



(12) 发明专利申请

(10) 申请公布号 CN 104303218 A

(43) 申请公布日 2015. 01. 21

(21) 申请号 201380019203. 9

(51) Int. Cl.

(22) 申请日 2013. 03. 15

G08B 21/02 (2006. 01)

(30) 优先权数据

G08B 21/18 (2006. 01)

61/611, 284 2012. 03. 15 US

G08B 13/196 (2006. 01)

G06T 7/20 (2006. 01)

(85) PCT国际申请进入国家阶段日

2014. 10. 09

(86) PCT国际申请的申请数据

PCT/US2013/032075 2013. 03. 15

(87) PCT国际申请的公布数据

W02013/138719 EN 2013. 09. 19

(71) 申请人 行为识别系统公司

地址 美国德克萨斯州

(72) 发明人 韦斯利·肯内特·科布 明中·萧

徐刚 基肖尔·阿迪纳斯·塞特沃

安东尼·埃金斯 克里·约瑟夫

丹尼斯·乌雷奇

(74) 专利代理机构 北京律诚同业知识产权代理

有限公司 11006

代理人 徐金国 吴启超

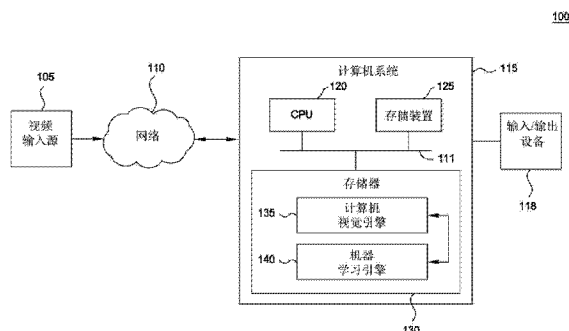
权利要求书1页 说明书11页 附图6页

(54) 发明名称

行为识别系统中的警报指令和聚焦警报指令

(57) 摘要

本发明公开了警报指令和聚焦警报指令,所述警报指令和聚焦警报指令允许用户向行为识别系统提供反馈,以便始终或从不发布针对某些事件的警报。这种方法绕过了行为识别系统的正常发布方法,然而并不妨碍系统的学习过程。



1. 一种用于针对对应于警报指令的行为向用户发出警报的方法,所述方法包括:
从场景中的所观察事件获得特性值,
就具有标准值的范围的匹配警报指令来剖析警报指令的列表,其中所述特性值在所述标准值的所述范围内,以及
在确认所述匹配警报指令后,针对所述所观察事件向所述用户发出警报。
2. 如权利要求1所述的方法,其中所观察事件的特性值是前景对象的像素-高度值、像素-宽度值以及x坐标和y坐标中心位置。
3. 如权利要求1所述的方法,其中所观察事件的特性值是对应于前景对象轨迹的一组x坐标和y坐标。
4. 如权利要求1所述的方法,其中所述匹配警报指令具有聚焦遮罩,所述聚焦遮罩与所观察事件发生的所述场景中的区域相交。
5. 如权利要求1所述的方法,其中所述标准值可以在图形编辑器中加以调整。
6. 如权利要求1所述的方法,其中针对所观察事件向所述用户发出警报包括基于所述匹配警报指令来产生一个或多个警报。
7. 如权利要求1所述的方法,其中就匹配警报指令来剖析警报指令的列表包括使对象类型和行为匹配。
8. 一种系统,其包括:
处理器,以及
托管应用程序的存储器,当所述应用程序在所述处理器上执行时,所述应用程序进行用于针对对应于警报指令的行为向用户发出警报的操作,所述操作包括:
从场景中的所观察事件获得特性值,
就具有标准值的范围的匹配警报指令来剖析警报指令的列表,其中所述特性值在所述标准值的所述范围内,以及
在确认所述匹配警报指令后,针对所观察事件向所述用户发出警报。
9. 如权利要求8所述的系统,其中所观察事件的特性值是前景对象的像素-高度值、像素-宽度值以及x坐标和y坐标中心位置。
10. 如权利要求8所述的系统,其中所观察事件的特性值是对应于前景对象轨迹的一组x坐标和y坐标。
11. 如权利要求8所述的系统,其中所述匹配警报指令具有聚焦遮罩,所述聚焦遮罩与所观察事件发生的所述场景中的区域相交。
12. 如权利要求8所述的系统,其中所述标准值可以在图形编辑器中加以调整。
13. 如权利要求8所述的系统,其中针对所观察事件向所述用户发出警报包括基于所述匹配警报指令来产生一个或多个警报。
14. 如权利要求8所述的系统,其中就匹配警报指令来剖析警报指令的列表包括使对象类型和行为匹配。

行为识别系统中的警报指令和聚焦警报指令

技术领域

[0001] 本发明的实施方案总体上涉及配置基于行为识别的视频监控系统来产生针对某些事件的警报。更确切来说,实施方案提供多种技术,这些技术允许行为识别系统确认应始终或从不产生警报的事件,而不会阻碍监控系统的非监督学习进程。

背景技术

[0002] 一些当前可用的视频监控系统提供了简单的对象识别能力。例如,视频监控系统可以被配置来将给定帧中的一组像素(称为“色斑(blob)”)分类为具体的对象(例如人或车辆)。一旦确认,就可以逐帧地跟踪“色斑”,以便追踪随时间的推移在场景中移动的“色斑”,例如,在视频监控摄像机的视野中穿行的人。另外,这类系统可以被配置来确定对象何时参与某些预定义的行为。例如,系统可以包括用于识别许多预定义事件的发生的定义,例如,系统可以评估分类为描绘汽车的对象的出现(车辆出现事件),所述对象在数帧内停止(车辆停车事件)。此后,新的前景对象可能出现并且被分类为人(人出现事件),人随后从帧中离开(人消失事件)。另外,系统可以将前两个事件的组合识别为“停车事件”。这类监视系统通常要求可由系统识别的对象和/或行为应预先定义。因此,在实践中,这些系统依赖于对象和/或行为的预定义的定义来评估视频序列。更一般来说,这类系统依赖于预定义的规则和静态模式,并因此往往不能动态地确认对象、事件、行为或模式,甚至更不能将它们分类为正常或异常的。

[0003] 另一方面,行为识别系统是一类视频监控系统,所述视频监控系统可以被配置来通过观察单个帧的序列(其它情况下称为视频流)来学习、确认和识别行为的模式。不同于基于规则的视频监控系统,行为识别系统通过归纳视频输入并建立对所观察事物的存储内容来学习对象和行为模式。随着时间的推移,行为识别系统使用这些存储内容来在视频流的视场中捕获的正常行为与异常行为之间进行区分。在检测到异常行为后,行为识别系统向用户发布警报来向用户通知所述行为。在具体事件的数次再现之后,行为识别系统得知事件是非异常的,并且停止发布后续警报。例如,聚焦在建筑物走廊上的行为识别系统最初可以在一天中的某一时间、在每次有人出现在摄像机视场内的走廊中时发布警报。如果这个事件发生足够多次,那么行为识别系统可以得知这是非异常行为,并停止就这个事件向用户发出警报。

[0004] 然而,虽然在多种情况下这是用户希望这种系统工作的方式,但是在一些情形下,用户可能希望行为识别系统始终针对具体的行为事件发布警报。回到前面的实例,如果走廊的通行权限受到限制,那么保安人员可能希望每次某人出现在走廊中时获得通知,以便确定走廊中的人只能是有权出现在那里的人。相反,用户可能从不希望行为识别系统针对具体的行为发布警报。当事件时常出现,但频率不足以产生警报时,就会出现这种情况。例如,聚焦于建筑工地附近的建筑物中的房间上的行为识别系统可以在施工车辆通过房间窗户外部的摄像机视场的任何情况下产生警报。在这种情形下,保安人员可能不希望行为识别系统每次都对这些事件发出警报。

