



(12)发明专利

(10)授权公告号 CN 104123115 B

(45)授权公告日 2017.05.24

(21)申请号 201410364822.8

(22)申请日 2014.07.28

(65)同一申请的已公布的文献号  
申请公布号 CN 104123115 A

(43)申请公布日 2014.10.29

(73)专利权人 联想(北京)有限公司  
地址 100085 北京市海淀区上地创业路6号

(72)发明人 高扬

(74)专利代理机构 北京同达信恒知识产权代理  
有限公司 11291  
代理人 黄志华

(51)Int.Cl.  
G06F 3/16(2006.01)

(56)对比文件

CN 103035247 A,2013.04.10,  
CN 103218763 A,2013.07.24,  
CN 103700371 A,2014.04.02,  
JP 2014092777 A,2014.05.19,

审查员 刘念

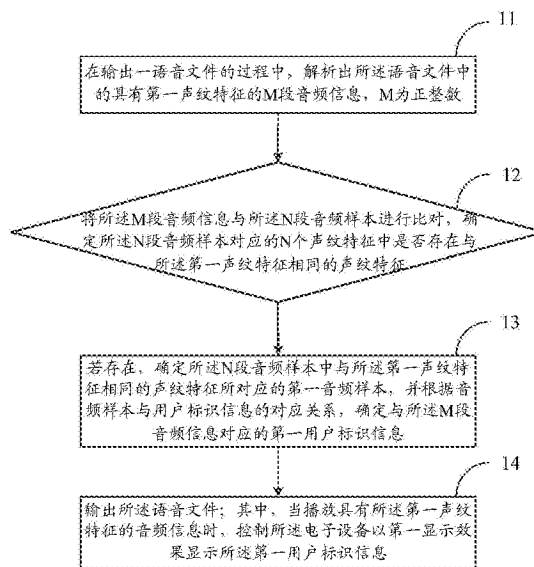
权利要求书3页 说明书12页 附图2页

(54)发明名称

一种音频信息处理方法及电子设备

(57)摘要

本发明公开了一种音频信息处理方法,用于解决现有技术中电子设备显示效果较差的技术问题。该方法包括:在输出一语音文件的过程中,解析出所述语音文件中的具有第一声纹特征的M段音频信息;将所述M段音频信息与所述N段音频样本进行对比,确定所述N段音频样本中与所述第一声纹特征相同的声纹特征所对应的第一音频样本,并根据音频样本与用户标识信息的对应关系,确定与所述M段音频信息对应的第一用户标识信息;输出所述语音文件;其中,当播放具有所述第一声纹特征的音频信息时,控制所述电子设备以第一显示效果显示所述第一用户标识信息。本发明还公开了用于实现该方法的电子设备。



1. 一种音频信息处理方法,应用于电子设备中,所述电子设备中存储有N段音频样本,所述N段音频样本中的每段音频样本分别对应一个用户标识信息,所述用户标识信息包含能够用于表征与音频信息对应的音频对象的信息,N为正整数,所述方法包括:

在输出一语音文件的过程中,解析出所述语音文件中的具有第一声纹特征的M段音频信息,M为正整数;

将所述M段音频信息与所述N段音频样本进行比对,确定所述N段音频样本对应的N个声纹特征中是否存在与所述第一声纹特征相同的声纹特征;

若存在,确定所述N段音频样本中与所述第一声纹特征相同的声纹特征所对应的第一音频样本,并根据音频样本与用户标识信息的对应关系,确定与所述M段音频信息对应的第一用户标识信息;

输出所述语音文件;其中,当播放具有所述第一声纹特征的音频信息时,控制所述电子设备以第一显示效果显示所述第一用户标识信息;

将所述M段音频信息与所述N段音频样本进行比对,确定所述N段音频样本对应的N个声纹特征中是否存在与所述第一声纹特征相同的声纹特征,还包括:

若所述N段音频样本对应的N个声纹特征中不存在与所述第一声纹特征相同的声纹特征,判断所述M段音频信息是否为关键音频信息;其中,所述关键音频信息为与所述电子设备中存储的联系对象相关的音频信息;

若所述M段音频信息为所述关键音频信息,根据所述联系对象建立与所述M段音频信息对应的用户标识信息;或者

若所述M段音频信息不为所述关键音频信息,设置第一特定标识信息作为与所述M段音频信息对应的用户标识信息;其中,所述第一特定标识信息为所述电子设备中特定图像信息、特定文字信息及特定语音信息中任一信息或组合信息。

2. 如权利要求1所述的方法,其特征在于,所述方法还包括:

检测到所述语音文件中包含的音频信息段中同时具有第二声纹特征及第三声纹特征时,根据所述第二声纹特征及所述第三声纹特征从所述音频信息段中分离出具有所述第二声纹特征的第二音频信息,以及具有所述第三声纹特征的第三音频信息;

通过将所述第二音频信息及所述第三音频信息分别与所述N段音频样本进行比对,确定出与所述第二声纹特征相对应的第二音频样本,以及与第三声纹特征相对应的第三音频样本;并根据音频样本与用户标识信息的对应关系,确定出与所述第二声纹特征相对应的第二用户标识信息,以及与所述第三声纹特征相对应的第三用户标识信息;

控制所述电子设备在播放所述音频信息的过程中,同时显示所述第二用户标识信息和所述第三用户标识信息。

3. 如权利要求2所述的方法,其特征在于,控制所述电子设备在播放所述音频信息段的过程中,同时显示所述第二用户标识信息和所述第三用户标识信息,还包括:

检测具有所述第二声纹特征的音频信息对应的第二音频强度,及具有所述第三声纹特征的音频信息对应的第三音频强度;

比较所述第二音频强度和所述第三音频强度,将所述第二音频强度和所述第三音频强度中音频强度大的音频信息确定为主音频信息,及将所述第二音频强度和所述第三音频强度中音频强度小的音频信息确定为副音频信息;

根据音频强度与显示效果的对应关系,控制所述电子设备以第一显示效果显示与所述主音频信息对应的用户标识信息,及以第二显示效果显示与所述副音频信息对应的用户标识信息。

4.如权利要求1所述的方法,其特征在于,在若所述M段音频信息为所述关键音频信息,根据所述联系对象建立与所述M段音频信息对应的用户标识信息的同时或之后,所述方法还包括:

根据所述M段音频信息,获取第一音频片断;

将所述第一音频片断作为第N+1段音频样本进行存储;其中,所述第N+1段音频样本与所述M段音频信息对应于同一用户标识信息。

5.一种电子设备,所述电子设备中存储有N段音频样本,所述N段音频样本中的每段音频样本分别对应一个用户标识信息,所述用户标识信息包含能够用于表征与音频信息对应的音频对象的信息,N为正整数,所述电子设备包括:

解析模块,用于在输出一语音文件的过程中,解析出所述语音文件中的具有第一声纹特征的M段音频信息,M为正整数;

比对模块,用于将所述M段音频信息与所述N段音频样本进行比对,确定所述N段音频样本对应的N个声纹特征中是否存在与所述第一声纹特征相同的声纹特征;

