(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 105391990 A
(43)申请公布日 2016.03.09

(21)申请号 201510954472.5

(22)申请日 2006.04.05

(30)优先权数据
11/098,385 2005.04.05 US

(62)分案原申请数据
200680019911.2 2006.04.05

(71)申请人 威智伦富智堡公司
地址 加拿大不列颠哥伦比亚省温哥华

(72)发明人 彼得·L·韦奈蒂阿奈尔
艾伦·J·利普顿 安德鲁·J·肖克
马休·F·弗拉吉尔 尼尔斯·黑林
普瑞·梅耶斯 尹卫红 张忠

(74)专利代理机构 中科专利商标代理有限责任公司 11021
代理人 杨静

(51)Int.Cl.
H04N 7/18(2006.01)
H04N 21/234(2011.01)
H04N 21/44(2011.01)
G06F 17/30(2006.01)
G06K 9/00(2006.01)
G06T 7/20(2006.01)
G08B 13/196(2006.01)
H04N 5/272(2006.01)

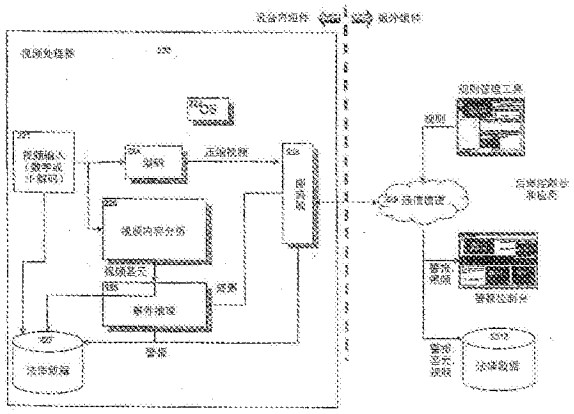
权利要求书6页 说明书20页 附图18页

(54)发明名称

采用视频基元的视频监视系统

(57)摘要

一种视频监视系统，提取视频基元，并使用事件鉴别器来从视频基元中提取事件发生。该系统可基于所提取的事件发生，做出响应，例如警报。



CN 105391990 A

1. 一种视频系统，包括：

第一处理器，所述第一处理器对视频进行分析以确定在所述视频中检测到的对象的属性，所述第一处理器与第一通信链路进行通信以通过所述通信链路传送所确定的属性；以及

第二处理器，所述第二处理器与所述第一处理器分离，所述第二处理器与所述第一通信链路进行通信以接收通过所述第一通信链路从所述第一处理器传送的所确定的属性，所述第二处理器通过对接收到的所确定的属性的组合进行分析来确定不是所确定的属性之一的第一事件，并且响应于所述第一事件的确定，提供以下至少一项：针对用户的提醒、报告的信息、以及用于采取行动的指令，

其中，所述第一处理器独立于所述第二处理器对所述第一事件的选择来确定属性，以及

所述第二处理器在无需对所述第一处理器分析的视频进行重新处理的情况下确定所述第一事件。

2. 根据权利要求 1 所述的视频系统，其中，所述第一通信链路包括网络。

3. 根据权利要求 1 所述的视频系统，其中，所述第二处理器通过仅对通过所述通信链路传送的属性进行分析来确定所述第一事件。

4. 根据权利要求 1 所述的视频系统，其中，所述第二处理器对所述属性进行分析以实时地检测所述第一事件。

5. 一种视频系统，包括：

输入端，所述输入端与通信信道进行通信；

处理器，被配置为从所述输入端接收通过所述通信信道接收的检测属性流，所述属性是在视频中检测到的一个或多个对象的属性，所述处理器被配置为通过对所接收的属性的组合进行分析来确定不是检测属性之一的事件，并且被配置为在确定所述事件时提供以下至少一项：针对用户的提醒、报告的信息、以及用于采取行动的指令，

其中，通过所述通信信道接收到的属性独立于要由所述处理器确定的事件；以及

所述处理器被配置为在无需对所述视频进行重新处理的情况下确定所述事件。

6. 根据权利要求 5 所述的视频系统，其中，所述通信信道包括网络。

7. 根据权利要求 5 所述的视频系统，其中，所述处理器能够操作以通过仅对所接收的属性流的属性进行分析来确定事件。

8. 一种根据视频来检测事件的方法，包括：

通过通信信道接收检测属性流，检测属性表示先前在远端位置处在所述视频中检测到的对象的属性；

对所述检测属性的组合进行分析，以在无需对所述视频进行重新处理的情况下检测不是所述检测属性之一的事件；

在检测到所述事件时，提供以下至少一项：针对用户的提醒、报告的信息、以及用于采取行动的指令，

其中，在所述属性流中接收到的检测属性独立于对要检测的事件的选择。

9. 根据权利要求 8 所述的方法，其中，所述通信信道包括网络。

10. 根据权利要求 8 所述的方法，其中，被执行以检测事件的分析通过仅对在所述检测

属性流中接收到的属性进行分析来确定事件。

11. 一种方法，包括：

对视频进行分析以检测对象；

在第一位置处通过对所述视频进行分析以确定检测到的对象的属性来创建属性流；

向第二位置发送从所述第一位置移除的所述属性流以进行后续分析，

其中，所述属性流是通过通信信道发送到所述第二位置的；以及

所述属性流足以允许所述后续分析以检测所述视频的事件从而提供以下至少一项：针对用户的提醒、报告的信息、以及用于采取行动的指令，所述事件不是所确定的属性之一，

所述属性流足以允许在无需对所述第一位置的视频进行重新处理的情况下检测不是所确定的属性之一的事件。

12. 根据权利要求 11 所述的方法，还包括：

使用视频捕获装置来获得所述视频。

13. 根据权利要求 11 所述的方法，其中，所述通信信道包括网络。

14. 根据权利要求 11 所述的方法，其中，所述属性流的属性是独立于所述后续分析来创建的。

15. 根据权利要求 11 所述的方法，其中，所述属性流足以允许通过对所述属性的组合进行分析来检测不是所确定的属性之一的事件。

16. 根据权利要求 11 所述的方法，其中，所述属性流是在无需在所述第一位置处检测事件的情况下通过通信信道发送的。

17. 一种视频设备，包括：

第一位置处的处理器，所述处理器对视频进行分析以检测对象并且确定在所述视频中检测到的所述对象的属性；

输出端，被配置为通过通信链路发送由所述处理器确定的属性，

其中，所述输出端被配置为向第二位置发送从所述处理器移除的属性以用于在所述第二位置处对所述属性的组合进行后续分析，

所述处理器独立于对属性的组合的后续分析以确定不是所确定的属性之一的事件来确定属性，以及

所述属性足以允许检测事件以提供以下至少一项：针对用户的提醒、报告的信息、以及用于采取行动的指令，所述事件不是所确定的属性之一并且能够通过对所述属性的组合进行分析来确定，

所述属性足以允许在无需对所述第一位置的所述视频进行重新处理的情况下检测事件。

18. 根据权利要求 17 所述的视频设备，还包括：

视频捕获装置，用于向所述处理器提供所述视频。

19. 根据权利要求 17 所述的视频设备，其中，所述输出端通过所述通信链路发送检测属性流。

20. 根据权利要求 19 所述的视频设备，其中，所述通信链路包括网络。

21. 根据权利要求 17 所述的视频设备，其中，所述属性是在无需由所述处理器对事件进行检测的情况下通过所述通信信道发送的。

22. 一种视频系统，包括：

第一处理器，所述第一处理器对视频进行分析以确定在所述视频中检测到的对象的属性，所述第一处理器与第一通信链路进行通信以通过所述通信链路传送所确定的属性；以及

第二处理器，所述第二处理器与所述第一处理器分离，所述第二处理器与所述第一通信链路进行通信以接收通过所述第一通信链路从所述第一处理器传送的所确定的属性，所述第二处理器通过对接收到的所确定的属性的组合进行分析来确定不是所确定的属性之一的第一事件，并且响应于所述第一事件的确定，提供以下至少一项：针对用户的提醒、报告的信息、以及用于采取行动的指令，其中，对接收到的所确定的属性的组合进行分析包括滤波，

其中，所述第一处理器独立于所述第二处理器对所述第一事件的选择来确定属性，以及

所述第二处理器在无需对所述第一处理器分析的视频进行重新处理的情况下确定所述第一事件。

23. 根据权利要求 22 所述的视频系统，其中，所述滤波包括：对接收到的所确定的属性的组合进行滤波以确定是否发生所述第一事件。

24. 一种视频系统，包括：

输入端，所述输入端与通信信道进行通信；

处理器，被配置为从所述输入端接收通过所述通信信道接收的检测属性流，所述属性是在视频中检测到的一个或多个对象的属性，所述处理器被配置为通过对所接收的属性的组合进行分析来确定不是检测属性之一的事件，并且被配置为在确定所述事件时提供以下至少一项：针对用户的提醒、报告的信息、以及用于采取行动的指令，其中，对接收到的所确定的属性的组合进行分析包括滤波，

其中，通过所述通信信道接收到的属性独立于要由所述处理器确定的事件；以及

所述处理器被配置为在无需对所述视频进行重新处理的情况下确定所述事件。

25. 根据权利要求 24 所述的视频系统，其中，所述滤波包括：对接收到的属性的组合进行滤波以确定是否发生所述事件。

26. 一种根据视频检测事件的方法，包括：

通过通信信道接收检测属性流，检测属性表示先前在远端位置处在所述视频中检测到的对象的属性；

对所述检测属性的组合进行分析，以在无需对所述视频进行重新处理的情况下检测不是所述检测属性之一的事件，其中，对所述检测属性的组合进行分析包括滤波；

在检测到所述事件时，提供以下至少一项：针对用户的提醒、报告的信息、以及用于采取行动的指令，

其中，在所述属性流中接收到的检测属性独立于对要检测的事件的选择。

27. 根据权利要求 26 所述的方法，其中，所述滤波包括：对所述检测属性的组合进行滤波以确定是否发生所述事件。

28. 一种视频系统，包括：

第一处理器，所述第一处理器对视频进行分析以确定在所述视频中检测到的对象的属

性，所述第一处理器与第一通信链路进行通信以通过所述通信链路传送所确定的属性；以及

第二处理器，所述第二处理器与所述第一处理器分离，所述第二处理器与所述第一通信链路进行通信以接收通过所述第一通信链路从所述第一处理器传送的所确定的属性，所述第二处理器通过对接收到的所确定的属性的组合进行分析来确定不是所确定的属性之一的第一事件，并且响应于所述第一事件的确定，提供以下至少一项：针对用户的提醒、报告的信息、以及用于采取行动的指令，

其中，在所述视频中检测到的对象的属性包括第一对象和第二对象，并且所述第一事件是所述第一对象和所述第二对象一起到来，

所述第一处理器独立于所述第二处理器对所述第一事件的选择来确定属性，以及

所述第二处理器在无需对所述第一处理器分析的视频进行重新处理的情况下确定所述第一事件。

29. 一种视频系统，包括：

输入端，所述输入端与通信信道进行通信；

处理器，被配置为从所述输入端接收通过所述通信信道接收的检测属性流，所述属性是在视频中检测到的第一对象和第二对象的属性，所述处理器被配置为通过对所接收的属性的组合进行分析来确定不是检测属性之一的事件，并且被配置为在确定所述事件时提供以下至少一项：针对用户的提醒、报告的信息、以及用于采取行动的指令，所述事件是所述第一对象和所述第二对象一起到来，

其中，通过所述通信信道接收到的属性独立于要由所述处理器确定的事件；以及

所述处理器被配置为在无需对所述视频进行重新处理的情况下确定所述事件。

30. 一种根据视频来检测事件的方法，包括：

通过通信信道接收检测属性流，检测属性表示先前在远端位置处在所述视频中检测到的第一对象和第二对象的属性；

对所述检测属性的组合进行分析，以在无需对所述视频进行重新处理的情况下检测不是所述检测属性之一的事件，所述事件是所述第一对象和所述第二对象一起到来；

在检测到所述事件时，提供以下至少一项：针对用户的提醒、报告的信息、以及用于采取行动的指令，

其中，在所述属性流中接收到的检测属性独立于对要检测的事件的选择。

31. 一种方法，包括：

对视频进行分析以检测第一对象和第二对象；

在第一位置处通过对所述视频进行分析以确定检测到的第一对象的属性和检测到的第二对象的属性来创建属性流；

向第二位置发送从所述第一位置移除的所述属性流以进行后续分析，

其中，所述属性流是通过通信信道发送到所述第二位置的；以及

所述属性流足以允许所述后续分析以检测所述视频的事件从而提供以下至少一项：针对用户的提醒、报告的信息、以及用于采取行动的指令，所述事件不是所确定的属性之一，所述事件是所述第一对象和所述第二对象一起到来，

所述属性流足以允许在无需对所述第一位置的视频进行重新处理的情况下检测不是

所确定的属性之一的事件。

32. 一种视频设备，包括：

第一位置处的处理器，所述处理器对视频进行分析以检测第一对象和第二对象并且确定在所述视频中检测到的所述第一对象的属性和在所述视频中检测到的所述第二对象的属性；

输出端，被配置为通过通信链路发送由所述处理器确定的属性，

其中，所述输出端被配置为向第二位置发送从所述处理器移除的属性以用于在所述第二位置处对所述属性的组合进行后续分析，

所述处理器独立于对属性的组合的后续分析以确定不是所确定的属性之一的事件来确定属性，所述事件是所述第一对象和所述第二对象一起到来，以及

所述属性足以允许检测事件以提供以下至少一项：针对用户的提醒、报告的信息、以及用于采取行动的指令，所述事件不是所确定的属性之一并且能够通过对所述属性的组合进行分析来确定，

所述属性足以允许在无需对所述第一位置的所述视频进行重新处理的情况下检测事件。

33. 一种视频系统，包括：

第一处理器，所述第一处理器对视频进行分析以确定在所述视频中检测到的对象的属性，所述第一处理器与第一通信链路进行通信以通过所述通信链路传送所确定的属性；以及

第二处理器，所述第二处理器与所述第一处理器分离，所述第二处理器与所述第一通信链路进行通信以接收通过所述第一通信链路从所述第一处理器传送的所确定的属性，所述第二处理器通过对接收到的所确定的属性的组合进行分析来确定不是所确定的属性之一的第一事件，并且响应于所述第一事件的确定，提供以下至少一项：针对用户的提醒、报告的信息、以及用于采取行动的指令，其中，对接收到的所确定的属性的组合进行分析包括滤波，

其中，在所述视频中检测到的对象的属性包括第一对象和第二对象，并且所述第一事件是所述第一对象和所述第二对象一起到来，

所述第一处理器独立于所述第二处理器对所述第一事件的选择来确定属性，以及

所述第二处理器在无需对所述第一处理器分析的视频进行重新处理的情况下确定所述第一事件。

34. 一种视频系统，包括：

输入端，所述输入端与通信信道进行通信；

处理器，被配置为从所述输入端接收通过所述通信信道接收的检测属性流，所述属性是在视频中检测到的第一对象和第二对象的属性，所述处理器被配置为通过对所接收的属性的组合进行分析来确定不是检测属性之一的事件，并且被配置为在确定所述事件时提供以下至少一项：针对用户的提醒、报告的信息、以及用于采取行动的指令，所述事件是所述第一对象和所述第二对象一起到来，其中，对接收到的属性的组合进行分析包括滤波，

其中，通过所述通信信道接收到的属性独立于要由所述处理器确定的事件；以及

所述处理器被配置为在无需对所述视频进行重新处理的情况下确定所述事件。

35. 一种根据视频来检测事件的方法，包括：

通过通信信道接收检测属性流，检测属性表示先前在远端位置处在所述视频中检测到的第一对象和第二对象的属性；

对所述检测属性的组合进行分析，以在无需对所述视频进行重新处理的情况下检测不是所述检测属性之一的事件，其中，对所述检测属性的组合进行分析包括滤波，所述事件是所述第一对象和所述第二对象一起到来；

在检测到所述事件时，提供以下至少一项：针对用户的提醒、报告的信息、以及用于采取行动的指令，

其中，在所述属性流中接收到的检测属性独立于对要检测的事件的选择。

采用视频基元的视频监视系统

技术领域

[0001] 本发明涉及一种采用视频基元的自动视频监视系统。

[0002] 参考文献

[0003] 为了方便读者，下面列出在此引用的参考文献。在说明书中，括号内的数字表示各个参考文献。将所列的参考文献合并在此以作参考。

[0004] 下面的参考文献描述了运动目标检测：

[0005] {1}A.Lipton，H.Fujiyoshi 和 R.S.Patil，″ Moving Target Detection and Classification from Real-Time Video，″ Proceedings of IEEE WACV′ 98.Princeton，NJ，1998，pp.8-14.

[0006] {2}W.E.L.Grimson，et al.，″ Using Adaptive Tracking to Classify and Monitor Activities in a Site″，CVPR.pp.22-29，June 1998.

[0007] {3}AJ.Lipton，H.Fujiyoshi，R.S.Patil，″ Moving Target Classification and Tracking from Real-time Video.″ IUW.pp.129-136，1998.

[0008] {4}TJ.Olson 和 F.Z.Brill，″ Moving Object Detection and Event Recognition Algorithm for Smart Cameras，″ IUW.pp.159-175，May1997.

[0009] 下面的参考文献描述了检测和人物跟踪：

[0010] {5}A.J.Lipton，″ Local Application of Optical Flow to Analyze Rigid Versus Non-Rigid Motion，″ International Conference on Computer Vision.Corfu，Greece，September 1999.

[0011] {6}F.Bartolini，V.Cappellini，和 A.Mecocci，″ Counting people getting in and out of a bus by real-time image-sequence processing，″ IVC.12(1)：36-41，January 1994.

[0012] {7}M.Rossi 和 A.Bozzoli，″ Tracking and counting moving people，″ ICIP94.pp.212-216，1994；

[0013] {8}CR.Wren，A.Azarbayejani，T.Darrell，和 A.Pentland，″ Pfinder：Real-time tracking of the human body，″ Vismod.1995.

[0014] {9}L.Khoudour，L.Duvieubourg，J.P.Deparis，″ Real-Time Pedestrian Counting by Active Linear Cameras，″ JEI，5(4)：452-459，October 1996.

