



(12) 发明专利申请

(10) 申请公布号 CN 111920420 A

(43) 申请公布日 2020. 11. 13

(21) 申请号 202010740444.4

(22) 申请日 2020.07.28

(71) 申请人 复旦大学

地址 200433 上海市杨浦区邯郸路220号

(72) 发明人 张立华 杨鼎康 邝昊鹏 林野

(74) 专利代理机构 上海科盛知识产权代理有限公司 31225

代理人 赵继明

(51) Int. Cl.

A61B 5/11 (2006.01)

A61B 5/16 (2006.01)

G06K 9/62 (2006.01)

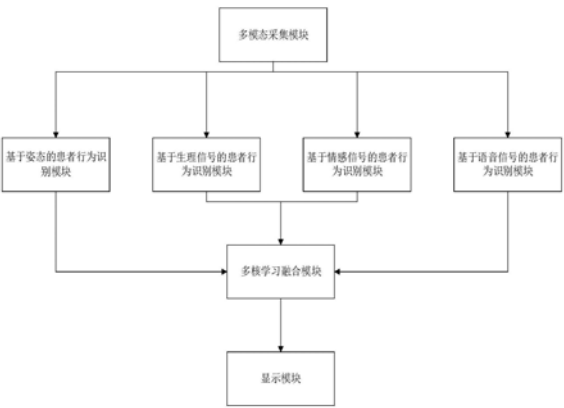
权利要求书2页 说明书9页 附图6页

(54) 发明名称

一种基于统计学习的患者行为多模态分析与预测系统

(57) 摘要

本发明涉及一种基于统计学习的患者行为多模态分析与预测系统,包括多模态采集模块、基于姿态的患者行为识别模块、基于生理信号的患者行为识别模块、基于情感信号的患者行为识别模块、基于语音信号的患者行为识别模块和基于多核学习的融合模块;基于多核学习的融合模块包括多核分类器,其训练过程具体为,对各个核函数分别进行训练并组合后,进行整体训练,得到每个核函数的权重系数。与现有技术相比,本发明所运用的多模态数据更接近云计算和大数据背景下患者信息流真实的形态,具有综合性和复杂性,利用多特征多核学习方法,在面对复杂的非线性多模态信息处理时减少了大量信息的损耗,且在处理模态跨度较大的数据时表现较好。



1. 一种基于统计学习的患者行为多模态分析与预测系统,其特征在于,包括:多模态采集模块、基于姿态的患者行为识别模块、基于生理信号的患者行为识别模块、基于情感信号的患者行为识别模块、基于语音信号的患者行为识别模块和基于多核学习的融合模块;所述基于姿态的患者行为识别模块、基于生理信号的患者行为识别模块、基于情感信号的患者行为识别模块和基于语音信号的患者行为识别模块均分别连接所述多模态采集模块和基于多核学习的融合模块;

所述患者行为多模态分析与预测系统采用预先建立并训练好的基于多核学习的融合模块进行数据处理,所述基于多核学习的融合模块包括多核分类器,所述基于多核学习的融合模块的训练过程具体为,首先对所述多核分类器的各个核函数分别进行训练并组合,然后对所述多核分类器进行整体训练,得到每个核函数的权重系数,从而获取训练后的所述基于多核学习的融合模块。

2. 根据权利要求1所述的一种基于统计学习的患者行为多模态分析与预测系统,其特征在于,所述多模态采集模块用于获取多模态数据,该多模态数据包括人体姿态数据、生理信号数据、情感表情表达数据和语音音频数据。

3. 根据权利要求2所述的一种基于统计学习的患者行为多模态分析与预测系统,其特征在于,所述基于姿态的患者行为识别模块包括依次连接的数据预处理单元、特征提取表示单元、分类器选择单元和标签输出单元;

所述数据预处理单元,通过线性插值的方法对所述人体姿态数据进行对齐操作。

4. 根据权利要求3所述的一种基于统计学习的患者行为多模态分析与预测系统,其特征在于,所述特征提取表示单元,通过快速傅里叶变换系数和离散余弦变换的方法进行频域特征的表示,通过小波分析进行时域特征的表示。

5. 根据权利要求2所述的一种基于统计学习的患者行为多模态分析与预测系统,其特征在于,所述基于生理信号的患者行为识别模块包括依次连接的生理信号采集单元、生理信号预处理单元、信号特征提取单元和模态输出单元;

所述生理信号采集单元,通过低通滤波和高通滤波的方法对所述生理信号数据进行滤波处理,删除所述生理信号数据中的死信道数据和坏信道数据,并采用被删除数据相邻通道的有效数据进行插值替换。

6. 根据权利要求2所述的一种基于统计学习的患者行为多模态分析与预测系统,其特征在于,所述基于情感信号的患者行为识别模块包括维度情感模型单元、维度情感标注单元和性能指标评价单元;

所述维度情感模型单元构建有多模态情绪的维度模型,该维度模型包括效力维、唤醒维和能力维,所述效力维为患者情绪程度的度量,所述唤醒维为患者生理活动和心理警觉水平的度量,所述能力维为影响周围环境及他人或受周围环境及他人影响的度量;

所述维度情感标注单元,用于基于所述维度模型进行情感标注;

所述性能指标评价单元,用于根据所述情感标注的结果,评价所述维度情感标注单元的估计性能。

7. 根据权利要求6所述的一种基于统计学习的患者行为多模态分析与预测系统,其特征在于,所述维度情感标注单元通过卡通小人进行情感标注,所述维度情感标注单元通过所述卡通小人面部表情的变化表示所述效力维的取值、通过所述卡通小人心脏位置出现的

震动程度以及眼睛的有神程度表示唤醒维的取值、通过所述卡通小人图片的大小表示能力维的取值。

8. 根据权利要求7所述的一种基于统计学习的患者行为多模态分析与预测系统, 其特征在于, 通过多个所述卡通小人, 对所述效力维、唤醒维和能力维进行标注。

9. 根据权利要求2所述的一种基于统计学习的患者行为多模态分析与预测系统, 其特征在于, 所述基于语音信号的患者行为识别模块包括依次连接的WAVE文件提取单元、语音预处理单元和MFCC特征提取单元;