[0005] 由于行为识别系统本身的性质,这些系统只要有可能就会避免使用预定义的规则以便有利于非监督学习。因此,解决这些问题需要一种用于向行为识别系统提供反馈的自然方法,所述反馈是关于所述系统应针对哪些行为始终或从不产生警报。

发明内容

[0006] 本发明的一个实施方案提供就对应于警报指令的行为向用户发出警报的方法。这种方法总体上可以包括从场景中的所观察事件获得特性值。这种方法还可以包括就具有标准值的范围的匹配警报指令来剖析警报指令的列表。如果特性值在标准值的范围内,那么所观察事件对应于匹配警报指令。这种方法还可以包括在确认匹配警报指令后,针对所观察事件向用户发出警报。

[0007] 此外,所观察事件的特性值可以是前景对象的像素-高度值、像素-宽度值以及 x 坐标和 y 坐标中心位置。特性值也可以是对应于前景对象的轨迹的一组 x 坐标和 y 坐标。另外,匹配警报指令可以具有聚焦遮罩 (focus mask),所述聚焦遮罩与所观察事件发生的场景中的区域相交。

[0008] 其它实施方案包括但不限于:计算机可读介质,所述计算机可读介质包括使处理单元能够实现所公开方法的一个或多个方面的指令;以及系统,所述系统具有被配置来实现所公开方法的一个或多个方面的处理器、存储器和应用程序。

附图说明

[0009] 因此,获得或可详细地理解本发明的上述特征、优点和目标的方式,即上文简要概述的较为具体的描述可参照附图中例示的实施方案进行。

[0010] 然而,应注意,附图仅例示本发明的典型实施方案,且因此不应被视为对本发明范围的限制,因为本公开可允许其他同等有效的实施方案。

[0011] 图 1 例示根据一个实施方案的视频分析系统的部件。

[0012] 图 2 进一步例示根据一个实施方案的在图 1 中所示的视频分析系统的部件。

[0013] 图 3 例示根据一个实施方案的客户端设备中的警报数据库的实例。

[0014] 图 4 例示根据一个实施方案的用于在配置有警报指令和聚焦警报指令的行为识别系统中发布警报的方法。

[0015] 图 5 例示根据一个实施方案的适用于轨迹警报的一组公差의示例性图形表示。

[0016] 图 6 例示根据一个实施方案的适用于具体警报的警报指令和聚焦警报指令的示例性图形表示。

具体实施方式

[0017] 本文公开的本发明的实施方案提供用于在基于行为识别的视频监控系统中创建警报指令和聚焦警报指令的技术。也就是说,所公开的技术允许行为识别系统的用户确认先前发过警报的行为,所述行为应始终或从不产生后续警报。由于警报指令仅优先于行为识别系统的正常警报发布过程(其在系统已经进行它的学习过程后发生),因此这种方法不会破坏行为识别系统的非监督学习。

[0018] 在一个实施方案中,行为识别系统包括计算机视觉引擎和机器学习引擎。计算机

视觉引擎可以被配置来处理视频流中捕获的视场。这个视场通常被称为“场景”。在处理中，计算机视觉引擎将前景对象（例如，类似于人、车辆等的对象）与背景对象（例如，类似于路面、天空等的对象）分离。处理场景后，计算机视觉引擎可产生所观察活动的信息流（例如，外观特征、运动特征等），并且将所述流传送到机器学习引擎。继而，机器学习引擎可被配置来使用所述信息学习场景中的对象行为。除了基于学习的行为之外，机器学习引擎可被配置来建立场景内某些行为的模型，并且确定观察结果是否指示对象的行为相对于模型是异常的。在检测到异常行为后，机器学习引擎产生警报。在确定应发布警报后，行为识别系统将警报发布至用户界面。用户界面可以包含系统的用户通常可访问的先前所发出警报的数据库。用户可以查看呈列表形式的这些警报，其中每个列表项显示警报的信息，并且可以包括对应的视频或图像数据。

[0019] 在发布针对具体行为事件的足够数量的警报后，机器学习引擎得知所述事件是非异常发生，并且停止发布针对所述事件的后续警报。在一个实施方案中，用户可以创建警报指令来优先于正常的警报发布进程。警报指令允许用户向机器学习引擎提供反馈，以便始终或从不针对某一行为事件产生警报。机器学习引擎在学习有关事件的信息流之后，并且在评估事件的异常行为之前查询警报指令定义的列表。因此，警报指令并不阻碍机器学习引擎的学习过程。

[0020] 为了创建警报指令，用户选择先前由系统产生的事件发生或警报来用作模板。例如，用户可以剖析针对场景的基于时间、类型、名称、事件或其它来表征的先前警报数据库，以及查看引起警报的基本活动视频。在一个实施方案中，用户可以通过用户界面上的警报浏览器中的对话框来执行这个操作。选择警报之后，用户定义警报指令匹配标准。在一个实施方案中，标准可以包括行为是否应始终或从不产生警报，警报应发布的频繁程度（例如，在行为导致在短时间跨度内的大量警报的情况下），以及是否机器学习引擎应匹配行为或对象类型（或两者兼有）。一旦用户已经定义匹配标准，用户界面就在警报数据库中创建警报指令，其中引用符（reference）往回指向用于创建所述警报指令的原始警报。此后，用户界面将关于警报指令的信息发送到机器学习引擎。

[0021] 注意，通过机器学习引擎将警报指令与警报行为匹配可以取决于警报类型和由用户指定的一系列参数。例如，一旦用户选择要用于警报指令的警报，所述警报的对应视频就可以显示安全门前方的人连同边界框，所述边界框指示由系统分类为描绘那个人的像素。在这种情况下，图形编辑器可以允许用户调整选定警报中的人周围的边界框，以便调整警报指令的公差——创建例如前景对象的中心（x，y）位置的范围等。类似地，用户可以指定用于门前方的人的相对位置的另一边界框，即对象位置的公差。通过调整公差大小和人的姿势以及这个人的位置，警报指令可以用于指定安全门前方的区域，以使得在观察到任何人存在的任何情况下，机器学习引擎产生警报。因此，这种方法允许所观察对象的高度、位置、宽度、速度等的变化，而仍然满足警报指令定义。

[0022] 在另一实施方案中，警报指令可以被扩展来提供聚焦警报指令。这种方法扩展了机器学习引擎可以将警报指令应用于场景中的行为事件的情况。例如，摄像机可以聚焦于具有多个安全门的建筑物走廊。在这种情况下，用户必须对每个门创建并且调整单独的警报指令，以便在某人出现在门前的任何情况下发出警报。作为替代方案，聚焦警报指令允许用户创建警报指令（例如，针对出现在安全门前方的人来创建）以及聚焦遮罩，所述聚焦遮

罩指定场景中应产生警报的不同区域。

[0023] 也就是说,用户可以将位置的公差扩展至全视场,而不是指定安全门前方的人的位置公差。用户定义了警报指令公差内其它处观察到的前景对象时,场景中应产生警报的一个或多个区域。例如,出现在第一个门前方(在高度、宽度和姿势的警报指令公差内)的人的警报定义了警报指令,但所述位置被延伸到与用户定义区域相交的整个摄像机视场。因此,为了在摄像机聚焦于具有多个安全门的区域上的给定实例中创建聚焦警报指令,用户针对任何一个门前方的人选择“人出现”警报,使用图形编辑器来指定所述人的外观周围的公差,并且随后创建“人出现”警报的位置的遮罩以便包括大致在每一个门前方的区域。

