(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 105940407 A
(43)申请公布日 2016.09.14

(21)申请号 201580006253.2
(22)申请日 2015.01.27
(30)优先权数据
14/172,619 2014.02.04 US
(85)PCT国际申请进入国家阶段日
2016.07.28
(86)PCT国际申请的申请数据
PCT/US2015/013126 2015.01.27
(87)PCT国际申请的公布数据
WO2015/119806 EN 2015.08.13
(71)申请人 高通股份有限公司
地址 美国加利福尼亚州
(72)发明人 金莱轩 南尤汉 埃里克·维瑟
(74)专利代理机构 北京律盟知识产权代理有限责任公司 11287
代理人 宋献涛
(51)Int.Cl.
G06F 21/46(2006.01)

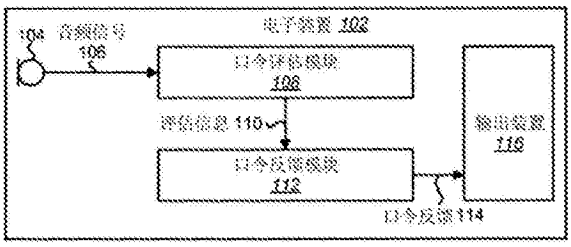
权利要求书2页 说明书27页 附图13页

(54)发明名称

用于评估音频口令的强度的系统和方法

(57)摘要

本发明描述一种用于通过电子装置来评估音频口令的强度的方法。所述方法包含获得由一或多个麦克风捕获的音频信号。所述音频信号包含音频口令。所述方法还包含基于测量所述音频信号的一或多个唯一特性来评估所述音频口令的所述强度。所述方法进一步包含基于对所述音频口令的所述强度的所述评估来告知用户所述音频口令较弱。



CN 105940407 A

1.一种用于通过电子装置来评估音频口令的强度的方法,其包括:

获得一或多个麦克风所捕获的音频信号,其中所述音频信号包含音频口令;

基于测量所述音频信号的一或多个唯一特性来评估所述音频口令的所述强度;以及

基于所述音频口令的所述强度的所述评估来告知用户所述音频口令较弱。

2.根据权利要求1所述的方法,其中所述音频信号包含至少一个语音分量。

3.根据权利要求1所述的方法,其中所述测量所述音频信号的一或多个唯一特性是基于通用语音模型。

4.根据权利要求1所述的方法,其中告知所述用户包括显示与所述音频口令的所述强度相关联的标记。

5.根据权利要求1所述的方法,其中告知所述用户包括显示口令强度得分。

6.根据权利要求1所述的方法,其进一步包括将口令强度得分与另一值进行比较。

7.根据权利要求6所述的方法,其中所述另一值是阈值或先前口令强度得分。

8.根据权利要求1所述的方法,其中告知所述用户包括显示至少一个候选语音分量。

9.根据权利要求1所述的方法,其进一步包括获得至少一个额外验证输入。

10.根据权利要求9所述的方法,其进一步包括使所述音频信号和所述额外验证输入中的至少一者降级。

11.根据权利要求1所述的方法,其进一步包括基于地理位置、用户年龄、用户性别、用户语言和地方方言中的一或多者更新通用语音模型。

12.一种用于评估音频口令的强度的电子装置,其包括:

一或多个麦克风,其捕获音频信号,其中所述音频信号包含音频口令;

口令评估电路,其耦合到所述一或多个麦克风,其中所述口令评估电路基于测量所述音频信号的一或多个唯一特性来评估所述音频口令的所述强度;以及

口令反馈电路,其耦合到所述口令评估电路,其中所述口令反馈电路基于所述音频口令的所述强度的所述评估来告知用户所述音频口令较弱。

13.根据权利要求12所述的电子装置,其中所述音频信号包含至少一个语音分量。

14.根据权利要求12所述的电子装置,其中所述测量所述音频信号的一或多个唯一特性是基于通用语音模型。

15.根据权利要求12所述的电子装置,其中告知所述用户包括显示与所述音频口令的所述强度相关联的标记。

16.根据权利要求12所述的电子装置,其中告知所述用户包括显示口令强度得分。

17.根据权利要求12所述的电子装置,其中所述口令评估电路进一步将口令强度得分与另一值进行比较。

18.根据权利要求17所述的电子装置,其中所述另一值是阈值或先前口令强度得分。

19.根据权利要求12所述的电子装置,其中告知所述用户包括显示至少一个候选语音分量。

20.根据权利要求12所述的电子装置,其进一步包括耦合到所述口令评估电路的一或多个输入装置,其中所述一或多个输入装置获得至少一个额外验证输入。

21.根据权利要求20所述的电子装置,其中所述口令评估电路使所述音频信号和所述额外验证输入中的至少一者进一步降级。

22.根据权利要求12所述的电子装置，其中所述口令评估电路进一步基于地理位置、用户年龄、用户性别、用户语言和地方方言中的一或多者来更新通用语音模型。

23.一种用于评估音频口令的强度的计算机程序产品，其包括上面具有指令的非暂时性有形计算机可读媒体，所述指令包括：

用于致使电子装置获得一或多个麦克风所捕获的音频信号的代码，其中所述音频信号包含音频口令；

用于致使所述电子装置基于测量所述音频信号的一或多个唯一特性来评估所述音频口令的所述强度的代码；以及

用于致使所述电子装置基于对所述音频口令的所述强度的所述评估来告知用户所述音频口令较弱的代码。

24.根据权利要求23所述的计算机程序产品，其中告知所述用户包括显示与所述音频口令的所述强度相关联的标记。

25.根据权利要求23所述的计算机程序产品，其中告知所述用户包括显示至少一个候选语音分量。

26.根据权利要求23所述的计算机程序产品，其进一步包括用于致使所述电子装置获得至少一个额外验证输入的代码。

27.一种用于评估音频口令的强度的设备，其包括：

用于获得音频信号的装置，其中所述音频信号包含音频口令；

用于基于测量所述音频信号的一或多个唯一特性来评估所述音频口令的所述强度的装置；以及

用于基于所述音频口令的所述强度的所述评估来告知用户所述音频口令较弱的装置。

28.根据权利要求27所述的设备，其中告知所述用户包括显示与所述音频口令的所述强度相关联的标记。

29.根据权利要求27所述的设备，其中告知所述用户包括显示至少一个候选语音分量。

30.根据权利要求27所述的设备，其进一步包括用于获得至少一个额外验证输入的装置。

用于评估音频口令的强度的系统和方法

技术领域

[0001] 本发明大体上涉及电子装置。更具体地说，本发明涉及用评估音频口令的强度的系统和方法。

背景技术

[0002] 在最近几十年中，电子装置的使用已变得普遍。明确地说，电子技术中的进步已减少了越来越复杂且有用的电子装置的成本。成本降低和消费者需求已使电子装置的使用剧增，使得其在现代社会中几乎随处可见。由于电子装置的使用已推广开来，因此具有对电子装置的新的且改进的特征的需求。更具体来说，人们常常寻求执行新功能和/或更快、更有效或以更高质量执行功能的电子装置。

[0003] 一些电子装置(例如，蜂窝式电话、智能电话、音频记录器、摄录影机、计算机等)利用音频信号。这些电子装置可捕获、编码、存储和/或发射所述音频信号。举例来说，智能电话可获得、编码和发射用于电话呼叫的语音信号，同时另一智能电话可接收所述语音信号并对其进行解码。

[0004] 然而，将音频信号用于安全目的的电子装置可能产生特定挑战。举例来说，许多音频信号可能不足以充分确保电子装置接入的安全。如从此论述可观察到，改进安全性的系统和方法可为有益的。

发明内容

[0005] 描述一种用于通过电子装置来评估音频口令的强度的方法。所述方法包含获得由一或多个麦克风捕获的音频信号。所述音频信号包含音频口令。所述方法还包含基于测量所述音频信号的一或多个唯一特性来评估所述音频口令的所述强度。所述方法进一步包含基于对音频口令的强度的评估来告知用户音频口令较弱。所述音频信号可包含至少一个语音分量。测量所述音频信号的一或多个唯一特性可基于通用语音模型。

[0006] 告知所述用户可包含显示与所述音频口令的强度相关联的标记。告知所述用户可包含显示口令强度得分。告知所述用户可包含显示至少一个候选语音分量。

[0007] 所述方法可包含将口令强度得分与另一值进行比较。所述另一值可为阈值或先前口令强度得分。

[0008] 所述方法可包含获得至少一个额外验证输入。所述方法可包含使所述音频信号和所述额外验证输入中的至少一者降级。所述方法可包含基于地理位置、用户年龄、用户性别、用户语言和地方方言中的一或多者来更新通用语音模型。

[0009] 还描述一种用于评估音频口令的强度的电子装置。所述电子装置包含捕获音频信号的一或多个麦克风。所述音频信号包含音频口令。所述还包含耦合到所述一或多个麦克风的口令评估电路。所述口令评估电路基于测量所述音频信号的一或多个唯一特性来评估音频口令的强度。所述电子装置进一步包含耦合到所述口令评估电路的口令反馈电路。所述口令反馈电路基于对音频口令的强度的评估而告知用户音频口令较弱。

[0010] 还描述一种用于评估音频口令的强度的计算机程序产品。所述计算机程序产品包含上面具有指令的非暂时性有形计算机可读媒体。所述指令包含用于致使电子装置获得由一或多个麦克风捕获的音频信号的代码。所述音频信号包含音频口令。所述指令还包含用于致使所述电子装置基于测量所述音频信号的一或多个唯一特性来评估音频口令的强度的代码。所述指令进一步包含用于致使所述电子装置基于对音频口令的强度的评估来告知用户音频口令较弱的代码。

[0011] 还描述一种用于评估音频口令的强度的设备。所述设备包含用于获得音频信号的装置。所述音频信号包含音频口令。所述设备还包含用于基于测量所述音频信号的一或多个唯一特性来评估音频口令的强度的装置。所述设备进一步包含用于基于对音频口令的强度的评估来告知用户音频口令较弱的装置。

附图说明

[0012] 图1是说明其中可实施用于评估音频口令的强度的系统和方法的电子装置的一个配置的框图；

[0013] 图2是说明用于评估音频口令的强度的方法的一个配置的流程图；

[0014] 图3包含说明唯一性量度的实例的图表；

[0015] 图4是说明其中可实施用于评估音频口令的强度的系统和方法的电子装置的更具体配置的框图；

[0016] 图5是说明用于评估音频口令的强度的方法的更具体配置的流程图；

[0017] 图6是说明用于评估音频口令的强度的方法的另一更具体配置的流程图；

[0018] 图7是说明用于评估音频口令的强度的方法的另一更具体配置的流程图；

[0019] 图8是说明用于评估音频口令的强度的方法的另一更具体配置的流程图；

[0020] 图9是说明扬声器(例如,用户)辨识模型的一个实例的框图；

[0021] 图10是说明用于基于预训练提供一或多个候选语音分量的方法的一个配置的流程图；

[0022] 图11是说明其中可实施用于评估音频口令的强度的系统和方法的电子装置的另一更具体配置的框图；

[0023] 图12是说明用于评估音频口令的强度的方法的更具体配置的流程图；

[0024] 图13是说明其中可实施用于评估音频口令的强度的系统和方法的无线通信装置的一个配置的框图;以及

[0025] 图14说明可在电子装置中利用的各种组件。

具体实施方式

[0026] 本文所揭示的系统和方法的一些配置提供口令强度评估以及对基于语音的生物计量验证的建议。当出于验证的目的使用话音时,用户可能想要将口令设定成说出。然而,可能难以知晓所述口令在话音音色方面是否将足够唯一,使得当正好说出同一口令时,其他任何人无法打破所述系统。如果说出的口令含有用户自身的与任意设定口令不同的生物计量差异,那么将更好。如果额外手段可用,那么其可恰当地用来加强安全性。

[0027] 本文所揭示的系统和方法可提供途径来评估“唯一性”的强度,使得用户可选择足

够唯一的口令。在一些配置中，本文所揭示的系统和方法可使用保留用户的增强型唯一性的发声来建议一些候选者。本文所揭示的系统和方法可建议一些候选者，不仅通过使用保留用户自身的增强型唯一性的发声，并且通过在一些配置中利用一或多个其它可用模态。

[0028] 一些扬声器检验系统通过使扬声器数据适合于通用背景模型(UBM)来训练扬声器模型。在检验阶段中，可计算在扬声器模型与UBM之间观察到的帧的似然比。可计算整个话语/句子帧上的概述统计，以确定语音帧是否来自真实扬声器。然而，每话语/音素/音节或甚至每帧的“局部”可能性指示一些具有高区别，但一些并不具有。可将不具有多高区别的部分解释也从其它模型阐述的部分，意味着其将污染检验性能。或者，可将其阐述为目标模型看不见的数据，意味着其可能难以被用户重复。因此，具有足够强且可容易再现的口令可为有益的。

[0029] 现在参考图式描述各种配置，其中相同的参考标号可指示功能上相似的元件。可以广泛多种不同配置来布置和设计如本文中在各图中大体描述和说明的系统和方法。因此，对如各图中所表示的若干配置的以下更详细描述无意限制如所主张的范围，而仅表示系统和方法。

[0030] 图1是说明其中可实施用于评估音频口令的强度的系统和方法的电子装置102的一个配置的框图。电子装置102的实例包含智能电话、蜂窝式电话、平板装置、计算机(例如，膝上型计算机、桌上型计算机等)、游戏系统、电子汽车控制台、个人数字助理(PDA)等。

[0031] 电子装置102包含一或多个麦克风104、口令评估模块108、口令反馈模块112和一或多个输出装置116。麦克风104可为将声信号转换为电子信号的一或多个变换器。所述一或多个输出装置116可为用于提供来自电子装置102的输出的装置。所述一或多个输出装置116的实例包含显示器(例如，显示面板、触摸屏)、扬声器(例如，将电子信号转换为声信号的变换器)、触觉装置(例如，产生力、运动和/或振动的装置)等。“模块”可在硬件(例如，电路)中或在硬件与软件的组合(例如，具有指令的处理器)中实施。举例来说，口令评估模块108和/或口令反馈模块112可在硬件中或在硬件与软件的组合中实施。

[0032] 一或多个麦克风104可耦合到口令评估模块108。口令评估模块108可耦合到口令反馈模块112。口令反馈模块112可耦合到一或多个输出装置116。如本文中所使用，术语“耦合”和相关术语可意味着一个组件直接连接(例如，无介入组件)或间接连接(例如，具有一或多个介入组件)到另一组件。图式中所描绘的箭头和/或线可表示耦合。

[0033] 一或多个麦克风104可捕获音频信号106。举例来说，一或多个麦克风104可捕获声学信号，并将其转换为电子音频信号106。音频信号106可包含音频口令。音频口令可包含用于检验用户的身份的一或多个声音(例如，一或多个语音分量，例如音素、音节、词语、短语、语句、发声等)。举例来说，音频口令可包含一或多个特性(例如，生物计量特性、音色等)，其可用于识别用户。可将音频信号106提供到口令评估模块108。

[0034] 口令评估模块108可获得(例如，接收)一或多个麦克风104所捕获的音频信号106。如上文所描述，音频信号106可包含音频口令。口令评估模块108可基于测量音频信号106的一或多个唯一特性来评估音频口令的强度。口令“强度”可为指示所述口令的安全程度的属性。举例来说，强音频口令(例如，具有高强度的音频口令)对于冒名顶替者来说非常难以或几乎不可能自然地模仿或复写，在冒名顶替者可自然地模仿或复写的情况下，所述冒名顶替者被不当地识别为真实用户。然而，对于冒名顶替者来说，弱音频口令(例如，具有低强度

的音频口令)可能更容易自然地模仿或复写,其中冒名顶替者被不当地识别为真实用户。在一些配置中,音频口令强度可依据唯一性来表达。举例来说,音频口令的一或多个语音分量越唯一,所述口令越强。然而,音频口令的一或多个语音组件越不唯一,所述口令越弱。因此,可对音频口令强度进行定量,且程度范围从弱到强。举例来说,较唯一的语音分量得分可比较不唯一的语音分量高(例如,强)。

[0035] 在一些配置中,口令评估模块108可用唯一性程度或与一或多个通用语音模型(例如,UBM)的区别来评估音频口令的一或多个语音分量(例如,发声、音素等)的强度。通用语音模型可为表示一群人的语音的语音模型(例如,统计语音模型)。一或多个UBM是通用语音模型的实例。

[0036] 在一些配置中,口令评估模块108可利用多个通用语音模型(例如,UBM)。举例来说,可基于用户的输入和/或特性(例如地理位置(例如,邮政编码、城市、县、州、国家等)、性别、年龄、语言、地方方言等)来采用(例如,选择和/或适应等)多个通用语音模型。用户的特性可影响用户语音的声学特性。在一些配置中,如果用户提供的信息与所存储的通用语音模型不匹配,那么电子装置102可通知用户和/或可根据用户的肯定应答改为使用恰当的模型。通过使用更具体匹配的通用语音模型(例如,UBM)来测量唯一性,电子装置102(例如,口令评估模块108)可提供更准确的唯一性量度和/或得分。在一些配置中,电子装置102(例如,口令评估模块108)可基于参与的一或多个用户的数据来更新对应的通用语音模型(例如,UBM)。

[0037] 在一些配置中,口令评估模块108可基于如下测量音频信号106的一或多个特性(例如,唯一特性)来评估音频口令的强度。口令评估模块108可从音频信号106提取一或多个特性(例如,特征向量)。举例来说,口令评估模块108可基于所述音频信号106确定一或多个梅尔频率倒谱系数(MFCC)。在一些配置中,MFCC可为通过对音频信号106的梅尔频率经平滑谱的记录量值应用离散余弦变换(DCT)而获得的系数。根据本文所揭示的系统和方法,可提取可用于扬声器/语音辨识的任何或所有特征来使用。MFCC是作为一实例而给出,因为其可为用于此类应用的相关特征向量。在一些配置中,根据本文所揭示的系统和方法而提取和/或利用的特征可不限于确定性特征(意味着例如不管数据如何,获得特征的方式可为固定的)。举例来说,可使用数据驱动的方法(例如在一些方法中,深神经网络)来提取(例如,习得)特征向量。

[0038] 口令评估模块108可基于一或多个通用语音模型(例如,UMB)获得音频信号106的唯一性量度。唯一性量度可指示音频信号106(例如,音频口令)上的唯一性。举例来说,唯一性量度可随音频信号106(例如,音频口令)的时间周期而变化。在一些配置中,可在每一语音分量(例如,音素、音节、词语等)和/或音频信号106(例如,音频口令)的帧上获得唯一性量度。在一些配置中,可将音频信号106(例如,输入波)转换为特征向量(例如,MFCC),其可用于获得唯一性量度和/或口令强度得分。

[0039] 在一些配置中,唯一性量度可为音频信号106与通用语音模型之间的似然比。举例来说,可根据等式(1)来确定似然比。

[0040] $$\sum_{t}\log\left(p\left(X|\lambda_{target}\right)\right)-\log\left(p\left(X|\lambda_{generic}\right)\right) \qquad (1)$$

