



(12) 发明专利

(10) 授权公告号 CN 114974312 B

(45) 授权公告日 2022.10.21

(21) 申请号 202210902311.1

G10L 25/24 (2013.01)

(22) 申请日 2022.07.29

G10L 25/03 (2013.01)

G06N 20/00 (2019.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 114974312 A

(43) 申请公布日 2022.08.30

(73) 专利权人 环球数科集团有限公司

地址 518063 广东省深圳市南山区粤海街道高新南九道10号深圳湾科技生态园10栋B座17层01-03号

(72) 发明人 张卫平 丁烨 彭中华 张思琪

米小武 丁园

(74) 专利代理机构 北京清控智云知识产权代理

事务所(特殊普通合伙)
11919

专利代理师 马肃 林淡如

(51) Int. Cl.

G10L 25/63 (2013.01)

G10L 25/27 (2013.01)

G10L 15/22 (2006.01)

G10L 15/02 (2006.01)

(56) 对比文件

CN 114003643 A, 2022.02.01

CN 111273764 A, 2020.06.12

US 2017206095 A1, 2017.07.20

US 9812151 B1, 2017.11.07

CN 112927721 A, 2021.06.08

CN 108009573 A, 2018.05.08

CN 113436622 A, 2021.09.24

CN 107329990 A, 2017.11.07

吴伟国等.PAD情感空间内人工情感建模及人机交互实验.《哈尔滨工业大学学报》.2019,第51卷(第1期),

Fung,P等.Towards Empathetic Human-Robot Interactions.《COMPUTATIONAL LINGUISTICS AND INTELLIGENT TEXT PROCESSING》.2018,

审查员 可杨

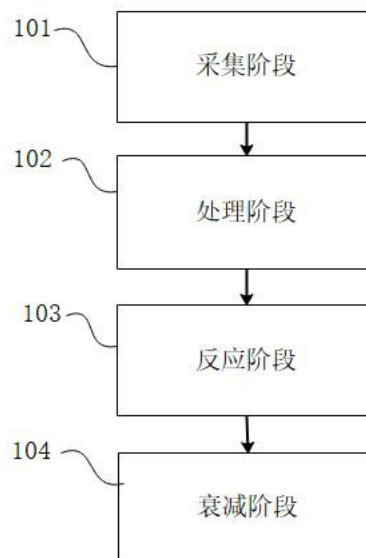
权利要求书2页 说明书9页 附图2页

(54) 发明名称

一种虚拟人情绪生成方法与系统

(57) 摘要

本发明涉及一种虚拟人情绪生成方法以及系统;所述生成方法通过采集模块记录用户与虚拟人交流过程中的声音音频,并从声音音频中提取多个特征以及特征值;进一步的,使用分类模块对多个特征以及特征值进行基于n项情绪的分类,并计算n项情绪中每一项的情绪能量;进一步的,获取虚拟人实时的情绪指数,根据实时的情绪指数获取对应的情绪反应函数,并计算情绪增量;该情绪反应函数用于反映基于当前情绪状况下,特定情绪能量能产生对情绪的影响效果;进一步的,根据情绪增量更新所述情绪参数后,驱动虚拟人的情绪发生相应变化。



1. 一种虚拟人情绪生成方法,所述生成方法运行在一种虚拟人情绪生成系统,其特征在于,所述生成系统包括采集模块、特征提取模块、分类模块以及情绪驱动模块;其中,

所述采集模块,被配置为用于采集用户与虚拟人交流时的声音,并将声音转换为可被特征提取模块识别的数字音频信息;

所述特征提取模块,被配置为从一段音频信息中进行特征提取,从而获得多个特征以及对应的特征值;

所述分类模块,被配置为使用一段音频信息中包括的多个特征以及对应特征值中的至少一部分进行情绪分类,从而计算该段音频信息中提供的n项情绪参数能量 q_1, q_2, \dots, q_n ,并将计算结果上传到情绪驱动模块;

所述情绪驱动模块,被配置为保存情绪参数,根据所述分类模块的计算结果,更新所述情绪参数,并根据所述情绪参数生成虚拟人的实时情绪表现;其中,所述生成方法包括采集阶段、处理阶段以及反应阶段;

在所述采集阶段,包括以下步骤:

S1:采集并存储当前与虚拟人交流的用户音频信息;

S2:提取音频信息中的语气和语义中包含的多个特征,并分析每个特征对应的特征值 $f_i, i=1, 2, 3, \dots$;

S3:基于多个特征值 f_i 对音频进行基于n项情绪的分类,并且计算分类后的n项情绪每项的强度,从而获得该音频提供的n项情绪参数能量 q_1, q_2, \dots, q_n ;

在所述处理阶段,包括以下步骤:

E1:获取驱动虚拟人情绪表现的情绪参数EM及其数值,且 $EM=[em_1, em_2, \dots, em_n]$,其中 em_1, em_2, \dots, em_n 为所述情绪参数EM中的n项情绪子参数的数值;

E2:基于多个情绪子参数的数值,确定每项情绪的情绪等级 $EL=[e_{l_1}, e_{l_2}, \dots, e_{l_n}]$,并调用对应该等级的情绪反应函数 $FEL()$,即 $FEL()=[Fel_1(), Fel_2(), \dots, Fel_n()]$;

E3:将n项情绪参数能量作为激励,输入到对应的情绪反应函数中,获取情绪增量 ΔEM ;所述情绪增量 ΔEM 包括n项情绪子增量,即 $\Delta EM=[\Delta em_1, \Delta em_2, \dots, \Delta em_n]$,并且:

$$\Delta em_i = Fel_i(q_i), \text{式1};$$

式1中, $i=1, 2, \dots, n$;

