

(12) 发明专利申请

(10) 申请公布号 CN 101753992 A

(43) 申请公布日 2010.06.23

(21) 申请号 200810241284.8

(22) 申请日 2008.12.17

(71) 申请人 深圳市先进智能技术研究所

地址 518040 广东省深圳市福田区深南大道  
6007 号安徽大厦 1808 室

(72) 发明人 丁宁 覃强

(74) 专利代理机构 深圳市顺天达专利商标代理  
有限公司 44217

代理人 郭伟刚

(51) Int. Cl.

H04N 7/18(2006.01)

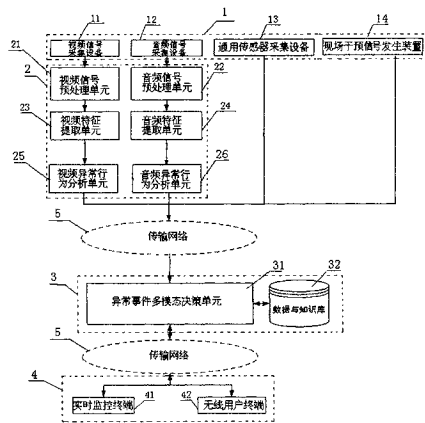
权利要求书 2 页 说明书 8 页 附图 4 页

(54) 发明名称

一种多模态智能监控系统和方法

(57) 摘要

本发明涉及一种多模态智能监控系统,包括信号输入系统(1)、信号预处理系统(2)、异常事件决策系统(3)、客户终端(4)和传输网络(5),所述信号输入系统(1)将采集的音/视频信号送往信号预处理系统(2)进行预处理后将信号送往异常事件决策系统(3)识别判断,从而将所产生的相关信息通过传输网络(5)发送给客户终端(4),实现多模态监控的目的。同时异常事件决策系统(3)根据判别的结论,实时发送事件干预指令,触发安装在现场的设备或其他模态通道的驱动机构进行协同联动与及时干预。采用本发明所述系统,能同时针对视频信息、音频信息和传感器多模态信息进行响应和处理。



1. 一种多模态智能监控系统,其特征在于:包括信号输入系统(1)、信号预处理系统(2)、异常事件决策系统(3)、客户终端(4)和传输网络(5);所述信号输入系统(1)将采集的音/视频信号送往信号预处理系统(2)进行预处理,然后将预处理后的信号送往异常事件决策系统(3),由异常事件决策系统(3)进行识别判断,从而将所产生的相关信息通过传输网络(5)发送给客户终端(4),实现多模态监控的目的。

2. 如权利要求1所述的多模态智能监控系统,其特征在于:所述信号输入系统(1)包括视频信号采集设备(11)、音频信号采集设备(12)、通用传感器采集设备(13)和现场干预信号发生装置(14)。

3. 如权利要求1所述的多模态智能监控系统,其特征在于:所述信号预处理系统(2)包括视频信号预处理单元(21)、音频信号预处理单元(22)、视频特征提取单元(23)、音频特征提取单元(24)、视频异常行为分析单元(25)和音频异常行为分析单元(26);所述视频信号预处理单元(21)接收从视频信号采集设备(11)传送过来的视频信号进行预处理,然后送往视频特征提取单元(23)进行视频信号特征提取处理;所述音频信号预处理单元(22)接收从音频信号采集设备(12)传送过来的音频信号进行预处理,然后送往音频特征提取单元(24)进行音频信号特征提取处理。

4. 如权利要求1所述的多模态智能监控系统,其特征在于:所述异常事件决策系统(3)包括异常事件多模态决策单元(31)和数据与知识库(32),经过信号预处理系统(2)处理过的音视频信号通过传输网络(5)传送给异常事件决策系统(3)进行识别判断。

5. 如权利要求1所述的多模态智能监控系统,其特征在于:所述客户终端(4)包括实时监控终端(41)、无线用户终端(42);所述传输网络(5)为TCP/IP网络;所述通用传感器采集设备(13)包括振动传感器、烟雾传感器、人体红外传感器、气体传感器、压力传感器、红外对射传感器和/或温湿度传感器,传感器类型、形式和数量可根据实际需要进行扩展。

6. 如权利要求3所述的多模态智能监控系统,其特征在于:所述异常事件决策系统(3)通过运行在內的算法对接收的信号进行识别判断;所述视频信号预处理单元(21)、音频信号预处理单元(22)、视频特征提取单元(23)和音频特征提取单元(24)采用基于PC或DSP平台的智能处理设备实现。

7. 如权利要求3所述多模态智能监控系统,其特征在于:所述视频信号采集设备(11)、视频信号预处理单元(21)、视频特征提取单元(23)和视频异常行为分析单元(25)组成视频智能分析子系统;所述音频信号采集设备(12)、音频信号预处理单元(22)、音频特征提取单元(24)和音频异常行为分析单元(26)组成音频智能分析子系统;所述通用传感器采集设备(13)自身组成通用传感器平台分析子系统;所述视频智能分析子系统、音频智能分析子系统和通用传感器平台分析子系统将各自信息通过各自通道送往异常事件决策系统(3)处理。

8. 一种多模态智能监控方法,基于一种多模态智能监控系统,该监控系统包括信号输入系统(1)、信号预处理系统(2)、异常事件决策系统(3)、客户终端(4)和传输网络(5),所述信号输入系统(1)将采集的音视频信号送往信号预处理系统(2)进行预处理,然后将预处理后的信号送往异常事件决策系统(3),由异常事件决策系统(3)进行识别判断,从而将所产生的相关信息通过传输网络(5)发送给客户终端(4),实现多模态监控的目的,其特征在于,所述方法包括步骤:

A、信号输入系统 (1) 采集信号；  
B、将采集到的信号送往信号预处理系统 (2) 进行预处理和特征提取；  
C、异常事件决策系统 (3) 对信号进行识别判断；  
D、将步骤C判断为报警信息的信息通过传输网络 (5) 传输给客户终端 (4)，实现对信息的智能监控。