所述语音预处理单元采用数字滤波器实现预加重, 并通过分帧操作, 获取由每一帧特征参数组成的特征参数时间序列。

10. 根据权利要求9所述的一种基于统计学习的患者行为多模态分析与预测系统, 其特征在于, 所述语音预处理单元还加窗操作。

## 一种基于统计学习的患者行为多模态分析与预测系统

### 技术领域

[0001] 本发明涉及患者行为识别领域,尤其是涉及一种基于统计学习的患者行为多模态分析与预测系统。

### 背景技术

[0002] 患者行为,指的是一种以医疗场景下的患者认知、情感、过去的经验等主观感受为中介的行为反应。具体指通过医疗诊断设备,例如B超、计算机断层扫描(CT)、核磁共振和定点的三维运动捕捉系统等采集到的患者图像信息、语音信息、生理信号等。

[0003] 面向医疗诊断场景和患者行为分析领域,在人工智能高速发展的当今,通过统计学习的手段采取数据驱动的方式,对多模态患者行为数据分析预测,能够为下一步的医疗结局相关性分析提出有力的数据支撑,为患者认知和康复治疗取得更为全面的解决方案。基于统计学习的多模态数据处理方法被视为一种多模态学习方法,目前在跨模态数据分类、多维度情感分析、多模态语义计算和异构数据的样本匹配方面取得了令人瞩目的效果。

[0004] 针对智慧医院建立基础的患者行为数据研究,目前统计学习研究对象是患者数据。通过提取患者行为数据特征,抽象出数据模型,发现数据的深层知识语义关系,又回到数据本身的分析和预测中去。作为机器学习的对象,患者行为数据是多模态和多样的,包括医疗诊断设备采集到的各种数字、文字、图像、视频、音频信号等。现有的统计学习方法中,核学习方法将低维不可分的样本通过核映射的方式映射到高维非线性空间,完成模态样本的有效分类;典型相关性分析(CCA),在多模态领域内被应用于度量两种模态信息之间的相关特征,并保持模态之间的相关性;还有共享子空间的学习方式,通过底层特征的表征特点,对多模态数据源的相关关系进行挖掘,得到多模态特征的一致性表示。

[0005] 目前大部分医疗场所和研究机构对于患者行为的信息采集和分析往往针对某一特定领域或者单方面,即只关注了单模态特定数据的分析和处理,相较于单一通道处理,多模态数据更接近云计算和大数据背景下患者信息流真实的形态,具有综合性和复杂性。传统的统计学习方法面对复杂的非线性多模态信息处理时往往容易造成大量信息的损耗,且在处理模态跨度较大的数据时表现不佳。

[0006] 受医疗科技水平和计算资源等的制约,现有的统计学习方法处理样本往往以中小规模的数据集居多,在庞大的患者行为多模态数据集处理上有所欠缺,例如病例记录文本信息处理、视频监控系統所提取的图像信息处理等任务中,大多需要技术人员参与特征提取和处理的过程,准确度不高的同时造成了大量的人力资源浪费。

[0007] 尽管基于单模态的患者行为识别研究已经取得了较好的成果,但是这些方法一般都局限在特定的场景下以及单一的数据集上,且很容易受到数据噪声的干扰,同时模型的迁移能力和数据集的泛化能力较弱。

[0008] 综上所述,基于改进统计学习的方法,利用患者行为的多模态数据,充分表达异构特征的判别力,克服现有的单模态数据研究困难,构建多模态患者行为感知和分析系统成为了本研究领域技术人员亟待解决的问题。

## 发明内容

[0009] 本发明的目的就是为了解决上述现有技术存在的缺陷而提供一种基于改进统计学习的方法,利用患者行为的多模态数据,充分表达异构特征的判别力,克服现有的单模态数据研究困难的基于统计学习的患者行为多模态分析与预测系统。

[0010] 本发明的目的可以通过以下技术方案来实现:

[0011] 一种基于统计学习的患者行为多模态分析与预测系统,包括:多模态采集模块、基于姿态的患者行为识别模块、基于生理信号的患者行为识别模块、基于情感信号的患者行为识别模块、基于语音信号的患者行为识别模块和基于多核学习的融合模块;所述基于姿态的患者行为识别模块、基于生理信号的患者行为识别模块、基于情感信号的患者行为识别模块和基于语音信号的患者行为识别模块均分别连接所述多模态采集模块和基于多核学习的融合模块;

[0012] 所述患者行为多模态分析与预测系统采用预先建立并训练好的基于多核学习的融合模块进行数据处理,所述基于多核学习的融合模块包括多核分类器,所述基于多核学习的融合模块的训练过程具体为,对所述多核分类器的各个核函数分别进行训练,然后对所述多核分类器整体进行训练,得到每个核函数的权重系数,从而获取训练后的所述基于多核学习的融合模块。

[0013] 进一步地,所述多模态采集模块用于获取多模态数据,该多模态数据包括人体姿态数据、生理信号数据、情感表情表达数据和语音音频数据。

[0014] 进一步地,所述基于姿态的患者行为识别模块包括依次连接的数据预处理单元、特征提取表示单元、分类器选择单元和标签输出单元;

[0015] 所述数据预处理单元,通过线性插值的方法对所述人体姿态数据进行对齐操作;

[0016] 所述特征提取表示单元,通过快速傅里叶变换系数和离散余弦变换的方法进行频域特征的表示,通过小波分析进行时域特征的表示。

[0017] 进一步地,所述基于生理信号的患者行为识别模块包括依次连接的生理信号采集单元、生理信号预处理单元、信号特征提取单元和模态输出单元;

