



(12) 发明专利申请

(10) 申请公布号 CN 115494941 A

(43) 申请公布日 2022. 12. 20

(21) 申请号 202211006866.4

(22) 申请日 2022.08.22

(71) 申请人 同济大学

地址 200092 上海市杨浦区四平路1239号

(72) 发明人 刘富强 胡新宇 王俊元 李雷

贺茂恩

(74) 专利代理机构 上海德昭知识产权代理有限

公司 31204

专利代理师 陈龙梅

(51) Int. Cl.

G06F 3/01 (2006.01)

G06F 16/332 (2019.01)

G06N 3/08 (2006.01)

G06T 17/00 (2006.01)

G10L 13/02 (2013.01)

权利要求书3页 说明书16页 附图16页

(54) 发明名称

基于神经网络的元宇宙情感陪护虚拟人实现方法及系统

(57) 摘要

本发明提供一种基于神经网络的元宇宙情感陪护虚拟人实现方法及系统,陪护委托人构建客制化的元宇宙虚拟人及空间以托管与老人情感沟通的时间,提高了被陪护人在情感陪护过程中虚拟人对象的智能性,增强其在接受陪护时的真实感、亲切感与沉浸感;通过智能技术建立虚拟人聊天后端和学习系统,使虚拟人从形象到语音再到聊天内容特点都具有高度可客制化、个性化特性,解决了传统机器人情感陪护的机械死板且单一的问题,使机器情感陪护可以进行量身定制与随使用自适应学习改进;通过图像与建模技术建立虚拟人所处的场景,在场景营造、细节还原等方面逼近现实,进一步增强情感陪护过程的观感真实性,并与信息反馈机制结合,提高信息反馈效率与准确性。



1. 一种基于神经网络的元宇宙情感陪护虚拟人系统,其特征在于,包括:

虚拟人学习模块,基于陪护委托人提交的个人影像资料、多模态语料数据、指定领域的公共信息,学习生成元宇宙情感陪护虚拟人的客制化元宇宙视听形象及客制化语音聊天内容;

环境录入与配置模块,基于所述陪护委托人提供的环境影像资料,生成客制化元宇宙空间环境三维模型及环境细节;

情感沟通陪护模块,根据所述虚拟人学习模块、所述环境录入与配置模块的结果生成所述陪护委托人期望提供给被陪护人的元宇宙情感陪护虚拟人;以及

信息反馈模块,在陪护过程中为所述陪护委托人提供反馈信息,

其中,所述虚拟人学习模块包括:

个体形象客制化单元,基于所述个人影像资料,对该个人影像资料中的所述陪护委托人形象进行三维重建,并从该个人影像资料中提取各种状态下的神态表情、口型特征进行存储和学习,从而得到所述客制化元宇宙视听形象;以及

聊天内容客制化单元,基于所述多模态语料数据以及所述指定领域的公共信息,进行语义情感分析以及语句合成,从而得到所述客制化语音聊天内容。

2. 根据权利要求1所述的基于神经网络的元宇宙情感陪护虚拟人系统,其特征在于:

其中,所述个体形象客制化单元包括三维重建神经网络模型、语言与特征神经网络模型、姿态与神情生成神经网络模型以及口型生成神经网络模型,分别用于生成所述客制化元宇宙视听形象中的人物与环境三维模型、聊天时的声音、姿态与神情以及口型,

所述个体形象客制化单元还提供预览功能,并根据所述陪护委托人对预览的反馈对所述客制化元宇宙视听形象进行修正或重新学习,

所述聊天内容客制化单元包括语音聊天内容生成神经网络模型,所述多模态语料数据经语音识别和情感语义分析后,语义和情感识别结果交由所述语音聊天内容生成神经网络模型生成客制化的语音聊天内容,

所述聊天内容客制化单元还提供预览功能,并根据所述陪护委托人对预览的反馈对所述语音聊天内容进行修正或重新生成。

3. 根据权利要求2所述的基于神经网络的元宇宙情感陪护虚拟人系统,其特征在于:

其中,所述三维重建神经网络模型为神经网络的结构化堆叠,

所述语言与特征神经网络模型为训练好的迁移神经网络,所述迁移神经网络为经过预训练的具有泛化能力的可在模型应用过程中继续训练改善的神经网络的结构化堆叠,

所述姿态与神情生成神经网络模型为训练好的姿态学习神经网络,

所述口型生成神经网络模型为训练好的口型学习神经网络,

所述姿态学习神经网络和所述口型学习神经网络均为生成对抗网络,

所述语音聊天内容生成神经网络模型为神经网络的结构化堆叠。

4. 根据权利要求1所述的基于神经网络的元宇宙情感陪护虚拟人系统,其特征在于:

其中,所述环境录入与配置模块包括:

预设模型存储单元,预存有若干聊天背景环境的三维模型以及若干环境细节效果;

影像资料提交接口,供所述陪护委托人提交希望录入的环境的照片或全景照片;

环境模型生成单元,用于生成环境三维模型,提供预设方式以及客制化方式,在预设方

式下,所述陪护委托人从所述预设模型存储单元中选择预存的所述三维模型作为所述环境三维模型,在客制化方式下,基于所述照片或所述全景照片生成所述环境三维模型;

环境细节配置单元,用于配置环境细节,提供预设方式以及客制化方式,在预设方式下,所述陪护委托人从所述预设模型存储单元中选择预存的所述环境细节效果作为所述环境细节,在客制化方式下,基于所述陪护委托人输入的地理位置信息和/或环境时刻配置作为环境生成依据。

5. 根据权利要求1所述的基于神经网络的元宇宙情感陪护虚拟人系统,其特征在于:

其中,在所述陪护委托人有条件提供三维数据时,所述虚拟人学习模块还接受所述三维数据,并以所述三维数据为依据对客制化的虚拟人三维模型进行修正,

所述环境录入与配置模块还接受所述三维数据,并以所述三维数据为依据对客制化的环境三维模型进行修正,

所述三维数据为三维点云数据、景深数据中的任意一种或两种的组合。

6. 根据权利要求1所述的基于神经网络的元宇宙情感陪护虚拟人系统,其特征在于:

其中,所述情感沟通陪护模块包括:

自然语言处理与生成单元,在陪护过程中监听所述被陪护人的说话行为,并对监听到的被陪护人语言进行语音识别和情感语义分析,获得所述被陪护人的主要意图,并根据所述主要意图完成相应的处理动作;以及

可视化单元,根据所述虚拟人学习模块以及所述环境录入与配置模块的结果生成客制化的元宇宙虚拟人个体半身形象及其所处的元宇宙环境,并进一步根据所述自然语言处理与生成单元的结果生成包含所述元宇宙情感陪护虚拟人个体的情感陪护元宇宙空间。

7. 根据权利要求6所述的基于神经网络的元宇宙情感陪护虚拟人系统,其特征在于:

其中,所述自然语言处理与生成单元包括:

自然语言处理神经网络模型,用于进行所述情感语义分析,提取出语言信息中的语义要点与所述主要意图;以及

内容生成部分,

其中,所述内容生成部分包括:

内容生成神经网络模型或第三方内容生成服务,基于所述语义要点与所述主要意图生成文本回复内容;以及

从文本到语音神经网络模型,基于所述文本回复内容以及所述虚拟人学习模块的配置,生成所述元宇宙情感陪护虚拟人的声音音频流,

所述可视化单元生成所述元宇宙情感陪护虚拟人在所述情感陪护元宇宙中的三维半身像视频流。

8. 根据权利要求6所述的基于神经网络的元宇宙情感陪护虚拟人系统,其特征在于:

其中,所述被陪护人的所述主要意图分为信息服务需求与情感沟通需求,

当所述主要意图为所述信息服务需求时,所述情感沟通陪护模块通过互联网进行对应的信息处理以满足该需求;

当所述主要意图为所述情感沟通需求时,所述情感沟通陪护模块根据所述虚拟人学习模块的配置,生成符合语境的个性化回复内容,并根据所述虚拟人学习模块的配置以及生成的所述个性化回复内容进一步确定情感和语调,以特定声纹、音色、方言声音特征进行合

成,得到合成语音,并将所述合成语音传递至所述可视化单元进行协同输出。

9. 根据权利要求6所述的基于神经网络的元宇宙情感陪护虚拟人系统,其特征在于:
其中,所述信息反馈模块包括:

智能评价单元,为以自然语言处理神经网络为基础构建的评分网络,用于在情感陪护过程中提取语义信息,并对所述语义信息的价值与重要程度进行实时评估;

资料暂存单元,根据所述实时评估结果选择性地存储所述语义信息,并配套存储提取对应的语义信息时所利用到的音频数据、该音频数据对应时间片的视频数据;以及

人机交互接口,供所述陪护委托人查看存储的所述语义信息、对应的所述音频数据以及所述视频数据。

10. 一种采用如权利要求1-9中任意一项所述的基于神经网络的元宇宙情感陪护虚拟人系统实现情感陪护的实现方法,其特征在于,包括以下步骤:

步骤S1,虚拟人学习模块采集陪护委托人的多模态数据,基于多模态数据生成元宇宙情感陪护虚拟人的客制化元宇宙视听形象,显示预览效果,并循环地根据所述陪护委托人对该预览效果的反馈调整配置或重复训练以改善学习效果,最终得到客制化元宇宙视听形象;

步骤S2,虚拟人学习模块获取所述陪护委托人录入的多模态形式语料与选择包含的公开信息领域,基于所述多模态形式语料与所述包含的公开信息领域生成语音聊天内容,并根据所述陪护委托人的试用反馈增进录入训练以提高客制化效果,最终得到客制化语音聊天内容;

步骤S3,环境录入与配置模块显示虚拟环境配置选项,采集所述陪护委托人的环境图像与三维信息,基于所述环境图像与所述三维信息生成环境三维模型和环境细节,并显示预览效果,根据所述陪护委托人对该预览效果的反馈调整所述环境三维模型和所述环境细节,得到客制化元宇宙空间环境三维模型及环境细节;

步骤S4,情感沟通陪护模块基于所述客制化元宇宙视听形象、所述客制化语音聊天内容、所述客制化元宇宙空间环境三维模型及环境细节合成元宇宙情感陪护虚拟人的音视频流,基于被陪护人的语音,对所述被陪护人的情感语义与需求进行智能化分析,根据智能化分析结果满足所述被陪护人的服务需求与情感沟通需求,并提供所述陪护委托人实时接入、与所述被陪护人即时通讯的功能;

步骤S5,信息反馈模块基于情感语义分析,存储情感陪护过程中智能提取的所述被陪护人的情感语义与需求要点,并配套存储对应时间段的音视频数据,响应于所述陪护委托人开始查看所述被陪护人近期陪护情况的操作,向所述陪护委托人提供反馈信息。

基于神经网络的元宇宙情感陪护虚拟人实现方法及系统

技术领域

[0001] 本发明属于元宇宙与人工智能技术领域,具体涉及一种基于神经网络的元宇宙情感陪护虚拟人实现方法及系统。

背景技术

[0002] 随着经济社会的发展,人口快速老龄化导致的老年群体的各种需求在迅速增加。其中,老人情感需求由于目前主要依赖于人的陪伴来解决,因受到现代社会生活节奏等的限制而很难得到切实的满足。目前,市场上存在一些简单的利用电子信息技术满足老人基本精神需求的机器人,例如中国专利公开文献CN112223318A公开的一种情感陪护智能机器人,通过人形机器以交互的形式为老人提供基本陪护服务,使得老人可以通过机器人获得查询信息、健康管理、收听/看节目、与亲友聊天等服务,而无需子女或护工等的辅助,实现基本的精神陪护功能。