9. 如权利要求 8 所述的多模态智能监控方法，其特征在于：

上述步骤 C 中，所述异常事件决策系统 (3) 对信息的识别判断还包括以下步骤：

a、将信号预处理系统 (2) 传送过来的信号通过融合算法进行融合处理，所述融合算法包括：小波变换、加权平均、产生式规则和卡尔曼滤波；

b、对经 a 步骤处理后的信息通过算法进行特征融合处理，将低维的特征描述向量融合形成更高维的联合特征向量参数，所述算法包括：参量模板法、聚类分析法、模糊集理论、自适应神经网络、物理模型、黑板模型和逻辑模板法；

c、对经步骤 b 处理过的信号通过算法进行识别判断处理，所述算法包括：经典推理、Bayes 推理、统计决策、D-S 证据理论、模糊理论、神经网络、专家系统、粗糙集理论、广义证据推理理论、直接推断法、可能理论和缺省推理。

10. 如权利要求 8 所述的多模态智能监控方法，其特征在于：

所述异常事件决策系统 (3) 包括数据与知识库 (32) 根据算法对信息自动学习的步骤。

## 一种多模态智能监控系统和方法

### 技术领域

[0001] 本发明涉及一种监控系统,更具体地说,涉及一种多模态智能监控系统和方法。

### 背景技术

[0002] 随着社会对公共安全问题的重视,实时监控设备得到越来越广泛的应用。在目前大量应用的人工方式监控的监控系统中,如何对大量监控信息进行快速、有效的处理成为一个突出问题。应用计算机技术对监控信息进行处理、识别异常事件进而报警是解决这一问题的有效方法之一。

[0003] 传统的视频监控是一种事件历史记录系统,不是监控预警工具。传统的视频监控系统将多台摄像机的视频信息传回到监控中心存储起来,同时在视频监视墙上实时显示给安全人员。基本上是在事件发生之后再查询录像记录,不能在事件发生时,实时通知安全人员,致使造成的损失无法挽回。而为了满足人们对安防质量越来越高的要求,就需要视频监控系统由“事后查询”变成“预警”,由“被动”变“主动”。

[0004] 智能视频监控就是在不需要人为干预的情况下,通过对摄像机拍摄的图像序列进行自动分析,实现对动态场景中目标的定位、识别和跟踪,并在此基础上分析和判断目标的行为,从而做到既能完成日常管理又能在异常情况发生的时候及时做出反应。

[0005] 人类具有将人体的各种感觉器官(眼、耳、鼻、四肢等)所得到的信息(包括图像、声音、气味、触觉等)结合先念知识进行综合处理的能力,就可以对周围环境和正在发生的事情作出估计和判断。这一处理过程是复杂的,也是自适应的,它将各种信息转化为对环境的有价值的解释。本发明涉及的多模态智能监控系统的核心就是充分利用多个信息源提供的信息,期望像人脑综合处理各种信息的过程一样,通过对这些信息的合理使用,把冗余或者互补的信息依据某种规则来进行组合,以获得对被监控对象及其事件的一致性解释或描述。

[0006] 上述现有技术的监控系统和方法的不足在于,传统的视频监控系统只是一个事件历史记录系统,不能提供预警工具,无法实时监控事件的发生,造成了极大的安全漏洞。

### 发明内容

[0007] 本发明要解决的技术问题在于,针对现有技术的上述不能提供实时预警的缺陷,提供一种多模态智能监控系统和方法。

[0008] 本发明解决所述技术问题可以通过采用以下技术方案来实现:提供一种多模态智能监控系统,尤其是,包括信号输入系统、信号预处理系统、异常事件决策系统、客户终端和传输网络;

[0009] 所述信号输入系统将采集的音视频信号送往信号预处理系统进行预处理,然后将预处理后的信号送往异常事件决策系统,由异常事件决策系统进行识别判断,从而将所产生的相关信息通过传输网络发送给客户终端,实现多模态监控的目的。

[0010] 所述信号输入系统包括视频信号采集设备、音频信号采集设备、通用传感器采集

设备和现场干预信号发生装置。

[0011] 所述信号预处理系统包括视频信号预处理单元、音频信号预处理单元、视频特征提取单元、音频特征提取单元、视频异常行为分析单元和音频异常行为分析单元；所述视频信号预处理单元接收从视频信号采集设备所传送过来的视频信号进行预处理，然后送往视频特征提取单元进行视频信号特征提取处理；所述音频信号预处理单元接收从音频信号采集设备传送过来的音频信号进行预处理，然后送往音频特征提取单元进行音频信号特征提取处理。

[0012] 所述异常事件决策系统包括异常事件多模态决策单元和数据与知识库，经过信号预处理系统处理过的音视频信号通过传输网络传送给异常事件决策系统进行识别判断。

[0013] 所述客户终端包括实时监控终端、无线用户终端；所述传输网络为 TCP/IP 或其他网络；所述通用传感器采集设备包括振动传感器、烟雾传感器、人体红外传感器、气体传感器、压力传感器、红外对射传感器、温湿度传感器，传感器类型、形式和数量可根据实际需要扩展。

[0014] 所述异常事件决策系统通过运行在的算法对接收的信号进行识别判断；

[0015] 所述视频信号预处理单元、音频信号预处理单元、视频特征提取单元和音频特征提取单元为基于 PC 平台或 DSP 平台的智能处理设备。

[0016] 所述视频信号采集设备、视频信号预处理单元、视频特征提取单元和视频异常行为分析单元组成视频智能分析子系统；所述音频信号采集设备、音频信号预处理单元、音频特征提取单元和音频异常行为分析单元组成音频智能分析子系统；所述通用传感器采集设备自身组成通用传感器平台分析子系统；所述视频智能分析子系统、音频智能分析子系统和通用传感器平台分析子系统将各自信息通过各自通道送往异常事件决策系统处理；