[0015] {10}S.Ioffe，D.A.Forsyth，″ Probabilistic Methods for Finding People，″ IJCV.43(1)：45-68，June 2001.

[0016] {11}M.Isard 和 J.MacCormick，″ BraMBLe：A Bayesian Multiple-Blob Tracker，″ ICCV.2001.The following references describe blob analysis：

[0017] 下面的参考文献描述了团迹(blob)分析：

[0018] {12}D.M.Gavrila，″ The Visual Analysis of Human Movement：A Survey，″ CVIU.73(1)：82-98，January 1999.

[0019] {13}Niels Haering 和 Niels da Vitoria Lobo，″ Visual Event

Detection, " Video Computing Series,Editor Mubarak Shah,2001.

[0020] 下面的参考文献描述了针对卡车、汽车和人物的团迹分析：

[0021] {14}Collins,Lipton,Kanade,Fujiyoshi,Duggins,Tsin,Tolliver,Enomoto,和Hasegawa, " A System for Video Surveillance and Monitoring:V SAM Final Report, " Technical Report CMU-RI-TR-00-12,Robotics Institute,Carnegie Mellon University,May 2000.

[0022] {15}Lipton,Fujiyoshi,和Patil, " Moving Target Classification and Tracking from Real-time Video, " 98 Darpa IUW.Nov.20-23,1998.

[0023] 下面的参考文献描述了对单人的团迹及其轮廓的分析：

[0024] {16}CR.Wren,A.Azarbayejani,T.Darrell,和A.P.Pentland." Pf[iota]nder：Real-Time Tracking of the Human Body, " PAMI.vol 19,pp.780-784,1997.

[0025] 下面的参考文献描述了团迹的内部运动,包括任何基于运动的分割：

[0026] {17}M.Allmen和C.Dyer," Long-Range Spatiotemporal Motion Understanding Using Spatiotemporal Flow Curves, " Proc.IEEE CVPR.Lahaina,Maui,Hawaii,pp.303-309,1991.

[0027] {18}L.Wixson," Detecting Salient Motion by Accumulating Directionally Consistent Flow",IEEE Trans.Pattern Anal.Mach.Intell.,vol.22,pp.774-781,Aug,2000.

背景技术

[0028] 公共场所的视频监视变得非常普遍,并被普通大众所接受。遗憾的是,传统的视频监视系统产生非常大量的数据,从而在视频监视数据的分析中造成难以处理的问题。

[0029] 需要减少视频监视数据量,以便可以进行视频监视数据的分析。

[0030] 需要过滤视频监视数据,以便识别出视频监视数据中需要的部分。

发明内容

[0031] 本发明的目的是减少视频监视数据量,以便可以进行视频监视数据的分析。

[0032] 本发明的目的是过滤视频监视数据,以便识别出视频监视数据中需要的部分。

[0033] 本发明的目的是根据视频监视数据,基于对事件的自动检测,产生实时警报。

[0034] 本发明的目的是整合来自监视传感器的数据而不是视频,以提高搜索能力。

[0035] 本发明的目的是整合来自监视传感器的数据而不是视频,以提高事件检测能力。

[0036] 本发明包括用于视频监视的制造产品、方法、系统和设备。

[0037] 本发明的制造产品包括计算机可读介质,计算机可读介质包括用于视频监视系统的软件,并包括用于基于视频基元来操作视频监视系统的代码段。

[0038] 本发明的制造产品包括计算机可读介质,计算机可读介质包括用于视频监视系统的软件,并包括用于访问存档的视频基元的代码段以及用于从所访问的存档视频基元中提取事件发生的代码段。

[0039] 本发明的系统包括计算机系统,计算机系统包括计算机可读介质,计算机可读介质具有软件,用于根据本发明来操作计算机。

[0040] 本发明的设备包括计算机，计算机包括计算机可读介质，计算机可读介质具有软件，用于根据本发明来操作计算机。

[0041] 本发明的制造产品包括计算机可读介质，计算机可读介质具有软件，用于根据本发明来操作计算机。

[0042] 此外，本发明的上述目的和优点是本发明可实现的目的和优点的示例，而不是穷尽列举的。因此，根据这里的说明，这里所体现的以及考虑到对于本领域技术人员而言显而易见的任意改变而做出的修改的本发明的这些和其它目的和优点将是显然的。

[0043] 定义

[0044] “视频”是指以模拟和 / 或数字形式呈现的运动图像。视频的示例包括 :电视、电影、来自摄像机或其它观察器的图像序列以及计算机产生的图像序列。

[0045] “帧”是指视频内的特定图像或其它离散单元。

[0046] “对象”是指视频中感兴趣的项目。对象的示例包括 :人、车辆、动物和物理对象。

[0047] “活动”是指一个或多个对象的一个或多个动作和 / 或一个或多个动作的一个或多个合成。活动的示例包括 :进入、退出、停止、移动、上升、下降、膨胀和收缩。

[0048] “位置”是指活动可能发生的空间。例如，位置可以是基于场景的或基于图像的。基于场景的位置的示例包括 :公共空间、商店、零售商店、办公室、仓库、酒店房间、酒店大厅、建筑物大厅、娱乐场、公共汽车站、火车站、飞机场、港口、公共汽车、火车、飞机以及轮船。基于图像的位置的示例包括 :视频图像 ;视频图像中的行 ;视频图像中的区域 ;视频图像的矩形部分 ;视频图像的多边形部分。

[0049] “事件”是指活动所涉及的一个或多个对象。可以相对于位置和 / 或时间而提及事件。

[0050] “计算机”是指能够接受结构化输入、根据预定规则来处理结构化输入、并产生处理结果作为输出的任意设备。计算机的示例包括 :计算机、通用计算机、超型计算机、大型机、超小型计算机、小型计算机、工作站、微型计算机、服务器、交互式电视、计算机和交互式电视的混合组合、以及对计算机和 / 或软件进行仿真的专用硬件。计算机可具有单个处理器或多个处理器，这些处理器可并行操作也可不并行操作。计算机还指通过网络连接在一起、用于在计算机之间发送或接收信息的两个或多个计算机。这种计算机的示例包括分布式计算机系统，用于经由通过网络链接的计算机来处理信息。

[0051] “计算机可读介质”是指用于存储可由计算机访问的数据的任何存储设备。计算机可读介质的示例包括 :磁硬盘、软盘、例如 CD-ROM 和 DVD 的光盘、磁带、存储芯片、以及用于载有计算机可读电信号的载波，例如用在发送和接收电子邮件或访问网络中的载波。

[0052] “软件”是指操作计算机的预定规则。软件的示例包括 :软件、代码段、指令、计算机程序和可编程逻辑。

[0053] “计算机系统”是指具有计算机的系统，其中计算机包括包含有操作计算机的软件的计算机可读介质。

[0054] “网络”是指通过通信设施连接的多个计算机和相关的设备。网络包括例如电缆的永久性连接和例如通过电话或其它通信链路而进行的暂时性连接。网络的示例包括 :例如因特网的互联网、内联网、局域网 (LAN)、广域网 (WAN) 以及例如互联网和内联网等网络的组合。

附图说明

[0055] 通过附图来更加详细地解释本发明的实施例，附图中相同的附图标记指代相同的特征。

[0056] 图 1 示出了本发明的视频监视系统的平面图。

[0057] 图 2 示出了本发明的视频监视系统的流程图。

[0058] 图 3 示出了用于给视频监视系统分派任务的流程图。

[0059] 图 4 示出了用于操作视频监视系统的流程图。

[0060] 图 5 示出了用于提取视频监视系统的视频基元的流程图。

[0061] 图 6 示出了对视频监视系统采取行动的流程图。

[0062] 图 7 示出了视频监视系统的半自动校准的流程图。

[0063] 图 8 示出了视频监视系统的自动校准的流程图。

[0064] 图 9 示出了本发明的视频监视系统的附加流程图。

[0065] 图 10-15 示出了本发明的视频监视系统用于监视杂货商店的示例。

[0066] 图 16a 示出了根据本发明实施例的视频分析子系统的流程图。

[0067] 图 16b 示出了根据本发明实施例的事件发生检测和响应子系统的流程图。

[0068] 图 17 示出了示范性的数据库查询。

[0069] 图 18 示出了根据本发明多个实施例的三个示范性活动检测器 :检测横穿绊网 (图 18a)、闲逛 (图 18b) 和偷窃 (图 18c)。

[0070] 图 19 示出了根据本发明实施例的活动检测器查询。

[0071] 图 20 示出了根据本发明实施例的使用活动检测器以及具有修饰符的布尔算子的示范性查询。

[0072] 图 21a 和 21b 示出了使用多级组合算符 (combinator)、活动检测器以及属性查询的示范性查询。

[0073] 图 22 示出了根据本发明实施例的视频监视系统的示范性配置。

[0074] 图 23 示出了根据本发明实施例的视频监视系统的另一示范性配置。

[0075] 图 24 示出了根据本发明实施例的视频监视系统的另一示范性配置。

[0076] 图 25 示出了可在本发明实施例的示范性配置中使用的网络。

具体实施方式

[0077] 本发明的自动视频监视系统用于地点监视，以用于例如市场研究或安全目的。该系统可以是安装有为特定目的构造的监视组件的专用视频监视装置，或者该系统可以是对现有背负 (piggyback off) 监视视频馈给的视频监视装置的改进。该系统能够分析来自现场源或来自记录媒体的视频数据。该系统能够实时处理视频数据，并存储所提取的视频基元以允许以后进行非常高速的法律 (forensic) 事件检测。该系统可具有针对该分析的规定响应，例如 :记录数据，激活警报机制，或者激活另一个传感器系统。该系统还能够集成有其它监视系统组件。该系统可被用于制作例如可根据操作者的需要而定制的安全或市场研究报告，以及可选地，可通过基于交互式网络的界面或者其它报告机制而提供的报告。

[0078] 通过使用事件鉴别器来配置系统，给操作者提供最大灵活性。事件鉴别器由一个

或多个对象（其描述是基于视频基元的）、以及一个或多个可选空间属性和/或一个或多个可选时间属性来标识。例如，操作者可将事件鉴别器（在该示例中被称为“闲逛”事件）定义为“在 10 :00p.m. 和 6 :00a.m. 之间”“人物”对象在“自动售货机”空间“长于 15 分钟”。事件鉴别器可与修饰的布尔算子组合以形成更复杂的查询。

[0079] 尽管本发明的视频监视系统利用来自公共领域的已知的计算机图像技术，但是本发明的视频监视系统具有当前并不可用的多个独特且新颖的特征。例如，当前的视频监视系统将大量视频图像用作信息交换的初级产品。本发明的系统将视频基元用作初级产品，而将代表性视频图像用作附属证据。可以（手动、半自动或自动）校准本发明的系统，从而本发明的系统可自动地从视频图像中推断出视频基元。该系统还可分析先前处理的视频，而不需要完全重新处理视频。通过分析先前处理的视频，该系统可基于先前记录的视频基元来执行推理分析，这极大地提高了计算机系统的分析速度。

[0080] 视频基元的使用还可显著地降低对视频的存储需求。这是因为事件检测和响应子系统仅使用视频来对检测进行示例。因此，可以以低质量来存储或传输视频。在可能的实施例中，仅在检测到活动时才存储或传输视频，而不是总在存储或传输视频。在另一可能的实施例中，所存储或传输的视频的质量可取决于是否检测到活动：在检测到活动时以较高质量（更高的帧频和/或比特率）来存储或传输视频，并且在其它时候以低质量来存储或传输视频。在另一示范性实施例中，例如，可由数字视频记录器 (DVR) 独立地处理视频存储和数据库，并且视频处理子系统可仅控制是否且以何种质量来存储数据。在另一实施例中，视频监视系统（或其组件）可以处于例如数字摄像机、网络视频服务器、DVR 或网络视频记录器 (NVR) 的视频管理设备上载有的处理设备（例如通用处理器、DSP、微控制器、ASIC、FPGA 或其它设备）上，并且系统可对来自设备的视频的带宽进行调制。仅在检测到感兴趣的活动时，才通过 IP 视频网络传输高质量视频（高比特率或帧频）。在该实施例中，可通过网络将来自智能化设备的基元广播到位于不同物理位置处的多个活动推理应用，以使单摄像机网络能够通过分布式处理来提供多用途的应用。

[0081] 图 22 示出了实现视频监视系统的一种配置。块 221 代表原始（未压缩的）数字视频输入。例如，这可通过模拟视频信号的模 - 数捕获或数字视频信号的解码而获得。块 222 代表容纳视频监视系统的主要组件（视频分析块 225 和活动推理块 226）的硬件平台。该硬件平台可包含其它组件，例如：操作系统（块 223）；视频编码器（块 224），用于使用任何可用压缩方案 (JPEG、MJPEG、MPEG1、MPEG2、MPEG4、H.263、H.264、小波或其它）来压缩原始数字视频以便流传输或存储视频；存储机制（块 227），用于保持例如视频、压缩视频、警报和视频基元的数据，该存储设备可以是例如硬盘、板上 RAM、板上 FLASH 存储器或其它存储介质；以及通信层（块 228），例如可以将数据分组和/或数字化，以在通信信道（块 229）上传输。

[0082] 可能有其它软件组件驻留在与通信信道 229 连接的网络的其它节点处的计算平台上。块 2210 示出了用作创建视频监视规则的用户接口的规则管理工具。块 2211 示出了用于向用户显示警报并向用户做出报告的警报控制台。块 2212 示出了用于存储警报、基元和视频以用于进一步的事后处理的存储设备（例如 DVR、NVR 或 PC）。

[0083] 硬件平台（块 222）上的组件可在例如摄像机、数码摄像机、IP 摄像机、IP 视频服务器、数字视频记录器 (DVR)、网络视频记录器 (NVR)、PC、膝上型计算机或其它设备的任何

视频捕获、处理或管理设备上的任何处理硬件（通用处理器、微控制器、DSP、ASIC、FPGA 或其它处理设备）上实现。对于该配置，存在多种不同的可能的操作模式。

[0084] 在一种模式下，系统被编程为查找特定事件。在发生这些事件时，通过通信信道（块 229）将警报传输给其它系统。

[0085] 在另一种模式下，在视频设备对视频数据进行分析的同时从视频设备中流传输视频。当发生事件时，通过通信信道（块 229）来传输警报。

[0086] 在另一种模式下，通过内容分析和活动推理来对视频编码和流传输进行调制。当没有活动（没有产生基元）时，没有流传输的视频（或者低质量、比特率、帧频、分辨率）。当存在某种活动（产生了基元）时，流传输高质量、比特率、分辨率的视频。当通过事件推理检测到感兴趣的事件时，流传输非常高的质量、比特率、帧频、分辨率的视频。

[0087] 在另一种操作模式下，将信息存储在板上存储设备（块 227）中。所存储的数据可包括数字视频（原始或压缩的）、视频基元、警报或其它信息。所存储的视频质量也可由基元或警报的存在与否来控制。当存在基元和警报时，可存储高质量、比特率、帧频、分辨率的视频。

[0088] 图 23 示出了实现视频监视系统的另一种配置。块 231 代表原始（未压缩的）数字视频输入。例如，这可通过模拟视频信号的模 - 数捕获或数字视频信号的解码而获得。块 232 代表容纳视频监视系统的主要组件（块 235）的硬件平台。该硬件平台可包含其它组件，例如：操作系统（块 233）；视频编码器（块 234），用于使用任何可用压缩方案（JPEG、MJPEG、MPEG1、MPEG2、MPEG4、H.263、H.264、小波或其它）来压缩原始数字视频以便流传输或存储视频；存储机制（块 236），用于保持例如视频、压缩视频、警报和视频基元的数据，该存储设备可以是例如硬盘、板上 RAM、板上 FLASH 存储器或其它存储介质；以及通信层（块 237），例如可以将数据分组和 / 或数字化，以在通信信道（块 238）上传输。在图 23 所示的本发明实施例中，活动推理组件（块 2311）被示出为在独立的硬件组件（块 239）上，该独立的硬件组件与通信信道 238 所连接的网络相连。

[0089] 可能有其它软件组件驻留在该网络的其它节点（块 239）处的计算平台上。块 2310 示出了用作创建视频监视规则的用户接口的规则管理工具。块 2312 示出了用于向用户显示警报并向用户做出报告的警报控制台。块 2313 示出了可位于物理上相同硬件平台上的存储设备（例如硬盘、软盘、其它磁盘、CD、DVD、其它光盘、MD 或其它磁光盘、例如 RAM 或 FLASH RAM 的固态存储设备或其它存储设备），或者块 2313 也可以是独立的存储设备（例如外部磁盘驱动器、PC、膝上型计算机、DVR、NVR 或其它存储设备）。

[0090] 硬件平台（块 222）上的组件可在例如摄像机、数码摄像机、IP 摄像机、IP 视频服务器、数字视频记录器 (DVR)、网络视频记录器 (NVR)、PC、膝上型计算机或其它设备的任何视频捕获、处理或管理设备上的任何处理平台（通用处理器、微控制器、DSP、ASIC、FPGA 或其它处理设备）上实现。后端硬件平台（块 239）上的组件可在例如 PC、膝上型计算机、单板计算机、DVR、NVR、视频服务器、网络路由器、手持设备（例如视频电话、寻呼机或 PDA）的任何处理设备上的任何处理硬件（通用处理器、微处理器、DSP、FPGA、ASIC、或任意其它设备）上实现。对于该配置，存在多种不同的可能的操作模式。

[0091] 在一种模式下，系统被编程在后端设备（或任意其它与后端设备相连的设备）上，用于查找特定事件。视频处理平台（块 232）上的内容分析模块（块 235）产生传输到后端

处理平台（块 239）的基元。事件推理模块（块 2311）确定是否违反了规则，并产生可显示在警报控制台（块 2312）上或存储在存储设备（2313）中的警报以用于后面的分析。

[0092] 在另一种模式下，可将视频基元和视频存储在后端平台（2313）上的存储设备中以用于后面的分析。

[0093] 在另一种模式下，可由警报来调制所存储的视频的质量、比特率、帧频、分辨率。当存在警报时，可以以较高的质量、比特率、帧频、分辨率来存储视频。

[0094] 在另一种模式下，可通过通信信道将视频基元存储在视频处理设备（块 232 中的块 236）上，以用于后面的分析。

[0095] 在另一种模式下，可由基元的存在与否来调制存储在视频处理设备（块 232 中的块 236）上的视频的质量。当存在基元时（当出现某些事时），可提高所存储的视频的质量、比特率、帧频、分辨率。