[0018] 所述生理信号采集单元,通过低通滤波和高通滤波的方法对所述生理信号数据进行滤波处理,删除所述生理信号数据中的死信道数据和坏信道数据,并采用被删除数据相邻通道的有效数据进行插值替换。

[0019] 进一步地,所述基于情感信号的患者行为识别模块包括维度情感模型单元、维度情感标注单元和性能指标评价单元;

[0020] 所述维度情感模型单元构建有多模态情绪的维度模型,该维度模型包括效力维、唤醒维和能力维,所述效力维为患者情绪程度的度量,所述唤醒维为患者生理活动和心理警觉水平的度量,所述能力维为影响周围环境及他人或受周围环境及他人影响的度量;

[0021] 所述维度情感标注单元,用于基于所述维度模型进行情感标注;

[0022] 所述性能指标评价单元,用于根据所述情感标注的结果,评价所述维度情感标注单元的估计性能。

[0023] 进一步地,所述维度情感标注单元通过卡通小人进行情感标注,所述维度情感标注单元通过所述卡通小人面部表情的变化表示所述效力维的取值、通过所述卡通小人心脏位置出现的震动程度以及眼睛的有神程度表示唤醒维的取值、通过所述卡通小人图片的大

小表示能力维的取值。

[0024] 进一步地,通过设置所述卡通小人的数目,分别对所述效力维、唤醒维和能力维进行量化比较。

[0025] 进一步地,所述基于语音信号的患者行为识别模块包括依次连接的WAVE文件提取单元、语音预处理单元和MFCC特征提取单元;

[0026] 所述语音预处理单元采用数字滤波器实现预加重,并通过分帧操作,获取由每一帧特征参数组成的特征参数时间序列。

[0027] 进一步地,所述语音预处理单元还加窗操作。

[0028] 与现有技术相比,本发明具有以下优点:

[0029] (1) 本发明所提供的基于统计学习的患者行为多模态分析和预测系统,包含采集模块、基于姿态的患者行为识别模块、基于生理信号的患者行为识别模块、基于情感信号的患者行为识别模块、基于语音信号的患者行为识别模块、融合模块和显示模块;区别于传统患者行为数据采集模式,本发明的采集模块通过医疗影像技术和多传感器实现对患者行为的多模态数据的收集操作;通过不同的识别模块实现患者行为的多模态数据处理和表示,本发明所运用的多模态数据更接近云计算和大数据背景下患者信息流真实的形态,具有综合性和复杂性;同时融合模块不同于传统的统计学习方法,利用多特征多核学习方法,在面对复杂的非线性多模态信息处理时减少了大量信息的损耗,且在处理模态跨度较大的数据时表现较好。

[0030] (2) 基于多核学习的融合模块采用多核分类器,通过将多模态数据投影到不同的核空间进行映射,能够使数据在新的合成的空间进行更好的表达,从而提高数据的表达准确精度。

[0031] (3) 基于情感信号的患者行为识别模块构建的多模态情绪的维度模型,通过用几个取值连续的维度将情感刻画为一个多维信号,可以对情感的细微差别进行区分,同时可以通过对情感状态的实时标注跟踪情感状态的演变过程。

[0032] (4) 通过卡通小人进行情感标注的方法,具有简单、快速、直观的优点,并且避免了不同人对同一词语的不同理解造成的差异,从而获得的标注结果方差较小、不同标注者间的一致性较高。

## 附图说明

[0033] 图1为本发明提供的基于统计学习的患者行为多模态分析与预测系统的结构示意图;

[0034] 图2为本发明提供的基于姿态的患者行为识别模块的原理框图;

[0035] 图3为本发明提供的基于生理信号的患者行为识别模块的原理框图;

[0036] 图4为本发明提供的基于情感信号的患者行为识别模块的原理框图;

[0037] 图5为本发明提供的基于语音信号的患者行为识别模块的原理框图;

[0038] 图6为本发明提供的基于多核学习的多模态特征融合流程示意图。

## 具体实施方式

[0039] 下面结合附图和具体实施例对本发明进行详细说明。本实施例以本发明技术方案

为前提进行实施,给出了详细的实施方式和具体的操作过程,但本发明的保护范围不限于下述的实施例。

#### [0040] 实施例1

[0041] 本实施例提供一种基于统计学习的患者行为多模态分析与预测系统,包括:多模态采集模块、基于姿态的患者行为识别模块、基于生理信号的患者行为识别模块、基于情感信号的患者行为识别模块、基于语音信号的患者行为识别模块和基于多核学习的融合模块;多模态采集模块分别连接基于姿态的患者行为识别模块、基于生理信号的患者行为识别模块、基于情感信号的患者行为识别模块和基于语音信号的患者行为识别模块;基于姿态的患者行为识别模块、基于生理信号的患者行为识别模块、基于情感信号的患者行为识别模块和基于语音信号的患者行为识别模块均连接基于多核学习的融合模块。

[0042] 多模态采集模块包括患者姿态采集装置、生理信号采集装置、情绪信号采集装置和语音信号的采集装置等,通过医疗影像技术和多传感器实现对患者行为的多模态数据的收集操作。

[0043] 具体的,多模态数据主要来源于患者的人体姿态信息、生理信号信息、情感表情表达信息和语音音频信息。在识别模块通过不同模态的识别单元实现对多模态信息的预处理、数据特征提取以及指标评价和模态输出等操作。在融合模块通过一种基于多核学习的患者行为多模态特征融合与识别方法,通过将不同模态的数据和属性选取不同的核函数,进而采用特定方法对不同的核函数进行融合,实现了特征层的患者行为多模态特征融合,并在显示模块展示患者行为的诊断分析和康复治疗预测结果。