[0041] 在等式(1)中,t是时间,X是音频信号(或基于所述音频信号的特征向量,例如),

λ_{target}是目标(例如,真实用户)模型,$\lambda_{generic}$是通用语音模型(例如,UBM),$p(X|\lambda_{target})$是X对应于真实用户的概率,且$p(X|\lambda_{generic})$是X对应于通用用户(例如,冒名顶替者、非真实用户等)的概率。通用术语(例如,$\lambda_{generic}$)可为冒名顶替者和/或非真实用户等的模型。冒名顶替者和/或非真实用户的模型可用于比较实际用户模型。比较实际用户模型可计算密集型和/或穷尽性的,因此可利用一些层级来限定搜索范围(例如,性别、年龄、位置等)。另外或替代地,通用术语(例如,$\lambda_{generic}$)可为非用户相依模型(例如,通用扬声器模型)。非用户相依模型可用于简化所述比较,其中可仅需要一个模型来用于比较。应注意,可更新(如果需要,例如)电子装置102中和/或远程装置(例如,远程服务器)中的通用模型(例如,$\lambda_{generic}$)。在一些实例中,可通过更新一或多个模型参数(例如,平均和/或混合权重)来更新通用模型。可周期性地(例如,定期)和/或不定期地(例如,按需、基于更新确定等)执行更新。

[0042] 在其它配置中,唯一性量度(例如,似然比)可一般化为任意非递减函数f。举例来说,可根据等式(2)来确定唯一性量度。

[0043] $$\sum_{t} f\left(\frac{p\left(X|\lambda_{target}\right)}{p\left(X|\lambda_{generic}\right)}\right) \qquad (2)$$

[0044] 在一些配置中,可如下获得和/或更新通用语音模型。通用语音模型可为(例如,不同于真实用户的)其它用户的语音进行建模。在一些配置中,通用语音模型可为其它用户的“始终适应模型”。另外或替代地,可(例如,通过电子装置102或远程装置)将音频信号106(例如,音频口令)与其它用户的模型进行比较,如果它们使用同一系统(例如,具有同一远程服务器)的话。在一些配置中,代替于将音频信号106与UBM进行比较来执行此步骤。

[0045] 复杂性可为此方法的一个问题,但可通过缩小搜索范围来减轻复杂性。举例来说,可首先执行基本信息检索,例如性别、年龄、语言(包含地方方言)等。另外或替代地,电子装置102或远程装置(例如,服务器)可尝试定位用户的物理住宅区或其一些历史。接着可将音频信号106(例如,音频口令)与具有同一类别(例如,性别、年龄、语言、地方方言、物理区等)的其它音频信号的实际模型的小得多的集合进行比较,其可正静态或动态地变化。电子装置102可动态地(例如,取决于住宅区或他/她讲的语言等)(向用户)提供对口令的不同建议。

[0046] 在一些配置中,通用语音模型可基于多个模型。举例来说,通用语音模型可基于基于具有从原始单个UBM更新的高可能性的高斯混合模型(GMM)状态来群集多个UBM。另外或替代地,通用语音模型可基于分组,所述分组基于可使用的物理区(例如,92121,圣地亚哥),且可将用户的模型与同一区中的人的模型进行比较。

[0047] 口令评估模块108可基于唯一性量度确定一或多个口令强度得分。口令强度得分可指示音频口令的强度。举例来说,口令强度得分可为整个音频口令的强度的指示。另外或替代地,可确定一或多个子级口令强度得分。在一些配置中,可基于唯一性量度的概述统计来确定口令强度得分。

[0048] 在一些配置中,口令强度得分可为唯一性量度本身。另外或替代地,确定口令强度得分可包含组合(例如,求和)唯一性量度的若干部分。另外或替代地,确定口令强度得分可包含映射唯一性量度、映射唯一性量度的一或多个部分和/或映射一或多个概述统计到数值(例如,百分比)、到词语(例如,“弱”、“适中”、“强”等)和/或到一些其它指示符(例如,色

彩、形状等)。

[0049] 在一些配置中,口令强度得分可为唯一性量度。举例来说,可利用等式(1)和/或等式(2)来获得口令强度得分。应注意,t可确定概述统计的长度。举例来说,可利用一些小常数t(例如,帧长度)来获得唯一性量度(例如,连续得分)。结合图3描述以小常数t获得的唯一性量度的一个实例。

[0050] 在一些配置中,确定口令强度得分可包含组合(例如,求和、求平均等)唯一性量度的若干部分。举例来说,口令评估模块108可在唯一性量度的某一周期上组合(例如,求和、求平均等),以确定口令强度得分。举例来说,口令评估模块108可使用整个唯一性量度或所述唯一性量度的一或多个足够长的时间帧来获得经平滑的得分。此经平滑的得分可为口令强度得分的一个实例。

[0051] 在一些配置中,如果t足够长,那么口令强度得分可为唯一性量度本身,而不组合唯一性量度的若干部分。然而,获得唯一性量度的对应于一或多个语音分量(例如,在音素级)的部分可为有益的,其可用于推荐和/或接入语音分量级(例如,音素级)唯一性。接着可组合唯一性量度的这些部分,以确定总口令强度得分。

[0052] 在一些配置中,可获得一或多个子级口令强度。举例来说,所述子级口令强度中的每一者可或可基于唯一性量度的所述部分。这可有益于使唯一性量度变窄到语音分量(例如,音素)级。另外或替代地,口令评估模块108可通过组合(例如,求和、求平均等)唯一性量度的若干部分(但不是所有唯一性量度,举例来说)来获得一或多个子级口令强度。举例来说,口令评估模块108可组合唯一性量度的分别对应于语音分量的部分。在一种方法中,口令评估模块108可对唯一性量度的对应于较大集合内的音素(例如,词语、短语、句子等)的部分求和和/或求平均。以此方式,可确定一或多个较高级(例如,词语级、短语级、句子级等)口令强度得分。

[0053] 在一些配置中,确定口令强度得分可包含将口令强度得分表达为和/或将口令强度得分映射到数值(例如,10%、43%、65%、90%等)、词语(例如“弱”、“适中”、“强”等)和/或一些其它指示符(例如红色、黄色、绿色等)。举例来说,口令评估模块108可使唯一性量度(和/或所述唯一性量度的若干部分)的概述统计乘以某一因子(例如,100),以确定口令强度得分。另外或替代地,口令评估模块108可基于唯一性量度、所述唯一性量度的若干部分和/或所述唯一性量度的概述统计来选择(例如,查找)特定数值、词语和/或某一其它指示符,以确定口令强度得分。举例来说,口令评估模块108可基于唯一性量度、所述唯一性量度的一个或若干部分和/或基于所述唯一性量度的一或多个量(例如,综合、平均值、统计等)来确定口令强度得分。可将这些量中的一或多者与一或多个阈值进行比较,以确定口令强度得分,和/或可基于这些量中的一或多者来(例如,在表中)查找口令强度得分。

[0054] 在一些配置中,口令评估模块108可确定音频口令是否足够强(例如,根据任意概率,根据用户偏好和/或足以使冒名顶替者非常不可能借助于发出音频口令而作为真实用户通过)。举例来说,口令评估模块108可将口令强度得分与值进行比较。举例来说,所述值可为先前口令强度得分和/或阈值。所述值可为静态(例如,预定)的和/或动态的。在一些配置中,所述值可由制造商设定和/或由用户配置。所述值可表达为数值(例如,60%、80%、90%等)和/或表达为词语(例如,“适中”、“强”等)。所述值可建立描绘口令强度被认为是充分还是不充分的决策点。

[0055] 在一些配置中，口令强度得分可结合音频口令考虑一或多个额外验证输入。举例来说，如果结合字母数字代码或指纹扫描使用音频口令，那么强度得分可反映音频口令与一或多个额外验证输入(如果利用)的组合所提供的额外验证强度。

[0056] 在一些配置中，电子装置102(例如，口令评估模块108)可接收一或多个额外验证输入。举例来说，一些配置可允许使用其它模态，例如视频陀螺/加速计传感器，键盘，指纹传感器等。在一些方法中，一或多个此类模态可用于具有较少唯一性或辨别强度的(短语、句子等)的一或多个部分。举例来说，当用户发出具有低唯一性的词语(例如，具有较小可辨别得分的词语“学校”)时，电子装置102可获得或接收一或多个额外验证输入。

[0057] 所述一或多个额外验证输入的实例如下给出。在电子装置102具有手势辨识的配置中，电子装置102可接收用户所输入的示意动作(例如，触摸屏图案、触摸垫图案、相机所捕获的视觉手势图案等)。所述示意动作可为用户创建或预定义的。在电子装置102包含相机的配置中，电子装置102可捕获用户的一或多个图像，例如用户的脸部、眼睛、鼻子、嘴唇、面部形状和/或更多的唯一信息，例如具有音频信号106的虹膜。举例来说，包含于电子装置102中的相机可(例如，通过用户)瞄准以捕获用户的脸部的全部或部分。

[0058] 在电子装置102包含一或多个运动和/或定向传感器(例如，陀螺仪、加速计、倾斜传感器等)的配置中，电子装置102可获得运动和/或定向信息。举例来说，用户可以用户创建或预定义的方式来定向和/或移动电子装置102(例如，电话)。举例来说，电子装置102可连同音频信号106编码陀螺和/或加速计传感器信息。

[0059] 在电子装置102包含物理或软件小键盘或键盘的配置中，电子装置102可连同音频信号106接收数值代码、文本和/或字母数字串(例如，由用户键入)。在电子装置102包含指纹传感器的配置中，电子装置102可接收指纹(例如，当用户触摸或握持指纹传感器时)。

[0060] 在电子装置102包含多个麦克风104的配置中，电子装置102可获得(例如，接收和/或确定)音频信号106的空间方向性信息。举例来说，用户可在相对于电子装置102的一序列方向(例如，顶部、底部、左、右、前、后、右上、左下等)上说出音频口令。举例来说，用户可朝电子装置102的底部说出第一个字，朝电子装置102的顶部说出第二个字，朝电子装置102的左侧说出第三个字，且朝电子装置102的右侧说出第四个字。

[0061] 可利用一或多个额外验证输入，而无时序和/或序列限制。在一些实例中，电子装置102可在接收到音频信号106之前、期间或之后的任何时间获得一或多个额外验证输入。

[0062] 在其它实例中，电子装置102可要求(或经配置以要求)相对于音频信号106的接收以某一时序约束条件和/或以某一序列接收所述一或多个额外验证输入。在一实例中，电子装置102可要求(或经配置以要求)在接收到音频信号106之前、期间和/或之后的某一时间周期内接收一或多个额外验证输入。举例来说，电子装置102可要求在音频口令的较弱语音分量期间接收额外验证输入。举例来说，假定对于音频口令“绿洲是海市蜃楼”，与音频口令的另一部分相比，“是”部分可为较不唯一或较弱。当用户发出“是”时，电子装置102可要求(或经配置以要求)接收额外验证输入(例如，文本、数值代码、字母数字串、空间方向性和/或额外生物计量(例如指纹扫描、用户的脸部的相机图像或虹膜等))。另外或替代地，电子装置102可要求(或经配置以要求)以特定序列(例如，在语音分量之前、在语音分量之后、在语音分量之间、在具有其它额外验证输入的序列中等)接收额外验证输入。

[0063] 在一些配置中，电子装置102(例如，口令评估模块108)可使音频信号106和/或额

外验证输入降级。举例来说，电子装置102可将信息从音频信号106去除(例如，下取样，滤除所述音频信号的一或多个部分)。另外或替代地，电子装置102可将信息从指纹扫描或从用户的脸部或虹膜的图像去除。此方法的一个益处是出于其安全或隐私原因，用户可能不想要共享确切或高品质信息(例如，确切或高品质生物计量信息，例如话音样本、所扫描的指纹、图像等)。因此，降级的信息可为所捕获信息的简化或降级版本。在一些配置中，单个模态或输入类型(例如，话音或语音、指纹、虹膜扫描等)的降级的信息本身无法用于可靠的用户识别。然而，来自多个模态或输入类型的降级的信息的组合仍可提供强验证。因此，甚至“虹膜”或“指纹”扫描可利用额外模态，如话音口令，即使非降级版本本身可提供高唯一性强度。

[0064] 口令评估模块108可将评估信息110提供到口令反馈模块112。评估信息110可包含指示口令评估中获得的口令强度和/或信息的信息。举例来说，评估信息110可包含所提取特征、唯一性量度、口令强度得分和/或其它信息。

[0065] 口令反馈模块112可提供口令反馈114。举例来说，口令反馈模块112可基于对音频口令的强度的评估来告知用户音频口令较弱。提供口令反馈114可使用户能够确定(例如，选择、提供或创建)足够强的音频口令。口令反馈114可包含口令强度得分、一或多个语音分量候选者(例如，所推荐或建议的语音分量)、一或多个所建议动作和/或一或多个消息。举例来说，口令反馈114可包含指示音频口令较弱的口令强度得分和消息。另外或替代地，口令反馈114可包含用户可用于创建较强音频口令的一或多个建议语音分量。在一些配置中，电子装置102可提供由所建议的语音分量组成的所建议合成(例如未知)字作为口令反馈114。另外或替代地，口令反馈114可包含用户可提供额外验证输入(例如，文本、数值代码、字母数字串、空间方向性、额外生物计量(例如面部扫描、虹膜扫描、指纹等))的所建议动作。

[0066] 在一些配置中，口令反馈模块112可提供一或多个口令建议。举例来说，电子装置102(例如，口令反馈模块112)可识别具有足够高的唯一性或与一或多个其它模型(例如，通用语音模型、通用模型、UBM等)的区别来识别一或多个语音分量(例如，发声、音素等)。举例来说，口令反馈模块112可经由一对语音辨识和扬声器检验系统，基于用户针对每一音素的话音的唯一性来识别一或多个语音分量。接着，口令反馈模块112可产生一些可能候选语音分量(例如，音素、音节、发声、口令等)，其具有高“唯一性”，使得用户可选择一或多个候选语音分量来创建口令。举例来说，电子装置102可显示口令反馈114，例如：“你可使用/啊/、/k/、…、<三角形>、<高通>、…”。另外或替代地，可为用户发出的口令提供具体口令反馈114，以较多地加强所述口令(例如，“你的口令具有60％强度。话语/嗯/可被/啊/...代替”)。

[0067] 在一些配置中，电子装置102(例如，口令反馈模块112)可以多模态提供口令建议。如上文所描述，例如，口令反馈模块112可提供口令反馈114，其建议一或多个额外验证输入(例如，文本数值代码、字母数字串、空间方向性、额外生物计量(例如，面部扫描、虹膜扫描、指纹等))。

[0068] 在一些配置中，口令反馈模块112可执行以下操作中的一或多者，以产生口令反馈114。口令反馈模块112可基于一或多个所提取的特征执行语音辨识。举例来说，口令反馈模块112可基于一或多个所提取的特征来确定一或多个所辨识语音分量。可利用基于输入提供具有时间对准的一序列音素的任何已知语音辨识器来确定一或多个所辨识语音分量。可

利用的语音辨识器的一个实例是隐式马尔可夫模型工具包(HTK)。

[0069] 口令反馈模块112可使唯一性量度与一或多个所辨识语音分量对准。举例来说，口令反馈模块112可使一或多个所辨识语音分量的出现率与唯一性量度在时间上对准。在一些配置中，每一语音分量(例如，音素)边界的时间对准是语音辨识的副产品中的一者。明确地说，口令反馈模块112可利用所辨识语音分量(例如，音素)的边界信息以及对应时间周期内的唯一性量度来产生经对准语音和唯一性。举例来说，口令反馈模块112可指定唯一性量度的一或多个时间点作为语音分量边界，如由语音辨识所提供夫人的语音分量边界所指示。

[0070] 口令反馈模块112可基于唯一性量度对一或多个语音分量进行分类。举例来说，口令反馈模块112可确定一或多个语音分量中的每一者的唯一性(例如，强度或弱度)。在一些配置中，口令反馈模块112可将经对准语音分量中的每一者处的唯一性量度(或基于唯一性量度的一些值，例如平均值、最大值、最小值等)与一或多个阈值进行比较。如果对应于语音分量的唯一性量度(或基于唯一性量度的值)大于阈值，那么可将对应的语音分量分类为足够唯一或足够强。在一些配置中，分类为足够唯一或足够强(例如，大于阈值)的语音分量可作为建议在口令反馈114中提供。此外，包含所述语音分量或类似语音分量的类似语音分量和/或话语、字、短语和/或口令可作为建议在口令反馈114中提供。

[0071] 口令反馈模块112可将口令反馈114提供到一或多个输出装置116。一或多个输出装置116可因此向用户中继或传达口令反馈114。举例来说，输出装置116(例如，显示器、触摸屏、扬声器等)可中继与音频口令的强度相关联的标记。在一个方法中，显示面板可显示口令强度得分。另外或替代地，扬声器可输出声学信号(例如，文字到语音)，其指示口令强度得分(例如，“你的口令较弱”、“你的口令强度为60％”等)。

[0072] 在一些配置中，输出装置116可中继一或多个建议。举例来说，显示面板可显示一或多个所建议语音分量，例如音素、音节、字、发声和/或短语(例如“/啊/、/嗯/、/k/、/三角形/、/海市蜃楼/”)。另外或替代地，扬声器可输出声学信号以中继一或多个建议(例如，“请将/啊/、/嗯/、/k/、/三角形/、/海市蜃楼/和/或额外输入类型添加到你的口令”)。

[0073] 在一些配置中，可经由一或多个图形用户接口(GUI)提供口令反馈114。举例来说，标记(例如，口令强度得分)、一或多个建议和/或一或多个消息可在GUI上呈现。在一些配置中，GUI还可提供用于接收用户输入的接口。举例来说，用户可经由GUI选择一或多个建议(例如，一或多个候选语音分量、合成字、所建议口令、一或多个额外验证输入选项等)。

[0074] 在一些配置中，电子装置102可包含检验模块(未图示)。所述检验模块可基于音频口令检验说话的用户是否是真实用户。应注意，检验程序可不同于口令评估程序。举例来说，检验可不发生，直到口令(例如，音频口令和/或一或多个额外验证输入)被设定为止。因此，如本文所揭示的口令评估和建议可包含不同于口令检验的程序，例如其可仅在口令已设定之后发生。

[0075] 图2是说明用于评估音频口令的强度的方法200的一个配置的流程图。结合图1描述的电子装置102可执行方法200。

[0076] 电子装置102可获得(202)一或多个麦克风104所捕获的音频信号106。此操作可如上文结合图1所描述来实现。音频信号106可包含音频口令。

[0077] 电子装置102可基于测量音频信号106的一或多个特性(例如，唯一特性)来评估

(204)音频口令的强度。此操作可如上文结合图1所描述来实现。举例来说，电子装置102可用唯一性程度或与一或多个通用语音模型(例如，UBM)的区别来评估(204)音频口令的一或多个语音分量(例如，发声、音素等)的强度。在一些配置中，口令评估模块108可利用多个通用语音模型(例如，UBM)，如上文所描述。举例来说，可基于用户的输入和/或特性(例如地理位置(例如，邮政编码、城市、县、州、国家等)、性别、年龄、语言、地方方言等)来采用(例如，选择和/或适应等)多个通用语音模型。