[0096] 在另一种模式下，可经由通信信道 238，通过编码器（234）从视频处理器向网络上的其它设备流传输视频。

[0097] 在另一种模式下，可由内容分析模块（235）来调制视频质量。当没有基元（没有发生什么事）时，不流传输（或低质量、比特率、帧频、分辨率）视频。当存在活动时，流传输高质量、比特率、帧频、分辨率的视频。

[0098] 在另一种模式下，可由警报的存在与否来对流传输的视频的质量、比特率、帧频、分辨率进行调制。当后端事件推理模块（块 2311）检测到感兴趣的事件时，其可向视频处理组件（块 232）发送信号或命令，以请求视频（或较高质量、比特率、帧频、分辨率的视频）。在接收到该请求时，视频压缩组件（块 234）和通信层（块 237）可改变压缩和流传输参数。

[0099] 在另一种模式下，可由警报的存在与否来对存储在板上视频处理设备（块 232 中的块 236）上的视频的质量进行调制。当后端处理器（块 239）上的事件推理模块（块 2311）产生警报时，其经由通信信道（块 238）将消息发送到视频处理器硬件（块 232），以提高存储在板上存储设备（238）中的视频的质量、比特率、帧频、分辨率。

[0100] 图 24 示出了图 23 所述的配置的扩展。通过将视频内容分析和后端活动推理的功能分离，可以通过后期应用绑定的过程来实现多用途智能化视频监视系统。智能化摄像机的单个网络可广播单个视频基元流，来分离组织中不同部分（处于不同物理位置处）的后端应用，并实现多种功能。由于基元流包含与场景中正在进行的所有事情有关的信息，并且并不依赖于特定的应用领域，所以这是可能的。图 24 所示的示例与零售环境相关，但是示出了一般原理，并且适用于任何其它应用领域和任何其它监视功能。块 241 示出了设施内或多个设施间的一个或多个摄像机的智能化网络。内容分析组件可驻留在与网络相连的摄像机内的处理设备、视频服务器、网络路由器、DVR、NVR、PC、膝上型计算机或任意其它视频处理设备上。经由标准网络，从这些内容分析组件将基元流广播到驻留在用于不同目的的不同物理区域中的后端处理器（块 242-245）上的活动推理模块。后端处理器可以在计算机、膝上型计算机、DVR、NVR、网络路由器、手持设备（电话、寻呼机、PDA）或其它计算设备上。这种分布式的一个优点在于，不需要必须被编程为进行所有可能应用的所有处理的中央处理应用。另一个优点是安全性，因此组织的一个部分可基于本地存储的规则进行活动推理，而网络中的其它部分不能访问该信息。

[0101] 在块 242 中，针对物理安全性应用而分析来自智能化摄像机网络的基元流，以便：

确定是否存在周边侵犯、破坏，并保持关键资产。当然，这些应用仅仅是示例性的，任何其它应用都是可能的。

[0102] 在块 243 中，针对丢失防止应用来分析来自智能化摄像机网络的基元流，以便：监控装载码头；监视顾客或雇员的偷窃行为，监控仓库，并跟踪库存。当然，这些应用仅仅是示例性的，任何其它应用都是可能的。

[0103] 在块 244 中，针对公共安全和责任应用来分析来自智能化摄像机网络的基元流，以便：监控停车场中移动过快的人或车辆，监视人的滑到，并监控设施内和设施周围的人群。当然，这些应用仅仅是示例性的，任何其它应用都是可能的。

[0104] 在块 245 中，针对商业智能应用来分析来自智能化摄像机网络的基元流，以便：观察队列长度，跟踪顾客行为，了解行为模式，执行例如在无人时控制照明和加热的建筑物管理任务。当然，这些应用仅仅是示例性的，任何其它应用都是可能的。

[0105] 图 25 示出了有多个可能的智能化设备与之相连的网络（块 251）。块 252 是载有内容分析组件的 IP 摄像机，可在网络上流传输基元。块 253 是载有内容分析和活动推理组件的 IP 摄像机，可直接对其进行规则编程，从而该块 253 直接产生网络警报。块 254 是不载有智能组件的标准模拟摄像机；但是，它与执行视频数字化和压缩以及内容分析和活动推理的 IP 视频管理平台（块 256）相连。可用专用于查看的规则来对其进行编程，并且该块 254 能够经由网络来传输基元流和警报。块 255 是具有活动推理组件的 DVR，它能够从其它设备汲取基元流并产生警报。块 257 是具有无线网络通信能力的手持 PDA，载有活动推理算法，并能够从网络接受视频基元并显示警报。块 258 是完整的智能化视频分析系统，其能够接受模拟或数字视频流、执行内容分析和活动推理，并在一系列警报控制台上显示警报。

[0106] 作为另一示例，本发明的系统提供独特的系统任务分派。使用设备控制指示，当前视频系统允许用户定位视频传感器，并且在一些复杂的传统系统中，允许用户指出感兴趣或不感兴趣的区域。设备控制指示是控制摄像机的位置、朝向和焦点的指令。代替设备控制指示，本发明的系统使用基于视频基元的事件鉴别器，作为主要任务分派机制。利用事件鉴别器和视频基元，给操作者提供了比传统系统更加直观的方法，以从系统提取有用信息。并不是用例如“摄像机 A 向左摇转 45 度”的设备控制指示来给系统分派任务，而是，可以用基于视频基元的一个或多个事件鉴别器，例如“人进入了限制区域 A”，以人类直观方式给本发明的系统分派任务。

[0107] 将本发明用于市场研究，下面是本发明可执行的视频监视类型的示例：计算商店内的人；计算商店部分区域中的人；计算在商店的特定位置处停下来的人；测量人在商店待了多长时间；测量人在部分区域中待了多长时间；以及测量商店中的队列长度。

[0108] 将本发明用于保安目的，下面是本发明可执行的视频监视类型的示例：确定何时有人进入限制区域并存储有关图像；确定何时有人在异常时间进入区域；确定何时发生也许是未授权的对货架空间和存储空间的改变；确定何时机上乘客接近座舱；确定何时人们穿过安全入口；确定在机场是否有无人看管的包；以及确定是否有偷窃财物行为。

[0109] 示范性的应用可以是访问控制，这可包括例如：检测是否有人爬过围墙，或者进入禁止区域；检测是否有人沿错误的方向移动（例如在机场，通过出口进入安全区域）；确定如果在感兴趣区域检测到的对象数目并不与基于 RFID 标签或刷卡进入而预计的数目匹配，指示未授权人员的存在。这在住宅应用中也是有用的，其中，视频监视系统能够区分人

和宠物的动作，因此消除大多数错误的警报。注意，在多数住宅应用中，隐私是重要问题；例如，业主也许不希望有人远程监视家里并能够看到家中的物品和家中发生的事务。因此，在这些应用中使用的一些实施例中，可在本地执行视频处理，并且仅在需要时（例如但不局限于对动物活动或其它危险情况的检测）才将可选视频或镜头发送到一个或多个远程监控站。

[0110] 另一个示范性应用领域可以是资产监控。这意味着，检测对象是否离开场景，例如，展品是否被从展览馆移走。在零售环境中，资产监控可具有多个方面，例如可包括：检测是否有单个可疑人员偷偷摸摸地拿了大量的指定物品；确定是否有人通过入口退出，尤其是是否推着购物车通过入口退出；确定是否有人将不匹配价格标签贴到物品上，例如，装满一包很贵的咖啡，但是使用了低价咖啡的价格标签；或者检测是否有人在装载码头留下了大箱子。

[0111] 另一个示范性应用领域可以是针对安全目的。例如，这可以包括：检测是否有人滑倒，例如在商店或在停车场；检测是否有车在停车场开得很快；检测在站台没有车时是否有人离火车或地铁的站台边缘距离太近；检测是否有人在铁轨上；检测在火车开动时是否有人被车门夹住；或者计算进入和离开设施的人数，因此保持精确的人头数，这在紧急情况下非常重要。

[0112] 另一个示范性应用领域可以是交通监控。这可包括：检测是否有车辆停住了，尤其是在诸如桥梁或隧道之类的地点，或者检测是否有车辆停在非停车区域。

[0113] 另一个示范性应用领域可以是反恐怖主义。除了上面提到的一些应用之外，这还可包括：检测是否有物体被留在机场，是否有物体被扔过围墙，或者是否有物体被留在铁轨上；检测围绕关键基础设施移动的人或车辆；或者检测向港口或开放水域的轮船快速接近的小船。

[0114] 另一个示范性应用领域可以照顾老弱病残，尤其是在家里。例如，这可以包括：检测是否有人跌倒；或者检测异常行为，例如在过长时间段上无人进入厨房。

[0115] 图 1 示出了本发明的视频监视系统的平面图。计算机系统 11 包括计算机 12，计算机 12 具有计算机可读介质 13，计算机可读介质 13 包含用于根据本发明来操作计算机 12 的软件。计算机系统 11 连接到一个或多个视频传感器 14、一个或多个视频记录器 15 以及一个或多个输入 / 输出 (I/O) 设备 16。视频传感器 14 还可选地连接到视频记录器 15，用于直接记录视频监视数据。该计算机系统可选地连接到其它传感器 17。

[0116] 视频传感器 14 向计算机系统 11 提供源视频。每个视频传感器 14 可通过例如直接连接（例如火线数码摄像机接口）或网络连接到计算机系统 11。视频传感器 14 可在本发明的安装之前就存在，或者可作为本发明的一部分而安装。视频传感器 14 的示例包括：视频摄像机、数码摄像机、彩色摄像机、单色摄像机、摄像机、可携式摄像机、PC 摄像头、网络摄像机 (webcam)、红外视频摄像机和 CCTV 摄像机。

[0117] 视频记录器 15 从计算机系统 11 接收视频监视数据以用于记录和 / 或将源视频提供给计算机系统 11。每个视频记录器 15 可使用例如直接连接或网络连接到计算机系统 11。视频记录器 15 可在本发明的安装之前就存在，或者可作为本发明的一部分而安装。计算机系统 11 中的视频监视系统可控制何时以何种质量来设置视频记录器 15 记录视频。视频记录器 15 的示例包括：磁带录像机、数字视频记录器、网络视频记录器、录像盘、DVD 和计算机

可读介质。该系统还可通过控制视频编码器和流传输协议，来调制在网络上流传输的视频的带宽和质量。当检测到感兴趣的活动时，可编码并流传输较高比特率、帧频或分辨率的图像。

[0118] I/O 设备 16 向计算机系统 11 提供输入并从计算机系统 11 接收输出。I/O 设备 16 可用于给计算机系统 11 分派任务，并从计算机系统 11 产生报告。I/O 设备 16 的示例包括：键盘、鼠标、手写笔、监视器、打印机、其他计算机系统、网络和警报器。

[0119] 其它传感器 17 给计算机系统 11 提供附加输入。每个其它传感器 17 可使用例如直接连接或网络连接到计算机系统 11。其它传感器 17 可在本发明的安装之前就存在，或者作为本发明的一部分而安装。其它传感器 17 的示例包括但不局限于：运动传感器、光学绊网、生物测定传感器、RFID 传感器和基于卡或基于键盘的授权系统。其它传感器 17 的输出可由计算机系统 11、记录设备和 / 或记录系统记录。

[0120] 图 2 示出了本发明的视频监视系统的流程图。参考图 10-15，以本发明的多个方面为例，示出了用于监控杂货店的本发明视频监视系统的示例。

[0121] 在块 21 中，如针对图 1 所述的，建立视频监视系统。每个视频传感器 14 朝向视频监视的位置。计算机系统 11 与来自视频装置 14 和 15 的视频馈送相连。可使用该位置的已有装置或新安装的装置来实现视频监视系统。

[0122] 在块 22 中，校准视频监视系统。一旦视频监视系统从块 21 就位了，就进行校准。块 22 的结果是视频监视系统能够确定在视频传感器所提供的视频图像的各个位置处的特定对象（例如人）的近似绝对大小和速度。。可使用手动校准、半自动校准和自动校准来校准系统。在块 24 的论述之后进一步描述校准。

[0123] 在图 2 的块 23 中，给视频监视系统分派任务。分派任务发生在块 22 的校准之后，并且是可选的。给视频监视系统分派任务包括指定一个或多个事件鉴别器。在没有分派任务的情况下，视频监视系统通过检测并存档视频基元和有关的视频图像，而不像图 4 中块 45 所示地采取任何动作，来进行操作。

[0124] 图 3 示出了用于给视频监视系统分派任务以确定事件鉴别器的流程图。事件鉴别器是指可选地与一个或多个空间属性和 / 或一个或多个时间属性相互作用的一个或多个对象。根据视频基元（也称为活动描述元数据）来描述事件鉴别器。一些视频基元设计标准包括如下：能够被实时从视频流中提取；包括来自视频的所有有关信息；以及表征的简洁性。

[0125] 希望从视频流实时提取视频基元，以便使系统能够产生实时警报。并且通过这么做，由于视频提供了连续的输入流，该系统不会落后。

[0126] 视频基元还应该包含来自视频的所有有关信息，因为在提取视频基元时，系统不知道用户定义的规则。因此，视频基元应该包含信息以能够检测用户所指定的任何事件，而不需要返回视频并重新分析视频。

[0127] 由于多种原因，也希望有简洁表征。所提出的本发明的一个目的是扩展监视系统的存储循环使用时间。如上所述，这可以通过存储活动描述元数据和质量取决于活动存在与否的视频来取代总是存储高质量的视频而实现。因此，视频基元越简洁，可存储的数据越多。此外，视频基元表征越简洁，数据访问越快，因此，可加速法律搜索。

[0128] 视频基元的精确内容可取决于应用和可能的感兴趣事件。下面描述一些示范性实

施例。

[0129] 视频基元的示范性实施例可包括场景 / 视频描述符，描述整个场景和视频。通常，这可包括场景出现的详细描述，例如天空、植物、人造物体、水等的定位；和 / 或气象条件，例如降雨、雾等的存在 / 不存在。对于视频监视应用，例如，全景中的变化可能非常重要。示范性描述符可描述突然的照明变化；它们可以指示摄像机的运动，尤其是摄像机开始或停止移动，并且在后一种情况下，指示摄像机是否返回到前一景象或者至少返回到前一已知景象；它们可指示视频馈送的质量变化，例如，馈送是否突然变得噪声很大或变暗，可能指示出对馈送的篡改；或者它们可以示出沿着水体的水线的改变（对于该后一问题的特定方法的其它信息，例如，可以查阅 2004 年 10 月 1 日申请的待审美国专利申请 No. 10/954,479，并通过参考将其合并在此）。

[0130] 视频基元的另一个示范性实施例可包括对在视频馈送中看到的对象的可观测属性进行指代的对象描述符。存储何种对象有关信息可取决于应用领域和可用的处理能力。示范性对象描述符可包括一般属性，包括但不局限于大小、形状、周长、位置、运动轨迹、速度和方向、运动突出性及其特征、颜色、刚性、纹理和 / 或类别。对象描述符还可包含一些对于应用和类别更加特定的信息：对于人，这可包括肤色的存在和比例、性别和种族信息、描述体形和姿态的一些人体模型；或者对于车辆，这可包括类别（例如卡车、SUV、轿车、自行车等）、品牌、型号、许可牌号。对象描述符还可包含活动，包括但不局限于携带对象、跑动、走动、站起或举手。例如谈话、打架或冲突的一些活动也可涉及到其它对象。对象描述符还可包含识别信息，包括但不局限于脸部或步态。

[0131] 视频基元的另一个示范性实施例可包括描述视频的每个区域的运动方向的动向 (flow) 描述符。例如，可通过检测禁止方向上的任何运动，将这种描述符用于检测回传 (passback) 事件（对于该后一问题的特定方法的其它信息，例如，可以查阅 2004 年 1 月 30 日申请的待审美国专利申请 No. 10/766,949，并通过参考将其合并在此）。

[0132] 基元还可来自非视频源，例如音频传感器、热传感器、压力传感器、读卡器、RFID 标签、生物测定传感器等。

[0133] 分类是指对对象属于特定种类或类别的标识。分类的示例包括：人、狗、车辆、警车、个人和特定种类的对象。

[0134] 大小是指对象的尺寸属性。大小的示例包括：大、中、小、平、高于 6 英尺、短于 1 英尺、比 3 英尺宽、比 4 英尺薄、大约人体尺寸、比人大、比人小、大约车的大小、图像中具有近似像素大小的矩形以及多个图像像素。

[0135] 位置是指对象的空间属性。例如，位置可以是：像素坐标形式的图像位置、一些世界坐标系中的绝对真实世界位置、或者相对于地理标志或另一对象的位置。

[0136] 颜色是指对象的色彩属性。颜色的示例包括：白色、黑色、灰色、红色、HSV 值的范围、YUV 值的范围、RGB 值的范围、平均 RGB 值、平均 YUV 值、RGB 值的直方图。

[0137] 刚性是指对象的形状一致性属性。非刚性对象（例如人或动物）的形状可在不同帧之间变化，而刚性对象（例如车辆或房子）的形状可在不同帧之间保持大体上不变（除非可能由于转向而稍有改变）。

[0138] 纹理是指对象的图案属性。纹理特征的示例包括：自相似性、谱功率、线性和粗糙程度。

[0139] 内部运动是指对象刚性的度量。完全刚性的对象的示例是车，其并不表现出大量内部运动。完全非刚性的对象的示例是四肢摆动的人，其表现出大量的内部运动。

[0140] 运动是指可自动检测到的任何运动。运动的示例包括：对象的出现；对象的消失；对象的垂直运动；对象的水平运动；以及对象的周期运动。

[0141] 突出（salient）运动是指可被自动检测到并在某个时间段上被跟踪的任何运动。这种运动的对象表现出明显有目的的运动。突出运动的示例包括：从一个地方移动到另一个地方；以及移动以与另一个对象相互作用。

