



(12) 发明专利

(10) 授权公告号 CN 110648670 B

(45) 授权公告日 2021. 11. 26

(21) 申请号 201911006499.6

(22) 申请日 2019.10.22

(65) 同一申请的已公布的文献号
申请公布号 CN 110648670 A

(43) 申请公布日 2020.01.03

(73) 专利权人 中信银行股份有限公司
地址 100010 北京市东城区朝阳门北大街9号

(72) 发明人 赖勇铨 贺亚运 林春 李美玲

(74) 专利代理机构 北京市兰台律师事务所
11354

代理人 张峰

(51) Int. Cl.

G10L 17/02 (2013.01)

G10L 17/04 (2013.01)

G10L 17/18 (2013.01)

G10L 17/00 (2013.01)

G06K 9/62 (2006.01)

G06Q 40/02 (2012.01)

(56) 对比文件

CN 108417226 A, 2018.08.17

CN 110718228 A, 2020.01.21

CN 109493882 A, 2019.03.19

CN 106157135 A, 2016.11.23

CN 110223165 A, 2019.09.10

CN 102231277 A, 2011.11.02

CN 1283843 A, 2001.02.14

US 2021037136 A1, 2021.02.04

US 10003688 B1, 2018.06.19

EP 2808866 A1, 2014.12.03

刘弘胤. AI赋能下的声纹识别技术在公共安全领域的深度应用.《中国安防》. 2019, 于嫻. 声纹识别在微信中的模式匹配研究.《中国优秀硕士学位论文全文数据库 信息科技辑》. 2016,

Jian Hua. Voiceprint identification based on model clustering.《2013 3rd International Conference on Consumer Electronics, Communications and Networks》. 2014,

审查员 孙筱逸

权利要求书2页 说明书12页 附图3页

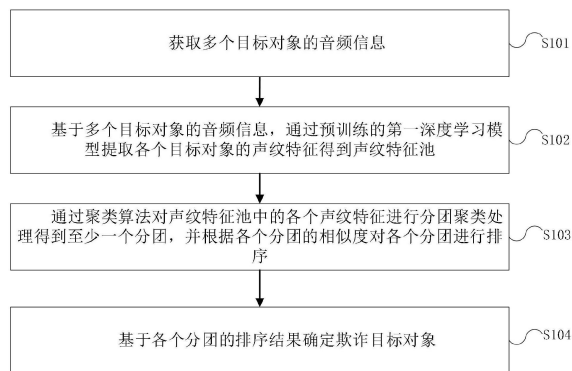
(54) 发明名称

欺诈识别方法、装置、电子设备及计算机可读存储介质

(57) 摘要

本申请提供了一种欺诈识别方法、装置、电子设备及计算机可读存储介质,应用于音频处理技术领域,其中该方法包括:通过聚类算法对声纹特征池中的各个声纹特征进行分团聚类处理得到至少一个分团,并根据各个分团的相似度对各个分团进行排序,然后根据排序结果确定欺诈识别对象,解决了如何识别一人冒充多人身份进行欺诈申请的问题,实现了欺诈识别的自动化处理;此外,基于当前申请人的声纹特征与历史申请人的声纹特征的聚类分析,解决了历史申请中欺诈申请的识别问题,提升了欺诈申请识别的准

确度。



CN 110648670 B

1. 一种欺诈识别方法,其特征在于,包括:

获取多个目标对象的音频信息;

基于所述多个目标对象的音频信息,通过预训练的第一深度学习模型提取各个所述目标对象的声纹特征得到声纹特征池;

通过聚类算法对所述声纹特征池中的各个声纹特征进行分团聚类处理得到至少一个分团,并以分团作为整体计算各个分团分别对应的相似度值,根据各个分团的相似度对所述各个分团进行排序;

基于各个分团的排序结果确定对应分团中包含的目标对象为欺诈目标对象;

所述分团聚类处理包括:步骤1、从特征向量池中随机不放回的抽取一个未成团的特征点作为一个团的初始成员;步骤2、依次检测特征向量池中每个特征点与这个团里所有成员的平均距离,如果距离小于指定阈值则加入到此团中来,如果某个特征点和其它所有其它特征点的距离都大于指定阈值,则此特征点单独成团;步骤3、之后再从特征向量池中随机不放回抽取一个未成团的特征点,重复1,2过程,直到特征向量池中所有成员都已成团,聚类结束。

2. 根据权利要求1所述的方法,其特征在于,所述获取多个目标对象的音频信息,之前包括:

获取所述目标对象与相应客服之间的通话语音信息;

基于相应的语音分离算法将所述目标对象与相应客服的通话语音信息进行语音分离;

基于所述目标对象与相应客服之间的通话语音信息语音分离后的音频,确定所述目标对象的音频信息。

3. 根据权利要求2所述的方法,其特征在于,所述基于相应的语音分离算法将所述目标对象与相应客服的通话语音信息进行语音分离,包括以下至少一项:

通过卷积神经网络与循环神经网络,将各个所述目标对象与相应客服的通话语音信息进行语音分离;

通过滑窗法与K-Means聚类算法,将各个所述目标对象与相应客服的通话语音信息进行语音分离。

4. 根据权利要求3所述的方法,其特征在于,所述通过卷积神经网络与循环神经网络,将各个所述目标对象与相应客服的通话语音信息进行语音分离,包括:

获取所述目标对象与相应客服的通话语音信息对应的声谱图;

基于预训练的卷积神经网络提取所述声谱图的三维特征,所述三维特征包括时间维度特征、频率维度特征、通道维度特征;

对所述三维特征在频率维度进行平均池化处理,得到池化处理后的特征;

将所述池化处理后的特征输入至预训练的循环神经网络,得到时间维度上的分割标签;

基于所述时间维度上的分割标签对所述目标通话音频进行语音分离。

5. 根据权利要求3所述的方法,其特征在于,所述通过滑窗法与K-Means聚类算法,将各个所述目标对象与相应客服的通话语音信息进行语音分离,包括:

通过滑窗法基于预训练的第二深度学习模型提取所述目标对象与相应客服的通话语音信息的语音特征,得到特征向量数组;

对所述特征向量数组进行K-Means聚类,得到各个特征向量的标签;

基于所述各个特征向量的标签,将各个所述目标对象与相应客服的通话语音信息进行语音分离。

6. 根据权利要求2-5任一项所述的方法,其特征在于,所述基于所述目标对象与相应客服之间的通话语音信息语音分离后的音频,确定所述目标对象的音频信息,包括:

提取所述目标对象与相应客服之间的通话语音信息语音分离后的音频的声纹特征;

计算语音分离后的音频的声纹特征与预存储的至少一个客服的声纹特征的相似度;

