



(12) 发明专利申请

(10) 申请公布号 CN 104023208 A

(43) 申请公布日 2014. 09. 03

(21) 申请号 201410247458. 7

(22) 申请日 2014. 06. 05

(71) 申请人 北京小鱼儿科技有限公司

地址 100088 北京市海淀区花园路 5 号院 2
号楼 2031 室

(72) 发明人 宋晨枫

(74) 专利代理机构 北京汉昊知识产权代理事务
所（普通合伙） 11370

代理人 朱海波

(51) Int. Cl.

H04N 7/18 (2006. 01)

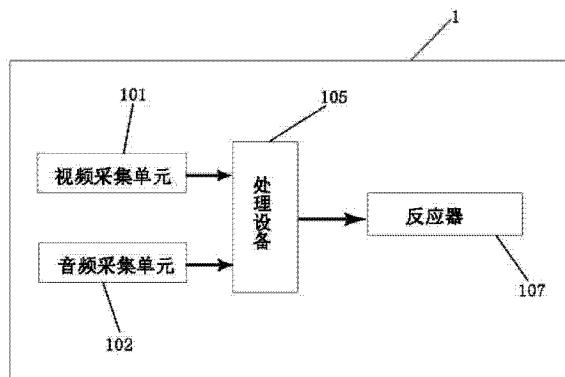
权利要求书1页 说明书8页 附图4页

(54) 发明名称

一种自动监测与自主反应的装置及方法

(57) 摘要

本发明公开了一种自动监测与自主反应的装置及方法。该装置包括：视频采集单元，用于实时采集并发送视频信息；音频采集单元，用于实时采集并发送音频信息；处理设备，用于响应于所接收的视频、音频信息，对视频、音频信息的内容进行识别并发出指令信息；反应器，用于接收所述指令信息并按照所述指令信息作出反应。与现有技术相比，本发明使得摄像监控系统监视到某些情况时能够自动处理和自主反应，而不需要人的干预。



1. 一种自动监测与自主反应装置 (1), 包括 :

视频采集单元 (101), 用于实时采集并发送视频信息 ;

音频采集单元 (102), 用于实时采集并发送音频信息 ;

处理设备 (105), 用于响应于所接收的视频、音频信息, 对视频、音频信息的内容进行识别并发出指令信息 ;

反应器 (107), 用于接收所述指令信息并按照所述指令信息作出反应。

2. 根据权利要求 1 所述的自动监测与自主反应装置 (1), 其中所述反应器 (107) 包括以下中的一个或多个 :

致动器 (1071), 用于响应于所述指令信息是要求产生特定反应动作的信息, 产生特定反应动作 ;

扬声器 (1072), 用于响应于所述指令信息是要求发出特定反应声音的信息, 发出特定反应声音 ;

显示器 (1073), 用于响应于所述指令信息是要求显示特定反应画面的信息, 显示特定反应画面。

3. 根据权利要求 2 所述的自动监测与自主反应装置 (1), 所述处理设备 (105) 體合或包含有存储器 (1051), 用于存储监测和反应记录, 且响应于调取请求, 处理设备 (105) 调取存储器 (1051) 上的相关记录。

4. 根据权利要求 3 所述的自动监测与自主反应装置 (1), 所述处理设备 (105) 从视频采集单元 (101)、音频采集单元 (102) 采集到的视频、音频信息中识别出人的特定动作或特定人的特定动作, 并发出与该人的特定动作或特定人的特定动作对应的指令信息。

5. 根据权利要求 4 所述的自动监测与自主反应装置 (1), 其中所述人或特定人是基于人脸识别、身高识别、声音识别中的一个或多个来识别的。

6. 根据权利要求 4 所述的自动监测与自主反应装置 (1), 其中处理设备 (105) 还接收携带手机发出的无线信号, 基于该无线信号中标明的携带手机的身份, 来识别人或特定人。

7. 根据权利要求 4 所述的自动监测与自主反应装置 (1), 其中特定动作是通过事先为特定动作建立模型, 并从视频采集单元 (101)、音频采集单元 (102) 分别采集到的视频、音频信息中搜索与建立的模型的匹配识别的。

8. 根据权利要求 7 所述的自动监测与自主反应装置 (1), 其中所述模型是通过自学习的方式产生的。

9. 根据权利要求 7 所述的自动监测与自主反应装置 (1), 其中所述模型是预先输入的标准化的模型。

10. 根据权利要求 4 所述的自动监测与自主反应装置 (1), 还包括深度传感器, 所述特定动作是基于视频采集单元 (101)、音频采集单元 (102) 分别采集到的视频、音频以及深度传感器感测的深度识别的。

一种自动监测与自主反应的装置及方法

技术领域

[0001] 本发明涉及通信技术，尤其涉及一种自动监测与自主反应的装置及方法。

背景技术

[0002] 现有技术中，存在着一种家庭摄像监控系统。在家庭中安装有摄像头，在监视端的屏幕上显示采集到的视频。但是，这种家庭摄像监控系统仅能执行监视功能。如果监视到某种情况需要处理时，仍需要人来处理。

发明内容

[0003] 本发明解决的技术问题之一是使得摄像监控系统监视到某些情况时能够自动处理和自主反应，而不需要人的干预。

[0004] 根据本发明的一个方面的一个实施例，提供了一种自动监测与自主反应装置，包括：视频采集单元、音频采集单元、处理设备以及反应器。视频采集单元、音频采集单元分别采集并发送视、音频信息。处理设备响应于所接收的视、音频信息，对视、音频信息的内容进行识别并向反应器发出指令信息。反应器接收到指令信息，按照指令信息作出反应。

[0005] 根据本发明的一个实施例，所述反应器包括以下中的一个或多个：致动器，用于响应于所述指令信息是要求产生特定反应动作的信息，产生特定反应动作；扬声器，用于响应于所述指令信息是要求发出特定反应声音的信息，发出特定反应声音；显示器，用于响应于所述指令信息是要求显示特定反应画面的信息，显示特定反应画面。