[0142] 突出运动的特征是指突出运动的性质。突出运动的特征的示例包括：轨迹；图像空间中轨迹的长度；轨迹在环境的三维表征中的近似长度；对象在图像空间随时间变化的位置；对象在环境的三维表征中随时间变化的近似位置；轨迹的持续时间；图像空间中的速度（例如速率和方向）；环境的三维表征中的近似速度（例如速率和方向）；在某一速度上的持续时间；图像空间中速度的改变；环境的三维表征中速度的近似改变；速度改变的持续时间；运动的停止；以及运动停止的持续时间。速度是指在特定时间出对象的速率和方向。轨迹是指在可跟踪对象的时间段或某一时间段上对象的（位置，速度）对集合。

[0143] 场景变化是指可检测到随时间段而变化的场景的任何区域。场景变化的示例包括：静态对象离开场景；对象进入场景并变为静态的；对象在场景中改变位置；以及对象改变外观（例如颜色、形状或大小）。

[0144] 场景变化的特征是指场景变化的性质。场景变化的特征的示例包括：图像空间中场景变化的大小；环境的三维表征中场景变化的近似大小；场景变化发生的时间；图像空间中场景变化的位置；以及环境的三维表征中场景变化的近似位置。

[0145] 预定义模型是指对象的先验已知模型。预定义模型的示例可包括：成人；小孩；车辆和半拖车。

[0146] 图 16a 示出了根据本发明实施例的视频监视系统的示范性视频分析部分。在图 16a 中，视频传感器（例如但不局限于视频摄像机）1601 可向视频分析子系统 1603 提供视频流 1602。视频分析子系统 1603 然后可执行视频流 1602 的分析，以导出视频基元，可将视频基元存储在基元存储器 1605 中。基元存储器 1605 还可用于存储非视频基元。视频分析子系统 1603 还可控制将视频流 1602 的所有或部分（例如上述的视频的质量和 / 或量）存储在视频存储器 1604 中。

[0147] 现在参考图 16b，一旦视频和（如果有其它传感器的话）非视频基元 161 可用，则系统可检测事件。用户通过使用规则和响应定义接口 162，定义规则 163 和相应的响应 164，来给系统分派任务。将规则转译为事件鉴别器，并且系统提取相应的事件发生 165。所检测到的事件发生 166 触发用户定义的响应 167。响应可包括来自视频存储器 168（可以与图 16a 中的视频存储器 1604 相同，也可以与之不同）的检测到的事件的视频快照（snapshot）。视频存储器 168 可以是视频监视系统的一部分，或者可以是独立的记录设备 15。响应的示例包括但不一定局限于如下：在系统显示器上激活视觉和 / 或听觉警报；现场激活视觉和 / 或听觉警报系统；激活静音警报；激活快速响应机制；锁门；联系保安服务；通过网络（例如但据不限于互联网）向另一计算机系统转发或流传输数据（例如图像数据、视频数据、视频基元和 / 或分析数据）；将这种数据保存到指定的计算机可读介质；激活另一个传感器或监视系统；给计算机系统 11 和 / 或另一计算机系统分派任务；和 / 或指挥

计算机系统 11 和 / 或另一计算机系统。

[0148] 基元数据可以被当作存储在数据库中的数据。为了在其中检测事件发生，需要有效的查询语言。本发明系统的实施例可包括活动推理语言，这将在下面进行描述。

[0149] 传统的关系数据库查询方案通常按照布尔二叉树结构，以允许用户对各种存储数据创建灵活的查询。叶节点通常具有“性质关系值”的形式，而性质是数据的某个关键特征（例如时间或名称）；关系通常是数值运算符（“>”、“<”、“=”等）；值是该性质的有效状态。分支节点通常代表一元或二元布尔逻辑算子，例如“与”、“或”和“非”。

[0150] 这可以形成活动查询公式方案的基础，如在本发明的实施例中一样。在视频监视应用的情况下，性质可以是视频流中检测到的对象的特征，例如大小、速率、颜色、类别（人、车），或者性质可以是场景变化性质。图 17 给出了使用这种查询的示例。在图 17a 中，提出了查询“向我显示任何红色车辆”171。这可以被分解为两个“性质关系值”（或简称为“性质”）查询，即测试对象的类别是否是车辆 173 以及其颜色是否主要为红色 174。这两个子查询可利用布尔算子“与”172 来组合。类似地，在图 17b 中，查询“向我显示摄像机何时开始或停止移动”可表达为性质子查询“摄像机是否开始移动”177 和“摄像机是否停止移动”178 的布尔“或”176 的组合。

[0151] 本发明的实施例可以以两个示范性方式来扩展这种数据库查询方案：(1) 基本叶节点可增加用于描述场景内空间活动的活动检测符；以及 (2) 布尔算子分支节点可增加用于指定空间、时间和对象相互关系的修饰符。

[0152] 活动检测符对应于与视频场景的区域有关的行为。它们描述了对象怎样与场景中的位置发生作用。图 18 示出了三个示范性活动检测符。图 18a 表示沿特定方向穿过使用虚拟视频绊网的周边的行为（对于关于怎样实现这种虚拟视频绊网的进一步信息，可以查阅例如美国专利 No.6,696,945）。图 18b 表示在铁路轨道上闲逛一段时间的行为。图 18c 表示从墙壁部分拿走某东西的行为（对于怎样可以实现的示范性方法，可以查阅 2003 年 1 月 30 日申请的题为“Video Scene Background Maintenance-Change Detection&Classification”美国专利申请 NO.10/331,778）。其它示范性活动检测符可包括检测人跌倒、检测人改变方向或速率、检测人进入区域或检测人沿错误方向前进。

[0153] 图 19 示出了活动检测符叶节点（这里是指绊网横穿）怎样与简单的性质查询结合以便检测红色车辆横穿视频绊网 191 的示例。用布尔“与”算子 192 组合性质查询 172、173、174 和活动检测符 193。

[0154] 将查询与修饰的布尔算子相组合（组合算符）可进一步增加灵活性。示范性修饰符包括空间、时间、对象和计数修饰符。

[0155] 空间修饰符可使布尔算子仅对在场景中邻近 / 非邻近的子活动进行操作（即，布尔算子的自变量，例如在图 19 中在布尔算子下示出的自变量）。例如，“与 - 在 50 像素内”可用于表示仅在活动之间的距离小于 50 像素时才适用“与”。

[0156] 时间修饰符可使布尔算子仅对发生在彼此的特定时段内、这种时段外或时间范围内的子活动进行操作。还可指定事件的时间顺序。例如，“与 - 第一在第二的 10 秒内”可用于表示仅在第二子活动发生在第一子活动之后的 10 秒之内时才适用“与”。

[0157] 对象修饰符可使布尔算子仅对涉及相同或不同对象的子活动进行操作。例如，“和 - 涉及相同对象”可用于表示仅在两个子活动涉及相同特定对象时才适用。

[0158] 计数修饰符可使布尔算子仅在以预定次数满足条件时才被触发。计数修饰符通常可包括数值关系，例如“至少 n 次”、“精确 n 次”、“至多 n 次”等。例如，“或 - 至少两次”可用于表示至少两个“或”算子的子查询必须为真。计数修饰符的另一种使用可实现例如“如果同一个人从货架上取了至少五个物品则警报”的规则。

[0159] 图 20 示出了使用组合算符的示例。这里，所需活动查询用于“找到做出非法左转的红色汽车”201。可通过活动描述符和修饰的布尔算子的组合来捕获非法左转。可使用一个虚拟绊网来检测对象出了边道 193，并且可使用另一个虚拟绊网来检测对象沿着道路前进到左边 205。这些可通过修饰的“与”算子 202 来组合。标准布尔“与”算子确保了必须检测活动 193 和 205。对象修饰符 203 检查同一个对象穿过两个绊网，而时间修饰符 204 检查首先穿过从下到上绊网 193，然后在 10 秒内穿过从右到左绊网 205。

[0160] 该示例还指示了组合算符的能力。理论上，可以针对左转而定义独立的活动检测符，而不依赖于简单的活动检测符和组合算符。然而，检测符是不灵活的，使得难以提供任意的转角和方向，并且对于所有可能的事件而写出独立的检测符也是麻烦的。相反地，使用组合算符和简单的检测符提供了极大的灵活性。

[0161] 可作为简单活动的组合而被检测到的复杂活动的其它示例包括停车和人下车、或多个人形成一组并彼此紧随。这些组合算符还可组合不同类型和来源的基元。示例可包括如下规则，例如“示出关灯之前室内的人”、“示出之前没有刷卡就进门的人”、或“示出感兴趣的区域是否具有比 RFID 标签读取器所预计的更多的对象”，即是否有无 RFID 标签的非法对象处于区域内。

[0162] 组合算符可以任何深度组合任意数目的子查询，并且甚至可以组合其它组合算符。图 21a 和 21b 所示的示例可以是如下规则 :检测车辆是否左转 2101、然后右转 2104。可以由方向性绊网 2102 和 2103 来检测左转 2101，而由方向性绊网 2105 和 2106 来检测右转 2101。左转可被表达为绊网活动检测符 2112 和 2113，分别对应于绊网 2102 和 2103，并通过“与”组合算符 2111 与对象修饰符“相同”2117 和时间修饰符“2112 在 2113 之前”2118 结合。类似地，右转可被表达为绊网活动检测符 2115 和 2116，分别对应于绊网 2105 和 2106，并通过“与”组合算符 2114 与对象修饰符“相同”2119 和时间修饰符“2115 在 2116 之前”2120 结合。为了检测相同对象首先左转然后右转，左转检测符 211 和右转检测符 2114 通过“与”组合算符 2121 与对象修饰符“相同”2122 和时间修饰符“2111 在 2114 之前”2133 结合。最终，为了确保被检测的对象是汽车，利用布尔“与”算子 2125 将左转和右转检测符 2121 和性质查询 2124 组合。

[0163] 可选地，所有这些检测符均可与时间属性组合。时间属性的示例包括 :每 15 分钟 ;在 9 :00pm 和 6 :30am 之间 ;小于 5 分钟 ;长于 30 秒 ;以及在周末。

[0164] 在图 2 的块 24 中，操作视频监视系统。本发明的视频监视系统自动地操作，检测并存档场景中对象的视频基元，并使用事件鉴别器来实时检测事件发生。此外，在适当的情况下实时采取行动，例如激活警报、产生报告以及产生输出。可在系统本地或者经由网络（例如互联网）在其他地方显示和 / 或存储报告和输出。图 4 示出了用于操作视频监视系统的流程图。

[0165] 在块 41 中，计算机系统 11 从视频传感器 14 和 / 或视频记录器 15 获得源视频。

[0166] 在块 42 中，实时从源视频中提取视频基元。可选地，可从一个或多个其他传感器

17 中获得和 / 或提取非视频基元，并用于本发明。图 5 示出了视频基元的提取。

[0167] 图 5 示出了用于提取视频监视系统的视频基元的流程图。块 51 和 52 并行操作，并且可按照任意顺序或者同时执行。在块 51 中，通过运动来检测对象。对于该块，可使用在像素级上对帧间运动进行检测的任何运动检测算法。例如，可使用三帧差分技术，这在 {1} 中有所讨论。将检测到的对象转发到块 53。

[0168] 在块 52 中，通过改变来检测对象。对于该块，可使用用于检测相对于背景模型的变化的任何变化检测算法。如果由于一个或多个像素并不与帧的背景模型一致而将它们认为处于帧的前景中，则在该块中检测到对象。例如，可使用例如动态自适应背景消减的随机背景建模技术，这在 {1} 和 2000 年 10 月 24 日申请的美国专利申请 No. 09/694, 712 中有所描述，将检测到的对象转发到块 53。

[0169] 块 51 的运动检测技术和块 52 的变化检测技术是互补技术，其中，每个技术有利地解决了另一技术中的缺陷。可选地，对于针对块 51 和 52 所讨论的技术，可使用附加的和 / 或备选的检测方案。附加的和 / 或备选的检测方案的示例包括如下 ：如 {8} 所述的用于找到人的 Pfinder 检测方案 ；肤色检测方案 ；人脸检测方案 ；以及基于模型的检测方案。将这种附加的和 / 或备选的检测方案的结果提供给块 53。

[0170] 可选地，如果视频传感器 14 运动 (视频摄像机扫拍、变焦和 / 或平移)，则可在块 51 和 52 之间的块之前插入附加块，以给块 51 和 52 提供输入，以用于视频稳定。可通过仿射或射影的全局运动补偿来实现视频稳定。例如，在 2000 年 7 月 3 日申请的美国专利申请 No. 09/609, 919 (现在的美国专利 No. 6, 738, 424) 中描述的图像对准可用于获得视频稳定，通过参考将其合并在此。

[0171] 在块 53 中，产生团迹。通常，团迹是帧中的任何对象。团迹的示例包括 ：运动对象，例如人或车辆 ；以及消费类产品，例如家具、衣物或零售货物。使用从块 32 和 33 中检测到的对象来产生团迹。对于该块，可使用用于产生团迹的任何技术。用于根据运动检测和变化检测来产生团迹的示范性技术使用连接组成部分 (connected components) 方案。例如，可使用形态学和连接组成部分算法，这在 {1} 中有所描述。

[0172] 在块 54 中，跟踪团迹。对于该块，可使用用于跟踪团迹的任意技术。例如，可使用 Kalman 滤波或 CONDENSATION 算法。作为另一个示例，可使用如在 {1} 中所述的模版匹配技术。作为另一个示例，可使用如 {5} 中所述的多重假设 Kalman 跟踪器。作为另一个示例，可使用在 2000 年 10 月 24 日申请的美国专利申请 No. 09/694, 712 中所述的逐帧跟踪技术。对于杂货店位置处的示例，可跟踪的对象的示例包括移动的人、库存物品和库存移动工具，例如购物车或手推车。

[0173] 可选地，可用对于本领域技术人员而言是公知的任意检测和跟踪方案来代替块 51-54。在 {1} 中描述了这种检测和跟踪方案的示例。

[0174] 在块 55 中，分析所跟踪的对象的每个轨迹，以确定轨迹是否突出。如果轨迹不突出，则轨迹表示对象表现出不稳定的运动，或表示对象具有不稳定的大小或颜色，并且系统拒绝相应对象并不再分析该对象。如果轨迹是突出的，则轨迹表示可能对对象感兴趣。通过向轨迹应用突出性测量来确定轨迹是突出的还是不突出的。在 {13} 和 {18} 中描述了确定轨迹是突出的还是不突出的技术。

[0175] 在块 56 中，对每个对象进行分类。将每个对象的一般类别确定为该对象的种类。

可通过多种技术来执行分类，这种技术的示例包括使用神经网络分类器 {14} 和使用线性区别分类器 {14}。种类的示例与针对块 23 所讨论的相同。

[0176] 在块 57 中，使用来自块 51-57 的信息和必要的附加处理来识别视频基元。所识别的视频基元的示例与针对块 23 所讨论的相同。例如，对于大小，系统可将从块 22 中的校准所获得的信息用作视频基元。根据校准，系统具有足够的信息来确定对象的近似大小。作为另一示例，系统可将从块 54 测量的速度用作视频基元。

[0177] 在块 43 中，将来自块 42 的视频基元存档。可将视频基元存档在计算机可读介质 13 或另一计算机可读介质中。可与视频基元一起，将来自源视频的有关帧或视频图像存档。该存档步骤是可选的；如果仅将系统用于实时事件，则可跳过该存档步骤。

[0178] 在块 44 中，使用事件鉴别器，从视频基元中提取事件发生。在块 42 中确定视频基元，并且根据块 23 中的给系统分派任务来确定事件鉴别器。事件鉴别器用于对视频基元进行过滤以确定是否出现了事件发生。例如，事件鉴别器可查找由某人在上午 9 :00 和下午 5 :00 之间沿“错误路径”进入区域而定义的“错误路径”事件。事件鉴别器检查根据图 5 所产生的所有视频基元，并确定是否存在具有以下性质的任何视频基元 :在上午 9 :00 和下午 5 :00 之间的时戳、“人”或“人群”的类别、区域内的位置、以及“错误”运动方向。事件鉴别器还可使用上述的其它类型的基元，和 / 或组合来自多个视频源的视频基元以检测事件发生。

[0179] 在块 45 中，针对在块 44 中提取的每个事件发生，适当地采取行动。图 6 示出了对视频监视系统采取行动的流程图。

[0180] 在块 61 中，如检测到事件发生的事件鉴别器所指示的，做出响应。如果有的话，针对块 34 中的每个事件鉴别器，识别响应。

[0181] 在块 62 中，针对发生的每个事件发生，产生活动记录。活动记录包括例如 :对象轨迹的细节；对象检测的时间；对象检测的位置；以及所采用的事件鉴别器的描述或定义。活动记录可包括事件鉴别器所需的例如视频基元的信息。活动记录还可包括事件发生所涉及的对象和 / 或区域的代表性视频或静止图像。活动记录被存储在计算机可读介质上。

[0182] 在块 63 中，产生输出。该输出是基于在块 44 中提取的事件发生和来自块 41 的源视频的直接馈送的。该输出被存储在计算机可读介质上，显示在计算机系统 11 或另一计算机系统上，或者被转发到另一计算机系统。随着系统操作，收集关于事件发生的信息，并且操作者可在任何时间（包括实时）查看该信息。接收信息的格式的示例包括 :显示在计算机系统的监视器上；硬拷贝；计算机可读介质；以及交互式网页。

[0183] 该输出可包括经由模拟视频传输装置或经由网络视频流从块 41 传输来的源视频的直接馈送的显示。例如，源视频可被显示在计算机系统的监视器或闭路监视器的窗口中。此外，该输出可包括标出了图形以加亮事件发生所涉及的对象和 / 或区域的源视频。如果系统操作在法律分析模式下，则视频可来自视频记录器。

[0184] 该输出可包括基于操作者和 / 或事件发生的需求、针对操作者的一个或多个报告。报告的示例包括 :发生的事件发生的数目；场景中发生事件发生的位置；事件发生所发生的时间；每个事件发生的代表性图像；每个事件发生的代表性视频；原始统计数据；事件发生的统计（例如次数、频率、地点和时间）；和 / 或人类可读图形显示。