基于相似度计算结果,确定所述目标对象的音频信息。

7. 一种欺诈识别装置,其特征在于,包括:

第一获取模块,用于获取多个目标对象的音频信息;

提取模块,用于基于所述多个目标对象的音频信息,通过预训练的第一深度学习模型提取各个所述目标对象的声纹特征得到声纹特征池;

排序模块,用于通过聚类算法对所述声纹特征池中的各个声纹特征进行分团聚类处理得到至少一个分团,并以分团作为整体计算各个分团分别对应的相似度值,根据各个分团的相似度对所述各个分团进行排序;所述分团聚类处理包括:步骤1、从特征向量池中随机不放回的抽取一个未成团的特征点作为一个团的初始成员;步骤2、依次检测特征向量池中每个特征点与这个团里所有成员的平均距离,如果距离小于指定阈值则加入到此团中来,如果某个特征点和其它所有其它特征点的距离都大于指定阈值,则此特征点单独成团;步骤3、之后再从特征向量池中随机不放回抽取一个未成团的特征点,重复1,2过程,直到特征向量池中所有成员都已成团,聚类结束;

第一确定模块,用于基于各个分团的排序结果确定对应分团中包含的目标对象为欺诈目标对象。

8. 一种电子设备,其特征在于,包括:

一个或多个处理器;

存储器;

一个或多个应用程序,其中所述一个或多个应用程序被存储在所述存储器中并被配置为由所述一个或多个处理器执行,所述一个或多个程序配置用于:执行根据权利要求1至6任一项所述的欺诈识别方法。

9. 一种计算机可读存储介质,其特征在于,所述计算机存储介质用于存储计算机指令,当其在计算机上运行时,使得计算机可以执行上述权利要求1至6中任一项所述的欺诈识别方法。

欺诈识别方法、装置、电子设备及计算机可读存储介质

技术领域

[0001] 本申请涉及音频处理技术领域,具体而言,本申请涉及一种欺诈识别方法、装置、电子设备及计算机可读存储介质。

背景技术

[0002] 随着我国信用卡业务的发展和发卡量的大量增长,信用卡的风险也日益凸显,信用卡坏账率呈上升走势,其中欺诈申请造成的坏账占比很大,如何有效识别这些欺诈申请成为了一个问题。

[0003] 目前,信用卡欺诈申请是通过身份认证的方式实现的,即业务人员对信用卡申请人员提供的信息进行审核,如果提供的是虚假信息,则会被认定为欺诈申请。然而,现有的通过身份认证识别欺诈申请的方式,不能识别出一人冒充多人身份进行信用卡的欺诈申请情形。

发明内容

[0004] 本申请提供了一种欺诈识别方法、装置、电子设备及计算机可读存储介质,用于解决一人冒充多人身份的欺诈申请识别问题,以及历史申请中欺诈申请的识别问题本申请采用的技术方案如下:

[0005] 第一方面,提供了一种欺诈识别方法,该方法包括,

[0006] 获取多个目标对象的音频信息;

[0007] 基于多个目标对象的音频信息,通过预训练的第一深度学习模型提取各个目标对象的声纹特征得到声纹特征池;

[0008] 通过聚类算法对声纹特征池中的各个声纹特征进行分团聚类处理得到至少一个分团,并根据各个分团的相似度对各个分团进行排序;

[0009] 基于各个分团的排序结果确定欺诈目标对象。

[0010] 其中,聚类算法为贪心聚类算法。

[0011] 可选地,获取多个目标对象的音频信息,之前包括:

[0012] 获取目标对象与相应客服之间的通话语音信息;

[0013] 基于相应的语音分离算法将目标对象与相应客服的通话语音信息进行语音分离;

[0014] 基于目标对象与相应客服之间的通话语音信息语音分离后的音频,确定目标对象的音频信息。

[0015] 可选地,基于相应的语音分离算法将目标对象与相应客服的通话语音信息进行语音分离,包括以下至少一项:

[0016] 通过卷积神经网络与循环神经网络,将各个目标对象与相应客服的通话语音信息进行语音分离;

[0017] 通过滑窗法与K-Means聚类算法,将各个目标对象与相应客服的通话语音信息进行语音分离。

- [0018] 可选地,通过卷积神经网络与循环神经网络,将各个目标对象与相应客服的通话语音信息进行语音分离,包括:
- [0019] 获取目标对象与相应客服的通话语音信息对应的声谱图;
- [0020] 基于预训练的卷积神经网络提取声谱图的三维特征,三维特征包括时间维度特征、频率维度特征、通道维度特征;
- [0021] 对三维特征在频率维度进行平均池化处理,得到池化处理后的特征;
- [0022] 将池化处理后的特征输入至预训练的循环神经网络,得到时间维度上的分割标签;
- [0023] 基于时间维度上的分割标签对目标通话音频进行语音分离。
- [0024] 可选地,通过滑窗法与K-Means聚类算法,将各个目标对象与相应客服的通话语音信息进行语音分离,包括:
- [0025] 通过滑窗法基于预训练的第二深度学习模型提取目标对象与相应客服的通话语音信息的语音特征,得到特征向量数组;
- [0026] 对特征向量数组进行K-Means聚类,得到各个特征向量的标签;
- [0027] 基于各个特征向量的标签,将各个目标对象与相应客服的通话语音信息进行语音分离。
- [0028] 可选地,基于目标对象与相应客服之间的通话语音信息语音分离后的音频,确定目标对象的音频信息,包括:
- [0029] 提取目标对象与相应客服之间的通话语音信息语音分离后的音频的声纹特征;
- [0030] 计算语音分离后的音频的声纹特征与预存储的至少一个客服的声纹特征的相似度;
- [0031] 基于相似度计算结果,确定目标对象的音频信息。
- [0032] 第二方面,提供了一种欺诈识别装置,该装置包括,
- [0033] 第一获取模块,用于获取多个目标对象的音频信息;
- [0034] 提取模块,用于基于多个目标对象的音频信息,通过预训练的第一深度学习模型提取各个目标对象的声纹特征得到声纹特征池;
- [0035] 排序模块,用于通过聚类算法对声纹特征池中的各个声纹特征进行分团聚类处理得到至少一个分团,并根据各个分团的相似度对各个分团进行排序;
- [0036] 第一确定模块,用于基于各个分团的排序结果确定欺诈目标对象。
- [0037] 其中,聚类算法为贪心聚类算法。
- [0038] 可选地,该装置还包括:
- [0039] 第二获取模块,用于获取目标对象与相应客服之间的通话语音信息;
- [0040] 语音分离模块,用于基于相应的语音分离算法将目标对象与相应客服的通话语音信息进行语音分离;
- [0041] 第二确定模块,用于基于目标对象与相应客服之间的通话语音信息语音分离后的音频,确定目标对象的音频信息。
- [0042] 可选地,语音分离模块,用于通过卷积神经网络与循环神经网络,将各个目标对象与相应客服的通话语音信息进行语音分离;和/或用于通过滑窗法与K-Means聚类算法,将各个目标对象与相应客服的通话语音信息进行语音分离。