[0044] 下面对各模块进行详细描述。

#### [0045] 1、基于姿态的患者行为识别模块

[0046] 基于姿态的患者行为识别模块包括数据预处理,特征提取表示,分类器选择,标签输出四个单元。此模块通过医疗设备的微机电系统,利用加速度传感器实时的高频率采集人体关节点的加速度信号。基于患者人体姿态所呈现的每个动作序列长度和动作的快慢不同,特征提取之前进行对齐操作,即通过复制最后一个采样点的数据直到固定的长度,通过复制所有的采样点直到固定的长度和线性插值将实际数据长度进行对时齐到固定的长度。这里使用线性插值的方法进行原始数据对齐;之后通过快速傅里叶变换系数和离散余弦变换作为频域特征表示方法;利用小波分析作为时域特征表示方法。

[0047] 四个单元的具体步骤如下:

[0048] 1.1) 利用诊断监控设备获得姿态数据,通过位姿数据  $(x_0, y_0)$  和  $(x_1, y_1)$  计算  $[x_0, x_1]$  区间内某一位置  $x$  在直线上的  $y$  值。考虑加速度是采样率非常高的离散数据,在极短的时间内,通过如下公式进行线性插值得到的估计值和真实的数据将会非常接近。

$$[0049] \quad y = \frac{x_1 - x}{x_1 - x_0} y_0 + \frac{x - x_0}{x_1 - x_0} y_1$$

[0050] 通过数据分割使特征提取器能够捕捉到姿态信号的时间线索。利用滑动时间窗对加速度数据进行分割,实现滑动滤波的操作,使用10阶平均滑动滤波器对数据进行操作,其工作原理如下所示:

$$[0051] \quad y[n] = \frac{1}{2m+1} \left( \sum_{k=1}^m x(n-k) + \sum_{k=1}^m x[n+k] + x[n] \right)$$

[0052] 其中 $x[n]$ 表示当前采样点的原始数据,  $2m+1$ 表示滑动的窗口长度, 本发明中设置为10。本步骤中对加速度信号进行标准化和归一化的操作同时修正了原始信号的幅值。

[0053] 1.2) 对加速度传感器获得的一维人体姿态时间信号特征, 通过时频域特征进行特征提取。在对分割后的数据求取快速傅里叶变换系数时, 分别对数据的三维轴进行相应的FFT变换, 是抛弃作为地球加速的常亮的数据, 将三个轴的数据整理表示成一维特征的表示形式。在时域分析方向, 使用小波变换返回级别为 $n$ 的信号 $x$ 的小波分解。此模块输出分解结构由小波分解向量 $c$ 和薄记向量 $l$ 组成, 其中包含了按级别划分的系数。通过上述变化对分割后的三轴加速度数据进行WPD特征提取, 然后将三轴的数据整理成一维特征表示形式。

[0054] 1.3) 对输入的时频域姿态加速度数据分别进行不同的特征提取和表示, 之后利用合适的分类器进行训练, 通过完整的系统测试后对每个分类器进行决策级融合, 融合时选取两组分数的最大值作为最终模块的输出数据标签。基于姿态的患者行为识别模块对患者的慢跑、快走、起立和蹲下等多种运动姿态都有着良好的识别效果。

[0055] 2、基于生理信号的患者行为识别模块

[0056] 基于生理信号的患者行为识别模块包含了生理信号采集、生理信号预处理、信号特征提取和模态输出四个单元。生理信号采集方面通过专业的活性电极医疗设备和传感器采集脑电信号; 所述的脑电信号包括有肌电信号、心电信号和皮肤电信号等。生理信号预处理方面通过小波变换、滤波器滤波等进行脑电信号的预处理; 后续的通过对处理后的信号进行特征提取, 实现模态的输出。

[0057] 具体地, 基于生理信号的患者行为识别模块通过活性电极医疗设备和传感器采集到眼动、脸部肌肉活动、心电信号变化、呼吸频率后, 通过实验分割、光谱过滤和通道校正、特征提取四个步骤完成。

[0058] 2.1) 使用Fieldtrip处理, 消除环境和生理噪声。使用剃度计输出进行分析。本发明考虑将收集信号的每段数据的前4秒进行剔除, 去剩下的部分作为模块的有效数据。

[0059] 2.2) 在光谱过滤阶段, 将MEG信号的下采样到300Hz, 分别执行截止频率为95Hz和1Hz的低通滤波和高通滤波。利用高通滤波器可以去除信号中的低频环境噪声。一般为医疗诊断过程中所产生的周边环境噪声干扰。例如采集过程中医护人员使用设备时发出的干扰。相反, 低通滤波器可消除信号中的高频伪影噪声, 一般是指采集过程中患者产生的不可避免的生理噪声。例如在110Hz到150Hz之间的肌肉活动。

[0060] 2.3) 最后进行脑电信号的通道校正。从采集到的MEG数据中删除死信道和坏信道的数据、死信道数据指的是输出为零值的患者异常数据, 坏信道是指无效的数据, 即某些信号数据的方差和幅度 $z$ 分数随时间的度量的异常值。为了保持MEG数据在所有实验患者上的尺寸的一致性, 移除的通道将使用相邻通道的有效数据插值替换。计算钙通道上的MEG光谱功率的同时, 保留空间信息和减少数据维数。实现通道的校正。

[0061] 2.4) 根据时频域的特征方法, 对于EEG信号, 对每个通道分别提取具有代表性的四个频带光谱功率、对于MEG信号, 采用同样的方法提取具有代表性的四个频带中的脑磁图信号的离散余弦变换特征(DCT)。DCT特征可以有效的压缩MEG数据的时域特征而不影响数据的精度。

[0062] 3、基于情感信号的患者行为识别模块

[0063] 已知情感是一种常见的患者心理现象, 此模块通过收集患者的情感状态实现快速



准确地判断,可以进一步理解患者行为的情绪分量。此模块构建了情绪的维度模型,即“效力(Effect)-唤醒(Arousal)-能力(Capability)”模型,此模型可以看作情感具有效力维、唤醒维和能力维三个维度。效力维,是对患者的情绪程度的度量,用以表现患者情绪分量。之后通过维度情感标注,实现情感信息的处理和量化,最后通过整体分类准确率等标准实现维度情感信息的性能评估和模态指标评价。

