**一种基于变声技术制作智能表情及声感游戏的方法**

**技术领域**

**[0001]** 本发明涉及语音变声技术领域，尤其涉及一种基于变声技术制作智能表情及声感游戏的方法。

**背景技术**

**[0002]** 语音变声技术发展到今天已经相当的成熟，百度深度语音研究团队在2017年，使用30分钟时长的语音素材进行声音的模仿，将通过系统分析运算的语音转换成和目标人相似的声音。Adobe公司编写了一个名为VoCo的程序，它的模仿将声音素材时长缩短为20分钟，加拿大Lyrebird公司仅用一分钟时长的素材就能模仿声音。而现在百度深度语音研究团队进一步创新，将素材时长缩短为数秒钟。《新科学家》报道说语音变声克隆技术也有被滥用的风险，使用这样的技术能够做出声音模仿软件，在测试中准确率已经超过了九层相似，这使得AI可能成为语音诈骗的帮凶。一方面变声技术在对社会和人民生活有益的应用不多；另外一方面即时通讯及游戏中人与人的交流还是绝大数依靠文字，虽然qq、微信、一些大型游戏都提供了语音聊天功能，但使用率却很低，究其根本，人们不愿意让陌生人知道自己的语音，或出于害怕被熟人听出来泄漏自己的隐私，无法在聊天或游戏中随心所欲，导致体感等游戏中的人机及人与人在机器系统里的信息交换单调，限制了互动程度及娱乐性。

**发明内容**

**[0003]** 鉴于以上的背景技术改进的需要，本发明提出了将用户语音强制转换成虚拟角色特定的语音为基础，并以语音识别后的文本或互动交流的文本、表情图片、语言包等为依据，通过在情绪语库检索出最为相似的记录，并根据语境模式规则矫正得出情绪特征关键词等判定用户当前情绪，然后根据所述情绪特征关键词加载所对应的预设表情表达模式，让虚拟角色显示出和当前用户的情绪一致的表情外观，即语音—文本---情绪---角色外形模型一整套的技术方案，技术方案内容如下：

变声模块：前期预备：设置数据库软硬件环境，创建数据及表格，用于记录各种参数、信息内容等。

**[0004]** 为了保护用户隐私，增强用户在网络空间的娱乐性，本发明首先创建了变声模块，具体实施步骤如下：

步骤1、为了让变声等语音系统不影响主线程序的进程，避免造成卡顿等不良的用户体验，本发明将语音模块单独成可执行文件，通过同步通讯和主线程进行数据交换，独立的变声等语音模块实施例：使用vb语言通过配置Picture或text等控件的LinkTopic、LinkMode 参数 使得适合于同步通讯，主线程序中使用控件的LinkExecute命令将指定的信息传送给语音程序。主线程序、语音模块程序分别首先配置好工程的名称及程序界面的相关属性，使得符合结合上述通讯的接收条件，通过Form\_LinkExecute过程，接收通讯的信息内容。

**[0005]** 步骤2、用户通过录音设备或智能设备上的录音传感器，进行录音，在系统中配置语音信号处理模块，由系统监听来自于用户的语音，接收分析监听到的语音信息。

**[0006]** 步骤3、变声模块：本发明采用“变调不变速”变声方法，现有技术通常有：时域法、频域法、参量法，本质都是预判变调后的声音波形然后改变采样率，伸缩语音时长和原始语音时长相等的技术方案；也可以通过第三方平台或开放库来实现，常见的音频处理库有跨平台的音频处理引擎 Fmod 、 开源免费的音频处理库SoundTouch 等；变声的一种实施例：使用SoundTouch 现实变声，开发调用时需自行将音频文件解码成 PCM 脉冲编码，传入 SoundTouch 处理后，SoundTouch 会输出处理后的 PCM脉冲编码数据。基于这种使用方式，我们可以对其自由定制以适用于自己的应用。可以先用 Android音视频中相当重要的一个API：MediaCodec 将 用户语音 解码成 PCM，通过 SoundTouch 处理得到变声后的 PCM，这样就可以调用Android音频的AudioTrack接口api 实现播放等进一步的各种操作；SoundTouch提供的一个主类SoundStreamAudioPlayer是一个线程，异步处理的方式，可以在播放过程中我们可以实时地修改音频频率参数，改变接口参数有关pitch音调的值即可改变音调，具体的接口参数有：setSampleRate(uint) 设置采样率；、setRate(double) 指定播放速率，原始值为1.0，数值大表示快数值小表示慢；setTempo(double) 指定节拍，原始值为1.0，数值大表示快数值小表示慢； setPitch(double) 指定音调值，原始值为1.0

setPitchOctaves(double) 在原音调基础上以八度音为单位进行调整，取值为[-1.00,+1.00]；setPitchSemiTones(int) 在原音调基础上以半音为单位进行调整，取值为取值为[-12,+12]等等，根据这些参数给于符合系统要求的值，即可获得目标所需的变声效果。

