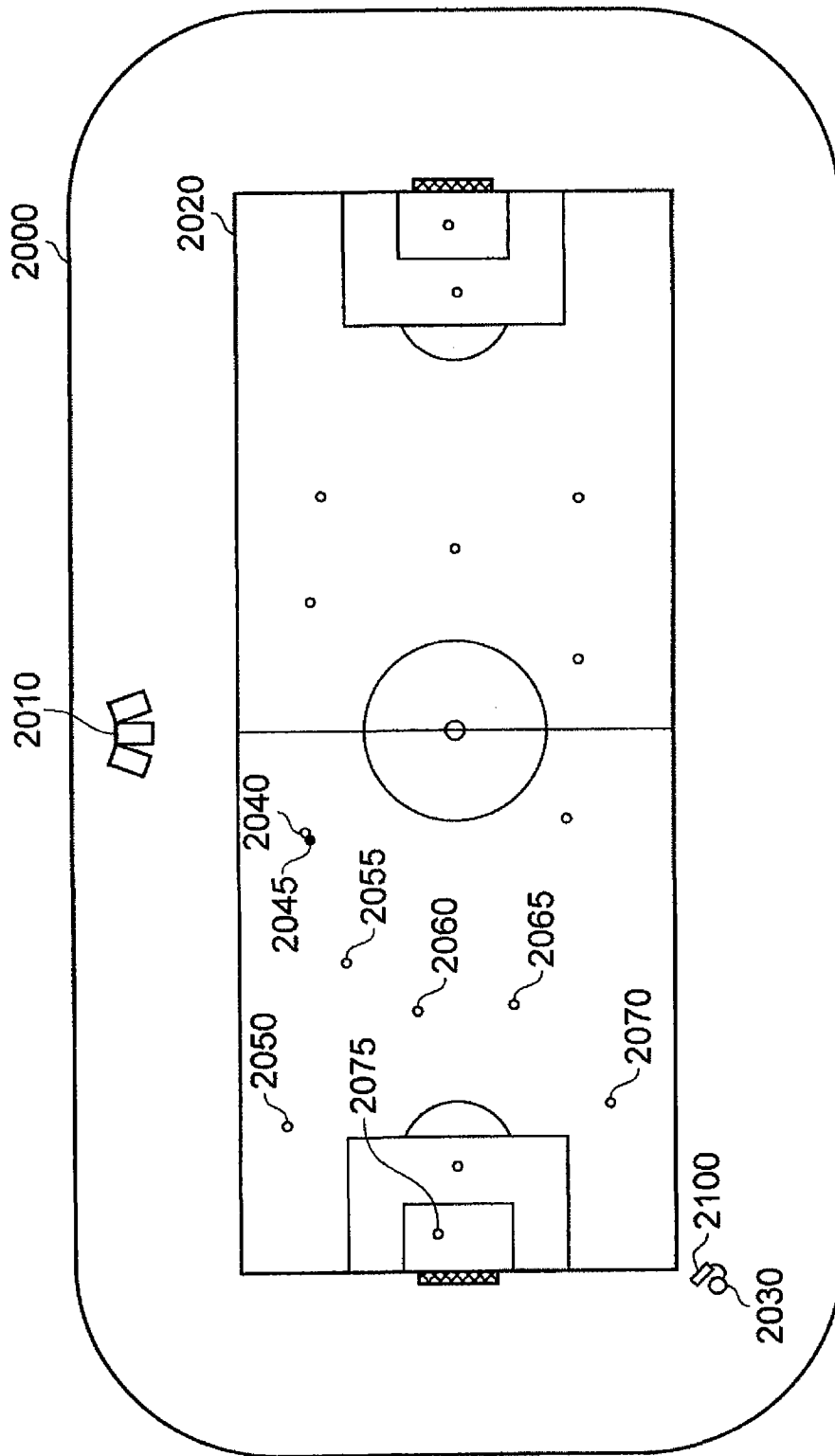


图 20