[0078] 在一些配置中，电子装置102可基于如下测量音频信号106的一或多个唯一特性来评估(204)音频口令的强度。电子装置102可从音频信号106提取一或多个特征。电子装置102可基于一或多个通用语音模型(例如，UMB)获得音频信号106的唯一性量度。电子装置102可基于所述唯一性量度来确定口令强度得分。

[0079] 在一些配置中，电子装置102可确定音频口令是否足够强(例如，根据任意概率，根据用户偏好和/或足以使冒名顶替者非常不可能借助于发出音频口令而作为真实用户通过)。举例来说，口令评估模块108可将口令强度得分与一值进行比较。所述值可为先前口令强度得分和/或阈值。

[0080] 电子装置102可提供口令反馈114。此操作可如上文结合图1所描述来实现。举例来说，电子装置102可基于音频口令的强度的评估(例如，当口令强度得分不大于值时)来告知(206)用户音频口令较弱。口令反馈114可包含口令强度得分、一或多个语音分量候选者(例如，所推荐或建议的语音分量)、一或多个所建议动作和/或一或多个消息。举例来说，口令反馈114可包含指示音频口令较弱的口令强度得分和消息。另外或替代地，口令反馈114可包含用户可用于创建较强音频口令的一或多个建议语音分量。另外或替代地，口令反馈114可包含用户可提供额外验证输入(例如，文本、数值代码、字母数字串、空间方向性、额外生物计量(例如面部扫描、虹膜扫描、指纹等))的所建议动作。

[0081] 可将口令反馈114提供到一或多个输出装置116。一或多个输出装置116可因此向用户中继或传达口令反馈114(例如，标记、一或多个所建议语音分量、一或多个所建议动作等)，如上文结合图1所描述。

[0082] 电子装置102可任选地检验用户输入。举例来说，电子装置102可在口令(例如，音频口令和/或额外验证输入)已设定之后接收用户输入。电子装置102可确定用户输入是否与口令充分匹配(例如，以足够高的概率)。音频口令检验的一种方法是结合图9所提供。如果用户输入与口令充分匹配(例如，与阈值概率和/或根据额外验证输入的一个或额外准则)，那么电子装置102可准予接入。举例来说，如果用户输入与口令充分匹配，那么电子装置102可允许用户接入一或多个功能(例如，应用程序、呼叫等)。

[0083] 图3包含说明唯一性量度的实例的图表。明确地说，图3包含图表A 318a、图表B318b和图表C 318c。图表A318a的垂直轴线梅尔频率标度说明，且图表A318a的水平轴以时间(帧)说明。图表B 318b的垂直轴线说明似然比，且图表B 318b的水平轴以时间(帧)说明。图表C 318c的垂直轴线说明似然比，且图表C 318c的水平轴以时间(帧)说明。

[0084] 图表A 318a说明随音频信号的时间过去的梅尔频率的频谱图。所述音频信号包含短语(例如，音频口令)“绿洲是海市蜃楼”。语音分量A 320包含话语“是”。语音分量B322在词语“海市蜃楼”中包含话语“啊”。

[0085] 图表B 318b说明随时间的过去，真实用户(例如，待验证的真实扬声器或用户)的

唯一性量度(例如,似然比)的一个实例。唯一性量度对应于图表A 318a。在此实例中,唯一性量度是真实用户的语音(例如,用户语音模型)与UBM之间的似然比。如在图表B318b中可观察到,语音分量A 320(例如,“是”)具有低唯一性。然而,语音分量B 322(例如,“海市蜃楼”中的“啊”)针对真实用户具有高唯一性。

[0086] 图表C 318c说明随时间过去冒名顶替者的唯一性量度(例如,似然比)的一个实例。唯一性量度对应于图表A 318a。在此实例中,唯一性量度是冒名顶替者的语音(例如,冒名顶替者语音模型)与UBM之间的似然比。如在图表C 318c中可观察到,语音分量A320(例如,“是”)和语音分量B 322具有低唯一性。如图3中所示,可利用提供真实用户的升高的唯一性(例如,似然比)但提供冒名顶替者的低似然比的语音分量(例如,音素、音节、字等)来创建较强口令。

[0087] 图4是说明其中可实施用于评估音频口令的强度的系统和方法的电子装置402的更具体配置的框图。结合图4描述的电子装置402可为结合图1描述的电子装置102的一个实例。

[0088] 电子装置402包含一或多个麦克风404、口令评估模块408、口令反馈模块412和一或多个输出装置416。包含于电子装置402中的组件中的一或多者可对应于包含于结合图1描述的电子装置102中的组件中的一或多者和/或可类似于其而起作用。

[0089] 电子装置402可任选地包含通信模块436。通信模块436可使电子装置402能够与一或多个远程装置(例如,其它电子装置、基站、服务器、计算机、网络基础设施等)通信。通信模块436可提供无线和/或有线通信。举例来说,通信模块436可根据一或多个无线规范(例如,第三代合作伙伴计划(3GPP)规范、电气电子工程师学会(IEEE)802.11规范等)与一或多个其它装置无线通信。另外或替代地,通信模块436可经由有线链路(例如,经由以太网、有线通信等)与其它装置通信。

[0090] 一或多个麦克风404可捕获音频信号406。音频信号406可包含音频口令。音频口令可包含用于检验用户的身份的一或多个声音(例如,一或多个语音分量,例如音素、音节、词语、短语、语句、发声等)。可将音频信号406提供到口令评估模块408。

[0091] 口令评估模块408可包含特征提取模块424、唯一性测量模块428和/或口令强度计分模块432。

[0092] 口令评估模块408(例如,特征提取模块424)可获得(例如,接收)一或多个麦克风404所捕获的音频信号406。特征提取模块424可从音频信号406提取一或多个特征以获得所提取特征426。此操作可如上文结合图1所描述来实现。举例来说,特征提取模块424可基于音频信号406确定一或多个MFCC。MFCC可为所提取特征426的一个实例。特征提取模块424可耦合到唯一性测量模块428。特征提取模块424可将所提取的特征426提供到唯一性测量模块428。

[0093] 唯一性测量模块428可基于一或多个通用语音模型(例如,UMB)获得音频信号406的唯一性量度430。在一些配置中,唯一性量度可为音频信号406与通用语音模型之间的似然比。图3中的图表B 318b说明唯一性量度430(例如,似然比)的一个实例。在一些配置中,电子装置402可本地确定(例如,计算)唯一性量度430。举例来说,电子装置402可本地存储一或多个通用语音模型,其可用来确定唯一性量度430。在其它配置中,电子装置402可从远程装置(例如,服务器、中央服务器)接收唯一性量度430。举例来说,远程装置(例如,服务

器、中央服务器)可存储一或多个通用语音模型,其可用于远程确定唯一性量度430。

[0094] 在一些配置中,可如上文结合图1所描述,可获得和/或更新通用语音模型。举例来说,电子装置402和/或远程装置(例如,服务器)可获得和/或更新通用语音模型。在一些配置中,电子装置402可获得和/或更新通用语音模型。举例来说,电子装置402(例如,唯一性测量模块428)可存储用于通用语音模型的预定数据。电子装置402可任选地通过经由通信模块436从远程装置(例如,服务器、中央服务器等)接收数据来更新通用语音模型。

[0095] 在一些配置中,电子装置402(例如,唯一性测量模块428)可接收和/或确定用户特性(例如,性别、年龄、位置等)。举例来说,唯一性测量模块428可获得如由用户经由一或多个输入装置输入的用户特性。电子装置402(例如,唯一性测量模块428)可任选地将通用语音模型(例如,UBM)更新请求发送到远程装置(例如,服务器、中央服务器等)。在一些方法中,通用语音模型更新请求可包含用户特性的一或多个指示符。所述远程装置可任选地(基于例如用户特性)为电子装置402的通用语音模型确定更新。所述远程装置可将通用语音模型(例如,UBM)更新数据发送到电子装置402。通用语音模型更新数据可基于用户特性,其可由所述电子装置402用于适应或修改电子装置402(例如,唯一性测量模块428)所使用的通用语音模型。

[0096] 在一些配置中,电子装置402可将唯一性量度请求发送到远程装置。举例来说,唯一性测量模块428可将唯一性量度请求提供到通信模块436,其可将唯一性量度请求发送到远程装置(例如,服务器)。唯一性量度请求可包含关于音频信号106的信息(例如,所提取特征426)。在此方法中,远程装置(例如,服务器)可基于一或多个通用语音模型(例如,UBM)确定(例如,计算)唯一性量度430(例如,似然比)。电子装置402(例如,通信模块436)可接收唯一性量度430,并将唯一性量度430提供到唯一性测量模块428。

[0097] 应注意,在一些配置中,远程装置可基于用户信息(例如,位置、年龄、性别等)获得、维持和/或适应其通用语音模型。所述用户信息可由远程装置从电子装置402、一或多个其它装置和/或一或多个第三方接收。远程装置接着可将唯一性量度发送到电子装置402。

[0098] 唯一性测量模块428可将唯一性量度430提供到口令强度计分模块432。口令强度计分模块432可基于唯一性量度430确定一或多个口令强度得分434。此操作可如上文结合图1所描述来实现。举例来说,口令强度得分可为唯一性量度,和/或确定口令强度得分可包含组合(例如,求和、求平均等)所述唯一性量度的若干部分。另外或替代地,确定口令强度得分可包含映射唯一性量度、映射所述唯一性量度的一或多个部分和/或映射一或多个概述统计到数值(例如,百分比)、到字(例如,“弱”、“适中”、“强”等)和/或到一些其它指示符(例如,色彩、形状等)。

[0099] 口令强度计分模块432可确定音频口令是否足够强,如上文结合图1所描述。举例来说,口令强度计分模块432可将口令强度得分434与一或多个值(例如,先前口令强度得分和/或阈值)进行比较。在一些配置中,口令强度得分可结合音频口令反映一或多个额外验证输入(例如,空间方向性、文本、数值代码、字母数字串、额外生物计量等)。在一些配置中,电子装置402(例如,口令评估模块408)可使音频信号406和/或额外验证输入降级。

[0100] 口令评估模块408可将评估信息提供到口令反馈模块412。举例来说,评估信息410可包含所提取特征426、唯一性量度430、口令强度得分434和/或其它信息。

[0101] 口令反馈模块412可任选地包含语音辨识模块438、对准模块442和/或语音分量分

类模块446语音辨识模块438可基于一或多个所提取的特征426执行语音辨识。举例来说，口令反馈模块412可基于一或多个所提取的特征426来确定一或多个所辨识语音分量440。此操作可如上文结合图1所描述来实现。语音辨识模块438可将所辨识的语音分量440提供到对准模块442。

[0102] 对准模块442可使唯一性量度430与一或多个所辨识语音分量440对准。举例来说，对准模块442可使一或多个所辨识语音分量440的出现与唯一性量度在时间上对准，以产生对准语音和唯一性444。此操作可如上文结合图1所描述来实现。对准模块442可将经对准的语音和唯一性444提供到语音分量分类模块446。

[0103] 语音分量分类模块446可基于唯一性量度430对一或多个语音分量(例如，所辨识语音分量440)进行分类。举例来说，口令反馈模块412可确定经对准语音和唯一性444中的一或多个所辨识语音分量中的每一者的唯一性(例如，强度或弱度)。在一些配置中，口令反馈模块412可将经对准语音分量中的每一者处的唯一性量度(或基于唯一性量度的一些值，例如平均值、最大值、最小值等)与一或多个阈值进行比较。如果对应于语音分量的唯一性量度(或基于唯一性量度的值)大于阈值，那么可将对应的语音分量分类为足够唯一或足够强。在一些配置中，分类为足够唯一或足够强(例如，大于阈值)的语音分量可作为建议在口令反馈414中提供。此外，包含所述语音分量或类似语音分量的类似语音分量和/或话语、字、短语和/或口令可作为建议在口令反馈414中提供。

[0104] 口令反馈模块412可将口令反馈414提供到一或多个输出装置416。口令反馈414可包含口令强度得分、一或多个语音分量候选者(例如，所推荐或所建议语音分量、一或多个所建议动作(例如，建议一或多个额外验证输入)和/或一或多个消息。一或多个输出装置416可因此向用户中继或传达口令反馈414。此操作可如上文结合图1所描述来实现。举例来说，输出装置416可输出口令反馈414作为文本、图像和/或声音。所述输出可中继标记(例如，口令强度得分)、一或多个语音分量候选者(例如，所推荐或所建议语音分量)、一或多个所建议动作(例如，建议一或多个额外验证输入)和/或一或多个消息。

[0105] 图5是说明用于评估音频口令的强度的方法500的更具体配置的流程图。结合图1和4描述的电子装置102、402中的一或多者可执行方法500。

[0106] 电子装置402可基于预训练任选地提供(502)一或多个候选语音分量。结合图10描述基于预训练提供(502)一或多个候选语音分量的实例。

[0107] 电子装置402可获得(504)一或多个麦克风404所捕获的音频信号406。这可如上文结合图1到2以及4中的一或多者所描述来实现。音频信号106可包含音频口令。音频口令可包含用于检验用户的身份的一或多个声音(例如，一或多个语音分量，例如音素、音节、词语、短语、语句、发声等)。

[0108] 电子装置402可从音频信号406提取(506)一或多个特征以获得所提取特征426。这可如上文结合图1和4中的一或多者所描述来实现。举例来说，电子装置402可基于音频信号406确定一或多个MFCC。MFCC可为所提取特征426的一个实例。

[0109] 电子装置402可基于一或多个通用语音模型(例如，UMB)获得(508)音频信号406的唯一性量度430。这可如上文结合图1到4中的一或多者所描述来实现。在一些配置中，唯一性量度可为音频信号406与通用语音模型之间的似然比。在一些配置中，电子装置402可本地确定(例如，计算)唯一性量度430。举例来说，电子装置402可本地存储一或多个通用语音

模型(例如,本地UBM),其可用来确定唯一性量度430。在其它配置中,电子装置402可从远程装置(例如,服务器、中央服务器)接收唯一性量度430。举例来说,远程装置(例如,服务器、中央服务器)可存储一或多个通用语音模型,其可用于远程确定唯一性量度430。在一些配置中,电子装置402可将唯一性量度请求发送到远程装置。唯一性量度请求可包含关于音频信号406的信息(例如,所提取特征426)。在此方法中,远程装置(例如,服务器)可基于一或多个通用语音模型(例如,UBM)确定(例如,计算)唯一性量度430(例如,似然比)。电子装置402可接收唯一性量度430。

[0110] 电子装置402可基于所述唯一性量度430来确定(510)口令强度得分434。这可如上文结合图1和4中的一或多者所描述来实现。

[0111] 电子装置402可确定(512)口令强度得分是否大于一值。这可如上文结合图1和4中的一或多者所描述来实现。举例来说,电子装置402可将口令强度得分434与一值(例如,先前口令强度得分和/或阈值)进行比较。

[0112] 如果口令强度得分434大于所述值(例如,先前口令强度得分和/或阈值),那么电子装置402可基于音频信号406设定(516)口令。在一些配置中,电子装置402可存储音频信号406和/或指定音频信号406作为口令。另外或替代地,电子装置402可存储和/或指定包含于作为口令的音频信号406中的所辨识语音分量的组合。

[0113] 如果口令强度得分434不大于所述值(例如,小于或等于所述值),那么电子装置402可提供(514)口令反馈。这可如上文结合图1到2以及4中的一或多者所描述来实现。举例来说,电子装置402可提供和/或输出口令反馈414。口令反馈414可包含口令强度得分、一或多个语音分量候选者(例如,所推荐或所建议语音分量、一或多个所建议动作(例如,建议一或多个额外验证输入)和/或一或多个消息。举例来说,电子装置402可输出口令反馈414作为文本、图像和/或声音。所述输出可中继标记(例如,口令强度得分)、一或多个语音分量候选者(例如,所推荐或所建议语音分量)、一或多个所建议动作(例如,建议一或多个额外验证输入)和/或一或多个消息。

[0114] 图6是说明用于评估音频口令的强度的方法600的另一更具体配置的流程图。明确地说,这种配置提供可执行以便提供一或多个建议的操作的实例。结合图1和4描述的电子装置102、402中的一或多者可执行方法600。

[0115] 电子装置402可基于预训练任选地提供(602)一或多个候选语音分量。结合图10描述基于预训练提供(602)一或多个候选语音分量的实例。

[0116] 电子装置402可获得(604)一或多个麦克风404所捕获的音频信号406。这可如上文结合图1到2以及4到5中的一或多者所描述来实现。

[0117] 电子装置402可从音频信号406提取(606)一或多个特征以获得所提取特征426。这可如上文结合图1以及4到5中的一或多者所描述来实现。

[0118] 电子装置402可基于一或多个通用语音模型(例如,UMB)获得(608)音频信号406的唯一性量度430。这可如上文结合图1以及4到5中的一或多者所描述来实现。

[0119] 电子装置402可基于所述唯一性量度430来确定(610)口令强度得分434。这可如上文结合图1以及4到5中的一或多者所描述来实现。

[0120] 电子装置402可确定(612)口令强度得分是否大于一值。这可如上文结合图1以及4到5中的一或多者所描述来实现。

[0121] 如果口令强度得分434大于所述值(例如,先前口令强度得分和/或阈值),那么电子装置402可基于音频信号406来设定(622)口令。此操作可如上文结合图5所描述来实现。

[0122] 如果口令强度得分434不大于所述值(例如,小于或等于所述值),那么电子装置402可基于一或多个所提取的特征426来执行(614)语音辨识。举例来说,电子装置402可基于一或多个所提取的特征426来确定一或多个所辨识语音分量440。此操作可如上文结合图1所描述来实现。

[0123] 电子装置402可使唯一性量度430与一或多个所辨识语音分量440对准(616)。举例来说,电子装置402可使一或多个所辨识语音分量的出现与所述唯一性量度在时间上对准,以产生经对准的语音和唯一性444。这可如上文结合图1和4中的一或多者所描述来实现。

[0124] 电子装置402可基于唯一性量度430对一或多个语音分量(例如,所辨识语音分量440)进行分类(618)。举例来说,电子装置402可确定经对准语音和唯一性444中的一或多个所辨识语音分量中的每一者的唯一性(例如,强度或弱度)。在一些配置中,口令反馈模块412可将经对准语音分量中的每一者处的唯一性量度(或基于唯一性量度的一些值,例如平均值、最大值、最小值等)与一或多个阈值进行比较。如果对应于语音分量的唯一性量度(或基于唯一性量度的值)大于阈值,那么可将对应的语音分量分类为足够唯一或足够强。在一些配置中,分类为足够唯一或足够强(例如,大于阈值)的语音分量可作为建议在口令反馈414中提供(620)。此外,包含所述语音分量或类似语音分量的类似语音分量和/或话语、字、短语和/或口令可作为建议在口令反馈414中提供(620)。