[0024] 一旦用户创建警报指令(或聚焦警报指令),用户界面就将警报指令(和聚焦遮罩,若适用)的信息发送到机器学习引擎。当机器学习引擎处理匹配警报指令的匹配标准和公差的后继事件的信息时,机器学习引擎绕过行为识别系统的正常发布方法,并且立即发布警报或舍弃事件(在给定匹配标准的情况下),与机器学习引擎是否认为所观察行为是异常无关。这种方法并不改变关于具体场景的已学习状态或影响机器学习引擎的非指导学习。在所有情况下,机器学习引擎在应用警报指令之前已经进行了它的学习过程。

[0025] 在下文中,参考了本公开的实施方案。然而,应理解的是,本公开不限于任何特定描述的实施方案。相反,无论是否涉及不同的实施方案,以下特征和要素的任意组合被涵盖来实施和实践所公开的内容。此外,在各种实施方案中,本发明提供优于现有技术的许多优点。然而,尽管实施方案可以实现优于其它可能的解决方案和/或优于现有技术的优点,但是给定的实施方案是否实现具体的优点不具有限制性。因此,以下各方面、特征、实施方案和优点仅仅是说明性的,并且不视为附加权利要求书的要素或限制,除非在权利要求书中有明确记载。同样,对“本发明”或“本公开”的任何提及不应解释为对本文公开的任何发明主题的概括,并且不应视为附加权利要求书的要素或限制,除非在权利要求书中有明确记载。

[0026] 本发明的一个实施方案被实现为供计算机系统使用的程序产品。程序产品的程序定义了实施方案的功能(包括本文描述的方法)并且可以包含在各种计算机可读存储介质上。计算机可读存储介质的实例包括:(i)不可写存储介质(例如,计算机内的只读存储器设备,如可通过光学介质驱动器读取的CD-ROM或DVD-ROM盘),信息被永久存储在这些存储介质上;(ii)可写存储介质(例如,硬盘驱动器),可变的信息存储在这些存储介质上。在带有引导本发明的功能的计算机可读指令时,这类计算机可读存储介质是本发明的实施方案。其它示例性介质包括借以将信息(如通过计算机或电话网络)传输到计算机的通信介质,包括无线通信网络。

[0027] 一般来说,被执行来实施本发明的实施方案的例行程序可以是操作系统或特定应用程序、部件、程序、模块、对象或指令序列的一部分。本发明的计算机程序通常包括大量的指令,这些指令将通过本机计算机翻译成机器可读格式并且由此得到可执行指令。此外,程序包括驻留在本地程序中或存在于存储器中或存储设备上的变量和数据结构。此外,本文中描述的各种程序可基于这些程序在本公开的特定实施方案中针对性实施的应用来确认。然而,应了解的是,下文任何具体的程序命名仅仅是为了方便而使用,且因此本公开不应限于仅在由这样的术语所确认和/或暗示的任何特定应用中使用。

[0028] 图 1 例示根据一个实施方案的视频分析与行为识别系统 100 的部件。如图所示, 行为识别系统 100 包括视频输入源 105、网络 110、计算机系统 115 以及输入与输出设备 118(例如, 监视器、键盘、鼠标、打印机等等)。网络 110 可将由视频输入 105 所记录的视频数据发送至计算机系统 115。说明性地, 计算机系统 115 包括 CPU 120、存储装置 125(例如, 磁盘驱动器、光盘驱动器、软盘驱动器等等), 以及包含计算机视觉引擎 135 和机器学习引擎 140 的存储器 130。如下文更详细地描述, 计算机视觉引擎 135 和机器学习引擎 140 可以提供软件应用程序, 所述软件应用程序被配置来分析由视频输入 105 提供的视频帧的序列。

[0029] 网络 110 接收来自视频输入源 105 的视频数据(例如, 视频流、视频图像等)。视频输入源 105 可以是视频摄像机、VCR、DVR、DVD、计算机、网络摄像设备等。例如, 视频输入源 105 可以是瞄准某个区域(例如, 地铁站、停车场、建筑物入口/出口等等)的固定视频摄像机, 所述固定视频摄像机记录这个区域中正在发生的事件。一般来说, 摄像机视场中的区域被称为场景。视频输入源 105 可以被配置来在指定帧率(例如, 每秒 24 帧)下将场景记录为单个视频帧的序列, 其中每一帧包括固定数量的像素(例如, 320×240 个)。每一帧的每一像素可以指定颜色值(例如, RGB 值)或灰度值(例如, 0 至 255 之间的亮度值)。另外, 视频流可以使用已知的例如 MPEG2、MJPEG、MPEG4、H. 263、H. 264 等的这类格式来进行格式化。

[0030] 如上所述, 计算机视觉引擎 135 可以被配置来分析这个原始信息以确认视频流中的活动对象、确认由机器学习引擎 140 用来得出对象分类的各种外观和运动特征、得出关于这些对象的动作和相互作用的各种元数据, 并将这个信息提供给机器学习引擎 140。并且继而, 机器学习引擎 140 可以被配置来评估、观察、学习和存储关于在场景内随着时间的推移而发生的事件(和事件类型)的细节。

[0031] 在一个实施方案中, 机器学习引擎 140 接收由计算机视觉引擎 135 产生的视频帧和数据。机器学习引擎 140 可以被配置来分析所接收的数据、使具有相似视觉和/或运动特征的对象聚集、建立视频帧中描绘的事件的语义表示。机器学习引擎 140 学习映射到群集的对象的前期行为模式。因此, 随着时间的推移, 机器学习引擎学习这些观察到的模式, 以便确认正常和/或异常事件。也就是说, 机器学习引擎 140 会建立自身的已观察到的不同对象类型的模型(例如, 基于运动和或外观特征的群集), 以及建立给定对象类型的前期行为的模型, 而不是利用预先定义的模式、对象、对象类型或活动。随后, 机器学习引擎可以基于先前的学习来决定所观察事件的行为是否异常。

[0032] 描述是否已确定正常/异常行为/事件和/或这个行为/事件的具体内容的数据可以被提供给输出设备 118 以发出警报, 例如, 在 GUI 界面屏幕上呈现的具有对应视频和图像数据的警报消息。这类输出设备也可以配置有先前所发出警报的数据库, 用户可根据所述数据库来创建警报指令。

[0033] 一般来说, 计算机视觉引擎 135 和机器学习引擎 140 两者都实时处理视频数据。然而, 计算机视觉引擎 135 和机器学习引擎 140 处理信息的时间尺度可以是不同的。例如, 在一个实施方案中, 计算机视觉引擎 135 逐帧地处理所接收的视频数据, 而机器学习引擎 140 每 N 个帧处理一次数据。换句话说, 虽然计算机视觉引擎 135 可以实时分析每一帧来得出与帧中观察到的对象相关的一组运动和外观数据, 但是机器学习引擎 140 不受视频输入的

实时帧率限制。

[0034] 然而,应注意,图 1 仅例示行为识别系统 100 的一个可能的布置。例如,虽然视频输入源 105 展示为经由网络 110 连接到计算机系统 115,但是网络 110 并不总是存在或是需要的(例如,视频输入源 105 可以直接连接到计算机系统 115)。另外,行为识别系统 100 的各种部件和模块可以在其它系统中实现。例如,在一个实施方案中,计算机视觉引擎 135 可以实现为视频输入设备的一部分(例如,实现为直接有线连接到视频摄像机的固件部件)。在这样的情况下,视频摄像机的输出可以被提供给机器学习引擎 140 来进行分析。类似地,来自计算机视觉引擎 135 和机器学习引擎 140 的输出可以通过计算机网络 110 上提供到其它计算机系统。例如,计算机视觉引擎 135 和机器学习引擎 140 可以安装在服务器系统上,并且被配置来处理来自多个输入源(即,来自多个摄像机)的视频。在这样的情况下,在另一计算机系统上运行的客户端应用程序 250 可以通过网络 110 来请求(或接收)结果。