[0006] 根据本发明的一个实施例，所述处理设备耦合或包含有存储器，用于存储监测和反应记录，且响应于调取请求，处理设备调取存储器上的相关记录。

[0007] 根据本发明的一个实施例，所述处理设备从视频采集单元、音频采集单元采集到的视频、音频信息中识别出人的特定动作或特定人的特定动作，并发出与该人的特定动作或特定人的特定动作对应的指令信息。

[0008] 根据本发明的一个实施例，所述人或特定人是基于人脸识别、身高识别、声音识别中的一个或多个来识别的。

[0009] 根据本发明的一个实施例，处理设备还接收携带手机发出的无线信号，基于该无线信号中标明的携带手机的身份，来识别人或特定人。

[0010] 根据本发明的一个实施例，特定动作是通过事先为特定动作建立模型，并从视频采集单元、音频采集单元分别采集到的视频、音频信息中搜索与建立的模型的匹配识别的。

[0011] 根据本发明的一个实施例，所述模型是通过自学习的方式产生的。

[0012] 根据本发明的一个实施例，所述模型是预先输入的标准化的模型。

[0013] 根据本发明的一个实施例，自动监测与自主反应装置还包括深度传感器，所述特定动作是基于视频采集单元、音频采集单元分别采集到的视频、音频以及深度传感器感测的深度识别的。

[0014] 根据本发明的一个实施例，处理设备响应于识别出的特定动作和 / 或特定人的特

定动作，匹配存储器中的模型，预测下一步可能产生的动作，向致动器发出与该预测的动作对应的指令信息，致动器基于该指令信息产生该预测的动作。

[0015] 根据本发明的一个实施例，存储器还存储与特定动作和 / 或特定人的特定动作对应的限制条件、以及如果该限制条件被满足应采取的强制动作，处理设备响应于识别出的特定动作和 / 或特定人的特定动作，检索与该特定动作和 / 或特定人的特定动作对应的限制条件，判断该限制条件是否被满足，且响应于判断出该限制条件未被满足，向致动器发出对应于该强制动作的指令信息，致动器基于该指令信息产生该强制的动作。

[0016] 根据本发明的一个实施例，存储器还存储与特定动作和 / 或特定人的特定动作对应的限制条件，处理设备响应于识别出的特定动作和 / 或特定人的特定动作，检索与该特定动作和 / 或特定人的特定动作对应的限制条件，判断该限制条件是否被满足，且响应于判断出该限制条件未被满足，向扬声器或显示器发出指示提醒的指令信息，扬声器或显示器基于该指令信息发出提醒声音或显示提醒画面。

[0017] 根据本发明的一个实施例，存储器还存储与特定动作和 / 或特定人的特定动作对应的互动声音或画面，处理设备响应于识别出的特定动作和 / 或特定人的特定动作，检索与该特定动作和 / 或特定人的特定动作对应的互动声音或画面，向扬声器或显示器发出对应于该互动声音或画面的指令信息，由扬声器基于该指令信息产生该互动声音，或由显示器基于该指令信息显示该互动画面。

[0018] 根据本发明的另一个方面的一个实施例，提供了一种自动监测与自主反应的方法，包括以下步骤：采集环境的视频、音频信息；响应于所采集的视频、音频信息，对视频、音频信息的内容进行识别；基于识别结果做出反应。

[0019] 根据本发明的一个实施例，所述反应包括以下中的一个或多个：产生特定反应动作；发出特定反应声音；显示特定反应画面。

[0020] 根据本发明的一个实施例，响应于所采集的视频、音频信息，对视频、音频信息的内容进行识别的内容包括：从采集到的视频、音频信息中识别出特定动作和 / 或特定人的特定动作。

[0021] 根据本发明的一个实施例，所述人或特定人是基于人脸识别、身高识别、声音识别中的一个或多个来识别的。

[0022] 根据本发明的一个实施例，该方法还包括接收携带手机发出的无线信号，且所述人或特定人是基于该无线信号中标明的携带手机的身份来识别的。

[0023] 根据本发明的一个实施例，所述特定动作基于所采集的视频、音频信息和深度传感器感测的深度进行识别。

[0024] 根据本发明的一个实施例，特定动作是通过事先为特定动作建立模型，并从采集到的视频、音频信息中搜索与建立的模型的匹配识别的。

[0025] 根据本发明的一个实施例，所述模型是通过自学习的方式产生的。

[0026] 根据本发明的一个实施例，所述模型是预先输入的标准化的模型。

[0027] 由于本发明的一个实施例能够对采集到的视、音频信息内容进行识别，并按照识别出的视、音频信息的内容作出反应，因此实现了摄像监控系统监视到某些情况时能够自动处理和反应而不需要人的干预的效果。

[0028] 本领域普通技术人员将了解，虽然下面的详细说明将参考图示实施例、附图进行，

但本发明并不仅限于这些实施例。而是，本发明的范围是广泛的，且意在仅通过后附的权利要求限定本发明的范围。

附图说明

[0029] 通过阅读参照以下附图所作的对非限制性实施例所作的详细描述，本发明的其它特征、目的和优点将会变得更明显：

- [0030] 图 1 示出了根据本发明一个实施例的自动监测与自主反应装置的示意性框图；
- [0031] 图 2 示出了根据本发明一个实施例的自动监测与自主反应装置的外部正视图；
- [0032] 图 3 示出了根据本发明一个实施例的自动监测与自主反应装置的外部左视图；
- [0033] 图 4 示出了根据本发明一个实施例的自动监测与自主反应方法的流程图；
- [0034] 图 5 示出了根据本发明的一个实施例的行为模式触发的预测执行系统的示意性框图；
- [0035] 图 6 示出了根据本发明的一个实施例的行为模式触发的纠正致动系统的示意性框图；
- [0036] 图 7 示出了根据本发明的一个实施例的行为模式触发的报警系统的示意性框图；
- [0037] 图 8 示出了根据本发明的一个实施例的行为模式触发的互动系统的示意性框图。
- [0038] 附图中相同或相似的附图标记代表相同或相似的部件。