- [0043] 可选地,语音分离模块包括:
- [0044] 获取单元,用于获取目标对象与相应客服的通话语音信息对应的声谱图;
- [0045] 第一提取单元,用于基于预训练的卷积神经网络提取声谱图的三维特征,三维特征包括时间维度特征、频率维度特征、通道维度特征;
- [0046] 池化单元,用于对三维特征在频率维度进行平均池化处理,得到池化处理后的特征;
- [0047] 第一标签单元,用于将池化处理后的特征输入至预训练的循环神经网络,得到时间维度上的分割标签;
- [0048] 第一分离单元,用于基于时间维度上的分割标签对目标通话音频进行语音分离。
- [0049] 可选地,语音分离模块包括:
- [0050] 第二提取单元,通过滑窗法基于预训练的第二深度学习模型提取目标对象与相应客服的通话语音信息的语音特征,得到特征向量数组;
- [0051] 第二标签单元,用于对特征向量数组进行K-Means聚类,得到各个特征向量的标签;
- [0052] 第二分离单元,用于基于各个特征向量的标签,将各个目标对象与相应客服的通话语音信息进行语音分离。
- [0053] 可选地,第二确定模块包括:
- [0054] 第三提取单元,用于提取目标对象与相应客服之间的通话语音信息语音分离后的音频的声纹特征;
- [0055] 计算单元,用于计算语音分离后的音频的声纹特征与预存储的至少一个客服的声纹特征的相似度;
- [0056] 第三确定单元,用于基于相似度计算结果,确定目标对象的音频信息。
- [0057] 第三方面,提供了一种电子设备,该电子设备包括:
- [0058] 一个或多个处理器;
- [0059] 存储器;
- [0060] 一个或多个应用程序,其中一个或多个应用程序被存储在存储器中并被配置为由一个或多个处理器执行,一个或多个程序配置用于:执行第一方面所示的欺诈识别方法。
- [0061] 第四方面,提供了一种计算机可读存储介质,计算机存储介质用于存储计算机指令,当其在计算机上运行时,使得计算机可以执行第一方面所示的欺诈识别方法。
- [0062] 本申请提供了一种欺诈识别方法、装置、电子设备及计算机可读存储介质,与现有技术通过身份认证的方式实现欺诈申请的识别相比,本申请通过获取多个目标对象的音频信息,并基于多个目标对象的音频信息,通过预训练的第一深度学习模型提取各个目标对象的声纹特征得到声纹特征池,继而通过聚类算法对声纹特征池中的各个声纹特征进行分团聚类处理得到至少一个分团,并根据各个分团的相似度对各个分团进行排序,最后基于各个分团的排序结果确定欺诈目标对象。即通过聚类算法对声纹特征池中的各个声纹特征进行分团聚类处理得到至少一个分团,并根据各个分团的相似度对各个分团进行排序,然后根据排序结果确定欺诈识别对象,解决了如何识别一人冒充多人身份进行欺诈申请的问题,实现了欺诈识别的自动化处理;此外,基于当前申请人的声纹特征与历史申请人的声纹特征的聚类分析,解决了历史申请中欺诈申请的识别问题,提升了欺诈申请识别的准确度。

[0063] 本申请附加的方面和优点将在下面的描述中部分给出,这些将从下面的描述中变得明显,或通过本申请的实践了解到。

附图说明

[0064] 本申请上述的和/或附加的方面和优点从下面结合附图对实施例的描述中将变得明显和容易理解,其中:

[0065] 图1为本申请实施例的一种欺诈识别方法的流程示意图;

[0066] 图2为本申请实施例的一种欺诈识别装置的结构示意图;

[0067] 图3为本申请实施例的另一种欺诈识别装置的结构示意图;

[0068] 图4为本申请实施例的一种电子设备的结构示意图。

具体实施方式

[0069] 下面详细描述本申请的实施例,各实施例的示例在附图中示出,其中自始至终相同或类似的标号表示相同或类似的元件或具有相同或类似功能的元件。下面通过参考附图描述的实施例是示例性的,仅用于解释本申请,而不能解释为对本申请的限制。

[0070] 本技术领域技术人员可以理解,除非特意声明,这里使用的单数形式“一”、“一个”和“该”也可包括复数形式。应该进一步理解的是,本申请的说明书中使用的措辞“包括”是指存在特征、整数、步骤、操作、元件和/或组件,但是并不排除存在或添加一个或多个其他特征、整数、步骤、操作、元件、组件和/或它们的组。这里使用的措辞“和/或”包括一个或多个相关联的列出项的全部或任一单元和全部组合。

[0071] 为使本申请的目的、技术方案和优点更加清楚,下面将结合附图对本申请实施方式作进一步地详细描述。

[0072] 下面以具体地实施例对本申请的技术方案以及本申请的技术方案如何解决上述技术问题进行详细说明。下面这几个具体的实施例可以相互结合,对于相同或相似的概念或过程可能在某些实施例中不再赘述。下面将结合附图,对本申请的实施例进行描述。

[0073] 本申请实施例提供了一种欺诈识别方法,如图1所示,该方法可以包括以下步骤:

[0074] 步骤S101,获取多个目标对象的音频信息;

[0075] 具体地,获取多个目标对象的音频信息,其中,该目标对象可以为信用卡申请人,该音频信息为信用卡申请人进行信用卡申请时,与客服的对话信息,但该音频信息仅包括信用卡申请人的音频,不包括客服的音频部分。

[0076] 步骤S102,基于多个目标对象的音频信息,通过预训练的第一深度学习模型提取各个目标对象的声纹特征得到声纹特征池;

[0077] 具体地,通过预训练的第一深度学习模型提取各个目标对象的声纹特征得到声纹特征池,其中,可以对该多个目标对象的音频信息进行标准化处理,并通过滑动窗口法以及FFT变换得到多个目标对象的声谱图,基于得到的多个声谱图,通过预训练的第一深度学习模型提取各个目标对象的声纹特征得到声纹特征池。

[0078] 步骤S103,通过聚类算法对声纹特征池中的各个声纹特征进行分团聚类处理得到至少一个分团,并根据各个分团的相似度对各个分团进行排序;