在所述反应阶段,包括以下步骤:

R1:获取情绪模拟单元中的所述情绪增量 ΔEM 的阈值,阈值包括上限 $\max(\Delta em_i)$ 和下限 $\min(\Delta em_i)$;

R2:根据以下条件式,计算每一项所述情绪子增量 Δem_i 的实际效用能量 $\Delta ef-em_i$:

$$\Delta ef - em_i = \begin{cases} \max(\Delta em_i), & \Delta em_i > \max(\Delta em_i) \\ \Delta em_i, & \min(\Delta em_i) < \Delta em_i \leq \max(\Delta em_i) \\ 0, & \Delta em_i \leq \min(\Delta em_i) \end{cases}$$

上式表明,若一项所述情绪子增量 Δem_i 超过其上限 $\max(\Delta em_i)$,则抑制该项所述情绪子增量 Δem_i ;另一方面若所述情绪子增量 Δem_i 未达到一阈值 $\min(\Delta em_i)$,则忽略其作用;其中,上限 $\max(\Delta em_i)$ 和下限 $\min(\Delta em_i)$ 决定虚拟人的情绪承受能力,由技术人员根据虚拟

人的情绪承受力设定；

R3:更新所述情绪参数EM,即使更新后的 $EM' = EM + \Delta EM$,从而由更新后的所述情绪参数 EM' 生成新的虚拟人最新的情绪表现。

2.如权利要求1所述一种虚拟人情绪生成方法,其特征在于,所述n项情绪分类的类别中至少包括中立情绪;并且分类的类别中还包括以下情绪的三项或三项以上:快乐、愤怒、恐惧、悲伤和厌恶。

3.如权利要求2所述一种虚拟人情绪生成方法,其特征在于,在采集阶段的步骤S2中,包括:

读取音频信息;

从音频信息中计算一组梅尔频率倒谱系数;

从音频信息中确定多个声学变量;

计算这些声学变量在音频信息中的平均值、标准差、最大值和最小值。

4.如权利要求3所述一种虚拟人情绪生成方法,其特征在于,在采集阶段中,包括采用基于机器学习的情绪分类模型进行情绪分类;所述情绪分类模型基于支持向量机内核而建立。

5.如权利要求4所述一种虚拟人情绪生成方法,其特征在于,所述生成方法还包括衰减阶段;当虚拟人与外部不存在任何语音交流时随即开始所述衰减阶段;在所述衰减阶段中的每个时间周期 Δt 中,所述情绪参数EM的每一项情绪子参数 em_i 基于以下衰减函数进行更新:

$$em'_i = em_i \cdot e^{-c_i \cdot \Delta t}, \text{式2};$$

上式中,e为自然常数, c_i 为第i项情绪的衰减系数且 $c_i > 0$,由技术人员根据虚拟人的性格决定; Δt 为单位时间周期。

一种虚拟人情绪生成方法与系统

技术领域

[0001] 本发明涉及电子数字数据处理技术领域。具体而言,涉及一种虚拟人情绪生成方法与系统。

背景技术

[0002] 虚拟人技术随着近年虚拟空间图像技术以及人工智能技术的高速发展,成为了一项社会重点研究和发展的领域,是体现以下两项技术的综合应用实力的一项最重要体现。而关于虚拟人的情绪驱动生成方面的研究目前还存在较大的开发困难。由于情绪具有较大的情境性和短暂性,真实人类本身的情绪研究本身仍存在较大的不确定性,因此虚拟人的情绪生成机制和系统目前尚有很大的开发空间。

[0003] 在一个方面,虚拟人在与人类进行互动的实质,是由虚拟人本身的个性,通过接受外部的刺激,并分析刺激的本质,从而作出合理的反应;因此需要建立虚拟人本身的个性属性,从而形成一个情绪机制启动的出发点和基础;另一方面,需要分析和处理外部环境对于虚拟人的影响,并计算虚拟人对于外部影响所产生的反应以及下一步动作的动机,从而生成兼具合理性、拟人性以及具有可持续发展的情绪表现。

[0004] 查阅相关已公开的技术方案,公开号为KR1020080074758A的技术方案通过检测虚拟空间的环境信息并根据对象之间的相互作用来检测发生的外部事件来生成传感器值,并通过该传感器值驱动虚拟人物的情绪或生理状态进行变化;公开号为JP2003071772A的技术方案中,对出一种利用电磁感应传感器进行外部感应的电子宠物,并通过所感应的物件生成该电子宠物的不同情绪与外部反应;公开号为W02021150492A1的技术方案通过读取并分析虚拟角色在游戏中的经历以及虚拟角色在游戏中的配置设定文件,从而基于经历中不同的特征值驱动虚拟角色的性格以及情感发生不变程度的改变。然而目前的技术方案所设定的虚拟人物情绪表现相对固化,还无法实时地基于虚拟人当前的经历展示出更近似于人类的情绪化表现。

[0005] 背景技术的前述论述仅意图便于理解本发明。此论述并不认可或承认提及的材料中的任一种公共常识的一部分。

发明内容

[0006] 本发明的目的在于,提供一种虚拟人情绪生成方法以及系统;所述生成方法通过采集模块记录用户与虚拟人交流过程中的声音音频,并从声音音频中提取多个特征以及特征值;进一步的,使用分类模块对多个特征以及特征值进行基于n项情绪的分类,并计算n项情绪中每一项的情绪能量;进一步的,获取虚拟人实时的情绪指数,根据实时的情绪指数获取对应的情绪反应函数,并计算情绪增量;该情绪反应函数用于反映基于当前情绪状况下,特定情绪能量能产生对情绪的影响效果;进一步的,根据情绪增量更新所述情绪参数后,驱动虚拟人的情绪发生相应变化。