[0017] 所述视频智能分析子系统、音频智能分析子系统和通用传感器平台分析子系统将各自信息通过各自通道送往异常事件决策系统处理。

[0018] 本发明解决所述的技术问题，还可以进一步采用以下技术方案来实现：提出一种多模态智能监控方法，基于一种多模态智能监控系统，该监控系统包括信号输入系统、信号预处理系统、异常事件决策系统、客户终端和传输网络，所述信号输入系统将采集的音视频信号送往信号预处理系统进行预处理，然后将预处理后的信号送往异常事件决策系统，由异常事件决策系统进行识别判断，从而将所产生的相关信息通过传输网络发送给客户终端，实现多模态监控的目的，尤其是，所述方法包括步骤：

[0019] A、信号输入系统采集信号；

[0020] B、将采集到的信号送往信号预处理系统进行预处理和特征提取；

[0021] C、将步骤 B 处理过的信号，送往异常事件决策系统进行识别判断；

[0022] D、将步骤 C 判断为报警信息的信息通过传输网络传输给客户终端，实现对信息的智能监控。

[0023] 所述异常事件决策系统对信息的识别判断包括以下步骤：

[0024] a、将信号预处理系统传送过来的低层信号通过融合算法进行融合处理，所述融合算法包括：小波变换、加权平均、产生式规则和卡尔曼滤波；

[0025] b、对经 a 步骤处理后的信息通过算法进行特征融合处理，将个低维的特征描述向量融合形成更高维的联合特征向量参数，所述算法包括：参量模板法、聚类分析法、模糊集

理论、可能性理论、自适应神经网络、物理模型、黑板模型和逻辑模板法；

[0026] c、对经步骤 b 处理过的信号通过算法进行识别判断处理，所述算法包括：决策层融合算法有：经典推理、Bayes 推理、统计决策、D-S 证据理论、模糊理论、神经网络、专家系统、粗糙集理论、广义证据推理理论、直接推断法和缺省推理。

[0027] 所述异常事件决策系统包括数据与知识库根据算法对信息自动学习的步骤。

[0028] 同现有技术相比较，本发明多模态智能监控系统和方法的有益效果在于：首先，能同时针对视频信息、音频信息和传感器多模态信息进行处理，利用互补的信息以降低错误率，利用多种来源的信息以增加稳定性，从视频、音频和通用传感器的多方面来获得信息，丰富监控系统的监控信息；以及，将“实时自动学习”引入本发明所述监控系统中，使监控系统能够不断积累和丰富自身的数据与知识库，通过自学习，增强和完善多模态信息融合与决策能力。

### 附图说明

[0029] 下面将结合附图及实施例对本发明作进一步说明，附图中：

[0030] 图 1 是本发明多模态智能监控系统实施例的原理示意图；

[0031] 图 2 是本发明多模态智能监控系统实施例的应用示意图；

[0032] 图 3 是本发明多模态智能监控系统实施例中异常事件决策系统层次逻辑示意图；

[0033] 图 4 是本发明多模态智能监控系统实施例中混合式系统结构图；

[0034] 图 5 是本发明多模态智能监控系统实施例中功能逻辑结构图。

### 具体实施方式

[0035] 以下结合各附图所示之最佳实施例作进一步详述：

[0036] 本发明构建了一套针对异常行为识别的多模态智能监控系统，该系统采用了视频监控、音频监控及通用传感器通信平台监控的复合监控方式，对预先设定的感兴趣的监控对象行为进行监控，根据基于规则的知识库，应用“实时自学习智能系统”与“多模态信息融合与决策”技术，自动识别监控区域内对象的行为，并做出实时记录及报警。

[0037] 按照本发明提供的一种多模态智能监控系统，如图 1 所示，包括信号输入系统 1、信号预处理系统 2、异常事件决策系统 3、客户终端 4 和传输网络 5；所述信号输入系统 1 将采集的音视频信号送往信号预处理系统 2 进行预处理，然后将预处理后的信号送往异常事件决策系统 3，由异常事件决策系统 3 进行识别判断，根据判别的结论，实时发送事件干预指令，触发安装在现场的设备或其他模态通道的驱动机构进行协同联动与及时干预，从而将所产生的相关信息通过传输网络 5 发送给客户终端 4，实现多模态监控的目的。

[0038] 如图 1 所示，所述信号输入系统 1 包括视频信号采集设备 11、音频信号采集设备 12 和通用传感器采集设备 13，分别用于采集或传感视频信号、音频信号和其他信号，此处的其他信号可以通过温感、压敏、加速度、位移等物理量或运动量传感器来采集，也可接收红外、烟雾、可燃气体等传感器监测到的信号。

[0039] 如图 1 所示，所述信号预处理系统 2 包括视频信号预处理单元 21、音频信号预处理单元 22、视频特征提取单元 23 和音频特征提取单元 24；所述视频信号预处理单元 21 接收从视频信号采集设备 11 传送过来的视频信号进行预处理，然后送往视频特征提取单元 23

进行视频信号特征提取处理;所述音频信号预处理单元 22 接收从音频信号采集设备 12 传送过来的音频信号进行预处理,然后送往音频特征提取单元 24 进行音频信号特征提取处理。

[0040] 所述异常事件决策系统 3 包括异常事件多模态决策单元 31 和数据与知识库 32,经过信号预处理系统 2 处理过的音视频信号通过传输网络 5 传送给异常事件决策系统 3 进行识别判断。

[0041] 所述客户终端 4 包括实时监控终端 41、无线用户终端 42;所述传输网络 5 为 TCP/IP 或其他网络;所述通用传感器采集设备 13 包括振动传感器、烟雾传感器、人体红外传感器、气体传感器、压力传感器、红外对射传感器、温湿度传感器。