[0079] 具体地,通过相应的聚类算法对声纹特征池中的各个声纹特征进行分团聚类处理

得到至少一个分团,其中,该聚类算法可以是贪心聚类算法、K-Means聚类算法、层次聚类算法等;聚类处理后,可以根据计算得到的各个分团分别对应的相似度值,对各个分团进行排序,其中,可以根据相似度值由高到低进行排序。

[0080] 步骤S104,基于各个分团的排序结果确定欺诈目标对象。

[0081] 示例性地,各个分团按照相似度由高到低进行排序,可以设定排名靠前的N各分团中包含的申请人为欺诈申请人,也可以设定相似度超过预定阈值的分团中包含的申请人为欺诈申请人,其中,分团中相似度越高,说明该分团中一人冒充多人进行欺诈申请的概率越高。

[0082] 本申请实施例提供了一种欺诈识别方法,与现有技术通过身份认证的方式实现欺诈申请的识别相比,本申请实施例通过获取多个目标对象的音频信息,并基于多个目标对象的音频信息,通过预训练的第一深度学习模型提取各个目标对象的声纹特征得到声纹特征池,继而通过聚类算法对声纹特征池中的各个声纹特征进行分团聚类处理得到至少一个分团,并根据各个分团的相似度对各个分团进行排序,最后基于各个分团的排序结果确定欺诈目标对象。即通过聚类算法对声纹特征池中的各个声纹特征进行分团聚类处理得到至少一个分团,并根据各个分团的相似度对各个分团进行排序,然后根据排序结果确定欺诈识别对象,解决了如何识别一人冒充多人身份进行欺诈申请的问题,实现了欺诈识别的自动化处理;此外,基于当前申请人的声纹特征与历史申请人的声纹特征的聚类分析,解决了历史申请中欺诈申请的识别问题,提升了欺诈申请识别的准确度。

[0083] 本申请实施例提供了一种可能的实现方式,其中,聚类算法为贪心聚类算法。

[0084] 具体地,聚类算法为贪心聚类算法,多个目标申请人的声纹特征组成了一个特征向量池,使用贪心算法进行特征点的分团和聚类。其中,运用贪心算法进行聚类处理的步骤如下:

[0085] 步骤1、从特征向量池中随机不放回的抽取一个未成团的特征点作为一个团的初始成员;步骤2、依次检测特征向量池中每个特征点与这个团里所有成员的平均距离,如果距离小于指定阈值则加入到此团中来,如果某个特征点和其它所有其它特征点的距离都大于指定阈值,则此特征点单独成团;步骤3、之后再从特征向量池中随机不放回抽取一个未成团的特征点,重复1,2过程,直到特征向量池中所有成员都已成团,聚类结束。

[0086] 对于本申请实施例,运用贪心聚类算法进行声纹特征的聚类处理,相比于K-Means或谱聚类等算法,计算复杂度小,且不用事先指定类别数,操作更加方便。

[0087] 本申请实施例提供了一种可能的实现方式,进一步地,步骤S101之前包括:

[0088] 步骤105(图中未示出),获取目标对象与相应客服之间的通话语音信息;

[0089] 具体地,可以是在目标对象进行信用卡申请时,通过手机等终端设备的音频获取装置,获取目标对象与相应客服之间的通话语音信息。其中,对该通话语音信息,可以使用VAD进行静音切除,之后通过预训练能够识别音乐片段的深度学习模型,将通话语音信息中的手机铃声等音乐部分切除。

[0090] 步骤106(图中未示出),基于相应的语音分离算法将目标对象与相应客服的通话语音信息进行语音分离;

[0091] 具体地,通过相应的语音分割算法将目标对象与相应客服的通话语音信息进行语音分离;其中,语音分离技术是在没有任何先验知识的情况下,从一段音频资料中区分不

同说话者的说话区段,并且一一标注出来。

[0092] 步骤107(图中未示出),基于目标对象与相应客服之间的通话语音信息语音分离后的音频,确定目标对象的音频信息。

[0093] 示例性地,对目标对象与相应客服之间的通话语音信息进行语音分离后得到两种音频,其中一种代表目标对象的音频,另一种代表客服的音频,基于该两种音频确定出何种音频为目标对象的音频。

[0094] 对于本申请实施例,解决了目标对象的音频信息的确定问题。

[0095] 本申请实施例提供了一种可能的实现方式,步骤S106(图中未示出)包括:

[0096] 步骤S1061(图中未示出),通过卷积神经网络与循环神经网络,将各个目标对象与相应客服的通话语音信息进行语音分离;

[0097] 步骤S1062(图中未示出),通过滑窗法与K-Means聚类算法,将各个目标对象与相应客服的通话语音信息进行语音分离。

[0098] 具体地,相应的语音分离可以通过通过卷积神经网络与循环神经网络的方法,或者通过滑窗法与K-Means聚类算法实现的。

[0099] 对于本申请实施例,解决了目标对象与相应客服之间的通话语音信息的语音分离问题。

[0100] 本申请实施例提供了一种可能的实现方式,具体地,步骤S1061(图中未示出)包括:

[0101] 步骤S10611(图中未示出),获取目标对象与相应客服的通话语音信息对应的声谱图;

[0102] 通过相应的数据获取方法获取目标对象与相应客服的通话语音信息对应的声谱图,其中,该声谱图可以通过对目标对象与相应客服的通话语音信息进行标准化处理、滑动窗口处理以及FFT处理得到的。

[0103] 步骤S10612(图中未示出),基于预训练的卷积神经网络提取声谱图的三维特征,三维特征包括时间维度特征、频率维度特征、通道维度特征;

[0104] 具体地,通过预训练的卷积神经网络(如ResNet-18、VGG-16等)提取声谱图的三维特征;其中,可以将二维声谱图(f, t)输入卷积神经网络得到三维的特征图(f, t, c),其中, f 为频率维度, t 为时间维度, c 为channel(通道)维度,通道维度与卷积层的卷积核数量有关。

[0105] 步骤S10613(图中未示出),对三维特征在频率维度进行平均池化处理,得到池化处理后的特征;

[0106] 具体地,对三维特征在频率维度进行平均池化处理,得到池化处理后的特征,从而将三维的特征图转为二维特征图,同时保留了时间上的维度长度,而二维特征图的另一个维度尺寸,则由卷积网络输出特征图的卷积核数量决定,从而可以处理不同长度的通话语音。

[0107] 示例性地, $channel=6, f=28, t=28$,将三维特征图转为二维特征图的过程可以是:沿频率 f 方向做mean pooling,也就是沿频率方向取各个时间点上28个数值的平均值,从而得到一个一维的28位数的向量;然后,将6个channel处理得到的一维向量堆叠在一起,得到最终的二维特征(6行28列)。