[0007] 本发明采用如下技术方案:

[0008] 一种生成虚拟人情绪生成系统;所述生成系统包括:

[0009] 一种虚拟人情绪生成系统,所述生成系统包括采集模块、特征提取模块、分类模块以及情绪驱动模块;其中

[0010] 所述采集模块,被配置为用于采集用户与虚拟人交流时的声音,并将声音转换为可被所述特征提取模块识别的数字音频信息;

[0011] 所述特征提取模块,被配置为从一段音频信息中进行特征提取,从而获得多个特征以及对应的特征值;

[0012] 所述分类模块,被配置为使用一段音频信息中包括的多个特征以及对应特征值中的至少一部分进行情绪分类,从而计算该段音频信息中提供的n项情绪参数能量 $q_1, q_2 \dots q_n$,并将计算结果上传到情绪驱动模块;

[0013] 所述情绪驱动模块,被配置为保存情绪参数,根据所述分类模块的计算结果,更新所述情绪参数,并根据所述情绪参数生成虚拟人的实时情绪表现;

[0014] 进一步的,所述生成系统包括一种虚拟人情绪生成方法;所述生成方法包括采集阶段、处理阶段以及反应阶段;

[0015] 在所述采集阶段,包括以下步骤:

[0016] S1:采集并存储当前与虚拟人交流的用户的音频信息;

[0017] S2:提取音频信息中的语气和语义中包含的多个特征,并分析每个特征对应的特征值 $f_i, i=1, 2, 3 \dots$;

[0018] S3:基于多个特征值 f_i 对音频进行基于n项情绪的分类,并且计算分类后的n项情绪每项的强度,从而获得该音频提供的n项情绪参数能量 $q_1, q_2 \dots q_n$;

[0019] 在所述处理阶段,包括以下步骤:

[0020] E1:获取驱动虚拟人情绪表现的每项情绪的情绪参数EM,即 $EM=[em_1, em_2 \dots em_n]$,其中 $em_1, em_2 \dots em_n$ 为所述情绪参数EM中的n项情绪子参数;

[0021] E2:基于多个情绪子参数的当前值,确定每项情绪的情绪等级 $EL=[el_1, el_2 \dots el_n]$,并调用对应该等级的情绪反应函数 $FEL()$,即 $FEL()=[Fel_1(), Fel_2() \dots Fel_n()]$;

[0022] E3:将n项情绪参数能量作为激励,输入到对应的情绪反应函数中,获取情绪增量 ΔEM ;所述情绪增量 ΔEM 包括n项情绪子增量,即 $\Delta EM=[\Delta em_1, \Delta em_2 \dots \Delta em_n]$,并且:

[0023] $\Delta em_i = Fel_i(q_i)$,式1;

[0024] 式1中, $i=1, 2, \dots n$;

[0025] 在所述反应阶段,包括以下步骤:

[0026] R1:获取情绪模拟单元中的所述情绪增量 ΔEM 的阈值,阈值包括上限 $\max(\Delta em_i)$ 和下限 $\min(\Delta em_i)$;

[0027] R2:根据以下条件式,计算每一项所述情绪子增量 Δem_i 的实际效用能量 $\Delta ef-em_i$:

$$[0028] \quad \Delta ef - em_i = \begin{cases} \max(\Delta em_i), & \Delta em_i > \max(\Delta em_i) \\ \Delta em_i & , \min(\Delta em_i) < \Delta em_i \leq \max(\Delta em_i) \\ 0, & \Delta em_i \leq \min(\Delta em_i) \end{cases}$$

[0029] 上式表明,若一项所述情绪子增量 Δem_i 超过其上限 $\max(\Delta em_i)$,则抑制该项所述

情绪子增量 Δem_i ; 另一方面若所述情绪子增量 Δem_i 未达到一阈值 $\min(\Delta em_i)$, 则忽略其作用; 其中, 上限 $\max(\Delta em_i)$ 和下限 $\min(\Delta em_i)$ 决定虚拟人的情绪承受能力, 由技术人员根据虚拟人的情绪承受力设定。

[0030] R3: 更新所述情绪参数 EM, 即使更新后的 $EM' = EM + \Delta EM$, 从而由更新后的所述情绪参数 EM' 生成新的虚拟人最新的情绪表现;

[0031] 可选地, 所述 n 项情绪分类的类别中至少包括中立情绪; 并且分类的类别中还包括以下情绪的三项或三项以上: 快乐、愤怒、恐惧、悲伤和厌恶;

[0032] 可选地, 在采集阶段的步骤 S2 中, 包括:

[0033] 读取音频信息;

[0034] 从音频信息中计算一组梅尔频率倒谱系数;

[0035] 从音频信息中确定多个声学变量;

[0036] 计算这些声学变量在音频信息中的平均值、标准差、最大值和最小值;

[0037] 可选地, 在采集阶段中, 包括采用基于机器学习的情绪分类模型进行情绪分类; 所述情绪分类模型基于支持向量机内核而建立;

[0038] 可选地, 所述生成方法还包括衰减阶段; 当虚拟人与外部不存在任何语音交流时随即开始所述衰减阶段; 在所述衰减阶段中的每个时间周期 Δt 中, 所述情绪参数 EM 的每一项情绪子参数 em_i 基于以下衰减函数进行更新:

[0039] $em'_i = em_i \cdot e^{-c_i \cdot \Delta t}$, 式 2;