[0185] 图 13 和 14 示出了针对图 15 的杂货店的走道的示范性报告。在图 13 和 14 中，在

块 22 中识别多个区域，并相应地在图像中标记出。图 13 中的区域与图 12 中的区域匹配，并且图 14 中的区域是不同的区域。给系统分派任务以查找在区域中停留的人。

[0186] 在途 13 中，示范性报告是来自被标记以包括标签、图形、统计信息和统计信息分析的视频的图像。例如，被识别为咖啡的区域具有在区域中的平均顾客人数为 2 人 / 小时并且区域中的平均逗留时间为 5 秒的统计信息。系统确定该区域是“冷点”区域，这意味着在该区域没有多少商业活动。作为另一个示例，被识别为苏打水的区域具有在区域中的平均顾客人数为 15 人 / 小时并且区域中的平均逗留时间为 22 秒的统计信息。系统确定该区域是“热点”区域，这意味着在该区域有大量的商业活动。

[0187] 在图 14 中，示范性报告是来自被标记以包括标签、图形、统计信息和统计信息分析的视频的图像。例如，在走道背面的区域的平均顾客数人是 14 人 / 小时，并被确定为具有低流量。作为另一示例，在走道前面的区域的平均顾客人数是 83 人 / 小时，并被确定为具有高流量。

[0188] 对于图 13 或图 14，如果操作者想要关于任意特定区域的更多信息或任意特定区域，则指向 - 点击界面使得操作者可以在系统检测到并存档的区域和 / 或活动的代表性静止画面和视频图像中导航。

[0189] 图 15 示出了对于杂货店的走道的另一示范性报告。该示范性报告包括来自被标记以包括标签和轨迹指示以及描述该带标记图像的文本的视频的图像。该示例的系统被分派了搜索多个区域的任务 :对象轨迹的长度、位置和时间 ;对象静止的时间和位置 ;操作员所指定的轨迹与区域的相关 ;以及对象的类别，例如不是人、一个人、两个人和三个或更多个人。

[0190] 图 15 的视频图像来自记录轨迹的时间段。在三个对象中，两个对象中的每个均被分类为一个人，而一个对象被分类为不是人。每个对象被分配有标签，即 Person ID 1032、Person ID 1033 以及 Object ID 32001。对于 Person ID 1032，系统确定该人在区域逗留了 52 秒，并在圆圈所指定的位置处逗留了 18 秒。对于 Person ID 1033，系统确定该人在区域逗留了 1 分 18 秒，并在圆圈所指定的位置处逗留了 12 秒。在标记的图像中包括 Person Id 1032 和 Person ID 1033 的轨迹。对于 Object ID 32001，系统并不进一步分析该对象，并用 X 指示该对象的位置。

[0191] 转向图 2 中的块 22，校准可以是 (1) 手动的 ;(2) 使用来自视频传感器或视频记录器的图像的半自动的 ;或者 (3) 使用来自视频传感器或视频记录器的图像的自动的。如果需要图像，则假定要由计算机系统 11 分析的源视频是来自获得用于校准的源视频的视频传感器的。

[0192] 对于手动校准，操作者给计算机系统 11 提供每个视频传感器 14 的朝向和内部参数，以及每个视频传感器 14 相对于位置的放置。计算机系统 11 可选地保持该位置的地图，并且可在地图上指示出视频传感器 14 的放置。该地图可以是环境的二维或三维表征。此外，手动校准给系统提供了足够的信息以确定对象的近似大小和相对位置。

[0193] 可选地，对于手动校准，操作者可用表示已知大小对象（例如人）的外观的图像来标记来自传感器的视频图像。如果操作者可在至少两个不同位置标记图像，则系统可推断出近似摄像机校准信息。

[0194] 对于半自动和自动校准，不需要知道摄像机参数或场景几何形状。根据半自动和

自动校准，产生查找表以对场景中多个区域处的对象的大小进行近似，或者推断出摄像机的内部和外部摄像机校准参数。

[0195] 对于半自动校准，使用组合有来自操作者的输入的视频源来校准视频监视系统。单个人位于要半自动校准的视频传感器的视野范围内。计算机系统 11 接收关于该单个人的源视频，并基于该数据来自动地推断出人的大小。随着视频传感器的视野范围中看到该人的位置数目增加，并且随着视频传感器的视野范围中看到该人的时间段增加，半自动校准的精确度也提高。

[0196] 图 7 示出了视频监视系统的半自动校准的流程图。除了典型对象沿各种轨迹移动通过场景之外，块 71 与块 41 相同。该典型对象具有各种速度，并在各个位置处静止。例如，该典型对象移动到尽可能地靠近视频传感器，然后移动到尽可能地远离视频传感器。典型对象的这种运动可按照需要而重复。

[0197] 块 72-75 分别与块 51-54 相同。

[0198] 在块 76 中，在整个场景中监控典型对象。假定仅（或者至多）被跟踪的稳定对象是场景中的校准对象（即电信网对象移动通过该场景）。针对场景中所观察的每个点，收集稳定对象的大小，并且将该信息用于产生校准信息。

[0199] 在块 77 中，针对整个场景的不同区域，识别典型对象的大小。典型对象的大小被用于确定场景中各个区域处类似对象的近似大小。利用该信息，产生查找表，以匹配图像的各个区域中的典型对象的典型外观大小，或者推断出内部和外部摄像机校准参数。作为采样输出，图像的各个区域中的棍状图的显示指示出系统将其确定为适当高度。图 11 示出了这种棍状图。

[0200] 对于自动校准，执行学习阶段，其中，计算机系统 11 确定关于每个视频传感器的视野范围中的位置的信息。在自动校准中，计算机系统 11 在足以获得场景的典型对象的统计显著采样的代表性时段（例如分钟，小时或天）内接收位置的源视频，并因此推理出典型外观大小和位置。

[0201] 图 8 示出了视频监视系统的自动校准的流程图。块 81-86 与图 7 中的块 71-76 相同。

[0202] 在块 87 中，识别视频传感器的视野范围中的可跟踪区域。可跟踪区域是指视频传感器的视野范围中可容易且 / 或精确地跟踪对象的区域。不可跟踪区域是指视频传感器的视野范围中不能够容易且 / 或精确地跟踪对象的区域。不可跟踪区域可被称为不稳定或不突出区域。可能由于对象太小（例如小于预定阈值）、出现太短的时间（例如小于预定阈值）、或者表现出不突出的运动（例如无目的的）而难以跟踪对象。例如，可使用 {13} 中所述的技术来识别可跟踪区域。

[0203] 图 10 示出了针对杂货店的走道所确定的可跟踪区域。由于在走道的远端的区域处出现很多混乱物，所以该区域被确定为不突出的。混乱物是指视频中使跟踪方案混乱的某种东西。混乱物的示例包括 :飘落的叶子 ;部分被遮挡的对象 ;以及出现时间太短而难以精确地跟踪的对象。相反地，由于对于走道的近端的区域，确定了良好的跟踪，所以该区域被确定为突出的。

[0204] 在块 88 中，针对整个场景中的不同区域，识别对象的大小。对象的大小被用于确定在场景的多个区域处类似对象的近似大小。例如使用直方图或统计中值的技术被用于确

定根据场景中的位置而变的对象的典型外观高度和宽度。在场景的图像的一个部分中，典型对象可具有典型的外观高度和宽度。利用该信息，产生查找表，以匹配图像的各个区域中的对象的典型外观大小，或者推理出内部和外部摄像机校准参数。

[0205] 图 11 示出了针对杂货店走道中的典型对象而从图 10 中识别典型大小。假定典型对象是人，并相应地由标签标识。通过突出区域中检测到的人的平均高度和平均宽度的图示，确定人的典型大小。在示例中，针对平常人的平均高度，确定图 A，并针对一个人、两个人和三个人的平均宽度，确定图 B。

[0206] 对于图 A，x 轴示出了以像素为单位的团迹的高度，y 轴示出了在 x 轴上所识别的特定高度处出现的实例个数。图 A 的线的峰值对应于场景的指定区域中的团迹的最常见高度，对于该示例，峰值对应于站在指定区域的人的平均高度。

[0207] 假定人以松散组织成的组行进，对于宽度，产生与图 A 类似的图 B。对于图 B，x 轴示出了以像素为单位的团迹的宽度，y 轴示出了在 x 轴所识别的特定宽度处出现的实例个数。图 B 的线的峰值对应于多个团迹的平均宽度。假定大多数组仅包含一个人，则最大峰值对应于最常见的宽度，其对应于指定区域中单个人的平均宽度。类似地，第二最大峰值对应于指定区域中两个人的平均宽度，第三最大峰值对应于指定区域中三个人的平均宽度。

[0208] 图 9 示出了本发明的视频监视系统的附加流程图。在该附加实施例中，系统利用事件鉴别器来分析存档的视频基元，以产生附加报告，而例如不需要重新查看整个源视频。在根据本发明处理了视频源之后的任意时刻，在图 4 的块 43 中对源视频的视频基元进行存档。由于仅重新查看视频基元，并且因为不处理视频源，所以附加实施例可在相对较短的时间内分析视频内容。这相对于现有系统极大地提高效率，因为处理视频图像数据的计算成本极高，而分析从视频提取的小尺寸视频基元的计算则成本很低。例如，可产生下面的事件鉴别器：“在上两个月内在区域 A 中停留长于 10 分钟的人数”。利用附加实施例，不需要重新查看上两个月的源视频。而是，仅需要重新查看来自上两个月的视频基元，显然这是更加有效的处理。

[0209] 块 91 与图 2 中的块 23 相同。

[0210] 在块 92 中，访问所存档的视频基元。在图 4 的块 43 中对视频基元进行了存档。

[0211] 块 93 和 94 与图 4 中的块 44 和 45 相同。

[0212] 作为示范性应用，本发明可用于通过测量零售展示的功效而分析零售市场空间。将大量金钱注入零售展示，使之尽可能令人注目，以促进展示物品以及附属物品的销售。本发明的视频监视系统可配置用于测量这些零售展示的效力。

[0213] 对于该示范性应用，通过使视频传感器的视野范围朝向所希望的零售展示附近的空间，来建立视频监视系统。在分派任务期间，操作者选择代表所希望的零售展示附近的空间的区域。作为鉴别器，操作者定义他或她希望监控的人体大小的对象，该对象进入该区域并表现出可测量到的速率降低或停止了相当长时间。

[0214] 在操作一段时间之后，视频监视系统可提供报告以供市场分析。该报告可包括：在零售展示附近减速的人数；在零售展示处停下来的人数；对零售展示感兴趣的人作为时间函数的分解，例如在周末有多少人感兴趣，在晚上有多少人感兴趣；以及对零售展示表示出兴趣的人的视频镜头。从视频监视系统所获得的市场研究信息可与来自商店的销售信息以及来自商店的顾客记录组合，以提高对零售展示的效力的分析理解。

[0215] 这里所述的实施例和示例是非限制性的示例。

[0216] 关于优选实施例而详细描述了本发明，从上述说明，对于本领域技术人员而言显而易见的是，在不背离本发明的情况下可在广泛方面做出改变和修改，因此，权利要求所限定的本发明意欲涵盖落入本发明的真实精神内的所有这种改变和修改。