[0042] 如图 1 所示,视频信号采集设备 11 用于采集视频信号、音频信号采集设备 12 用于采集音频信号,而通用传感器采集设备 13 用于采集其它各类传感器信号,异常事件决策系统 3 对所采集的视频信号、音频信号和其它传感器信号进行判别决策,得出对异常事件的决策结果,再将决策结果经传输网络 5 发送至客户终端 4 的实时监控终端 41 或无线用户终端 42,客户终端 4 也可以是报警器。

[0043] 所述异常事件决策系统 3 通过运行在內的算法对接收的信号进行识别判断:

[0044] 所述视频信号预处理单元 21、音频信号预处理单元 22、视频特征提取单元 23 和音频特征提取单元 24 为计算机或 DSP 平台处理设备。

[0045] 参照图 1,所述视频信号采集设备 11、视频信号预处理单元 21 和视频特征提取单元 23 组成视频智能分析子系统;所述音频信号采集设备 12、音频信号预处理单元 22 和音频特征提取单元 24 组成音频智能分析子系统;所述通用传感器采集设备 13 自身组成通用传感器平台分析子系统;所述视频智能分析子系统、音频智能分析子系统和通用传感器平台分析子系统将各自信息通过各自通道送往异常事件决策系统 3 处理。

[0046] 所述视频智能分析子系统通过对视频信号的,实现以下功能:目标提取与跟踪;徘徊行为;异常运动行为。

[0047] 所述音频智能分析子系统特征:1) 异常声音识别:异常声音识别模块通过对大量含有异常声音如,尖叫声、呼救声、玻璃破碎声、哭声等的样本进行学习,与正常声音如,笑声、鼓掌声、城市噪声等进行模式分类,从而具备了对异常声音的识别能力;2) 异常声音定位:通过麦克风阵列对异常声音的来源进行识别,实现对声源的初步定位。该功能用于与 PTZ 相机协同联动,引导 PTZ 相机转向声源位置,再进一步利用视频分析功能进行判别。

[0048] 所述传感器与通用传感器平台特征:除视频、音频信号以外,通用传感器平台是多模态智能监控系统的另一主要输入信号,它作为视音频信号的有效扩展,可以充分发挥传感器种类多、相应快速和性能可靠等特点,特别适用于一些特殊监控场所。例如:在 ATM 自助银行监控时可应用振动传感器判别对设备的在通用传感器平台上可扩展振动、烟雾、人体红外等多种传感器。

[0049] 可扩展的传感器有振动传感器、烟雾传感器、人体红外传感器、气体传感器、压力传感器、红外对射传感器、温湿度传感器等,传感器类型、形式和数量可根据实际需要进行扩展。

[0050] 应用视频信号采集设备 11 的机器视觉与机器学习技术,所述视频智能分析子系统对视频信号进行实时分析,根据运动目标的运动特征,识别出视野内各类异常行为。

[0051] 通过对数万条各类音频样本特征学习,音频智能分析子系统建立了针对几类典型异常声音的识别模型,该子系统对音频信号进行实时分类,判别出异常声音。

[0052] 异常事件决策系统 3 是整个系统的核心,决策系统 3 接收音 / 视频特征信息及传感器信息,根据异常事件判别规则库在知识层面上对事件性质进行判别,该系统第一时间将识别结果提示给后台监控人员,同时触发安装在现场的干预装置,对事件进行主动干预,从而真正实现对事件的监与控。

[0053] 下面就结合具体实施例,作进一步说明:

[0054] 图 2 是本实施例的硬件构架图,该实例系统采用了串并行共存的系统构架,信号输入系统 1 中的视频信号采集设备 11 为摄像机;音频信号采集设备 12 为拾音器,用于采集音频信号;通用传感器采集设备 13 为各类传感器;信号预处理系统 2 由计算机组成或由 DSP 平台的智能处理设备来实现。

[0055] 如图 2 所示,摄像机、拾音器、传感器等设备是本系统的信号输入设备,用来采集视频、音频及其他感兴趣的信,信号预处理系统 2 可用计算机架构或 DSP 平台的智能处理设备实现,用于对系统输入的原始数据进行分析 and 特征提取,其中计算机硬件构架主要用于多路视频的中低密度情况下精确目标跟踪结果,DSP 芯片构架处理设备,主要用于人群状态、音频分析、通用传感器等基础数据提取;图 2 中,异常事件多模态决策单元 31 依据数据与知识库 32 中已有的规则,对经信号预处理系统 2 的视频、音频和传感器特征数据进行综合判别,识别出异常事件,进而输出客户终端 4,同时进行存储、报警等操作。摄像机、拾音器及其他传感器设备、信号预处理系统 2、异常事件多模态决策单元 31 及客户终端 4 之间通过 TCP/IP 网络或者其他通信方式组成网络进行通信。

[0056] 串行构架是将特征提取与模式识别拆分为两个独立部分,分别有信号预处理系统 2 与异常事件多模态决策单元 31 分别完成;并行构架是每个摄像机配置一个处理单元进行同步独立处理,提取出来的特征和抽象出来的信息仅需要较少的数据量来表示,降低了信息传输网络的负担,对于提取出的基础数据对于分析服务器的要求相对较低。复杂的分析任务进行了功能性分解,缓解了处理装置的负担,实现了基础数据的有效共享,并降低了整个系统的成本。

[0057] 通过如图 2 所示的系统结构,系统在进行多模态智能监控时,采用的方法如下:

[0058] 本发明还提供了一种多模态智能监控方法,如图 1 所示,该方法基于一种多模态智能监控系统,该监控系统包括信号输入系统 1、信号预处理系统 2、异常事件决策系统 3、客户终端 4 和传输网络 5,所述信号输入系统 1 将采集的音视频信号送往信号预处理系统 2 进行预处理,然后将预处理后的信号送往异常事件决策系统 3,由异常事件决策系统 3 进行识别判断,从而将所产生的相关信息通过传输网络 5 发送给客户终端 4,实现多模态监控的目的,所述方法包括步骤:

[0059] A、信号输入系统 1 采集信号;