**[0007]** 步骤4、对游戏人物或角色进行语音模式的设定，并配置相应的数据库表来记载所述角色的语音声学特征参数等信息。当用户启动语音并启动角色智能语音功能的时候，通过步骤3将用户语音转化成具有觉得语音声学特征的语音；将用户的语音强制变化成游戏人物设定的语音模式。

**[0008]** 步骤5、系统将录音通过语音识别模块进行语音识别转化成文本文字，可以通过语音识别模块也可以通过提交给语音识别开放平台获得语音对应的文本，百度讯飞等公司都提供在线的语音识别api接口，能将语音转换成文字，只需要按照语音开放平台提高的借口参数，将语音文件和相关参数post提交到指定的网址，就可以返回得到语音对应的文本。

**[0009]** 步骤6，将步骤5获得的文本及用户直接输入的文本、发送的表情、语音包等记载到系统中等待调用，并执行相应的运算分析后，进入情绪表情的模块。

**[0010]** 情绪表情模块：让虚拟角色和用户表现一致的情绪外观表情，就得从大量的用户表达的信息进行运算分析判定，掌握心理学在人情绪方面总结出规律的研究成果，结合电学软件算法，进行无缝对接应用，本发明也是基于此根本理念创造了如下的技术方案步骤：

步骤1、创建情绪分类表、根据国际著名的心理学家，保罗·艾克曼，经过多年的研究与于1976年所创造的面部表情编码系统，对人类现实中的表情做出的细致的分类： 高兴、 悲伤、 恐惧 、 愤怒、 惊讶 、厌恶 、轻蔑，本发明在此基础上还增加了“忧愁”、“喜欢”、“沉思”的分类项，并创建记录相关情绪分类项对应的相关性信息，诸如：表情特征等的表格栏。

**[0011]** 步骤2、创建情绪二级情绪特征词句分类表，记载情绪类别所属对应的情绪特征关键词句，诸如：喜欢：喜爱、喜悦、喜好、喜欢、高兴、快乐等情感；愤怒：包括愤怒、恼怒、发怒、怨恨、愤恨等情感；忧愁：包括悲伤、悲痛、悲哀、怜悯、哀怜、哀愁、哀悯、哀怨、哀思等情感；高兴：指欢乐，身心愉悦，充满幸福的一种情感；惊讶：指惊咤、惊愕、惊慌、惊悸、惊奇、惊叹、惊喜、惊讶等情感；恐惧：指恐慌、恐惧、害怕、担心、担忧、畏惧等情感；沉思：指思念、想念、思慕等情感。

**[0012]** 步骤3、为步骤1、步骤2的分类分别设定相应的代表表情图片、语音等素材的名称及路径等信息，并记录FACS面部行为编码系统相匹配的编码组合。FACS（Facial Action Coding System）是国际著名心理学家Paul Ekman和研究伙伴W.V.Friesen作了深入的研究，(通过观察和生物反馈)他们描绘出了不同的脸部肌肉动作和不同表情的对应关系，经过多年研究于1976年所创制的“面部表情编码系统；本发明FACS面部行为编码系统，但不局限于使用该编码系统系统。