第一确定模块,用于若存在,确定所述N段音频样本中与所述第一声纹特征相同的声纹特征所对应的第一音频样本,并根据音频样本与用户标识信息的对应关系,确定与所述M段音频信息对应的第一用户标识信息;

输出模块,用于输出所述语音文件;其中,当播放具有所述第一声纹特征的音频信息时,控制所述电子设备以第一显示效果显示所述第一用户标识信息;

所述电子设备还包括:

判断模块,用于若所述N段音频样本对应的N个声纹特征中不存在与所述第一声纹特征相同的声纹特征,判断所述M段音频信息是否为关键音频信息;其中,所述关键音频信息为与所述电子设备中存储的联系对象相关的音频信息;

第二处理模块,用于若所述M段音频信息为所述关键音频信息,根据所述联系对象建立与所述M段音频信息对应的用户标识信息;或者,若所述M段音频信息不为所述关键音频信息,设置第一特定标识信息作为与所述M段音频信息对应的用户标识信息;其中,所述第一特定标识信息为所述电子设备中特定图像信息、特定文字信息及特定语音信息中任一信息或组合信息。

6.如权利要求5所述的电子设备,其特征在于,所述电子设备还包括:

分离模块,用于检测到所述语音文件中包含的音频信息段中同时具有第二声纹特征及第三声纹特征时,根据所述第二声纹特征及所述第三声纹特征从所述音频信息段中分离出具有所述第二声纹特征的第二音频信息,以及具有所述第三声纹特征的第三音频信息;

第二确定模块,用于通过将所述第二音频信息及所述第三音频信息分别与所述N段音频样本进行比对,确定出与所述第二声纹特征相对应的第二音频样本,以及与第三声纹特征相对应的第三音频样本;并根据音频样本与用户标识信息的对应关系,确定出与所述第二声纹特征相对应的第二用户标识信息,以及与所述第三声纹特征相对应的第三用户标识信息;

控制模块,用于控制所述电子设备在播放所述音频信息的过程中,同时显示所述第二用户标识信息和所述第三用户标识信息。

7.如权利要求6所述的电子设备,其特征在于,所述电子设备还包括:

检测模块,用于检测具有所述第二声纹特征的音频信息对应的第二音频强度,及具有所述第三声纹特征的音频信息对应的第三音频强度;

比较模块,用于比较所述第二音频强度和所述第三音频强度,将所述第二音频强度和所述第三音频强度中音频强度大的音频信息确定为主音频信息,及将所述第二音频强度和所述第三音频强度中音频强度小的音频信息确定为副音频信息;

第一处理模块,用于根据音频强度与显示效果的对应关系,控制所述电子设备以第一显示效果显示与所述主音频信息对应的用户标识信息,及以第二显示效果显示与所述副音频信息对应的用户标识信息。

8.如权利要求5所述的电子设备,其特征在于,所述电子设备还包括:

获取模块,用于根据所述M段音频信息,获取第一音频片断;

存储模块,将所述第一音频片断作为第N+1段音频样本进行存储;其中,所述第N+1段音频样本与所述M段音频信息对应于同一用户标识信息。

## 一种音频信息处理方法及电子设备

### 技术领域

[0001] 本发明涉及计算机技术领域,特别涉及一种音频信息处理方法及电子设备。

### 背景技术

[0002] 随着科技的迅速发展和市场竞争的日益激烈,电子设备的性能及外观得到了大力提升,其中笔记本电脑以其小巧轻便、便于携带、娱乐性强等优点正受到越来越多的人的喜爱,成为学习和生活中不可缺少的一部分。用户利用电子设备可以做的事也越来越多,如:用户可以通过具有语音功能的手机或平板电脑进行通信、录音等。

[0003] 目前,大部分的电子设备均有录音功能,能够满足多种场景的录音需求,例如会议、课堂的录音等。通常,由于录音场景的复杂性,在用户使用电子设备获取录音后,会导致播放录音时不容易区分语音内容具体对应的说话者,特别是针对声音比较近似的说话者,或者收听者还不熟悉的说话者,都会造成收听时的分辨难度。例如,在会议中,用户使用电子设备对会议内容进行录音,在后期播放录音进行回顾时,若存在多人同时讨论的情况时,可能会出现播放声音很嘈杂,无法较快地区分具体是哪些参会人在说话,则收听者在听取录音的过程中还需用心辨别播放录音对应的发话人,并且为了能够快速反应与录音内容对应的发话人,可能需要反复播放录音,从而使得电子设备的负担较重,用户体验也较差。

[0004] 综上可知,现有技术中存在电子设备录音效果较差的技术问题。

### 发明内容

[0005] 本发明实施例提供一种音频信息处理方法及电子设备,用于解决电子设备录音效果较差的技术问题。

[0006] 一种音频信息处理方法,应用于电子设备中,所述电子设备中存储有N段音频样本,所述N段音频样本中的每段音频样本分别对应一个用户标识信息,所述用户标识信息包含能够用于表征与音频信息对应的音频对象的信息,N为正整数,所述方法包括:

[0007] 在输出一语音文件的过程中,解析出所述语音文件中的具有第一声纹特征的M段音频信息,M为正整数;

[0008] 将所述M段音频信息与所述N段音频样本进行比对,确定所述N段音频样本对应的N个声纹特征中是否存在与所述第一声纹特征相同的声纹特征;

[0009] 若存在,确定所述N段音频样本中与所述第一声纹特征相同的声纹特征所对应的第一音频样本,并根据音频样本与用户标识信息的对应关系,确定与所述M段音频信息对应的第一用户标识信息;

[0010] 输出所述语音文件;其中,当播放具有所述第一声纹特征的音频信息时,控制所述电子设备以第一显示效果显示所述第一用户标识信息。

[0011] 可选的,所述方法还包括:

[0012] 检测到所述语音文件中包含的音频信息段中同时具有第二声纹特征及第三声纹特征时,根据所述第二声纹特征及所述第三声纹特征从所述音频信息段中分离出具有所述

第二声纹特征的第二音频信息,以及具有所述第三声纹特征的第三音频信息;

[0013] 通过将所述第二音频信息及所述第三音频信息分别与所述N段音频样本进行比对,确定出与所述第二声纹特征相对应的第二音频样本,以及与第三声纹特征相对应的第三音频样本;并根据音频样本与用户标识信息的对应关系,确定出与所述第二声纹特征相对应第二用户标识信息,以及与所述第三声纹特征相对应的第三用户标识信息;

[0014] 控制所述电子设备在播放所述音频信息的过程中,同时显示所述第二用户标识信息和所述第三用户标识信息。

[0015] 可选的,控制所述电子设备在播放所述音频信息端的过程中,同时显示所述第二用户标识信息和所述第三用户标识信息,还包括:

[0016] 检测具有所述第二声纹特征的音频信息对应的第二音频强度,及具有所述第三声纹特征的音频信息对应的第三音频强度;