具体实施方式

[0039] 下面结合附图对本发明作进一步详细描述。

[0040] 图 1 示出了根据本发明一个实施例的自动监测与自主反应装置 (nemo) 1 的示意性框图。根据本发明一个实施例的自动监测与自主反应装置 1 包括视频采集单元 101、音频采集单元 102、处理设备 105 以及反应器 107。视频采集单元 101、音频采集单元 102 分别实时采集视、音频并将其发送给处理设备 105。处理设备 105 响应于所接收的视、音频信息，对视、音频信息的内容进行识别并向反应器 107 发出指令信息。反应器 107 接收到指令信息，按照指令信息作出反应。

[0041] 视频采集单元是指用于视频采集的装置，例如摄像头、摄像机、有摄像功能的手机等。音频采集单元是指用于音频采集的装置，例如麦克风、录音机、有录音功能的手机等。处理设备是指具有数据处理、分析功能的设备，用来接收视频采集单元、音频采集单元发送来的视、音频信息，并对所述视、音频信息进行处理、识别，发出对应指令信息，例如一个 CPU 芯片、一台计算机或多台计算机组的处理中心。

[0042] 所述处理设备耦合或包含有存储器 1051，用于存储监测和反应记录。响应于调取请求，处理设备 105 调取存储器 1051 上的相关记录。

[0043] 如图 5 所示，反应器 107 可以包括有以下中的一个或多个：致动器 1071、扬声器 1072、显示器 1073。致动器 1071 例如是窗帘开关、电灯开关、热水器开关等，用于响应于所述指令信息是要求产生特定反应动作的信息，产生特定反应动作。所述扬声器 1072 例如是电脑音箱、外设独立音箱、扩音器等，用于响应于所述指令信息是要求发出特定反应声音的信息，发出特定反应声音。所述显示器 1073 例如是电脑屏幕、电视屏幕、外接显示屏等，用于响应于所述指令信息是要求显示特定反应画面的信息，显示特定反应画面。

[0044] 反应器中包括的多种类型增加了反应的种类和灵活性,提高了用户体验。

[0045] 图2示出了根据本发明一个实施例的自动监测与自主反应装置的外部正视图。在该实施例中,视频采集单元101是位于显示器1073上端的摄像头。处理设备105封闭在底座中。

[0046] 在工作过程中,视频采集单元101、音频采集单元102将其采集到的视、音频信息发送给处理设备105,处理设备105对接收到的视、音频信息进行识别,识别出人的特定动作或特定人的特定动作,向反应器107发出对应的指令信息,指令致动器1071、扬声器1072、显示器1073中的一个或多个完成反应动作。所述指令信息可以是产生特定动作的指令信息,发出特定声音的指令信息或者是显示特定画面的指令信息。

[0047] 这样,就可以实现自动监视环境中是否出现了特定动作,或特定人是否执行了特定动作,一旦特定动作发生,就产生相应的反应的有益效果。

[0048] 所述自动监测与自主反应装置1可以通过视频采集单元101、音频采集单元102以及其他装置或单元,基于人脸识别、身高识别、声音识别、携带手机发出的无线信号表明的身份中的一个或多个来识别人或特定人。多种识别人的方式结合相比于只靠单一的方式识别,增加了识别人的准确度。

[0049] 在识别人的情况下,由于人脸的模式是很像的,绝大多数人的身高也是在特定范围内、人的声音频率也是在特定范围内,因此,例如当拍摄的图像中的某一区域与存储的人脸的模式类似,且 / 或结合位置传感器和 / 或深度传感器感测到的人脸与自动监测与反应装置1的距离判断出其身高在特定范围内,且 / 或音频采集单元102采集到的音频也在特定范围内,可识别出人的存在。

[0050] 在识别特定人的情况下,可以预先将特定人的人脸的模式和 / 或身高和 / 或声音频率存储在存储器中。当拍摄的图像中的某一区域与存储的该特定人脸的模式匹配,且 / 或结合位置传感器和 / 或深度传感器感测到的该特定人脸与自动监测与自主反应装置1的距离判断出其身高与存储的身高匹配,且 / 或音频采集单元102采集到的音频的匹配与存储的改特定人的声音的频率匹配时,可识别出特定人的存在。

[0051] 识别人或特定人的存在也可以采用自学习的方法。例如,如果拍摄的图像中的某个模式与采集到的声音的某个频率总是同时出现,可以在显示器上显示提示,即识别到了人,请自动监测与自主反应装置1旁的人判断并命名。如果实况自动监测与自主反应装置1旁的人发现识别错误,则在显示器的界面上反馈。接收到这种反馈后,在下一次拍摄的图像中的这种模式与采集到的声音的这种频率同时出现时就不认为出现了人或特定人。在自学习的方式下,也可以预先不将特定人的人脸的模式和 / 或身高和 / 或声音频率存储在存储器中。

[0052] 另外,也可以基于携带手机发出的无线信号表明的身份识别人或特定人。例如自动监测与自主反应装置1具有蓝牙设备,用户的手机中也具有蓝牙无线单元。当自动监测与自主反应装置1识别出特定身份的蓝牙无线单元出现在一定距离内时,则认为识别出了特定人。