[0060] B、将采集到的信号送往信号预处理系统 2 进行预处理和特征提取;

[0061] C、异常事件决策系统 3 对信号进行识别判断;

[0062] D、将步骤 C 判断为报警信息的信息通过传输网络 5 传输给客户终端 4,实现对信息的智能监控。

[0063] 上述步骤 C 中,所述异常事件决策系统 3 对信息的识别判断包括以下步骤:



[0064] a、将信号预处理系统 2 传送过来的低层信号通过融合算法进行融合处理,所述融合算法包括:小波变换、加权平均、产生式规则和卡尔曼滤波;

[0065] b、对经 a 步骤处理后的信息通过算法进行特征融合处理,将多个低维的特征描述向量融合形成更高维的联合特征向量参数,所述算法包括:参量模板法、聚类分析法、模糊集理论、可能性理论、自适应神经网络、物理模型、黑板模型和逻辑模板法;

[0066] c、对经步骤 b 处理过的信号通过算法进行识别判断处理,所述算法包括:经典推理、Bayes 推理、统计决策、D-S 证据理论、模糊理论、神经网络、专家系统、粗糙集理论、广义证据推理理论、直接推断法、可能理论和缺省推理。

[0067] 所述异常事件决策系统 3 能够根据算法对数据与知识库 32 中没有存储的信息进行学习。

[0068] 异常事件决策系统 3 主要完成对信息的识别判断,是本发明所述监控系统的核心,在本实施例中,该异常事件决策系统 3 主要功能及结构主要为:

[0069] (1) 所述决策系统 3 逻辑上呈层级结构,该结构参照图 3 所示,各层结构及功能为:

[0070] a. 元数据层:用于采集、传输各多模态通道;

[0071] b. 参数与特征层:用于提取和传输各模态通道相关参数以及特征提取;

[0072] c. 模型层:用于训练各通道的识别模型;存储与应用各通道的分析与识别功能;

[0073] d. 决策层:属于整个系统构架的最顶层,用于汇总各通道模型得到的识别结论,进行事件性质的综合判断。

[0074] (2) 结合上述 (1):异常事件决策系统 3 在接收到信息后,需要对各类型的信息进行融合处理,其过程可以概括为:

[0075] a. 数据级:最低层的数据级,直接包括了对原始特征最充分、最有效的描述,由于数据的大量性、特征的复杂性、以及数据之间的强相关性等,使得直接利用原始数据的融合几乎是不可能。因此数据级的融合还只能是层级框架的一个理论组成部分,数据层融合算法有:小波变换、加权平均、产生式规则、卡尔曼滤波等。

[0076] b. 参数、特征与模型级:参数包括各模态通道较为低层的信息,主要表征元数据中离散的元素之间的普遍关系,如图像参数。另一方面,也表征了各模态通道本身的参数,这些参数对于各模态的特征和处理方法都有影响。

[0077] 输入数据经过前端处理以后,对于各模态通道分别得到其特征描述向量,然后经过特征融合的处理,将多个低维的特征描述向量融合(合并)形成更高维的联合特征向量参数。在系统建模时,针对该高维向量进行模型的建立,匹配时,利用该高维向量进行匹配。参数、特征层的融合算法有:参量模板法、聚类分析法、模糊集理论、可能性理论、自适应神经网络、物理模型、黑板模型、逻辑模板法等。

[0078] c. 决策级融合:决策融合的系统,输入信号经过处理得到特征参数,然后分别进行单模态的建模与识别,将识别的中间结果参数,通过决策融合模块进行融合,然后通过多模态决策算法得到最终的鉴别结果。决策层融合算法有:经典推理、Bayes 推理、统计决策、D-S 证据理论、模糊理论、神经网络、专家系统、粗糙集理论、广义证据推理理论、直接推断法、可能理论、缺省推理等。

[0079] 信息融合三个层次的配合根据模态的特征进行调整,如各模态通道得到的数据相

匹配,则可进行数据层融合,反之则在特征层或者决策层进行融合。

[0080] (3) 采用混合式系统层次结构,如图 4 所示,每种类型的信号对应一种模态,视频信号、音频信号和其它传感器所采集的信号分别通过各自的通道进行传送,混合式层次结构包括以下三种信号传输结构:

[0081] a. 分散式:如图 4 中模态 2 通道所示,针对复杂的模态分析,如视频、音频,每种模态通道有各自的;

[0082] b. 集中式:如图 4 中模态 1 通道所示,用于各类传感器信号的汇总;

[0083] c. 反馈式:如图 4 模态 3 通道所示,决策用于各类模态信号之间的协同联动。

[0084] 本发明中,如图 5 所示,信号输入系统 1 根据获取的视频信号经处理可分析得出的参数和特征包括:

[0085] 物理参数:摄像机内部参数,摄像机相对于环境的外部参数(通过摄像机标定获得);

[0086] 图像参数:包括颜色、纹理和亮度;

[0087] 几何特征:包括边缘、特征点、轮廓与形状;

[0088] 运动特征:运动矢量(包括特征点的实时运动速度、方向)。

[0089] 本发明中,如图 5 所示,信号输入系统 1 根据获取的音频信号经处理可分析得出的参数和特征包括:

[0090] 音量参数:包括幅度、频率和短时能量;

[0091] 波动参数:包括过零率和基频;

[0092] 发音特征:包括 MFCC(梅尔倒谱系数)和 LPC(线性预测倒谱系数);

[0093] 韵律特征:包括语调、重音和停顿。

[0094] 本发明中,所述异常事件决策系统 3 中的数据与知识库 32 具备对信息自动学习的功能,数据与知识库 32 的学习模型在学习视频信号时采用神经网络、支持向量机等统计机器学习模型;学习音频信号时采用隐马尔科夫模型(HMM)学习模型。

[0095] 异常事件决策系统 3 在进行决策判断时,通过以下方式进行:

[0096] 1、基于知识的模型

[0097] 模仿人在目标识别中所采用的经验知识和分析推理过程,将目标的信息(包括分类、结构等)以知识的形式(句法规则、框架、逻辑关系)存放在预先建立好的知识库中,采用启发式的方法来进行目标的识别。其中的逻辑模板法和模糊集理论体现了对人的认识经验进行的数据描述,可通过规则制定的方式对已知的行为进行判别。

[0098] 2、基于特征的推理技术:

[0099] 对特征进行关联后,取联合特征完成识别各类目标的任务,采用基于统计计算的方法和应用信息论技术的方法对事件性质进行推理。

[0100] 3、基于物理模型的方法:

[0101] 根据物理模型直接计算目标的特征,按照目标属性的分类建立不同类别或具有不同类参数的目标属性模型,并通过目标测量信息与物理模型给出的预测信息进行匹配来识别目标。

[0102] 4、协同联动:

[0103] 协同是指多模态通道之间的协同配合、信息共享,共同完成监控任务的过程。与多

模态信息融合不同,这种功能更强调各模态之间的主次、辅助与实施等角色的分配。

[0104] 联动是指针对决策结论的干预实施,目的是为了阻止事件的进一步发展。

[0105] 数据与知识库用来存储专家知识,可由事实性知识和推理性知识组成,包含描述关系、现象和方法的规则,以及在系统专家范围内解决问题的知识。知识库对预先存储的知识按照预先指定的规则进行推理计算。

[0106] 数据库分为实时数据库和非实时数据库,实时数据库用于向系统提供当前观测结果及融合所需要的各种其它数据,并存储中间结果。非实时数据库存储历史数据及有关环境和目标的辅助信息。

[0107] 多模态智能监控系统对各个模态通道得到的结论,在知识层面上,根据不同场景和事件制定的规则库来进行判断事件的性质。例如,是个人行为还是群体行为,是轻微违规还是严重犯罪等。目前实现以下行为的判别:

[0108] (1) 入侵行为识别;

[0109] (2) 人群状态监控;

[0110] (3) 暴力冲突行为;

[0111] (4) 重要物品管理;

[0112] (5) 破坏设备行为;

[0113] (6) 恶意纵火行为。

[0114] 对于事件的正确理解是监控的一方面,监控系统更重要的任务是在于对于异常事件的主动干预,防止事件进一步恶化,以避免监控人员得到了及时的提示而无法对事件作及时处理,以至于失去对事态的控制。多模态智能监控除了要实现对于事件的全方位监视,更通过多种联动干预方式实现对于事件的控制。