[0064] 具体地,基于情绪信号的患者行为识别模块,利用医疗诊断数据库中收集到的基于患者行为的情绪表情数据集进行情绪数据的建模和分析,主要包括维度情感模型单元、维度情感标注单元、性能指标评价单元三个步骤:

[0065] 3.1) 维度情感模型单元构建了多模态情绪的维度模型,即“效力(Effect)-唤醒(Arousal)-能力(Capability)”模型或EAC模型,此模型认为情感具有效力维、唤醒维和能力维三个维度。效力维,是对患者的情绪程度的度量,从一个极端(苦恼)到另一个极端(狂喜);唤醒维也称为激活(Activation)维,是对患者生理活动和心理警觉水平的度量,如睡眠、厌倦等为低唤醒,清醒、紧张等为高唤醒;能力维,是指影响周围环境及他人(医护人员)或反过来受其影响的一种感受,高的支配度是一种有力、主宰感,而低的支配度是一种退缩、软弱感。

[0066] 具体来讲是用几个取值连续的维度将情感刻画为一个多维信号由于每个维度的取值可以连续变化,因此可以对情感的细微差别进行区分,同时可以通过对情感状态的实时标注跟踪情感状态的演变过程。

[0067] 3.2) 在维度情感标注单元中进行维度情感标注,其工作是基于情感量化理论完成的,此方法使用卡通小人的形象表示EAC模型中三个维度的取值,以卡通小人眉毛和嘴巴等面部表情的变化表示效力维的取值;以心脏位置出现的震动程度以及眼睛的有神程度表示唤醒维的取值;以图片的大小表示受控制的程度。在某个维度标注的过程中,只需从对应的卡通小人中选出一个最符合当前情感状态的即可。可使用多个小人对此维度进行更全面的描述,一般为5个或9个。此标注法具有简单、快速、直观的优点,并且避免了不同人对同一词语的不同理解造成的差异,从而获得的标注结果方差较小、不同标注者间的一致性较高。

[0068] 眼睛的有神程度可通过眼睛睁开的大小表示。

[0069] 3.3) 情感数据标注完成后在性能指标评价单元采用均方误差度量估计性能。设 $\hat{\theta}$ 是估计标签, $\theta$ 是真实标签, $n$ 为样本数目,将其公式定义如下:

$$[0070] \quad MSE = \frac{1}{n} \sum_{f=1}^n (\hat{\theta}(f) - \theta(f))^2$$

[0071] 通过此种方法很好的描述了情感数据预测值与真实值的偏差,反映了较好的模态吻合程度。

[0072] 4、基于语音信号的患者行为识别模块

[0073] 听觉模态是可用于患者行为识别的一个重要模态,声音信号中既有明确的语言信息又有非语言的声学信息,这些信息都可用于习惯行为状态的推断。很多临床实验都表明使用音频信息能够获得更好的行为预测效果。

[0074] 基于语音信号的患者行为识别模块中,通过语音识别模块对患者既往诊断过程中的对病情的叙述实现记录。包括了WAVE文件提取、语音预处理和MFCC特征提取三个单元。WAVE文件是基于患者行为的音频收集中一种广泛的声音标准格式。整个WAVE文件可以分成

两部分：前一部分为文件头，后一部分为数据块。根据其编码方式和采样数的不同，这两部分的大小有所不同。在WAVE文件中，所采用的编码方式有脉冲编码调制和自适应差分脉冲编码调制两种。在对语音信号进行分析和处理之前，语音预处理单元对其进行预加重、分帧、加窗等预处理操作。这些操作的目的是消除因为患者发声器官本身和由于采集语音信号的设备所带来的混叠、高次谐波失真、高频等等因素，对语音信号质量的影响。尽可能保证后续语音处理得到的信号更均匀、平滑，为信号参数提取提供优质的参数，提高语音处理质量。

[0075] 在语音信号中，包含着非常丰富的特征参数，不同的特征向量表征着不同的物理和声学意义。特征提取阶段就取出或削减语音信号中与识别无关的信息的影响，减少后续识别阶段需处理的数据量，生成表征语音信号中携带的患者说话信息的特征参数。根据语音特征的不同用途，需要提取不同的特征参数，从而保证识别的准确率。

[0076] 三个单元的具体步骤如下：

[0077] 4.1) WAVE文件提取单元中，WAVE文件是一种RIFF文件，通过将RIFF块分为两个子块，并将两个子块的ID分别命名为“fmt”和“data”，其中“fmt”子块由结构PCMWAVEFORMAT所组成，其子块的大小就是sizeof，数据组成就是其结构中的数据。

[0078] 4.2) 语音预处理单元中考虑到语音信号 $s(n)$ 的平均功率谱受声门激励和口鼻辐射的影响，首先进行预加重操作。为此将高频端大约在800Hz以上按6dB/oct（倍频程）衰减，即频率越高相应的成分越小，同时对语音信号 $s(n)$ 进行分析之前对其高频部分加以提升。本单元采用的措施是用数字滤波器实现预加重，预加重网络的输出和输入的语音信号 $s(n)$ 的关系为：

$$[0079] \quad \tilde{S}(n) = S(n) - aS(n-1)$$

[0080] 接下来进行分帧操作。本单元采用语音分析全过程中的“短时分析技术”。由于语音信号具有时变特性，但是在一个短时间范围内其特性基本保持不变即相对稳定，因而可以将其看作是一个准稳态过程，即语音信号具有短时平稳性。为此将语音信号分段来分析其特征参数，其中每一段称为一帧，帧长一般取为10~30ms。这样对于整体的语音信号来讲，分析出的是由每一帧特征参数组成的特征参数时间序列。