[0125] 电子装置402可提供(620)口令反馈。此操作可如上文结合图1以及4到5中的一或多者所描述而实现。举例来说,电子装置402可提供和/或输出口令反馈414。口令反馈414可包含口令强度得分、一或多个语音分量候选者(例如,所推荐或所建议语音分量、一或多个所建议动作(例如,建议一或多个额外验证输入)和/或一或多个消息。举例来说,电子装置402可输出口令反馈414作为文本、图像和/或声音。所述输出可中继标记(例如,口令强度得分)、一或多个语音分量候选者(例如,所推荐或所建议语音分量)、一或多个所建议动作(例如,建议一或多个额外验证输入)和/或一或多个消息。在一些配置中,电子装置402可提供(620)由所建议的语音分量组成的所建议合成(例如未知)字作为口令反馈。

[0126] 图7是说明用于评估音频口令的强度的方法700的另一更具体配置的流程图。明确地说,这种配置提供可为用其它用户的模型进行口令强度评估和建议执行的操作的实例。结合图1和4描述的电子装置102、402中的一或多者可执行方法700。

[0127] 电子装置402可基于预训练任选地提供(702)一或多个候选语音分量。结合图10描述基于预训练提供(702)一或多个候选语音分量的实例。

[0128] 电子装置402可获得(704)一或多个麦克风404所捕获的音频信号406。这可如上文结合图1到2以及4到6中的一或多者所描述来实现。

[0129] 电子装置402可从音频信号406提取(706)一或多个特征以获得所提取特征426。这可如上文结合图1以及4到6中的一或多者所描述来实现。

[0130] 电子装置402可将唯一性量度请求发送(708)(例如,到远程装置)。此操作可如上文结合图4所描述来实现。举例来说,电子装置402可经由有线和/或无线通信将唯一性量度请求发送到远程装置(例如,服务器)。唯一性量度请求可包含关于音频信号406的信息(例如,所提取特征426)。在此方法中,远程装置(例如,服务器)可基于一或多个通用语音模型

(例如,UBM、其它用户的语音模型等)确定(例如,计算)唯一性量度430(例如,似然比)。应注意,在一些配置中,远程装置可基于用户信息(例如,位置、年龄、性别等)获得、维持和/或适应其通用语音模型。所述用户信息可由远程装置从电子装置402、一或多个其它装置和/或一或多个第三方接收。远程装置接着可将唯一性量度发送到电子装置402。

[0131] 电子装置402(例如,通信模块436)可接收(710)唯一性量度430。举例来说,电子装置402可经由有线和/或无线通信从远程装置(例如,服务器)接收(710)唯一性量度430。

[0132] 电子装置402可基于所述唯一性量度430来确定(712)口令强度得分434。这可如上文结合图1以及4到6中的一或多者所描述来实现。

[0133] 电子装置402可确定(714)口令强度得分是否大于一值。这可如上文结合图1以及4到6中的一或多者所描述来实现。

[0134] 如果口令强度得分434大于所述值(例如,先前口令强度得分和/或阈值),那么电子装置402可基于音频信号406来设定(724)口令。这可如上文结合图5到6中的一或多者所描述来实现。

[0135] 如果口令强度得分434不大于所述值(例如,小于或等于所述值),那么电子装置402可基于一或多个所提取的特征426任选地执行(716)语音辨识。这可如上文结合图1到6中的一或多者所描述来实现。

[0136] 电子装置402可任选地使唯一性量度430与一或多个所辨识语音分量440对准(718)。这可如上文结合图1、4和6中的一或多者所描述来实现。

[0137] 电子装置402可任选地基于唯一性量度430对一或多个语音分量(例如,所辨识语音分量440)进行分类(720)。这可如上文结合图1、4和6中的一或多者所描述来实现。

[0138] 电子装置402可提供(722)口令反馈。这可如上文结合图1以及4到6中的一或多者所描述来实现。

[0139] 图8是说明用于评估音频口令的强度的方法800的另一更具体配置的流程图。明确地说,这种配置提供可执行以用于更新通用语音模型的操作的实例。结合图1和4描述的电子装置102、402中的一或多者可执行方法800。

[0140] 电子装置402可基于预训练任选地提供(802)一或多个候选语音分量。结合图10描述基于预训练提供(802)一或多个候选语音分量的实例。

[0141] 电子装置402可获得(804)一或多个麦克风404所捕获的音频信号406。这可如上文结合图1到2以及4到7中的一或多者所描述来实现。

[0142] 电子装置402可从音频信号406提取(806)一或多个特征以获得所提取特征426。这可如上文结合图1以及4到7中的一或多者所描述来实现。

[0143] 电子装置402可获得(808)一或多个用户特性。用户特性的实例包含地理位置(例如,邮政编码、城市、县、州、国家等)、性别、年龄、语言和/或地方方言等。举例来说,电子装置402可(例如,从用户)接收指示一或多个用户特性的一或多个输入。另外或替代地,电子装置402可从一或多个传感器获得(808)一或多个用户特性。举例来说,电子装置402可基于从麦克风404捕获的音频来确定用户的性别、语言和/或地方方言。另外或替代地,电子装置402可基于从麦克风404捕获的音频来估计用户年龄。另外或替代地,电子装置402可基于来自全球定位系统(GPS)模块的数据确定地理位置。另外或替代地,电子装置402可从远程装置(例如,服务提供商服务器)请求一或多个用户特性。

[0144] 电子装置402可基于一或多个用户特性更新(810)通用语音模型。这可如上文结合图1和4中的一或多者所描述来实现。举例来说，电子装置402和/或远程装置(例如，服务器)可更新(810)通用语音模型。在一些配置中，电子装置402可基于用户特性来本地更新(810)通用语音模型。举例来说，电子装置402可任选地存储用于通用语音模型的预定数据，电子装置402可通过仅包含具有类似于所述用户的特性的特性的其它用户的数据来本地更新(810)所述预定数据。

[0145] 电子装置402可通过经由通信模块436将用户特性发送到远程装置(例如，服务器)和/或从远程装置(例如，服务器、中央服务器等)接收数据，基于用户特性来任选地更新(810)通用语音模型。举例来说，电子装置402可将通用语音模型(例如，UBM)更新请求发送到远程装置(例如，服务器、中央服务器等)。在一些方法中，通用语音模型更新请求可包含用户特性的一或多个指示符。在一些配置中，远程装置可基于用户特性来更新存储在远程装置上的一或多个通用语音模型。另外或替代地，远程装置可(例如，基于用户特性)任选地确定对电子装置402的通用语音模型的更新。所述远程装置可将通用语音模型(例如，UBM)更新数据发送到电子装置402。

[0146] 电子装置402可基于一或多个通用语音模型(例如，UMB)获得(812)音频信号406的唯一性量度430。这可如上文结合图1以及4到7中的一或多者所描述来实现。

[0147] 电子装置402可基于所述唯一性量度430来确定(814)口令强度得分434。这可如上文结合图1以及4到7中的一或多者所描述来实现。

[0148] 电子装置402可确定(816)口令强度得分是否大于一值。这可如上文结合图1以及4到7中的一或多者所描述来实现。

[0149] 如果口令强度得分434大于所述值(例如，先前口令强度得分和/或阈值)，那么电子装置402可基于音频信号406来设定(826)口令。这可如上文结合图5到7中的一或多者所描述来实现。

[0150] 如果口令强度得分434不大于所述值(例如，小于或等于所述值)，那么电子装置402可任选地基于一或多个所提取的特征426执行(818)语音辨识。这可如上文结合图1以及6到7中的一或多者所描述来实现。

[0151] 电子装置402可任选地使唯一性量度430与一或多个所辨识语音分量440对准(820)。这可如上文结合图1、4以及6到7中的一或多者所描述来实现。

[0152] 电子装置402可任选地基于唯一性量度430对一或多个语音分量(例如，所辨识语音分量440)进行分类(822)。这可如上文结合图1、4以及6到7中的一或多者所描述来实现。

[0153] 电子装置402可提供(824)口令反馈。这可如上文结合图1以及4到7中的一或多者所描述来实现。

[0154] 图9是说明扬声器(例如，用户)辨识模型的一个实例的框图。扬声器辨识模型可基于文本无关扬声器辨识。一个模型是基于MFCC和UBM-GMM。这包含使用GMM来训练UBM。如图9中所示，训练948可包含将训练语音950用于通用语音模型产生952。

[0155] 在一些方法中，可使用对通用语音模型(例如，UBM)的最大后验概率(MAP)适应来执行扬声器登记954。如图9中所示，登记954(例如，适应)可包含将用户话语956用于用户语音模型产生958。

[0156] 在一些方法中，可通过比较通用语音模型(例如，UBM)与每一所登记扬声器模型之

间的似然比来检验每一语音话语962。如图9中所示，可在检验(964)程序中利用每一话语962。举例来说，可根据等式(1)和/或等式(2)执行检验(964)程序。举例来说，检验(964)程序可根据$\sum_t \log(p(X|\lambda_{target})) - \log(p(X|\lambda_{generic})) > \theta$执行，其中t是时间，X是话语962或音频信号，$\lambda_{target}$是目标(例如，真实用户话语)模型，$\lambda_{generic}$是通用语音模型(例如，UBM)，$p(X|\lambda_{target})$是X对应于真实用户的概率，$p(X|\lambda_{generic})$是X对应于通用用户(例如，冒名顶替者、非真实用户、非用户相依模型或通用扬声器模型)的概率，且θ是检验阈值。当识别多个扬声器时，可选择产生最高可能性的那个扬声器。另外或替代地，可利用其它分类器(例如，支持向量机或神经网络)。

[0157] 图10是说明用于基于预训练提供一或多个候选语音分量的方法1000的一个配置的流程图。举例来说，结合图10描述的程序中的一或多者可用于针对登记的预训练中。举例来说，针对登记的预训练可在接收到用于评估(例如，在结合图5到8中的一或多者描述的步骤502、602、702和802中的一或多者中)的音频口令之前发生。

[0158] 下文给出关于登记和比较的更多细节。登记用户的一种方法可包含让用户说一会话，以提供足够的音素来从通用语音模型(例如，UBM)适应所述用户的模型。在一些配置中，电子装置102、402可提供一些预定义的在语音学上平衡的语句来最小化训练时间。另外或替代地，用户可读足够长的提词(例如，以充分地训练，使通用语音模型适应所述用户的语音模型)。

[0159] 另外或替代地，电子装置102、402可收集呼叫期间的用户数据(例如，语音)，假定所述用户是所述装置的属主(例如，真实用户)。一旦达到数据大小方面的某一层级，电子装置102、402就可通知或告知(例如，显示消息，输出提供所述消息的语音)用户可启用话音口令。在一些配置中，电子装置可继续更新用户的语音模型。以此方式，可监视用户随时间的音色改变(例如，年龄相关改变)。

[0160] 结合图1和4中的一或多者描述的电子装置102、402中的一或多者可执行方法1000。应注意，尽管如结合图10所描述的预训练或登记期间所执行的程序中的一或多者可类似于在获得和评估音频口令(例如，如结合图1到2以及4到8中的一或多者所描述)后即刻执行的程序中的一或多者，结合图10所描述的程序中的一或多者可与在如上文所描述获得音频口令后即刻进行的程序分开和/或在其之前进行。

[0161] 电子装置402可接收(1002)用户音频信号406。举例来说，用户音频信号406可由一或多个麦克风404捕获。举例来说，当用户读提词或打电话时，可接收用户音频信号406。

[0162] 电子装置402可确定(1004)是否在良好声学条件下接收到用户音频信号406。举例来说，电子装置402可确定用户音频信号406的信噪比(SNR)。如果SNR高于SNR阈值，那么电子装置402可确定(1004)在良好声学条件下接收到用户音频信号406。如果SNR不高于(例如，小于或等于)SNR阈值，那么电子装置402可确定(1004)未在良好声学条件下接收到用户音频信号406。如果未在良好声学条件下接收到用户音频信号406，那么电子装置402可丢弃接收到的用户音频信号406并返回以接收(1002)后续用户音频信号406。

[0163] 如果在良好声学条件下接收到用户音频信号406，那么电子装置402可从音频信号406提取(1006)一或多个特征，以获得所提取特征426。举例来说，电子装置402可基于音频信号406确定一或多个MFCC。

[0164] 电子装置402可基于一或多个通用语音模型(例如,UMB)确定(1008)音频信号406的唯一性量度430。在一些配置中,唯一性量度可为音频信号406与通用语音模型之间的似然比。在一些配置中,电子装置402可本地确定(例如,计算)唯一性量度430。在其它配置中,电子装置402可从远程装置(例如,服务器、中央服务器)请求和接收唯一性量度430。

[0165] 电子装置402可基于一或多个所提取的特征426执行(1010)语音辨识。举例来说,电子装置402可基于一或多个所提取的特征426来确定一或多个所辨识语音分量440。

[0166] 电子装置402可使唯一性量度430与一或多个所辨识语音分量440对准(1012)。举例来说,电子装置402可使一或多个所辨识语音分量的出现与所述唯一性量度在时间上对准,以产生经对准的语音和唯一性444。

[0167] 电子装置402可更新(1014)一或多个语音分量(例如,所辨识语音分量)的唯一性统计。举例来说,电子装置402可基于对应于语音分量的唯一性量度来更新(1014)语音分量的唯一性统计。在一些配置中,电子装置402可存储当捕获和辨识时对应于一或多个所辨识语音分量得唯一性量度(或基于唯一性量度的值,例如最大值、最小值或平均值)。其后在获得所辨识语音分量时的每一后续时刻,电子装置402可更新唯一性统计。举例来说,电子装置402可基于所存储的唯一性量度(或值)以及当前唯一性量度(或值)来计算一些统计量度(例如,平均值等)。电子装置402接着可存储经更新的统计量度。

[0168] 电子装置402可登记(1016)一或多个语音分量。举例来说,电子装置402可为一或多个所辨识语音分量中的每一者存储数据。另外或替代地,电子装置402可将所辨识语音分量中的一或多者指定为对于口令建议来说足够唯一或强(例如,如果语音分量具有大于阈值的对应唯一性量度或唯一性统计)。举例来说,在一些配置中,在最初接收到对口令评估的音频口令之前,电子装置402可提供一或多个所建议语音分量。

[0169] 电子装置402可适应(1018)用户语音模型。举例来说,电子装置402可通过更新用户语音模型的音素数据和/或权重来适应或修改用户语音模型(例如,其可最初基于通用语音模型)。在一些配置中,适应(1018)用户语音模型可包含更新一或多个模型参数(例如,GMM分量)。具体地说,适应(1018)可通过更新GMM的平均值和/或混录权重来执行。

[0170] 电子装置402可确定(1020)是否存在充分的数据供用户语音模型准确地描述用户的语音。举例来说,电子装置402可确定是否已捕获阈值数目和/或某些音素,使得用户语音模型足够细化以准确地反映真实用户的语音。如果不存在充分的数据,那么电子装置402可继续接收(1002)用户音频信号。

[0171] 如果存在充分的数据,那么电子装置402可提供(1022)用户语音模型。举例来说,电子装置402可使用户语音模型可用于音频口令强度评估和/或建议,如上文所描述。应注意,尽管可提供(1022)用户语音模型来使用,但方法1000可反复数次和/或连续,以便进一步适应和/或细化用户语音模型。

[0172] 图11是说明其中可实施用于评估音频口令的强度的系统和方法的电子装置1102的另一更具体配置的框图。结合图11描述的电子装置1102可为结合图1和4描述的电子装置102、402中的一或多者的实例。

[0173] 电子装置1102包含一或多个麦克风1104、口令评估模块1108、口令反馈模块1112和一或多个输出装置1116。包含于电子装置1102中的分量中的一或多者可对应于包含于结合图1和4中的一或多者描述的电子装置102、402中的一或多者中的组件中的一或多者和/

或可类似于其而起作用。

[0174] 电子装置1102可包含一或多个输入装置1166。输入装置1166的实例包含触摸屏、触控板、图像传感器(例如,相机)、键盘(例如,物理和/或软件键盘)、小键盘(例如,物理和/或软件小键盘、指纹扫描器、额外麦克风、定向传感器(例如,倾斜传感器)、运动传感器(例如,加速计)、GPS模块、压力传感器等。一或多个输入装置1166可获得或接收一或多个输入1168。可将所述一或多个输入1168提供到口令评估模块1108。

[0175] 一或多个麦克风1104可捕获音频信号1106。音频信号1106可包含音频口令。可将音频信号1106提供到口令评估模块1108。

[0176] 口令评估模块1108可获得(例如,接收)一或多个麦克风1104所捕获的音频信号1106。如上文所描述,音频信号1106可包含音频口令。口令评估模块1108可基于测量音频信号1106的一或多个唯一特性来评估音频口令的强度。这可如上文结合图1到2以及4到8中的一或多者所描述来实现。

[0177] 口令评估模块1108可任选地包含额外验证输入评估模块1170。额外验证输入评估模块1170可结合音频口令考虑一或多个额外验证输入1168。举例来说,如果结合字母数字代码或指纹扫描使用音频口令,那么强度得分可反映音频口令与一或多个额外验证输入(如果利用)的组合所提供的额外验证强度。在一些配置中,电子装置1102(例如,口令评估模块1108)可获得一或多个额外验证输入1168。举例来说,一些配置可允许使用其它模态,例如视频陀螺/加速计传感器,键盘,指纹传感器等。在一些方法中,一或多个此类模态可用于具有较少唯一性或辨别强度的(短语、句子等)的一或多个部分。举例来说,当用户发出具有低唯一性的词语(例如,具有较小可辨别得分的词语“学校”)时,电子装置1102可获得或接收一或多个额外验证输入1168。

[0178] 所述一或多个额外验证输入1168的实例如下给出。在电子装置1102具有手势辨识的配置中,电子装置1102可接收用户所输入的示意动作(例如,触摸屏图案、触摸垫图案、相机所捕获的视觉手势图案等)。所述示意动作可为用户创建或预定义的。在电子装置1102包含相机的配置中,电子装置1102可捕获用户的一或多个图像,例如用户的脸部、眼睛、鼻子、嘴唇、面部形状和/或更多的唯一信息,例如具有音频信号1106的虹膜。举例来说,包含于电子装置1102中的相机可(例如,通过用户)瞄准以捕获用户的脸部的全部或部分。

[0179] 在电子装置1102包含一或多个运动和/或定向传感器(例如,陀螺仪、加速计、倾斜传感器等)的配置中,电子装置1102可获得运动和/或定向信息。举例来说,用户可以用户创建或预定义的方式来定向和/或移动电子装置1102(例如,电话)。举例来说,电子装置1102可连同音频信号1106编码陀螺和/或加速计传感器信息。

[0180] 在电子装置1102包含物理或软件(例如,触摸屏或显示器上)小键盘或键盘的配置中,电子装置1102可连同音频信号1106接收数值代码、文本和/或字母数字串(例如,由用户键入)。在电子装置1102包含指纹传感器的配置中,电子装置1102可接收指纹(例如,当用户触摸或握持指纹传感器时)。