**[0013]** 步骤4，创建语境模式分类表；同样的说话内容在不同的语境下却有不同的意思，有时候甚至是相反的意思，例如“我很高兴”，在关系一般的同学，师徒等聊天中一般指真实的“我很高兴”，但在情侣赌气吵架的时候说“是啊 是啊，我很高兴好了吧”，这语句中“我很高兴”不但不能理解成高兴而得理解成生气的情绪，因此本发明使用语境模式进行统筹分析语句中情绪关键词的真实含义及语音识别形成本文在特定语库特定类中释意，具体的将语境分类，例如：同学模式、恋人模式、师徒模式、玩友模式、父女模式、母儿模式、兄弟模式、姐妹模式、家人模式，亲戚模式、闺蜜模式、损友模式、同事模式、系统设置模式等等分类，系统处于语音监视状态时候，语音识别所获的文本在语库中只检索当天语境模式类别的记录内容，并根据符合记录的内容执行相应的指令或获得特定的信息等待系统调用，形成语音操作系统，一种实施例：语境模式分类中系统切换模式、人机对话模式、聊天模式、运动控制模式，表情表达模式，交流场景模式、设置场景模式等等，并分别设置对应的程序处理模块，定义不同的权重，每个语境模式程序处理模块所需的参数都给于一个缺省值，在运行缺少参数时候，以缺省值与随机函数计算的值合并运算（相加相乘根据设计）为参数参与运算，一种实施例：系统切换模式为最高权重，其关键词为系统保护的关键词如“对话开启”，系统检测到此关键词，则开启全部的语音识别成文本的分析，“退出对话”系统对非系统切换模式的关键词一律忽视，系统切换模式下，系统只根据对应语库中该模式类别的记录的特征关键词进行监视，只当检测到该语境模式类别下有符合记录的语音识别后的文本关键词时，进行该记录相应的操作或调用信息，如系统切换模式类别下 有“运动控制模式”，从语音识别的文本含有“运动控制”的文本时候，系统自动进入运动控制模式的语境，而在运动控制模式下：语音识别出的文本只检索语库中该模式类的记录，通过运动控制的模块相应对应符合记录的操作，例如检测到“奔跑”而对应的记录为运动控制处理模块中角色速度参数为最大值、“左前”，根据现有方向 的角度 与随机函数计算的值合并运算（相加相乘根据设计），形成一个新的角度，角色就依照这两点参数和模块产生的指令进行奔跑。

**[0014]** 步骤5、为各个语境模式中涉及情绪识别的类别设置相应的根据文本中关键词及文本本身识别情绪的规则。

**[0015]**  步骤6、创建语境模式预判规则。根据关键词在指定时间断内出现的频率结合双方的用户资料，辨别互动者大致的关系初步确定语境模式，根据当前聊天起始的时间为开始时间，关键词出现的频率等预先判断的规则。实施例：恋人模式：两用户必须是一男一女，年纪差的绝对值小于15，指定特征关键词“想你”、“喜欢”等在指定时间段内出现概率的条件，实施例：instr（文本记录，“喜欢”）>0 or instr（文本记录，“爱”）>2。

**[0016]** 步骤7、修正语境模式：通过检索指定时间段内互动形成的文本中称呼的特征关键词及相应的矫正规则验证或修正步骤6预判的语境模式，实施例：检索到“爸爸”，且前后句中根据正则表达式验证人的姓名：/^[\u4E00-\u9FA5\uf900-\ufa2d]{2,4}$/ 、/^[\u4E00-\u9FA5\uf900-\ufa2d·s]{2,20}$/（表达式中20数字是考虑少数民族人的姓名）等判定没有人的姓名时，即确认为 父子模式，如果步骤6的预判记录不一致，则修正编辑对应的记录并保存。

**[0017]** 步骤8、创建情绪语库，记载蕴含各种情绪的常用语句和词语的文本信息，设定好所述文本所属的特征情绪关键词（步骤2所创建的情绪特征情绪关键词句分类）、语境模式等特性分类及其他相关信息,实施例：针对情绪二级情绪特征词句分类中“高兴”，配置记录常用词句：“真的吗 太好啦”；“呵呵 真棒”；“谢谢 我请客”；等等记录，以此类推创建完整的情绪语库。

**[0018]** 步骤9、创建拓补情绪语库，记载用户互动时对方语句中蕴含各种推定表达用户情绪的常用语句和词语的文本信息，并记载设定好所述文本所属的情绪特征关键词（步骤2所创建的情绪特征关键词句分类）、语境模式等特性分类及其他相关信息,在用户产生的信息没有能分析到用户情绪时，使用互动时对方推送给用户的文本、表情图片、flash、语言包等内容到拓补情绪语库中检索分析推定用户情绪，并使用户的虚拟角色产生对应的表情；实施例：针对情绪二级情绪特征词句分类中“高兴”，配置记录常用词句：“你有什么好高兴的”；“恭喜恭喜 必须请客”；“我也为你高兴”；等等记录，以此类推创建完整的拓补情绪语库。

**[0019]** 基于情绪语库及拓补情绪语库需要大量的数据和个性化精准的数据，方能提高情绪识别及表情推定，因此本发明创造了一种自训练自完善情绪语库及拓补情绪语库的技术方法，使得本发明系统在使用的过程中不断自我修正，自我改善综合性能，实施步骤也就是紧接上述步骤的如下几步。