[0003] 但是,现有的情感陪护机器人仍存在以下弊端:第一、情感陪护机器人提供的服务核心仍与传统老年医护产品相似,仅能提供一些机械的、单一的功能的集成,其情感陪护仍需要人的深度参与,与人亲自进行情感陪护本质上并无二异,无法实现相对更智能化、自助化的情感陪护服务,更无进行陪护个性化的可能。第二、情感陪护机器人以实体形式存在,制造者和购买方都需要付出大量的成本投入到硬件设备中;且日常使用会受到硬件形式的限制,老人需要花费大量精力学习陪护机器人的交互方法,产品使用寿命也会受到硬件寿命的限制。

[0004] 近年来,人工智能技术迅速发展,自然语言处理NLP网络的相关研究取得重大突破,人工语音合成在发音音调、音色、流畅性与准确性等多个方面都能做到基本与人类发音无异,且进一步地还可以在情感、方言发音、声纹等方面做到差异化的模仿与学习模拟;此外,元宇宙概念的兴起为传统互联与智能体验增加了空间性维度,它将赋予用户时空拓展层面上的全新体验、价值,为用户创造沉浸式、交互式、更多感官维度的体验将是元宇宙的技术主脉络,在人工智能技术的加持下,元宇宙在场景营造、细节还原等方面的性能将得到更进一步的提升。因此,可基于元宇宙及人工智能技术来对现有的情感陪护机器人、陪护方法进行改进。

发明内容

[0005] 本发明是为解决上述问题而进行的,目的在于提供一种能够实现更智能化、自助化情感陪护的情感陪护虚拟人系统及采用该系统实现情感陪护的方法,本发明采用了如下技术方案:

[0006] 本发明提供了一种基于神经网络的元宇宙情感陪护虚拟人系统,其特征在于,包括:虚拟人学习模块,基于陪护委托人提交的个人影像资料、多模态语料数据、指定领域的公共信息,学习生成元宇宙情感陪护虚拟人的客制化元宇宙视听形象及客制化语音聊天内容;环境录入与配置模块,基于所述陪护委托人提供的环境影像资料,生成客制化元宇宙空

间环境三维模型及环境细节;情感沟通陪护模块,根据所述虚拟人学习模块、所述环境录入与配置模块的结果生成所述陪护委托人期望提供给被陪护人的元宇宙情感陪护虚拟人;以及信息反馈模块,在陪护过程中为所述陪护委托人提供反馈信息,其中,所述虚拟人学习模块包括:个体形象客制化单元,基于所述个人影像资料,对该个人影像资料中的所述陪护委托人形象进行三维重建,并从该个人影像资料中提取各种状态下的神态表情、口型特征进行存储和学习,从而得到所述客制化元宇宙视听形象;以及聊天内容客制化单元,基于所述多模态语料数据以及所述指定领域的公共信息,进行语义情感分析以及语句合成,从而得到所述客制化语音聊天内容。

[0007] 本发明提供的基于神经网络的元宇宙情感陪护虚拟人系统,还可以具有这样的技术特征,其中,所述个体形象客制化单元包括三维重建神经网络模型、语言与特征神经网络模型、姿态与神情生成神经网络模型以及口型生成神经网络模型,分别用于生成所述客制化元宇宙视听形象中的人物与环境三维模型、聊天时的声音、姿态与神情以及口型,所述个体形象客制化单元还提供预览功能,并根据所述陪护委托人对预览的反馈对所述视听形象的生成模型进行修正或重新学习,所述聊天内容客制化单元包括语音聊天内容生成神经网络模型,所述多模态语料数据经语音识别和情感语义分析后,语义和情感识别结果交由所述语音聊天内容生成神经网络模型生成客制化的语音聊天内容,所述聊天内容客制化单元还提供预览功能,并根据所述陪护委托人对预览的反馈对所述语音聊天内容的生成模型进行修正或重新生成。

[0008] 本发明提供的基于神经网络的元宇宙情感陪护虚拟人系统,还可以具有这样的技术特征,其中,所述三维重建神经网络模型为神经网络的结构化堆叠经训练后的结果模型,所述语言与特征神经网络模型为训练好的迁移神经网络的结果模型,所述迁移神经网络为经过预训练的具有泛化能力的可在模型应用过程中继续训练改善的神经网络的结构化堆叠,所述姿态与神情生成神经网络模型为训练好的姿态学习神经网络的结果模型,所述口型生成神经网络模型为训练好的口型学习神经网络的结果模型,所述姿态学习神经网络和所述口型学习神经网络均为生成对抗网络或其它改进的无监督学习方法中的一种或几种的结构化堆叠,所述语音聊天内容生成神经网络模型为神经网络的结构化堆叠经训练后的结果模型。

[0009] 本发明提供的基于神经网络的元宇宙情感陪护虚拟人系统,还可以具有这样的技术特征,其中,所述环境录入与配置模块包括:预设模型存储单元,预存有若干聊天背景环境的三维模型以及若干环境细节效果;影像资料提交接口,供所述陪护委托人提交希望录入的环境的照片或全景照片或视频影像;环境模型生成单元,用于生成环境三维模型,提供预设方式以及客制化方式,在预设方式下,所述陪护委托人从所述预设模型存储单元中选择预存的所述三维模型作为所述环境三维模型,在客制化方式下,基于所述照片或所述全景照片生成所述环境三维模型;环境细节配置单元,用于配置环境细节,提供预设方式以及客制化方式,在预设方式下,所述陪护委托人从所述预设模型存储单元中选择预存的所述环境细节效果作为所述环境细节,在客制化方式下,基于所述陪护委托人输入的地理位置信息和/或环境时刻配置作为环境生成依据。

[0010] 本发明提供的基于神经网络的元宇宙情感陪护虚拟人系统,还可以具有这样的技术特征,其中,在所述陪护委托人有条件提供三维数据时,所述虚拟人学习模块还接受所述

三维数据,并以所述三维数据为依据对定制化的虚拟人三维模型进行修正,所述环境录入与配置模块还接受所述三维数据,并以所述三维数据为依据对定制化的环境三维模型进行修正,所述三维数据为三维点云数据、景深数据中的任意一种或两种的组合。

[0011] 本发明提供的基于神经网络的元宇宙情感陪护虚拟人系统,还可以具有这样的技术特征,其中,所述情感沟通陪护模块包括:自然语言处理与生成单元,在陪护过程中监听所述被陪护人的说话行为,并对监听到的被陪护人语言进行语音识别和情感语义分析,获得所述被陪护人的主要意图,并根据所述主要意图完成相应的处理动作;以及可视化单元,根据所述虚拟人学习模块以及所述环境录入与配置模块的结果生成定制化的元宇宙虚拟人个体半身形象及其所处的元宇宙环境,并进一步根据所述自然语言处理与生成单元的结果生成包含所述元宇宙情感陪护虚拟人个体的情感陪护元宇宙空间。

[0012] 本发明提供的基于神经网络的元宇宙情感陪护虚拟人系统,还可以具有这样的技术特征,其中,所述自然语言处理与生成单元包括:自然语言处理神经网络模型,用于进行所述情感语义分析,提取出语言信息中的语义要点与所述主要意图;以及内容生成部分,其中,所述内容生成部分包括:内容生成神经网络模型或第三方内容生成服务,基于所述语义要点与所述主要意图生成文本回复内容;以及从文本到语音神经网络模型,基于所述文本回复内容以及所述虚拟人学习模块的配置,生成所述元宇宙情感陪护虚拟人的声音音频流,所述可视化单元生成所述元宇宙情感陪护虚拟人在所述情感陪护元宇宙中的三维半身像视频流。

[0013] 本发明提供的基于神经网络的元宇宙情感陪护虚拟人系统,还可以具有这样的技术特征,其中,所述被陪护人的所述主要意图分为信息服务需求与情感沟通需求,当所述主要意图为所述信息服务需求时,所述情感沟通陪护模块通过互联网进行对应的信息处理以满足该需求;当所述主要意图为所述情感沟通需求时,所述情感沟通陪护模块根据所述虚拟人学习模块的配置,生成符合语境的个性化回复内容,并根据所述虚拟人学习模块的配置以及生成的所述个性化回复内容进一步确定情感和语调,以特定声纹、音色、方言声音特征进行合成,得到合成语音,并将所述合成语音传递至所述可视化单元进行协同输出。

[0014] 本发明提供的基于神经网络的元宇宙情感陪护虚拟人系统,还可以具有这样的技术特征,其中,所述信息反馈模块包括:智能评价单元,为以自然语言处理神经网络为基础构建的评分网络,用于在情感陪护过程中提取语义信息,并对所述语义信息的价值与重要程度进行实时评估;资料暂存单元,根据所述实时评估结果选择性地存储所述语义信息,并配套存储提取对应的语义信息时所利用到的音频数据、该音频数据对应时间片的视频数据;以及人机交互接口,供所述陪护委托人查看存储的所述语义信息、对应的所述音频数据以及所述视频数据。

[0015] 本发明提供了一种采用上述基于神经网络的元宇宙情感陪护虚拟人系统实现情感陪护的实现方法,其特征在于,包括以下步骤:

[0016] 步骤S1,虚拟人学习模块采集陪护委托人的多模态数据,基于多模态数据生成元宇宙情感陪护虚拟人的视听形象,显示预览效果,并循环地根据所述陪护委托人对该预览效果的反馈调整配置或重复训练以改善学习效果,最终得到定制化视听形象;

[0017] 步骤S2,虚拟人学习模块获取所述陪护委托人录入的多模态形式语料与选择包含的公开信息领域,基于所述多模态形式语料与所述包含的公开信息领域生成语音聊天内

容,并根据所述陪护委托人的试用反馈增进录入训练以提高客制化效果,最终得到客制化语音聊天内容;

[0018] 步骤S3,环境录入与配置模块显示虚拟环境配置选项,采集所述陪护委托人的环境图像与三维信息,基于所述环境图像与所述三维信息生成环境三维模型和环境细节,并显示预览效果,根据所述陪护委托人对该预览效果的反馈调整所述环境三维模型和所述环境细节,得到客制化元宇宙空间环境三维模型及环境细节;

[0019] 步骤S4,情感沟通陪护模块基于所述客制化视听形象、所述客制化语音聊天内容、所述客制化元宇宙空间环境三维模型及环境细节合成元宇宙情感陪护虚拟人的音视频流,基于被陪护人的语音,对所述被陪护人的情感语义与需求进行智能化分析,根据智能化分析结果满足所述被陪护人的服务需求与情感沟通需求,并提供所述陪护委托人实时接入、与所述被陪护人即时通讯的功能;

[0020] 步骤S5,信息反馈模块基于情感语义分析,存储情感陪护过程中智能提取的所述被陪护人的情感语义与需求要点,并配套存储对应时间段的音视频数据,响应于所述陪护委托人开始查看所述被陪护人近期陪护情况的操作,向所述陪护委托人提供反馈信息。

[0021] 发明作用与效果

[0022] 根据本发明的基于神经网络的元宇宙情感陪护虚拟人实现方法及系统,采用元宇宙方式以神经网络技术为基础建立起用于老年人情感陪护的虚拟人系统,陪护委托人和被陪护人可以采用移动终端、XR等任一种方式在系统中完成虚拟人的配置和情感陪护过程,陪护委托人构建客制化的元宇宙虚拟人以托管与老人情感沟通的时间,被陪护人可以与陪护委托人的元宇宙虚拟人角色进行趋近于真实的沟通交流,弥补了陪护委托人由于工作而未能足量投入的在陪护老人方面的精力,同时提高了被陪护人在情感陪护过程中虚拟人对象的智能性,增强被陪护人在接受陪护时的真实感、亲切感与沉浸感,以理解被陪护人感情与解决感性需求和实用需求作为情感陪护沟通的首要目标。通过智能技术建立的虚拟人聊天后端和学习系统,使虚拟人从形象到语音再到聊天内容特点都具有高度可客制化、个性化特性,并着重强调聊天的情感关怀特性,解决了传统机器人情感陪护的机械死板且单一的问题,使机器情感陪护可以进行量身定制与随使用自适应学习改进;通过图像与建模技术建立虚拟人所处的场景,在场景营造、细节还原等方面逼近现实,进一步增强情感陪护过程的观感真实性,并与信息反馈机制结合,提高信息反馈效率与准确性。

