



(12) 发明专利

(10) 授权公告号 CN 106663430 B

(45) 授权公告日 2021.02.26

(21) 申请号 201580047427.X  
 (22) 申请日 2015.08.18  
 (65) 同一申请的已公布的文献号  
 申请公布号 CN 106663430 A  
 (43) 申请公布日 2017.05.10  
 (30) 优先权数据  
 14/479,708 2014.09.08 US  
 (85) PCT国际申请进入国家阶段日  
 2017.03.03  
 (86) PCT国际申请的申请数据  
 PCT/US2015/045661 2015.08.18  
 (87) PCT国际申请的公布数据  
 WO2016/039945 EN 2016.03.17  
 (73) 专利权人 高通股份有限公司  
 地址 美国加利福尼亚州  
 (72) 发明人 李敏秀 金泰殊 尹宋克  
 (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

(56) 对比文件  
 CN 101027716 A, 2007.08.29  
 EP 1215661 A1, 2002.06.19  
 CN 102510426 A, 2012.06.20  
 EP 0945851 A2, 1999.09.29  
 US 2004/0162728 A1, 2004.08.19  
 CN 101124623 A, 2008.02.13  
 US 2010/0082342 A1, 2010.04.01  
 CN 101123428 A, 2008.02.13  
 US 2004/0088162 A1, 2004.05.06  
 CN 101145341 A, 2008.03.19  
 US 5832063 A, 1998.11.03  
 CN 101661754 A, 2010.03.03  
 CN 101599270 A, 2009.12.09  
 WO 2013/110125 A1, 2013.08.01  
 TW 405109 B, 2000.09.11  
 CN 1421846 A, 2003.06.04  
 CN 101136199 A, 2008.03.05  
 WO 01/20597 A1, 2001.03.22  
 CN 1746970 A, 2006.03.15  
 US 2011/0066433 A1, 2011.03.17 (续)  
 审查员 宋玉香

(51) Int. Cl.

G10L 15/30 (2013.01)

权利要求书3页 说明书18页 附图12页

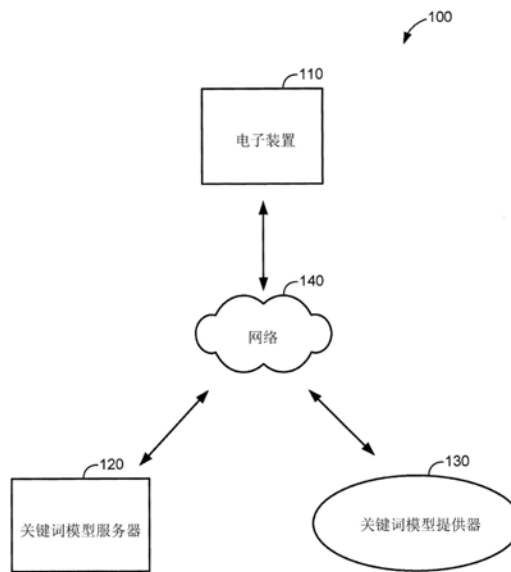
(54) 发明名称

使用用户指定关键词的说话者不相关关键词模型的关键词检测

(57) 摘要

本发明揭示一种由电子装置执行的用于获得由用户指定的关键词的说话者不相关关键词模型的方法。所述方法可包含从所述用户接收指示所述关键词的至少一个样本声音。所述方法还可基于所述至少一个样本声音而产生所述关键词的说话者相依关键词模型、响应于产生所述说话者相依关键词模型而向服务器发送对于所述关键词的所述说话者不相关关键词模型请求，和从所述服务器接收适于检测由多个用户说出的所述关键词的所述说话者不相关关键词模型。

CN 106663430 B



[接上页]

(56) 对比文件

US 6487530 B1, 2002.11.26

Tetsuo Kosaka et al..SPEAKER-  
INDEPENDENT PHONE MODELING BASED ON

SPEAKER-DEPENDENT HMMs" COMPOSITION AND  
CLUSTERING.《1995 INTERNATIONAL CONFERENCE  
ON ACOUSTICS, SPEECH, AND SIGNAL  
PROCESSING》.2002,

1. 一种由电子装置执行的用于获得与由用户指定的关键词相关联的说话者不相依关键词模型的方法,所述方法包括:

从所述用户接收指示所述关键词的至少一个样本声音;

基于所述至少一个样本声音而产生与所述关键词相关联的说话者相依关键词模型;

在产生与所述关键词相关联的说话者相依关键词模型后,向服务器发送对获取与所述关键词相关联的所述说话者不相依关键词模型的请求;

从所述服务器接收所述说话者不相依关键词模型;以及

基于关键词的类型或与所述关键词相关联的功能来选择说话者相依关键词模型或说话者不相依关键词模型中之一来检测所述关键词。

2. 根据权利要求1所述的方法,其进一步包括:

确定所述关键词是否包括私用关键词;以及

在确定所述关键词包括所述私用关键词之后基于所述说话者相依关键词模型而从输入声音检测所述关键词。

3. 根据权利要求1所述的方法,其进一步包括:

确定所述关键词是否包括非私用关键词;以及

在确定所述关键词包括所述非私用关键词之后基于所述说话者不相依关键词模型而从输入声音检测所述关键词。

4. 根据权利要求1所述的方法,其进一步包括:

确定与所述关键词相关联的功能的安全等级;以及

基于以下各者中的至少一者而从输入声音检测所述关键词:

在确定所述功能的所述安全等级低于预定安全等级之后基于所述说话者不相依关键词模型而从输入声音检测所述关键词;或

在确定所述功能的所述安全等级大于或等于所述预定安全等级之后基于所述说话者相依关键词模型而从输入声音检测所述关键词。

5. 根据权利要求1所述的方法,其进一步包括:

接收输入声音;

基于所述说话者相依关键词模型而从所述输入声音确定所述关键词的置信度评分;

在确定所述置信度评分大于或等于第一阈值之后从所述输入声音检测所述关键词;以及

在确定所述置信度评分大于或等于第二阈值且小于所述第一阈值之后更新指示被遗漏检测数目的计数。

6. 根据权利要求5所述的方法,其进一步包括:

在确定所述计数大于或等于预定数值后,基于所述说话者不相依关键词模型从后续输入声音检测所述关键词。

7. 根据权利要求1所述的方法,其进一步包括:

接收输入声音;

基于所述说话者相依关键词模型而确定所述输入声音的关键词评分;

基于所述说话者不相依关键词模型而确定所述输入声音的非关键词评分;以及

基于所述输入声音的所述关键词评分和所述非关键词评分而将所述输入声音验证为

指示所述用户。

8. 根据权利要求1所述的方法,其中所述请求包含由所述用户指定的所述关键词、所述至少一个样本声音、从所述至少一个样本声音提取的至少一个声音特征和所述说话者相依关键词模型中的至少一者。