[0040] 上式中, e 为自然常数, c_i 为第 i 项情绪的衰减系数且 $c_i > 0$, 由技术人员根据虚拟人的性格决定; Δt 为单位时间周期, 取值为 0.1 秒, 0.5 秒, 1 秒或者其他时间数值, 由技术人员根据虚拟人的情绪稳定性决定。

[0041] 本发明所取得的有益效果是:

[0042] 1. 本发明的生成方法基于用户与虚拟人进行声音交流时的语音特征, 分析用户本身的语气以及语义所具有的内在情绪表现, 从而驱动虚拟人作对应的情绪反应;

[0043] 2. 本发明的生成法基于情绪参数驱动虚拟人的实时情绪反应; 通过设定在不同情绪参数下具有不同的情绪反应函数, 模拟人类处理不同情绪状态下针对不同的外部刺激具有超出一般表现的情绪反应; 通过设定情绪增量的上、下限阈值, 模拟虚拟人对于不同情绪的承受能力以及敏感程度;

[0044] 3. 本发明的生成系统适用于纯数字形象的虚拟人系统, 同时亦可适用接入诸如虚拟机器人或者其他虚拟-现实结合的虚拟人系统;

[0045] 本发明的生成系统其硬件模块以及装置采用模块化设计和配合, 后期可通过软件、硬件进行灵活优化和变更, 节省了大量后期维护升级成本。

附图说明

[0046] 从以下结合附图的描述可以进一步理解本发明。图中的部件不一定按比例绘制, 而是将重点放在示出实施例的原理上。在不同的视图中, 相同的附图标记指定对应的部分。

[0047] 图 1 为本发明所述生成方法的原理示意图;

[0048] 图 2 为本发明中所述采集阶段的步骤示意图;

[0049] 图3为本发明实施例中对情绪进行分类的示意图；

[0050] 图4为本发明中三个不同的情绪等级下的情绪反应函数的示意图。

[0051] 附图图例说明：101-采集阶段；102-处理阶段；103-反应阶段；104-衰减阶段；301-音频信息；302-一阶段分类；303-类别1的二阶段分类；304-类别2的二阶段分类。

具体实施方式

[0052] 为了使得本发明的目的技术方案及优点更加清楚明白，以下结合其实施例，对本发明进行进一步详细说明；应当理解，此处所描述的具体实施例仅用于解释本发明，并不用于限定本发明。对于本领域技术人员而言，在查阅以下详细描述之后，本实施例的其它系统、方法和/或特征将变得显而易见。旨在所有此类附加的系统、方法、特征和优点都包括在本说明书内，包括在本发明的范围内，并且受所附权利要求书的保护。在以下详细描述描述了所公开的实施例的另外的特征，并且这些特征根据以下将详细描述将是显而易见的。

[0053] 本发明实施例的附图中相同或相似的标号对应相同或相似的部件；在本发明的描述中，需要理解的是，若有术语“上”、“下”、“左”、“右”等指示的方位或位置关系为基于附图所示的方位或位置关系，仅是为了便于描述本发明和简化描述，而不是指示或暗示所指的装置或组件必须具有特定的方位，以特定的方位构造和操作，因此附图中描述位置关系的用语仅用于示例性说明，不能理解为对本专利的限制，对于本领域的普通技术人员而言，可以根据具体情况理解上述术语的具体含义。

[0054] 实施例一：

[0055] 一种生成虚拟人情绪生成系统；所述生成系统包括：

[0056] 一种虚拟人情绪生成系统，所述生成系统包括采集模块、特征提取模块、分类模块以及情绪驱动模块；其中

[0057] 所述采集模块，被配置为用于采集用户与虚拟人交流时的声音，并将声音转换为可被所述特征提取模块识别的数字音频信息；

[0058] 所述特征提取模块，被配置为从一段音频信息中进行特征提取，从而获得多个特征以及对应的特征值；

[0059] 所述分类模块，被配置为使用一段音频信息中包括的多个特征以及对应特征值中的至少一部分进行情绪分类，从而计算该段音频信息中提供的n项情绪参数能量 $q_1, q_2 \dots q_n$ ，并将计算结果上传到情绪驱动模块；

[0060] 所述情绪驱动模块，被配置为保存情绪参数，根据所述分类模块的计算结果，更新所述情绪参数，并根据所述情绪参数生成虚拟人的实时情绪表现；

[0061] 如附图1所示，进一步的，所述生成系统包括一种虚拟人情绪生成方法；所述生成方法包括采集阶段、处理阶段以及反应阶段；

[0062] 在所述采集阶段，包括以下步骤：

[0063] S1:采集并存储当前与虚拟人交流的用户音频信息；

[0064] S2:提取音频信息中的语气和语义中包含的多个特征，并分析每个特征对应的特征值 $f_i, i=1, 2, 3 \dots$ ；

[0065] S3:基于多个特征值 f_i 对音频进行基于n项情绪的分类，并且计算分类后的n项情绪每项的强度，从而获得该音频提供的n项情绪参数能量 $q_1, q_2 \dots q_n$ ；

[0066] 在所述处理阶段,包括以下步骤:

[0067] E1:获取驱动虚拟人情绪表现的每项情绪的情绪参数EM,即 $EM=[em_1, em_2 \cdots em_n]$,其中 $em_1, em_2 \cdots em_n$ 为所述情绪参数EM中的n项情绪子参数;

[0068] E2:基于多个情绪子参数的当前值,确定每项情绪的情绪等级 $EL=[el_1, el_2 \cdots el_n]$,并调用对应该等级的情绪反应函数 $FEL()$,即 $FEL()=[Fel_1(), Fel_2() \cdots Fel_n()]$;