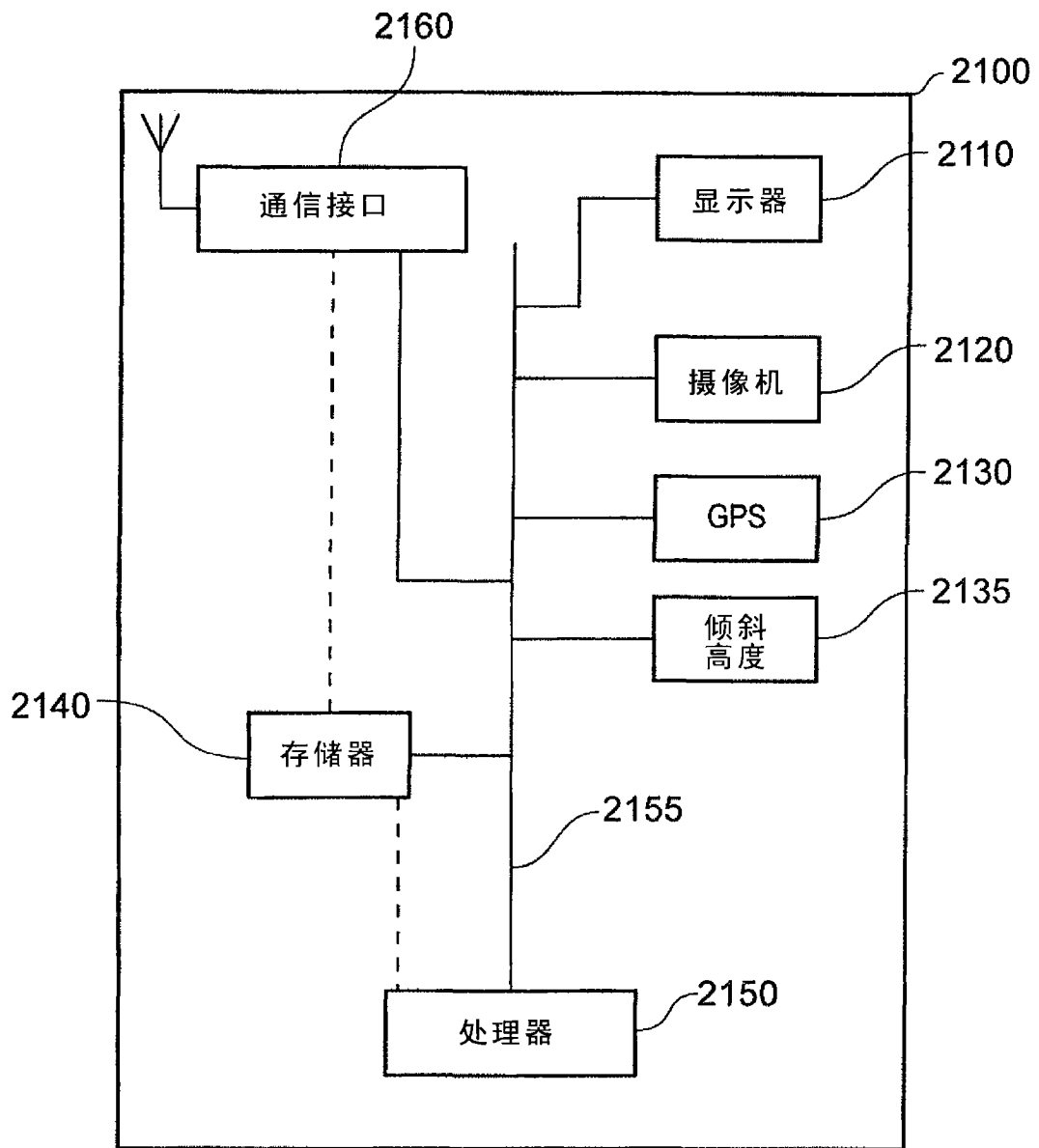


图 21

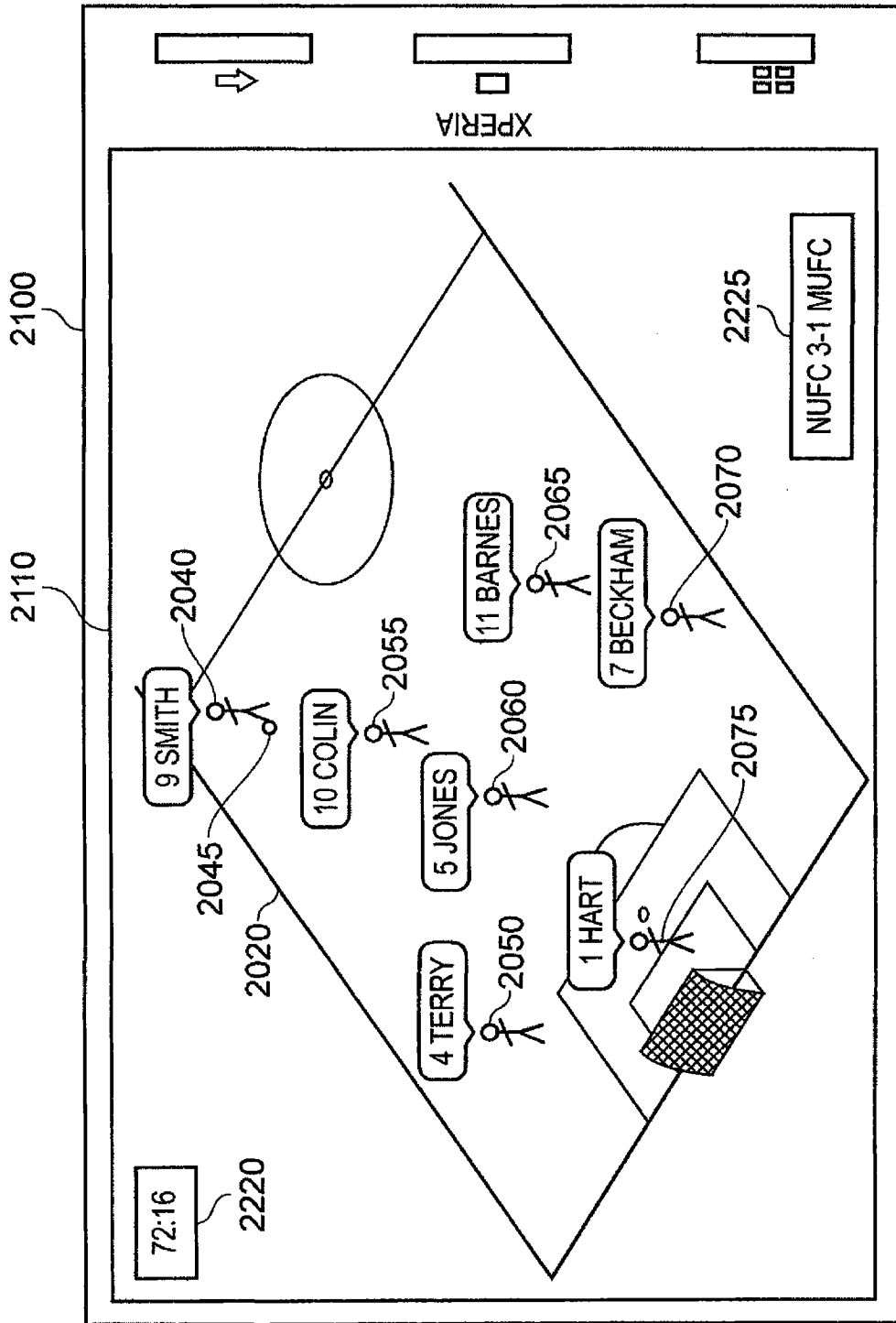