9. 根据权利要求1所述的方法,其中,接收所述说话者不相依关键词模型是在包括无线通信设备的设备中执行的。

10. 根据权利要求1所述的方法,其中,接收所述说话者不相依关键词模型是在包括基站执行的系统中执行的。

11. 根据权利要求1所述的方法,其进一步包括将所述说话者相依关键词模型存储在数据库中。

12. 根据权利要求1所述的方法,其中,所述服务器是远程服务器。

13. 一种电子装置,其包括:

声音传感器,其经配置以接收指示由用户指定的关键词的至少一个样本声音;

关键词模型产生器,其经配置以基于所述至少一个样本声音而产生与所述关键词相关联的说话者相依关键词模型;以及

关键词模型管理单元,其经配置以:

在产生与所述关键词相关联的所述说话者相依关键词模型后,向服务器发送获取与所述关键词相关联的说话者不相依关键词模型的请求;

接收所述说话者不相依关键词模型;以及

基于关键词的类型和与所述关键词相关联的功能选择所述说话者相依关键词模型和所述说话者不相依关键词模型中之一来检测所述关键词。

14. 根据权利要求13所述的电子装置,其中所述关键词模型管理单元经进一步配置以确定所述关键词是否包括私用关键词,且其中所述电子装置进一步包括关键词检测单元,所述关键词检测单元经配置以在确定所述关键词包括所述私用关键词之后基于所述说话者相依关键词模型而从输入声音检测所述关键词。

15. 根据权利要求13所述的电子装置,其中所述关键词模型管理单元经进一步配置以确定所述关键词是否包括非私用关键词,且其中所述电子装置进一步包括关键词检测单元,所述关键词检测单元经配置以在确定所述关键词包括所述非私用关键词之后基于所述说话者不相依关键词模型而从输入声音检测所述关键词。

16. 根据权利要求13所述的电子装置,其中所述关键词模型管理单元经进一步配置以确定与所述关键词相关联的功能的安全等级;且其中所述电子装置进一步包括经配置以基于以下各者中的至少一者而从输入声音检测所述关键词的关键词检测单元:

在确定所述功能的所述安全等级低于预定安全等级之后基于所述说话者不相依关键词模型而从输入声音检测所述关键词;或

在确定所述功能的所述安全等级大于或等于所述预定安全等级之后基于所述说话者相依关键词模型而从输入声音检测所述关键词。

17. 根据权利要求13所述的电子装置,其中所述关键词模型管理单元经进一步配置以接收输入声音,且其中所述电子装置进一步包括:关键词检测单元,其经配置以:

基于所述说话者相依关键词模型而从所述输入声音确定所述关键词的置信度评分;

在确定所述置信度评分大于或等于第一阈值之后从所述输入声音检测所述关键词；以及

在确定所述置信度评分大于或等于第二阈值且小于所述第一阈值之后更新指示被遗漏检测数目的计数。

18. 根据权利要求17所述的电子装置,其中所述关键词检测单元经配置以在确定所述计数大于或等于预定数值之后基于所述说话者不相依关键词模型而从后续输入声音检测所述关键词。

19. 根据权利要求13所述的电子装置,其中所述声音传感器经进一步配置以接收输入声音,且其中所述电子装置进一步包括:关键词检测单元,其经配置以:

基于所述说话者相依关键词模型而确定所述输入声音的关键词评分;

基于所述说话者不相依关键词模型而确定所述输入声音的非关键词评分;以及

基于所述输入声音的所述关键词评分和所述非关键词评分而将所述输入声音验证为指示所述用户。

20. 根据权利要求13所述的电子装置,还包括:

天线;

接收器,耦合到所述天线并配置成接收经译码音频信号;以及

处理器,耦合到所述接收器,所述处理器包括所述关键词模型管理单元。

21. 根据权利要求20所述的电子装置,其中,所述接收器和所述处理器集成到移动通信设备中。

22. 根据权利要求20所述的电子装置,其中,所述接收器和所述处理器集成到包括基站系统中。

23. 根据权利要求13所述的电子装置,其中,所述请求包括所述用户指定的关键词、至少一个样本声音、从所述至少一个样本声音中提取的至少一个声音特征或所述关键词的说话者相依关键词模型。

24. 根据权利要求13所述的电子装置,其中,所述关键词管理单元进一步配置成在数据库中存储所述说话者相依关键词模型。

25. 根据权利要求13所述的电子装置,其中,所述服务器是远程服务器。

## 使用用户指定关键词的说话者不相依关键词模型的关键词检测

[0001] 优先权要求

[0002] 本申请案主张2014年9月8日申请的名称为“使用用户指定关键词的说话者不相依关键词模型的关键词检测 (KEYWORD DETECTION USING SPEAKER-INDEPENDENT KEYWORD MODELS FOR USER-DESIGNATED KEYWORDS)”的第14/479,708号美国申请案的优先权,所述美国申请案的内容以全文引用的方式并入本文中。

### 技术领域

[0003] 本发明大体上涉及电子装置中的语音辨识,且更确切地说,涉及基于说话者相依关键词模型和说话者不相依关键词模型中的至少一者而检测关键词。

### 背景技术

[0004] 最近,例如智能电话、平板计算机和智能TV和类似者的电子装置的使用已变得普遍。这些装置常常通过无线或有线网络提供话音和/或数据通信功能性。此外,这些装置可包含用于提供经设计以增强用户便利性(例如声音处理、图像或视频处理、导航、音乐或多媒体文件的再现等)的多种功能的其它特征。

[0005] 在这些特征当中,习知电子装置常常装备有语音辨识功能。此功能允许电子装置在辨识来自用户的话音命令时执行各种功能。举例来说,电子装置可响应于来自用户的话音命令而激活应用、播放音频文件或获取图片。

[0006] 一些电子装置可由多个用户使用。举例来说,智能TV或平板计算机可由多个家庭成员使用。可响应于用户的话音命令而操作这些电子装置。这些电子装置常常装备有可用于检测来自多个用户的关键词的有限数目个预定义关键词。由于这些关键词的有限数目,电子装置还可允许定义界定话音命令的关键词以执行电子装置的特定功能。然而,使用由用户定义的此关键词可能由于用户的话音特性或发音的差别而不合适于辨识另一用户的话音命令中的关键词。

### 发明内容

[0007] 本发明提供用于获得由用户指定的关键词的说话者不相依关键词模型的方法和设备。

[0008] 根据本发明的一个方面,揭示一种由电子装置执行的用于获得由用户指定的关键词的说话者不相依关键词模型的方法。所述方法可包含从所述用户接收指示所述关键词的至少一个样本声音。所述方法还可基于所述至少一个样本声音而产生所述关键词的说话者相依关键词模型、响应于产生所述说话者相依关键词模型而向服务器发送对于所述关键词的所述说话者不相依关键词模型请求、且接收适于检测由多个用户说出的所述关键词的所述说话者不相依关键词模型。本发明还描述一种与此方法相关的装置。