[0069] E3:将n项情绪参数能量作为激励,输入到对应的情绪反应函数中,获取情绪增量 ΔEM ;所述情绪增量 ΔEM 包括n项情绪子增量,即 $\Delta EM=[\Delta em_1, \Delta em_2 \cdots \Delta em_n]$,并且:

[0070] $\Delta em_i = Fel_i(q_i)$,式1;

[0071] 式1中, $i=1, 2, \cdots n$;

[0072] 在所述反应阶段,包括以下步骤:

[0073] R1:获取情绪模拟单元中的所述情绪增量 ΔEM 的阈值,阈值包括上限 $\max(\Delta em_i)$ 和下限 $\min(\Delta em_i)$;

[0074] R2:根据以下条件式,计算每一项所述情绪子增量 Δem_i 的实际效用能量 $\Delta ef-em_i$:

$$[0075] \quad \Delta ef-em_i = \begin{cases} \max(\Delta em_i), & \Delta em_i > \max(\Delta em_i) \\ \Delta em_i, & \min(\Delta em_i) < \Delta em_i \leq \max(\Delta em_i) \\ 0, & \Delta em_i \leq \min(\Delta em_i) \end{cases}$$

[0076] 上式表明,若一项所述情绪子增量 Δem_i 超过其上限 $\max(\Delta em_i)$,则抑制该项所述情绪子增量 Δem_i ;另一方面若所述情绪子增量 Δem_i 未达到一阈值 $\min(\Delta em_i)$,则忽略其作用;其中,上限 $\max(\Delta em_i)$ 和下限 $\min(\Delta em_i)$ 决定虚拟人的情绪承受能力,由技术人员根据虚拟人的情绪承受力设定。

[0077] R3:更新所述情绪参数EM,即使更新后的 $EM'=EM+\Delta EM$,从而由更新后的所述情绪参数 EM' 生成新的虚拟人最新的情绪表现;

[0078] 可选地,所述n项情绪分类的类别中至少包括中立情绪;并且分类的类别中还包括以下情绪的三项或三项以上:快乐、愤怒、恐惧、悲伤和厌恶。

[0079] 可选地,在采集阶段的步骤S2中,包括:

[0080] 读取音频信息;

[0081] 从音频信息中计算一组梅尔频率倒谱系数;

[0082] 从音频信息中确定多个声学变量;

[0083] 计算这些声学变量在音频信息中的平均值、标准差、最大值和最小值;

[0084] 可选地,在采集阶段中,包括采用基于机器学习的情绪分类模型进行情绪分类;所述情绪分类模型基于支持向量机内核而建立;

[0085] 可选地,所述生成方法还包括衰减阶段;当虚拟人与外部不存在任何语音交流时随即开始所述衰减阶段;在所述衰减阶段中的每个时间周期 Δt 中,所述情绪参数EM的每一项情绪子参数 em_i 基于以下衰减函数进行更新:

[0086] $em'_i = em_i \cdot e^{-c_i \cdot \Delta t}$,式2;

[0087] 上式中, e 为自然常数, c_i 为第i项情绪的衰减系数且 $c_i > 0$,由技术人员根据虚拟人

的性格决定； Δt 为单位时间周期，取值为0.1秒，0.5秒，1秒或者其他时间数值，由技术人员根据虚拟人的情绪稳定性决定；

[0088] 其中，虚拟人的当前情绪由所述情绪参数EM中n项所述情绪子参数驱动从而作出情绪表现；例如，当所述情绪子参数中代表快乐的子参数具有较高数值时，则驱动虚拟人产生快乐的情绪；当所述情绪子参数中代表中立的子参数与悲伤的子参数较高时，则可以驱动虚拟人具有悲伤情绪的冲动，而同时受自身中立情绪压制着悲伤的情绪；

[0089] 在一种实施方式中，如附图2所示的采集阶段，采集模块可以包括各类音频捕获设备，例如麦克风；音频信息由麦克风或其他音频捕获设备捕获的语音中获得的数据；在一些实施方式中，音频信息包括的无声部分被删除；通过截断音频信息并丢弃无助于情绪识别的数据，去除音频的无声部分可以提高系统的速度和效率；此外，从语音信号和滤波器中去除静音间隔，从而减少来自活动语音段的级联的失真；这在计算与语音相关的平均量时很有用，因为它消除了单词和音节之间的静默停顿，这在人与人之间可能会发生很大变化并影响性能计算；

[0090] 但相对地在一些实施方式中，无声部分同样可作为一种情绪的表达，例如等待或者无奈的表现；这取决于虚拟人系统本身的处理能力或者需要接收外界信息的细节程度来决定是否去除无声部分；

[0091] 进一步的，对采集后的音频信息进行模数转换以及重采样步骤，从而实现将采集到的原有模拟信号转化为具有目标采样频率的数字信号；

[0092] 进一步的，对音频信息提取计算语音特征；常见的一种算法为梅尔频率倒谱系数(MFCC)；MFCC系数是通过将信号与宽度随着频率增加而增加的三角形箱进行分箱来生成的；梅尔频率倒谱系数通常用于语音和情感分类；因此，本领域技术人员将理解可以使用许多计算系数的方法；

[0093] 在优选的实施方法中，使用15个韵律特征和45个语音特征，用于描述音频信息的基频和幅度；韵律特征在实时情感分类中很有用，通过使用话语能够准确地反映在话语音频中的情感状态；