[0181] 在电子装置1102包含多个麦克风1104的配置中,电子装置1102可获得(例如,接收和/或确定)音频信号1106的空间方向性信息。举例来说,用户可在相对于电子装置1102的一序列方向(例如,顶部、底部、左、右、前、后、右上、左下等)上说出音频口令。举例来说,用户可朝电子装置1102的底部说出第一个字,朝电子装置1102的顶部说出第二个字,朝电子

装置1102的左侧说出第三个字，且朝电子装置1102的右侧说出第四个字。

[0182] 下文提供关于空间方向性信息的额外细节。在一些配置中，电子装置1102可利用空间音频的整合来获得安全性。举例来说，为了解锁电子装置1102，用户可向某一空间扇区或不同空间扇区(例如，相对于电子装置1102(例如，电话))中发出一序列。

[0183] 电子装置1102(例如，图11中未图示的检验模块)可识别用户(利用扬声器辨识)，且识别空间说话方向序列是否正确。仅充分高的扬声器辨识可能性与正确空间序列的组合将解锁电子装置1102。举例来说，在一些配置中，电子装置1102可如下执行空间音频/扬声器辨识特征的检验。电子装置1102可初始化提示，接收来自电子装置1102前面的话语，接收来自电子装置1102左侧的话语，接收来自电子装置1102顶部的话语，且接收来自电子装置1102左侧的话语。在初始提示之后，电子装置1102(具有多个麦克风)提供预定义序列的空间音频拾取。在这些配置中，用户可需要知晓向正确的空间扇区中发出音频口令(例如，语句)的序列。举例来说，用户可说：“我最喜欢的”-切换扇区-“宠物的”-切换扇区-“名字是”-切换扇区-“巴尼”)。

[0184] 在一些配置中，每一空间扇区中的话语的时序和/或持续时间可为检验程序的一部分(例如，在前扇区中2秒，在顶部扇区中5秒，在右扇区中3秒等)。举例来说，电子装置1102可经由话音提示或通过检测按钮或屏幕的推动而起始话音记录过程。电子装置1102可根据预定义序列(例如，激活的空间扇区和/或每一空间扇区的时序(持续时间)的序列)，在不同空间扇区中起始收听。如果电子装置1102在每一空间扇区(上下文相依或独立发声)中识别到真实用户，那么电子装置1102准予接入。

[0185] 更具体地说，电子装置1102可根据以下方法或程序来操作。电子装置1102可用话音提示和/或在接收到(例如，按钮或触摸屏的)输入时起始语音记录。电子装置1102可根据预定义序列在不同空间扇区中起始收听。举例来说，电子装置1102可在一序列所激活空间扇区中接收音频。在一些配置中，电子装置1102可根据每一空间扇区中的时序(例如，持续时间)序列来接收音频。

[0186] 如果电子装置1102在每一空间扇区(上下文相依或独立发声)中识别到真实用户(例如，所要扬声器)，那么电子装置1102准予接入。举例来说，电子装置1102可允许用户接入电子装置1102的较多功能性(例如，应用程序、话音呼叫等)。

[0187] 在一个实例中，用户可从相对于所述装置的一个特定方向发出口令、密码或词语序列(例如，“句子”)。在另一实例中，用户可在一序列方向上发出一句子的若干部分。另外或替代地，可要求用户以某一时序发出所述句子的不同部分。另外或替代地，可利用多个用户的话音。举例来说，第一用户可从电子保险箱的左侧发出口令，同时第二用户可从电子保险箱的右侧发出口令，以便解锁所述保险箱。可独立地或结合其它量度(例如，人脸辨识、指纹辨识等)实施空间音频安全特征。

[0188] 在一些配置中，可需要音频口令结合一或多个额外验证输入来通过多个准则，以设定口令(例如，具有一或多个额外验证输入1168的组合音频口令)。举例来说，口令评估模块1108可要求音频口令提供最小唯一性，且一或多个额外验证输入1168满足一或多个额外准则。可对唯一性阈值和/或一或多个额外准则进行加权。

[0189] 在一些配置中，额外验证输入评估模块1170可基于音频信号1106和/或一或多个额外验证输入1168来忽视一或多个阈值。举例来说，如果指纹扫描提供额外验证强度，那么

口令评估模块1108可需要较低唯一性阈值或音频口令强度。另外或替代地，如果音频信号1106提供高唯一性，那么口令评估模块1108可需要额外验证输入1168所贡献的较低强度。举例来说，如果音频信号1106提供相对良好的唯一性，那么口令评估模块1108可建议利用2位数值代码。然而，如果音频信号1106提供相对较弱的唯一性，那么口令评估模块1108可建议利用4位数值代码和/或指纹扫描。

[0190] 可利用一或多个额外验证输入1168，而无时序和/或序列限制。在一些实例中，电子装置1102可在接收到音频信号1106之前、期间或之后的任何时间获得一或多个额外验证输入1168。

[0191] 在其它实例中，电子装置1102可要求(或经配置以要求)相对于音频信号1106的接收以某一时序约束条件和/或以某一序列接收所述一或多个额外验证输入1168。在一实例中，电子装置1102可要求(或经配置以要求)在接收到音频信号1106之前、期间和/或之后的某一时间周期内接收一或多个额外验证输入1168。举例来说，电子装置1102可要求在音频口令的较弱语音分量期间接收额外验证输入1168。另外或替代地，电子装置1102可要求(或经配置以要求)以特定序列(例如，在语音分量之前、在语音分量之后、在语音分量之间，以具有其它额外验证输入的序列等)接收额外验证输入1168。在一些配置中，电子装置1102可以增加复杂性的次序添加(和/或建议添加)一或多个额外验证输入1168。另外或替代地，电子装置1102可要求添加一或多个额外验证输入1168，直到口令(例如，结合一或多个额外验证输入1168的音频口令)超过最小所需强度为止。

[0192] 在一些配置中，口令评估模块1108可任选地包含输入降级模块1172。输入降级模块1172可使音频信号1106和/或额外验证输入1168降级。举例来说，口令评估模块1108可将信息从音频信号1106去除(例如，下取样、滤除其一或多个部分)。另外或替代地，口令评估模块1108可将信息从指纹扫描或从用户的脸部的图像或虹膜去除。

[0193] 口令评估模块1108可将评估信息1110提供到口令反馈模块1112。评估信息1110可包含指示口令评估中获得的口令强度和/或信息的信息。举例来说，评估信息1110可包含所提取特征、唯一性量度、口令强度得分和/或其它信息。

[0194] 口令反馈模块1112可提供口令反馈1114。举例来说，口令反馈模块1112可基于对音频口令的强度的评估来告知用户音频口令较弱。提供口令反馈1114可使用户能够确定(例如，选择、提供或创建)足够强的音频口令。口令反馈1114可包含口令强度得分、一或多个语音分量候选者(例如，所推荐或建议的语音分量)、一或多个所建议动作和/或一或多个消息。举例来说，口令反馈1114可包含口令强度得分和指示音频口令较弱的消息。另外或替代地，口令反馈1114可包含用户可用于创建较强音频口令的一或多个所建议语音分量。在一些配置中，电子装置1102可提供由所建议的语音分量组成的所建议合成(例如未知)字作为口令反馈1114。另外或替代地，口令反馈1114可包含用户可提供额外验证输入(例如，文本、数值代码、字母数字串、空间方向性、额外生物计量(例如面部扫描、虹膜扫描、指纹等))的所建议动作。

[0195] 在一些配置中，口令反馈模块1112可提供一或多个口令建议。举例来说，电子装置1102(例如，口令反馈模块1112)可识别具有足够高的唯一性或与一或多个其它模型(例如，通用语音模型、通用模型、UBM等)的区别来识别一或多个语音分量(例如，发声、音素等)。举例来说，口令反馈模块1112可经由一对语音辨识和扬声器检验系统，基于用户针对每一音

素的话音的唯一性来识别一或多个语音分量。接着，口令反馈模块1112可产生一些可能候选语音分量(例如，音素、音节、发声、口令等)，其具有高“唯一性”，使得用户可选择一或多个候选语音分量来创建口令。举例来说，电子装置1102可显示口令反馈1114，例如：“你可使用/啊/、/k/、…、<三角形>、<高通>、…、”。另外或替代地，可为用户发出的口令提供具体口令反馈1114，以较多地加强所述口令(例如，“你的口令具有60％强度。话语/嗯/可被/啊/...代替”)。

[0196] 在一些配置中，电子装置1102(例如，口令反馈模块1112)可以多模态提供口令建议。如上文所描述，例如，口令反馈模块1112可提供口令反馈1114，其建议一或多个额外验证输入1168(例如，文本、数值代码、字母数字串、空间方向性、额外生物计量(例如，面部扫描、虹膜扫描、指纹等))。

[0197] 口令反馈模块1112可将口令反馈1114提供到一或多个输出装置1116。一或多个输出装置1116可因此向用户中继或传达口令反馈1114。举例来说，输出装置1116(例如，显示器、触摸屏、扬声器等)可中继与音频口令的强度相关联的标记1174。在一些配置中，这可经由如结合图1所描述的一或多个GUI来实现。在一个方法中，显示面板可显示口令强度得分。另外或替代地，扬声器可输出声学信号(例如，文字到语音)，其指示口令强度得分(例如，“你的口令较弱”、“你的口令强度为60％”等)。

[0198] 在一些配置中，输出装置1116可中继一或多个建议(例如，候选语音分量1176、额外验证输入选项1178等)。举例来说，显示面板可显示一或多个候选语音分量1176，例如音素、音节、字、发声和/或短语(例如“/啊/、/嗯/、/k/、/三角形/、/海市蜃楼/”)。另外或替代地，扬声器可输出声学信号以中继一或多个建议(例如，“请将/啊/、/嗯/、/k/、/三角形/、/海市蜃楼/和/或额外输入类型添加到你的口令”)。

[0199] 使用音频口令(例如，独立音频口令和/或具有一或多个额外验证输入1168(例如空间方向性)的音频口令等)来获得安全可应用于许多不同类型的电子装置1102(例如，其可包含麦克风阵列1104)。举例来说，此安全特征可应用于智能电话、平板装置、电子门锁、门传感器、相机、智能按键、膝上型计算机、桌上型计算机、游戏系统、汽车、缴费查询一体机(例如，作为验证交易的一种方式)，电视机、音频装置(例如，mp3播放器、iPod、压缩光盘(CD)播放器等)、音频/视频装置(例如，数字视频记录器(DVR)、蓝光播放器、数字视频光盘(DVD)播放器等)、家用电器、恒温器、保险箱等。另外或替代地，此安全特征可远程应用(例如，应用于远程装置)。举例来说，用户可在智能电话上提供音频口令(例如，句子、密码、口令等)，其可将验证凭证或命令提供到电子门锁，来解锁/锁定门(例如，家门、车门、办公室门等)。在另一实例中，用户可在智能电话、膝上型计算机或平板计算机上提供空间音频代码，以向远程服务器验证来进行网站验证、交易(例如，购买、银行业务)验证等。

[0200] 图12是说明用于评估音频口令的强度的方法1200的更具体配置的流程图。结合图1、4和11描述的电子装置102、402、1102中的一或多者可执行方法1200。

[0201] 电子装置1102可获得(1202)一或多个麦克风1104所捕获的音频信号1106。这可如上文结合图1到2、4到8以及11中的一或多者所描述来实现。音频信号1106可包含音频口令。

[0202] 电子装置1102可获得至少一个额外验证输入1168。此操作可如上文结合图(例如，图1、4和11)中的一或多者所描述来实现。举例来说，电子装置可获得(1204)一或多个额外验证输入1168，例如文本、数值代码、字母数字串、空间方向性和/或额外生物计量(例如指

纹扫描、用户脸部的相机图像或虹膜等)。

[0203] 电子装置1102可任选地使音频信号1106和/或额外验证输入1168降级(1206)。此操作可如上文结合图(例如,图1、4和11)中的一或多者所描述来实现。举例来说,电子装置1102可将信息从音频信号1106去除(例如,下取样、滤除其一或多个部分)。另外或替代地,口令评估模块1108可将信息从指纹扫描或从用户的脸部的图像或虹膜去除。

[0204] 电子装置1102可结合至少一个额外验证输入1168来评估(1208)音频口令的强度。举例来说,电子装置1102可结合音频口令考虑一或多个额外验证输入1168。举例来说,如果结合字母数字代码或指纹扫描使用音频口令,那么强度得分可反映音频口令与一或多个额外验证输入的组合所提供的额外验证强度。

[0205] 如果结合至少一个额外验证输入1168的音频口令的强度较弱,那么电子装置1102可提供(1210)口令反馈1114。这可如上文结合图1到2、4到8以及11中的一或多者所描述来实现。举例来说,电子装置1102可基于对结合至少一个额外验证输入1168的音频口令的强度的评估(例如,当口令强度得分不大于一值时),告知(1206)用户音频口令较弱。口令反馈1114可包含口令强度得分、一或多个语音分量候选者(例如,所推荐或建议的语音分量)、一或多个所建议动作和/或一或多个消息。举例来说,口令反馈1114可包含口令强度得分和指示音频口令较弱的消息。另外或替代地,口令反馈1114可包含用户可用于创建较强音频口令的一或多个所建议语音分量。另外或替代地,口令反馈1114可包含用户可提供额外验证输入1168(例如,文本、数值代码、字母数字串、空间方向性、额外生物计量(例如面部扫描、虹膜扫描、指纹等))的所建议动作。

[0206] 图13是说明其中可实施用于评估音频口令的强度的系统和方法的无线通信装置1302的一个配置的框图。图13中说明的无线通信装置1302可为本文所述的电子装置102、402、1102中的一或多者的实例。无线通信装置1302可包含应用处理器1384。应用程序处理器1384通常处理指令(例如,运行程序)以执行无线通信装置1302上的功能。应用程序处理器1384可耦合到音频译码器/解码器(编解码器)1382。

[0207] 音频编解码器1382可用于对音频信号进行译码和/或解码。音频编解码器1382可耦合到至少一个扬声器1335、耳机1337、输出插孔1339和/或至少一个麦克风1380。扬声器1335可包含一或多个将电或电子信号转换为声学信号的电声转换器。举例来说,扬声器1335可用于播放音乐或输出扬声器电话对话等。耳机1337可为可用于向用户输出声学信号(例如,话语信号)的另一扬声器或电声转换器。举例来说,可使用听筒1337使得仅用户可确实地听到声学信号。输出插孔1339可用于将其它装置(例如头戴式耳机)耦合到无线通信装置1302以用于输出音频。扬声器1335、听筒1337和/或输出插孔1339可通常用于从音频编解码器1382输出音频信号。至少一个麦克风1380可为将声学信号(例如用户的话音)转换为提供至音频编解码器1382的电或电子信号的声电转换器。

[0208] 在一些配置中,音频编解码器1382可包含口令评估模块1308a和/或口令反馈模块1312a。另外或替代地,应用程序处理器1384可包含口令评估模块1308b和/或口令反馈模块1312b。口令评估模块1308a-b和/或口令反馈模块1312a-b可为上文结合图1、4和11中的一或多者描述的口令评估模块108、408、1108和/或口令反馈模块112、412、1112的实例。在其它配置中,口令评估模块1308a和口令反馈模块1312a中的一或多者可分别从音频编解码器1382和应用程序处理器1384在无线通信装置1302上实施。

[0209] 应用处理器1384还可耦合到电力管理电路1394。电力管理电路1394的一个实例是电力管理集成电路(PMIC),其可用于管理无线通信装置1302的电力消耗。电力管理电路1394可耦合到电池1396。电池1396可通常将电力提供到无线通信装置1302。举例来说,电池1396和/或功率管理电路1394可耦合到包含于无线通信装置1302中的元件中的至少一者。

[0210] 应用处理器1384可耦合到至少一个输入装置1398以用于接收输入。输入装置1398的实例包含红外传感器、图像传感器、加速计、触摸传感器、小键盘等。输入装置1398可允许用户与无线通信装置1302交互。应用程序处理器1384还可耦合到一或多个输出装置1301。输出装置1301的实例包含打印机、投影仪、屏幕、触觉装置等。输出装置1301可允许无线通信装置1302产生可由用户体验的输出。

[0211] 应用程序处理器1384可耦合到应用程序存储器1303。应用程序存储器1303可为能够存储电子信息的任何电子装置。应用存储器1303的实例包含双数据速率同步动态随机存取存储器(DDRAM)、同步动态随机存取存储器(SDRAM)、快闪存储器等。应用存储器1303可为应用处理器1384提供存储。举例来说,应用存储器1303可存储在应用程序处理器1384上运行的程序的功能的数据和/或指令。

[0212] 应用程序处理器1384可耦合到显示控制器1305,所述显示控制器又可耦合到显示器1307。显示控制器1305可为用于在显示器1307上产生图像的硬件块。举例来说,显示器控制器1305可将来自应用程序处理器1384的指令和/或数据转译为可呈现在显示器1307上的图像。显示器1307的实例包含液晶显示器(LCD)面板、发光二极管(LED)面板、阴极射线管(CRT)显示器、等离子显示器等。

[0213] 应用程序处理器1384可耦合到基带处理器1386。基带处理器1386通常处理通信信号。举例来说,基带处理器1386可对接收到的信号进行解调和/或解码。另外或或者,基带处理器1386可对信号进行编码和/或调制以准备发射。

[0214] 基带处理器1386可耦合到基带存储器1309。基带存储器1309可为能够存储电子信息的任何电子装置,例如SDRAM、DDRAM、快闪存储器等。基带处理器1386可从基带存储器1309读取信息(例如,指令和/或数据)和/或将信息写入到基带存储器1309。另外或或者,基带处理器1386可使用存储在基带存储器1309中的指令和/或数据来执行通信操作。

[0215] 基带处理器1386可耦合到射频(RF)收发器1388。RF收发器1388可耦合到功率放大器1390和一或多个天线1392。RF收发器1388可发射和/或接收射频信号。举例来说,RF收发器1388可使用功率放大器1390和至少一个天线1392发射RF信号。RF收发器1388还可使用一或多个天线1392接收RF信号。

[0216] 图14说明可在电子装置1402中利用的各种组件。所说明的组件可位于同一实体结构内或位于单独外壳或结构中。结合图14所描述的电子装置1402可根据本文中所描述的电子装置102、402、1102和无线通信装置1302中的一或多者来实施。电子装置1402包含处理器1417。处理器1417可为通用单芯片或多芯片微处理器(例如,ARM)专用微处理器(例如,数字信号处理器(DSP))、微控制器、可编程门阵列等。处理器1417可被称作中央处理单元(CPU)。尽管在图14的电子装置1402中仅示出单个处理器1417,但在替代配置中,可使用处理器(例如ARM与DSP)的组合。

[0217] 电子装置1402还包含与处理器1417进行电子通信的存储器1411。也就是说,处理器1417可从存储器1411读取信息和/或将信息写入到存储器1411。存储器1411可为能够存

储电子信息的任何电子组件。存储器1411可为随机存取存储器(RAM)、只读存储器(ROM)、磁盘存储媒体、光学存储媒体、RAM中的快闪存储器装置、随处理器一起包含的机载存储器、可编程只读存储器(PROM)、可擦除可编程只读存储器(EPROM)、电可擦除PROM(EEPROM)、寄存器等,包含其组合。