[0035] 图 2 进一步例示根据本发明的一个实施方案的首先在图 1 中例示的计算机视觉引擎 135 和机器学习引擎 140 的部件。如图所示,计算机视觉引擎 135 包括数据摄入器 205、检测器 215、跟踪器 215、情境事件发生器 220、警报发生器 225 和事件总线 230。共同来说,部件 205、210、215 和 220 提供用于处理由视频输入源 105 提供的传入视频帧序列的流水线(由连接部件的实线箭头指示)。在一个实施方案中,部件 210、215 和 220 可以各自提供被配置来提供本文描述的功能的软件模块。当然,本领域技术人员将认识到,部件 205、210、215 和 220 可加以组合(或进一步细分)以适应具体情况的需要,并且应进一步认识到其它部件可以添加至视频监控系统(或可以从所述系统除去一些部件)。

[0036] 在一个实施方案中,数据摄入器 205 接收来自视频输入源 105 的视频输入。数据摄入器 205 可以被配置来预处理输入数据,然后将数据发送到检测器 210。检测器 210 可以被配置来将所提供的视频的每一帧分成固定或静态部分(场景背景)和易变部分的集合(场景前景)。帧本身可以包括多个通道(例如,彩色视频的 RGB 通道或黑白视频的灰度通道或亮度通道)的两维像素值阵列。在一个实施方案中,检测器 210 可以使用自适应共振理论(ART)网络来将每一像素的背景状态建模。也就是说,可以使用将给定像素建模的 ART 网络将每一像素分类为描绘场景前景或场景背景。当然,可以使用在场景前景与场景背景之间进行区分的其它方法。

[0037] 此外,检测器 210 可以被配置来产生遮罩,所述遮罩用于确认场景的哪些像素被分类为描绘前景,以及相反地,哪些像素被分类为描绘场景背景。然后,检测器 210 确认场景中包含场景前景的一部分(称为前景“色斑”或“斑块(patch)”)的区域,并且将这个信息提供给流水线的后续阶段。此外,分类为描绘场景背景的像素可用于产生将场景建模的背景图像。

[0038] 在一个实施方案中,检测器 210 可以被配置来检测场景的流(flow)。一旦前景斑块已经分离,检测器 210 就逐帧地检查所有前景斑块的任何边缘和拐角。检测器 210 将在相似运动流中移动的前景斑块确认为最有可能属于单一对象或单一运动关联,并且将这个信息发送到跟踪器 215。

[0039] 跟踪器 215 可以接收由检测器 210 产生的前景斑块并且针对所述斑块产生计算模型。跟踪器 215 可以被配置来使用这个信息和原始视频的每一连续帧,以试图在由给定前景斑块描绘的对象在场景中移动时跟踪所述对象的运动。也就是说,跟踪器 215 通过逐帧

地跟踪给定对象来为系统的其它元件提供连续性。所述跟踪器进一步计算前景对象的各种运动和/或外观特征,例如,大小、高度、宽度和面积(以像素为单位)、反射率、光泽度刚性、速度(speed velocity)等等。

[0040] 情境事件发生器 220 可以接收来自流水线的其它阶段的输出。使用这个信息,情境处理器 220 可以被配置来产生关于(由跟踪器部件 210)所跟踪对象的情境事件流。例如,情境事件发生器 220 可将对象的微特征矢量和动态观察结果的流封包并且例如以 5Hz 的速率将这个封包输出给机器学习引擎 140。在一个实施方案中,情境事件以轨迹形式来封包。如本文中所使用,轨迹总体上是指将连续帧或样本中的具体前景对象的运动数据封包的矢量。轨迹中的每一要素表示在具体时间点捕获到的所述对象的运动数据。通常,完整的轨迹包括对象在视频帧中首次被观察到时所获得的运动数据,连同所述对象在离开场景(或变为固定于融入帧背景中的点)前的每一连续观察结果。因此,假定计算机视觉引擎 135 以 5Hz 的速率工作,对象的轨迹每 200 毫秒更新一次,直到完成。情境事件发生器 220 也可以计算和封包每一跟踪对象的外观数据,方式是评估对象的如形状、宽度和其它物理特征的各种外观属性,并且对每一属性分配数值评分。

[0041] 计算机视觉引擎 135 可以从部件 205、210、215 和 220 获取描述场景中跟踪对象的运动和动作的输出,并且通过事件总线 230 将这个信息提供给机器学习引擎 140。说明性地,机器学习引擎 140 包括分类器 235、语义模块 240、映射器 245、认知模块 250 和归一化模块 265。

[0042] 分类器 235 从计算机视觉引擎 135 接收如运动数据和外观数据的情境事件并且将数据映射在神经网络上。在一个实施方案中,神经网络是自组织映射(SOM)和 ART 网络的组合,在图 2 中展示为 SOM-ART 分类器 236。数据根据重复出现的彼此相关联的特征来聚集并组合。然后,基于这些再现类型,分类器 235 定义出对象的类型。例如,分类器 235 可以将具有例如高光泽度刚性和反射率的前景斑块定义为 1 型对象。然后,这些被定义的类型沿用到系统的全部其余部分。

[0043] 映射器 240 可以通过在系统中针对斑块搜索空间和时间相关性和行为来使用这些类型,从而建立事件可能或不可能发生的地点和时间的映射。在一个实施方案中,映射器 240 包括时间存储器 ART 网络 241、空间存储器 ART 网络 242 和统计引擎 243。例如,映射器 240 可以寻找 1 型对象的斑块。空间存储器 ART 网络 242 使用统计引擎 243 来建立这些对象的统计数据,如场景中何处会出现这些斑块、这些斑块倾向于在什么方向上行进、这些斑块行进的快速程度如何、这些斑块是否改变方向等。然后,映射器 240 建立这个信息的神经网络,所述神经网络成为用以比较对象行为的存储模板。时间存储器 ART 网络 241 使用统计引擎 243 来创建基于时间片采样的统计数据。在一个实施方案中,初始采样以每三十分钟的间隔发生。如果在时间片内发生许多事件,那么时间分辨率可以被动态地改变成更精细的分辨率。反之,如果在时间片内发生较少事件,那么时间分辨率可以被动态地改变成较粗略的分辨率。

[0044] 在一个实施方案中,语义模块 245 包括相空间分隔部件 246。语义模块 245 确认场景中的运动或轨迹的模式,并且通过概括法来分析场景的异常行为。通过划分场景并且将前景斑块分成许多不同的嵌块(tessera),语义模块 245 可跟踪对象的轨迹并且从轨迹中学习模式。语义模块 245 分析这些模式,并且将它们与其它模式进行比较。当对象进入场

景时,相空间分隔部件 246 建立自适应栅格并且将对象和它们的轨迹映射到栅格上。在更多特征和轨迹被填充到栅格上时,机器学习引擎就学习对于场景来说常见的轨迹,并且进一步将正常行为与异常行为区分开。