[0094] 其中，使用基于次谐波与谐波比的音调确定算法来确定音频信息的基频的水平音高、音域范围和轮廓，以上声学变量均与情感特征密切相关，可以根据以上声学变量的特征用于分析音频所包含的发言人需要表现的情绪；例如，在快乐情绪下的说话音频的基频会随之增加，并且平均语音强度增加，音频的特征具有较高的可变性；而无聊或中立情绪下的与基频会随之降低，并且声学变量的标准差较小；

[0095] 与仅使用韵律特征相比，一起结合使用韵律特征和语音特征有助于实现更高的分类准确度；在一些实施方法中，特征统计更侧重于各项声学变量的统计参数，例如均值、标准差、最小值、最大值和范围；

[0096] 进一步的，特征被提取后继续用于对音频信息进行分类；

[0097] 在一种实施方法中，采集阶段的步骤S3如附图3所示，所述分析模块分为两个阶段为音频信息进行分类，在两阶段分类中，在步骤301中，音频信息输入到所述分析模块；在步骤302，数据被分类为类别1和类别2两种情绪；在优选的一种实施方法中，类别1包括中性和积极的情绪，例如快乐的情绪；类别2包括消极的情绪，例如愤怒和悲伤；并且在第二阶段的步骤303中，进一步将类别1中的中立或积极情绪作进一步的区分；同样地，在步骤304中，进

一步将类别2中的消极情绪作进一步的区分;优选地,以上步骤302、303、304中,均配置一个基于支持向量机的分类器,用于每一阶段的情绪类别的分类操作;

[0098] 出于分类的目的并且基于支持向量机(SVM)的特性,使用具有线性、二次和径向基函数内核的支持向量机来生成用于优化分类的超平面;根据虚拟人的特定应用,可以针对不同的内核使用不同的参数运行优化,并且为每个分类存储最佳性能模型及其参数,以便稍后与虚拟人一起使用;

[0099] 例如基于一种与年龄较小的用户进行互动的虚拟人,其情绪分类的模型及其参数则基于该年龄段的语言特点进行调教,且明显区别于成年人,或者专业性更强的使用场合;

[0100] 进一步的,需要训练对于情绪的分类模型以运行于支持向量机的分类操作;其中每个分类器必须运行分类模型,并且优选地可以基于所需要的分类任务,配置不同的分类模型;在一种训练方法中,对模型的训练集使用十折交叉验证方法,即将每种情绪对应的音频集合随机分组为10份等大小的数据组;轮流将其中9份作为训练数据,1份作为测试数据,进行试验;每次试验都会得出相应的正确率(或差错率);10次的结果的正确率(或差错率)的平均值作为对算法精度的估计,一般还需要进行多次十折交叉验证,例如进行5次或者10次的十折交叉验证,再求其均值,作为对算法准确性的估计;

[0101] 进一步的,基于以上方法,包括设定更多的分类器以及分类模型,对每种情绪进行情绪能量等级的划分;在一些实施例中,包括将除“中立”以外的具有明显特征的情绪,例如“快乐”或“愤怒”分为5个等级程度:非常高、比较高、一般、较为轻微、极轻微;而在使用更高运算能力的分类器中,可以将情绪能量等级划分为8级或者10级;

[0102] 进一步的,并且对对应不同数据的情绪能量等级进行数值化,以映射到不同的情绪参数能量;例如在非常高等级,其情绪参数能量为10;而在一般等级,其情绪参数能量为5;从而将用户的音频信息,转化为具有n项情绪参数能量的输入量,进入下一阶段,即所述处理阶段。

[0103] 实施例二:

[0104] 本实施例应当理解为至少包含前述任意一个实施例的全部特征,并在其基础上进一步改进;

[0105] 人类在不同情绪状态下,接受同一种外界刺激后,会产生不同的应激反应;例如,在比较快乐情绪下,接收一中等程度快乐的信息时,其表现不会出现太大改变;而在一中立情绪下,接收与上述相同的中等程度快乐的信息时,则其情绪相对地产生较大的起伏,从而表现出明显的快乐情绪;因此,为了进一步模拟人类在一定情绪基调下,在受到情绪刺激时产生具有差异化情绪反应,所述生成方法包括在所述处理阶段对虚拟人的情绪生成作进一步的细化计算;

[0106] 其中,在步骤E1中,需要获取虚拟人当前的所述情绪参数,亦即确定虚拟人当前的情绪状态;

[0107] 从而在步骤E2中,针对不同的情绪状态,可以由所述情绪驱动模块输出对应该情绪状态对于外界刺激的一个反应;

[0108] 如附图4所示,该函数曲线示意图包括了对应于第i项情绪中,三个不同的情绪等级下的所述情绪反应函数 $Fe1_i()$;其中第一曲线401对应于一较低的情绪等级 $e1_i=1$,第二曲线402对应于一中等程度的情绪等级 $e1_i=3$,第三曲线403对应于一较低的情绪等级 $e1_i=5$;

[0109] 可见,对于一个确定的情绪能量 q 的刺激,由于在不同的情绪等级下,其情绪反应函数具有不同的输出值,因此,将会获得不同的情绪子增量 Δem_i ;亦即获得不同的情绪变化的驱动;

[0110] 进一步的,所述情绪反应函数根据设计人员对虚拟人的情感特性、性格特性等属性进行设定;所述情绪反应函数在一些实施方式中,可以为单调函数,即随着情绪能量的增加,情绪子增量亦随之增加;然而在一些实施方式中,所述情绪反应函数可以为常量,即设定虚拟人对某项情绪作出的表现恒定,该设定方式适合于一些较为严肃性格的虚拟人,或者应用于一些特殊需要的用户或应用场景。