[0009] 根据本发明的另一方面,一种电子装置可包含声音传感器和关键词模型产生器。

所述声音传感器可经配置以接收指示由用户指定的关键词的至少一个样本声音。此外,所述关键词模型产生器可经配置以基于所述至少一个样本声音而产生所述关键词的说话者相依关键词模型。此外,所述电子装置可经配置以响应于产生所述说话者相依关键词模型而向服务器发送对于所述关键词的说话者不相依关键词模型的请求和从所述服务器接收适于检测由多个用户说出的关键词的所述说话者不相依关键词模型。

[0010] 根据本发明的又一方面,揭示一种由服务器执行的用于识别与由用户指定的关键词相关联的说话者不相依关键词模型的方法。所述方法可包含从电子装置接收提供与所述关键词相关联的所述说话者不相依关键词模型请求。所述方法还可确定与多个关键词相关联的多个说话者不相依关键词模型是否包含与所述关键词相关联的所述说话者不相依关键词模型。所述多个说话者不相依关键词模型可适于检测由多个用户说出的多个关键词。此外,所述方法可包含在确定所述多个说话者不相依关键词模型包含所述说话者不相依关键词模型之后即刻向所述电子装置提供所述说话者不相依关键词模型,和在确定所述多个说话者不相依关键词模型不包含所述说话者不相依关键词模型之后,即刻将所述关键词添加到未经识别关键词的列表。本发明还描述一种与此方法相关的装置。

[0011] 根据本发明的又一方面,一种服务器可包含通信单元和关键词匹配单元。所述通信单元可经配置以从电子装置接收提供与关键词相关联的说话者不相依关键词模型请求。此外,所述关键词匹配单元可经配置以确定与多个关键词相关联的多个说话者不相依关键词模型是否包含与所述关键词相关联的所述说话者不相依关键词模型,和在确定所述多个说话者不相依关键词模型不包含所述说话者不相依关键词模型之后,即刻将所述关键词添加到未经识别关键词的列表。所述多个说话者不相依关键词模型可适于检测由多个用户说出的多个关键词。此外,所述服务器可经配置以在确定所述多个说话者不相依关键词模型包含所述说话者不相依关键词模型之后即刻向所述电子装置提供所述说话者不相依关键词模型。

## 附图说明

[0012] 当结合附图理解时,将通过参考以下详细描述理解本发明的发明性方面的实施例。

[0013] 图1说明根据本发明的一个实施例的用于提供与关键词相关联的说话者不相依关键词模型的系统,关键词可由用户指定,所述模型供用于从电子装置的输入语音或声音检测关键词。

[0014] 图2说明根据本发明的一个实施例的经配置以识别与由用户指定的关键词相关联的说话者不相依关键词模型的关键词模型服务器的框图。

[0015] 图3说明根据本发明的一个实施例的经配置以获得由用户指定的关键词的说话者不相依关键词模型的电子装置的框图。

[0016] 图4说明根据本发明的一个实施例的包含关键词模型数据库、关键词到功能数据库、安全关键词数据库和安全等级数据库的存储单元的框图。

[0017] 图5说明根据本发明的一个实施例的在关键词模型服务器中执行的用于识别与由用户指定的关键词相关联的说话者不相依关键词模型的方法的实例的流程图。

[0018] 图6A说明根据本发明的一个实施例的在电子装置中执行的用于获得与由用户指

定的关键词相关联的说话者不相依关键词模型的方法的实例的流程图。

[0019] 图6B说明根据本发明的一个实施例的在电子装置中执行的用于使用说话者相依关键词模型和说话者不相依关键词模型中的至少一者来检测用户指定关键词的方法的实例的流程图。

[0020] 图7说明根据本发明的一个实施例的电子装置的展示用于将可由用户指定的关键词指定到功能的多个菜单的显示屏幕。

[0021] 图8A说明根据本发明的一个实施例的经配置以在从由产生关键词的说话者相依关键词模型的用户说出的话语检测相关联关键词之后即刻执行展示天气的功能的电子装置。

[0022] 图8B说明根据本发明的一个实施例的经配置以在从由不同用户说出的话语检测相关联关键词之后即刻执行展示天气的功能的电子装置。

[0023] 图9A说明根据本发明的一个实施例的经配置以在从由产生关键词的说话者相依关键词模型的用户说出的话语检测相关联关键词之后即刻执行展示联系人列表的功能的电子装置。

[0024] 图9B说明根据本发明的一个实施例的经配置在并不产生说话者相依关键词模型的用户说出相关联关键词时不执行展示联系人列表的功能的电子装置。

[0025] 图10说明根据本发明的一些实施例的可实施用于获得由用户指定的关键词的说话者不相依关键词模型的方法和装置的无线通信系统中的电子装置的框图。

[0026] 图11说明根据本发明的一些实施例的可实施用于识别与由用户指定的关键词相关联的说话者不相依关键词模型的方法和装置的服务器系统的框图。

## 具体实施方式

[0027] 现在将详细参考各种实施例,在附图中图示所述实施例的实例。在以下详细描述中,陈述众多具体细节以便提供对本发明的透彻理解。然而,对于所属领域的一般技术人员将是显而易见的是,可在没有这些具体细节的情况下实践本发明。在其它情况下,未详细描述众所周知的方法、程序、系统和组件,以免不必要地混淆各种实施例的各方面。

[0028] 图1说明根据本发明的一个实施例的用于提供与关键词相关联的说话者不相依关键词模型的系统100,关键词可由用户指定,所述模型供用于从电子装置110的输入语音或声音检测关键词。系统100可包含电子装置110、关键词模型服务器120和关键词模型提供者130,所述各者可经配置以经由通信网络140通信。电子装置110可为装备有声音俘获和处理能力的任何合适装置,例如蜂窝式电话、智能电话、膝上型计算机、可携佩式计算机、平板装置、游戏装置、多媒体播放器、智能TV等。虽然图1说明电子装置110,但是系统100可包含经由通信网络140与关键词模型服务器120通信的任何合适数目个这些电子装置。

[0029] 如本文所使用,术语“关键词”可指代指示可用以执行或激活电子装置110中的功能或应用的一或多个词语或声音的任何数字或模拟信号或数据。如果从电子装置110的输入语音或声音检测到这个关键词,那么可在电子装置110中执行或激活与关键词相关联的功能(例如,特征、应用等)。举例来说,当从输入语音检测到关键词“你好助理”时,电子装置110可激活与关键词相关联的话音助理应用。