**[0020]** 步骤10**、**通过用户平日聊天的文本和所发表情图片、语音包等等记录对应，分析结果；训练文本和情绪映射对应关系库，即通过用户所发表情对应的表情图片、fash、语音包中的语音，记录其特征或特别的动作，并对用户所发这些表情要素相关的素材前后一个设定时间断的聊天文本进行分析整理**，**从中获得文本特征要素，和微表情的衍射关系修正情绪语库中相关的记录或作为新记录存储到情绪语库对应的表格中，一个实施例：用户互动时候发送文本内容“我发达啦发达啦”，而紧接其后发送一张表情图片，该图片的图片说明为高兴，则以“我发达啦发达啦”文本为检索单元到情绪语库中检索相似记录，检索到符合条件的记录，将该记录的情绪特征关键词修改为“高兴”，并将高兴所对应的表情编码等一并修改到记录并保存下来，如果没有检索到符合条件的记录，则在情绪语库中以文本内容“我发达啦 发达啦”，情绪特征关键词为“高兴”，附件相关的信息为内容增加一条记录。

**[0021]** 步骤11**、**采用步骤10的方法，将映射逻辑改为旁证的逻辑关系，例如文本内容“你有什么好高兴的”映射用户“高兴”的情绪状态，对拓补情绪语库进行日常训练和完善；通过互动时对方推送给用户的文本、表情图片、语音包等信息记录对应，分析结果；训练文本和用户情绪映射对应关系库，即通过用户互动对象所发表情对应的表情图片、fash、语音包中的语音，记录其特征或特别的动作，；和用户情绪、微表情的衍射关系修正拓补情绪语库中相关的记录或作为新记录存储到拓补情绪语库。

**[0022]** 步骤12**、**表情识别对情绪语库、拓补情绪语库进行训练自我完善优化；用户视频时候的聊天文本，或聊天语音识别后的文本，进行对应，训练文本和情绪映射对应关系库，对视频、图像的中人微表情的识别现在有许多成熟的技术方案，也有商用的相关平台和产品，表情识别实施例：美国微软公司对外公布的 emotion recognition (表情识别)的api，其中只需要将指定内容作为post参数附加在网址“https://api.projectoxford.ai/emotion/v1.0/recognize”之后，并根据api的要求和提交的需要附加其他的相关参数，例如：\*.setRequestHeader("Content-Type","application/json");\*.setRequestHeader("Ocp-Apim-Subscription-Key","{subscription key}");等，就能获得表情识别的文本等结果。通过这些技术获得微表情的有关情绪的文字描述，并从微表情对应时间段内，分析此期间用户聊天文本内容，所发表情包的内容，从中获得文本特征要素，和微表情的衍射关系修正情绪语库或拓补情绪语库中相关的记录或作为新记录存储到情绪语库或拓补情绪语库对应的表格中。

**[0023]** 步骤13、文本相似度检索识别当前用户的情绪，设定一个语义阀值，系统在用户启用情绪模块时，实时对语音识别出的文本或用户输入的文本内容进行分析，即在情绪语库使用like等函数检索所述文本内容相近的记录，并将两者进行进一步的相似度计算，如结果小于系统设定的语义阀值则忽视进入下一步的逻辑线程，否则以检索到记录中特征情绪关键词为内容，到特征情绪关键词及语句表检索记录，并读取记录中的对应的表情图片、语音等素材的名称及路径等信息，FACS编码组合等内容，待系统调用。至于文本、语义相似度计算，现有技术有许多成熟的技术方案，例如杰卡德相似系数、编辑距离算法；文本、语义相似度计算实施例：采用杰卡德相似系数Jaccard index, 又称为Jaccard相似系数（Jaccard similarity coefficient）用于比较有限样本集之间的相似性与差异性。Jaccard系数值越大，样本相似度越高，两个文本变量及其值为str=“ http://9vitcn.218899.com/n/218803/”，str1=“http://9vitcn.218899.com/n/218847/”， 按照公式：

P=Convert.ToInt32(str.Intersect(str1) / str.Union(str1) + Math.Abs(str1.Length - str.Length)) \* 100)

计算结果p值则为两个文本str、srt1之间的相似度；当然关于语义相似度还有其他许多算法，但这些语义相似度算法不是本发明技术方案的特征要素，本实施例只是构成本步骤具体实施的方式及说明技术步骤的可行性。