[0108] 步骤S10614(图中未示出),将池化处理后的特征输入至预训练的循环神经网络,得到时间维度上的分割标签;

[0109] 具体地,将池化处理后的特征输入预训练的循环神经网络,得到时间维度上的分割标签,采用循环神经网络可以对卷积神经网络提取的时间维度上的特征做序列分析,通过关联前后一段时间的输入来计算,可以有效地准确输出相应时间点上的说话者标签,进而可以通过这些标签达到说话者语音分离的目的;其中,该循环神经网络可以是LSTM网络(Long Short Memory Network,长短时记忆网络),也可以是能够实现本申请的其他循环神经网络,本申请此处不做限定。

[0110] 步骤S10615(图中未示出),基于时间维度上的分割标签对目标通话音频进行语音分离。

[0111] 具体地,基于时间维度上的分割标签对目标通话音频进行语音分离,可以基于分割标签与相应的音频段之间的映射关系,将目标通话音频进行语音分离。

[0112] 对于本申请实施例,提取目标对象与相应客服的通话语音信息对应的声谱图的三维特征,充分利用了说话者的信息,能够提升语音切分的准确性;此外,通过预训练的循环神经网络自动输出目标通话音频的分割标签,对转换点的判断更加精准,进而提升语音切分的准确性,再者,不需要另外的聚类算法进行聚类处理,从而能够实现语音分离的端到端处理。

[0113] 本申请实施例提供了一种可能的实现方式,具体地,步骤S1062(图中未示出)包括:

[0114] 步骤S10621(图中未示出)通过滑窗法基于预训练的第二深度学习模型提取目标对象与相应客服的通话语音信息的语音特征,得到特征向量数组;

[0115] 具体地,使用滑窗法(窗体长为n秒,步长为s秒)在音频数据上使用训练好的深度学习模型进行特征提取,得到特征向量数组。滑窗在滑动过程中会处理目标对象与客服各自的语音片段,同时滑窗也会划过两个人语音的重叠区(重叠意为滑窗的窗体同时覆盖了两个人的语音),滑窗的滑动在特征空间中表现为特征点在两个特征团之间的往复运动。

[0116] 步骤S10622(图中未示出)对特征向量数组进行K-Means聚类,得到各个特征向量的标签;

[0117] 步骤S10623(图中未示出)基于各个特征向量的标签,将各个目标对象与相应客服的通话语音信息进行语音分离。

[0118] 具体地,对提取的特征向量进行K-means聚类,特征点被打上0或1的标签,随着特征点的移动,在0和1转变的交界点就是不同人语音的分割点。此时的分割点对应的是一个滑窗的窗体,即为一段长度为n秒的音频,取这段n秒音频的中点作为真正的分割点,有了分割点数据后,便可以同方法1一样去提取出客户音频。

[0119] 对于本申请实施例,通过K-means聚类算法,解决了目标对象与客服的通话语音的语音分离问题。

[0120] 本申请实施例提供了一种可能的实现方式,具体地,步骤S107(图中未示出)包括:

[0121] 步骤S1071(图中未示出)提取目标对象与相应客服之间的通话语音信息语音分离后的音频的声纹特征;

[0122] 步骤S1072(图中未示出)计算语音分离后的音频的声纹特征与预存储的至少一个

客服的声纹特征的相似度；

[0123] 步骤S1073(图中未示出)基于相似度计算结果,确定目标对象的音频信息。

[0124] 具体地,目标对象与相应客服的通话语音信息语音分离后,得到两个人(目标对象、客服)的音频,通过相应的声纹特征提取方法(如深度学习的方法),提取进行分离处理后的音频的声纹特征。

[0125] 具体地,当预存储的只有一名客服的声纹特征时,如果两个人中其中一人的声纹特征与预存储的该客服的声纹特征的相似度小于预定的阈值时,则该其中一人为目标对象(即信用卡申请人的音频);如果其中一人的声纹特征与预存储的该客服的声纹特征的相似度大于预定的阈值时,则该其中一人的音频为客服的音频,另一人对应的音频为目标对象(即信用卡申请人)的音频。

[0126] 具体地,当预存储有多名客服的声纹特征时,分别计算其中一人的声纹特征与预存储的各个客服的声纹特征的相似度,得到多个相似度值;如果有任一相似度值大于预定的阈值,则说明该其中一人的音频为客服的音频;如果所有的相似度值皆小于预定的阈值,则说明该其中一人对应的音频为目标对象(即信用卡申请者)的音频。

[0127] 对于本申请实施例,解决了目标对象的音频的确定问题。

[0128] 图2为本申请实施例提供的一种欺诈识别装置,该装置20包括:第一获取模块201、提取模块202、排序模块203以及第一确定模块204,其中,

[0129] 第一获取模块201,用于获取多个目标对象的音频信息;

[0130] 提取模块202,用于基于多个目标对象的音频信息,通过预训练的第一深度学习模型提取各个目标对象的声纹特征得到声纹特征池;

[0131] 排序模块203,用于通过聚类算法对声纹特征池中的各个声纹特征进行分团聚类处理得到至少一个分团,并根据各个分团的相似度对各个分团进行排序;

[0132] 第一确定模块204,用于基于各个分团的排序结果确定欺诈目标对象。

[0133] 本申请实施例提供了一种欺诈识别装置,与现有技术通过身份认证的方式实现欺诈申请的识别相比,本申请实施例通过获取多个目标对象的音频信息,并基于多个目标对象的音频信息,通过预训练的第一深度学习模型提取各个目标对象的声纹特征得到声纹特征池,继而通过聚类算法对声纹特征池中的各个声纹特征进行分团聚类处理得到至少一个分团,并根据各个分团的相似度对各个分团进行排序,最后基于各个分团的排序结果确定欺诈目标对象。即通过聚类算法对声纹特征池中的各个声纹特征进行分团聚类处理得到至少一个分团,并根据各个分团的相似度对各个分团进行排序,然后根据排序结果确定欺诈识别对象,解决了如何识别一人冒充多人身份进行欺诈申请的问题,实现了欺诈识别的自动化处理;此外,基于当前申请人的声纹特征与历史申请人的声纹特征的聚类分析,解决了历史申请中欺诈申请的识别问题,提升了欺诈申请识别的准确度。

[0134] 本实施例的欺诈识别装置可执行本申请上述实施例中提供的一种欺诈识别方法,其实现原理相类似,此处不再赘述。

[0135] 如图3所示,本申请实施例提供了另一种欺诈识别装置,该装置30包括:第一获取模块301、提取模块302、排序模块303以及第一确定模块304,其中,

[0136] 第一获取模块301,用于获取多个目标对象的音频信息;