[0111] 实施例三:

[0112] 本实施例应当理解为至少包含前述任意一个实施例的全部特征,并在其基础上进一步改进;

[0113] 进一步的,在所述处理阶段以及所述衰减阶段,虚拟人的所述情绪参数随着各项情绪子增量的改变,从而情绪上出现相应的变化;

[0114] 其中,所述衰减函数用于表现虚拟人的情绪随时间的推移并出现平复;

[0115] 式2中, c_i 为第 i 项情绪的衰减系数,可选地,若设定虚拟人的某项情绪能够较快平复,则可以设定 c_i 为一个较大的值,令情绪参数的衰减随时间的推移更快,并且迅速地到达0;

[0116] 而若设定虚拟人的情绪更持续,例如具有较为感性的性格设定,则可以将 c_i 设为更接近于0的值,例如0.5,或者0.2;则如此,虚拟人的情绪可以保持的时间更长;

[0117] 进一步的,基于虚拟人的 n 项情绪子参数的每一项,可以设定为不同的 c_i 值,用于表现虚拟人在各项情绪上的特征,例如乐观性格的虚拟人,其快乐情绪相比悲伤应该维持得更久;而需要表现虚拟人为冲动个性的,可以令愤怒情绪的上限值相对较高,但是衰减速度亦相对较快,以表现其可以达到较大的愤怒程度同时恢复的时间更短;

[0118] 在上述实施例中,对各个实施例的描述都各有侧重,某个实施例中未详述或记载的部分,可以参见其它实施例的相关描述。

[0119] 虽然上面已经参考各种实施例描述了本发明,但是应当理解,在不脱离本发明的范围的情况下,可以进行许多改变和修改。也就是说上面讨论的方法,系统和设备是示例。各种配置可以适当地省略,替换或添加各种过程或组件。例如,在替代配置中,可以以与所描述的顺序不同的顺序执行方法,和/或可以添加,省略和/或组合各种部件。而且,关于某些配置描述的特征可以以各种其他配置组合,如可以以类似的方式组合配置的不同方面和元素。此外,随着技术发展其中的元素可以更新,即许多元素是示例,并不限制本公开或权利要求的范围。

[0120] 在说明书中给出了具体细节以提供对包括实现的示例性配置的透彻理解。然而,可以在没有这些具体细节的情况下实践配置例如,已经示出了众所周知的电路,过程,算法,结构和技术而没有不必要的细节,以避免模糊配置。该描述仅提供示例配置,并且不限制权利要求的范围,适用性或配置。相反,前面对配置的描述将为本领域技术人员提供用于实现所描述的技术的使能描述。在不脱离本公开的精神或范围的情况下,可以对元件的功能和布置进行各种改变。

[0121] 综上,其旨在上述详细描述被认为是例示性的而非限制性的,并且应当理解,以上

这些实施例应理解为仅用于说明本发明而不用于限制本发明的保护范围。在阅读了本发明的记载的内容之后,技术人员可以对本发明作各种改动或修改,这些等效变化和修饰同样落入本发明权利要求所限定的范围。