**[0024]** 如通过上述文本、语义相似度运算没有检索到合乎要求的记录，则设定一个语义阀值2，系统在用户启用情绪模块时，实时对互动时候对方的语音识别出的文本或对方传送来的文本内容进行分析，即在拓补情绪语库使用like等函数检索所述文本内容相近的记录，并将两者进行进一步的相似度计算，如结果小于系统设定的语义阀值则忽视进入下一步的逻辑线程，否则以检索到记录中特征情绪关键词为内容，到特征情绪关键词及语句表检索记录，并读取记录中的对应的表情图片、语音等素材的名称及路径等信息，FACS编码组合等内容，待系统调用。

**[0025]** 步骤14、一种虚拟角色情绪表情实现的方法，首先将网络虚拟角色或虚拟角色的脸部设定成场景中的图层或元件，根据步骤3中记录的FACS（面部行为编码系统）编码所代表的AU肌肉运动单元、运动的规则及外部表现形式，将标准角色图片或3d模型，进行修改成符合对应脸部表情编码的图片或模型，并将所述结果或将变化的过程作为图层或一组帧的动画单元作为元件保存在系统里，或者将表情动作的影片剪辑作为元件保存到系统中，同时在系统中记录所述图层、元件的路径、名称、和映射的情绪特征关键词等信息一并记录中系统，当系统因步骤3中所述的情绪特征关键词句调用表情请求时，在虚拟角色或角色脸部位置显示该图层或播放所述的元件；一种实施例：情绪特征关键词句中“愤怒”，其面部编码为：AU4+5+7+23,其中AU4的含义是皱眉，外部显示为：眉毛降低、眼皮向下使脸裂变狭窄、两眉毛靠拢、眉毛间产生皱纹等，根据这些外部显示的细节描述将角色的人物造型的图片或3d模型做相应的改变，并将改变的结果或改变的全过程打包成元件，这些元件是可反复取出使用的图形或影片剪辑段、元件中的小动画可以独立于虚拟角色所处的场景进行播放，元件可由多个独立由标准角色形象根据面部编码系统中对应的编码外部变化描述的变化中影片帧画面构成，需要调用或经过条件触发人物表情时候，表情模型元件或图片在角色脸部位置或角色的位置播放或现实对应的表情。

**[0026]** 基于以上的技术方案，将变声技术、心理学情绪研究出的规则、神经网络科学的基本原理、软件信息技术等专业技术有机融合在一起，充分在保护隐私的前提下将用户的情绪更充分地在虚拟角色上体现，可以在多个领域切实有效的进行运用，为了方便进一步理解和创造性应用，本发明创造了一个实际应用。

**[0027]** 实时和玩家情绪同步的网络个人形象装扮系统实施例：现在的网络个人形象装扮系统只是动漫图片加上各种对动画图片进行服饰，化妆道具等进行个性化形象代表，网络个人形象只和用户的喜好偏向相关，和用户本身的信息一致性很少，本发明个人形象装扮系统结合本发明的技术方案，创建基于动漫图片，3d角色模型、vr角色模型等赋予其脸部表情及相关特征动作，并和玩家当时的表情、情绪相一致。

**[0028]** 步骤1、基于上述的变声模块、情绪表情模块的技术方案，创建多个动漫图片，3d角色模型、vr角色模型等装扮系统基础形象单元，并增加了影片剪辑段、flash动画，gif动画图片等作为装扮系统基础形象单元， 并设置相对应的语音声学特征参数。

**[0029]** 步骤2、为步骤1创建形象现实的方案选择界面，并配置相应的程式，让用户自由选择固定静态形象、3d立体形象、vr模型、影片剪辑段、flash动画，gif动画图片等方案选项作为用户个人形象装扮的方案，分别设置语音开关、变声开关、情绪同步开关、文本转语音开关等各种选择开关及其他用户设置项目并配置相应的程序。

**[0030]** 步骤3、语音开关：用户开启关闭个人形象装扮系统的语音功能。

**[0031]** 步骤4、变声开关：切换个人形象装扮虚拟角色特定的语音和用户自己真实的语音。

**[0032]** 步骤5、情绪同步开关：开启关闭个人形象装扮虚拟角色启用通过情绪表情模块实时表达用户情绪所匹配的表情。

**[0033]** 步骤6、文本转语音开关：用户使用打字交流时，开启或关闭文本自动通过文本转语音模块转化成语音，在根据用户设置的相关参数和角色形象对应的声学特征参数，决定是否转化成所述形象匹配的声调，播放给用户交流的对象。

**[0034]** 步骤7 、为每个需要用户设置的项目预设一个参数，并建立用户资料分析规则，预判用户可能偏向的选项设定为默认值，实施例：用户资料为男、19岁，系统就预判个性形象装扮的角色为奥特曼，诸如此类建立系统预判参数的规则库。

**[0035]**     以上为一种基于变声技术制作智能表情及声感游戏的方法在个人形象装扮系统中的创造性应用，是本发明技术方案的一个实施例，同时，该实施例同样具备显著的技术进步和创造性、新颖性和实用性，但基于本发明核心技术方案的应用不局限于此。