[0137] 其中,图3中的第一获取模块301与图2中的第一获取模块201的功能相同或者相

似。

[0138] 提取模块302,用于基于多个目标对象的音频信息,通过预训练的第一深度学习模型提取各个目标对象的声纹特征得到声纹特征池;

[0139] 其中,图3中的提取模块302与图2中的提取模块202的功能相同或者相似。

[0140] 排序模块303,用于通过聚类算法对声纹特征池中的各个声纹特征进行分团聚类处理得到至少一个分团,并根据各个分团的相似度对各个分团进行排序;

[0141] 其中,图3中的排序模块303与图2中的排序模块203的功能相同或者相似。

[0142] 第一确定模块304,用于基于各个分团的排序结果确定欺诈目标对象。

[0143] 其中,图3中的第一确定模块304与图2中的第一确定模块204的功能相同或者相似。

[0144] 本申请实施例提供了一种可能的实现方式,其中,聚类算法为贪心聚类算法。

[0145] 本申请实施例提供了一种可能的实现方式,该装置30还包括:

[0146] 第二获取模块305,用于获取目标对象与相应客服之间的通话语音信息;

[0147] 语音分离模块306,用于基于相应的语音分离算法将目标对象与相应客服的通话语音信息进行语音分离;

[0148] 第二确定模块307,用于基于目标对象与相应客服之间的通话语音信息语音分离后的音频,确定目标对象的音频信息。

[0149] 对于本申请实施例,解决了目标对象的音频信息的确定问题。

[0150] 本申请实施例提供了一种可能的实现方式,具体地,语音分离模块306,用于通过卷积神经网络与循环神经网络,将各个目标对象与相应客服的通话语音信息进行语音分离;和/或用于通过滑窗法与K-Means聚类算法,将各个目标对象与相应客服的通话语音信息进行语音分离。

[0151] 对于本申请实施例,解决了目标对象与相应客服之间的通话语音信息的语音分离问题。

[0152] 本申请实施例提供了一种可能的实现方式,具体地,语音分离模块306包括:

[0153] 获取单元3061(图中未示出),用于获取目标对象与相应客服的通话语音信息对应的声谱图;

[0154] 第一提取单元获取单元3062(图中未示出),用于基于预训练的卷积神经网络提取声谱图的三维特征,三维特征包括时间维度特征、频率维度特征、通道维度特征;

[0155] 池化单元获取单元3063(图中未示出),用于对三维特征在频率维度进行平均池化处理,得到池化处理后的特征;

[0156] 第一标签单元获取单元3064(图中未示出),用于将池化处理后的特征输入至预训练的循环神经网络,得到时间维度上的分割标签;

[0157] 第一分离单元获取单元3065(图中未示出),用于基于时间维度上的分割标签对目标通话音频进行语音分离。

[0158] 对于本申请实施例,提取目标对象与相应客服的通话语音信息对应的声谱图的三维特征,充分利用了说话者的信息,能够提升语音切分的准确性;此外,通过预训练的循环神经网络自动输出目标通话音频的分割标签,对转换点的判断更加精准,进而提升语音切分的准确性,再者,不需要另外的聚类算法进行聚类处理,从而能够实现语音分离的端到端

处理。

[0159] 本申请实施例提供了一种可能的实现方式,具体地,语音分离模块306包括:

[0160] 第二提取单元3066(图中未示出),通过滑窗法基于预训练的第二深度学习模型提取目标对象与相应客服的通话语音信息的语音特征,得到特征向量数组;

[0161] 第二标签单元3067(图中未示出),用于对特征向量数组进行K-Means聚类,得到各个特征向量的标签;

[0162] 第二分离单元3068(图中未示出),用于基于各个特征向量的标签,将各个目标对象与相应客服的通话语音信息进行语音分离。

[0163] 对于本申请实施例,通过K-means聚类算法,解决了目标对象与客服的通话语音的语音分离问题。

[0164] 本申请实施例提供了一种可能的实现方式,具体地,第二确定模块305包括:

[0165] 第三提取单元3051(图中未示出),用于提取目标对象与相应客服之间的通话语音信息语音分离后的音频的声纹特征;

[0166] 计算单元3052(图中未示出),用于计算语音分离后的音频的声纹特征与预存储的至少一个客服的声纹特征的相似度;

[0167] 第三确定单元3053(图中未示出),用于基于相似度计算结果,确定目标对象的音频信息。

[0168] 对于本申请实施例,解决了目标对象的音频的确定问题。

[0169] 本申请实施例提供了一种欺诈识别装置,与现有技术通过身份认证的方式实现欺诈申请的识别相比,本申请实施例通过获取多个目标对象的音频信息,并基于多个目标对象的音频信息,通过预训练的第一深度学习模型提取各个目标对象的声纹特征得到声纹特征池,继而通过聚类算法对声纹特征池中的各个声纹特征进行分团聚类处理得到至少一个分团,并根据各个分团的相似度对各个分团进行排序,最后基于各个分团的排序结果确定欺诈目标对象。即通过聚类算法对声纹特征池中的各个声纹特征进行分团聚类处理得到至少一个分团,并根据各个分团的相似度对各个分团进行排序,然后根据排序结果确定欺诈识别对象,解决了如何识别一人冒充多人身份进行欺诈申请的问题,实现了欺诈识别的自动化处理;此外,基于当前申请人的声纹特征与历史申请人的声纹特征的聚类分析,解决了历史申请中欺诈申请的识别问题,提升了欺诈申请识别的准确度。

[0170] 本申请实施例提供了一种欺诈识别装置,适用于上述实施例所示的方法,在此不再赘述。

[0171] 本申请实施例提供了一种电子设备,如图4所示,图4所示的电子设备40包括:处理器401和存储器403。其中,处理器401和存储器403相连,如通过总线402相连。进一步地,电子设备40还可以包括收发器404。需要说明的是,实际应用中收发器404不限于一个,该电子设备40的结构并不构成对本申请实施例的限定。其中,处理器401应用于本申请实施例中,用于实现图2或图3所示的第一获取模块、提取模块、排序模块、第一确定模块的功能,以及图3所示的第二获取模块、语音分离模块、第二确定模块的功能。收发器404包括接收机和发射机。

[0172] 处理器401可以是CPU,通用处理器,DSP,ASIC,FPGA或者其他可编程逻辑器件、晶体管逻辑器件、硬件部件或者其任意组合。其可以实现或执行结合本申请公开内容所描述

的各种示例性的逻辑方框,模块和电路。处理器401也可以是实现计算功能的组合,例如包含一个或多个微处理器组合,DSP和微处理器的组合等。

[0173] 总线402可包括一通路,在上述组件之间传送信息。总线402可以是PCI总线或EISA总线等。总线402可以分为地址总线、数据总线、控制总线等。为便于表示,图4中仅用一条粗线表示,但并不表示仅有一根总线或一种类型的总线。