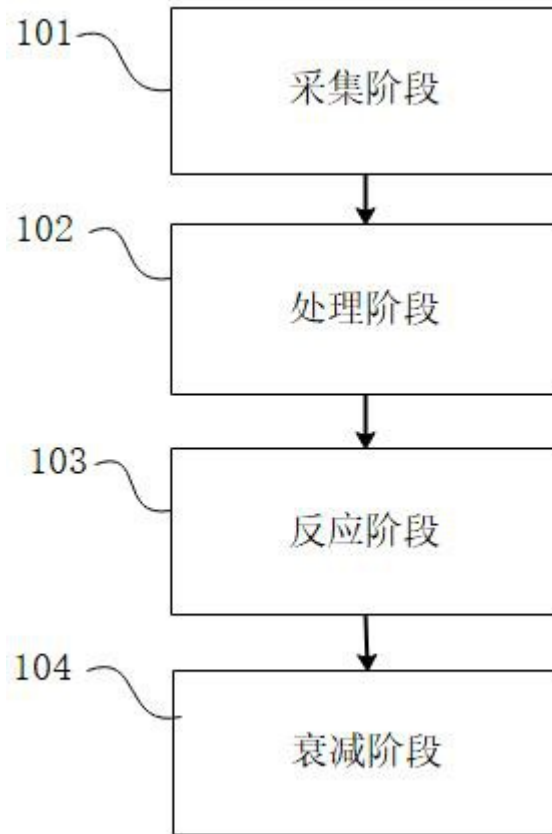


图1

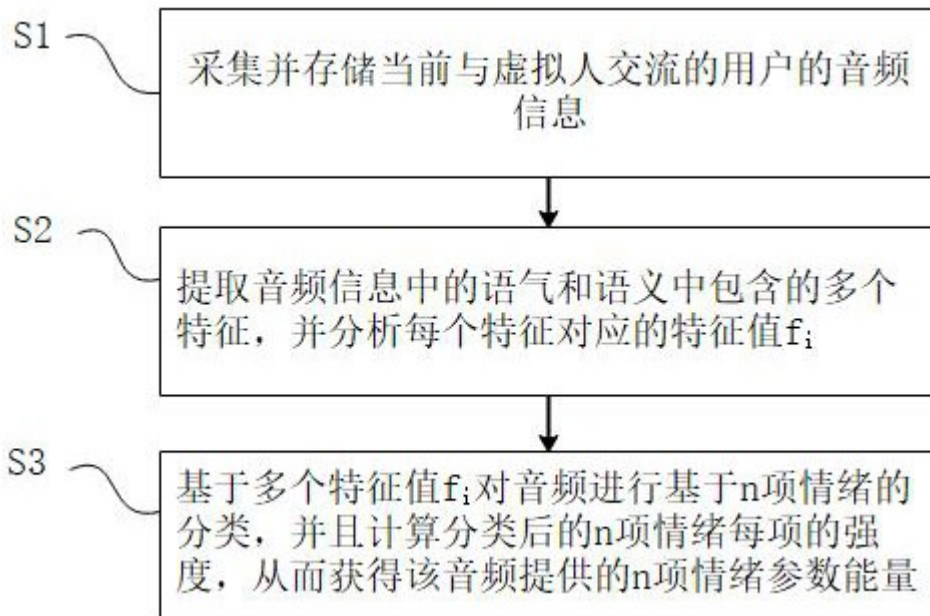


图2

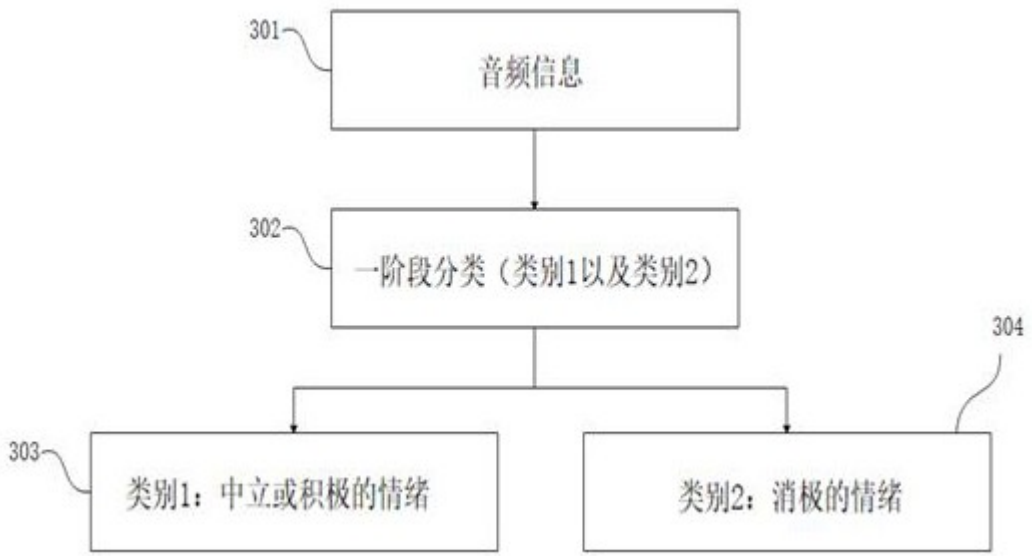


图3

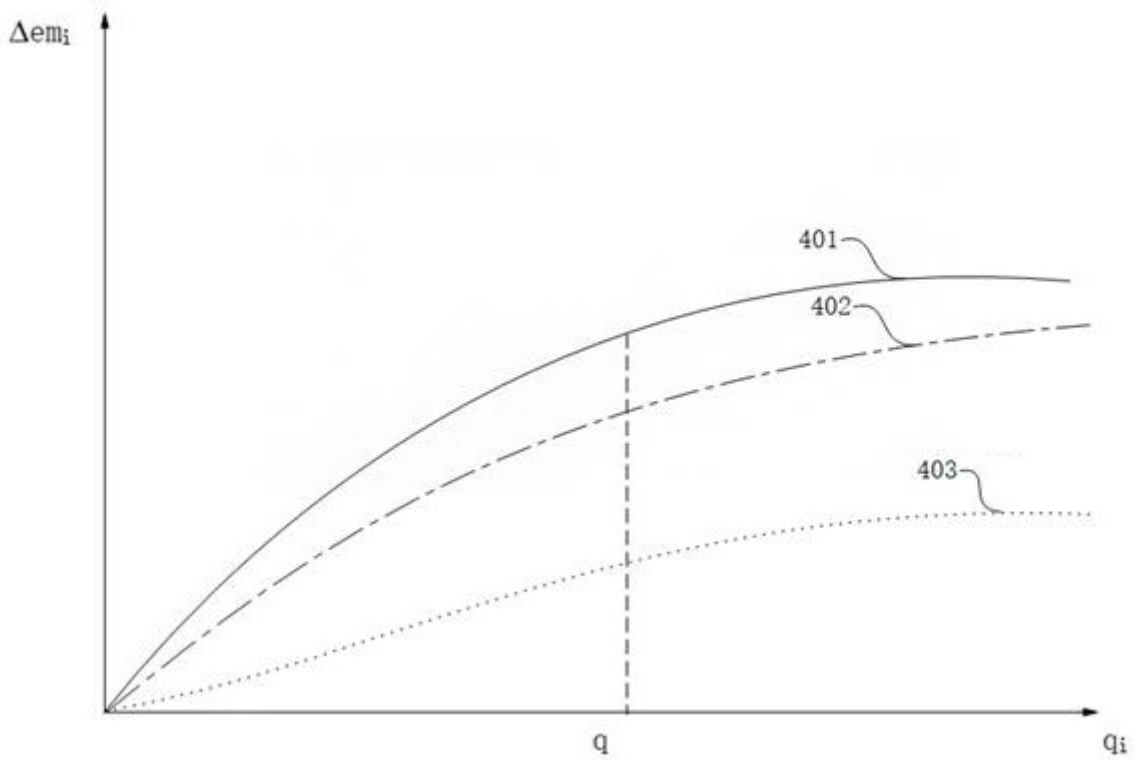


图4