[0017] 比较所述第二音频强度和所述第三音频强度,将其中音频强度大的音频信息确定为主音频信息,及将音频强度小的音频信息确定为副音频信息;

[0018] 根据音频强度与显示效果的对应关系,控制所述电子设备以第一显示效果显示与所述主音频信息对应的用户标识信息,及以第二显示效果显示与所述副音频信息对应的用户标识信息。

[0019] 可选的,将所述M段音频信息与所述N段音频样本进行比对,确定所述N段音频样本对应的N个声纹特征中是否存在与所述第一声纹特征相同的声纹特征,还包括:

[0020] 若所述N段音频样本对应的N个声纹特征中不存在与所述第一声纹特征相同的声纹特征,判断所述M段音频信息是否为关键音频信息;其中,所述关键音频信息为与所述电子设备中存储的联系对象相关的音频信息;

[0021] 若所述M段音频信息为所述关键音频信息,根据所述联系对象建立与所述M段音频信息对应的用户标识信息;或者

[0022] 若所述M段音频信息不为所述关键音频信息,设置第一特定标识信息作为与所述M段音频信息对应的用户标识信息;其中,所述第一特定标识信息为所述电子设备中特定图像信息、特定文字信息及特定语音信息中任一信息或组合信息。

[0023] 可选的,在若所述M段音频信息为所述关键音频信息,根据所述联系对象建立与所述M段音频信息对应的用户标识信息的同时或之后,所述方法还包括:

[0024] 根据所述M段音频信息,获取第一音频片断;

[0025] 将所述第一音频片断作为第N+1段音频样本进行存储;其中,所述N+1段音频样本与所述M段音频信息对应于同一用户标识信息。

[0026] 一种电子设备,所述电子设备中存储有N段音频样本,所述N段音频样本中的每段音频样本分别对应一个用户标识信息,所述用户标识信息包含能够用于表征与音频信息对应的音频对象的信息,N为正整数,所述电子设备包括:

[0027] 解析模块,用于在输出一语音文件的过程中,解析出所述语音文件中的具有第一声纹特征的M段音频信息,M为正整数;

[0028] 比对模块,用于将所述M段音频信息与所述N段音频样本进行比对,确定所述N段音频样本对应的N个声纹特征中是否存在与所述第一声纹特征相同的声纹特征;

[0029] 第一确定模块,用于若存在,确定所述N段音频样本中与所述第一声纹特征相同的

声纹特征所对应的第一音频样本,并根据音频样本与用户标识信息的对应关系,确定与  
所述M段音频信息对应的第一用户标识信息;

[0030] 输出模块,用于输出所述语音文件;其中,当播放具有所述第一声纹特征的音频信  
息时,控制所述电子设备以第一显示效果显示所述第一用户标识信息。

[0031] 可选的,所述电子设备还包括:

[0032] 分离模块,用于检测到所述语音文件中包含的音频信息段中同时具有第二声纹特  
征及第三声纹特征时,根据所述第二声纹特征及所述第三声纹特征从所述音频信息段中分  
离出具有所述第二声纹特征的第二音频信息,以及具有所述第三声纹特征的第三音频信  
息;

[0033] 第二确定模块,用于通过将所述第二音频信息及所述第三音频信息分别与所述N  
段音频样本进行比对,确定出与所述第二声纹特征相对应的第二音频样本,以及与第三声  
纹特征相对应的第三音频样本;并根据音频样本与用户标识信息的对应关系,确定出与所  
述第二声纹特征相对应第二用户标识信息,以及与所述第三声纹特征相对应的第三用户标  
识信息;

[0034] 控制模块,用于控制所述电子设备在播放所述音频信息的过程中,同时显示所述  
第二用户标识信息和所述第三用户标识信息。

[0035] 可选的,所述电子设备还包括:

[0036] 检测模块,用于检测具有所述第二声纹特征的音频信息对应的第二音频强度,及  
具有所述第三声纹特征的音频信息对应的第三音频强度;

[0037] 比较模块,用于比较所述第二音频强度和所述第三音频强度,将其中音频强度大  
的音频信息确定为主音频信息,及将音频强度小的音频信息确定为副音频信息;

[0038] 第一处理模块,用于根据音频强度与显示效果的对应关系,控制所述电子设备以  
第一显示效果显示与所述主音频信息对应的用户标识信息,及以第二显示效果显示与所  
述副音频信息对应的用户标识信息。

[0039] 可选的,所述电子设备还包括:

[0040] 判断模块,用于若所述N段音频样本对应的N个声纹特征中不存在与所述第一声  
纹特征相同的声纹特征,判断所述M段音频信息是否为关键音频信息;其中,所述关键音  
频信息为与所述电子设备中存储的联系对象相关的音频信息;

[0041] 第二处理模块,用于若所述M段音频信息为所述关键音频信息,根据所述联系对  
象建立与所述M段音频信息对应的用户标识信息;或者,若所述M段音频信息不为所述关键  
音频信息,设置第一特定标识信息作为与所述M段音频信息对应的用户标识信息;其中,  
所述第一特定标识信息为所述电子设备中特定图像信息、特定文字信息及特定语音信息中任  
一信息或组合信息。

[0042] 可选的,所述电子设备还包括:

[0043] 获取模块,用于根据所述M段音频信息,获取第一音频片断;

[0044] 存储模块,将所述第一音频片断作为第N+1段音频样本进行存储;其中,所述N+1段  
音频样本与所述M段音频信息对应于同一用户标识信息。

[0045] 本发明实施例中,由于所述电子设备中存储的所述N段音频样本均具有对应用  
户标识信息,且每个用户标识信息包含能够用于表征与音频信息对应的音频对象的信息,因

此在输出所述语音文件时,通过解析可以获知具有所述第一声纹特征的所述M段音频信息,并根据声纹特征,将所述M段音频信息与所述N段音频样本进行比对,则可以确定具有与所述第一声纹特征相同的声纹特征的所述第一音频样本,从而根据与所述第一音频样本对应的第一用户标识信息,从而可以使得在播放具有所述第一声纹特征的音频信息时,即播放至所述M段音频信息中的任一音频信息时,均可显示所述第一用户标识信息。因此,即便播放的录音内容具有多个发话人,那么由于每个发话人对应的声纹特征均不相同,故通过确定录音内容中具有相同声纹特征的多段音频信息后,通过比对确定对应的用户标识信息后,则在播放该音频信息时,可显示对应的用户标识信息,从而可以快速的获知当前播放的语音文件中对应的音频对象,而无需用户再花多余的时间进行辨别,故增强了电子设备的录音效果,也提高了用户的体验。

### 附图说明

- [0046] 图1为本发明实施例中音频信息处理方法的主要流程图;  
[0047] 图2为本发明实施例中显示第一用户标识信息的示意图;  
[0048] 图3为本发明实施例中显示第二用户标识信息和第三用户标识信息的示意图;  
[0049] 图4为本发明实施例中电子设备的主要模块图。