**附图说明**

**[0036]** 图1 为一种基于变声技术制作智能表情及声感游戏的方法整体逻辑框架图。

**[0037]** 特别申明：在本说明书中所述的 “实施例”等，指的是结合该实施例描述的具体特征、要素或者特点包括在本申请概括性描述的实施例中。在说明书中多个地方出现同种表述并非限定特指的是同一个实施例。也就是说，结合任一实施例描述一个具体特征、要素或者特点时，所要主张的是结合其他实施例来实现这种特征、要素或者特点被包含于本发明申请保护的权利要求范围中； 实施例是参照本发明逻辑架构及思路的多个解释性实施例对本发明进行了描述，但本发明的保护范围并不局限于此，本领域技术人员在本发明技术方案框架下可以设计出很多其他的修改和实施方式，可以对技术方案的要点变换组合/或布局进行多种非本质性变型和改进，对于本领域技术人员来说，其他的用途也将是明显的，可轻易想到实施的非实质性变化或替换，这些修改和实施方式将落在本申请公开的原则范围和精神之内。

**权利要求书**

**1.** . 一种基于变声技术制作智能表情互动及声感游戏的方法，其特征包含的步骤和要素有：

步骤1、首先为了让变声等语音系统不卡顿等，将语音模块单独成可执行文件，通过同步通讯和主线程进行数据交换；通过配置Picture或text等控件的LinkTopic、LinkMode 参数 使得适合于同步通讯；

步骤2、用户通过录音设备或智能设备上的录音传感器，进行录音，在系统中配置语音信号处理模块，由系统监听来自于用户的语音，接收分析监听到的语音信息；

步骤3、变声模块：本发明采用“变调不变速”变声方法，预判变调后的声音波形然后改变采样率，伸缩语音时长和原始语音时长相等；通过第三方平台或开放库来实现，使用SoundTouch 现实变声，开发调用时需自行将音频文件解码成 PCM 脉冲编码，传入 SoundTouch 处理后，SoundTouch 会输出处理后的 PCM脉冲编码数据，基于这种使用方式，我们可以对其自由定制以适用于自己的应用；可以先用 Android音视频中相当重要的一个API：MediaCodec 将 用户语音 解码成 PCM，通过 SoundTouch 处理得到变声后的 PCM，这样就可以调用Android音频的AudioTrack接口api 实现播放等进一步的各种操作；SoundTouch提供的一个主类SoundStreamAudioPlayer是一个线程，异步处理的方式，可以在播放过程中我们可以实时地修改音频频率参数，改变接口参数有关pitch音调的值即可改变音调，根据这些接口参数给于符合系统要求的值，获得目标所需的变声效果；

步骤4、对游戏人物或角色进行语音模式的设定，并配置相应的数据库表来记载所述角色的语音声学特征参数等信息，当用户启动语音并启动角色智能语音功能的时候，通过步骤3将用户语音转化成具有觉得语音声学特征的语音；将用户的语音强制变化成游戏人物设定的语音模式；