附图说明

[0023] 图1是本发明实施例一中元宇宙情感陪护虚拟人系统的应用场景框图;

[0024] 图2是本发明实施例一中元宇宙情感陪护虚拟人系统的结构框图;

[0025] 图3是本发明实施例一中元宇宙情感陪护虚拟人个体形象客制化的原理图1;

[0026] 图4是本发明实施例一中元宇宙情感陪护虚拟人个体形象客制化的原理图2;

[0027] 图5是本发明实施例一中元宇宙情感陪护虚拟人系统的交互示意图1;

[0028] 图6是本发明实施例一中元宇宙情感陪护虚拟人系统的交互示意图2;

[0029] 图7是本发明实施例一中元宇宙情感陪护虚拟人聊天内容客制化的原理图;

[0030] 图8是本发明实施例一中元宇宙情感陪护虚拟人系统的交互示意图3;

[0031] 图9是本发明实施例一中元宇宙情感陪护虚拟人系统的交互示意图4;

- [0032] 图10是本发明实施例一中元宇宙情感陪护虚拟人系统的交互示意图5；
- [0033] 图11是本发明实施例一中元宇宙情感陪护虚拟人系统的交互示意图6；
- [0034] 图12是本发明实施例一中元宇宙情感陪护虚拟人系统的交互示意图7；
- [0035] 图13是本发明实施例一中元宇宙情感陪护虚拟人陪护过程的原理图；
- [0036] 图14是本发明实施例一中陪护委托人与被陪护人实时交流的原理图；
- [0037] 图15是本发明实施例一中元宇宙情感陪护虚拟人系统的交互示意图8；
- [0038] 图16是本发明实施例一中元宇宙情感陪护虚拟人系统的交互示意图9；
- [0039] 图17是本发明实施例一中元宇宙情感陪护虚拟人系统的交互示意图10；
- [0040] 图18是本发明实施例一中元宇宙情感陪护虚拟人系统的交互示意图11；
- [0041] 图19是本发明实施例一中元宇宙情感陪护实现方法的流程图；
- [0042] 图20是本发明实施例二中元宇宙情感陪护虚拟人系统的应用场景框图。

具体实施方式

[0043] 为了使本发明实现的技术手段、创作特征、达成目的与功效易于明白了解，以下结合实施例及附图对本发明的基于神经网络的元宇宙情感陪护虚拟人实现方法及系统作具体阐述。

[0044] <实施例一>

[0045] 图1是本实施例中元宇宙情感陪护虚拟人系统的场景示意图。

[0046] 如图1所示，基于神经网络的元宇宙情感陪护虚拟人系统（以下简称虚拟人系统）集成在服务器10中，服务器10与多个终端通信连接并且接入互联网，通信连接是指通过有线或无线通信方式进行直接或间接地连接，在此不做限制。终端的数量至少为两台，陪护委托人与被陪护人的总数量与终端的数量一致。

[0047] 具体地，终端根据用户类型不同可分为两类，委托人终端21及被陪护人终端22，分别由陪护委托人及被陪护人持有。陪护委托人即需要委托、配置情感陪护虚拟人的用户；被陪护人即接受情感陪护过程的用户。根据用户类型，委托人终端21显示虚拟人配置功能界面，被陪护人终端22则显示情感陪护功能界面。

[0048] 服务器10可以是独立的物理服务器，也可以是多个物理服务器构成的服务器集群或者分布式系统，还可以是提供云计算服务的云服务器。委托人终端21为由陪护委托人持有的终端，被陪护人终端22为由被陪护人持有的终端，均可以是具有相机、麦克风及扬声器的智能手机、平板电脑、笔记本电脑、台式计算机、智能音箱、智能手表、智能电视等，但不局限于此。

[0049] 图2是本实施例中元宇宙情感陪护虚拟人系统的结构框图，由于该系统集成在服务器10中，因此在本实施例中，该图也即服务器10的结构框图。

[0050] 如图2所示，虚拟人系统（服务器10）包括用户信息管理模块11、情感沟通陪护模块12、信息反馈模块13、虚拟人学习模块14、环境录入与配置模块15以及控制模块16。

[0051] 其中，委托人终端21、被陪护人终端22两类终端均需首先在服务器10中注册，用户信息管理模块11用于对用户的注册信息进行存储和管理。作为示例，注册信息可以包括必需的用于确认终端用户唯一性的姓名、手机号、电子邮箱等等。陪护委托人与被陪护人之间通过唯一确定性信息，进行多对多关系的绑定，终端用户根据不同的绑定对象进行差异化

的功能使用。不同的终端显示的配置信息、合成音视频流等所需的素材数据可以由服务器10发送,服务器10可以根据存储的用户注册信息确定每一终端对应的用户及其类型,并向用户对应的终端传输对应类型的素材数据。

[0052] 此外,用户信息管理模块11还提供陪护委托人和被陪护人的双向互选、绑定及配置功能。陪护委托人和被陪护人用户间通过唯一身份识别信息实现互相的绑定,陪护委托人可绑定多个不同的被陪护人,同样地,被陪护人也可绑定多个不同的陪护委托人。进一步地,用户可根据唯一身份识别信息搜索并选中目标用户,发送绑定申请,针对收到的绑定申请,用户可以选择同意以建立绑定,或选择拒绝以避免骚扰。元宇宙情感陪护虚拟人的配置与元宇宙情感陪护沟通过程仅针对已绑定的用户进行。陪护委托人也可以针对不同的被陪护人指定不同的元宇宙情感陪护虚拟人状态的配置。

[0053] 图13是本实施例中元宇宙情感陪护虚拟人陪护过程的原理图。

[0054] 如图13所示,情感沟通陪护模块12根据虚拟人学习模块14、环境录入与配置模块14的结果生成陪护委托人期望使被陪护人看到的元宇宙情感陪护虚拟人状态,包括但不限于虚拟人的形象、使用的声音、方言、虚拟人所处环境细节、所处场景、天气情况、环境氛围等特点。情感沟通陪护模块12以其生成的元宇宙虚拟人角色及元宇宙环境为形式的视听交互为主要交互方式。

[0055] 情感沟通陪护模块12包括自然语言处理与生成单元121以及可视化单元122。

[0056] 自然语言处理与生成单元121监听被陪护人的说话行为,并对监听到的被陪护人语言进行语音识别和情感语义分析,获得被陪护人的主要意图,并根据主要意图完成相应的处理动作。主要意图(即被陪护人的需求)可以被分为信息服务需求与情感沟通聊天需求两类。若被陪护人有诸如查询公共信息、订餐订票等信息服务需求(功能性需求),情感沟通陪护模块12即通过互联网进行对应的信息处理以满足需求,并生成相应的人性化回复内容。若被陪护人意图进行沟通,则情感沟通陪护模块12将根据虚拟人学习模块14的配置,生成符合语境的个性化回复内容。完成回复内容的生成后,自然语言处理与生成单元121将按照虚拟人学习模块14的配置,根据生成的回复内容进一步确定情感和语调,以特定声纹、音色、方言声音特征进行合成,并将合成语音传递至可视化单元122进行协同输出。

[0057] 本实施例中,自然语言处理与生成单元121包括自然语言处理(NLP,Natural Language Processing)神经网络模型以及内容生成部分。其中自然语言处理神经网络的结构可以是经过针对性构建的神经网络堆叠,用于提取语言信息语义要点与被陪护人需求。内容生成部分可以包括训练好的内容生成神经网络模型以及从文本到语音(TTS,Text To Speech)网络模型,也可以是第三方的内容生成服务。

[0058] 需要说明的是,信息服务处理结果与信息详情的语句合成所传递至的内容合成部分与情感沟通需求语义要点所传递至的内容合成部分结构并不一致,同时语义要点也传递至信息反馈模块进行进一步处理;完成聊天回复内容合成后,内容一方面被传递至从文本到语音(TTS,Text To Speech)网络模型,其中文本到语音网络模型是以虚拟人配置数据库中经客制化的声音特征模型为基础的特定结构神经网络,生成元宇宙情感陪护虚拟人的声音音频流。另一方面,内容被传递至形象细节生成部分(即虚拟人学习模块14)进行相应形象细节的生成。

[0059] 可视化单元122根据虚拟人学习模块14以及环境录入与配置模块15的配置生成客

制化的元宇宙虚拟人个体半身形象及其所处的元宇宙环境。进一步地,可视化单元122以元宇宙虚拟人个体半身形象为基础,根据自然语言处理与生成单元121传递的合成语音及情感特点,生成包含元宇宙情感陪护虚拟人个体的情感陪护元宇宙空间。其中,元宇宙情感陪护虚拟人个体的面部表情根据对应时刻的合成语音的情感特征生成,口型由对应时刻合成语音的发音决定,肢体动作根据语义特点适当生成。进一步地,虚拟人所处的包括天气、光线氛围在内的元宇宙环境由环境录入与配置模块15决定。最终,可视化单元122输出虚拟人正面形象,背景视角由陪护委托人指定或在情感陪护过程中由被陪护人指定,以指定视角生成元宇宙虚拟人在情感陪护元宇宙中的3D半身像视频流。

[0060] 图14是本实施例中陪护委托人与被陪护人实时交流的原理图。

[0061] 如图14所示,可视化单元122还可以提供陪护委托人与被陪护人之间的实时交谈。陪护委托人可以选择预定即时通话,也可以选择在被陪护人进行情感陪护的过程中途接入(实时交谈)。其中,选择预定即时通话时,被陪护人终端22将收到对应的请求信息,以供被陪护人进行回应;在实时交谈模式下,被陪护人看到的虚拟人形象仍以客制化的虚拟人半身形象为基础,面部表情、肢体动作、口型等细节则通过委托人终端21的相机捕获的视频流资料进行原样重建。进一步地,在实时交谈模式下,陪护委托人可以看到被陪护人终端22的相机捕获的实时影像画面。陪护委托人也可以选择向被陪护人终端22直接输出委托人终端21相机捕获的实时影像,如图17所示。

[0062] 信息反馈模块13以自然语言处理与生成单元121为基础,在情感陪护过程中提取并记录被陪护人希望传达到对方(即陪护委托人)的信息。信息反馈模块13包括人机交互接口131、资料暂存单元132以及智能评价单元133。

[0063] 陪护委托人可以借助人机交互接口131查看情感陪护虚拟人在近期情感陪护过程中所提取、记录的被陪护人的近况、需求和期望等信息,如图16所示。进一步地,信息反馈模块13还将保留近若干天的信息提取依据材料(即被陪护人的音视频资料),存储在资料暂存单元132中,以备陪护委托人查询,以便陪护委托人更全面地了解被陪护人的身心状态。进一步地,智能评价单元133在情感陪护过程中对语义信息价值与重要程度进行实时评估,根据实时评估结果决定存储的重点与信息要点。

[0064] 本实施例中,智能评价单元133是以自然语言处理神经网络为基础构建的评分网络,根据价值评估得分着重存储被陪护人希望传达的信息要点,同时配套存储的还包括提取要点时所利用到的音频数据、音频数据对应时间片的视频数据。在陪护过程较短或存储空间富余的情况下,也可以完整地存储整个陪护过程的音视频数据。

[0065] 虚拟人学习模块14提供用户交互接口,陪护委托人可通过其委托人终端21,经用户交互接口提交语料、语音片段、音频资料及影像资料。在此基础上,虚拟人学习模块14进行学习训练与存储输出。虚拟人学习模块14包括个体形象客制化单元141、聊天内容客制化单元142。