### 具体实施方式

[0050] 本发明实施例公开了一种音频信息处理方法,应用于电子设备中,所述电子设备中存储有N段音频样本,所述N段音频样本中的每段音频样本分别对应一个用户标识信息,所述用户标识信息包含能够用于表征与音频信息对应的音频对象的信息,N为正整数,所述方法包括:在输出一语音文件的过程中,解析出所述语音文件中的具有第一声纹特征的M段音频信息,M为正整数;将所述M段音频信息与所述N段音频样本进行比对,确定所述N段音频样本对应的N个声纹特征中是否存在与所述第一声纹特征相同的声纹特征;若存在,确定所述N段音频样本中与所述第一声纹特征相同的声纹特征所对应的第一音频样本,并根据音频样本与用户标识信息的对应关系,确定与所述M段音频信息对应的第一用户标识信息;输出所述语音文件;其中,当播放具有所述第一声纹特征的音频信息时,控制所述电子设备以第一显示效果显示所述第一用户标识信息。

[0051] 本发明实施例中,由于所述电子设备中存储的所述N段音频样本均具有对应用户标识信息,且每个用户标识信息包含能够用于表征与音频信息对应的音频对象的信息,因此在输出所述语音文件时,通过解析可以获知具有所述第一声纹特征的所述M段音频信息,并根据声纹特征,将所述M段音频信息与所述N段音频样本进行比对,则可以确定具有与所述第一声纹特征相同的声纹特征的所述第一音频样本,从而根据与所述第一音频样本对应的第一用户标识信息,从而可以使得在播放具有所述第一声纹特征的音频信息时,即播放至所述M段音频信息中的任一音频信息时,均可显示所述第一用户标识信息。因此,即便播放的录音内容具有多个发话人,那么由于每个发话人对应的声纹特征均不相同,故通过确定录音内容中具有相同声纹特征的多段音频信息后,通过比对确定对应的用户标识信息后,则在播放该音频信息时,可显示对应的用户标识信息,从而可以快速的获知当前播放的语音文件中对应的音频对象,而无需用户再花多余的时间进行辨别,故增强了电子设备的



录音效果,也提高了用户的体验。

[0052] 请参见图1,本发明实施例公开了一种音频信息处理方法,应用于一具有显示单元的电子设备中,所述电子设备中存储有N段音频样本,所述N段音频样本中的每段音频样本分别对应一个用户标识信息,所述用户标识信息包含能够用于表征与音频信息对应的音频对象的信息,N为正整数,所述方法可以包括以下步骤:

[0053] 步骤11:在输出一语音文件的过程中,解析出所述语音文件中的具有第一声纹特征的M段音频信息,M为正整数。

[0054] 本发明实施例中,所述语音文件可以是录制的对应特殊场合的录音文件。例如,会议内容的录音文件或课堂的录音文件等。通常,所述语音文件可以是存储在本地的录音文件,例如将通过自身或其它设备录制的文件存储在本地,或者所述语音文件也可以是获取的来自其他电子设备或云端的录音文件。

[0055] 可选的,本发明实施例中,所述第一声纹特征可以是指在将所述语音文件进行输出的过程中,通过声纹识别确定的所述语音文件对应的声纹特征。

[0056] 通常,所谓声纹是指用电声学仪器显示的携带言语信息的声波频谱,且任何两个人的声纹图谱都有差异。因此,通过声纹识别,可以确定所述语音文件中各个音频信息对应的声纹特征,故可识别出具有相同声纹特征的音频信息,并且当所述语音文件是录制有多个说话人对应的发言内容时,所述语音文件可以对应有多个声纹特征。

[0057] 可选的,通过声纹识别可以确定所述语音文件中具有所述第一声纹特征的所述M段音频信息,故可以认为所述M段音频信息是来自于同一说话人的说话内容,且所述M段音频信息可以是处于所述语音文件中不同音频位置。例如,当该说话人与多个其它说话人处于同一场景下时,该说话人时不时的进行发言,则与其对应的所述M段音频信息是按照发言时间顺序被录制进所述语音文件,从而在播放所述语音文件时,会将录制的包括所有说话人的说话内容按照录制顺序进行播放,此时,所述M段音频信息就可能是穿插在所述语音文件中的多个位置的音频信息。

[0058] 步骤12:将所述M段音频信息与所述N段音频样本进行比对,确定所述N段音频样本对应的N个声纹特征中是否存在与所述第一声纹特征相同的声纹特征。

[0059] 本发明实施例中,由于每个人对应的声纹特征均不相同,因此在确定出所述M段音频信息,且将所述M段音频信息与所述N段音频样本进行比对时,可以通过声纹识别技术来进行判断,若能检测到与所述第一声纹特征相同的声纹特征,则说明所述N个声纹特征中存在与所述第一声纹特征相同的声纹特征,即具有与所述M段音频信息匹配的音频样本,否则,则不存在与所述M段音频信息对应的音频样本,故也不能通过当前存储的所述N段音频样本确定与所述M段音频信息对应的音频对象。

[0060] 本发明实施例中,所述N段音频样本可以是预先根据一个或多个录音文件设置的。例如,从预先录制或存储的录音文件中提取出与相关联系人对应的音频信息作为音频样本,或者,也可以根据联系人为其录制对应的音频片段作为与该联系人对应的音频样本。其中,所述N段音频样本中的每一段音频样本均可来自语音段的音频信息。例如,从语音文件的语音段中获取多个音频信息。

[0061] 可选的,本发明实施例中,所述N段音频样本中的每段音频样本分别对应一个用户标识信息,所述用户标识信息可以是包含能够用于表征与音频信息对应的音频对象的信

息。例如,所述用户标识信息可以包含联系人头像、姓名、工作性质等信息。

[0062] 步骤13:若存在,确定所述N段音频样本中与所述第一声纹特征相同的声纹特征所对应的第一音频样本,并根据音频样本与用户标识信息的对应关系,确定与所述M段音频信息对应的第一用户标识信息。

[0063] 本发明实施例中,由于每个音频信息具有不同的在通过声纹识别技术确定出与所述第一声纹特征相同的所述第一音频样本时,可以进一步确定与所述第一音频样本对应的所述第一用户标识信息,从而可以确定与所述M段音频信息对应的音频对象。

[0064] 可选的,本发明实施例中,音频样本与用户标识信息之间的对应关系可以是用户预先设置好的。例如,用户可以在设置所述N段音频样本时,将与每个音频样本相关的信息设置为与该音频样本对应的用户标识信息。例如,将与该段音频样本对应的音频对象的头像、名字等信息中的一个或组合确定为对应的用户标识信息。

[0065] 例如,用户手机中存储有说话人甲的第一音频样本,说话人甲的说话声音对应声纹特征1,且与所述第一音频样本对应的第一用户标识信息中包含说话人甲的头像信息、姓名信息,则当用户使用手机播放一个录音文件时,若该录音文件中包含有说话人甲的声音,则在播放该录音文件时,若确定该录音文件中识别出的声纹特征中具有与声纹特征1相同的声纹特征,则可以认为录音文件中具有声纹特征1的音频信息均为说话人甲对应的音频信息,故这些音频信息均可以为与第一用户标识信息相关联的音频信息。