步骤5、系统将录音通过语音识别模块进行语音识别转化成文本文字，可以通过语音识别模块也可以通过提交给语音识别开放平台获得语音对应的文本；

步骤6，将步骤5获得的文本及用户直接输入的文本、发送的表情、语音包等记载到系统中等待调用，并执行相应的运算分析后，进入情绪表情的模块；

情绪表情模块：让虚拟角色和用户表现一致的情绪外观表情，心理学在人情绪方面总结出规律的研究成果，结合电学软件算法，进行无缝对接应用，技术方案步骤：

步骤1、创建情绪分类表、分类： 高兴、 悲伤、 恐惧 、 愤怒、 惊讶 、厌恶 、轻蔑，，并创建记录相关情绪分类项对应的相关性信息；

步骤2、创建情绪二级情绪特征词句分类表，记载情绪类别所属对应的情绪特征关键词句；

步骤3、为步骤1、步骤2的分类分别设定相应的代表表情图片、语音等素材的名称及路径等信息，并记录FACS面部行为编码系统相匹配的编码组合；

步骤4，创建语境模式分类表；具体的将语境分类，例如：同学模式、恋人模式、师徒模式、玩友模式、父女模式、母儿模式、兄弟模式、姐妹模式、家人模式，亲戚模式、闺蜜模式、损友模式、同事模式、系统设置模式等等分类，系统处于语音监视状态时候，语音识别所获的文本在语库中只检索当天语境模式类别的记录内容，并根据符合记录的内容执行相应的指令或获得特定的信息等待系统调用，形成语音操作系统，语境模式分类中系统切换模式、人机对话模式、聊天模式、运动控制模式，表情表达模式，交流场景模式、设置场景模式等等，并分别设置对应的程序处理模块，定义不同的权重，每个语境模式程序处理模块所需的参数都给于一个缺省值，在运行缺少参数时候，以缺省值与随机函数计算的值合并运算（相加相乘根据设计）为参数参与运算，

步骤5、为各个语境模式设置相应的根据文本中关键词及文本本身识别情绪的规则；

步骤6、创建语境模式预判规则；

步骤7、验证或修正语境模式；

步骤8、创建情绪语库，记载蕴含各种情绪的常用语句和词语的文本信息，设定好所述文本所属的特征情绪关键词（步骤2所创建的情绪特征情绪关键词句分类）、语境模式等特性分类及其他相关信息,实施例：针对情绪二级情绪特征词句分类中“高兴”，配置记录常用词句：“真的吗 太好啦”；“呵呵 真棒”；“谢谢 我请客”；等等记录，以此类推创建完整的情绪语库；

步骤9**、**创建拓补情绪语库，记载用户互动时对方语句中蕴含各种推定表达用户情绪的常用语句和词语的文本信息；

步骤10**、**一种自训练自完善情绪语库及拓补情绪语库的技术方法，使得本发明系统在使用的过程中不断自我修正，自我改善综合性能；

步骤11、虚拟角色情绪表情实现。

**2.** 根据权利要求1所述一种基于变声技术制作智能表情互动及声感游戏的方法，其特征包含的步骤和要素有：创建情绪分类表、表情分类： 高兴、 悲伤、 恐惧 、 愤怒、 惊讶 、厌恶 、轻蔑，在此基础上还增加了“忧愁”、“喜欢”、“沉思”的分类项，并创建记录相关情绪分类项对应的相关性信息。

**3.** 根据权利要求1所述一种基于变声技术制作智能表情互动及声感游戏的方法，其特征包含的步骤和要素有：通过检索指定时间段内互动形成的文本中称呼的特征关键词及相应的矫正规则验证或修正权利要求1中预判的语境模式。

**4.** 根据权利要求1所述一种基于变声技术制作智能表情互动及声感游戏的方法，其特征包含的步骤和要素有：创建拓补情绪语库，记载用户互动时对方语句中蕴含各种推定表达用户情绪的常用语句和词语的文本信息，并记载设定好所述文本所属的情绪特征关键词、语境模式等特性分类及其他相关信息,在用户产生的信息没有能分析到用户情绪时，使用互动时对方推送给用户的文本、表情图片、flash、语言包等内容到拓补情绪语库中检索分析推定用户情绪，并使用户的虚拟角色产生对应的表情。

**5.** 一种自训练自完善情绪语库及拓补情绪语库的技术方法，其特征包含的步骤和要素有：基于权利要求1-4的基础上，自训练情绪语库及拓补情绪语库包含：

S1、通过用户平日聊天的文本和所发表情图片、语音包等等记录对应，分析结果；训练文本和情绪映射对应关系库，即通过用户所发表情对应的表情图片、fash、语音包中的语音，记录其特征或特别的动作，并对用户所发这些表情要素相关的素材前后一个设定时间断的聊天文本进行分析整理**，**从中获得文本特征要素，和微表情的衍射关系修正情绪语库中相关的记录或作为新记录存储到情绪语库对应的表格中；

S2**、**采用步骤S1中的方法，将映射逻辑改为旁证的逻辑关系，例如文本内容“你有什么好高兴的”映射用户“高兴”的情绪状态，对拓补情绪语库进行日常训练和完善；通过互动时对方推送给用户的文本、表情图片、语音包等信息记录对应，分析结果；训练文本和用户情绪映射对应关系库，即通过用户互动对象所发表情对应的表情图片、fash、语音包中的语音，记录其特征或特别的动作和用户情绪、微表情的衍射关系修正拓补情绪语库中相关的记录或作为新记录存储到拓补情绪语库。