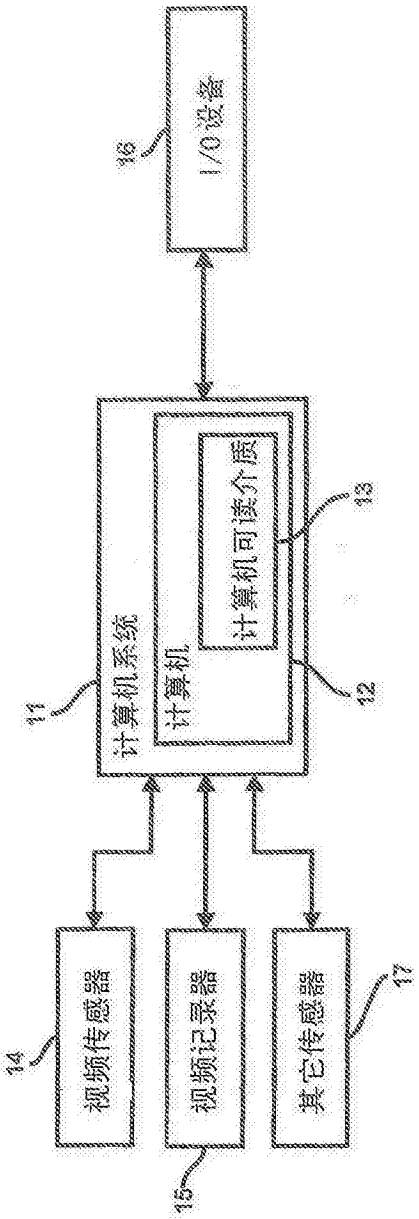

16
I/O 设备
11
计算机系统
计算机
计算机可读介质
13
12
14
视频传感器
视频记录器
15
其它传感器
17

图 1

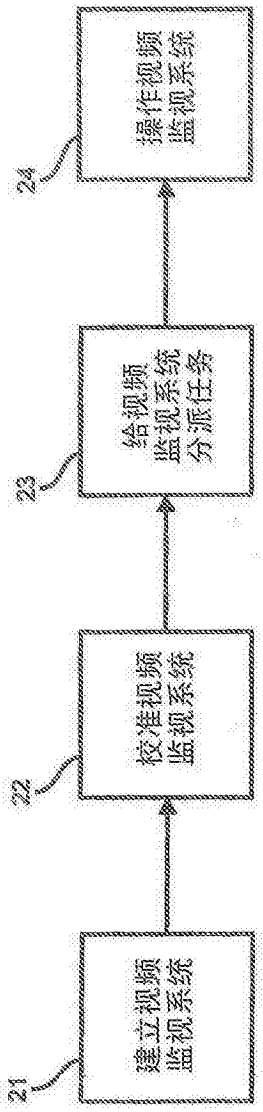

24
操作视频监视系统
23
给视频监视系统分派任务
22
校准视频监视系统
21
建立视频监视系统

图 2

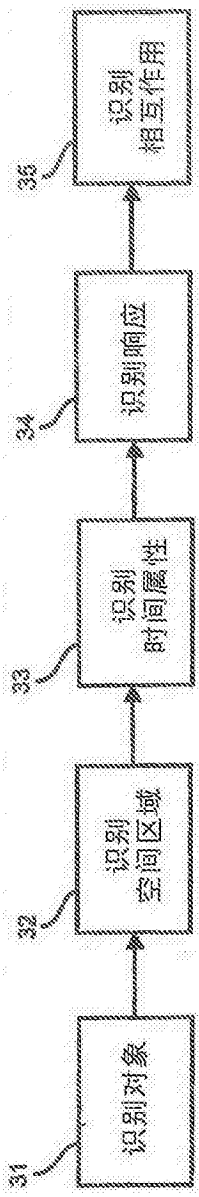

36
识别相互作用
34
识别响应
33
识别时间属性
32
识别空间区域
31
识别对象

图 3

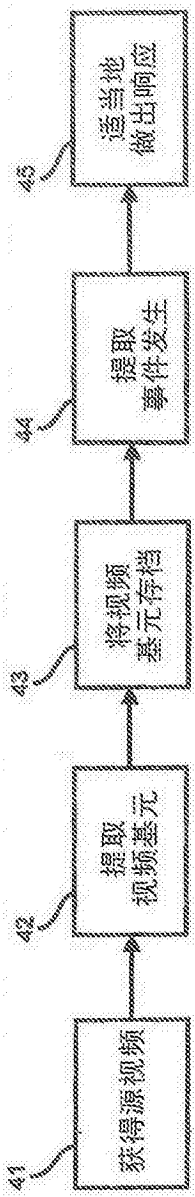

45
适当地做出响应
44
提取事件发生
43
将视频基元存档
42
提取视频基元
41
获得源视频

图 4

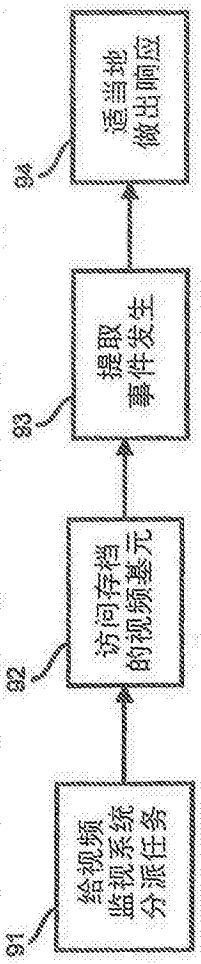

91
给视频监视系统分派任务
92
访问存档的视频基元
93
提取事件发生
94
适当地做出响应

图 9

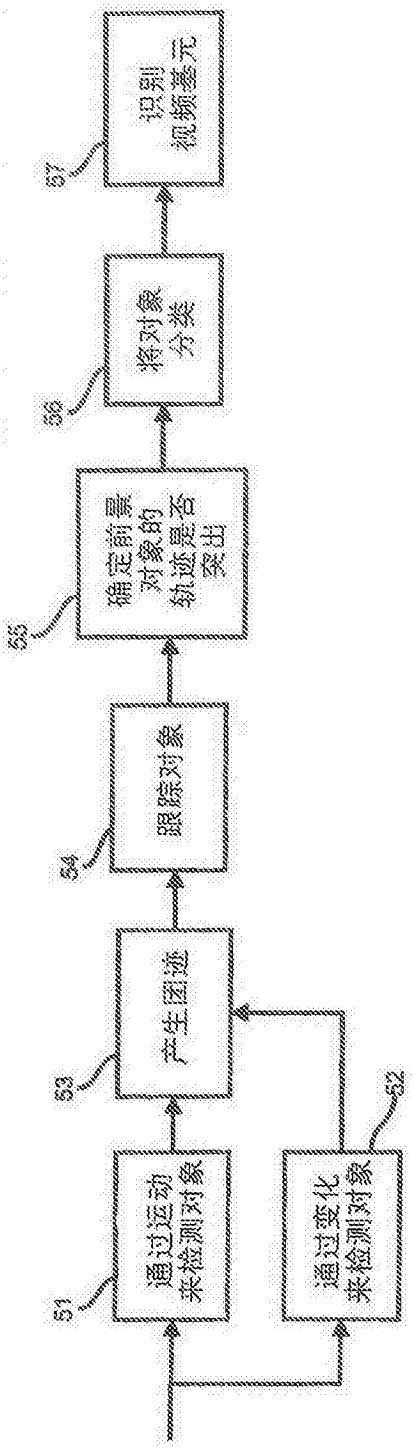

51
通过运动来检测对象
52
通过变化来检测对象
53
产生团迹
54
跟踪对象
55
确定前景对象的轨迹是否突出
56
将对象分类
57
识别视频基元

图 5

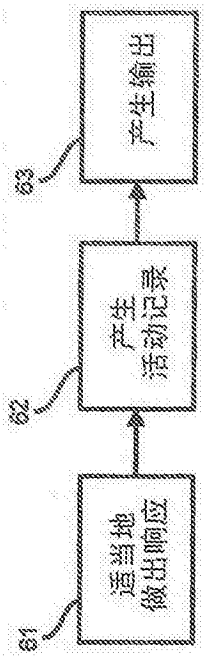

61
适当地做出响应
62
产生活动记录
63
产生输出

图 6

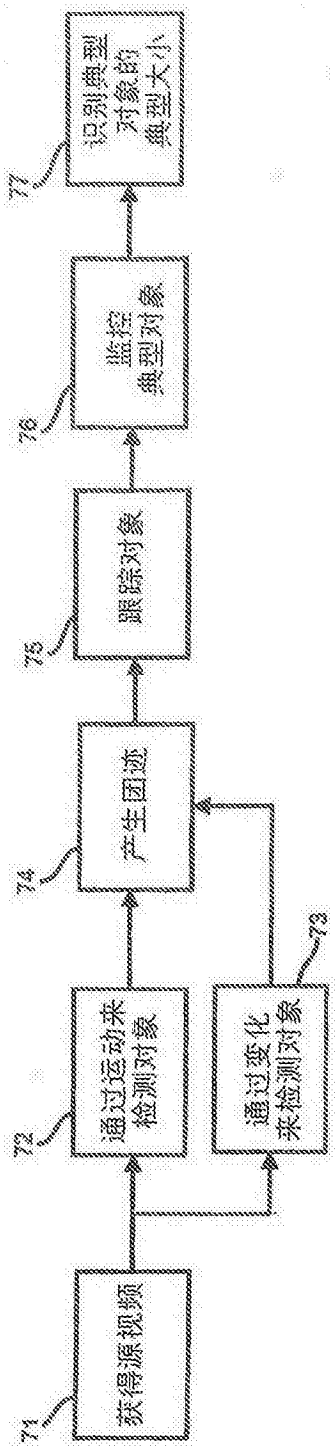

获得源视频
71
通过运动来检测对象
72
通过变化来检测对象
73
产生团迹
74
跟踪对象
75
监控典型对象
76
识别典型对象的典型大小
77

图 7

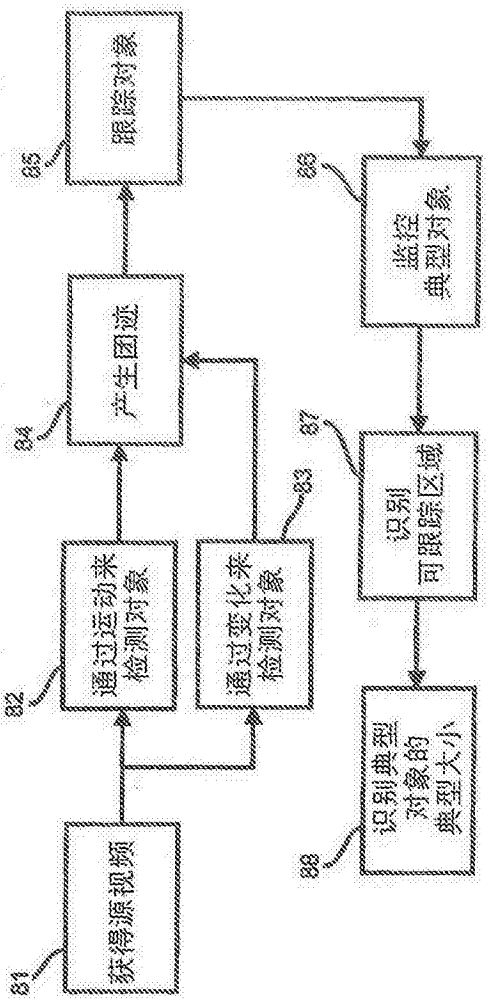

获得源视频
81
通过运动来检测对象
82
通过变化来检测对象
83
产生团迹
84
跟踪对象
85
识别典型对象的典型大小
88
识别可跟踪区域
87
监控典型对象
86

图 8

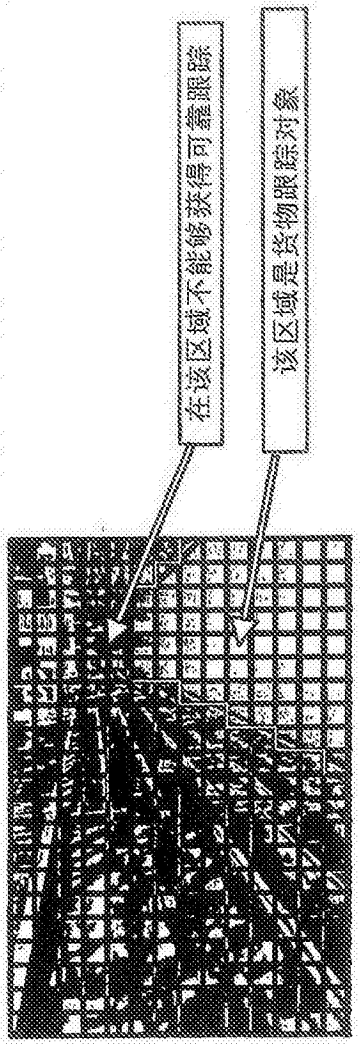


图 10

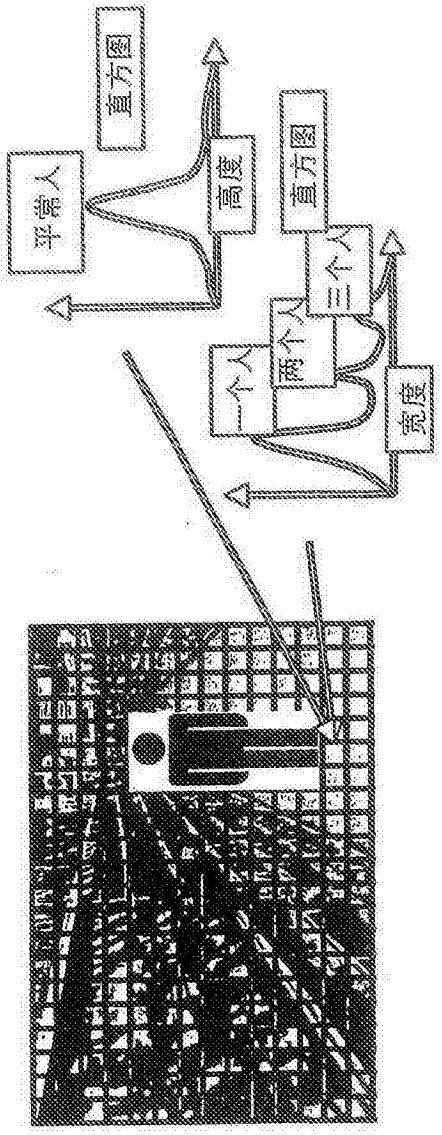


图 11

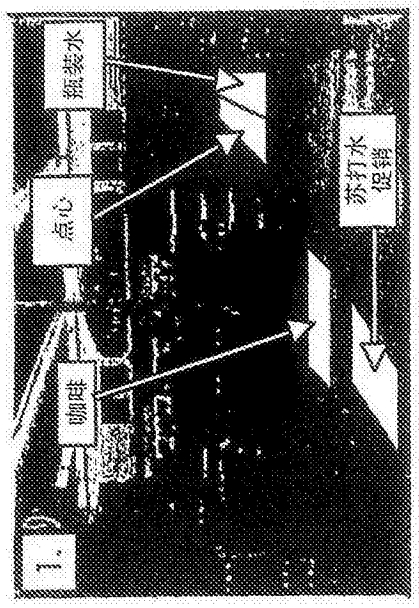

1.
瓶装水
点心
咖啡
苏打水
促销

图 12

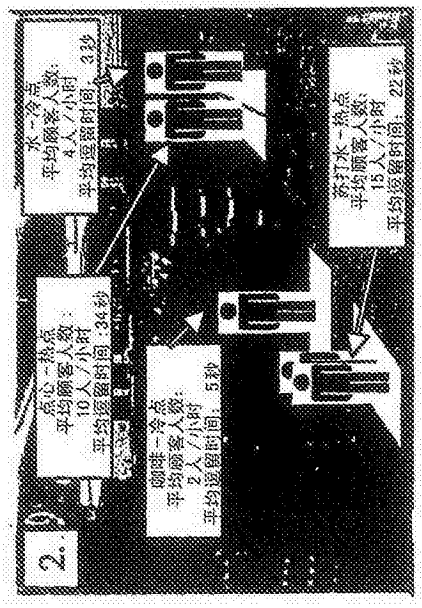

2.
水-冷点
平均顾客人数:
4人/小时
平均逗留时间: 32分
点心-热点
平均顾客人数:
10人/小时
平均逗留时间: 34分
咖啡-冷点
平均顾客人数:
2人/小时
平均逗留时间: 5分
苏打水-热点
平均顾客人数:
15人/小时
平均逗留时间: 22分

图 13

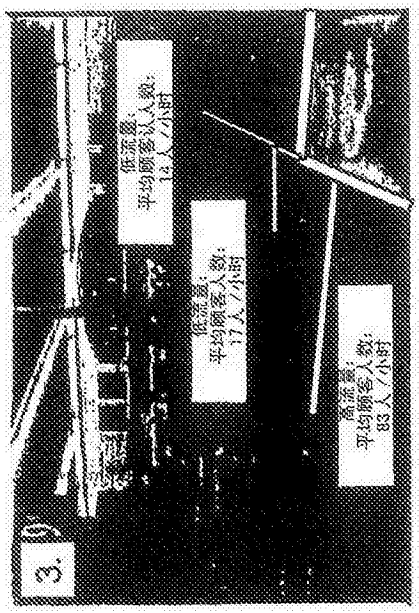

3.
低流量:
平均顾客人数:
14人/小时
低流量:
平均顾客人数:
17人/小时
高流量:
平均顾客人数:
83人/小时

图 14

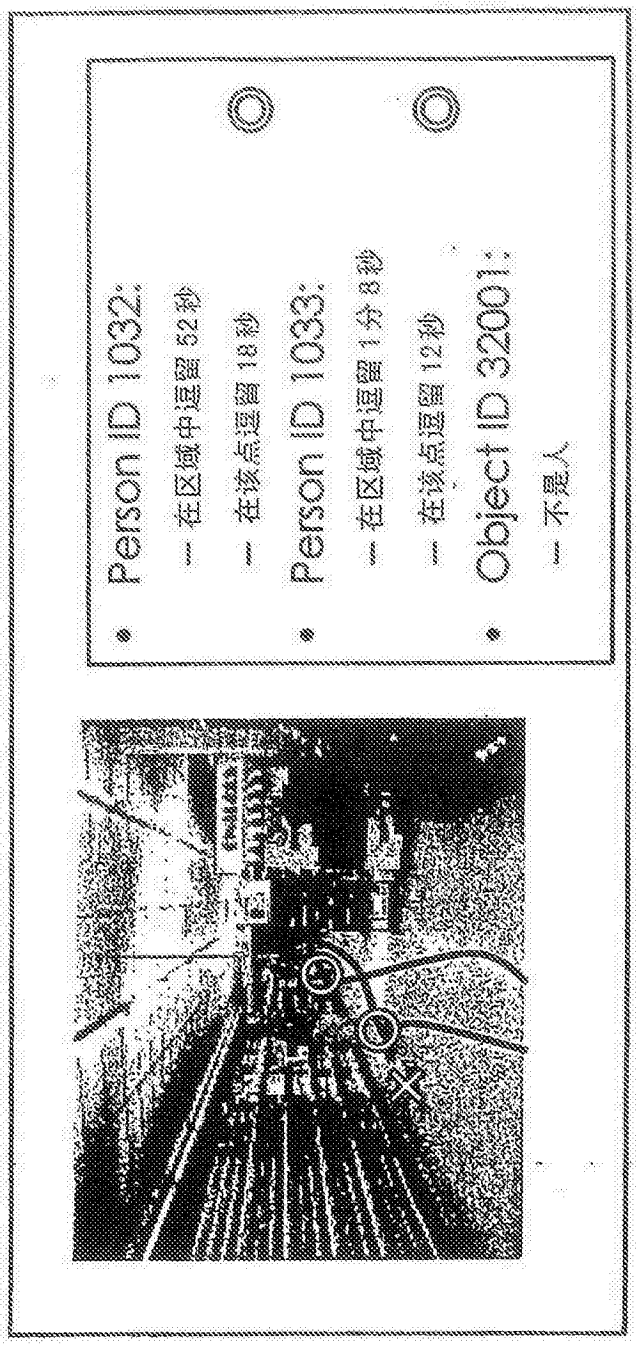

Person ID 1032:
– 在区域中逗留 52 秒
– 在该点逗留 18 秒
Person ID 1033:
– 在区域中逗留 1 分 8 秒
– 在该点逗留 12 秒
Object ID 32001:
– 不是人

图 15

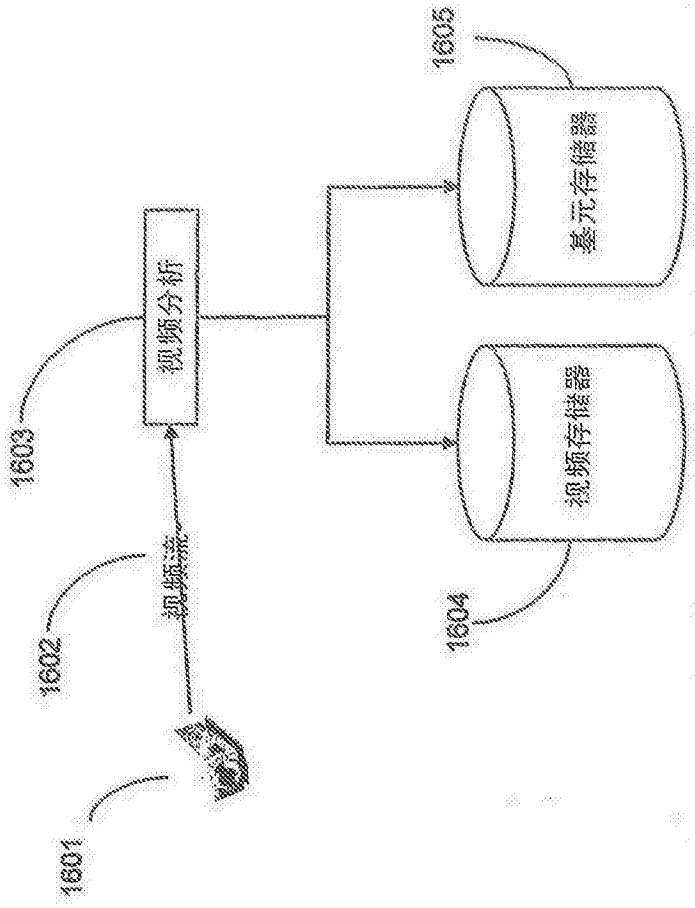


图 16a

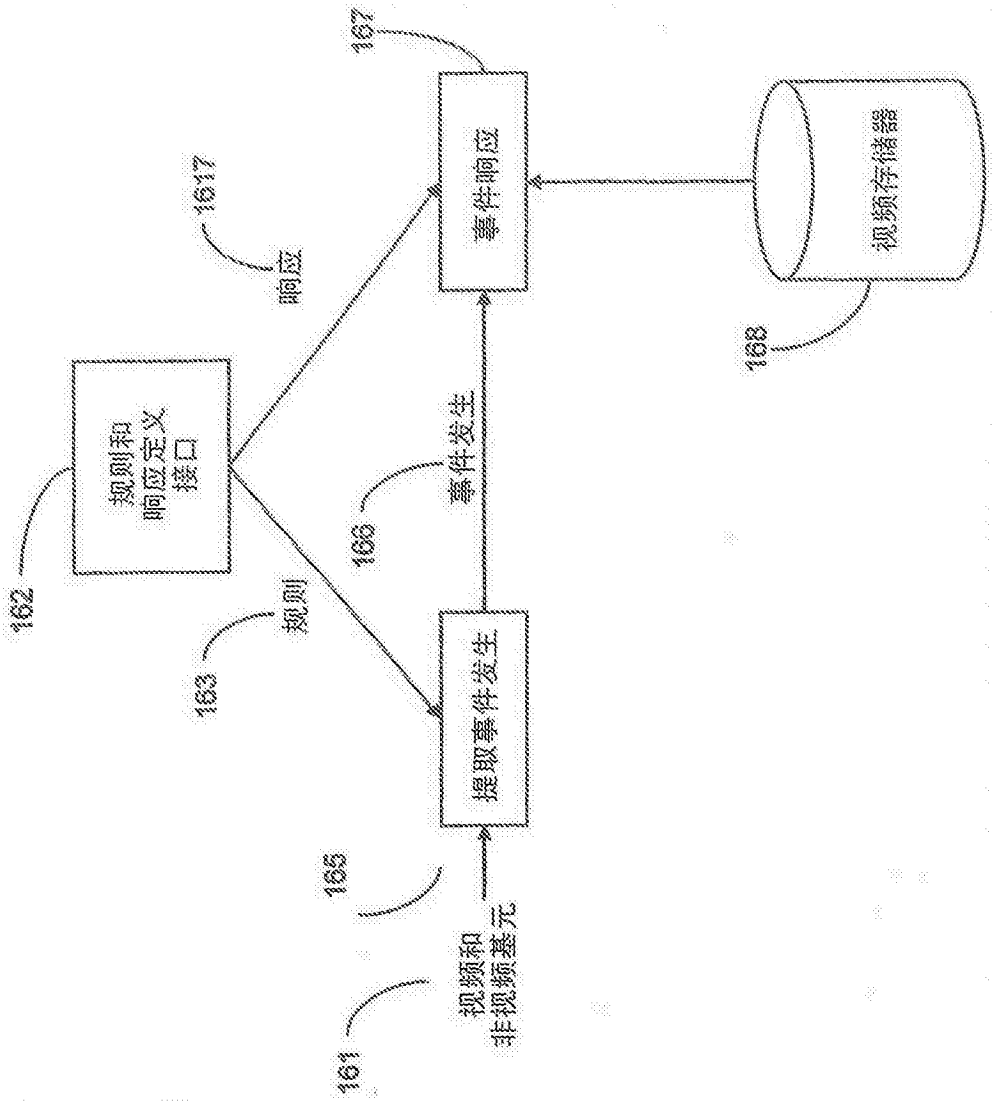

167
事件响应
视频存储器
1617
响应
168
规则和
响应定义
接口
事件发生
166
162
163
规则
提取事件发生
165
视频和
非视频基元
161