[0066] 在实际操作过程中,在将所述M段音频信息与所述N段音频样本进行比对,确定所述N段音频样本对应的N个声纹特征中是否存在与所述第一声纹特征相同的声纹特征时,还可以包括:若所述N段音频样本对应的N个声纹特征中不存在与所述第一声纹特征相同的声纹特征,判断所述M段音频信息是否为关键音频信息;其中,所述关键音频信息为与所述电子设备中存储的联系对象相关的音频信息;若所述M段音频信息为所述关键音频信息,根据所述联系对象建立与所述M段音频信息对应的用户标识信息;或者,若所述M段音频信息不为所述关键音频信息,设置第一特定标识信息作为与所述M段音频信息对应的用户标识信息;其中,所述第一特定标识信息为所述电子设备中特定图像信息、特定文字信息及特定语音信息中任一信息或组合信息。

[0067] 其中,判断所述M段音频信息是否为所述关键音频信息,可以具有以下两种实现判断的方法。

[0068] 第一种:通过用户进行判断。该过程可以是根据所述电子设备中存储的联系对象进行确定的,若存储的所述联系对象时未存储有对应的音频段,则以上所述判断过程可以由用户来进行实现的。例如,在播放的所述语音文件时,若确定播放的音频信息为未匹配成功的音频信息,则用户可以根据自己对联系人对应的声音的熟悉程度,辨别该音频信息是否为联系人对应的声音信息,若是,则可以将该音频信息确定为所述关键音频信息,否则,可以不对该段音频信息进行过多的设置。故通过用户自身来进行判断时,可以具有较大的自主选择性,提高了用户的体验度,同时也使得所述电子设备的录音效果具有较强的灵活性。

[0069] 第二种,通过电子设备进行判断。如果在存储所述联系对象的时候,也存储有与所述联系对象对应的音频信息,则判断所述M段音频信息是否为所述关键音频信息可以通过所述电子设备通过声纹识别及匹配来实现的。例如,若用户在建立所述联系对象的信息

的同时或之后,还为联系对象存储了一段对应的语音,从而在所述第一声纹特征与所述N段声纹特征未匹配成功时,可以通过将所述第一声纹特征与所述联系对象的语音段对应的声纹特征进行匹配,从而可以确定所述第一声纹特征是否与所述联系对象相关,进而确定所述M段音频信息是否为所述关键音频信息。

[0070] 本发明实施例中,若判断结果表明所述M段音频信息为所述关键音频信息,则可以根据所述联系对象建立与所述M段音频信息对应的用户标识信息。通常,用户在存储联系对象时,会包含相关的对象姓名、头像、工作单位等信息,若确定所述M段音频信息对应的所述联系对象为联系对象1时,则可将头像信息及姓名信息设置为与所述M段音频信息设置对应的用户标识信息所包含的内容。

[0071] 此外,在通过某些没有设置相应头像的联系对象建立与所述M段音频信息对应的用户标识信息时,可通过从本地或云端获取与该联系对象相关的图像进行设置,以便通过该用户标识信息可以快速进行辨别。例如,使用手机中存储的、与确定的联系对象相关的图像设置所述用户标识信息中的头像信息时,则可将该图像包含的头像部分进行截图,从而设置为该联系对象的头像信息,提高辨别度。

[0072] 或者,若判断表面所述M段音频信息不为所述关键音频信息,则可以设置第一特定标识信息作为与所述M段音频信息对应的用户标识信息;其中,所述第一特定标识信息为所述电子设备中特定图像信息、特定文字信息及特定语音信息中任一信息或组合信息。

[0073] 其中,所述特定图像可以是指所述电子设备默认或用户预先指定的、用于为声纹匹配不成功的音频信息设置的对应的图像,并且可以为该图像设置对应的文字信息,例如“未识别”、“未知”等。或者,所述特定图像也可以就是具有易识别的标识或图像,而不需要搭配特殊的文字,例如可以显示为未知人物头像的图像,使得用户一看便知此时播放的音频信息为与联系人不相关的信息。

[0074] 可选的,本发明实施例中,在若所述M段音频信息为所述关键音频信息,根据所述联系对象建立与所述M段音频信息对应的用户标识信息的同时或之后,所述方法还可以包括:根据所述M段音频信息,获取第一音频片断;将所述第一音频片断作为第N+1段音频样本进行存储;其中,所述N+1段音频样本与所述M段音频信息对应于同一用户标识信息。即在确定所述M段音频信息为所述关键音频信息时,可在所述M段音频信息中截取任意一个音频片断作为所述第一音频片断,并将所述第一音频片断作为所述第N+1段音频样本进行存储,从而不断地增加音频样本的数量,以使在进行声纹匹配时具有更多的可进行比对的声纹特征,以能够尽量多的识别出所述语音文件中的不同声纹特征对应的用户标识信息,从而获知对应的音频对象等,提高了所述电子设备对录音文件分析的准确性。

[0075] 步骤14:输出所述语音文件;其中,当播放具有所述第一声纹特征的音频信息时,控制所述电子设备以第一显示效果显示所述第一用户标识信息。

[0076] 本发明实施例中,在确定出所述语言文件中具有相同声纹特征的音频信息后,即可确定与所述音频信息对应的用户标识信息。从而在播放所述语音文件时,若通过声纹识别确定当前播放的音频信息在所述N段音频样本中具有对应的音频样本,则可将与具有该声纹特征的音频信息均显示相同的用户标识信息。例如,与该音频信息对应的音频对象的头像信息、姓名信息等。

[0077] 请参加图2,数字20代表所述电子设备,此处以手机为例;数字21代表所述电子设

备的显示单元,所述显示单元中正在播放所述语音文件,且当前播放的音频为所述M段音频信息中的任意一段,数字22代表所述用户标识信息,此处以用户头像信息为例,其中,标号为1的用户标识信息代表所述第一用户标识信息,其余用户标识表示与所述语音文件所包含的其他声纹特征对应的用户标识信息。

[0078] 本发明实施例中,所述音频信息处理方法还可以包括:检测到所述语音文件中包含的段音频信息段中同时具有第二声纹特征及第三声纹特征时,根据所述第二声纹特征及所述第三声纹特征的特征参数,从所述音频信息段中分离出具有所述第二声纹特征的第二音频信息,以及具有所述第三声纹特征的第三音频信息;通过将所述第二音频信息及所述第三音频信息分别与所述N段音频样本进行比对,确定出与所述第二声纹特征相对应的第二音频样本,以及与第三声纹特征相对应的第三音频样本;根据音频样本与用户标识信息的对应关系,确定出与所述第二声纹特征相对应第二用户标识信息,以及与所述第三声纹特征相对应的第三用户标识信息;控制所述电子设备在播放所述音频信息段的过程中,同时显示所述第二用户标识信息和所述第三用户标识信息。