[0115] 上述实现过程为本发明的优先实现过程,本领域的技术人员在本发明的基础上进行的通常变化和替换包含在本发明的保护范围之内。

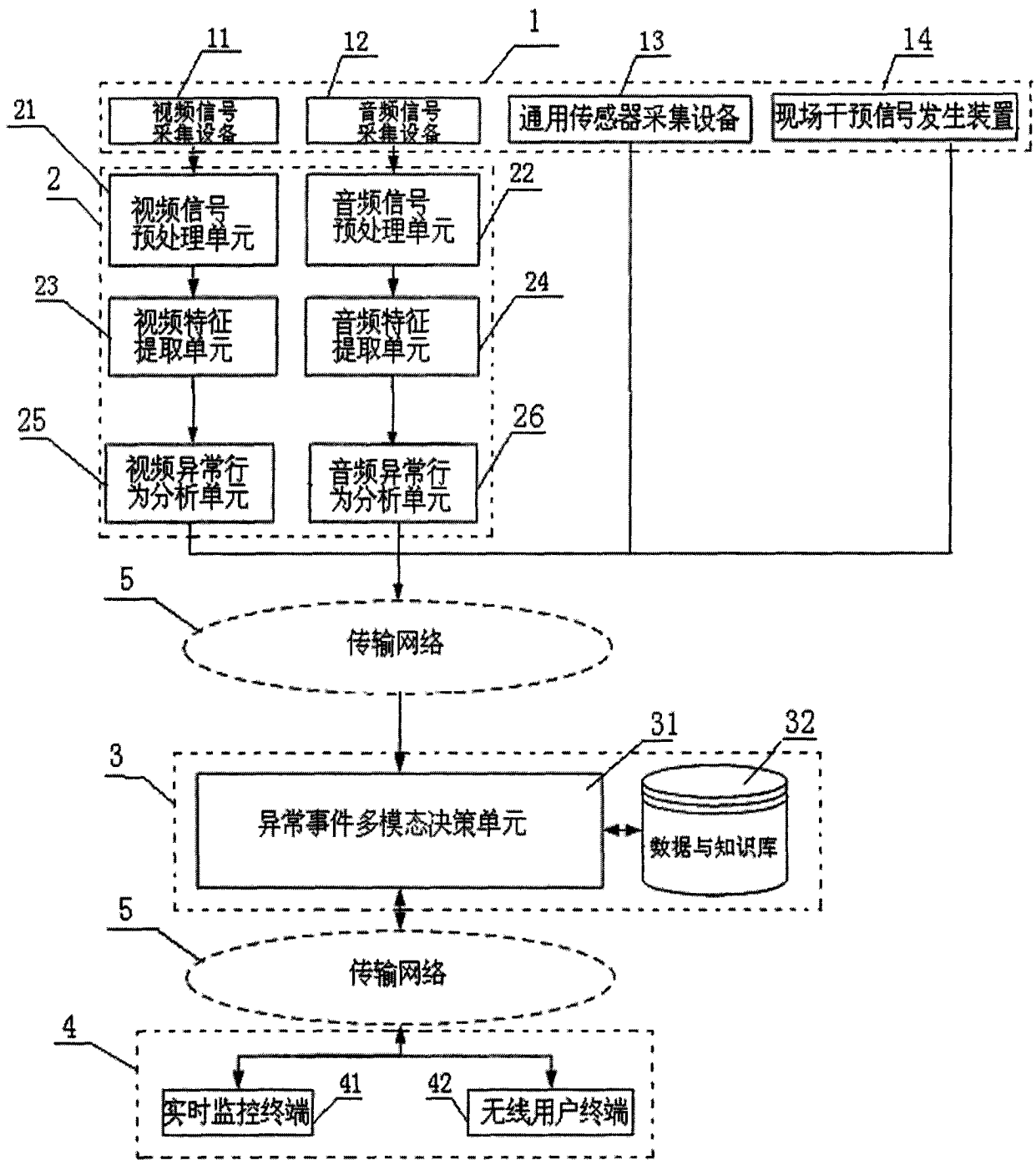


图 1

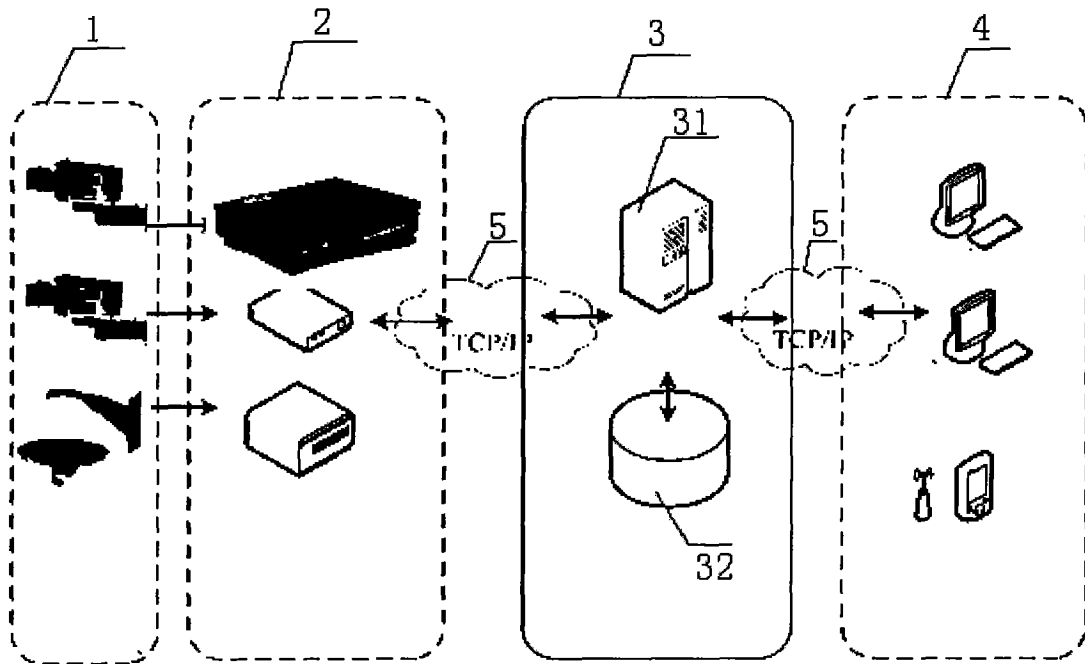