[0081] 4.3) MFCC特征提取单元中，MFCC参数是基于人的听觉特性利用人听觉的临界带效应，在Mel标度频率域提取出来的倒谱特征参数。其中Mel倒谱系数是根据人类听觉系统的特性提出的，模拟人耳对不同频率语音的感知。人耳分辨声音频率的过程就像一种取对数的操作。例如：在Mel频域内，人对音调的感知能力为线性关系，如果两段语音的Mel频率差两倍，则人在感知上也差两倍。对于MFCC算法，首先进行快速傅里叶变换如下：

$$[0082] \quad X[k] = \sum_{n=0}^{N-1} x[n] e^{-j \frac{2\pi}{N} nk}$$

[0083] 之后将实际频率尺度转换为Mel频率尺度，如下：

$$[0084] \quad Mel(f) = 2597 \lg(1 + \frac{f}{700})$$

[0085] 配置三角形滤波器组并计算每一个三角形滤波器对信号幅度谱滤波后的输出，对所有滤波器输出作对数运算，再进一步做离散余弦变换(DCT)，即可得到MFCC，如下：

$$[0086] \quad M(i) = \sqrt{\frac{2}{N}} \sum_{l=1}^L \log F(l) \cos\left[\left(l - \frac{1}{2}\right) \frac{i\pi}{L}\right]$$

[0087] 式中, N为信号截取帧的大小, F(l) 为中心频率函数, l为余弦角取值, L为MFCC系数阶数。

#### [0088] 5、基于多核学习的融合模块

[0089] 多核学习的融合模块, 提出了一种基于多核学习的患者行为多模态特征融合与识别方法, 即首先根据姿态、生理、情感和语音识别模块提取患者行为的多模态特征, 然后将每一种特征输入到一组核组合的SVM分类器中进行多核学习, 得到该特征的一组核的权重系数, 构成每个特征的核矩阵, 之后将每个核矩阵组合起来再进行多核权重学习, 得到每一组核的权重系数, 最后将得出的权重组合成新的核空间, 进行多核SVM分类。该方法有效地对患者行为的多模态特征进行融合, 实现对患者行为所展现的异构信息的分类和对患者行为综合体现信息的预测分析。

[0090] 如图6所示, 为基于多核学习的多模态特征融合流程示意图, 本发明提出的多核学习(MKL) 是支持向量机的扩展, 是鲁棒性和灵活性更强的基于核的学习模型。其通过将多模态数据投影到不同的核空间进行映射, 能够使数据在新的合成的空间进行更好的表达, 从而提高数据的表达准确精度。本发明采取的MKL将多核学习问题转化为光滑且凸的优化问题, 之后采用迭代的方法固定 $d_m$ , 然后结合支持向量机训练算法构建分类超平面, 根据梯度下降方向更新 $d_m$ , 将基于多核学习的特征融合转化为优化方向, 根据不同模态特征关系, 得到如下ST关系:

$$[0091] \quad \min_d J(d)$$

$$[0092] \quad s.t. \sum_{m=1}^M d_m = 1, d_m > 0, \forall m$$

$$[0093] \quad J(d) = \begin{cases} \min \frac{1}{2} \sum_{m=1}^M \frac{1}{d_m} \|f_m\|^2 + C \sum_{i=1}^n \zeta_i \\ s.t. y_i \sum_{m=1}^M f_m + y_i b \geq 1 - \zeta_i, \forall i \\ \zeta_i > 0, \forall i \end{cases}$$

[0094] 式中,  $d_m$ 为初始权重值, M为核矩阵度量,  $f_m$ 为松弛变量, C为正则化系数,  $\zeta_i$ 为损失系数,  $y_i$ 为超平面法向量, b为超平面截距, i为边界系数。

[0095] 融合模块中, 在一种特征中用一组核学习, 得到对应的一组权重参数, 在第二种特征值用一组核学习, 得到另一组权重参数, 以此类推运用到其他模态中进行核学习。然后将不同种类参数分别代入到对应的核矩阵组成新的矩阵, 再进行多核学习, 从而得到其他共四组权重, 在测试阶段, 四组权重组合成新的多核SVM分类器, 进行识别。

$$[0096] \quad K^* = \sum_{t=1}^T \lambda_t k_t$$

[0097] 其中T为患者多模态特征的种类数,  $\lambda$ 为对应特征的核矩阵的权重, k为特征t的核

矩阵,

[0098]  $K^*$ 为总的多特征多核SVM的核矩阵。在多模态数据训练阶段,先求出相应的初始权重系数 $d_m$ ,使 $K^*$ 满足如下关系式:

$$[0099] \quad K^* = \sum_{t=1}^T (\lambda_t (\sum_{i=1}^M d_m k_m))$$

[0100] 从而训练出数据特征的 $\lambda_m$ ,得到多特征融合的多核矩阵 $K$ 。由传统的SVM得到最后的判决函数如下:

$$[0101] \quad f(x) = \text{sign} \left\{ \sum_{i=1}^m a_i \cdot y_i \cdot K(x_i, x) + b \right\}$$

[0102] 式中,  $\text{sign}\{*\}$  为符号函数,  $K(x_i, x)$  表示从 $x_i$ 到 $x$ 维度的向量。

[0103] 判决函数的作用是找出满足到不同类训练数据的间距最大的超平面。

[0104] 融合模块的多核学习方法在初始阶段每个权重的值相等,为保证收敛速度,在开始是随机产生和为1的权重分量,使算法的收敛速度加快。

[0105] 作为进一步改进,本实施例基于统计学习的患者行为多模态分析和预测系统还包括显示模块,该显示模块连接基于多核学习的融合模块,用于对基于多核学习的机器学习算法分析和预测后的融合信息进行实时显示。

[0106] 本实施例展示的基于统计学习的患者行为多模态分析和预测系统很好的解决了患者行为异构数据的识别和预测问题,通过充分采集患者的生理信号、情感信息、语音信号和姿态信息等多模态信息,结合不同识别模块实现模态信息的预处理和特征提取,进一步利用多核学习方式实现决策层的患者行为模态信息融合,区别于传统单模态的分析方式,本发明的多核学习模型的迁移能力和相关患者行为数据集的泛化能力较强,可以综合的向医护人员反映患者的病理状态、预测患者的未来身心健康走势、评估患者行为在情绪、生理等方面的不同走向,为下一步基于患者行为的康复治疗 and 病理分析提供可靠保障。