[0079] 其中,所述音频信息段可以是指所述语音文件中的同时包含多段音频信息的语音段。例如,单位时间内,播放所述语音文件时可能同时包含多个说话人的发言内容,则根据每个人对应的音频信息可以确定出多个声纹特征。所述第二声纹特征和所述第三声纹特征可以是指不同说话对象的音频信息各自对应的声纹特征。

[0080] 在确定所述语音文件中包含的一段音频信息中同时存在所述第二声纹特征和所述第三声纹特征后,可根据所述第二声纹特征及所述第三声纹特征的特征参数对所述音频信息段进行提取,从而分离出具有所述第二声纹特征的第二音频信息,以及具有所述第三声纹特征的第三音频信息。其中,所述参数特征可以是声纹频谱中共振峰的频率值。一般来说,声纹频谱中共振峰的频率值及其走向是最稳定的特征参数,而且具有很强的特定性,而时长、音强、波形等特征参数稳定性较差,也可做参考。

[0081] 可选的,本发明实施例中,在确定出与所述第二声纹特征相对应第二用户标识信息,以及与所述第三声纹特征相对应的第三用户标识信息后,则在播放所述音频信息的过程中,可以同时显示所述第二用户标识信息和所述第三用户标识信息,以使得听者获知当前多人同时说话时对应的多人头像。例如,所述语音文件中包含同时发言的说话人甲和说话人乙的音频信息段1,则当播放至该音频信息段时,与说话人甲对应的头像a和说话人乙对应的头像b将同时进行显示,以表示当前播放的音频信息段为这两个头像分别对应的音频对象对应的声音。

[0082] 请参见图3,数字30代表所述电子设备,此处以手机为例;数字31代表所述电子设备的显示单元,所述显示单元中正在播放所述音频信息段,且所述音频信息段同时包含有所所述第二声纹特征对应的第二音频信息和所述第三声纹特征的第三音频信息,数字1和数字2分别代表所述第二用户标识信息和所述第三用户标识信息,且所述第二用户标识信息和所述第三用户标识信息相对于其它用户标识信息的状态为处于放大的状态,表示当前正在播放与所述第二用户标识信息和所述第三用户标识信息对应的音频信息。

[0083] 可选的,本发明实施例中,所述控制所述电子设备在播放所述音频信息段的过程中,同时显示所述第二用户标识信息和所述第三用户标识信息,还可以包括:检测具有所述第二声纹特征的音频信息对应的第二音频强度,及具有所述第三声纹特征的音频信息对应

的第三音频强度;比较所述第二音频强度和所述第三音频强度,将其中音频强度大的音频信息确定为主音频信息,及将音频强度小的音频信息确定为副音频信息;根据音频强度与显示效果的对应关系,控制所述电子设备以第一显示效果显示与所述主音频信息对应的用户标识信息,及以第二显示效果显示与所述副音频信息对应的用户标识信息。

[0084] 即在播放所述音频信息段时,由于同时显示有所述第二用户标识信息和所述第三用户标识信息,为了便于区分具体用户标识信息对应的音频信息,可以根据音频信息对应的音频强度确定对应的用户标识的显示效果。

[0085] 例如,与音频强度大的音频信息对应显示效果可以是用户标识信息以高频率进行跳动,而与音频强度小的音频信息对应显示效果可以是用户标识信息以低频率进行跳动,从而通过观察用户标识的跳动频率,可以将用户标识信息与说话人的声音强弱程度联系起来,从而在播放具有多人同时说话的音频信息段时,可以使得听者通过声音的响亮程度及用户标识信息的跳动频率来区分声音对应的用户标识信息,而避免了同时播放的录音文件中同时含有多个声音时导致不易区分的情况。

[0086] 请参见图4,基于同一发明构思,本发明实施例还提供一种电子设备,所述电子设备中存储有N段音频样本,所述N段音频样本中的每段音频样本分别对应一个用户标识信息,所述用户标识信息包含能够用于表征与音频信息对应的音频对象的信息,N为正整数,所述电子设备可以包括解析模块401、比对模块402、第一确定模块403和输出模块404。

[0087] 所述解析模块401可以用于在输出一语音文件的过程中,解析出所述语音文件中的具有第一声纹特征的M段音频信息,M为正整数。

[0088] 所述比对模块402可以用于将所述M段音频信息与所述N段音频样本进行比对,确定所述N段音频样本对应的N个声纹特征中是否存在与所述第一声纹特征相同的声纹特征。

[0089] 说是第一确定模块403可以用于若存在,确定所述N段音频样本中与所述第一声纹特征相同的声纹特征所对应的第一音频样本,并根据音频样本与用户标识信息的对应关系,确定与所述M段音频信息对应的第一用户标识信息;

[0090] 所述输出模块404可以用于输出所述语音文件;其中,当播放具有所述第一声纹特征的音频信息时,控制所述电子设备以第一显示效果显示所述第一用户标识信息。

[0091] 可选的,本发明实施例中,所述电子设备还包括:

[0092] 分离模块,用于检测到所述语音文件中包含的音频信息段中同时具有第二声纹特征及第三声纹特征时,根据所述第二声纹特征及所述第三声纹特征从所述音频信息段中分离出具有所述第二声纹特征的第二音频信息,以及具有所述第三声纹特征的第三音频信息;

[0093] 第二确定模块,用于通过将所述第二音频信息及所述第三音频信息分别与所述N段音频样本进行比对,确定出与所述第二声纹特征相对应的第二音频样本,以及与第三声纹特征相对应的第三音频样本;并根据音频样本与用户标识信息的对应关系,确定出与所述第二声纹特征相对应第二用户标识信息,以及与所述第三声纹特征相对应的第三用户标识信息;

[0094] 控制模块,用于控制所述电子设备在播放所述音频信息的过程中,同时显示所述第二用户标识信息和所述第三用户标识信息。

[0095] 可选的,本发明实施例中,所述电子设备还包括:

[0096] 检测模块,用于检测具有所述第二声纹特征的音频信息对应的第二音频强度,及具有所述第三声纹特征的音频信息对应的第三音频强度;

[0097] 比较模块,用于比较所述第二音频强度和所述第三音频强度,将其中音频强度大的音频信息确定为主音频信息,及将音频强度小的音频信息确定为副音频信息;

[0098] 第一处理模块,用于根据音频强度与显示效果的对应关系,控制所述电子设备以第一显示效果显示与所述主音频信息对应的用户标识信息,及以第二显示效果显示与所述副音频信息对应的用户标识信息。

[0099] 可选的,本发明实施例中,所述电子设备还包括:

[0100] 判断模块,用于若所述N段音频样本对应的N个声纹特征中不存在与所述第一声纹特征相同的声纹特征,判断所述M段音频信息是否为关键音频信息;其中,所述关键音频信息为与所述电子设备中存储的联系对象相关的音频信息;

[0101] 第二处理模块,用于若所述M段音频信息为所述关键音频信息,根据所述联系对象建立与所述M段音频信息对应的用户标识信息;或者,若所述M段音频信息不为所述关键音频信息,设置第一特定标识信息作为与所述M段音频信息对应的用户标识信息;其中,所述第一特定标识信息为所述电子设备中特定图像信息、特定文字信息及特定语音信息中任一信息或组合信息。