[0053] 特定动作是通过事先为特定动作建立模型,并从视频采集单元101、音频采集单元102分别采集到的视频、音频信息中搜索与建立的模型的匹配识别的。

[0054] 可选地,所述模型是预先输入的标准化的模型,即人为地事先设定并根据设定的

动作建立模型。例如,对于看电视这样一个动作,建立一个模型:识别出一个人坐在沙发上;顺着该人的目光方向看去,有一个物体;识别出该物体是电视;该人目光停留在电视上至少10秒。如果从视频采集单元101拍摄的图像中检测到人,然后检测到此人坐在沙发上(沙发的识别类似人脸识别,也可以通过模式匹配进行,也可以将人坐在沙发上的图像整体作为一个对象进行模式匹配识别),然后检测此人的目光方向,然后检测此人目光方向上的物体是否是电视(例如将电视作为一个对象进行模式匹配),如果是则计数10秒。如果达到10秒则认为检测到了看电视这样一个动作。

[0055]当然,所述处理设备105也可以通过机器学习等自学习的方式自动建立动作模型。例如,处理设备105从视频采集单元101、音频采集单元102所采集的视、音频中提取动作特征,并基于提取的特征建立动作模型。例如,从视频采集单元101、音频采集单元102所采集的视、音频中发现有一个人坐在沙发上、顺着此人目光看去的方向有一个电视、在此人目光停留在电视上的事件超过10秒的频率超过阈值,则认为这是一个特定动作的模型。在这种情况下,动作模型可以不预先存储在存储器中,而是根据从视频采集单元101、音频采集单元102所采集的视、音频以学习的方式提取动作的模型。

[0056]为了更准确地识别出特定动作,所述自动监测与自主反应装置1还包括深度传感器197,由视频采集单元101、音频采集单元102以及深度传感器通过采集的视、音频以及感测的深度共同识别出特定动作。虽然在图2中深度传感器197位于显示器上部边框中心偏左的位置,其也可以设置在其他合理的位置。

[0057]深度传感器197感测人或物体与自动监测与自主反应装置1的距离。当人或物体发生一个动作的时候,同样的动作幅度由于与自动监测与自主反应装置1的距离不同在拍摄到的图像中产生的变化幅度会是不同的。因此,结合深度传感器,对动作能够进行更准确的识别,从而提高识别精度。

[0058]图3示出了根据本发明一个实施例的自动监测与自主反应装置的外部左视图。如图3所示,为了更好地采集信息,自动监测与自主反应装置1还可包括转动装置199,用于使视频采集单元101转动。优选地,响应于从视频采集单元101、音频采集单元102分别采集到的视、音频中识别出以下要素中的一个,转动装置199使视频采集单元101向着面对识别出的要素的方向转动:人或特定人;特定动作;异常状况。

[0059]在一个实施例中,图3所示的视频采集单元101可以向着识别出的要素左右转动。在另一个实施例中,图3所示的视频采集单元101可以向着识别出的要素上下左右转动。

[0060]继续参考图2,如图2所示,自动监测与自主反应装置1还可包括:光线传感器198,用于感测自动监测与自主反应装置1周围环境光线的变化,其中显示器1073的显示亮度是根据所述光线的变化调整的。如果周围光线比较强,可以将显示器的显示亮度增加。如果周围光线比较弱,可以将显示器的显示亮度减少。这样,可以减少眼睛观看显示器的不舒适感。

[0061]虽然图2中的光线传感器位于显示器上边框的中心偏右的位置处,但其也可以设置在任何其他的合理的位置处。

[0062]以下是本发明实施例的几个典型应用场景。

[0063]I. 行为模式触发的预测执行系统

[0064]图5示出了根据本发明的一个实施例的行为模式触发的预测执行系统的示意性

框图。在本实施例中，用户每次回到家里，每次都是先换拖鞋，然后走到窗帘处将窗帘拉开。本实施例的自动监测与自主反应装置 1 就会学习用户的行为模式，在用户回到家里，换拖鞋后，帮用户自动将窗帘拉开。

[0065] 在此实施例中，致动器 1071 是窗帘开关。

[0066] 图 5 的行为模式触发的预测执行的具体过程如下。存储器 1051 中存储着视频采集单元 101、音频采集单元 102 等识别出的各个人的各个动作的记录。尽管一些动作不会产生反应器的反应，但这些动作仍然被记录在存储器 1051 中。当一些动作连续出现的次数超过一个预设的阈值，就把这些连续的动作记录在存储器 1051 中记录为模型。例如，当用户进门——换拖鞋——拉窗帘这一系列的动作出现超过 10 次，则将用户进门——换拖鞋——拉窗帘记录为模型。这时，当处理设备 105 根据视频采集单元 101、音频采集单元 102 等采集的视、音频识别出家里人或家里的特定人进门、然后换拖鞋，则认为与存储器中存储的用户进门——换拖鞋——拉窗帘这一模型匹配，从而预测出下一步可能产生的动作是拉窗帘。因此，处理设备 105 向窗帘开关发出拉窗帘的指令信息，窗帘开关基于该指令信息拉上窗帘。

[0067] 类型的应用场景还有例如，用户进门，走进洗手间的洗澡龙头下，摘下墙上挂着的洗澡用品，自动帮助用户打开洗澡龙头等。

[0068] II. 行为模式触发的纠正致动系统

[0069] 图 6 示出了根据本发明的一个实施例的行为模式触发的纠正致动系统的示意性框图。在本实施例中，家长预先设定小孩看电视不能超过 3 个小时。如果超过 3 个小时，电视自动关机。

[0070] 在此实施例中，致动器 1071 是电视开关。

[0071] 图 6 的行为模式触发的纠正致动系统的工作过程如下。

[0072] 存储器 1051 存储与特定动作和 / 或特定人的特定动作对应的限制条件、以及如果该限制条件被满足应采取的强制动作。例如，前面提到的小孩看电视——不超过 3 小时——自动关机。