**6.** 根据权利要求5所述一种自训练自完善情绪语库及拓补情绪语库的技术方法，其特征还包含的步骤和要素有：表情识别对情绪语库进行训练自我完善优化；用户视频时候的聊天文本，或聊天语音识别后的文本，进行对应，训练文本和情绪映射对应关系库，使用 emotion recognition (表情识别)的api，将指定内容作为post参数附加在网址“https://api.projectoxford.ai/emotion/v1.0/recognize”之后，并根据api的要求和提交的需要附加其他的相关参数，获得表情识别的文本等结果；通过这些技术获得微表情的有关情绪的文字描述，并从微表情对应时间段内，分析此期间用户聊天文本内容，所发表情包的内容，从中获得文本特征要素，和微表情的衍射关系修正情绪语库中相关的记录或作为新记录存储到情绪语库对应的表格中。

**7.** 根据权利要求5所述一种自训练自完善情绪语库及拓补情绪语库的技术方法，其特征还包含的步骤和要素有：表情识别对拓补情绪语库进行训练自我完善优化；基于权利要求6的方法，获得微表情的有关情绪的文字描述，并从微表情对应时间段内，分析此期间用户聊天文本内容，所发表情包的内容，从中获得文本特征要素，和微表情的衍射关系修正拓补情绪语库中相关的记录或作为新记录存储到拓补情绪语库对应的表格中。

**8.** 一种虚拟角色情绪表情实现的方法，其特征包含的步骤和要素有：基于权利要求1-7的基础上，首先将网络虚拟角色或虚拟角色的脸部设定成场景中的图层或元件，根据权利要求1-6的中记录的FACS（面部行为编码系统）编码所代表的AU肌肉运动单元、运动的规则及外部表现形式，将标准角色图片或3d模型，进行修改成符合对应脸部表情编码的图片或模型，并将所述结果或将变化的过程作为图层或一组帧的动画单元作为元件保存在系统里，或者将表情动作的影片剪辑作为元件保存到系统中，同时在系统中记录所述图层、元件的路径、名称、和映射的情绪特征关键词等信息一并记录中系统，当系统因步骤3中所述的情绪特征关键词句调用表情请求时，在虚拟角色或角色脸部位置显示该图层或播放所述的元件；FACS面部编含义是皱眉等，外部显示为：眉毛降低、眼皮向下使脸裂变狭窄、两眉毛靠拢、眉毛间产生皱纹等，根据这些外部显示的细节描述将角色的人物造型的图片或3d模型做相应的改变，并将改变的结果或改变的全过程打包成元件，这些元件是可反复取出使用的图形或影片剪辑段、元件中的小动画可以独立于虚拟角色所处的场景进行播放，元件可由多个独立由标准角色形象根据面部编码系统中对应的编码外部变化描述的变化中影片帧画面构成，需要调用或经过条件触发人物表情时候，表情模型元件或图片在角色脸部位置或角色的位置播放或现实对应的表情。

**9.** 一种实时和玩家情绪同步的网络个人形象的方法，其特征包含的步骤和要素有：基于权利要求1-8的基础上，创建基于动漫图片，3d角色模型、vr角色模型等赋予其脸部表情及相关特征动作，并可以做到和玩家当时的表情、情绪相一致的网络个人形象：

步骤1、基于权利要求1-8的技术方案，创建多个动漫图片，3d角色模型、vr角色模型等装扮系统基础形象单元，并增加了影片剪辑段、flash动画，gif动画图片等作为装扮系统基础形象单元， 并设置相对应的语音声学特征参数；

步骤2、为步骤1创建形象现实的方案选择界面，并配置相应的程式，让用户自由选择固定静态形象、3d立体形象、vr模型、影片剪辑段、flash动画，gif动画图片等方案选项作为用户个人形象装扮的方案，分别设置语音开关、变声开关、情绪同步开关、文本转语音开关等各种选择开关及其他用户设置项目并配置相应的程序；

步骤3、语音开关：用户开启关闭个人形象装扮系统的语音功能；

步骤4、变声开关：切换个人形象装扮虚拟角色特定的语音和用户自己真实的语音；

步骤5、情绪同步开关：开启关闭个人形象装扮虚拟角色启用通过情绪表情模块实时表达用户情绪所匹配的表情；

步骤6、文本转语音开关：用户使用打字交流时，开启或关闭文本自动通过文本转语音模块转化成语音，在根据用户设置的相关参数和角色形象对应的声学特征参数，决定是否转化成所述形象匹配的声调，播放给用户交流的对象；

步骤7 、为每个需要用户设置的项目预设一个参数，并建立用户资料分析规则，预判用户可能偏向的选项设定为默认值，建立系统预判参数的规则库。