[0107] 以上详细描述了本发明的较佳具体实施例。应当理解,本领域的普通技术人员无需创造性劳动就可以根据本发明的构思做出诸多修改和变化。因此,凡本技术领域技术人员依本发明的构思在现有技术的基础上通过逻辑分析、推理或者有限的实验可以得到的技术方案,皆应在由权利要求书所确定的保护范围内。

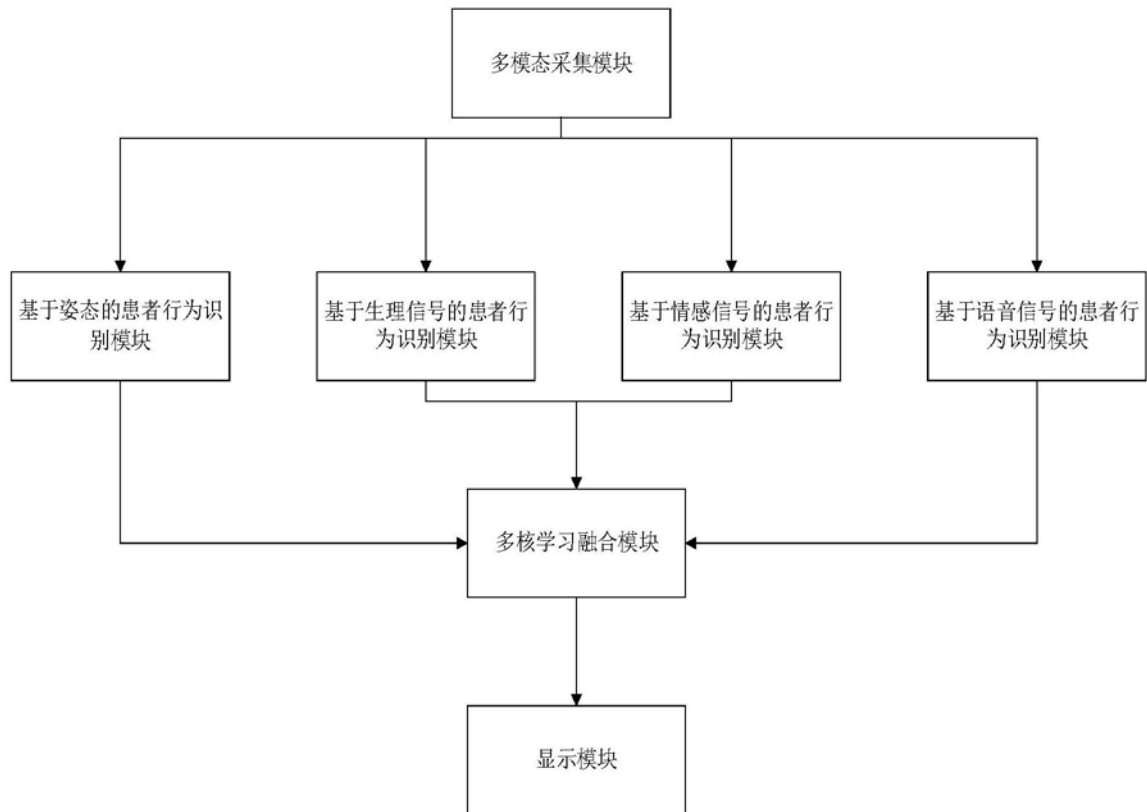


图1