图 16b

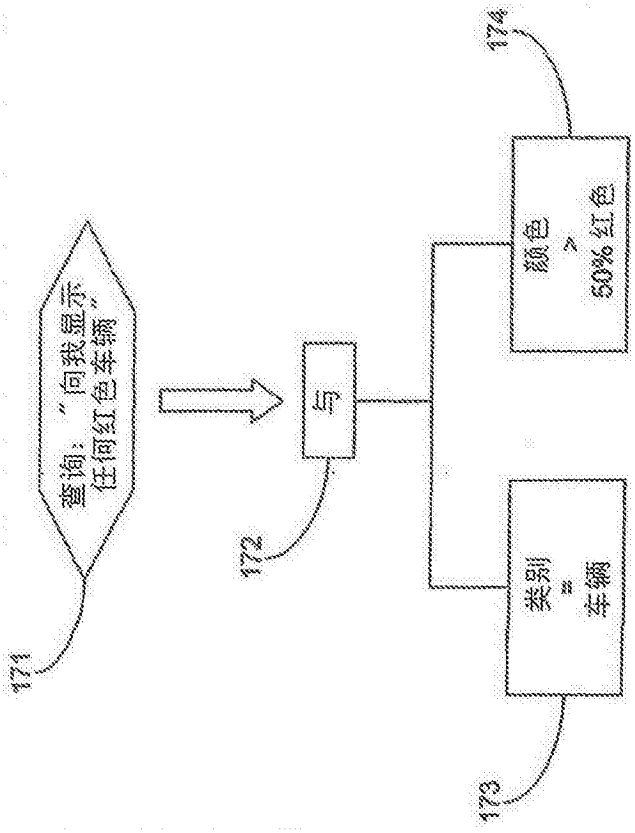


图 17a

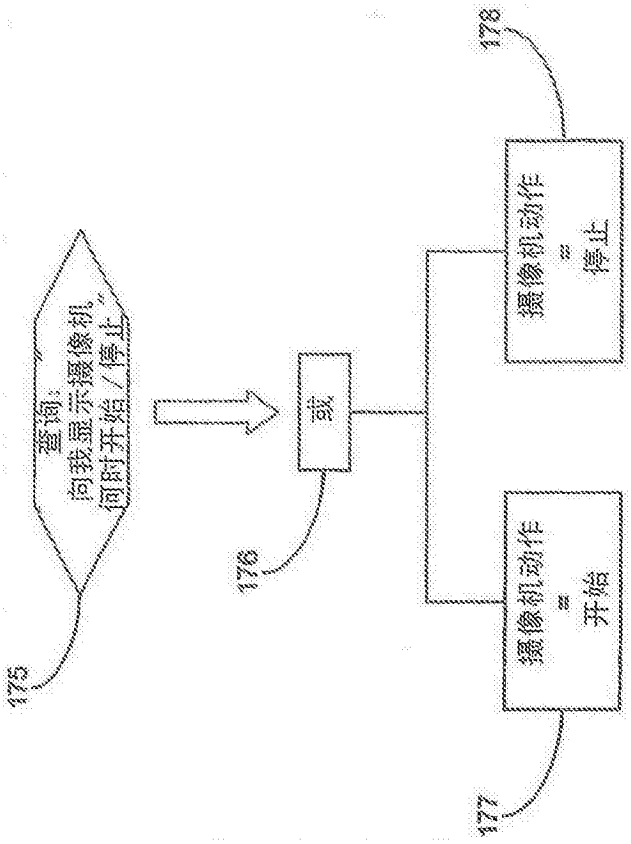


图 17b

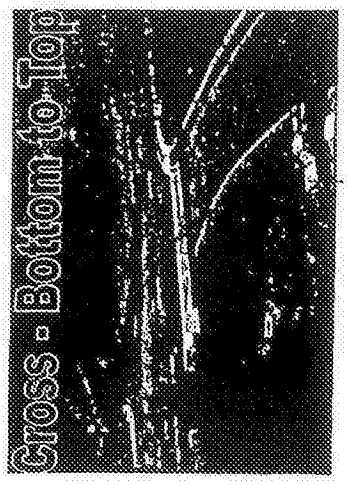


图 18a

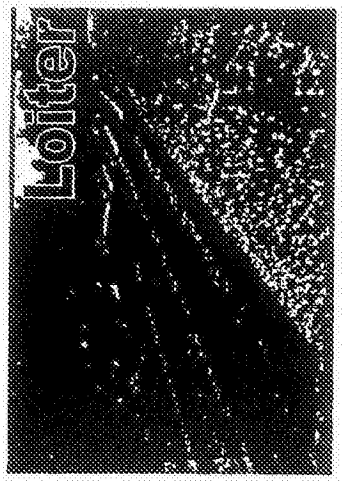


图 18b

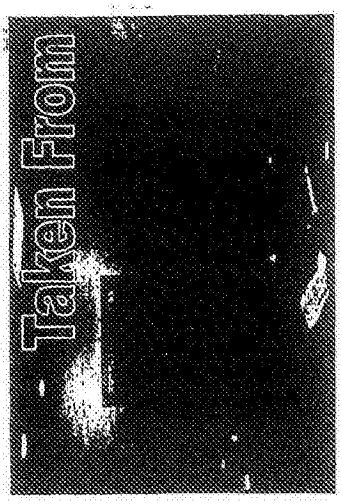


图 18c

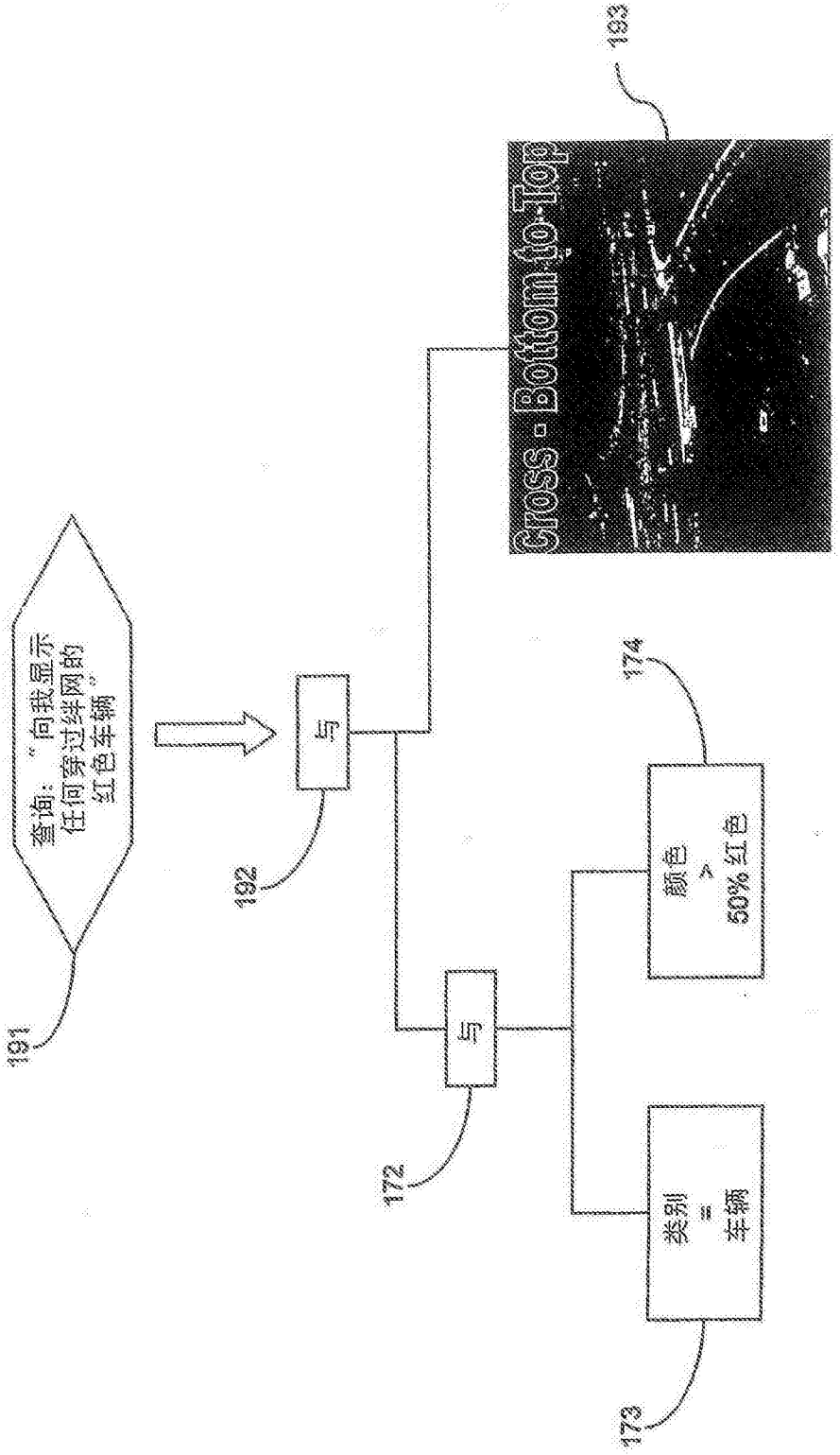

查询："向我显示
任何穿过绊网的
红色车辆"。
191
与
192
193
Cross - Bottom-to-Top
与
172
颜色
>
50% 红色
174
类别
=
车辆
173

图 19

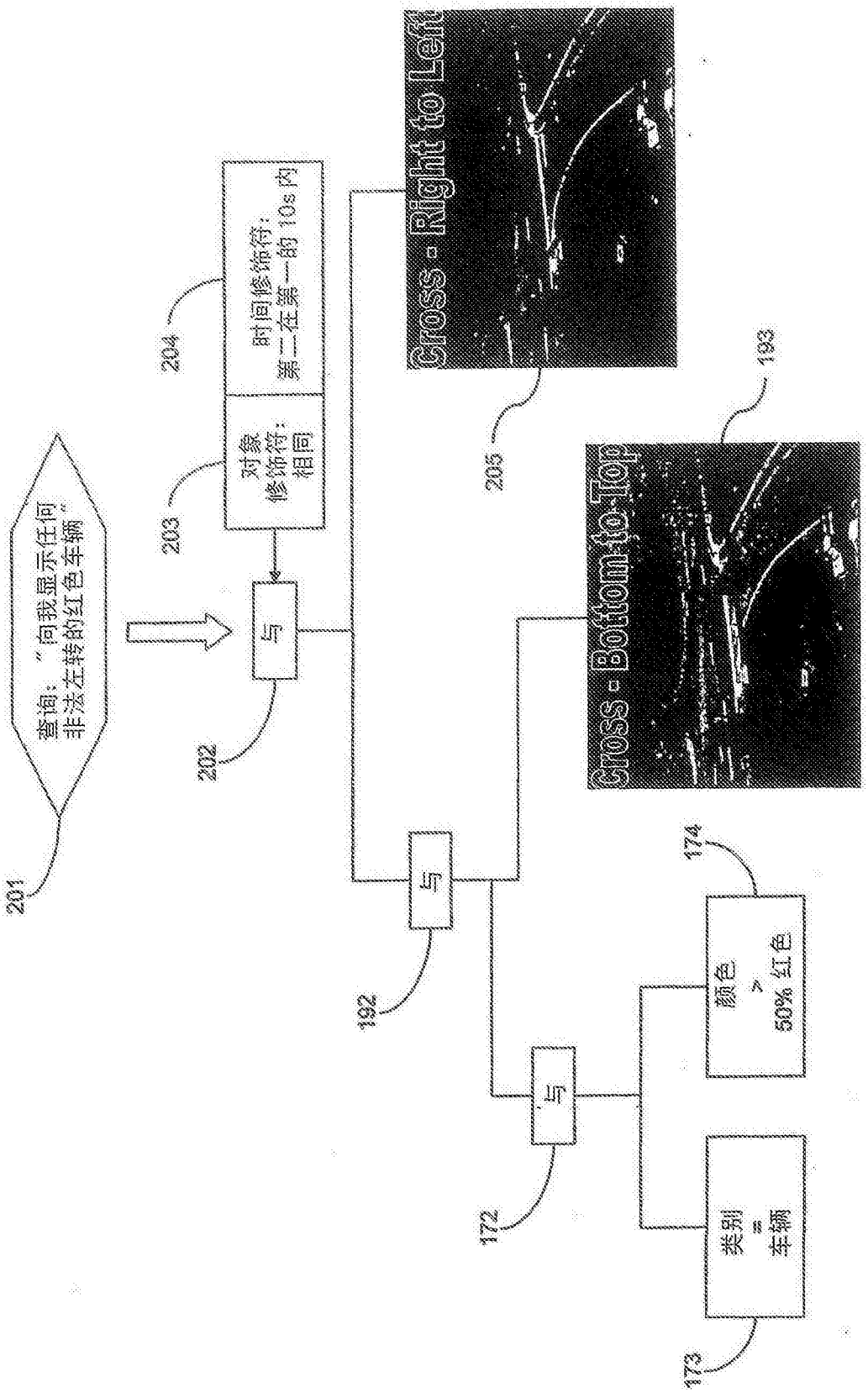

查询："向我显示任何非法左转的红色车辆"
201
与
202
对象修饰符：相同
203
时间修饰符：第二在第一的 10s 内
204
Cross - Right to Left
205
Cross - Bottom-to-Top
193
与
192
与
172
颜色 > 50% 红色
174
类别 = 车辆
173

图 20

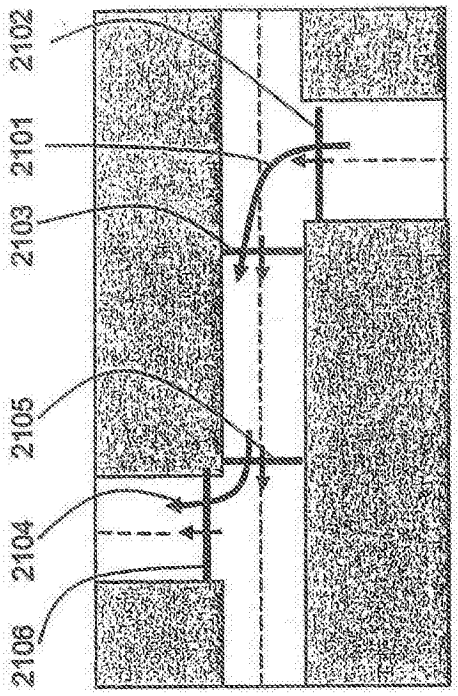


图 21a

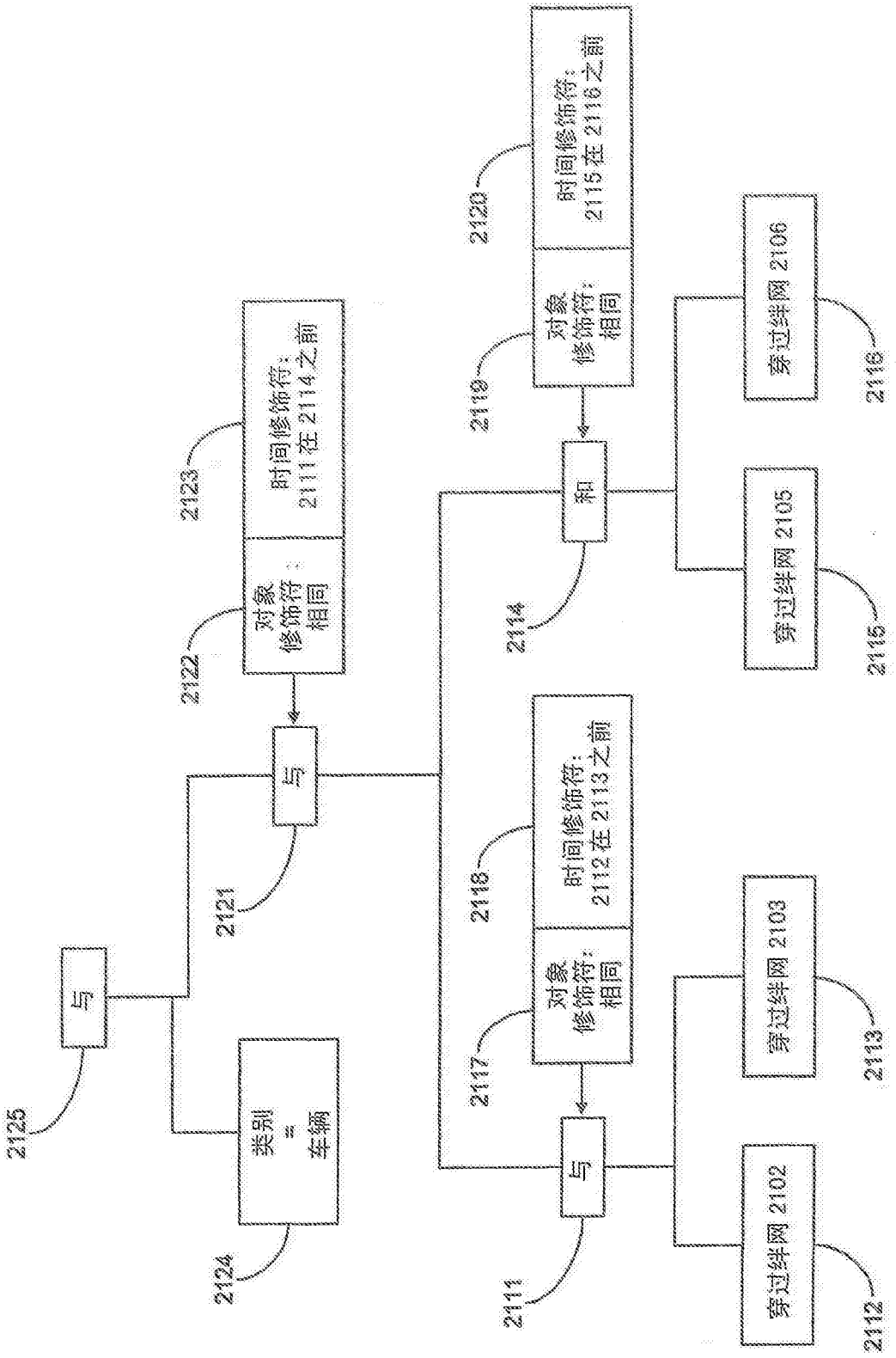

穿过绊网 2102
穿过绊网 2103
穿过绊网 2105
穿过绊网 2106
时间修饰符：2112 在 2113 之前
时间修饰符：2111 在 2114 之前
时间修饰符：2115 在 2116 之前
对象修饰符：相同
对象修饰符：相同
对象修饰符相同
与
与
与
和
类别 = 车辆
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125