[0102] 可选的,本发明实施例中,所述电子设备还包括:

[0103] 获取模块,用于根据所述M段音频信息,获取第一音频片断;

[0104] 存储模块,将所述第一音频片断作为第N+1段音频样本进行存储;其中,所述N+1段音频样本与所述M段音频信息对应于同一用户标识信息。

[0105] 本发明实施例公开了一种音频信息处理方法,应用于电子设备中,所述电子设备中存储有N段音频样本,所述N段音频样本中的每段音频样本分别对应一个用户标识信息,所述用户标识信息包含能够用于表征与音频信息对应的音频对象的信息,N为正整数,所述方法包括:在输出一语音文件的过程中,解析出所述语音文件中的具有第一声纹特征的M段音频信息,M为正整数;将所述M段音频信息与所述N段音频样本进行比对,确定所述N段音频样本对应的N个声纹特征中是否存在与所述第一声纹特征相同的声纹特征;若存在,确定所述N段音频样本中与所述第一声纹特征相同的声纹特征所对应的第一音频样本,并根据音频样本与用户标识信息的对应关系,确定与所述M段音频信息对应的第一用户标识信息;输出所述语音文件;其中,当播放具有所述第一声纹特征的音频信息时,控制所述电子设备以第一显示效果显示所述第一用户标识信息。

[0106] 本发明实施例中,由于所述电子设备中存储的所述N段音频样本均具有对应用户标识信息,且每个用户标识信息包含能够用于表征与音频信息对应的音频对象的信息,因此在输出所述语音文件时,通过解析可以获知具有所述第一声纹特征的所述M段音频信息,并根据声纹特征,将所述M段音频信息与所述N段音频样本进行比对,则可以确定具有与所述第一声纹特征相同的声纹特征的所述第一音频样本,从而根据与所述第一音频样本对应的第一用户标识信息,从而可以使得在播放具有所述第一声纹特征的音频信息时,即播放至所述M段音频信息中的任一音频信息时,均可显示所述第一用户标识信息。因此,即便播放的录音内容具有多个发话人,那么由于每个发话人对应的声纹特征均不相同,故通过确定录音内容中具有相同声纹特征的多段音频信息后,通过比对确定对应的用户标识信息

后,则在播放该音频信息时,可显示对应的用户标识信息,从而可以快速的获知当前播放的语音文件中对应的音频对象,而无需用户再花多余的时间进行辨别,故增强了电子设备的录音效果,也提高了用户的体验。

[0107] 具体来讲,本申请实施例中的信息处理方法对应的计算机程序指令可以被存储在光盘,硬盘,U盘等存储介质上,当存储介质中的与音频信息处理方法对应的计算机程序指令被一电子设备读取或被执行时,包括如下步骤:

[0108] 在输出一语音文件的过程中,解析出所述语音文件中的具有第一声纹特征的M段音频信息,M为正整数;

[0109] 将所述M段音频信息与所述N段音频样本进行比对,确定所述N段音频样本对应的N个声纹特征中是否存在与所述第一声纹特征相同的声纹特征;

[0110] 若存在,确定所述N段音频样本中与所述第一声纹特征相同的声纹特征所对应的第一音频样本,并根据音频样本与用户标识信息的对应关系,确定与所述M段音频信息对应的第一用户标识信息;

[0111] 输出所述语音文件;其中,当播放具有所述第一声纹特征的音频信息时,控制所述电子设备以第一显示效果显示所述第一用户标识信息。

[0112] 可选的,所述存储介质中还存储有另外一些计算机指令,这些计算机指令用于执行步骤:检测到所述语音文件中包含的音频信息段中同时具有第二声纹特征及第三声纹特征时,根据所述第二声纹特征及所述第三声纹特征从所述音频信息段中分离出具有所述第二声纹特征的第二音频信息,以及具有所述第三声纹特征的第三音频信息;

[0113] 通过将所述第二音频信息及所述第三音频信息分别与所述N段音频样本进行比对,确定出与所述第二声纹特征相对应的第二音频样本,以及与第三声纹特征相对应的第三音频样本;并根据音频样本与用户标识信息的对应关系,确定出与所述第二声纹特征相对应第二用户标识信息,以及与所述第三声纹特征相对应的第三用户标识信息;

[0114] 控制所述电子设备在播放所述音频信息的过程中,同时显示所述第二用户标识信息和所述第三用户标识信息。

[0115] 可选的,所述存储介质中存储的与步骤在控制所述电子设备在播放所述音频信息端,同时显示所述第二用户标识信息和所述第三用户标识信息对应的计算机指令在具体被执行过程中,还包括如下步骤:

[0116] 检测具有所述第二声纹特征的音频信息对应的第二音频强度,及具有所述第三声纹特征的音频信息对应的第三音频强度;

[0117] 比较所述第二声音强度和所述第三声音强度,将其中声音强度大的音频信息确定为主音频信息,及将声音强度小的音频信息确定为副音频信息;

[0118] 根据声音强度与显示效果的对应关系,控制所述电子设备以第一显示效果显示与所述主音频信息对应的用户标识信息,及以第二显示效果显示与所述副音频信息对应的用户标识信息。

[0119] 可选的,所述存储介质中存储的与步骤在将所述M段音频信息与所述N段音频样本进行比对,确定所述N段音频样本对应的N个声纹特征中是否存在与所述第一声纹特征相同的声纹特征对应的计算机指令在具体被执行过程中,还包括如下步骤:

[0120] 若所述N段音频样本对应的N个声纹特征中不存在与所述第一声纹特征相同的声

纹特征,判断所述M段音频信息是否为关键音频信息;其中,所述关键音频信息为与所述电子设备中存储的联系对象相关的音频信息;

[0121] 若所述M段音频信息为所述关键音频信息,根据所述联系对象建立与所述M段音频信息对应的用户标识信息;或者

[0122] 若所述M段音频信息不为所述关键音频信息,设置第一特定标识信息作为与所述M段音频信息对应的用户标识信息;其中,所述第一特定标识信息为所述电子设备中特定图像信息、特定文字信息及特定语音信息中任一信息或组合信息。

[0123] 可选的,所述存储介质中还存储有另外一些计算机指令,这些计算机指令在与步骤:若所述M段音频信息为所述关键音频信息,根据所述联系对象建立与所述M段音频信息对应的用户标识信息对应的计算机指令在具体被执行被执行的同时或之后被执行,在被执行时包括如下步骤:

[0124] 根据所述M段音频信息,获取第一音频片断;

[0125] 将所述第一音频片断作为第N+1段音频样本进行存储;其中,所述N+1段音频样本与所述M段音频信息对应于同一用户标识信息。

[0126] 显然,本领域的技术人员可以对本发明进行各种改动和变型而不脱离本发明的精神和范围。这样,倘若本发明的这些修改和变型属于本发明权利要求及其等同技术的范围之内,则本发明也意图包含这些改动和变型在内。