[0045] 在一个实施方案中,认知模块 250 包括感知存储器 251、情节存储器 252、长期存储器 253、工作区 254 和代码段 (codelet) 255。一般来说,工作区 254 为机器学习引擎 140 提供计算引擎。例如,工作区 240 可以被配置来复制来自感知存储器 251 的信息、从情节存储器 252 和长期存储器 253 中检索相关的存储内容、选择要执行哪些代码段 255。在一个实施方案中,每一代码段 255 是被配置来评估不同事件序列并且确定一个序列可以如何续接 (或以其它方式关联于) 另一序列的软件程序 (例如,有限状态机)。更一般来说,代码段可以提供被配置来从馈入机器学习引擎的数据流中检测感兴趣的模式的软件模块。继而,代码段 255 可以在情节存储器 252 和长期存储器 253 中创建、检索、加强或修改存储内容。通过重复地调用供执行用的代码段 255、复制去往 / 来自工作区 240 的存储内容与感知,机器学习引擎 140 执行用于观察和学习在场景中出现的行为模式的认知循环。

[0046] 在一个实施方案中,感知存储器 251、情节存储器 252 和长期存储器 253 用于确认行为模式、评估在场景中发生的事件,并且编码并存储观察结果。一般来说,感知存储器 251 接收计算机视觉引擎 135 的输出 (例如,情境事件流)。情节存储器 252 存储表示所观察事件以及与具体情节相关的细节的数据,例如,描述与事件相关的时间和空间细节的信息。也就是说,情节存储器 252 可以编码具体事件的特定细节,即场景中“何事和何地”发生某个事情,如具体车辆 (汽车 A) 在上午 9 时 43 分移动至确认为停车空间的位置 (停车空间 5)。

[0047] 对比来说,长期存储器 253 可以存储概括场景中所观察到的事件的数据。继续以车辆停车为例,长期存储器 253 可以编码采集通过分析场景中的对象行为来获知的观察结果和概括的信息,如“车辆倾向于停在场景的具体位置”、“停车时,车辆倾向于以某一速度移动”以及“车辆停车后,人会倾向于出现在场景中邻近车辆处”等。因此,长期存储器 253 存储关于在场景内发生何事的观察结果,其中许多具体的情节细节被去掉。在这种方式下,当新的事件发生时,来自情节存储器 252 和长期存储器 253 的存储内容可以用于关联和理解当前事件,即,新的事件可以与过去的经历进行比较,从而随着时间的推移引起对存储在长期存储器 253 中的信息的加强、衰减和调整。在具体的实施方案中,长期存储器 253 可以被实现为 ART 网络和稀疏分布式存储数据结构。然而,重要的是,这种方法不需要预先定义不同的对象类型分类。

[0048] 在一个实施方案中,模块 235、240、245 和 250 包括异常检测部件,如部件 237、244、247 和 256 所描绘。每一异常检测部件被配置来确认相对于场景的过去观察结果的异常行为。另外,每一部件被配置来接收来自警报数据库 270 的警报指令和聚焦遮罩信息。一般来说,如果任何异常检测部件确认异常行为,那么部件产生警报并且使警报通过归一化模块 265。例如,语义模块 245 中的异常检测器 247 使用所学习的模式和模型来检测反常的轨迹。如果前景对象表现出游荡行为,那么例如,异常检测部件 247 使用游荡模型来评估对象轨迹,随后产生警报,并将警报发送到归一化模块 265。在接收警报后,归一化模块 265 基于相对于所述警报类型的先前警报的警报罕见度来评估警报是否应予以发布。一旦归一化模块 265 确定应发布警报,那么所述模块就将警报传送到警报发生器 225 (通过事件总线

230)。

[0049] 然而,如果异常检测部件确认匹配警报指令的事件,那么异常检测器部件遵循警报指令的匹配标准而不是针对异常行为来评估所述事件。如果警报指令要求公布警报,那么异常检测部件将警报发送到警报发生器 225(通过事件总线 230)。否则,异常检测部件舍弃事件。注意,在任一情况下,如果事件数据匹配警报指令,那么异常检测部件不发送任何信息到归一化模块 265。

[0050] 在一个实施方案中,警报发生器 225 驻留在计算机视觉引擎 135 中。警报发生器 225 接收来自异常检测部件 237、244、247 和 256 以及归一化模块 265 的警报信息。警报发生器 225 将警报信息发布到 GUI/ 客户端设备 260。GUI/ 客户端设备将这个警报信息存储在警报数据库 270 中。警报数据库 270 包含先前所发出的警报,并且可以由 GUI/ 客户端设备 270 的用户来访问。

[0051] 图 3 例示根据一个实施方案的客户端设备中的警报数据库 300 的实例。警报数据库 300 存储先前所发出的警报,以使用户可以进行剖析来创建警报指令。如图所示,警报数据库 300 包括多个警报和警报指令列表 305。每个警报 310 包括标识符 311、指令标识符 312 和概要 313。标识符 311 是分配给警报 310 的唯一数字值。指令标识符 312 是指示警报 310 是否已分配警报指令的数字字段。

[0052] 概要 313 是数据有效载荷,其包含表征警报的数据的简要说明。概要 313 可以包括关于以下的信息:异常类型、异常何时发生、对象的高度值与宽度值以及 x 坐标和 y 坐标(如果异常在某个时间点发生)、对应于轨迹的一组 x 坐标和 y 坐标(如果异常在一系列帧中发生)等。警报指令评估匹配概要 313 中所提供的信息的对象行为或对象类型(或两者兼有)。

[0053] 警报指令列表 305 包括多个警报指令。每一警报指令 320 具有标识符 321、警报指针 322、匹配标准 323 和尾言 324。警报指令的标识符 321 是分配给警报指令的唯一数值。警报指针 322 是指向警报指令对应的原始警报的指针。通过指向原始警报,警报指令 320 可以访问由概要 313 提供的数据。在一个实施方案中,包含在概要 313 中的信息可以在对应警报指令 320 中存储为数据包。

[0054] 匹配标准 323 包含警报指令应如何处理某一事件的用户指定信息,如机器学习引擎是否应发布警报或舍弃行为,以及是否将警报指令与行为或对象类型(或两者)匹配。例如,如果用户选择忽略“反常位置”警报的匹配行为,那么机器学习引擎可以创建由警报指令指定的位置处的静止状态下的对象的警报,并且可以创建快速移动通过相同位置的对象的警报。另举一例,如果用户选择忽略“反常位置”警报的匹配类型,那么机器学习引擎可以创建位于所述位置处的对应于基于学习的分类类型 1(如汽车)的对象的警报,并且机器学习引擎还可以创建位于所述位置处的对应于基于学习的分类类型 2(如人)的对象的警报。

[0055] 尾言 324 是在由概要 313 所提供的数据中每一对应警报特性的公差值的阵列。公差为机器学习引擎提供将对象行为和类型与警报指令相匹配的灵活性,因为将场景中的具有相同特性(高度、宽度和中心(x, y)位置)的两个对象匹配的可能性是非常低的。在一个实施方案中,用户通过对选定的警报使用图形编辑器来定义这些公差。通过在触发警报的对象周围绘制边界框,用户可以调整警报指令的公差,创建选定警报的若干特征(例如,

对象的高度和宽度)的范围。

[0056] 图 4 是根据一个实施方案的在配置有警报指令的行为识别系统中发布警报的方法 400。方法 400 开始于步骤 405,其中机器学习引擎加载警报指令列表(和聚焦遮罩,若适用)。在一个实施方案中,机器学习引擎在系统启动时加载警报指令列表。另外,当用户在启动发生后创建警报指令时,用户界面将警报指令的信息发送到机器学习引擎。在步骤 410 处,机器学习引擎处理行为事件。例如,机器学习引擎可以处理由计算机视觉引擎产生的对应于站在场景中一点的人的信息。这时,机器引擎已经完成了它的学习过程。在步骤 415 处,机器学习引擎搜索警报指令列表,以便确定行为是否对应于基于匹配标准的警报指令。如果存在匹配警报指令(步骤 425),那么机器学习引擎绕过正常的发布进程并且将警报发布至用户界面。在目前的实例中,警报指令列表可以包括在给定人的高度、宽度和中心(x, y)位置的公差的情况下,始终针对站在场景中某个位置中的任何人(即对应于人的对象模型)来发出“反常位置”警报的指令。如果观察到的人的高度和宽度以及位置坐标与警报指令匹配,那么行为识别系统立即发布“反常位置”警报。然而,如果不存在匹配警报指令(步骤 430),那么机器学习引擎继续进行正常发布进程,并且评估事件的异常行为。