[0218] 数据1415a和指令1413a可存储在存储器1411中。指令1413a可包含一或多个程序、例程、子例程、功能、过程等。指令1413a可包含单个计算机可读语句或许多计算机可读语句。指令1413a可由处理器1417执行以实施上文所描述的方法、功能和程序中的一或多者。执行指令1413a可涉及使用存储在存储器1411中的数据1415a。图14示出一些指令1413b和数据1415b正加载到处理器1417中(其可来自指令1413a和数据1415a)。

[0219] 电子装置1402还可包含用于与其它电子装置通信的一或多个通信接口1421。通信接口1421可基于有线通信技术、无线通信技术或两者。不同类型的通信接口1421的实例包含串行端口、并行端口、通用串行总线(USB)、以太网配接器、电气电子工程师学会(IEEE)1494总线接口、小型计算机系统接口(SCSI)总线接口、红外(IR)通信端口、蓝牙无线通信配接器、第三代合作伙伴计划(3GPP)收发器、IEEE 802.11(“Wi-Fi”)收发器等。举例来说,通信接口1421可耦合到用于发射和接收无线信号的一或多个天线(未展示)。

[0220] 电子装置1402还可包含一或多个输入装置1423和一或多个输出装置1427。不同种类的输入装置1423的实例包含键盘、鼠标、麦克风、遥控器装置、按钮、操纵杆、跟踪球、触控板、光笔等。举例来说,电子装置1402可包含用于捕获声学信号的一或多个麦克风1425。在一种配置中,麦克风1425可为将声学信号(例如,话音、语音)转换成电或电子信号的变换器。不同种类的输出装置1427的实例包含扬声器、打印机等。举例来说,电子装置1402可包含一或多个扬声器1429。在一种配置中,扬声器1429可为将电或电子信号转换为声学信号的变换器。可通常包含在电子装置1402中的输出装置的一个特定类型为显示装置1431。与本文中所公开的配置一起使用的显示装置1431可利用任何合适的图像投影技术,例如阴极射线管(CRT)、液晶显示器(LCD)、发光二极管(LED)、气体等离子体、电致发光或类似者。还可提供显示器控制器1433,用于将存储在存储器1411中的数据转换为显示装置1431上示出的文本、图形和/或移动图像(按需要)。

[0221] 电子装置1402的各种组件可通过一或多个总线耦合在一起,所述总线可以包含电力总线、控制信号总线、状态信号总线、数据总线等。为简单起见,图14中将各种总线说明为总线系统1419。应注意,图14仅说明电子装置1402的一个可能配置。可利用各种其它架构和组件。

[0222] 在以上描述中,有时结合各种术语而使用参考标号。在术语结合参考数字使用的情况下,此可意味着指代图中的一或多者中示出的特定元件。在无参考标号而使用术语的情况下,此可意味着大体上指代所述术语,而不限于任何特定图。

[0223] 术语“确定”涵盖各种各样的动作,且因此“确定”可包含计算、运算、处理、导出、调查、查找(例如,在表、数据库或另一数据结构中查找)、查实等等。并且,“确定”可包含接收(例如,接收信息)、存取(例如,在存储器中存取数据)等。并且,“确定”可包括解析、选择、挑选、建立等等。

[0224] 除非另有明确指定,否则短语“基于”并不意味着“仅基于”。换句话说,短语“基于”描述“仅基于”以及“基于至少”两者。

[0225] 应注意，在相容的情况下，结合本文中所描述的配置中的任一者所描述的特征、功能、过程、组件、元件、结构等中的一或多者可与结合本文中所描述的其它配置中的任一者所描述的功能、过程、组件、元件、结构等中的一或多者进行组合。换句话说，可根据本文中揭示的系统和方法来实施本文中所描述的功能、程序、组件、元件等的任何相容的组合。

[0226] 可将本文中所描述的功能作为一或多个指令而存储在处理器可读或计算机可读媒体上。术语“计算机可读媒体”是指可由计算机或处理器存取的任何可用媒体。作为实例而非限制，此类媒体可包括随机存取存储器(RAM)、只读存储器(ROM)、电可擦除可编程只读存储器(EEPROM)、快闪存储器、压缩光盘只读存储器(CD-ROM)或其它光盘存储装置、磁盘存储器或其它磁性存储装置，或可用于以指令或数据结构的形式存储所要的程序代码且可由计算机存取的任何其它媒体。如本文中所使用，磁盘和光盘包含压缩光盘(CD)、激光光盘、光学光盘、数字影音光盘(DVD)、软性磁盘和Blu-ray®光盘，其中磁盘通常以磁性方式再现数据，而光盘利用激光以光学方式再现数据。应注意，计算机可读媒体可为有形且非暂时性的。术语“计算机程序产品”是指计算装置或处理器，其与可由计算装置或处理器执行、处理或计算的代码或指令(例如，“程序”)结合。如本文中所使用，术语“代码”可指可由计算装置或处理器执行的软件、指令、代码或数据。

[0227] 还可通过传输媒体来传输软件或指令。举例来说，如果使用同轴电缆、光纤电缆、双绞线、数字订户线路(DSL)或无线技术(例如，红外线、无线电和微波)从网站、服务器或其它远程源传输软件，那么同轴电缆、光纤电缆、双绞线、DSL或无线技术(例如，红外线、无线电和微波)包含在传输媒体的定义中。

[0228] 本文中所揭示的方法包括用于实现所描述的方法的一或多个步骤或动作。在不偏离权利要求书的范围的情况下，方法步骤和/或动作可彼此互换。换句话说，除非正描述的方法的适当操作需要步骤或动作的特定次序，否则，在不脱离权利要求书的范围的情况下，可修改特定步骤和/或动作的次序和/或使用。

[0229] 将理解，所附权利要求书不限于上文所说明的精确配置和组件。在不脱离权利要求书的范围的情况下，可在本文中所描述的系统、方法和设备的配置、操作和细节方面进行各种修改、改变和变更。

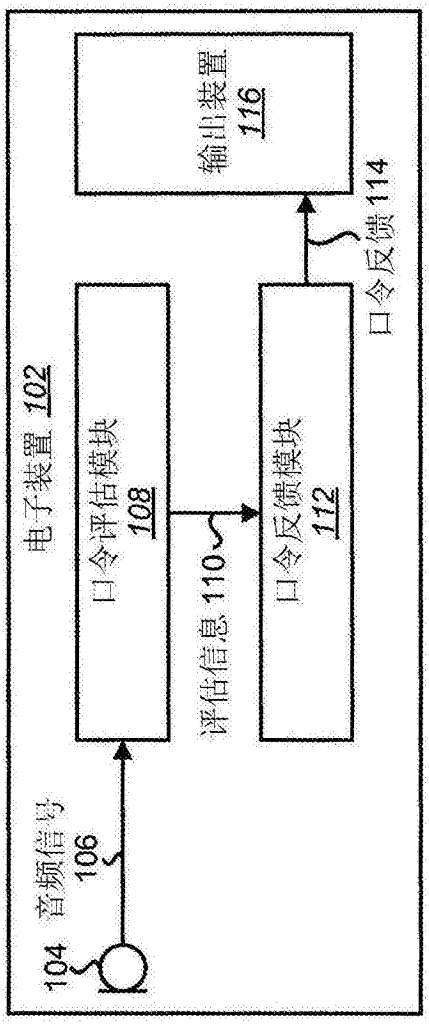

电子装置 102
输出装置 116
口令评估模块 108
口令反馈模块 112
口令反馈 114
评估信息 110
音频信号 106
104

图1

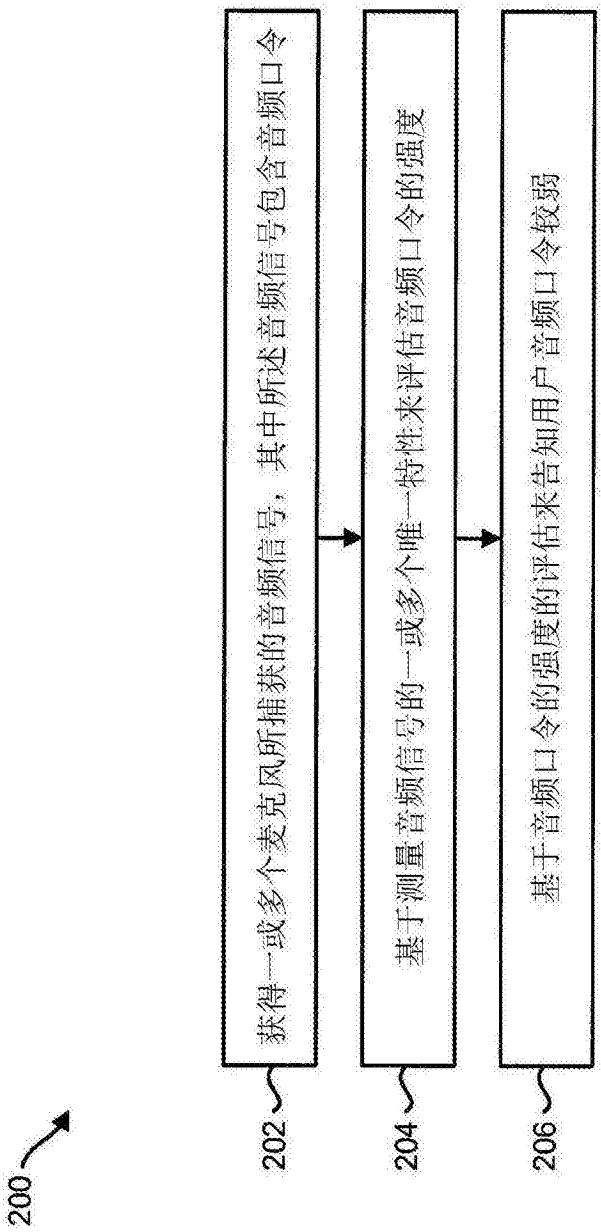

200
获得一或多个麦克风所捕获的音频信号，其中所述音频信号包含音频口令
基于测量音频信号的一或多个唯一特性来评估音频口令的强度
基于音频口令的强度的评估来告知用户音频口令较弱
202
204
206

图2

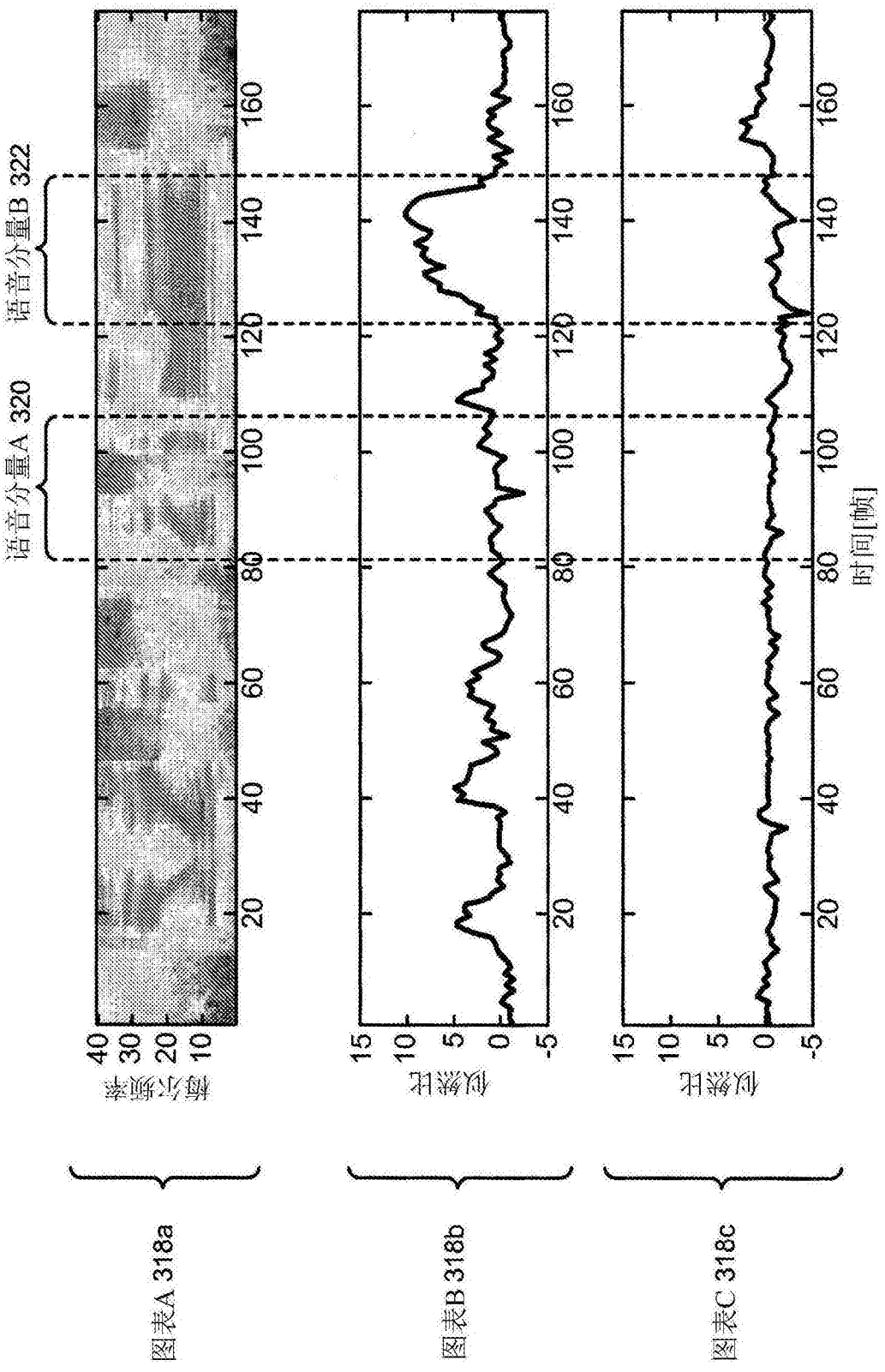

语音分量A 320
语音分量B 322
梅尔频率
似然比
似然比
时间[帧]
图表A 318a
图表B 318b
图表C 318c

图3

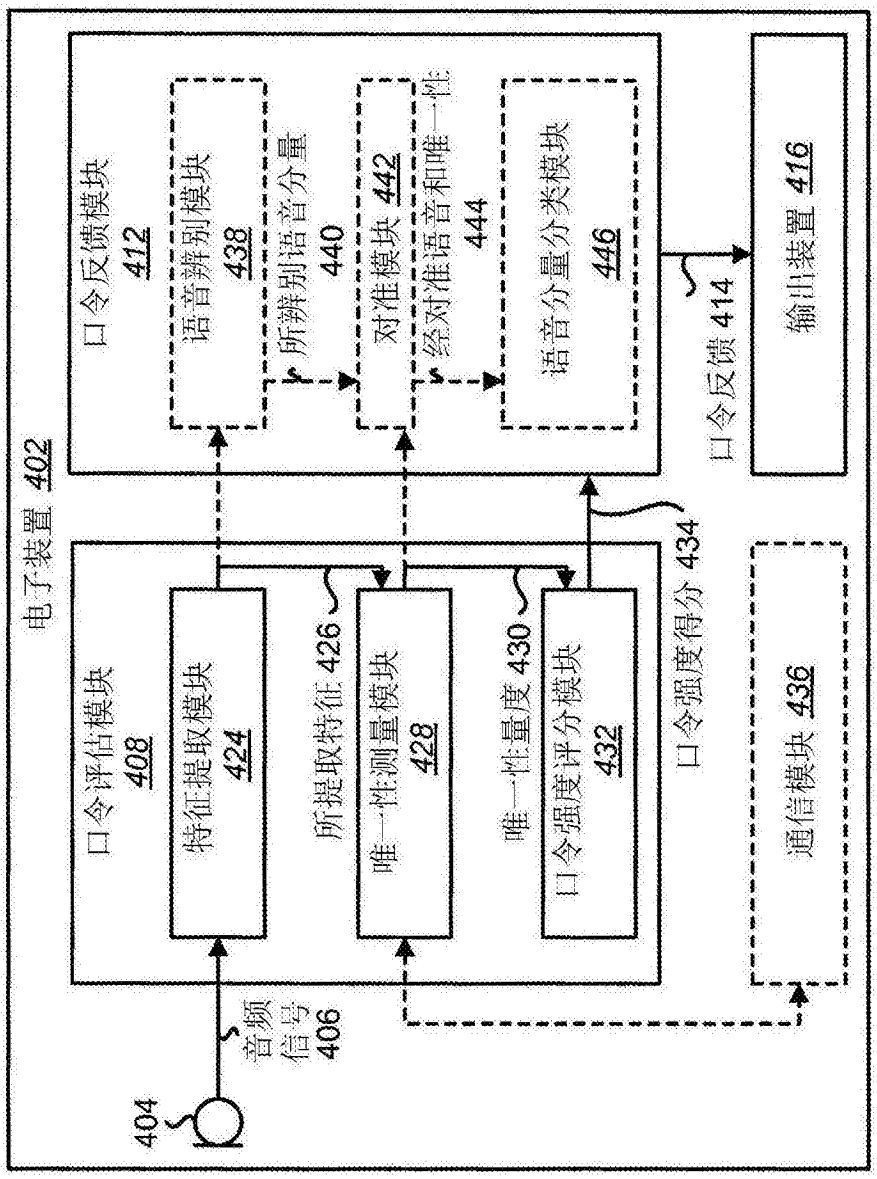

电子装置 402
口令反馈模块 412
语音辨别模块 438
所辨别语音分量 440
对准模块 442
经对准语音和唯一性 444
语音分量分类模块 446
口令反馈 414
输出装置 416
口令评估模块 408
特征提取模块 424
所提取特征 426
唯一性测量模块 428
唯一性量度 430
口令强度评分模块 432
口令强度得分 434
通信模块 436
音频信号 406
404