[0030] 电子装置110可经配置以通过使用关键词模型来从输入语音或声音检测关键词。



关键词模型可包含关于将由关键词模型检测到的关键词(例如,关键词的文本串或一或多个词语)、多个关键词模型参数和阈值的信息。在一个实施例中,关键词模型参数可指示相关联关键词的声音特性,且可包含关于与关键词的多个部分相关联的多个状态的状态信息,和用于确定指示输入声音包含关键词的概率的输入声音的置信度评分的概率模型。阈值可为用于确定输入声音包含相关联关键词的最小值。

[0031] 关键词模型可为说话者不相依关键词模型和说话者相依关键词模型。如本文所使用,“说话者不相依关键词模型”可与供用于检测由任何用户说出的关键词的关键词相关联,而无关于用户是谁,且可指代借由指示由特定用户说出的关键词的大数目个(例如,数千个)样本声音训练的关键词模型。术语“样本声音”可指代指示关键词的一连串一或多个声音信号或声音数据或类似者。由于说话者不相依特性,说话者不相依关键词模型可用于高度准确地从不同说话者的话语检测关键词。

[0032] 另一方面,“说话者相依关键词模型”可与供用于检测由特定用户说出的关键词相关联的关键词,且可指代借由指示由特定用户说出的关键词的小数目个(例如,五个)样本声音训练的关键词模型。由于说话者相依特性,虽然说话者相依关键词模型可适用于从由特定用户说出的话语检测用户指定关键词,但是其可能不较适用于从除特定用户以外的其它用户的话语检测关键词。在一个实施例中,说话者相依关键词模型可包含用于从输入声音检测关键词的第一阈值和用于检测由除特定用户以外的用户说出的关键词的第二阈值。在这个情况下第二阈值可小于第一阈值。

[0033] 关键词模型提供器130可产生关键词的基于指示由大量用户说出的关键词和/或非关键词的大量样本声音所训练的说话者不相依关键词模型。可针对检测来自不同用户的相关联关键词而优化以这个方式训练的说话者不相依关键词模型。关键词模型提供器130可为能够提供各种关键词的多种说话者不相依关键词模型的专用关键词模型提供器。可经由通信网络140将一或多个所产生说话者不相依关键词模型上载到关键词模型服务器120。

[0034] 关键词模型服务器120可将包含从关键词模型提供器130接收的一或多个说话者不相依关键词模型的多个关键词的多个说话者不相依关键词模型作为数据库存储。在一些实施例中,关键词模型服务器120可经配置以托管在线商店,其用于经由通信网络140将说话者不相依关键词模型提供到包含电子装置110的不同类型的电子装置。在线商店可为用于提供说话者不相依关键词模型的专用商店,或可实施为通用应用商店的一部分。

[0035] 在一些实施例中,电子装置110的用户可通过说出关键词若干次(例如,说出“你好助理”5次)作为样本声音来产生由用户指定的关键词的说话者相依关键词模型。如本文所使用,短语“由用户指定的关键词”与“用户指定关键词”互换地使用,且和可指代可由用户指定、指定、选择、识别或定义为关键词的一或多个词语或术语。在产生说话者相依关键词模型之后,电子装置110可即刻经由通信网络140向关键词模型服务器120发送对于与用户指定关键词相关联的说话者不相依关键词模型的请求。在一个实施例中,请求可包含用户指定关键词、关键词的一或多个样本声音、从一或多个样本声音提取的一或多个声音特征、和由电子装置110产生的用户指定关键词的说话者相依关键词模型中的至少一者。在用户指定关键词的情况下,请求中的关键词可为文本串或一或多个词语。

[0036] 在从电子装置110接收请求之后,关键词模型服务器120可即刻搜索多个说话者不相依关键词模型的数据库以发现与用户指定关键词相关联的说话者不相依关键词模型。如

果发现说话者不相依关键词模型,那么关键词模型服务器120可经由通信网络140向电子装置110提供说话者不相依关键词模型。一旦接收说话者不相依关键词模型,那么电子装置110可使用说话者不相依关键词模型来从由任何用户说出的话语检测用户指定关键词。通过使用说话者不相依关键词模型而非说话者相依关键词模型,电子装置110可高度准确地检测由不同用户说出的用户指定关键词。在一个实施例中,电子装置110可使用与用户指定关键词相关联的说话者相依和说话者不相依关键词模型两者,以将说话者验证为提供样本声音用于产生说话者相依关键词模型的用户。

[0037] 另一方面,如果关键词模型服务器120未发现与用户指定关键词相关联的说话者不相依关键词模型,那么关键词模型服务器120可将可包含于请求中的用户指定关键词添加到未经识别关键词的列表(例如,未经识别关键词的数据库),所述列表可包含一或多个未经识别关键词。或者或另外,关键词模型服务器120可在未经识别关键词的列表中存储关键词的至少一个样本声音、从至少一个样本声音提取的至少一个声音特征、和可包含于请求中的说话者相依关键词模型。此外,关键词模型服务器120可向电子装置110消息没有没有与关键词相关联的说话者不相依关键词模型可用的消息。作为响应,电子装置110可使用说话者相依关键词模型用于从由用户说出的话语检测用户指定关键词。

[0038] 关键词模型服务器120可在从关键词模型提供器130接收新说话者不相依关键词模型时更新多个说话者不相依关键词模型的数据库。在更新数据库之后,关键词模型服务器120可即刻确定新说话者不相依关键词模型是否与未经识别关键词的列表中的未经识别关键词相关联。如果新说话者不相依关键词模型与未经识别关键词中的一者相关联,那么关键词模型服务器120可确定请求未经识别关键词的电子装置(例如,电子装置110)且将新说话者不相依关键词模型提供到所确定电子装置。

[0039] 如上文所描述,电子装置110可产生用户指定关键词的用户相依关键词模型,且随后向关键词模型服务器120发送对于与关键词相关联的用户独立关键词模型的请求。或者,电子装置110可首先向关键词模型服务器120发送对于与用户指定关键词相关联的用户独立关键词模型的请求。如果并不从关键词模型服务器120接收用户独立关键词模型,那么可随后产生用户相依关键词模型。另一方面,如果从关键词模型服务器120接收用户独立关键词模型,那么电子装置110可使用用户独立关键词模型以从输入声音检测用户指定关键词。

[0040] 图2说明根据本发明的一个实施例的经配置以识别与由用户指定的关键词相关联的说话者不相依关键词模型的关键词模型服务器120的框图。关键词模型服务器120可包含控制单元210、存储单元220、关键词匹配单元230和通信单元240。控制单元210可控制且管理关键词模型服务器120的总体操作。如所展示,存储单元220可包含关键词模型数据库222、安全关键词数据库224和未经识别关键词数据库226。关键词模型服务器120可通过通信网络140经由通信单元240与电子装置110和关键词模型提供器130通信。