[0057] 图 5 是根据一个实施方案的应用于轨迹警报的一组公差的示例性图形表示。由于轨迹在一系列视频帧内发生,因此机器学习引擎将基于轨迹的事件与不同于场景中一点处发生的行为事件的警报指令相匹配,因而公差也以不同的方式创建。原始轨迹 515 表示产生警报的轨迹。如图所示,原始轨迹 515 包括起始点 505 和结束点 510,以及两个点之间的距离 525。除了这些部件之外,原始轨迹 515 的警报指令包括对应于路径的一组坐标。在一个实施方案中,用户可以通过图形界面对轨迹分配公差,以使得轨迹的未来发生不需要严格遵守原始轨迹 515 的坐标。通过在起始点 505 和结束点 510 周围创建边界框,用户指定轨迹发生以触发警报指令的公差区域(由虚线包围的区域表示)。因此,在替代轨迹 520 上行进的对象会触发警报指令,因为轨迹在公差设置范围内(由处于虚线包围的区域内示出)。

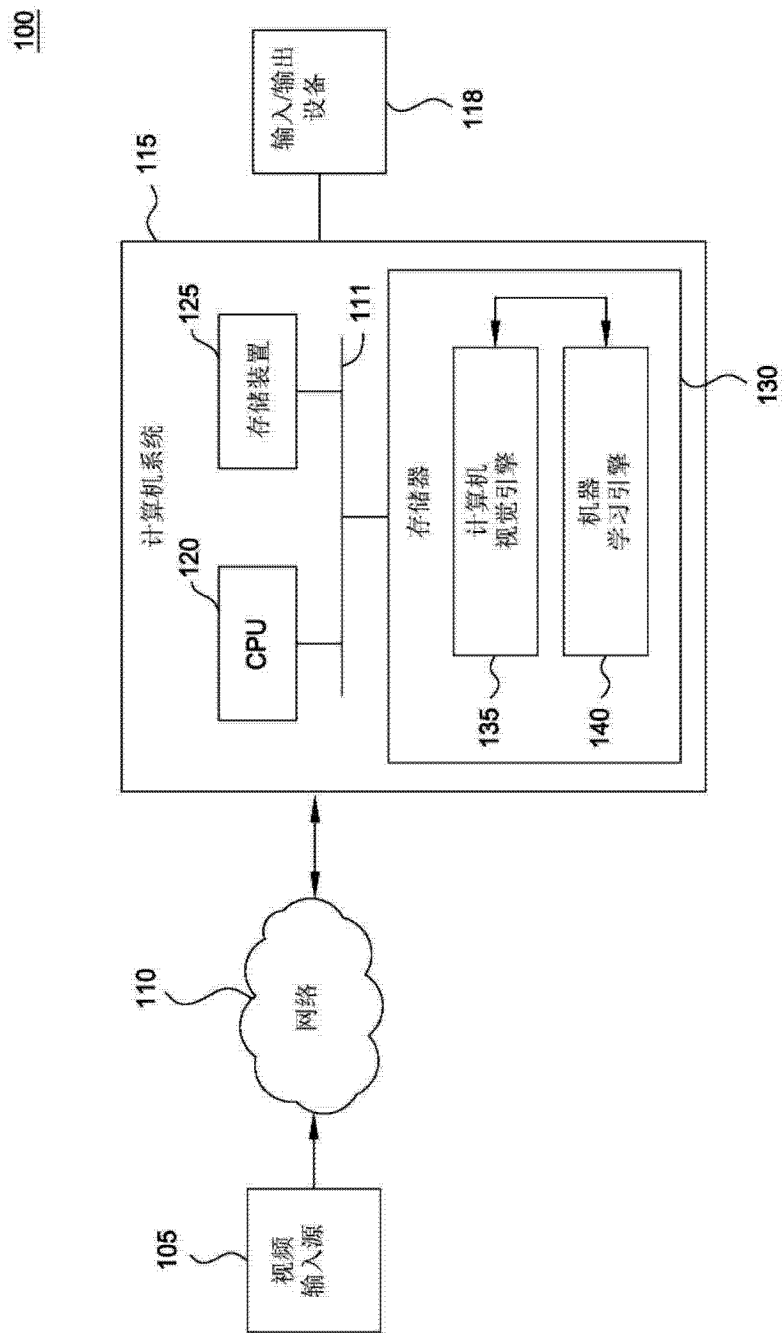
[0058] 图 6 是根据一个实施方案的应用于场景内的具体警报的警报指令和聚焦警报指令的示例性图形表示。在这个实例中,行为识别系统聚焦于火车站台上。图像 605、610、615 和 620 都表示提供给用户的相同警报的图像。图像 605 表示原始警报,其中边界框 606 在触发警报的人(即由机器学习引擎分类为人的像素)周围。对于这个实例的目的,假定警报是“反常位置”警报。在行为识别系统中,这个警报数据可以包括对象的高度和宽度像素值,以及对象的中心(x, y)位置。图像 610 表示通过在原始警报周围绘制较宽边界框 607 来创建警报指令的用户。通过创建较宽边界框 607,用户对机器学习引擎设置较大公差以便在处理所述区域内的类似出现时进行匹配。因此,出现在较宽边界框 607 中所描绘的场景的阴影部分中的人 would 触发“反常位置”警报的警报指令。

[0059] 用户可能希望相同“反常位置”警报指令适用于出现在对应于铁轨的场景区域内的对象。因此,用户可以创建聚焦警报指令来完成这个操作。图像 615 和 620 表示创建图像 605 中描绘的“反常位置”警报的聚焦警报指令的用户。为了从现有警报创建聚焦警报指令,用户首先在用户想要应用聚焦遮罩的部分上创建边界框 616。在所述边界框内,用户可以选择一个区域(多个区域),并且从边界框与所选择区域(多个区域)的相交部分产生聚焦遮罩。此后,如果人走到场景中的铁轨上,那么机器学习引擎使用聚焦警报指令来处理

这个行为并且发布警报。

[0060] 如上所述,本发明的实施方案提供了配置行为识别系统以产生警报的技术。更确切来说,通过创建警报指令(或聚焦警报指令)供机器学习引擎遵循,某些事件始终或从不产生警报。有利的是,这种方法并不会阻碍行为识别系统的非监督学习进程,因为当行为事件触发警报指令时,机器学习引擎已经完成了它的学习进程。