[0073] 处理设备响应于从视频采集单元 101、音频采集单元 102 采集的视、音频信息识别出特定动作和 / 或特定人的特定动作，检索与该特定动作和 / 或特定人的特定动作对应的限制条件，判断该限制条件是否被满足，且响应于判断出该限制条件未被满足，向致动器 1071 发出对应于该强制动作的指令信息，致动器 1071 基于该指令信息产生强制动作。例如，处理设备识别出小孩在看电视后，在存储器 1051 中检索出小孩看电视——不超过 3 小时——自动关机这样一条信息，然后监视小孩看电视的时长，当超过 3 小时后，向电视开关发出关电视的指令。电视开关基于该指令关电视。

[0074] 其它的应用场景例如还有，睡觉超过 8 小时自动开灯开窗让用户起床等。

[0075] III. 行为模式触发的报警系统

[0076] 图 7 示出了根据本发明的一个实施例的行为模式触发的报警系统的示意性框图。在本实施例中，当小孩看电视的时间超过 3 小时后，不是强制性关上电视，而是由扬声器发出提醒。

[0077] 行为模式触发的报警系统的工作过程如下。

[0078] 存储器 1051 存储与特定动作和 / 或特定人的特定动作对应的限制条件。例如，前面提到的小孩看电视——不超过 3 小时。

[0079] 处理设备 105 响应于识别出的特定动作和 / 或特定人的特定动作,检索与该特定动作和 / 或特定人的特定动作对应的限制条件,判断该限制条件是否被满足,且响应于判断出该限制条件未被满足,向扬声器 1072 发出指示提醒的指令信息。扬声器 1072 基于该指令信息发出提醒声音。例如,处理设备 105 识别出小孩在看电视后,在存储器 1051 中监视出小孩看电视——不超过 3 小时这一信息,监视小孩看电视的时长。当判断出超过 3 小时后,向扬声器发出指示,指示扬声器发出提醒。该提醒例如嘟嘟声。

[0080] 另外,也可以录制好一段专用于该场景的提醒,例如“你看电视已经超过 3 小时了,该歇歇了”,与小孩看电视——不超过 3 小时相关联地存储在存储器中。处理设备 105 识别出小孩在看电视后,在存储器 1051 中监视出小孩看电视——不超过 3 小时——“你看电视已经超过 3 小时了,该歇歇了”这一信息,监视小孩看电视的时长。当判断出超过 3 小时后,向扬声器发出指示,指示扬声器发出“你看电视已经超过 3 小时了,该歇歇了”的语音提醒。

[0081] 虽然图 7 中示出了扬声器 1072,但也可以用显示器 1073 取代扬声器 1072。当处理设备向显示器发出提示提醒的指令信息后,显示器基于该指令显示提醒画面。

[0082] IV. 行为模式触发的互动系统

[0083] 图 8 示出了根据本发明的一个实施例的行为模式触发的互动系统的示意性框图。在本实施例中,当家里的人走到洗衣机前时,显示器 1073 上自动显示洗衣机的操作步骤,供该人参考。

[0084] 行为模式触发的互动系统的工作过程如下。

[0085] 存储器 1051 存储与特定动作和 / 或特定人的特定动作对应的互动画面。例如人走到洗衣机面前——包含洗衣机的操作步骤的画面。

[0086] 处理设备 105 响应于从视频采集单元 101、音频采集单元 102 采集到的视、音频中识别到特定动作和 / 或特定人的特定动作,检索与该特定动作和 / 或特定人的特定动作对应的互动画面,向显示器 1073 发出对应于该互动画面的指令信息,由显示器 1073 基于该指令信息显示该互动画面。例如,当处理设备 105 识别出人走到洗衣机面前时,在存储器 1051 中检索到人走到洗衣机面前——包含洗衣机的操作步骤的画面这一条目,向显示器 1073 发指令信息,由显示器显示包含洗衣机的操作步骤的画面。

[0087] 当然,显示器 1073 可以被扬声器 1072 代替。即,存储器 1051 存储与特定动作和 / 或特定人的特定动作对应的互动声音。处理设备 105 响应于识别出的特定动作和 / 或特定人的特定动作,检索与该特定动作和 / 或特定人的特定动作对应的互动声音,向扬声器 1072 或显示器 1073 发出对应于该互动声音的指令信息,由扬声器 1072 基于该指令信息产生该互动声音。

[0088] 根据本发明的一个实施例,为了更好地采集信息,视频采集单元 101 是可转动的。优选地,所述自动监测与自主反应装置 1 基于视频采集单元 101、音频采集单元 102 所采集的视、音频识别出以下要素中的一个,视频采集单元 101 向着面对识别出的要素的方向转动:人或特定人;特定动作。

[0089] 在一个实施例中,图 3 所示的视频采集单元 101 可以向着识别出的要素左右转动。在另一个实施例,图 3 所示的视频采集单元 101 可以向着识别出的要素上下左右转动。