[0041] 存储单元220中的关键词模型数据库222可包含多个关键词的多个说话者不相依关键词模型。存储于关键词模型数据库222中的多个说话者不相依关键词模型可适于检测由多个用户说出的多个关键词。另外,安全关键词数据库224可包含指示高安全等级的多个私用关键词(例如,“银行”、“联系人”、“消息”、“日历”、“电子邮件”等)的列表和指示低安全等级的多个非私用关键词(例如,“因特网”、“TV”、“无线电”、“天气”、“音乐”、“时间”等)的列表。举例来说,私用关键词可指示与关键词相关联的功能可由指定用户接入。另一方面,

非私用关键词可指示与关键词相关联的功能可由任何用户接入。未经识别关键词数据库226可包含未经识别关键词的列表,可在未在关键词模型数据库222中发现用户指定关键词的说话者不相依关键词模型时更新所述列表。

[0042] 在关键词模型服务器120中,通信单元240可经由通信网络140从电子装置110接收对于与用户指定关键词相关联的说话者不相依关键词模型的请求。在一个实施例中,请求可包含用户指定关键词、用户指定关键词的一或多个样本声音、从一或多个样本声音提取的一或多个声音特征、和由电子装置110产生的用户指定关键词的说话者相依关键词模型中的至少一者。如果用户指定关键词包含于请求中,那么用户指定关键词可为文本串或一或多个词语。通信单元240可在从电子装置110接收请求之后即刻向关键词匹配单元230提供请求。

[0043] 在接收请求之后,关键词匹配单元230可即刻接入关键词模型数据库222以发现与用户指定关键词相关联的说话者不相依关键词模型。关键词匹配单元230可随后确定关键词模型数据库222是否包含与用户指定关键词相关联的说话者不相依关键词模型。如果确定关键词模型数据库222包含说话者不相依关键词模型,那么关键词模型服务器120可经由通信单元240将说话者不相依关键词模型提供到电子装置110。

[0044] 在一个实施例中,请求可包含用户指定关键词,可比较用户指定关键词与关键词模型数据库222中的多个说话者不相依关键词模型相关联的多个关键词。基于比较,关键词匹配单元230可在与关键词模型数据库222中的多个说话者不相依关键词模型相关联的多个关键词当中识别匹配请求中的用户指定关键词的关键词。可经由通信单元240随后将与所识别关键词相关联的说话者不相依关键词模型提供到电子装置110。

[0045] 在另一实施例中,请求可包含用户指定关键词的一或多个样本声音。在这个情况下,关键词匹配单元230可通过使用说话者不相依关键词模型中的至少一者以检测或辨识样本声音中的用户指定关键词来从关键词模型数据库222中的多个说话者不相依关键词模型当中识别说话者不相依关键词模型。举例来说,关键词匹配单元230可计算关键词模型数据库222中的每一说话者不相依关键词模型的样本声音的一或多个置信度评分,且将具有最高置信度评分的说话者不相依关键词模型选择为与用户指定关键词相关联的说话者不相依关键词模型。或者,请求可包含从一或多个样本声音提取的一或多个声音特征。在这个情况下,关键词匹配单元230可通过使用关键词模型数据库222中的说话者不相依关键词模型中的至少一者以检测或辨识一或多个声音特征中的用户指定关键词来识别说话者不相依关键词模型。

[0046] 在又一个实施例中,请求可包含由电子装置110产生的说话者相依关键词模型。在这个情况下,关键词匹配单元230可基于说话者相依关键词模型与多个说话者不相依关键词模型中的至少一者之间的相似度值而识别关键词模型数据库222中的说话者不相依关键词模型。举例来说,关键词匹配单元230可计算关键词模型数据库222中的说话者相依关键词模型中的多个关键词模型参数与说话者不相依关键词模型中的每一者中的多个关键词模型参数之间的相似度值。基于相似度值,关键词匹配单元230可将具有最高相似度值的说话者不相依关键词模型识别为与用户指定关键词相关联的说话者不相依关键词模型。

[0047] 在识别用户指定关键词的说话者不相依关键词模型之后,关键词模型服务器120可即刻经由通信单元240将说话者不相依关键词模型提供到电子装置110。在一个实施例

中,关键词模型服务器120可在识别到说话者不相依关键词模型之后即刻将说话者不相依关键词模型提供到电子装置110。在另一实施例中,关键词匹配单元230可确定用户指定关键词是否包含安全数据库224中的一或多个私用关键词(例如,“银行”、“联系人”、“消息”、“日历”、“电子邮件”等)。如果确定用户指定关键词包含一或多个私用关键词,那么关键词模型服务器120出于安全目的而不将说话者不相依关键词模型提供到电子装置110。举例来说,如果用户指定关键词为“展示我的日程”,那么关键词模型服务器120可确定用户指定关键词包含私用关键词“日程”且因此可向电子装置110发送出于安全目的而并未发送说话者不相依关键词模型的消息。

[0048] 或者或另外,关键词匹配单元230可确定用户指定关键词是否包含安全关键词数据库224中的一或多个非私用关键词(例如,“因特网”、“TV”、“无线电”、“天气”、“音乐”、“时间”等)。如果确定用户指定关键词包含一或多个非私用关键词,那么关键词模型服务器120可将说话者不相依关键词模型提供到电子装置110。举例来说,如果用户指定关键词为“开启TV”,那么关键词模型服务器120可确定用户指定关键词包含非私用关键词“TV”且因此可将说话者不相依关键词模型发送到电子装置110。在一些实施例中,可在确定用户指定关键词包含一或多个私用关键词和一或多个非私用关键词两者的情况中预先确定关键词模型服务器120是否可将说话者不相依关键词模型发送到电子装置110。

[0049] 另一方面,如果关键词匹配单元230不识别或发现与关键词模型数据库222中的用户指定关键词相关联的说话者不相依关键词模型,那么关键词匹配单元230可通过将可包括于请求中的用户指定关键词添加到未经识别关键词的列表来更新未经识别关键词数据库226。或者或另外,关键词匹配单元230可以用用户指定关键词的一或多个样本声音、从一或多个样本声音提取的一或多个声音特征和/或可包含于请求中的说话者相依关键词模型更新未经识别关键词数据库226。此外,关键词模型服务器120可向电子装置110发送关键词的说话者不相依关键词模型不可用的消息。

[0050] 关键词模型服务器120可以新说话者不相依关键词模型更新关键词模型数据库222,可经由通信单元240从关键词模型提供器130接收新说话者不相依关键词模型。在更新关键词模型数据库222之后,关键词匹配单元230可即刻确定新说话者不相依关键词模型是否与未经识别关键词数据库226的未经识别关键词列表中的未经识别关键词相关联。或者,关键词匹配单元230可确定新说话者不相依关键词模型是否与一或多个样本声音、一或多个声音特征和/或未经识别关键词数据库226中的说话者相依关键词模型相关联。如果确定新说话者不相依关键词模型与一或多个样本声音、一或多个声音特征和/或未经识别关键词数据库226中的说话者相依关键词模型中的任一者相关联,那么关键词模型服务器120可识别发送对于未经识别关键词的说话者不相依关键词模型的请求的电子装置。关键词模型服务器120可随后向经识别电子装置提供新说话者不相依关键词模型。