图 22

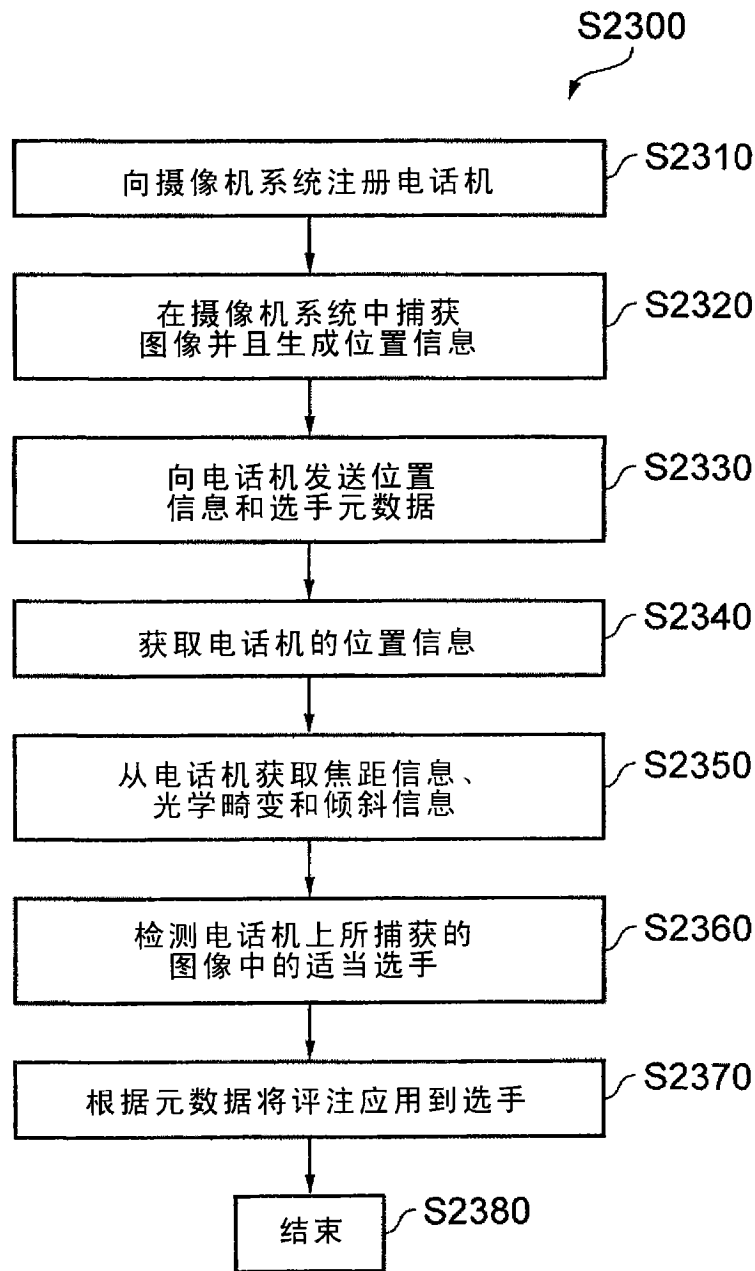


图 23