[0066] 图3、图4分别是本实施例中元宇宙情感陪护虚拟人个体形象客制化的原理图1、2,分别示出了虚拟人的三维模型客制化、声音、语言、姿态等客制化的原理。

[0067] 如图3-4所示,个体形象客制化单元141用于提供元宇宙虚拟人个体形象客制化功能,通过陪护委托人提交的影像资料,对影像资料中的陪护委托人形象进行三维重建,以陪护委托人形象复刻为主体完成虚拟人半身3D模型的建立,以最大程度还原陪护委托人的形

象特征,包括脸型、五官肤色、面貌特征等在内。进一步地,虚拟人学习模块14可以从陪护委托人提交的影像资料中提取各种状态下的神态表情并进行存储和学习,同时可以提取陪护委托人说话时的口型特征,以辅助提高可视化系统中口型生成的精确度。

[0068] 本实施例中,个体形象客制化单元141包括三维重建神经网络模型、语言与特征神经网络模型、姿态与神情生成神经网络模型以及口型生成神经网络模型。其中,三维重建神经网络模型可以是神经网络的结构化堆叠经训练后的结果模型,用于生成人像三维模型。语言与特征神经网络模型为训练好的迁移神经网络,迁移神经网络可以是经过预训练的具有泛化能力的可在模型应用过程中继续训练改善的神经网络结构化堆叠,将采集得到的音频数据作为训练样本输入到迁移学习神经网络进行训练。姿态与神情生成神经网络模型为训练好的姿态学习神经网络的结果模型,口型生成神经网络模型为训练好的口型学习神经网络的结果模型,姿态学习神经网络和口型学习神经网络可以是神经网络的结构化堆叠,例如生成对抗网络(GAN,Generative Adversarial Network),将采集得到的视频数据作为样本分别输入姿态学习神经网络和口型学习神经网络进行训练。

[0069] 图7是本实施例中元宇宙情感陪护虚拟人聊天内容客制化生成的原理图。

[0070] 如图7所示,聊天内容客制化单元142用于提供聊天内容生成客制化功能,可以利用特定神经网络技术允许客制化虚拟人生成的聊天内容时涉及的个性化元素,使内容重点回应被陪护人的情感关切。进一步地,虚拟人学习模块14具有预设的信息获取单元,可自动从互联网获取指定领域的公共信息以用于聊天内容的生成,陪护委托人可以指定所需要涵盖的领域,以达到调整聊天内容生成侧重点的目的,提高针对陪护委托人生活公共信息的回答准确性与完整度。进一步地,虚拟人学习模块14可以接受陪护委托人人工提供的多种信息,包括生活近况、生活要素、近期生活事件等,并存储这些信息作为语音聊天内容生成神经网络的基础。具体地,陪护委托人通过其委托人终端21录入多模态语料数据,然后聊天内容客制化单元142借助自然语言处理与生成单元121的语音识别和情感语义分析功能对多模态语料数据进行提取,得到语义和情感识别结果。语义和情感识别结果由陪护委托人审核后交由语音聊天内容生成神经网络模型进行生成尝试,生成聊天内容示例,并由陪护委托人对生成的示例进行评估,评估结果作为语音聊天内容生成神经网络对学习模型的改进的基础,最终进行客制化存储。进一步地,陪护委托人可以选择清除部分历史信息后重新训练模型以剔除过时信息。进一步地,对于客制化需求不高的陪护委托人,虚拟人学习模块12允许选择采用第三方聊天内容生成服务。

[0071] 本实施例中,聊天内容客制化单元142包括语音聊天内容生成神经网络模型,为训练好的内容生成神经网络,内容生成神经网络可以是结构化的神经网络堆叠。多模态语料数据传递至NLP网络进行语义分析处理以提取出语料要点,一方面输入数据库进行存储与反馈显示,向陪护委托人反馈已录入的语料信息要点,另一方面送入内容生成神经网络进行训练以得到客制化生成模型。

[0072] 环境录入与配置模块15提供虚拟人背景场景和环境氛围的录入与配置功能,为可视化单元122中虚拟人所处的虚拟环境提供支撑。环境录入与配置模块15包括预设模型存储单元151、影像资料提交接口152、环境模型生成单元153、环境细节配置单元154。

[0073] 环境录入与配置模块15提供预设与客制化两种方式供陪护委托人选择场景配置方法。在预设方式下,预设模型存储单元151中预先存储了若干经典聊天背景环境的三维模

型,如:家中、办公室、餐厅等常见场景,陪护委托人可直接预览并选定预设模型中的一种作为虚拟人的环境背景。在客制化方式下,环境录入与配置模块15向陪护委托人提供影像资料提交接口152,陪护委托人可通过相机拍摄并按要求提交希望录入的环境的若干照片或全景照片,环境录入与配置模块14接受后,其环境模型生成单元153以影像资料为基础进行环境三维重建并生成三维模型供预览,陪护委托人预览后可以针对生成模型中不理想的位置,重新拍摄照片并重新进行三维重建,直至满足期望为止,经陪护委托人确定后,生成环境三维模型存储并供可视化单元122使用。进一步,在环境图像与三维信息的录入过程中,委托人终端21实时输出合成的元宇宙虚拟环境,并智能分析合成完整度,根据合成情况向陪护委托人发出操作提示以提高录入完整度。进一步地,陪护委托人还可以选择限定出现在情感陪护过程中的背景范围。

[0074] 本实施例中,环境模型生成单元153包括预先训练好的三维重建神经网络模型,其中三维重建神经网络可以是结构化的神经网络的堆叠,用于生成环境三维模型。

[0075] 进一步地,环境细节配置单元154允许陪护委托人决定所生成的虚拟人背景场景的天气、光线等条件,提供固定预设与自动获取两种配置方式供陪护委托人选择。其中,在预设方式下,预设模型存储单元151预先存储了若干预设环境细节效果供陪护委托人预览并选择,包括晴天、多云、阴天、雨天、雪天等常见天气与柔光、强光、暗光等常见光效的多种组合搭配。在自动获取方式下,环境细节配置单元154接受陪护委托人输入的地理位置信息,通过互联网自动获取指定地理位置的天气状况,并传递给可视化单元122在情感陪护过程开始时作为环境生成依据。进一步地,环境细节配置单元154还提供环境时刻配置,陪护委托人可选择固定时差环境时刻实时生成或指定时刻。在固定时差环境时刻实时生成方式下,环境细节配置单元154接受陪护委托人指定的固定时差值,并在情感陪护过程开始时获取被陪护人所处地点时刻,进行运算后得到虚拟人环境时刻条件并传递至可视化单元122进行生成。在指定时刻方式下,环境细节配置单元144接受陪护委托人希望的虚拟人背景环境所处时刻并进行存储,在情感陪护过程开始时传递至可视化单元122进行生成。

[0076] 控制模块16用于对上述各模块的工作进行协调控制。

[0077] 图19是本实施例中基于神经网络的元宇宙情感陪护虚拟人实现方法的流程图。

[0078] 如图19所示,基于上述虚拟人系统,本实施例还提供一种基于神经网络的元宇宙情感陪护虚拟人实现方法,基于针对聊天内容客制化方面的虚拟人学习配置,提供聊天内容合成支持配置,允许客制化录入多模态形式语料与选择包含公开信息领域,根据陪护委托人试用反馈增进录入训练以提高客制化效果。

[0079] 该方法具体包括以下步骤:

[0080] 步骤S1,视听形象客制化方面的虚拟人学习配置;

[0081] 步骤S2,聊天内容客制化方面的虚拟人学习配置;

[0082] 步骤S3,环境客制化与配置;

[0083] 步骤S4,基于上述配置生成元宇宙情感陪护虚拟人,对被陪护人进行元宇宙情感陪护;

[0084] 步骤S5,在陪护过程中为陪护委托人提供信息反馈。

[0085] 以下将详细说明各步骤。

[0086] 步骤S1,视听形象客制化方面的虚拟人学习配置,虚拟人学习模块14采集陪护委

托人的多模态数据,基于多模态数据生成元宇宙情感陪护虚拟人的视听形象,显示预览效果,并循环地根据陪护委托人对预览效果的反馈调整配置或重复训练以改善学习效果,最终得到客制化视听形象。

[0087] 图5、图6分别是本实施例中元宇宙情感陪护虚拟人系统的交互示意图1、2,示出了个体形象客制化过程中的交互画面。

[0088] 如图5-6所示,针对虚拟人的特征配置可以包括选取预设自定义与采集训练客制化的配置方法,特征可以包括视觉特征与声音和语言特征,例如外貌、半身形象、姿态、口型、声纹、音调、方言等。虚拟人的特征配置可以包括多个子页面,每个子页面可以显示部分配置信息,每一子页面间的切换可以通过由陪护委托人的操作触发,例如点击切换按钮,也可以是根据配置情况自动识别跳转。

[0089] 多模态数据可以包括图像、视频、音频等各种形式数据,为了更准确地采集多模态数据,虚拟人特征配置页面在采集过程中可以为用户提供对应适当形式的反馈,例如为视频的采集提供实时的画面反馈,在采集过程中还可以根据扫描识别情况,显示对应的操作提示信息,引导用户正确完成采集扫描。

[0090] 为了提高虚拟人特征配置的准确性,完成配置后可以显示配置效果预览界面供陪护委托人评估,预览类型可以包括视觉类型和声音类型的预览。对于视觉类型数据的预览,试合成模型效果预览的显示视角可以通过由陪护委托人的操作自由变换,例如滑动屏幕、改变装备了陀螺仪的终端方向;对于声音类型数据的预览,试合成的声音可以通过陪护委托人指定文字做示例转换,并通过陪护委托人的操作进行播放;完成评估后,陪护委托人可以根据评估情况操作结束配置过程或返回重新配置预设选项或重新采集或进一步采集多模态信息。

[0091] 具体地,在采集训练客制化的配置方法下,对于视觉外貌特征学习后,如图3所示,终端多模态数据采集设备采集多模态数据,采集到的图像和视频数据作为样本输入到预先训练好的三维重建神经网络模型中进行生成,生成人像三维模型。完成生成后,陪护委托人可以查看合成的模型效果,若合成效果不佳,可以返回到多模态数据采集重新进行学习过程。对于声音类型特征学习以及动态特征学习,如图4所示,采集到的音视频数据作为样本分别输入到迁移学习神经、姿态学习神经网络和口型学习神经网络,经过训练后可以对应得到声音与语言特征神经网络模型、姿态与神情生成神经网络模型、口型生成神经网络模型并进行存储。

[0092] 比如,如图5所示,陪护委托人选择虚拟人特征配置的人像特征的采集训练客制化配置方法后,终端跳转至扫描界面,实时显示来终端相机的视频画面,并根据识别情况显示操作提示,识别录入完成后,后台处理完成采集数据后,自动跳转至预览子页面,显示采集到的人像模型,此时陪护委托人可以滑动预览区域屏幕或转动设备以改变视角进行预览和评估,若陪护委托人满意录入及合成效果,则可以选择完成以使终端响应存储该形象,若不满意,则可以选择重新录入以反复进行采集以完善录入至满意为止;陪护委托人完成客制化后点击完成控件以结束配置过程并对配置进行存储及应用。

[0093] 又比如,如图6所示,陪护委托人选择虚拟人特征配置的声音和语言特征的采集训练客制化配置方法后,终端跳转至录入界面,显示供陪护委托人阅读录入的语料信息,其中语料信息可以来自于互联网或预先定制,显示操作提示信息,用户点击代表开始录入的麦

克风按钮控件后,实时显示来自终端相机的视频画面。录入完成后,后台处理完成采集数据并完成训练后,自动跳转至子页面,显示可交互文本框与控件,陪护委托人可以通过文本框输入想要的测试的语料并点击控件进行语音合成尝试、预览和评估,若陪护委托人满意录入及合成效果,则可以选择完成以使终端响应存储该网络模型,若不满意,则可以选择继续录入以改善学习效果至满意为止;陪护委托人完成客制化后点击完成控件以结束配置过程并对配置进行存储及应用。