[0051] 图3说明根据本发明的一个实施例的经配置以获得由用户指定的关键词的说话者不相依关键词模型的电子装置110的框图。电子装置110可包含声音传感器310、处理器320、存储单元340、通信单元350和I/O单元360。电子装置110可装备有声音俘获和处理能力的任何合适的装置,例如蜂窝式电话、智能电话、膝上型计算机、可携佩式计算机、平板装置、游戏装置、多媒体播放器、智能TV等。在一些实施例中,电子装置110可由多个用户基于可包含由用户指定的一或多个关键词的语音命令而操作。举例来说,家庭成员可指定用于操作智

能电话或智能TV的关键词,且其它家庭成员也可使用所指定关键词以操作智能电话或智能TV。在这些情况下,说话者不相依关键词模型可更好的合适于从由与说话者相依关键词模型不同的用户说出的话语检测关键词。

[0052] 如所展示,处理器320可包含关键词模型产生器322、关键词设定单元324、关键词检测单元326、功能管理单元328和关键词模型管理单元330。处理器320可为用于管理且操作电子装置110的任何合适的处理器,例如应用处理器(AP)、中央处理单元(CPU)、数字信号处理器(DSP)、或类似者。处理器320中的关键词模型产生器322、关键词设定单元324、关键词检测单元326、功能管理单元328和关键词模型管理单元330可实施为一或多个软件模块或组件、硬件单元或组件和/或其任何组合。

[0053] 在一些实施例中,电子装置110的用户可经由声音传感器310指定语音命令的关键词。此外,用户可经由I/O单元360(例如键盘或触摸屏)输入关键词的文本(例如,关键词的文本串或一或多个词语)。举例来说,用户可经由声音传感器310输入关键词的一或多个样本声音(例如,说出关键词5次),且经由I/O单元360提供关键词的文本。声音传感器310可包含一或多个麦克风或可用以接收、俘获、感测、辨识和/或检测包含一或多个样本声音或语音命令的语音输入的任何其它类型的声音传感器。另外,声音传感器310可采用任何合适的软件或硬件以用于执行这些功能。

[0054] 关键词模型产生器322可从声音传感器310接收一或多个样本声音且基于样本声音而产生用户指定关键词的说话者相依关键词模型。在一些实施例中,关键词模型产生器322可执行对关键词的每一样本声音的子字辨识,且从子字辨识结果产生关键词的子字网络。基于子字网络,关键词模型产生器322可产生可包含多个关键词模型参数和阈值的关键词的说话者相依关键词模型。可将由关键词模型产生器322产生的说话者相依关键词模型添加到存储于存储单元340中的关键词模型数据库。

[0055] 在一个实施例中,关键词模型产生器322还可从I/O单元360接收关键词的文本。关键词的文本可随后与说话者相依关键词模型相关联。在这个情况下,关键词的文本还可作为说话者相依关键词模型的一部分存储。

[0056] 关键词设定单元324可经配置以将用户指定关键词与由用户经由I/O单元360选择的电子装置110中的功能(例如,特征、应用等)相关联。在将用户指定关键词与所选功能相关联之后,关键词设定单元324可即刻将用户指定关键词和相关联功能添加到存储单元340中的关键词到功能数据库。此外,可从数据库关键词模型撷取用户指定关键词的说话者相依关键词模型,且所述模型可存储于与关键词和功能相关联的关键词到功能数据库中。如果未在存储单元340中发现现有关键词到功能数据库,那么关键词设定单元324可产生将用户指定关键词映射到存储于存储单元340中的所选功能的关键词到功能数据库。在一些实施例中,关键词设定单元324可响应于对关键词的额外指定和对功能的选择或数据库中的现有项的删除而更新关键词到功能数据库。

[0057] 关键词检测单元326可经配置以通过接入包含用户指定关键词的说话者相依关键词模型的关键词到功能数据库中的关键词模型来从经由声音传感器310接收的输入声音检测用户指定关键词。在这个过程中,可针对关键词到功能数据库中的关键词模型中的每一者而计算输入声音的置信度评分。举例来说,可通过确定输入声音的关键词评分和非关键词评分和从关键词评分减去非关键词评分来计算置信度评分。

[0058] 在一个实施例中,关键词检测单元326可选择具有最高置信度评分的关键词模型。如果置信度评分大于或等于关键词模型中的第一阈值,那么关键词检测单元326可确定已从输入声音检测到与关键词模型相关联的关键词。另外,如果置信度评分大于或等于第二阈值且小于关键词模型中的第一阈值,那么关键词检测单元326可确定尚未从输入声音检测到与关键词模型相关联的关键词。在这个情况下,关键词检测单元326可更新指示被遗漏检测数目的计数且将计数存储于关键词到功能数据库中。计数可用于确定是否将使用说话者不相依关键词模型而非说话者相依关键词模型,如下文将详细描述。另一方面,如果置信度评分小于第二阈值,那么可确定未检测到关键词且可不更新计数。

[0059] 当已从输入声音检测到用户指定关键词时,关键词检测单元326可产生指示已检测到用户指定关键词的检测信号。功能管理单元328可接收检测信号且接入存储单元340中的关键词到功能数据库以识别与所检测用户指定关键词相关联的功能。功能管理单元328可随后执行与所检测用户指定关键词相关联的所识别功能。

[0060] 根据一实施例,当说话者相依关键词模型已由关键词模型产生器322产生时,关键词模型管理单元330可产生对于与可用于检测来自多个用户的关键词的用户指定关键词相关联的说话者不相依关键词模型的请求。请求可包含用户指定关键词(即,用户指定关键词的文本)、用户指定关键词的一或多个样本声音、从一或多个样本声音提取的至少一个声音特征和由关键词模型产生器322产生的说话者相依关键词模型中的至少一者。如果用户指定关键词包含于请求中,那么用户指定关键词可为文本串或一或多个词语。关键词模型管理单元330可通过通信网络140经由通信单元350向关键词模型服务器120发送请求。如果从关键词模型服务器120接收适于检测由多个用户说出的用户指定关键词的说话者不相依关键词模型,那么关键词模型管理单元330可通过添加所接收说话者不相依关键词模型来更新存储单元340中的关键词模型数据库。

[0061] 由于说话者相依特性,虽然由关键词模型产生器322产生的说话者相依关键词模型可适用于从由用户说出的话语检测用户指定关键词,但是其可能不较适用于从其它用户的话语检测关键词。在这些情况下,可使用从关键词模型服务器120接收的说话者不相依关键词模型,而非由关键词模型产生器322产生的用于更准确地检测由不同用户说出的用户指定关键词的说话者相依关键词模型。在一个实施例中,在从关键词模型服务器120接收说话者不相依关键词模型之后,关键词模型管理单元330可即刻以用户指定关键词的说话者不相依关键词模型替换关键词到功能数据库中的用户指定关键词的说话者相依关键词模型。