[0090] 图 4 示出了根据本发明一个实施例的自动监测与自主反应方法 2 的流程图。自动

监测与自主反应方法 2 包括：

- [0091] 步骤 S1：采集环境的视频、音频信息；
- [0092] 步骤 S2：响应于所采集的视频、音频信息，对视频、音频信息的内容进行识别；
- [0093] 步骤 S3：基于识别结果做出反应。
- [0094] 可选地，所述反应可包括以下中的一个或多个：产生特定反应动作；发出特定反应声音；显示特定反应画面。
- [0095] 可选地，响应于所采集的视频、音频信息，对视频、音频信息的内容进行识别的内容可包括：从采集到的视频、音频信息中识别出特定动作和 / 或特定人的特定动作。
- [0096] 可选地，所述人或特定人可以是基于人脸识别、身高识别、声音识别中的一个或多个来识别的。
- [0097] 可选地，该方法还可以包括接收携带手机发出的无线信号，且所述人或特定人是基于该无线信号中标明的携带手机的身份来识别的。
- [0098] 可选地，所述特定动作可基于所采集的视频、音频信息和深度传感器感测的深度进行识别。
- [0099] 可选地，特定动作可以是通过事先为特定动作建立模型，并从采集到的视频、音频信息中搜索与建立的模型的匹配识别的。
- [0100] 可选地，所述模型可以是通过自学习的方式产生的。
- [0101] 可选地，所述模型可以是预先输入的标准化的模型。
- [0102] 附图中的流程图和框图显示了根据本发明的多个实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上，流程图或框图中的每个方框可以代表一个模块、程序段或代码的一部分，所述模块、程序段或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意，在有些作为替换的实现中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个连续的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的是，框图和 / 或流程图中的每个方框、以及框图和 / 或流程图中的方框的组合，可以用执行规定的功能或操作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。
- [0103] 对于本领域技术人员而言，显然本发明不限于上述示范性实施例的细节，而且在不背离本发明的精神或基本特征的情况下，能够以其他的具体形式实现本发明。因此，无论从哪一点来看，均应将实施例看作是示范性的，而且是非限制性的，本发明的范围由所附权利要求而不是上述说明限定，因此旨在将落在权利要求的等同要件的含义和范围内的所有变化囊括在本发明内。不应将权利要求中的任何附图标记视为限制所涉及的权利要求。