[0094] 步骤S2,聊天内容客制化方面的虚拟人学习配置:虚拟人学习模块14提供聊天内容合成支持配置,允许陪护委托人客制化录入多模态形式语料与选择包含的公开信息领域,基于多模态形式语料与包含的公开信息领域生成语音聊天内容,并根据陪护委托人的试用反馈增进录入训练以提高客制化效果,最终得到客制化语音聊天内容。

[0095] 其中,针对聊天内容客制化方面的虚拟人学习配置可以包括聊天内容合成支持技术的配置选择,其中聊天内容合成支持技术可以是第三方公司提供的聊天内容智能合成服务,例如:科大讯飞公司的交互服务解决方案、微软小冰的聊天机器人服务等,也可以选择录入多模态形式语料对神经网络进行训练后完成内容合成的客制化。多模态形式的语料录入可以包括文本语料、音频语料以及自动获取的公开数据信息语料。聊天内容合成支持配置的页面可以分为多个子页面,每个子页面显示部分配置信息、录入反馈信息以及评估信息。

[0096] 如图7所示,用户通过多模态数据录入交互录入语料内容信息,例如音频、文本、领域信息,其中信息范围领域的录入可以通过在预设库中搜索并选择确定,领域信息进一步地传递至Web公共信息获取服务以获取所选的对应领域的具体信息数据;多模态数据进一步传递至NLP网络进行语义分析处理以提取出语料要点,一方面输入数据库进行存储与反馈显示,向陪护委托人反馈已录入的语料信息要点,另一方面送入内容生成神经网络进行训练以得到客制化生成模型。为了提高内容客制化准确性,完成录入与训练后,终端可以显示试生成评估界面,利用学习训练得到的模型与陪护委托人进行模拟对话,陪护委托人可以对内容合成情况进行评估,评估意见回传到神经网络以进行问答型生成任务的训练。

[0097] 历史录入的预料信息要点可以被删除,删除生效后无论是否录入新的语料要点,内容合成神经网络模型都将被重新训练。

[0098] 其中,神经网络对聊天内容的生成以人性化为原则,生成内容以回应被陪护人的情感需求和感性问题为重点,生成内容还可以包括陪护委托人通过语料录入意图分享给被陪护人的信息,信息可以是近况、趣闻、需要叮嘱的事项等,以满足被陪护人情感沟通需求为首要原则。

[0099] 图8、图9分别是本实施例中元宇宙情感陪护虚拟人系统的交互示意图3、4,示出了聊天内容客制化过程中的交互画面。

[0100] 如图8所示,进入针对聊天内容客制化方面的虚拟人学习配置页面后,首先显示聊天内容后端的配置子页面,包括自定义生成方法配置选项和第三方后端支持方法配置选项信息,其中选中第三方后端支持方法配置选项后还可以显示可选及待选的聊天服务提供商信息。如图9所示,选择自定义生成方法后跳转至公共信息涉及范围领域选择配置子页面,显示陪护委托人目前已添加的领域信息及删除控件,可以使用分页交互控件来确保大量领域信息的完整显示,显示搜索控件,其中陪护委托人在可以进行键入交互,进一步地可以显

示针对所键入的关键词在预设领域库中的所以匹配项,并允许陪护委托人查看添加状态信息及交互以选中进行添加;完成并选择下一步后进入语料录入配置子页面,显示文本信息录入交互控件,包括用于陪护委托人输入文本信息的文本框及解析并添加按钮控件,显示音频形式录入交互控件,显示目前已添加的语料信息要点及删除控件,可以使用分页交互控件来确保大量领域信息的完整显示;完成并选择下一步后进入试生成评估界面,显示操作提示,显示语音输入评估控件,显示试生成内容示例区域及评价与重新生成控件。陪护委托人完成客制化后点击完成控件以结束配置过程并对配置进行存储及应用。

[0101] 步骤S3,环境客制化与配置:环境录入与配置模块14显示虚拟人所处的虚拟环境配置选项,允许采集环境图像与三维信息,基于采集的信息生成元宇宙空间环境三维模型及环境细节,并实时显示配置预览效果,根据陪护委托人对预览效果的反馈情况改善元宇宙空间环境三维模型及环境细节。

[0102] 虚拟人所处的环境配置可以包括环境三维模型的配置,背景场景细节特征配置,场景细节可以是天气、光氛围、时刻,每个细节可以限定为只可在若干个预制效果中进行选择,也可以允许陪护委托人自由定义或指定规则,例如可以由陪护委托人通过下拉菜单在有限的选项中进行选择;虚拟人所处的环境配置可以包括多个子页面,每个子页面可以显示不同特征方面的配置信息与预览画面,每一子页面间的切换可以通过由陪护委托人的操作触发,例如点击切换按钮,也可以是根据配置情况自动识别跳转;显示配置预览效果可以包括:显示选取的预制模型的3D视角预览画面,显示根据采集的环境图像与三维信息重建的虚拟人环境三维模型的3D视角预览画面,显示配置不同场景细节下三维模型效果的3D视角预览画面,3D视角预览画面的显示视角可以通过由陪护委托人的操作自由变换,例如滑动屏幕、改变装备了陀螺仪的终端方向。

[0103] 其中环境三维模型的配置可以包括选取预设与采集扫描客制化的配置方法,预设配置反方法预制了若干种常见场景布局及陈设的三维模型,例如:客厅、办公室、餐厅等场景,预制的场景三维模型可以通过典型场地扫描录入合成,也可以通过建模工具人工建立;扫描客制化配置方法可以借助终端视频录入设备采集环境图像信息,其中环境图像信息采集的页面可以显示实时的采集画面,以及根据识别处理结果针对陪护委托人提出的操作提示,输入到预先训练好的三维重建神经网络模型中进行生成,其中三维重建神经网络可以是结构化的神经网络的堆叠,生成环境三维模型,显示预览画面由陪护委托人进行评估,陪护委托人可以根据评估结果选择针对某个生成效果较差的视角进行补录或完成录入以进入环境细节配置子页面。

[0104] 比如,如图10所示,陪护委托人选择环境配置后,显示需要选择环境三维模型的配置方法的界面,陪护委托人选择预设方法后,效果预览部分显示当前所选预设模型的预览视角;陪护委托人选择录入办法后,效果预览部分和预设效果选择插件被禁用,等待陪护委托人点击下一步按钮进行录入;参见图11,陪护委托人进入扫描录入界面后,页面显示实时采集的视频画面,并显示操作提示,识别所需的录入完成后,自动跳转至预览评估界面,预览区域视角画面根据陪护委托人操作进行变换,陪护委托人可以针对某一视角点击补录该视角按钮控件回到扫描录入界面进行补录,也可以点击下一步进入环境细节配置界面;参见图12,环境细节配置界面显示光氛围、天气、时间细节的下拉选择菜单,效果预览区域显示根据选择对生成的环境三维模型进行渲染后的3D视角预览画面;陪护委托人完成客制化

后点击完成控件以结束配置过程并对配置进行存储及应用。

[0105] 步骤S4,情感沟通陪护模块12基于步骤S1~S3得到的客制化视听形象、客制化语音聊天内容以及客制化元宇宙空间环境三维模型及环境细节,合成元宇宙情感陪护虚拟人的音视频流,基于被陪护人的语音,对被陪护人的情感语义与需求进行智能分析,根据分析结果差异化满足被陪护人的服务需求与情感沟通需求,并提供陪护委托人实时接入、与被陪护人即时通讯的功能。

[0106] 其中,情感陪护功能界面包括等待开始陪护的欢迎页面;响应于被陪护人选择开始后的情感陪护界面,包括显示元宇宙情感陪护虚拟人形象及播放其声音;显示陪护中辅助信息与提示,辅助信息与提示包括:被陪护人终端设备相机捕捉到的实时画面、是否正在与陪护委托人进行即时通话的状态信息、完成信息服务后的辅助信息显示等;显示结束陪护过程的按钮控件,陪护过程可以通过被陪护人点击按钮控件进行结束。

[0107] 陪护过程中元宇宙情感陪护虚拟人的形象以根据陪护委托人配置存储的数据库信息为基础进行生成,元宇宙情感陪护虚拟人与被陪护人的交互,其中交互包括语言交互、视觉变化交互等。

[0108] 如图13所示,情感陪护过程中,被陪护人终端设备借助语音输入设备实时获取被陪护人的语言信息,输入到自然语言处理神经网络进行情感语义分析,提取得到语言信息语义要点与被陪护人需求,其中被陪护人的需求可以被分为信息服务需求与情感沟通聊天需求,根据陪护人需求,信息陪护需求的语义将被传递至Web信息服务进行自动化处理以满足被陪护人需求,例如查询某商品信息、查询未来天气信息等,信息服务处理结果及信息详情将作为语料基础传递至内容生成部分进行语句合成,以完成回复内容的生成,对应地,情感沟通需求的语义要点将直接传递到内容生成部分,作为交互输入以进行答复语句的合成,其中内容生成部分可以是存储于虚拟人配置数据库的客制化内容生成模型驱动的内容生成神经网络,也可以是第三方的内容生成服务。需要说明的是,信息服务处理结果与信息详情的语句合成所传递至的内容合成部分与情感沟通需求语义要点所传递至的内容合成部分结构并不一致,同时语义要点也传递至信息反馈模块进行进一步处理;完成聊天回复内容合成后,内容一方面被传递至从文本到语音网络部分,其中文本到语音网络是以虚拟人配置数据库中经客制化的声音特征模型为基础的特定结构神经网络,生成元宇宙情感陪护虚拟人的声音音频流,另一方面,内容被传递至形象细节生成部分,可以包括口型生成神经网络、姿态与神情生成神经网络等,其中神经网络将以虚拟人配置库中经客制化训练的对应网络特征模型为核心,以经客制化的元宇宙情感陪护虚拟人三维形象模型为基础进行生成,得到具备细节及动态的系列元宇宙情感陪护虚拟人形象三维模型,输入三维模型合成部分,与环境模型合并得到元宇宙空间的视觉3D模型,其中环境模型是经陪护委托人客制化配置并存储于虚拟环境配置数据库的具备细节信息的环境三维模型,并确定根据进一步确定的观察视角得到3D模型视频流。最终3D模型视频流与对应时刻音频流合并得到情感陪护虚拟人的元宇宙个体交互响应,并经由被陪护人终端完成输出。

[0109] 元宇宙虚拟人的虚拟环境细节配置还可以取决于陪护委托人录入的现实信息,例如陪护委托人提交的所处地理位置信息等,元宇宙情感陪护虚拟人的场景构建可以首先通过Web信息获取服务获取对应的现实信息,例如所录入的地理位置处的天气情况与时刻。

[0110] 如图14所示,陪护委托人还可以通过实时对话功能接入被陪护人的情感沟通陪护

过程,此时陪护委托人通过终端视频输入设备实时录入视频图像,元宇宙情感陪护虚拟人形象三维模型及其所有细节与动态特征全部通过三维重建神经网络进行实时构建,其中三维重建神经网络可以是神经网络的结构化堆叠;元宇宙情感陪护虚拟人的语言音频全部来自陪护委托人终端的音频输入设备录入。需要注意的是,陪护委托人既可以选择预定即时通话,也可以选择在被陪护人进行情感陪护的过程中中途接入,其中选择预定即时通话时,被陪护人终端将收到对应的请求信息,被陪护人终端可以选择同意或拒绝。

[0111] 图15是本实施例中元宇宙情感陪护虚拟人系统的交互示意图8,示出了陪护委托人与被陪护人实时交流过程中的交互画面。

[0112] 如图15所示,通过被陪护人选择开始陪护过程后的情感陪护页面,被陪护人可以看到来自本人终端相机捕获的实时画面,接收情感陪护虚拟人的元宇宙个体交互响应,正在进行实时通话的状态信息。