[0062] 在另一实施例中,关键词模型管理单元330可确定用户指定关键词是否包含来自存储单元340中的安全关键词数据库的一或多个私用关键词(例如,“银行”、“联系人”、“消息”、“日历”、“电子邮件”等)。举例来说,私用关键词可指示与关键词相关联的功能可由指定用户接入。如果关键词模型管理单元330确定用户指定关键词包含一或多个私用关键词,那么关键词检测单元326可出于安全目的而继续通过使用关键词到功能数据库中的说话者相依关键词来从输入声音检测关键词。在这个情况下,可不以用户指定关键词的说话者不相依关键词模型替换关键词到功能数据库中的用户指定关键词的说话者相依关键词模型。

[0063] 或者或另外,关键词模型管理单元330可确定用户指定关键词是否包含来自安全关键词数据库的一或多个非私用关键词(例如,“因特网”、“TV”、“无线电”、“天气”、“音乐”、

“时间”等)。举例来说,非私用关键词可指示与关键词相关联的功能可由任何用户接入。如果确定用户指定关键词包含一或多个非私用关键词,那么关键词模型管理单元330可以用用户指定关键词的说话者不相依关键词模型替换关键词到功能数据库中的用户指定关键词的说话者相依关键词模型。关键词检测单元326可随后通过使用关键词到功能数据库中的说话者不相依关键词模型来从输入声音检测用户指定关键词。在一些实施例中,可在确定用户指定关键词包含一或多个私用关键词和一或多个非私用关键词两者的情况中预先确定说话者不相依关键词模型是否替换说话者相依关键词模型。

[0064] 根据一个实施例,关键词模型管理单元330可接入关键词到功能数据库以识别与用户指定关键词相关联的功能,且通过接入存储单元340中的安全等级数据库来确定相关联功能的安全等级。如果功能的安全等级确定为低安全等级,那么关键词模型管理单元330可以用用户指定关键词的说话者不相依关键词模型替换与关键词到功能数据库中的用户指定关键词的说话者相依关键词模型。关键词检测单元326可随后基于关键词到功能数据库中的说话者不相依关键词模型而从输入声音检测用户指定关键词。另一方面,如果功能的安全等级确定为高安全等级,那么关键词检测单元326可继续通过使用关键词到功能数据库中的说话者相依关键词模型来从输入声音检测用户指定关键词。

[0065] 在一个实施例中,关键词模型管理单元330可接入关键词到功能数据库且撷取指示用户指定关键词的被遗漏检测的数目的计数。如果确定计数大于或等于预定数目,那么关键词模型管理单元330可以用用户指定关键词的说话者不相依关键词模型替换关键词到功能数据库中的用户指定关键词的说话者相依关键词模型。在这个情况下,关键词检测单元326可使用关键词到功能数据库中的说话者不相依关键词模型以从输入声音检测用户指定关键词。根据一些实施例,关键词模型管理单元330可经配置以在满足上述条件(即,私用关键词、非私用关键词、安全等级和计数)中的一些或全部时以说话者不相依关键词模型替换关键词到功能数据库中的说话者相依关键词模型。

[0066] 或者或另外,关键词模型管理单元330可在接收说话者不相依关键词模型之后即刻将用户指定关键词的说话者不相依关键词模型添加到关键词到功能数据库。在这个情况下,关键词检测单元326可使用与用户指定关键词相关联的说话者相依和说话者不相依关键词模型两者,来将输入声音验证为指示提供用于产生说话者相依关键词模型的样本声音的用户。举例来说,关键词检测单元326可通过使用说话者相依关键词模型来确定输入声音的关键词评分,且通过使用说话者不相依关键词模型来确定输入声音的非关键词评分。基于关键词评分和非关键词评分,关键词检测单元326可将输入声音验证为指示用户。举例来说,如果关键词评分与非关键词评分之间的差大于或等于预定置信度值,那么关键词检测单元326可将输入声音验证为指示用户。

[0067] 如上文所描述,电子装置110可产生用户指定关键词的用户相依关键词模型且随后向关键词模型服务器120发送对于用户独立关键词模型的请求。或者,电子装置110可首先向关键词模型服务器120发送对于与用户指定关键词相关联的用户独立关键词模型的请求。如果并不从关键词模型服务器120接收用户独立关键词模型,那么用户相依关键词模型可随后由电子装置110产生。另一方面,如果从关键词模型服务器120接收用户独立关键词模型,那么电子装置110可使用用户独立关键词模型以从输入声音检测用户指定关键词。

[0068] 图4说明根据本发明的一个实施例的包含关键词模型数据库410、关键词到功能数

数据库420、安全关键词数据库430和安全等级数据库440的存储单元340的框图。存储单元340可使用例如RAM(随机接入存储器)、ROM(只读存储器)、EEPROM(电可擦除可编程只读存储器)、快闪存储器或SSD(固态驱动器)的任何合适存储装置或存储器装置予以实施。关键词模型数据库410可包含可由关键词模型产生器322产生的一或多个说话者相依关键词模型412。可从关键词模型服务器120接收的一或多个说话者不相依关键词模型414也可存储于关键词模型数据库410中。

[0069] 关键词到功能数据库420可包含关键词到功能表422,功能表422将一或多个关键词映射到一或多个功能。一或多个关键词的一或多个关键词模型424(例如,说话者相依关键词模型和说话者不相依关键词模型)还可存储于与关键词和功能相关联的关键词到功能数据库420中。关键词检测单元326可基于关键词模型424而从输入声音检测关键词,且功能管理单元328可通过接入关键词到功能表422来识别与所检测关键词相关联的功能。功能管理单元328可随后执行与所检测关键词相关联的所识别功能。

[0070] 关键词到功能数据库420还可包含包含每一关键词的计数的计数表426,所述计数可指示关键词的被遗漏检测的数目。如果与关键词相关联的说话者相依关键词模型的输入声音的置信度评分确定大于或等于第二阈值且小于说话者相依关键词模型的第一阈值,那么关键词检测单元326可更新(即,增大)关键词的计数。基于计数表426,关键词模型管理单元330可确定是否将以关键词的说话者不相依关键词模型替换关键词到功能数据库420中的关键词的说话者相依关键词模型。

[0071] 如所展示,安全关键词数据库430可包含指示高安全等级的私用关键词432(例如,“银行”、“联系人”、“消息”、“日历”、“电子邮件”等)的列表和指示低安全等级的非私用关键词434(例如,“因特网”、“TV”、“无线电”、“天气”、“音乐”、“时间”等)的列表。举例来说,私用关键词可指示与关键词相关联的功能可由指定用户接入。另一方面,非私用关键词可指示与关键词相关联的功能可由任何用户接入。