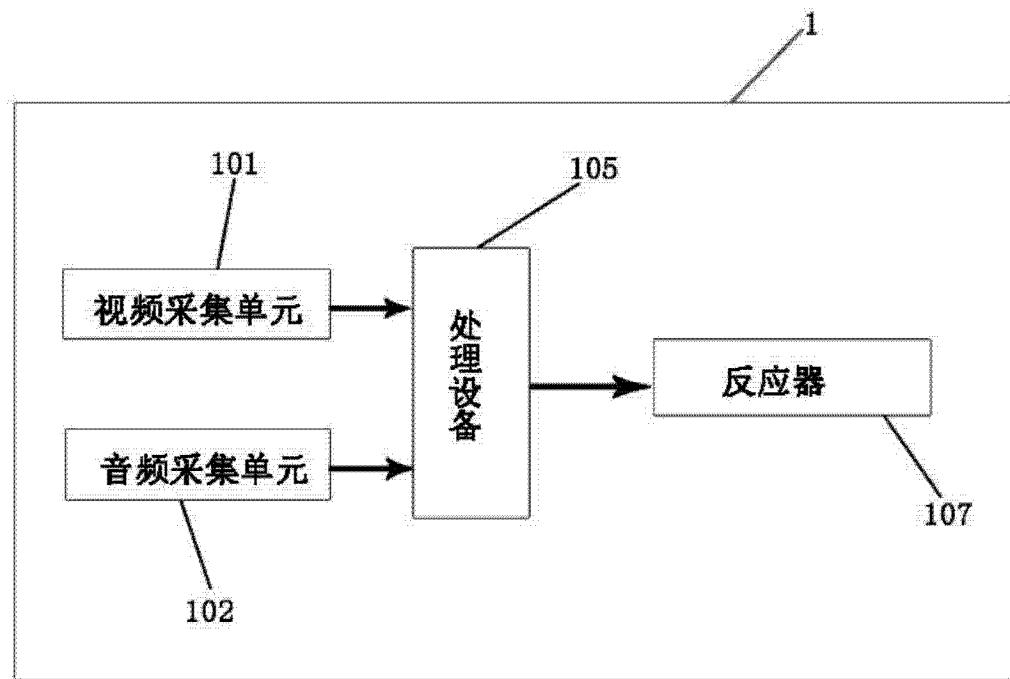


图 1

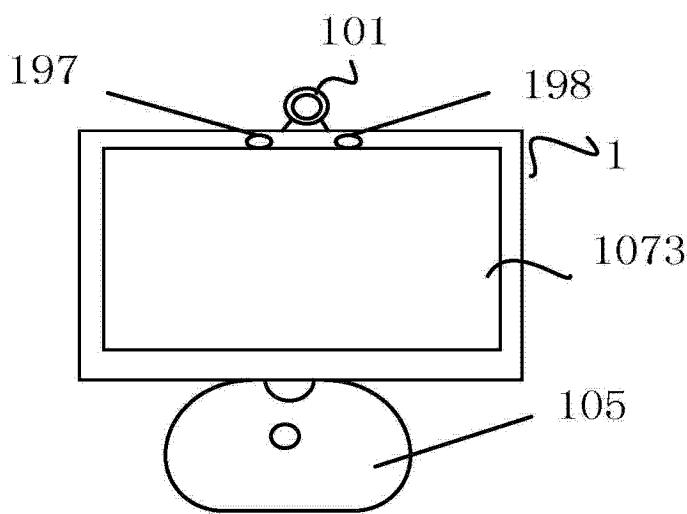


图 2

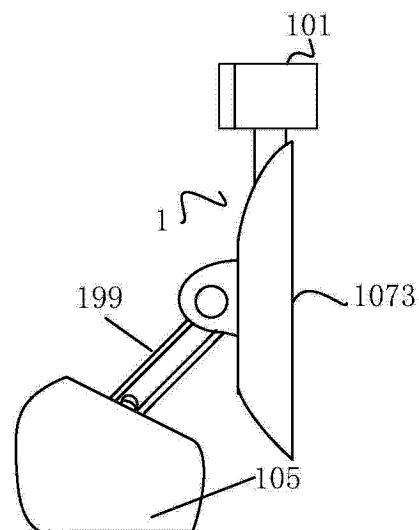


图 3

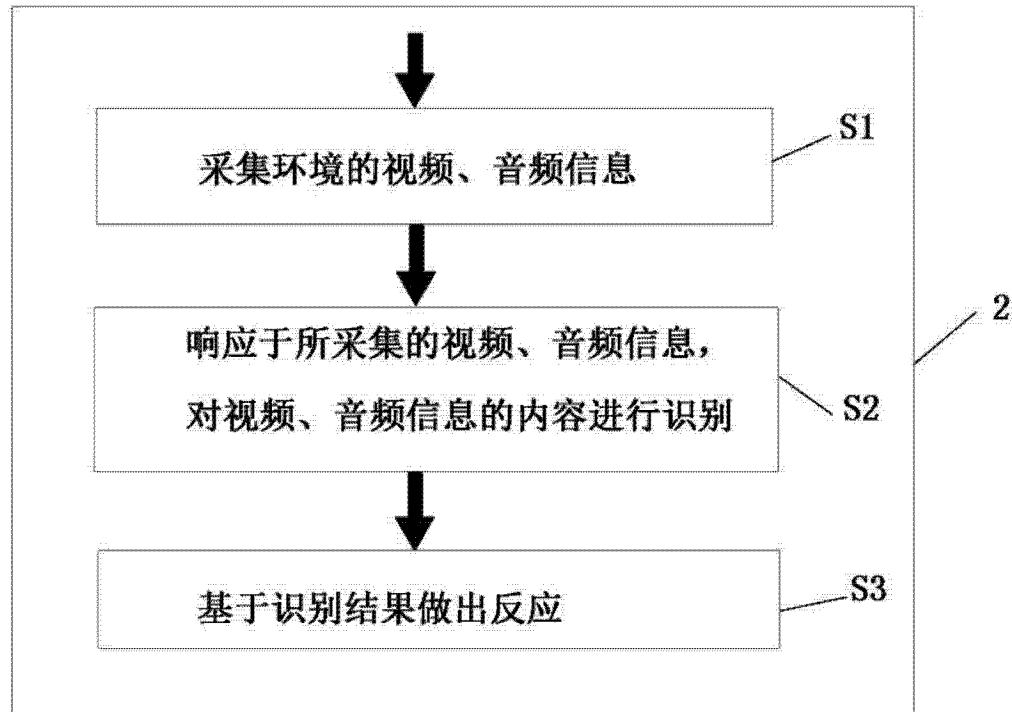


图 4