[0113] 此外,陪护中被陪护人还可以提出更换视角或进行其它元宇宙互动的要求,包括在陪护委托人配置的元宇宙环境中进行动态游览的要求、以及切换视角到其它环境氛围下的要求等,例如换到光线更好的地方等。其中被陪护人的要求由前述系统中的自然语言处理神经网络情感语义分析得到,通过作用于虚拟人的元宇宙个体交互响应生成过程完成满足。

[0114] 步骤S5,信息反馈模块13在情感陪护过程中为陪护委托人提供反馈信息。

[0115] 信息反馈模块13基于情感陪护过程中的情感语义分析,存储情感陪护过程中智能提取的被陪护人的情感语义与需求要点,配套存储对应时间段的音视频数据,响应于陪护委托人开始查看被陪护人近期陪护情况的操作,在委托人终端21显示信息反馈交互界面,包括陪护记录时间及所有记录数据详情显示等。

[0116] 信息反馈存储情感语义与需求要点的依据可以是根据智能评价体系对提取的被陪护人情感语义做出的价值评估得分,例如:智能评价体系可以是以自然语言处理神经网络为基础构建的评分网络,根据价值评估得分着重存储被陪护人希望传达的信息要点,同时配套存储的还包括系统提取要点时所利用到的音频数据、音频数据对应时间片的视频数据。

[0117] 信息反馈于陪护委托人类设备显示,交互界面可以由若干个子交互界面构成,每个子页面可以包括信息反馈系统存储的部分信息。

[0118] 图16是本实施例中元宇宙情感陪护虚拟人系统的交互示意图9,示出了陪护委托人查看陪护记录时的交互画面。

[0119] 如图16所示,陪护委托人进入信息反馈功能模块后,显示陪护记录总览界面,可以通过分页控件保证陪护记录显示的易读性,经陪护委托人操作选中,例如点选条目,展开显示陪护记录概览信息,例如陪护进行时间;经陪护委托人进一步操作选中,例如点击概览信息,跳转至陪护记录详览子页面,显示对应陪护记录的详细记录资料,包括显示对应陪护记录过程中记录的所有重要信息要点,显示重要节点视频与图像资料,经陪护委托人操作后可以播放关键记录片段音频,操作可以是点击音频播放控件。

[0120] 在陪护过程较短或存储控件富余,也可以完整记录存储整个陪护过程的音视频数据。

[0121] 此外,虚拟人配置的各个配置界面,在终端与服务器通信工程过程、神经网络模型

训练过程、神经网络处理过程中,终端有对应的示意等待交互界面,其中,对于部分耗时可能较长的过程,例如神经网络模型的训练过程,终端可以允许陪护委托人暂时离开对应界面进行其它操作,模型训练承载端可以进行异步继续进行训练任务,并在训练完成时通过终端给予陪护委托人某种形式的通知,例如弹窗提醒等。

[0122] 图17、图18分别是本实施例中元宇宙情感陪护虚拟人系统的交互示意图10、11,示出了用户终端的起始功能页面。

[0123] 各类型终端还根据终端用户类型拥有不同的起始功能页面,比如,如图17所示,被陪护人终端22起始界面显示欢迎信息与开始陪护过程控件,又比如,如图18所示,委托人终端21起始页面显示各功能模块的入口控件。

[0124] 此外,委托人终端21和被陪护人终端22功能页面还可以包括双向互选、绑定及配置功能页面,显示与当前终端注册者绑定的注册者信息,显示绑定注册者的管理交互页面,例如新增绑定、审核新的绑定申请、解除绑定,各绑定注册者的差异化配置页面,例如为多个绑定的被陪护人分配不同的元宇宙情感陪护虚拟人状态配置。

[0125] 委托人终端21界面还可以包括选择与被陪护人进行即时通话的选项及功能界面,显示来自被委托人终端21的实时画面与本人的元宇宙情感陪护虚拟人的状态画面,播放来自被陪护人终端22采集的音频信息。

[0126] <实施例二>

[0127] 图20是本实施例中基于神经网络技术的元宇宙情感陪护虚拟人系统的应用场景框图。

[0128] 如图20所示,与实施例一相比,区别之处在于,本实施例中,部分陪护委托人还持有辅助采集终端23,辅助采集终端23与相应的委托人终端21通信连接,采集得到的信息能够通过委托人终端21提交给虚拟人系统。

[0129] 本实施例中,进阶功能需要依靠辅助采集终端23提供的进一步数据来实现。辅助采集终端23为具备3D点云模型信息扫描、景深信息扫描或其他三维立体信息扫描功能的硬件设备。具有这些功能的硬件设备为现有技术,例如3D结构光传感器、激光雷达等。

[0130] 如图3所示,对于有条件录入3D点云数据或景深数据的终端,虚拟人学习模块13还允许陪护委托人采集并提交3D点云数据或景深数据,并以采集得到的包含三维信息的数据为依据对所生成的陪护委托人3D形象模型特征进行修正。在上述步骤S1中,可以将采集的景深信息等三维信息输入建模系统中并建立三维模型,并与三维重建模型进行合成以校准三维模型细节。

[0131] 同样地,对于有条件录入3D点云数据或景深数据的终端,环境录入与配置模块14还允许陪护委托人采集并提交3D点云数据或景深数据,并以采集得到的3D数据为依据对所生成的客制化环境三维模型进行修正。

[0132] 本实施例中,其他结构及功能与实施例一中相同,因此不再重复说明。

[0133] 实施例作用与效果

[0134] 根据本实施例提供的基于神经网络的元宇宙情感陪护虚拟人实现方法及系统,采用元宇宙方式以神经网络技术为基础建立起用于老年人情感陪护的虚拟人系统,陪护委托人和被陪护人可以采用移动终端、XR等任一种方式在系统中完成虚拟人的配置和情感陪护过程,陪护委托人构建客制化的元宇宙虚拟人以托管与老人情感沟通的时间,被陪护人可

以与陪护委托人的元宇宙虚拟人角色进行趋近于真实的沟通交流,弥补了陪护委托人由于工作而未能足量投入的在陪护老人方面的精力,同时提高了被陪护人在情感陪护过程中虚拟人对象的智能性,增强被陪护人在接受陪护时的真实感、亲切感与沉浸感,以理解被陪护人感情与解决感性需求和实用需求作为情感陪护沟通的首要目标。通过智能技术建立的虚拟人聊天后端和学习系统,使虚拟人从形象到语音再到聊天内容特点都具有高度可客制化、个性化特性,解决了传统机器人情感陪护的机械死板且单一的问题,使机器情感陪护可以进行量身定制与随使用自适应学习改进;通过图像与建模技术建立虚拟人所处的场景,在场景营造、细节还原等方面逼近现实,进一步增强情感陪护过程的观感真实性,并与信息反馈机制结合,提高信息反馈效率与准确性。

[0135] 进一步地,实施例二中,部分陪护委托人还持有辅助采集终端,有条件录入3D点云数据或景深数据,虚拟人系统还以采集得到的3D数据为依据对生成的客制化虚拟人模型及客制化环境模型进行修正,从而能够进一步在场景营造、细节还原等方面逼近现实,进一步增强情感陪护过程的观感真实性,给予被陪护人更为真实、良好的陪护体验。

[0136] 上述实施例仅用于举例说明本发明的具体实施方式,而本发明不限于上述实施例的描述范围。实施例描述的元宇宙情感陪护虚拟人实现方法以及场景仅是为了更加清楚地说明本发明实施例的技术方案,并不构成对技术方案的限定。本领域技术人员可知,随着元宇宙情感陪护虚拟人装置的演变和新业务场景的出现,本申请实施例提供的技术方案对于类似的技术问题,同样适用。