图 2

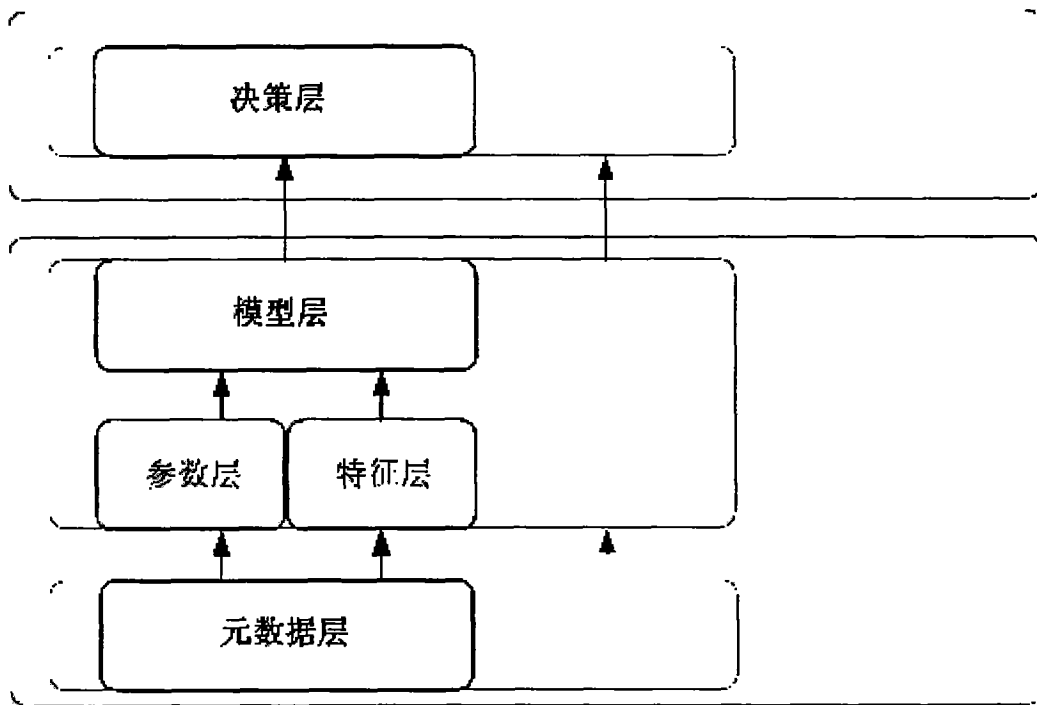


图 3

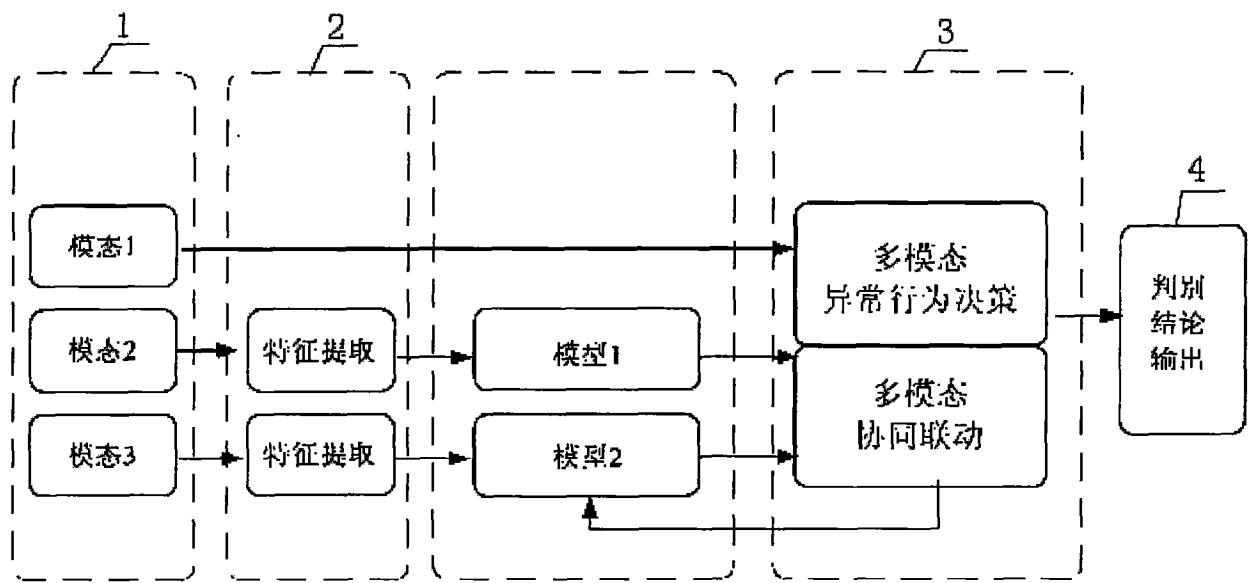


图 4

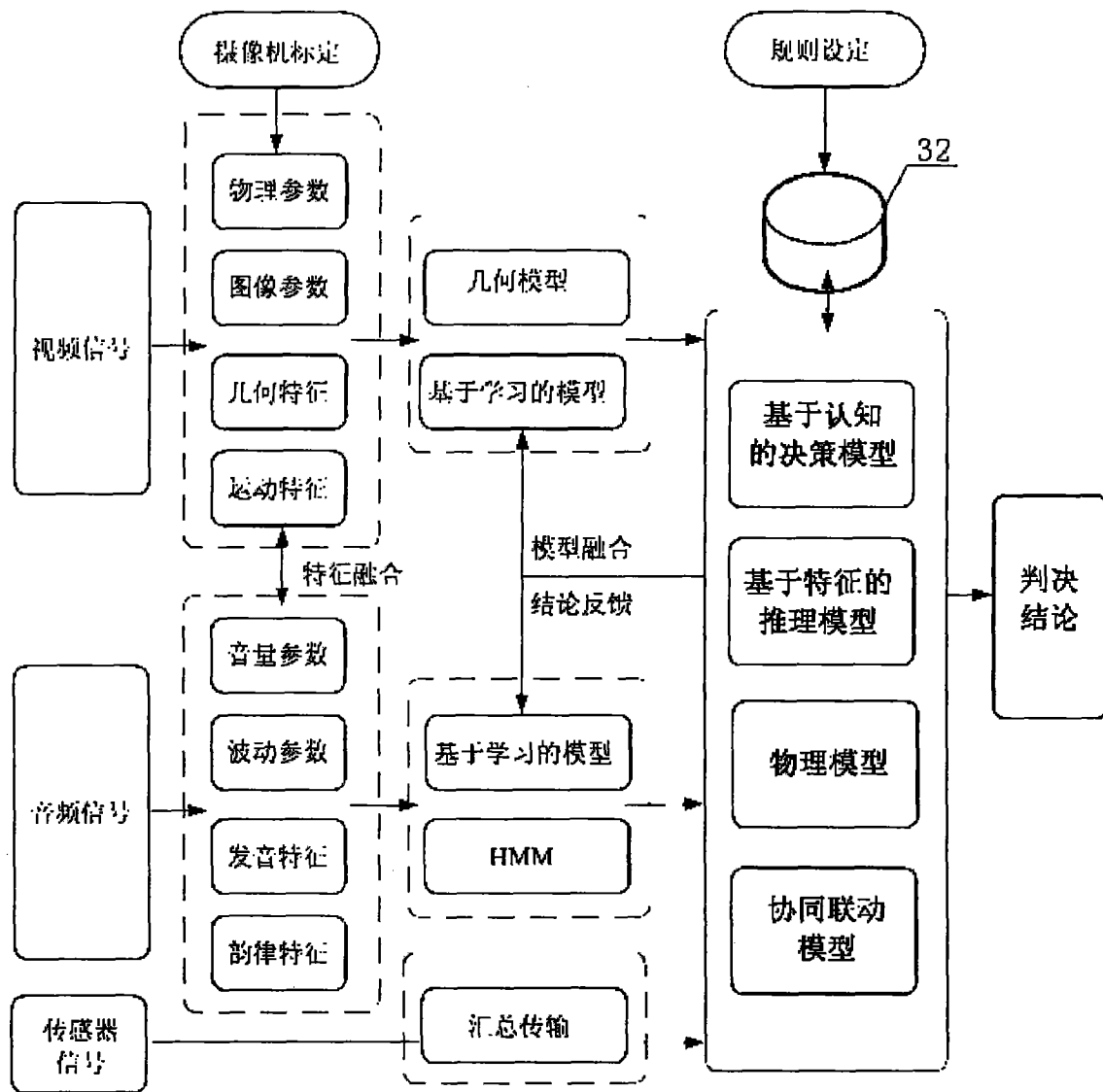


图 5