图4

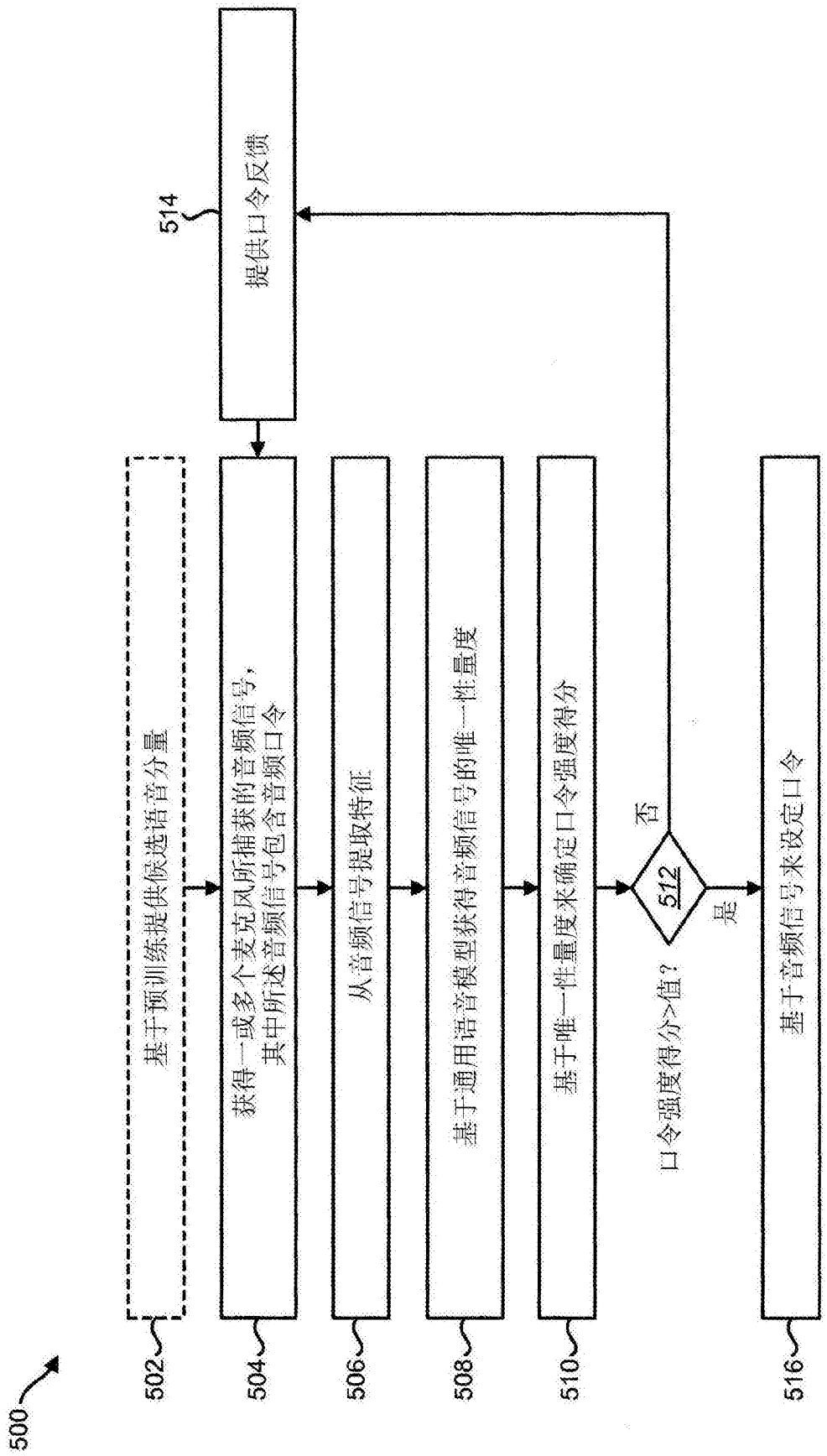

500
502
基于预训练提供候选语音分量
504
获得一或多个麦克风所捕获的音频信号，
其中所述音频信号包含音频口令
506
从音频信号提取特征
508
基于通用语音模型获得音频信号的唯一性量度
510
基于唯一性量度来确定口令强度得分
512
口令强度得分>值？
否
是
514
提供口令反馈
516
基于音频信号来设定口令

图5

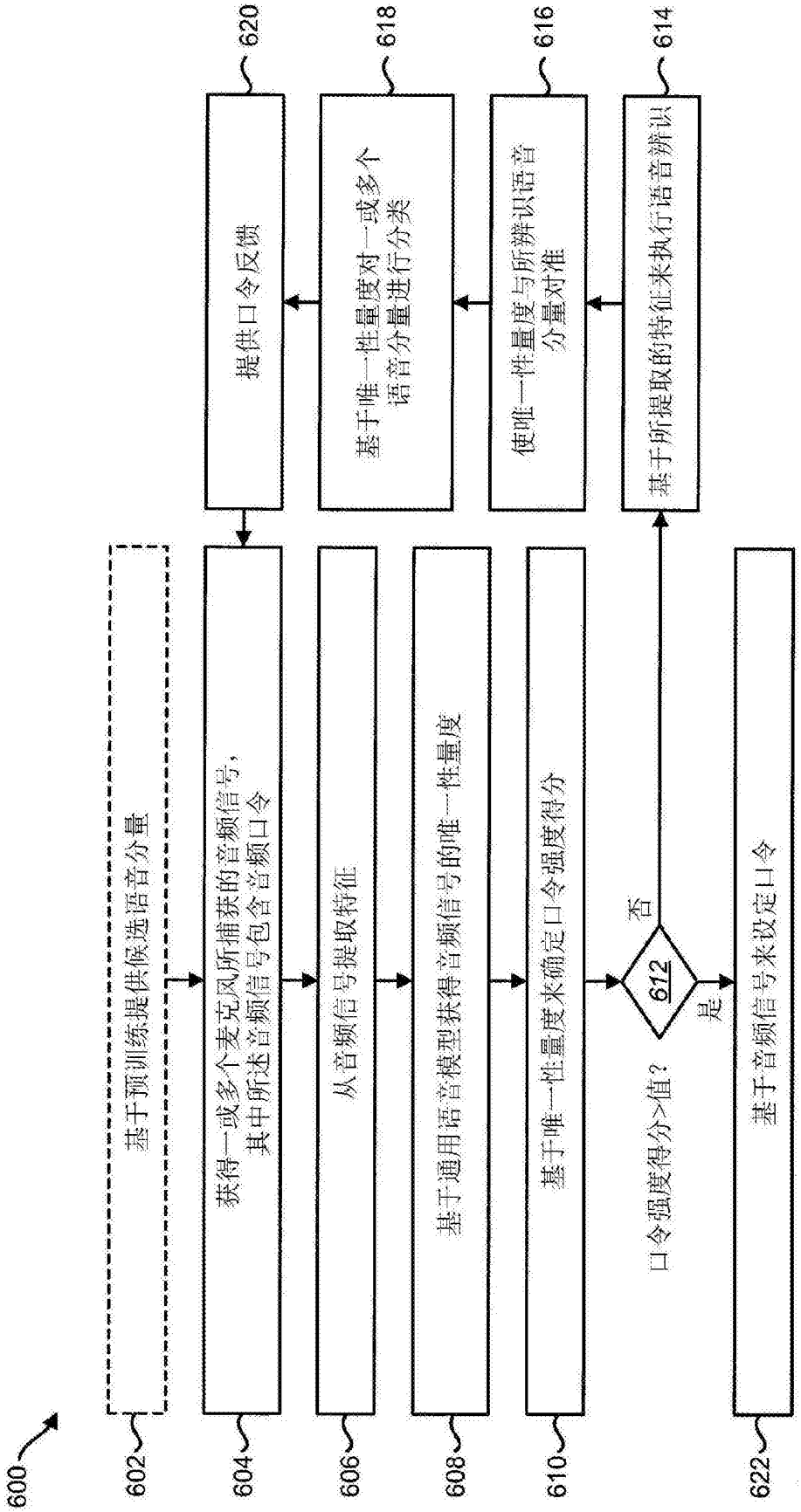

600
602
基于预训练提供候选语音分量
604
获得一或多个麦克风所捕获的音频信号，其中所述音频信号包含音频口令
606
从音频信号提取特征
608
基于通用语音模型获得音频信号的唯一性量度
610
基于唯一性量度来确定口令强度得分
612
口令强度得分>值？
否
是
614
基于所提取的特征来执行语音辨识
616
使唯一性量度与所辨识语音分量对准
618
基于唯一性量度对一或多个语音分量进行分类
620
提供口令反馈
622
基于音频信号来设定口令

图6

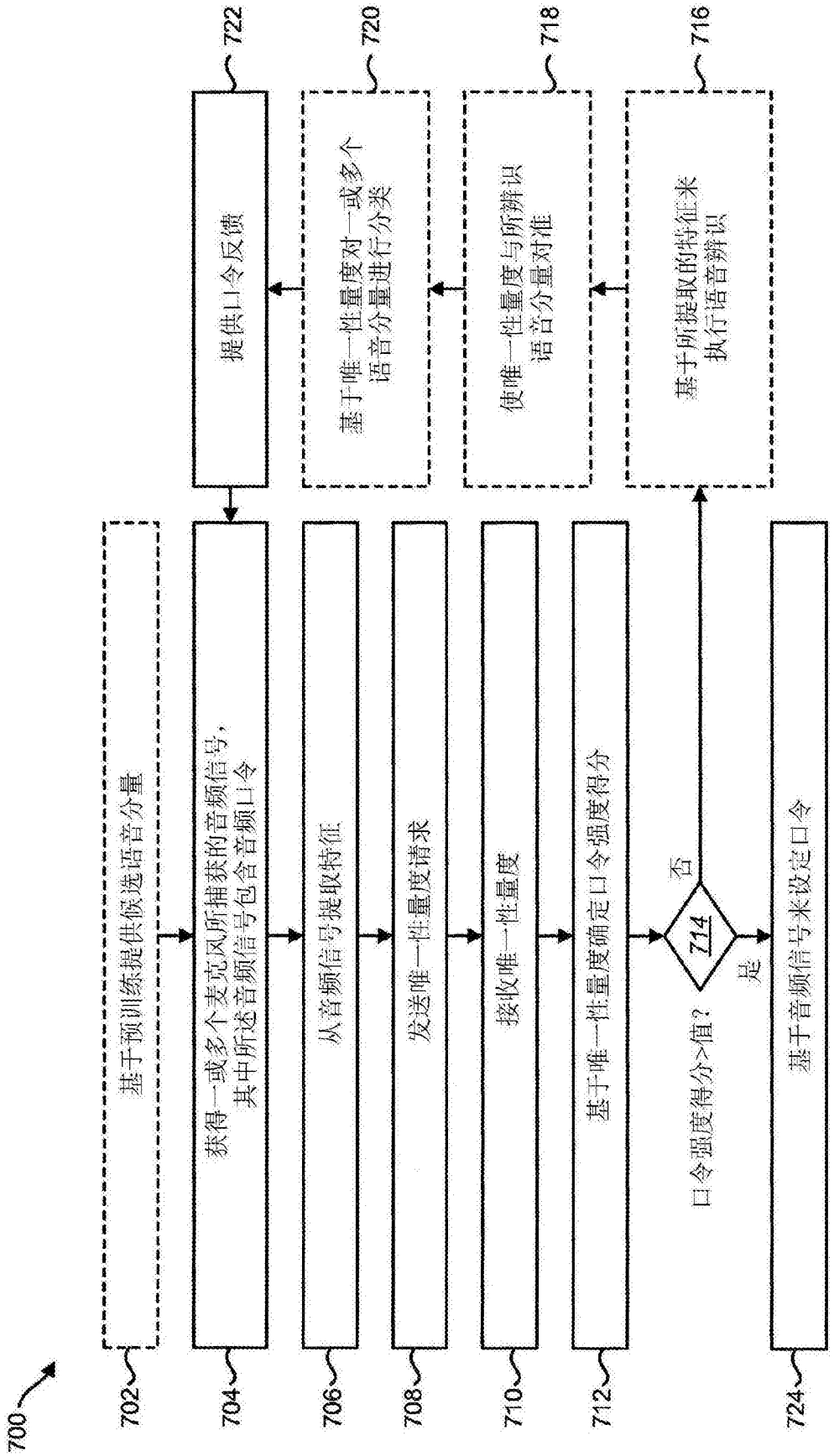

700
702
基于预训练提供候选语音分量
704
获得一或多个麦克风所捕获的音频信号，其中所述音频信号包含音频口令
706
从音频信号提取特征
708
发送唯一性量度请求
710
接收唯一性量度
712
基于唯一性量度确定口令强度得分
714
口令强度得分>值？
否
是
716
基于所提取的特征来执行语音辨识
718
使唯一性量度与所辨识语音分量对准
720
基于唯一性量度对一或多个语音分量进行分类
722
提供口令反馈
724
基于音频信号来设定口令

图7

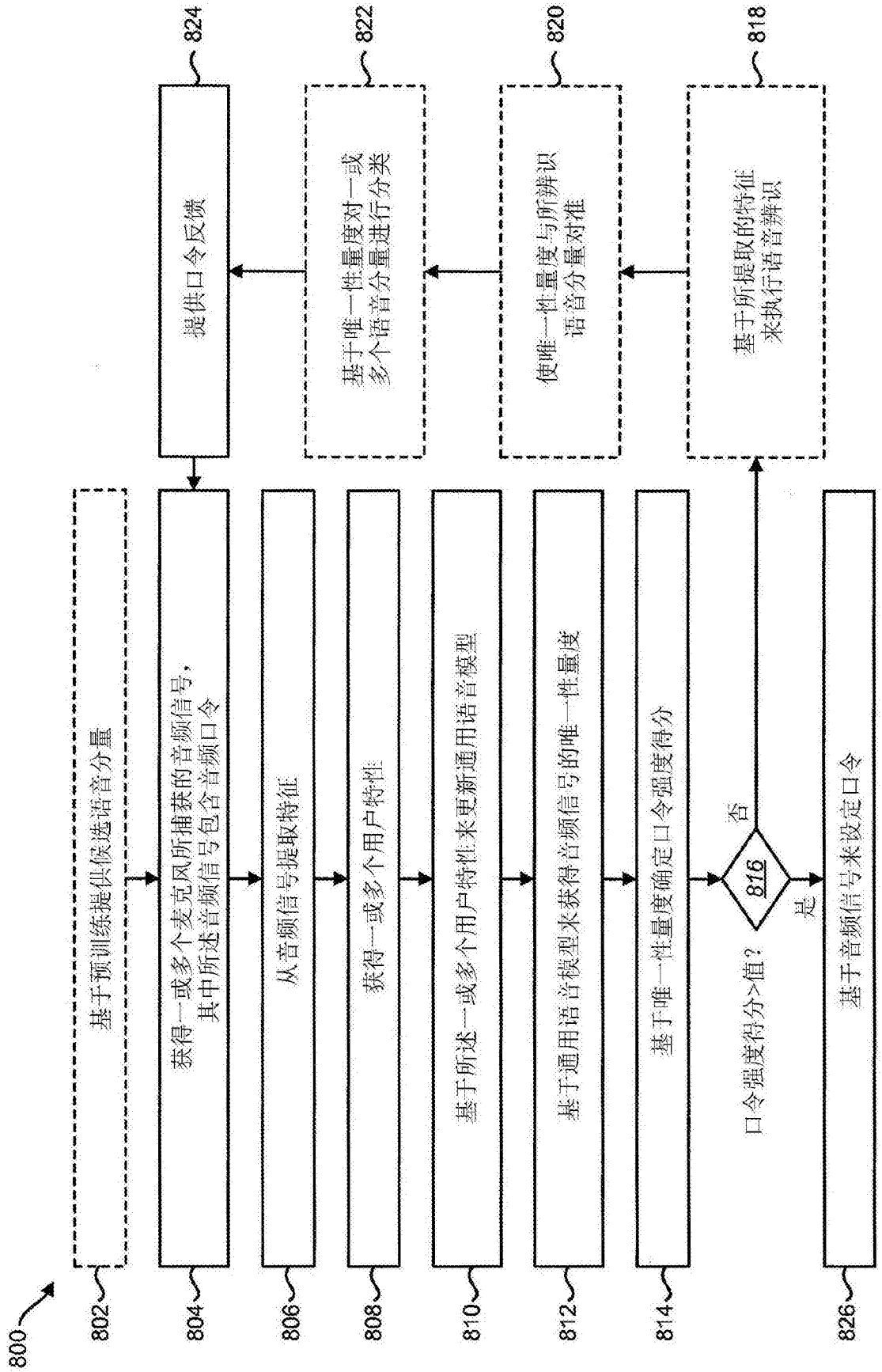

800
802
基于预训练提供候选语音分量
804
获得一或多个麦克风所捕获的音频信号，其中所述音频信号包含音频口令
806
从音频信号提取特征
808
获得一或多个用户特性
810
基于所述一或多个用户特性来更新通用语音模型
812
基于通用语音模型来获得音频信号的唯一性量度
814
基于唯一性量度确定口令强度得分
816
口令强度得分>值？
否
是
818
基于所提取的特征来执行语音辨识
820
使唯一性量度与所辨识语音分量对准
822
基于唯一性量度对一或多个语音分量进行分类
824
提供口令反馈
826
基于音频信号来设定口令

图8

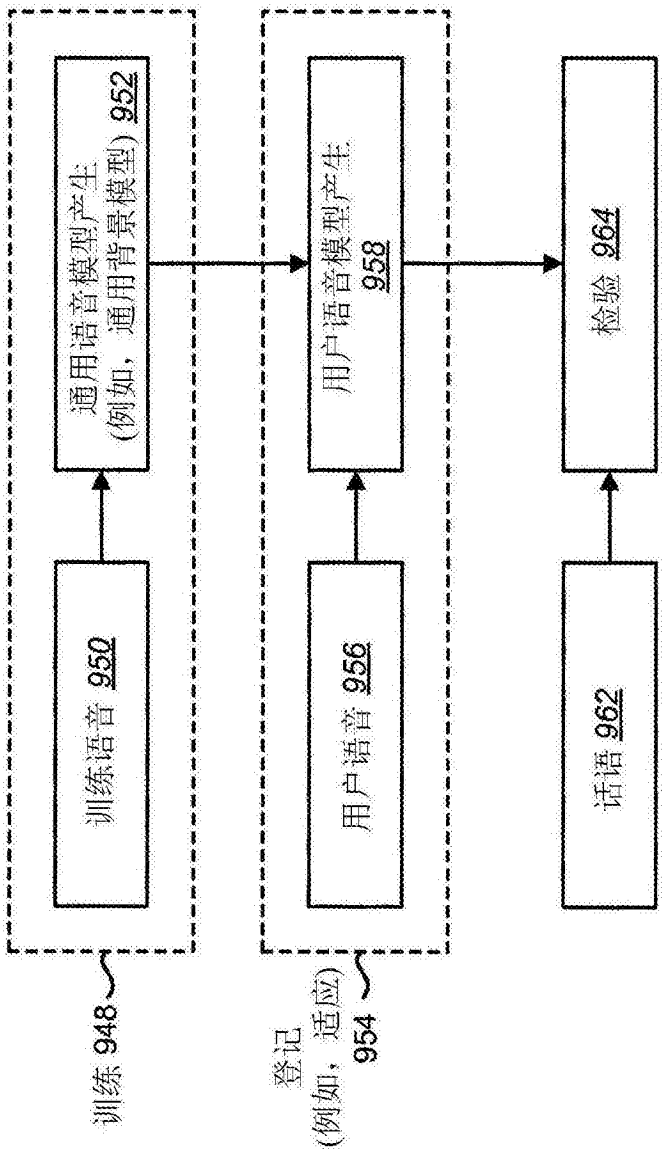

通用语音模型产生
(例如，通用背景模型) 952
用户语音模型产生
958
检验 964
训练语音 950
用户语音 956
话语 962
训练 948
登记
(例如，适应)
954