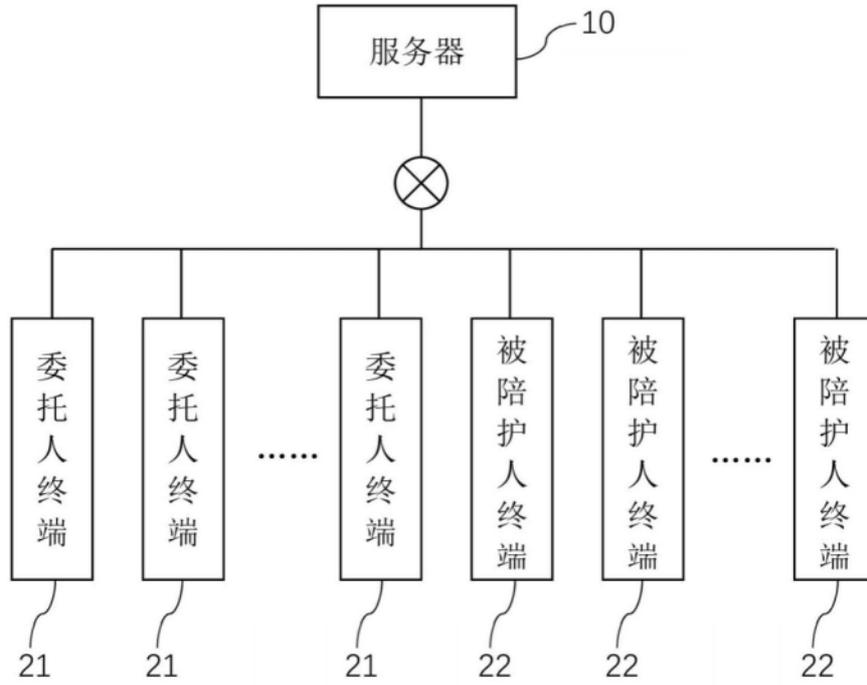


图1

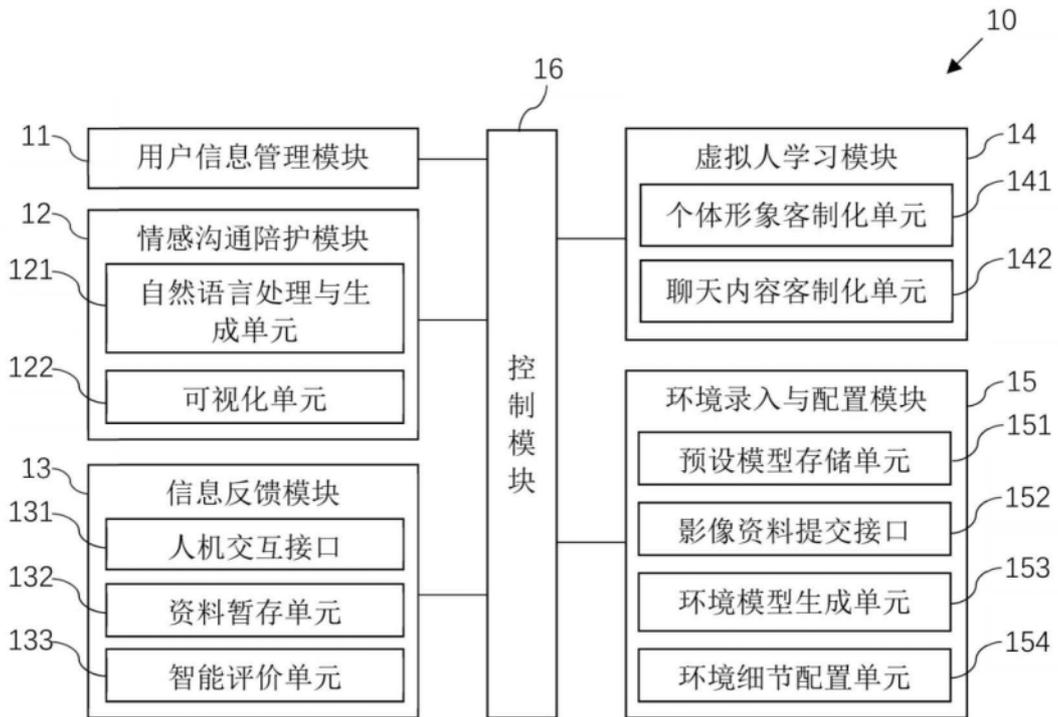


图2

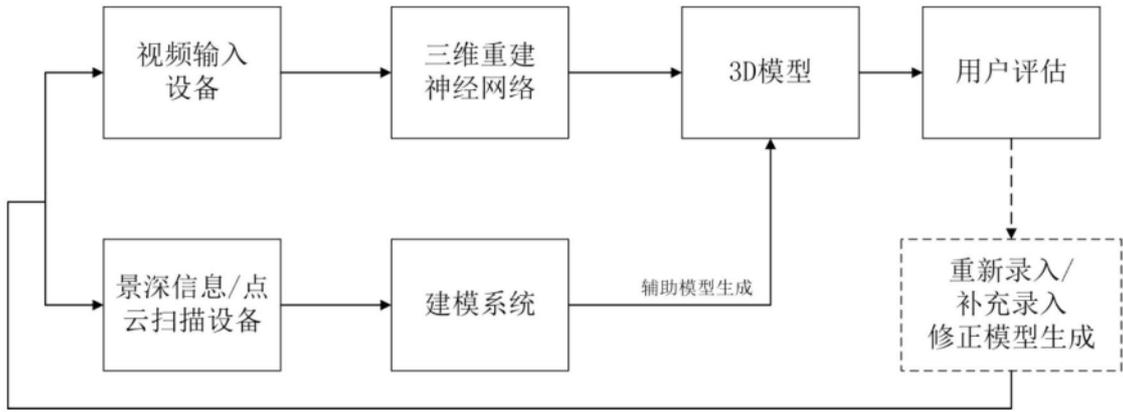


图3

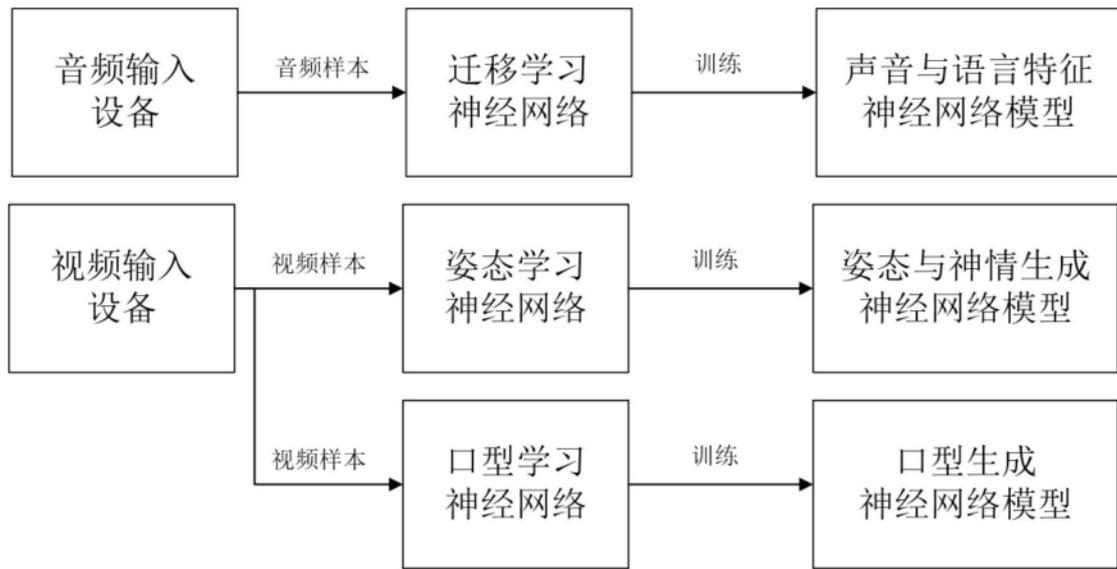


图4



图5



图6

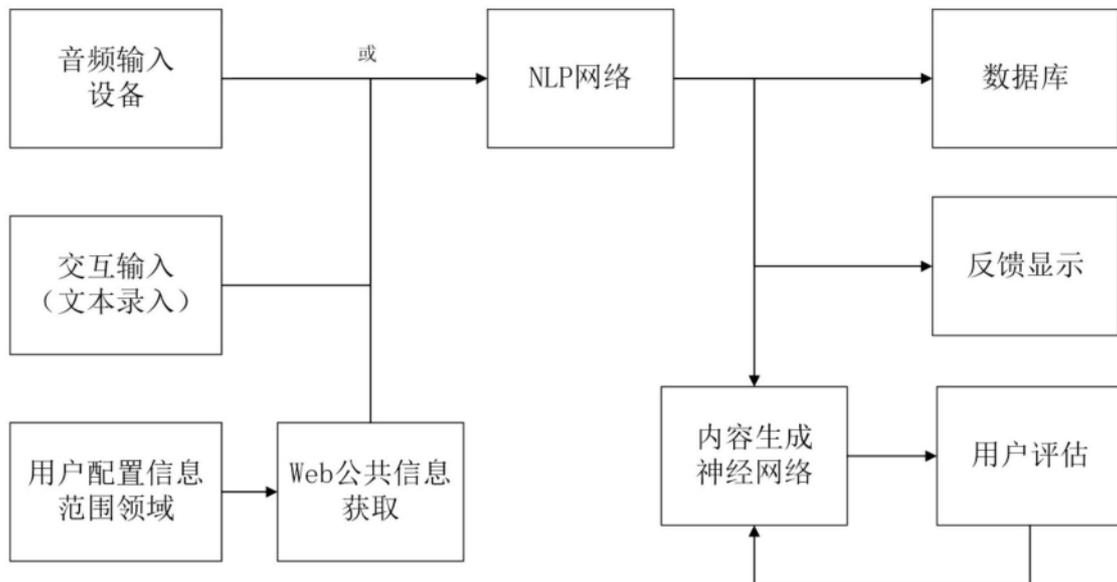


图7



图8



图9

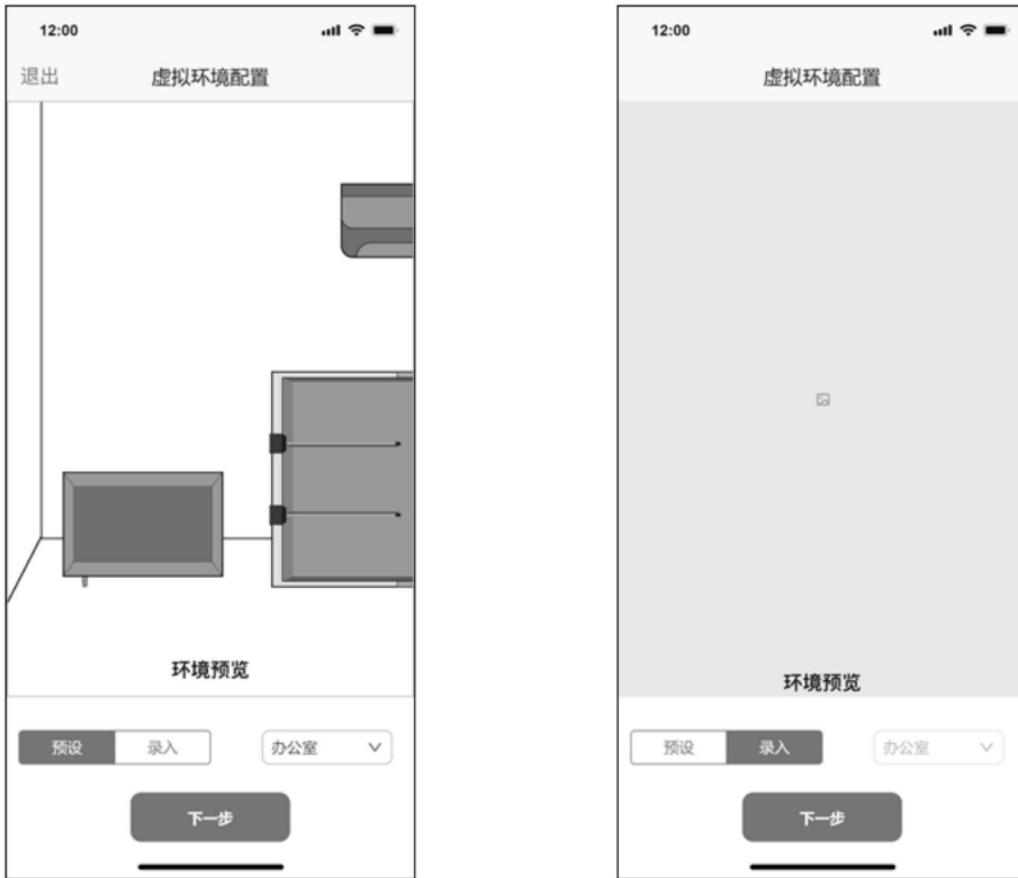


图10

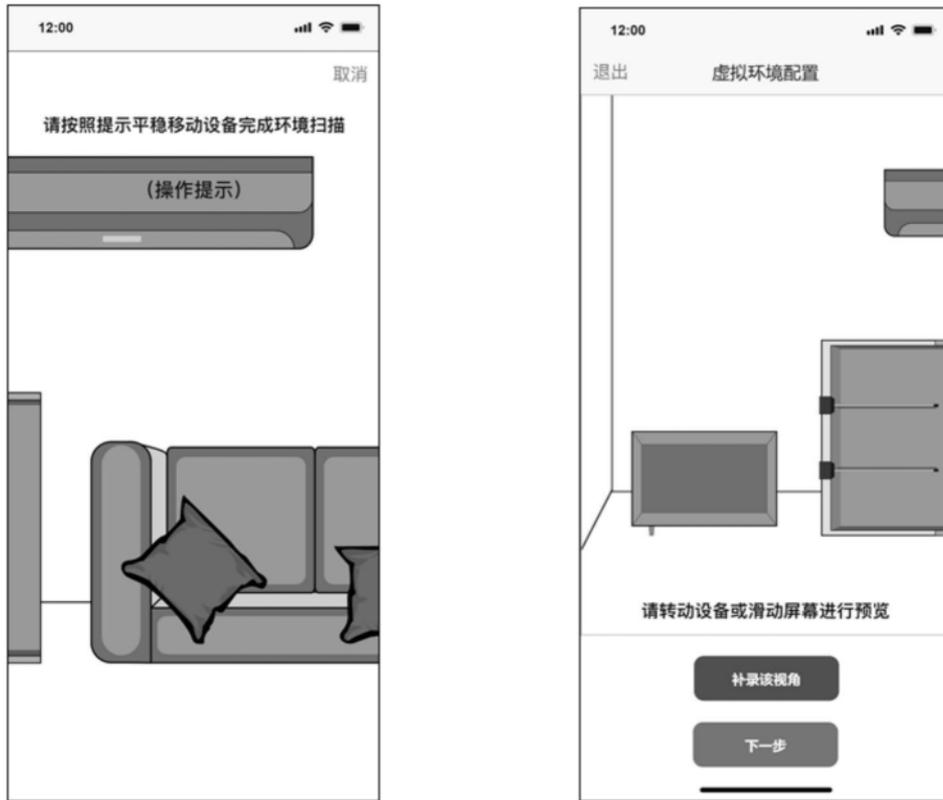