[0061] 尽管前述内容针对本公开的实施方案,但是可以设计本公开的其它和进一步的实施方案而不脱离本公开的基本范围,并且本发明的范围由所附权利要求书来决定。



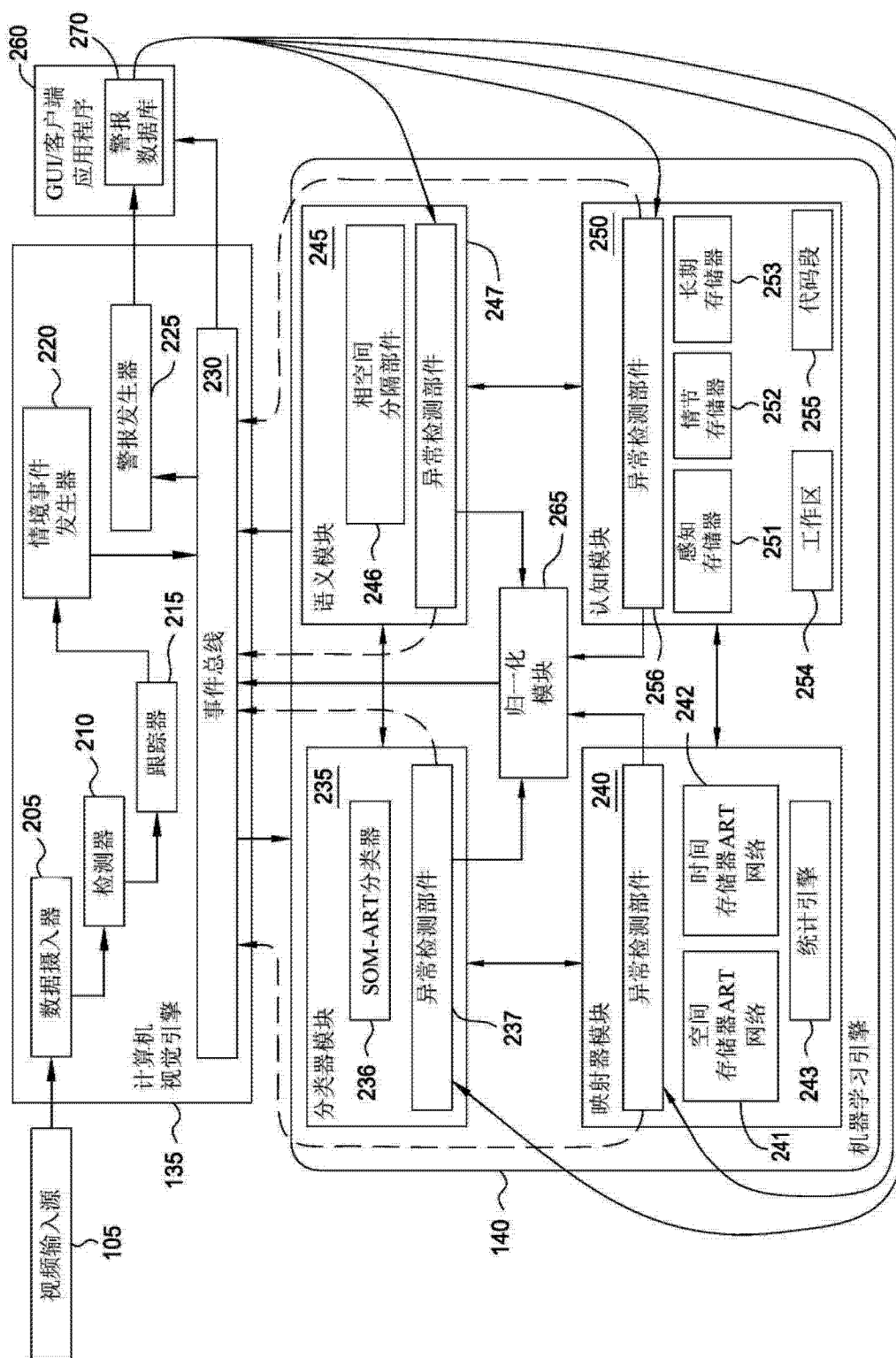


图 2

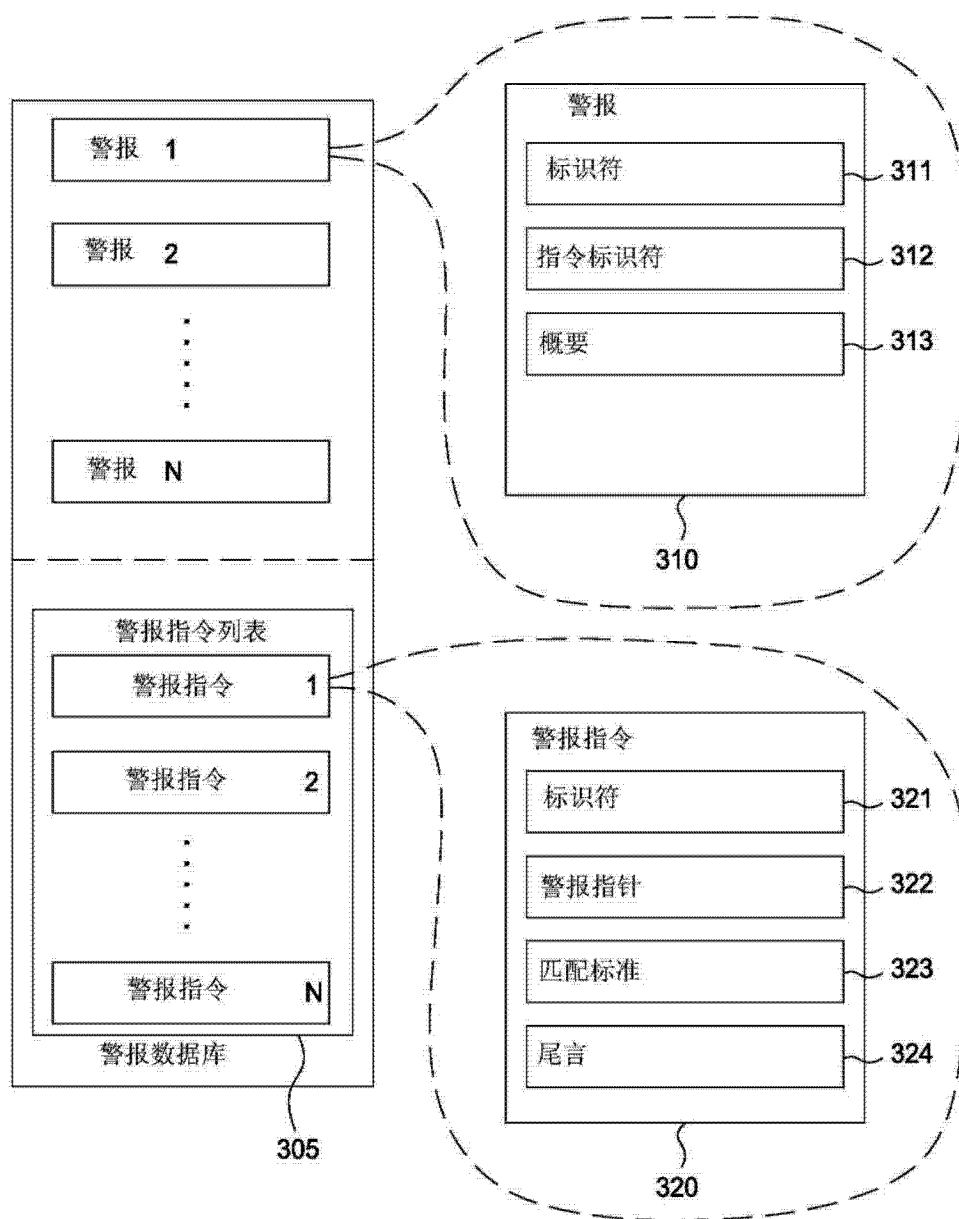