图 21b

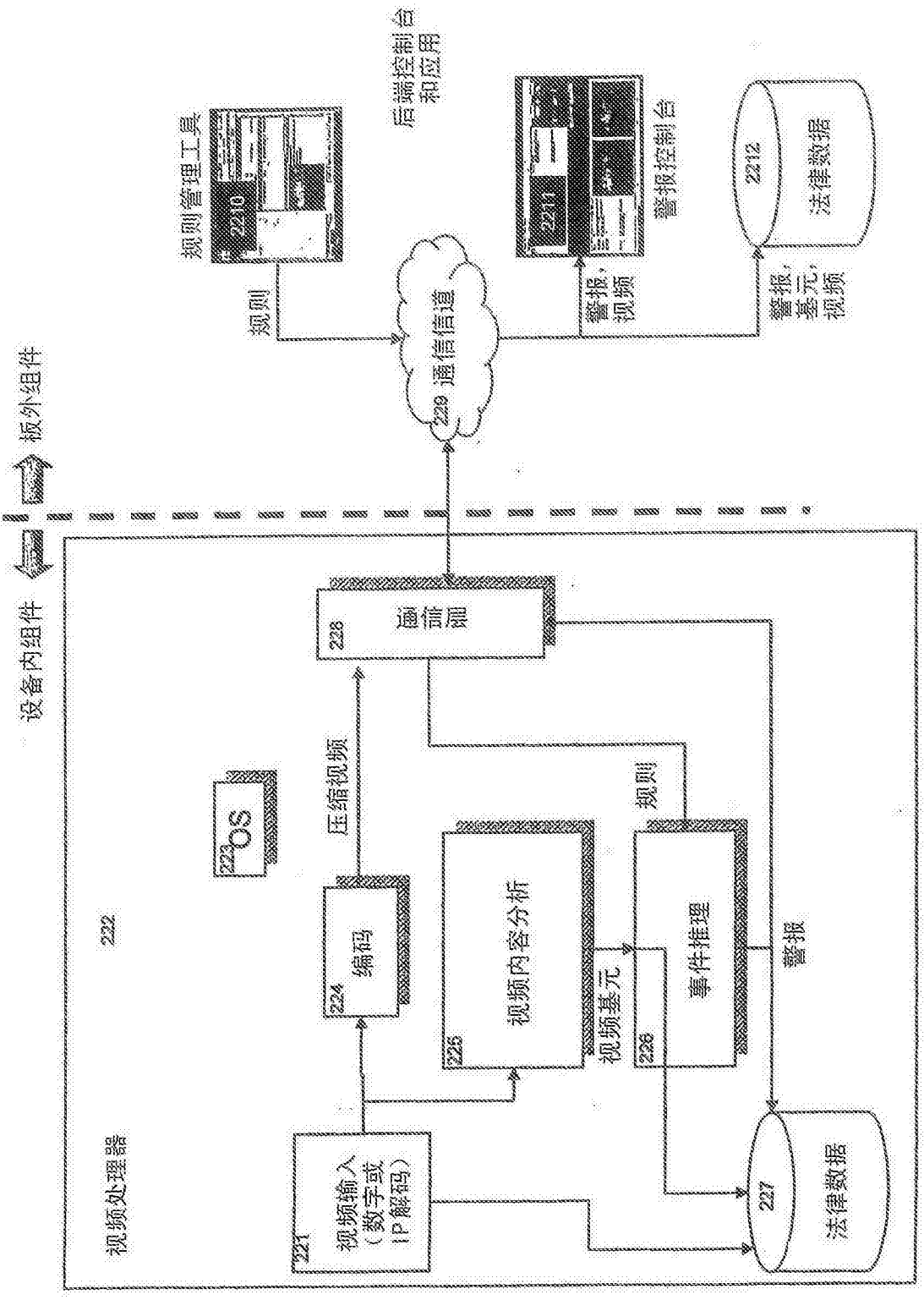

后端控制台和应用
规则管理工具
2210
规则
通信信道
239
警报控制台
2211
警报，视频
2212
法律数据
警报，基元，视频
板外组件
设备内组件
222
视频处理器
223
OS
通信层
228
压缩视频
规则
编码
224
视频内容分析
225
视频基元
事件推理
226
警报
视频输入（数字或IP解码）
221
227
法律数据

图 22

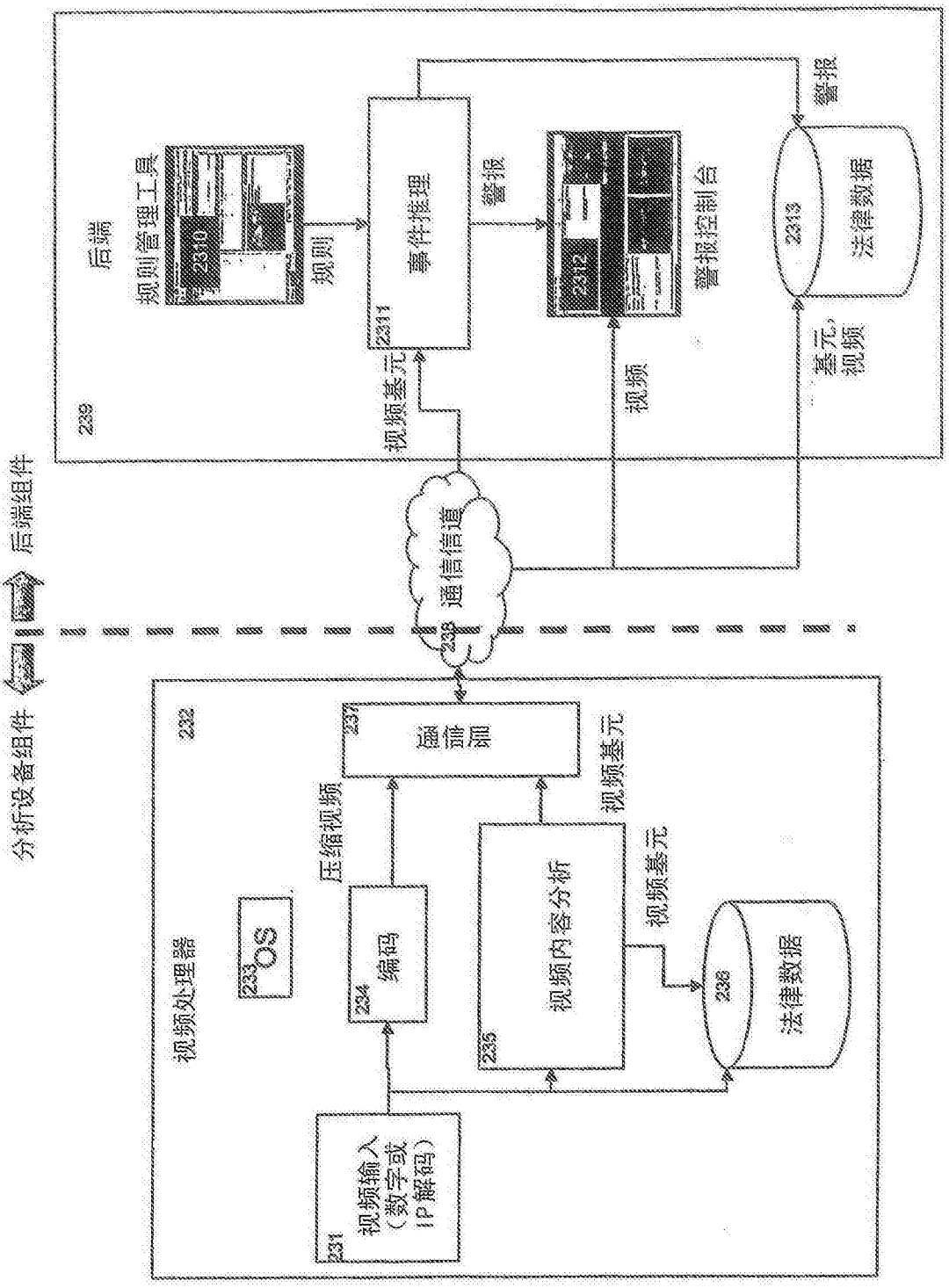

后端组件
分析设备组件
后端
规则管理工具
2310
规则
事件推理
2311
警报
警报控制台
2312
警报
2313
法律数据
239
视频基元
视频
基元，视频
通信信道
238
232
通信层
237
压缩视频
视频基元
视频处理器
OS
233
编码
234
视频内容分析
235
视频基元
236
法律数据
视频输入（数字或IP解码）
231

图 23

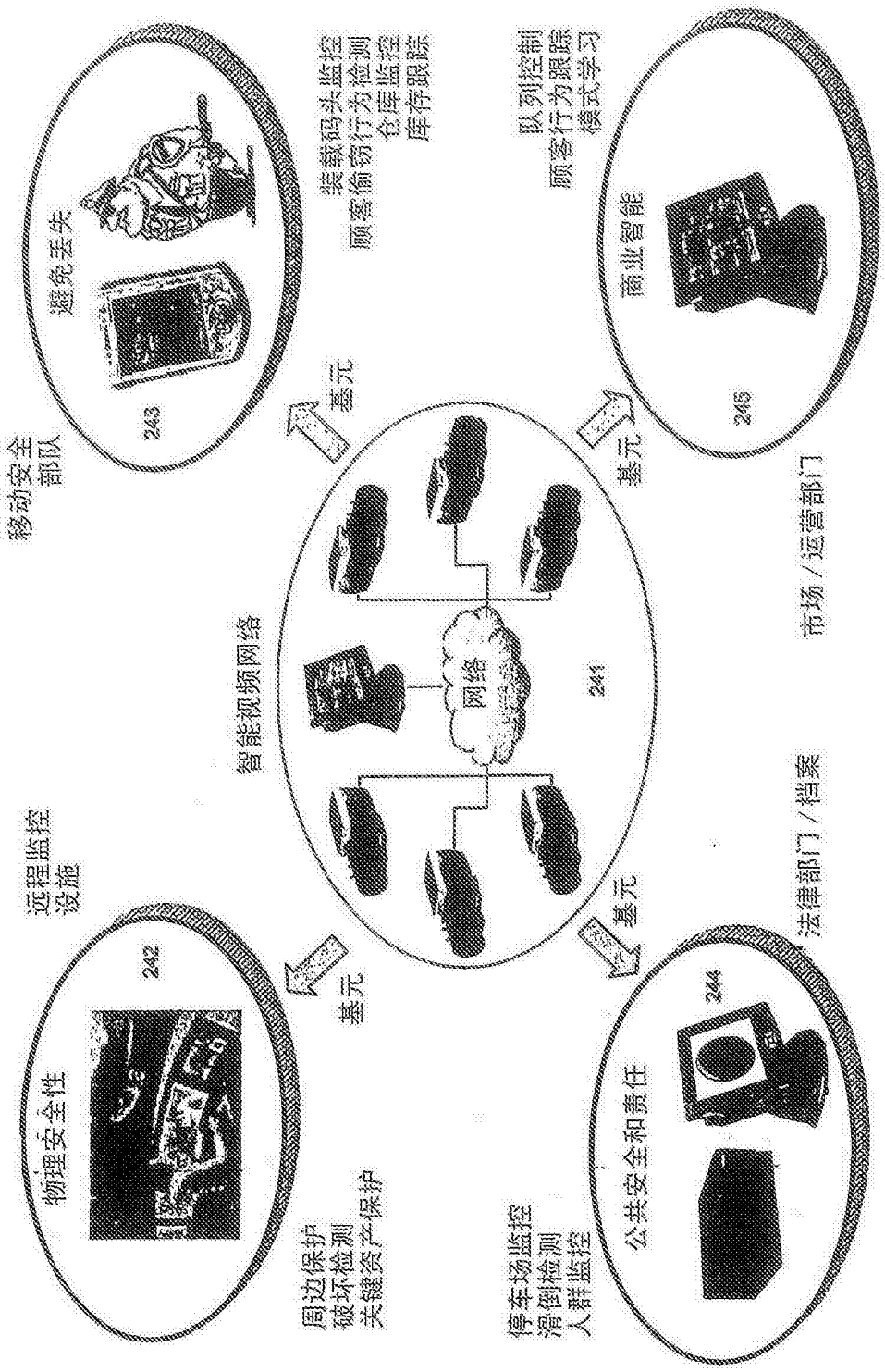

移动安全部队
避免丢失
243
装载码头监控
顾客偷窃行为检测
仓库监控
库存跟踪
队列控制
顾客行为跟踪
模式学习
商业智能
245
基元
基元
智能视频网络
网络
241
基元
基元
远程监控设施
物理安全性
242
周边保护
破坏检测
关键资产保护
停车场监控
滑倒检测
人群监控
公共安全和责任
244
法律部门 / 档案
市场 / 运营部门

图 24

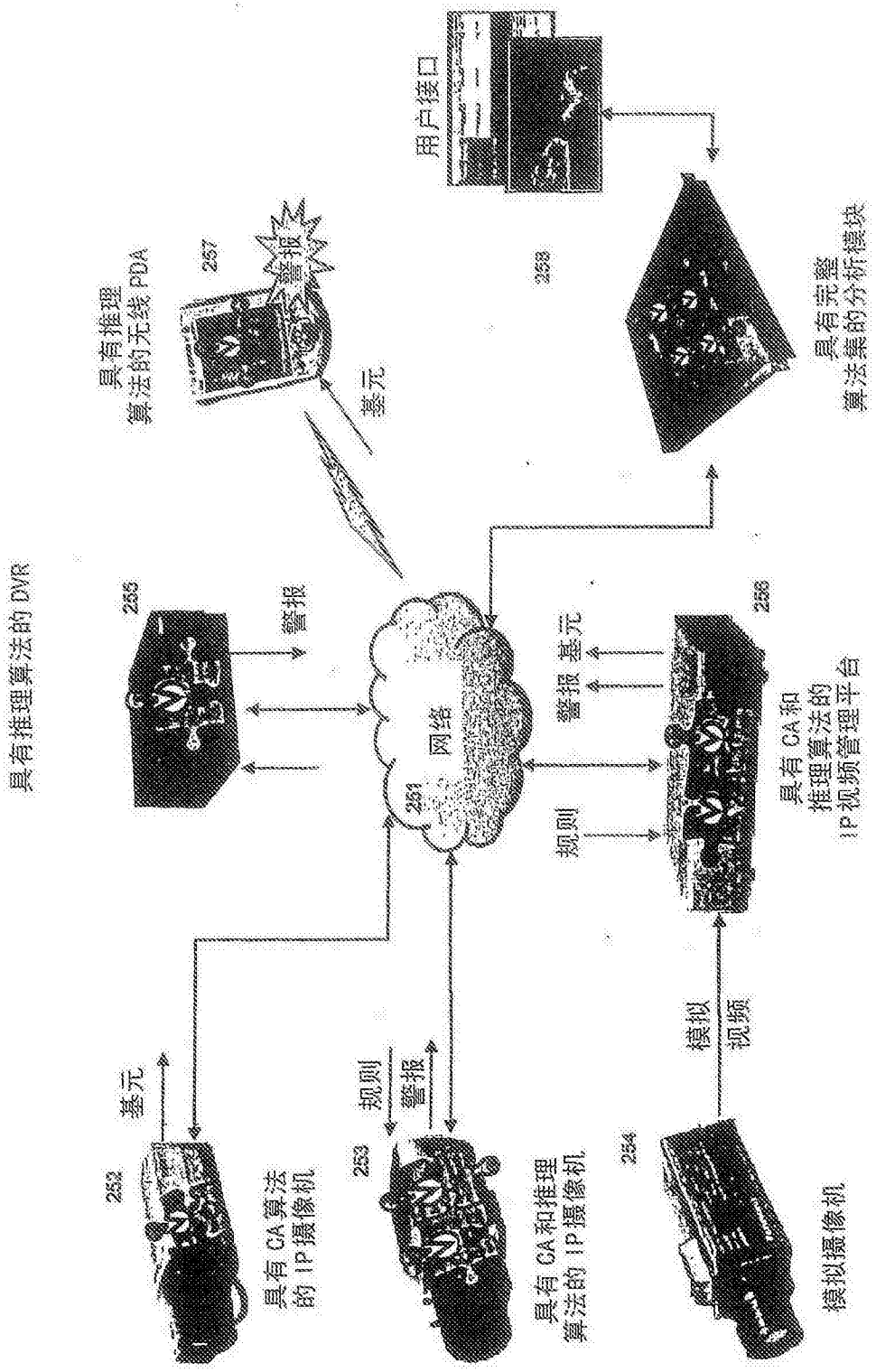

用户接口
具有推理
算法的无线 PDA
257
警报
基元
258
具有完整
算法集的分析模块
具有推理算法的 DVR
255
警报
网络
251
警报 基元
规则
256
具有 CA 和
推理算法的
IP 视频管理平台
基元
252
具有 CA 算法
的 IP 摄像机
规则
警报
253
具有 CA 和推理
算法的 IP 摄像机
模拟
视频
254
模拟摄像机

图 25