图11

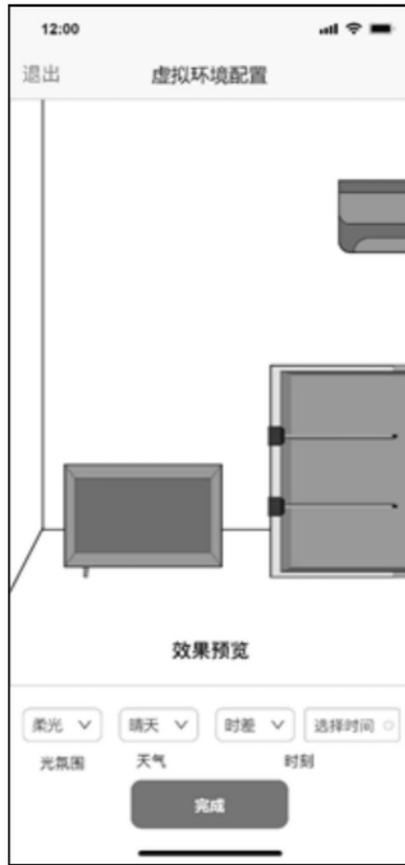


图12

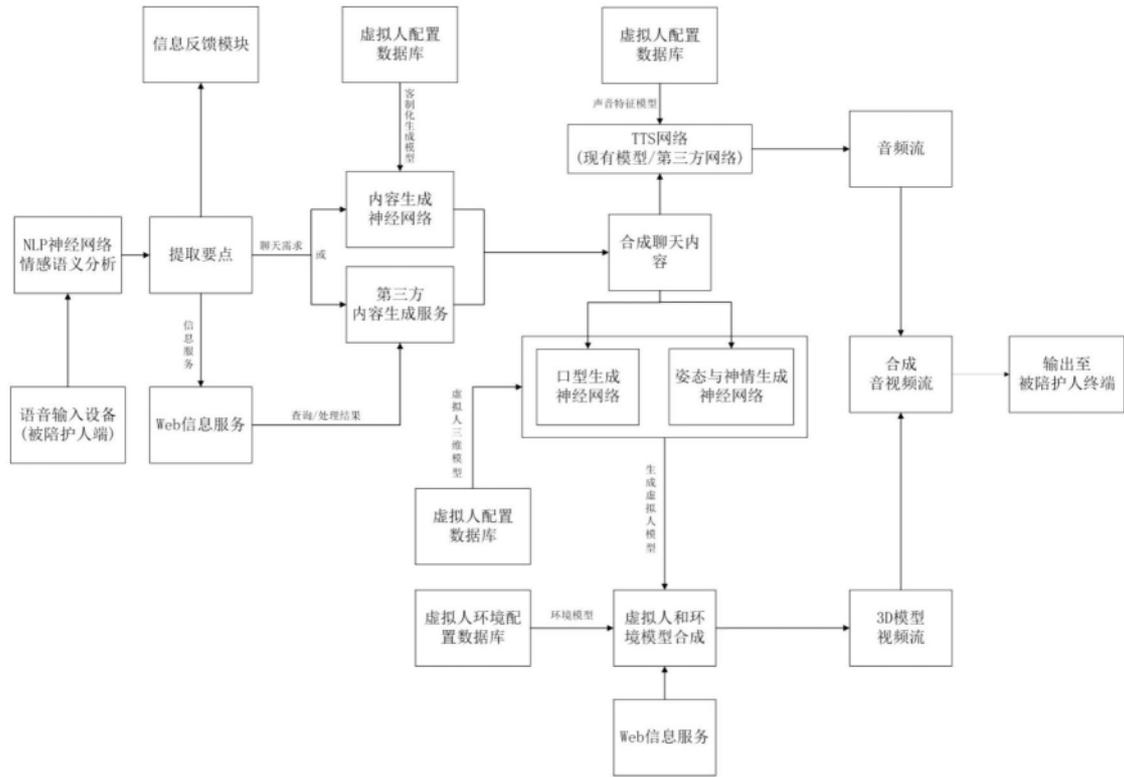


图13

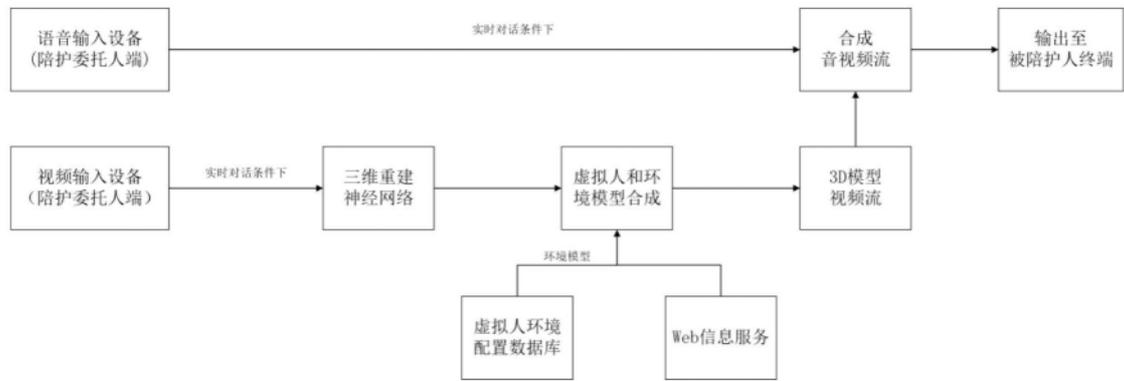


图14

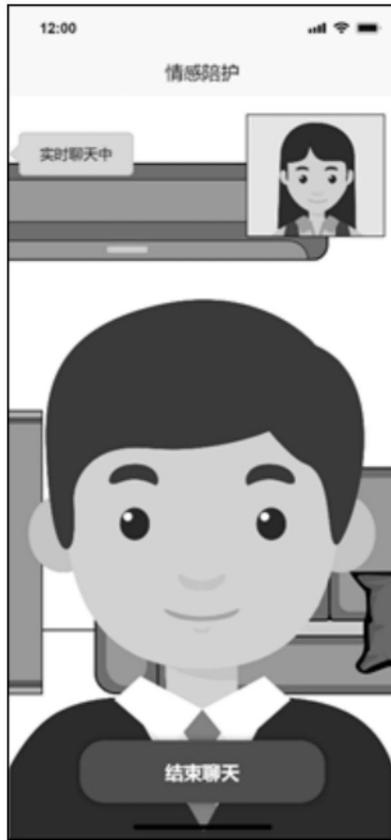


图15



图16



图17



图18

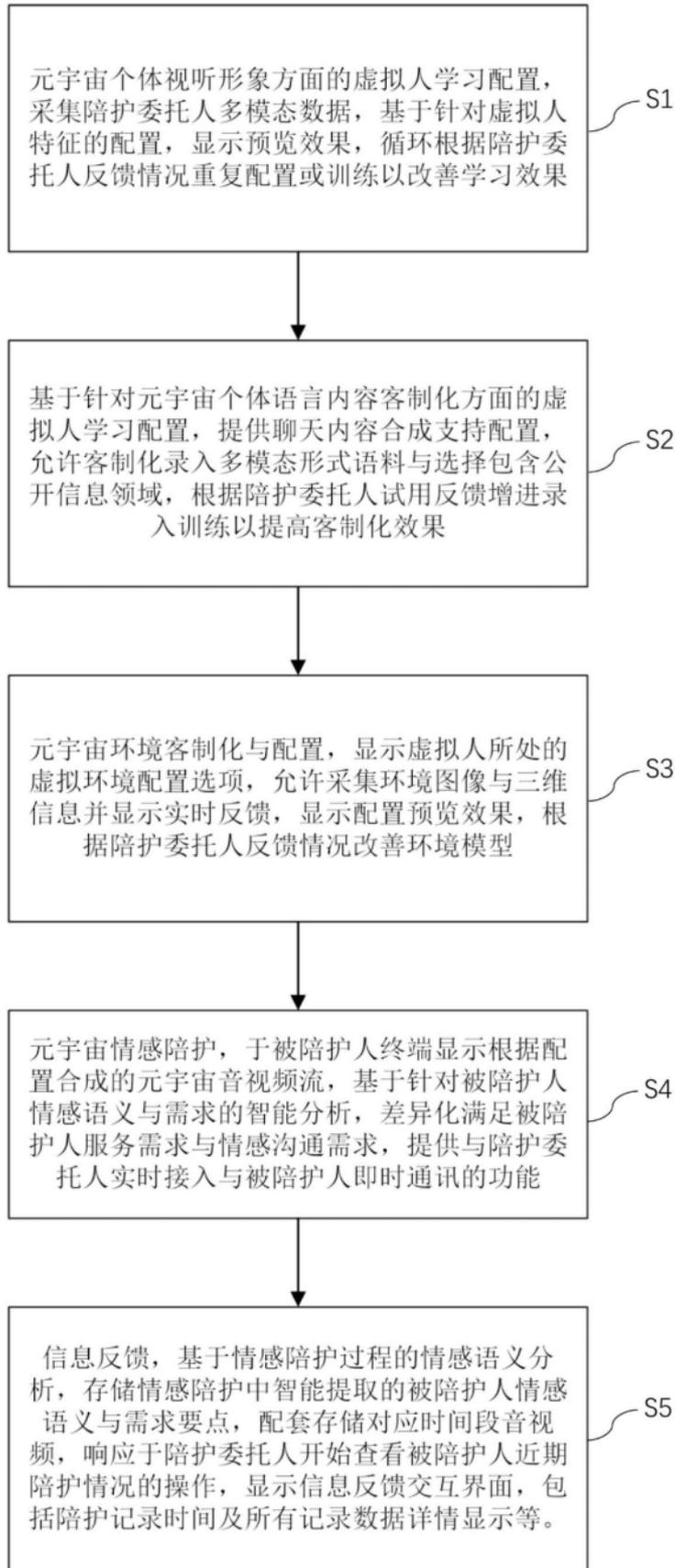


图19

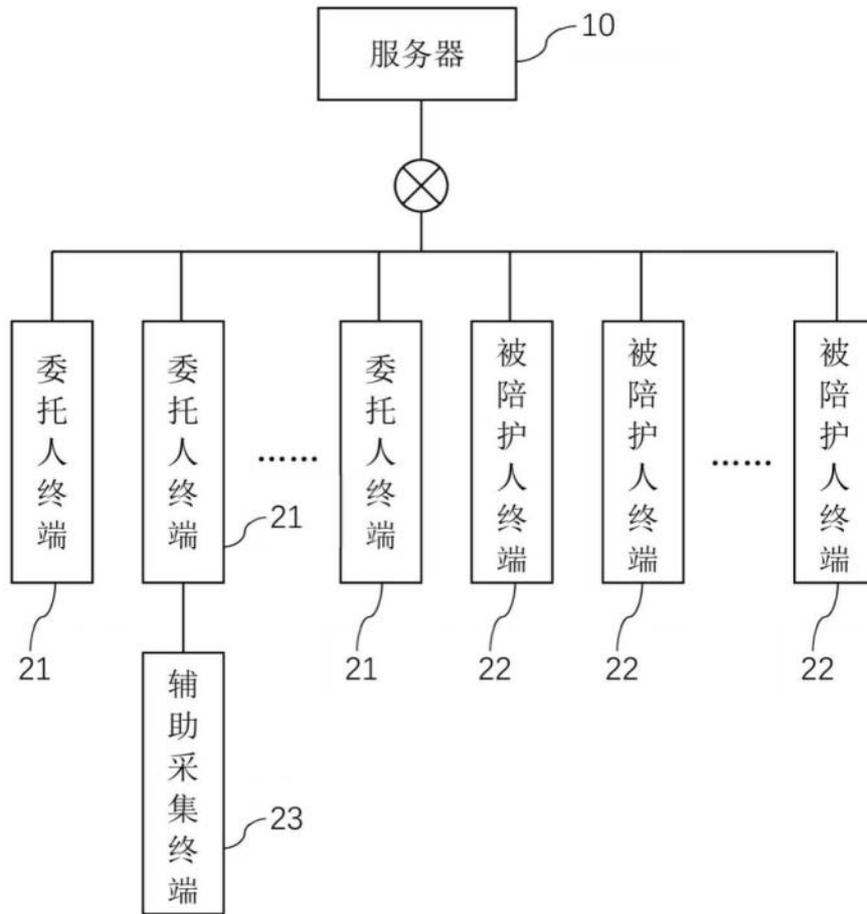


图20