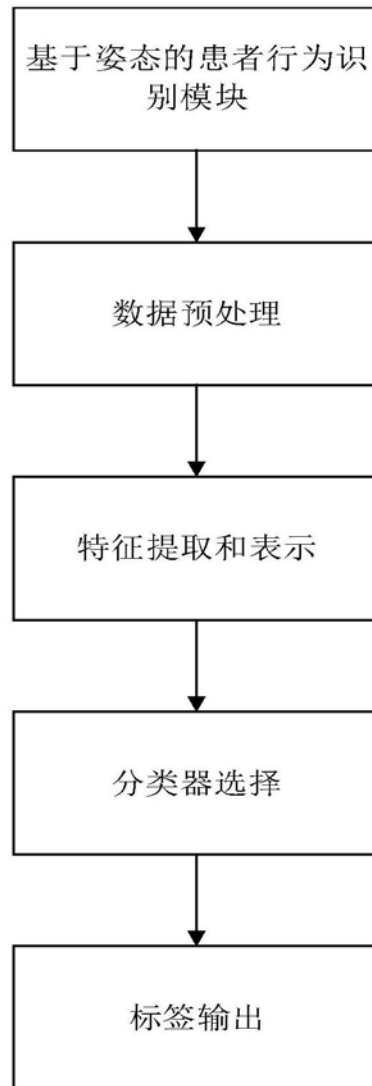


图2

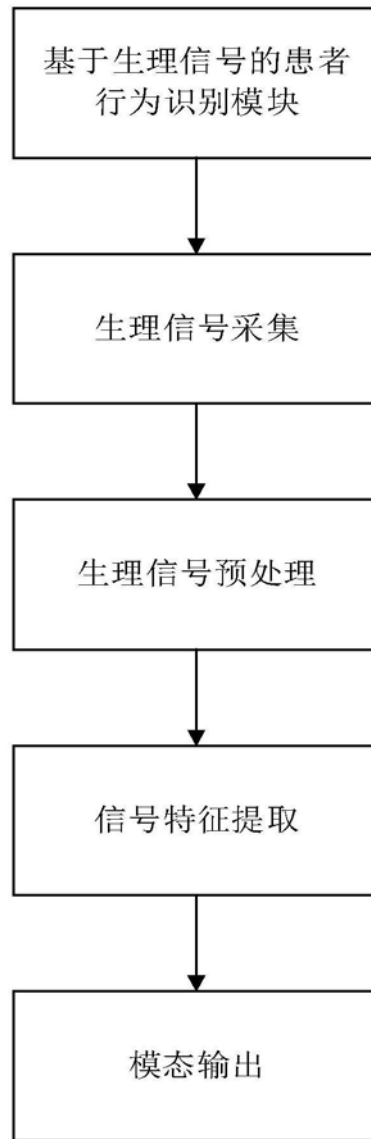


图3

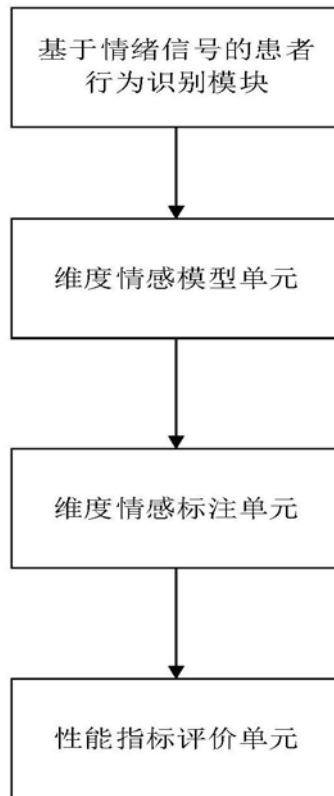


图4



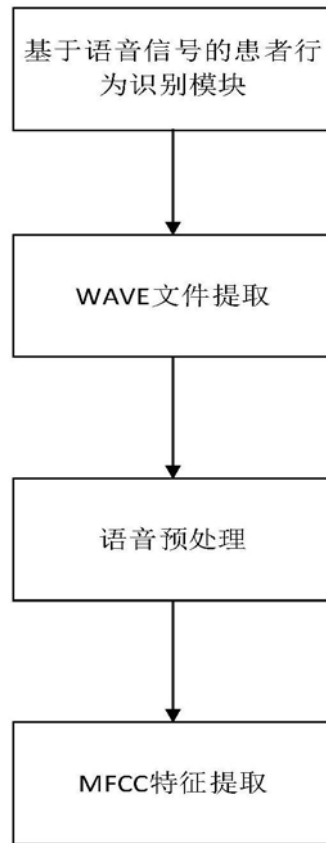


图5

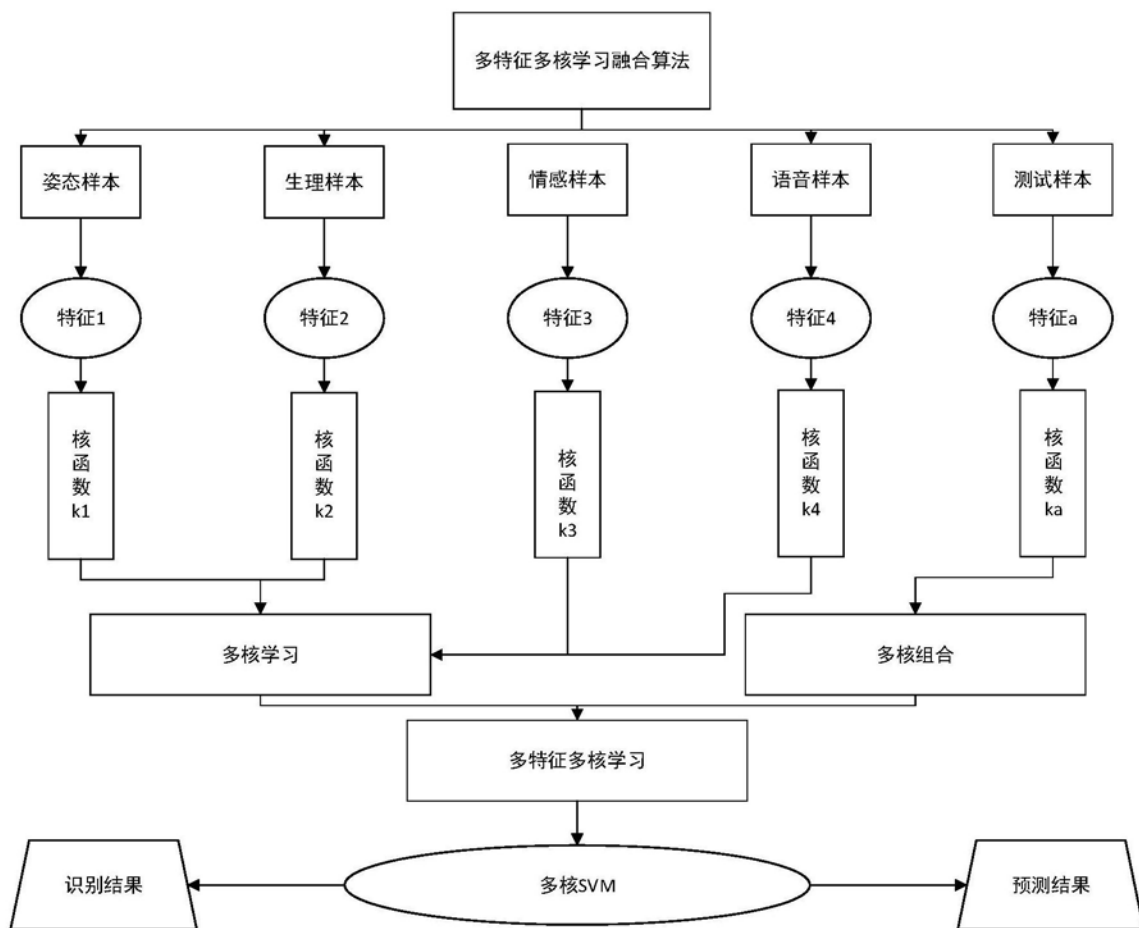


图6