图 3

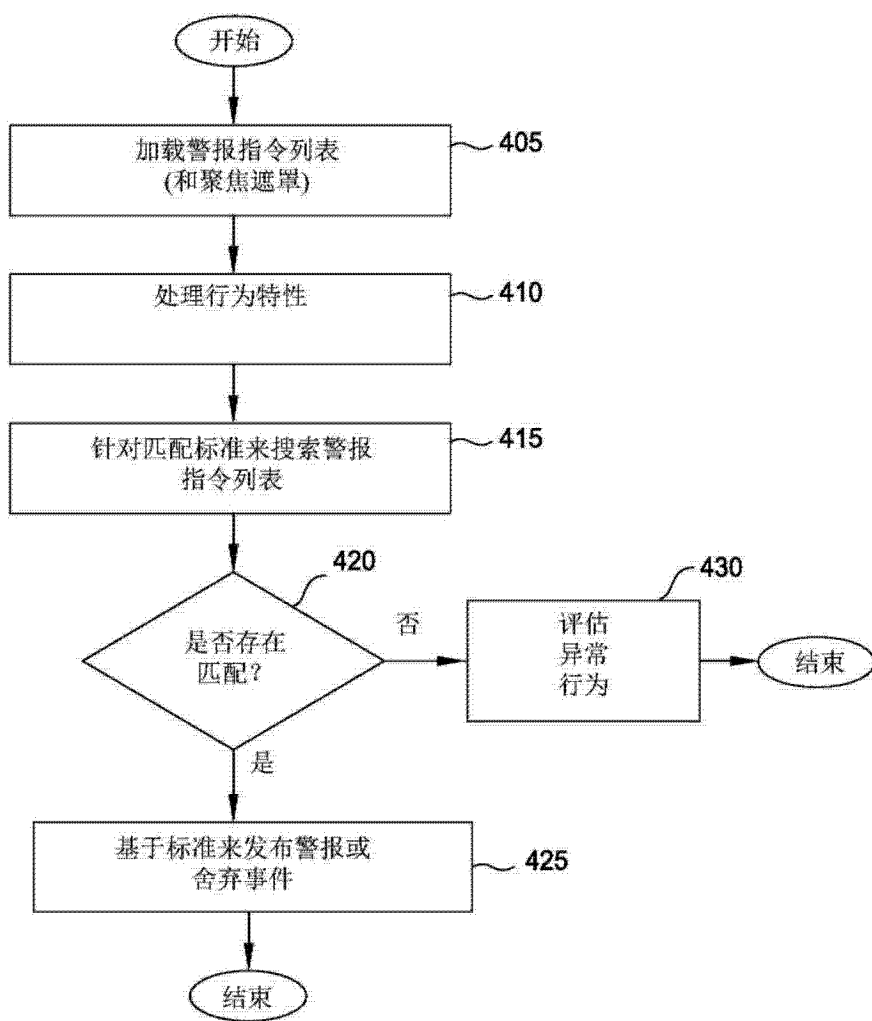


图 4

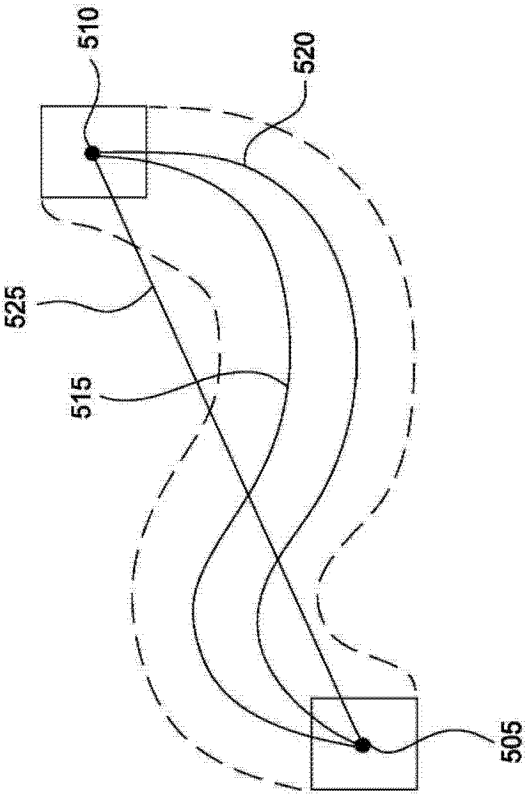


图 5

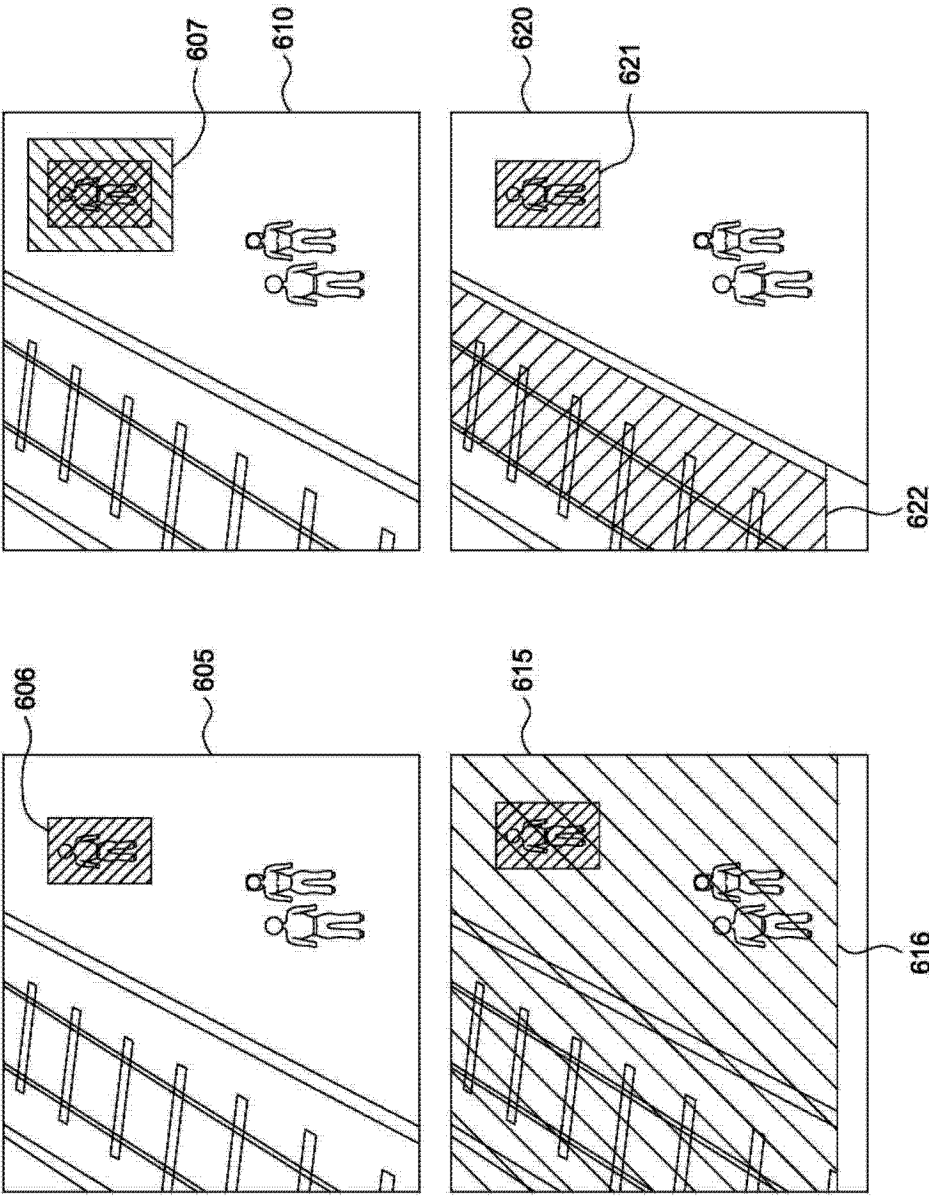


图 6