图9

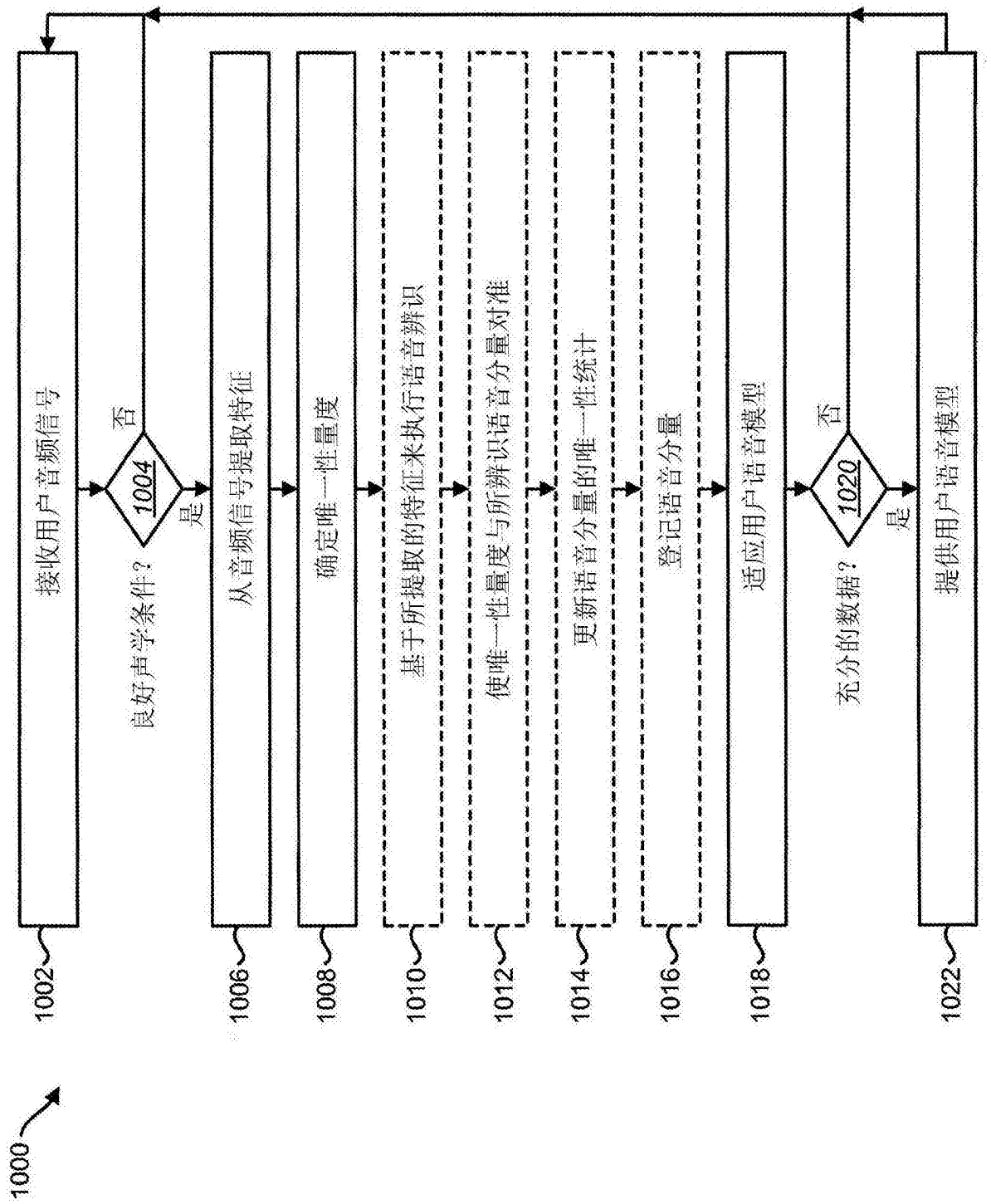

1000
接收用户音频信号
1002
良好声学条件？
1004
否
是
从音频信号提取特征
1006
确定唯一性量度
1008
基于所提取的特征来执行语音辨识
1010
使唯一性量度与所辨识语音分量对准
1012
更新语音分量的唯一性统计
1014
登记语音分量
1016
适应用户语音模型
1018
充分的数据？
1020
否
是
提供用户语音模型
1022

图10

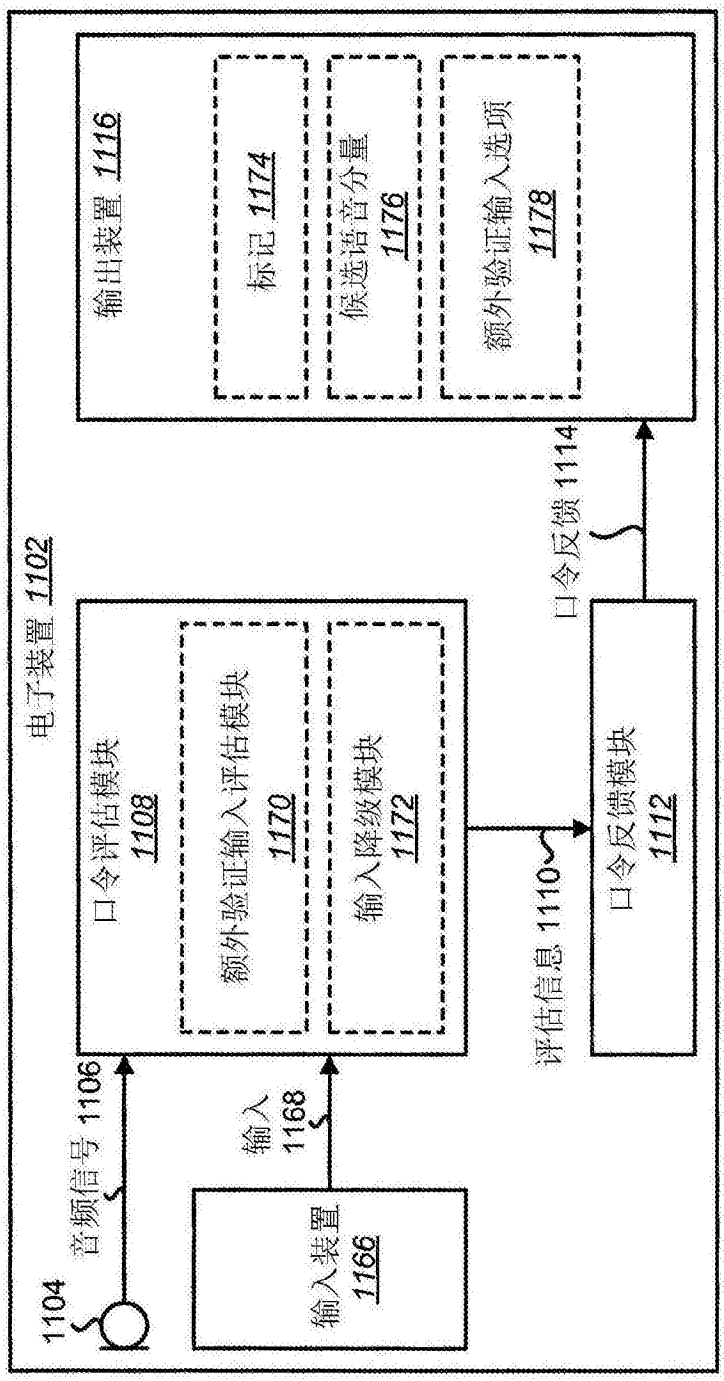

电子装置 1102
1104
音频信号 1106
输入装置 1166
输入 1168
口令评估模块 1108
额外验证输入评估模块 1170
输入降级模块 1172
评估信息 1110
口令反馈模块 1112
口令反馈 1114
输出装置 1116
标记 1174
候选语音分量 1176
额外验证输入选项 1178

图11

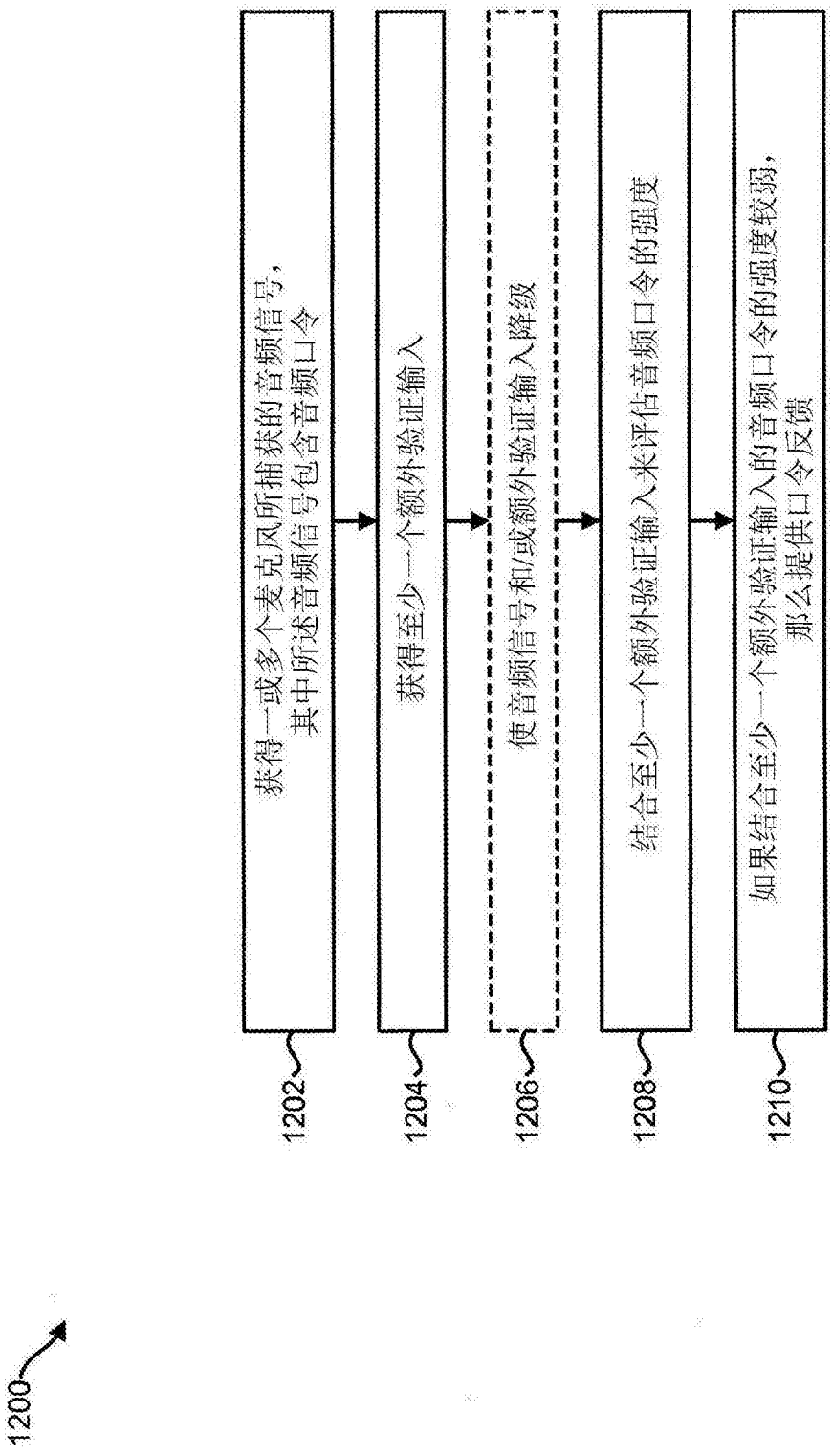

获得一或多个麦克风所捕获的音频信号，其中所述音频信号包含音频口令
获得至少一个额外验证输入
使音频信号和/或额外验证输入降级
结合至少一个额外验证输入来评估音频口令的强度
如果结合至少一个额外验证输入的音频口令的强度较弱，那么提供口令反馈
1202
1204
1206
1208
1210
1200

图12

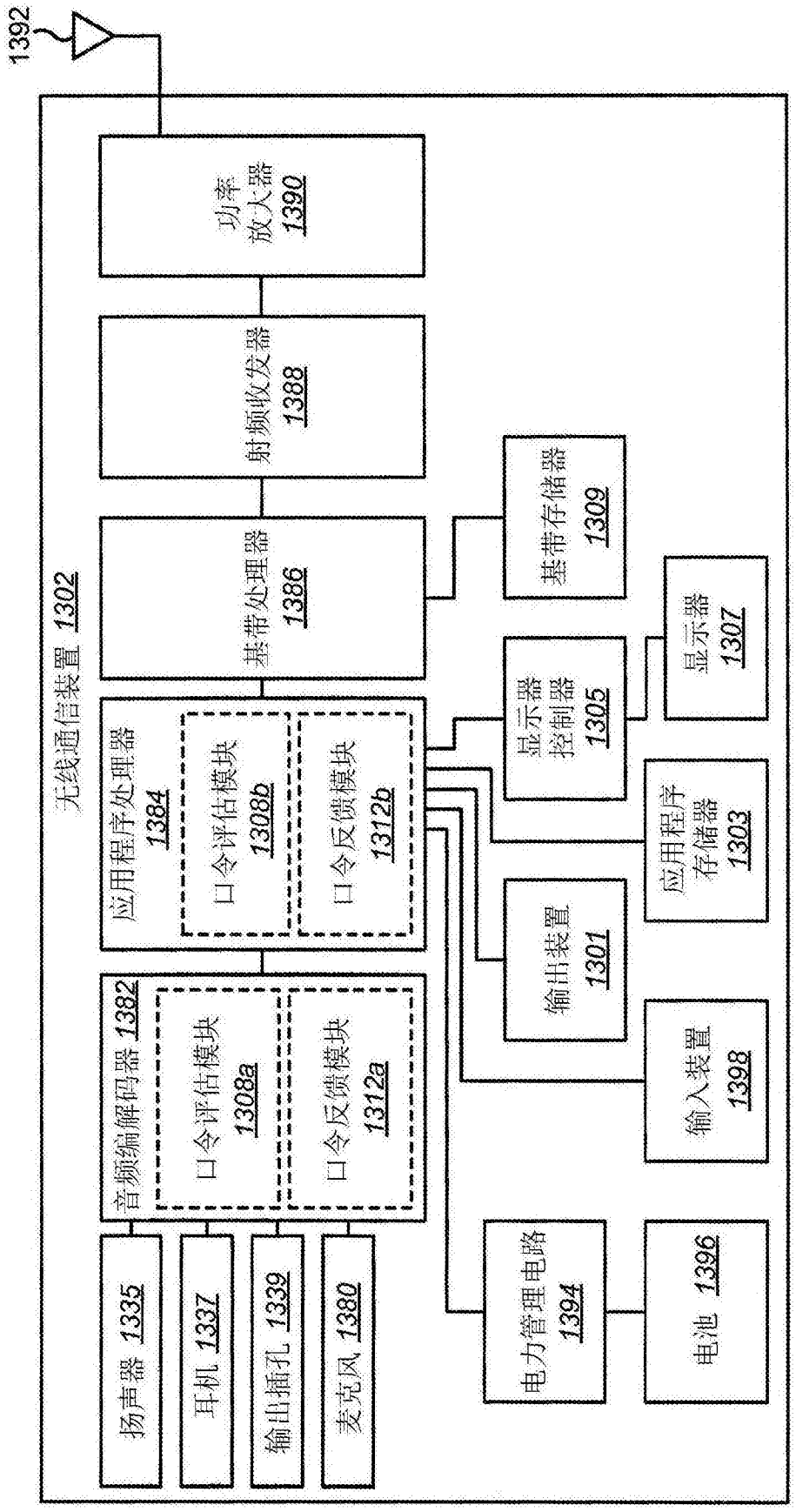

1392
无线通信装置 1302
功率放大器 1390
射频收发器 1388
基带处理器 1386
基带存储器 1309
应用程序处理器 1384
口令评估模块 1308b
口令反馈模块 1312b
显示器控制器 1305
显示器 1307
应用程序存储器 1303
输出装置 1301
输入装置 1398
音频编解码器 1382
口令评估模块 1308a
口令反馈模块 1312a
扬声器 1335
耳机 1337
输出插孔 1339
麦克风 1380
电力管理电路 1394
电池 1396

图13

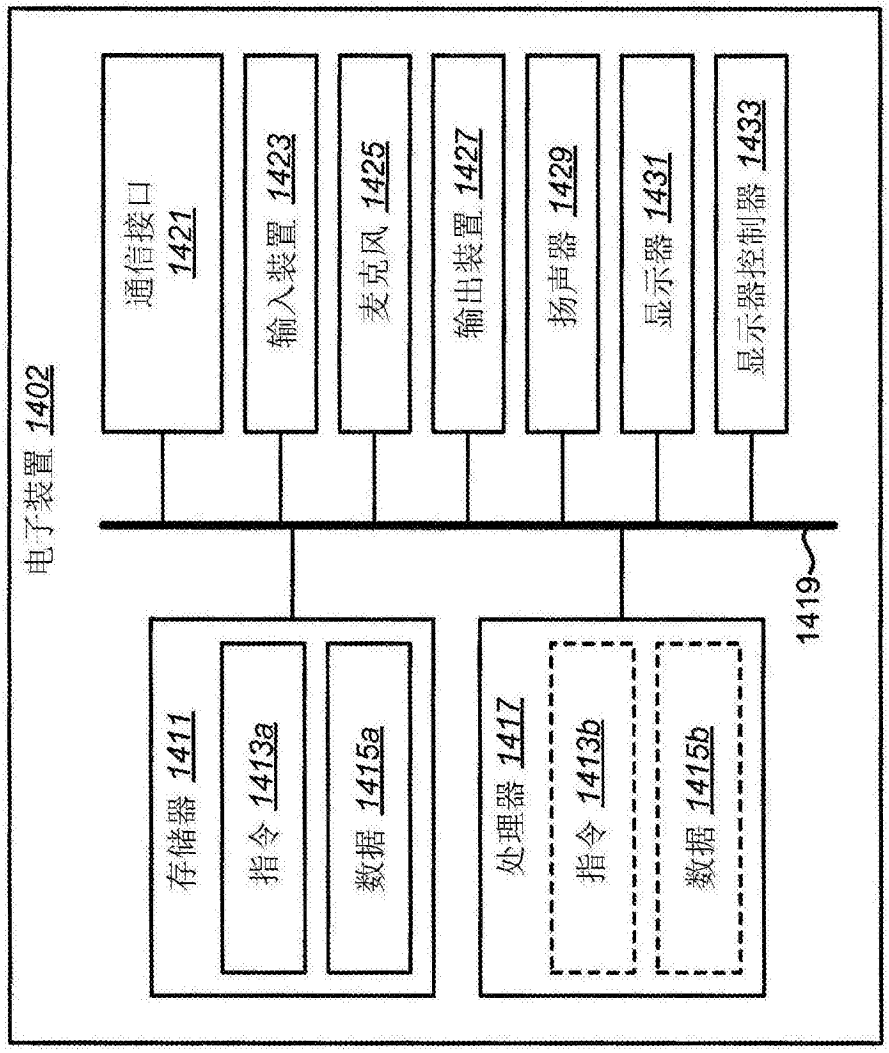

电子装置 1402
通信接口 1421
输入装置 1423
麦克风 1425
输出装置 1427
扬声器 1429
显示器 1431
显示器控制器 1433
1419
存储器 1411
指令 1413a
数据 1415a
处理器 1417
指令 1413b
数据 1415b

图14