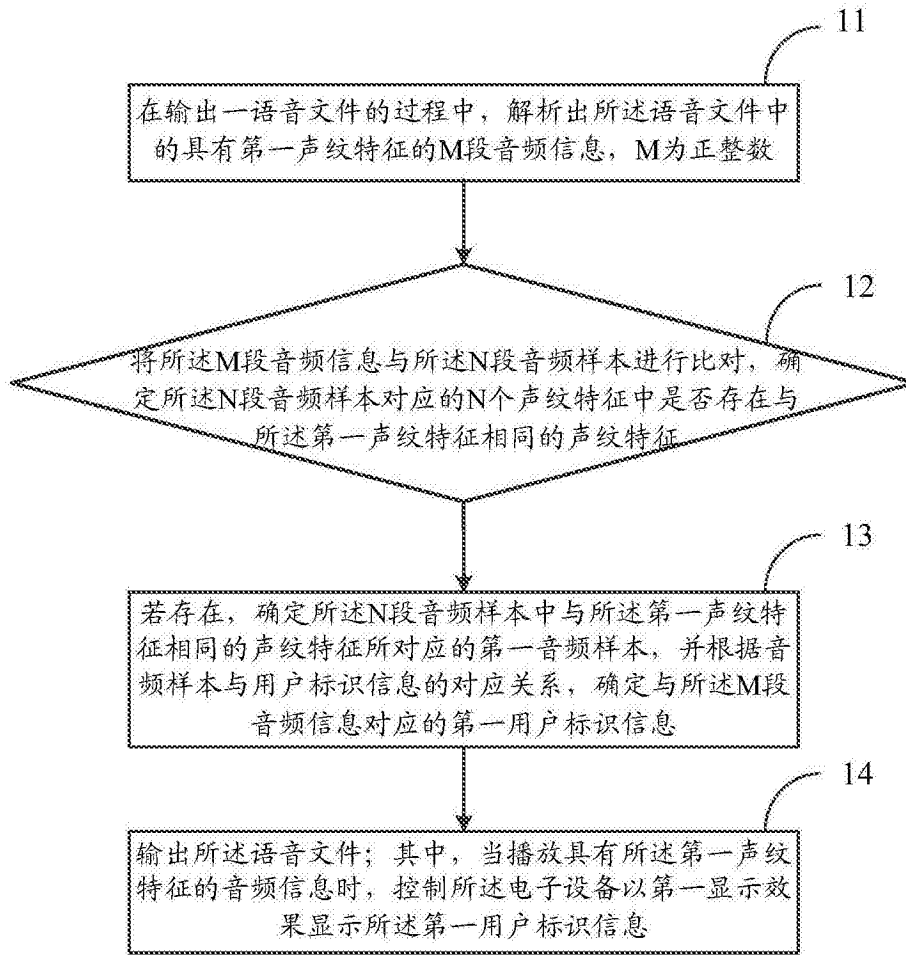


图1

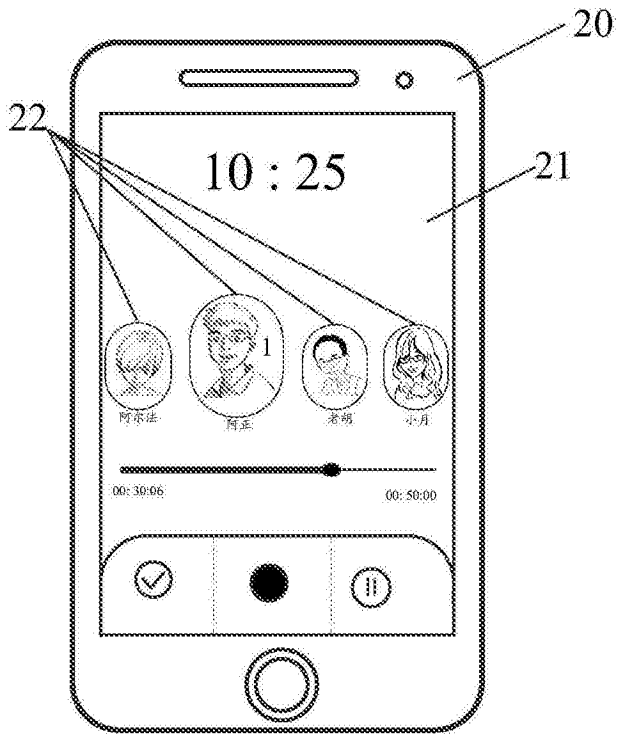


图2

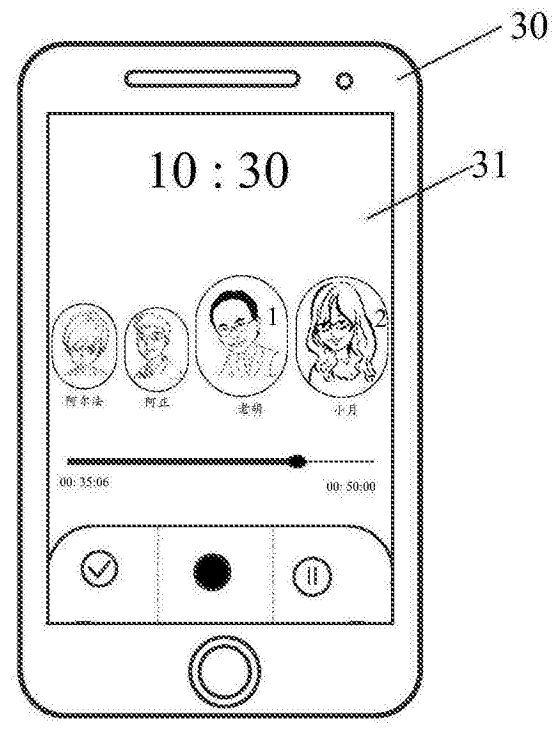


图3

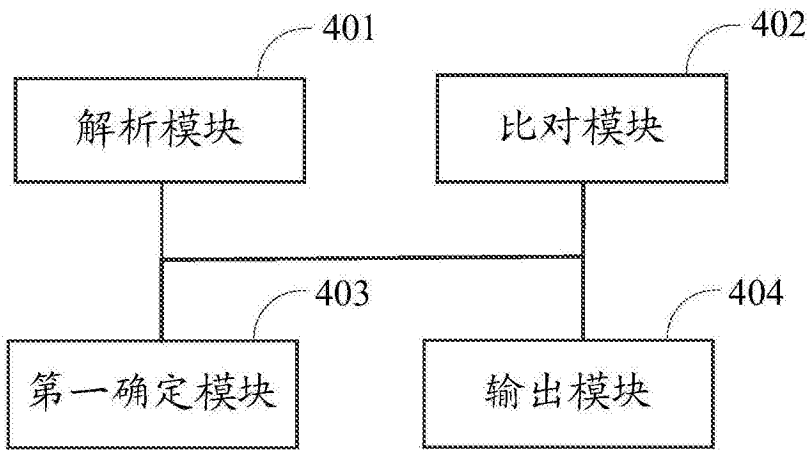


图4