[0174] 存储器403可以是ROM或可存储静态信息和指令的其他类型的静态存储设备,RAM或者可存储信息和指令的其他类型的动态存储设备,也可以是EEPROM、CD-ROM或其他光盘存储、光碟存储(包括压缩光碟、激光碟、光碟、数字通用光碟、蓝光光碟等)、磁盘存储介质或者其他磁存储设备、或者能够用于携带或存储具有指令或数据结构形式的期望的程序代码并能够由计算机存取的任何其他介质,但不限于此。

[0175] 存储器403用于存储执行本申请方案的应用程序代码,并由处理器401来控制执行。处理器401用于执行存储器403中存储的应用程序代码,以实现图2或图3所示实施例提供的欺诈识别装置的功能。

[0176] 本申请实施例提供了一种电子设备,与现有技术通过身份认证的方式实现欺诈申请的识别相比,本申请实施例通过获取多个目标对象的音频信息,并基于多个目标对象的音频信息,通过预训练的第一深度学习模型提取各个目标对象的声纹特征得到声纹特征池,继而通过聚类算法对声纹特征池中的各个声纹特征进行分团聚类处理得到至少一个分团,并根据各个分团的相似度对各个分团进行排序,最后基于各个分团的排序结果确定欺诈目标对象。即通过聚类算法对声纹特征池中的各个声纹特征进行分团聚类处理得到至少一个分团,并根据各个分团的相似度对各个分团进行排序,然后根据排序结果确定欺诈识别对象,解决了如何识别一人冒充多人身份进行欺诈申请的问题,实现了欺诈识别的自动化处理;此外,基于当前申请人的声纹特征与历史申请人的声纹特征的聚类分析,解决了历史申请中欺诈申请的识别问题,提升了欺诈申请识别的准确度。

[0177] 本申请实施例提供了一种电子设备适用于上述方法实施例。在此不再赘述。

[0178] 本申请实施例提供了一种计算机可读存储介质,该计算机可读存储介质上存储有计算机程序,该程序被处理器执行时实现上述实施例中所示的方法。

[0179] 本申请实施例提供了一种计算机可读存储介质,与现有技术通过身份认证的方式实现欺诈申请的识别相比,本申请实施例通过获取多个目标对象的音频信息,并基于多个目标对象的音频信息,通过预训练的第一深度学习模型提取各个目标对象的声纹特征得到声纹特征池,继而通过聚类算法对声纹特征池中的各个声纹特征进行分团聚类处理得到至少一个分团,并根据各个分团的相似度对各个分团进行排序,最后基于各个分团的排序结果确定欺诈目标对象。即通过聚类算法对声纹特征池中的各个声纹特征进行分团聚类处理得到至少一个分团,并根据各个分团的相似度对各个分团进行排序,然后根据排序结果确定欺诈识别对象,解决了如何识别一人冒充多人身份进行欺诈申请的问题,实现了欺诈识别的自动化处理;此外,基于当前申请人的声纹特征与历史申请人的声纹特征的聚类分析,解决了历史申请中欺诈申请的识别问题,提升了欺诈申请识别的准确度。

[0180] 本申请实施例提供了一种计算机可读存储介质适用于上述方法实施例。在此不再赘述。

[0181] 应该理解的是,虽然附图的流程图中的各个步骤按照箭头的指示依次显示,但是

这些步骤并不是必然按照箭头指示的顺序依次执行。除非本文中有明确的说明,这些步骤的执行并没有严格的顺序限制,其可以以其他的顺序执行。而且,附图的流程图中的至少一部分步骤可以包括多个子步骤或者多个阶段,这些子步骤或者阶段并不必然是在同一时刻执行完成,而是可以在不同的时刻执行,其执行顺序也不必然是依次进行,而是可以与其他步骤或者其他步骤的子步骤或者阶段的至少一部分轮流或者交替地执行。

[0182] 以上仅是本申请的部分实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本申请原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也应视为本申请的保护范围。