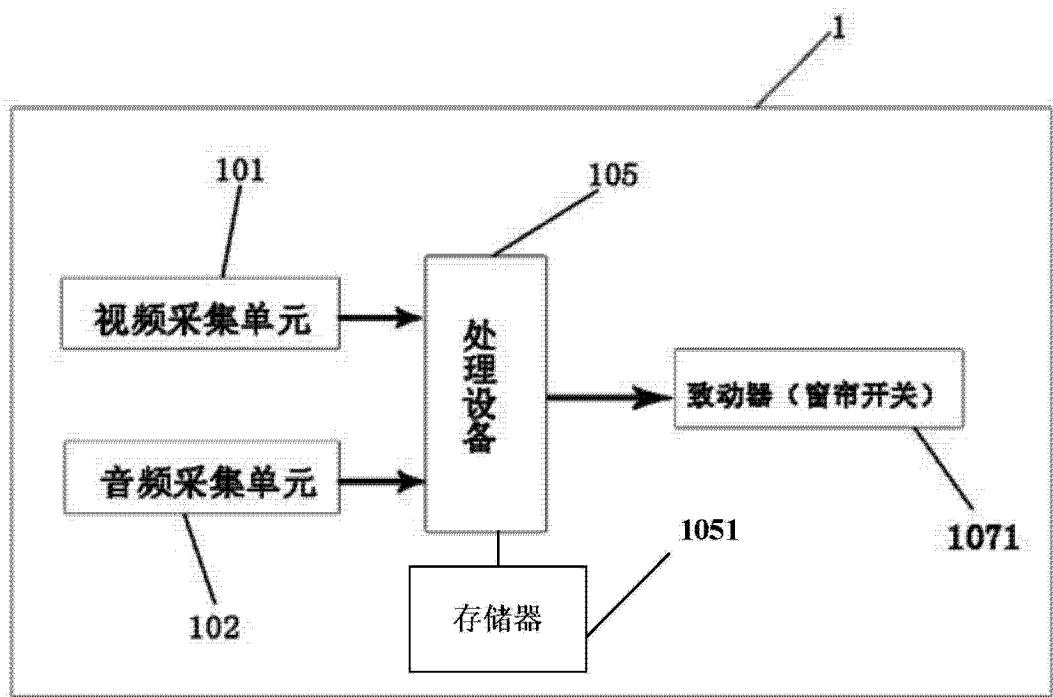


图 5

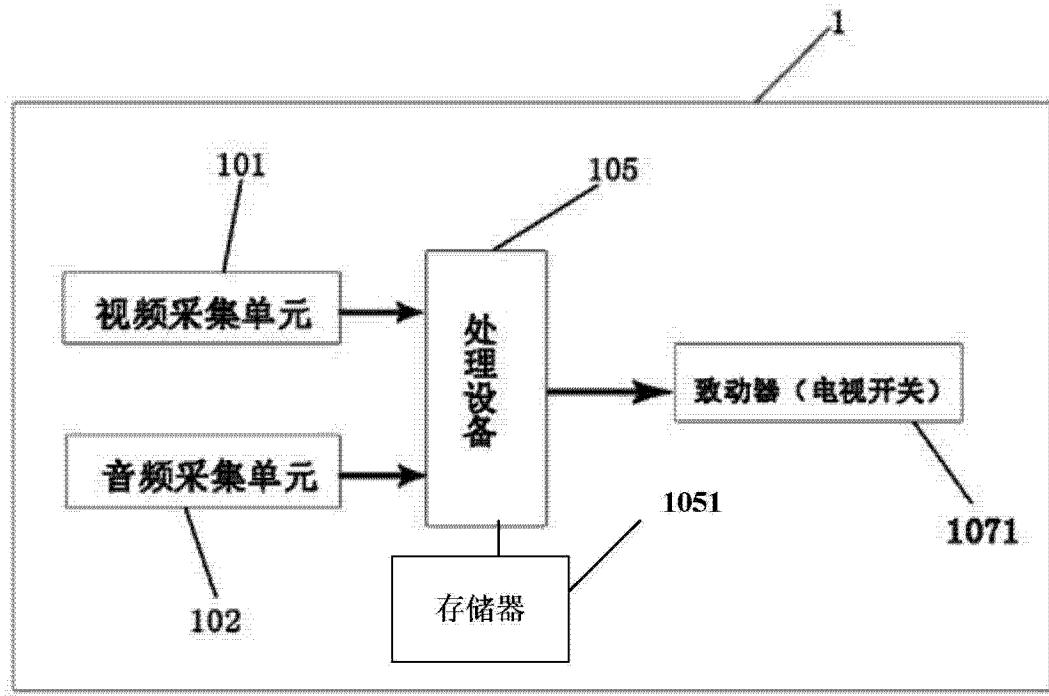


图 6

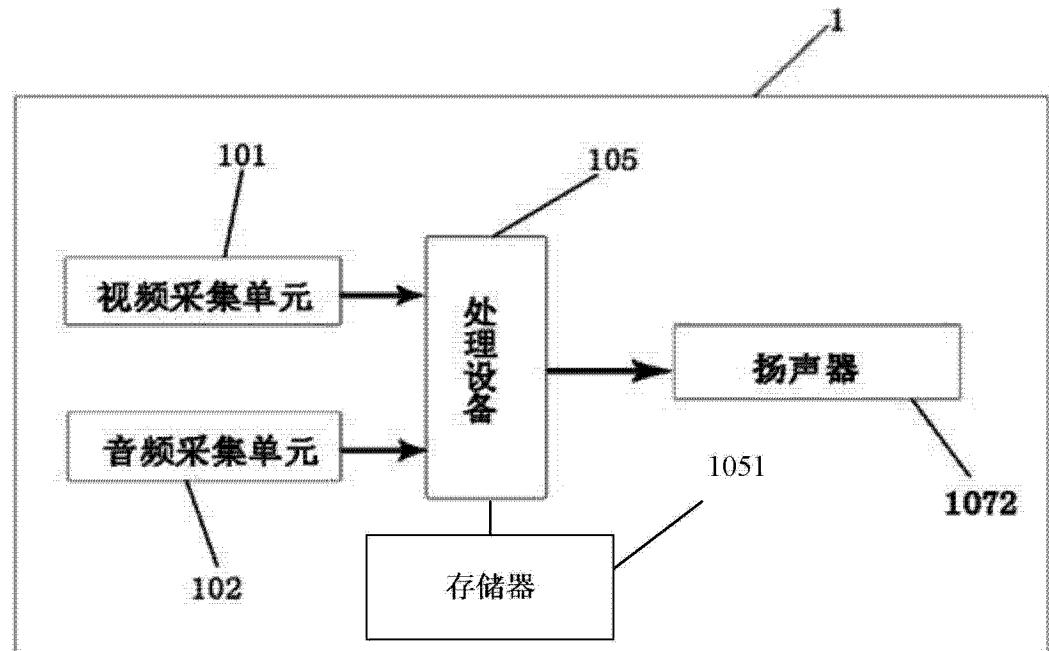


图 7

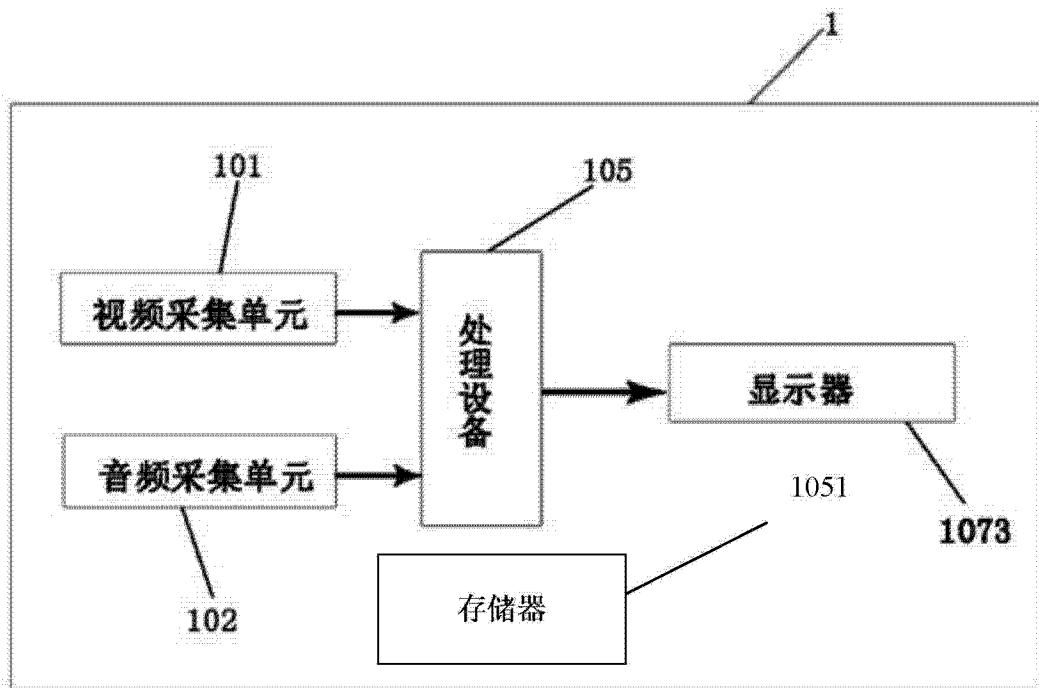


图 8