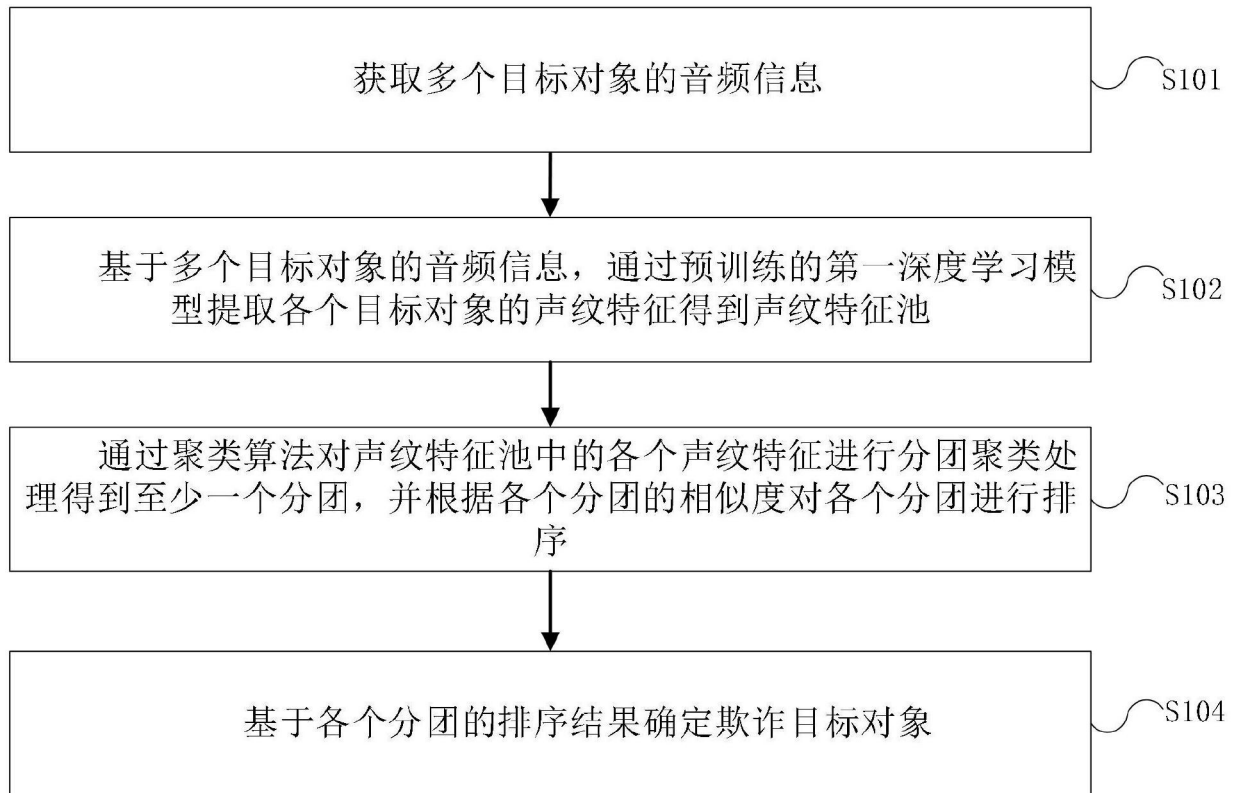


图1

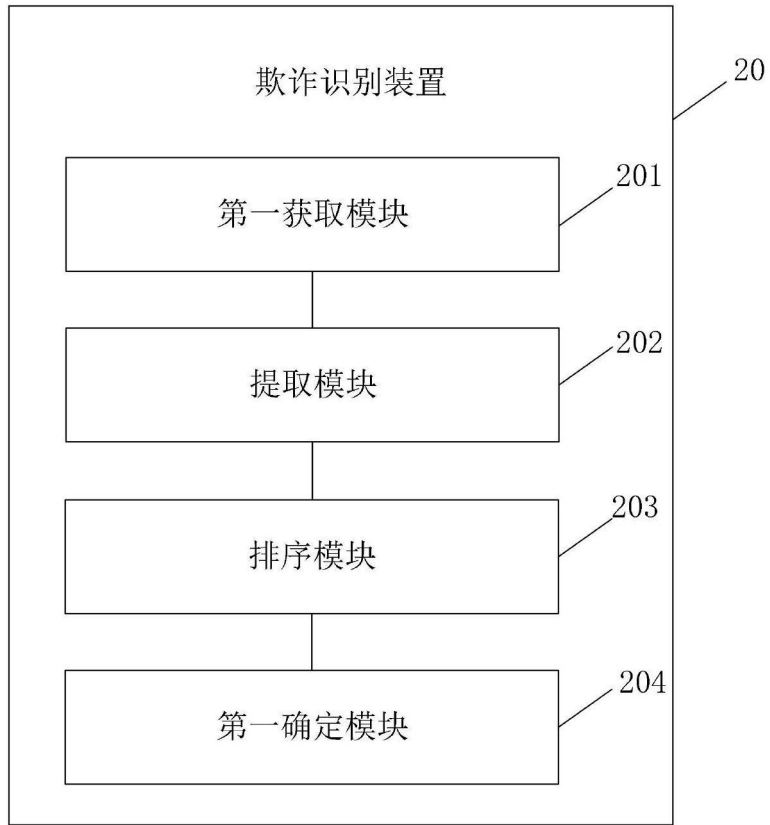


图2

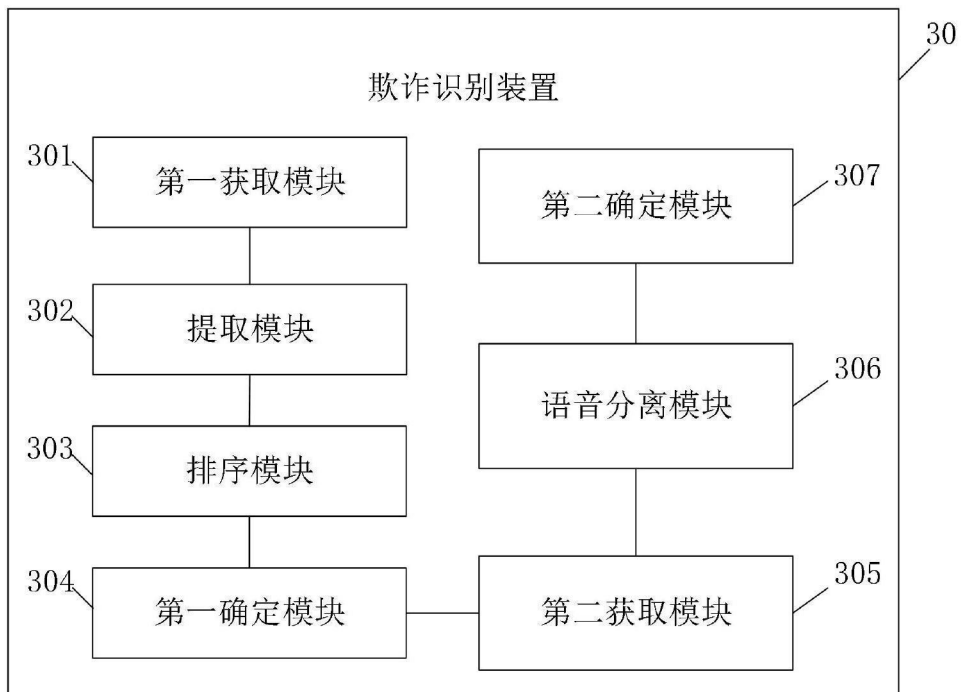


图3

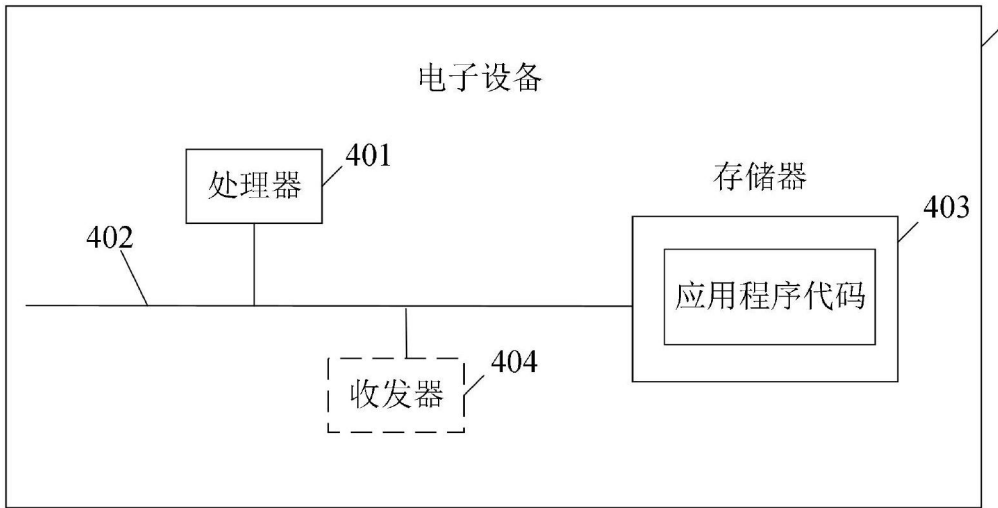


图4