[0072] 基于私用关键词432的列表和非私用关键词434的列表,关键词模型管理单元330可确定是否将以关键词的说话者不相依关键词模型替换关键词到功能数据库420中的关键词的说话者相依关键词模型。在一个实施例中,如果关键词包含一或多个私用关键词(例如,“银行”、“联系人”、“消息”、“日历”、“电子邮件”等),那么关键词模型管理单元330可确定不以说话者不相依关键词模型替换关键词到功能数据库420中的说话者相依关键词模型。在另一实施例中,如果关键词包含一或多个非私用关键词(例如,“因特网”、“TV”、“无线电”、“天气”、“音乐”、“时间”等),那么关键词模型管理单元330可确定以说话者不相依关键词模型替换关键词到功能数据库420中的说话者相依关键词模型。

[0073] 安全等级数据库440可包含功能到安全等级表442,安全等级表442将一或多个功能映射到一或多个安全等级。举例来说,每一功能可与高安全等级或低安全等级相关联。在这个情况下,高安全等级可指示功能可由指定用户接入。另一方面,低安全等级可指示功能可由任何用户接入。

[0074] 基于功能的安全等级,关键词模型管理单元330可确定是否将以说话者不相依关键词模型替换与功能相关联的说话者相依关键词模型。举例来说,如果关键词到功能表422中的功能的安全等级确定为低安全等级,那么关键词模型管理单元330可以说话者不相依关键词模型替换与功能相关联的关键词到功能数据库420中的说话者相依关键词模型。另



一方面,如果关键词到功能表422中的功能的安全等级确定为高安全等级,那么关键词检测单元326可继续通过使用与功能相关联的关键词到功能数据库420中的说话者相依关键词模型来从输入声音检测关键词。

[0075] 图5说明根据本发明的一个实施例的在关键词模型服务器120中执行的用于识别与由用户指定的关键词相关联的说话者不相依关键词模型的方法的实例的流程图。在510处,关键词模型服务器120可从电子装置110接收对与由用户指定的关键词相关联的说话者不相依关键词模型的请求。在一个实施例中,请求可包含用户指定关键词、用户指定关键词的一或多个样本声音、从一或多个样本声音提取的一或多个声音特征、和用户指定关键词的说话者相依关键词模型中的至少一者。如果用户指定关键词包含于请求中,那么用户指定关键词可为文本串或一或多个词语。

[0076] 在接收请求之后,关键词模型服务器120可即刻在520处接入其关键词模型数据库410以识别与用户指定关键词相关联的说话者不相依关键词模型。举例来说,关键词模型服务器120可确定关键词模型数据库410中的多个说话者不相依关键词模型是否包含与用户指定关键词相关联的说话者不相依关键词模型。在一些实施例中,关键词模型服务器120可基于请求中的用户指定关键词、至少一个样本声音、至少一个声音特征和说话者相依关键词模型中的至少一者而识别与用户指定关键词相关联的说话者不相依关键词模型。

[0077] 如果确定关键词模型数据库410中的多个说话者不相依关键词模型包含用户指定关键词的说话者不相依关键词模型(即,在520处为是),那么方法可前进到530且确定是否将用户指定关键词的说话者不相依关键词模型提供到电子装置110。在一个实施例中,关键词模型服务器120可确定用户指定关键词是否包含一或多个私用关键词。如果确定用户指定关键词包含一或多个私用关键词(即,在530处为否),那么关键词模型服务器120可在532处向电子装置110发送出于安全目的而并未发送说话者不相依关键词模型的消息。如果用户指定关键词不包含私用关键词(即,在530处为是),那么关键词模型服务器120可在534处将说话者不相依关键词模型提供到电子装置110。

[0078] 或者或另外,关键词模型服务器120可确定用户指定关键词是否包含一或多个非私用关键词。如果确定用户指定关键词包含一或多个非私用关键词(即,在530处为是),那么关键词模型服务器120可在534处将说话者不相依关键词模型提供到电子装置110。如果用户指定关键词不包含非私用关键词(即,在530处为否),那么关键词模型服务器120可在532处向电子装置110发送出于安全目的而不发送说话者不相依关键词模型的消息。

[0079] 另一方面,如果确定关键词模型数据库410中的多个说话者不相依关键词模型不包含与用户指定关键词相关联的说话者不相依关键词模型(即,在520处为否),那么关键词模型服务器120可在540处向电子装置110发送没有与关键词相关联的说话者不相依关键词模型可用的消息。关键词模型服务器120还可在550处以用户指定关键词更新未经识别关键词数据库226。或者或另外,关键词模型服务器120可以用用户指定关键词的一或多个样本声音、从一或多个样本声音提取的一或多个声音特征和/或可包含于请求中的说话者相依关键词模型更新未经识别关键词数据库226。

[0080] 在560处,关键词模型服务器120可确定可从关键词模型提供器接收的新说话者不相依关键词模型是否与未经识别关键词数据库226中的用户指定关键词相关联。如果新说话者不相依关键词模型并不与用户指定关键词相关联(即,在560处为否),那么方法前进返

回到560。另一方面,如果确定新说话者不相依关键词模型与用户指定关键词相关联(即,在560处为是),那么方法可前进到530以确定是否应将新说话者不相依关键词模型提供到电子装置110。

[0081] 图6A说明根据本发明的一个实施例的在电子装置110中执行的用于获得与由用户指定的关键词相关联的说话者不相依关键词模型的方法610的实例的流程图。最初,电子装置110可在612处从用户接收指示用户指定关键词的一或多个样本声音。基于一或多个样本声音,电子装置110可在614处产生用户指定关键词的说话者相依关键词模型。

[0082] 在产生说话者相依关键词模型之后,电子装置110可即刻在616处向关键词模型服务器120发送对于与用户指定关键词相关联的说话者不相依关键词模型的请求。在一个实施例中,请求可包含用户指定关键词、一或多个样本声音、从一或多个样本声音提取的一或多个声音特征和由电子装置110产生的说话者相依关键词模型中的至少一者。在用户指定关键词的情况下,请求中的关键词可为文本串或一或多个词语。

[0083] 在618处,电子装置110可确定是否从关键词模型服务器120接收与用户指定关键词相关联的说话者不相依关键词模型。如果并不从关键词模型服务器120接收说话者不相依关键词模型(例如,如果电子装置110接收出于安全目的而并未发送说话者不相依关键词模型的消息(即,在618处为否)),那么方法610可前进到620以通过使用说话者相依关键词模型来从输入声音检测用户指定关键词。另一方面,如果从关键词模型服务器120接收说话者不相依关键词模型(即,在618处为是),那么方法610可前进到图6B中的630以将所接收说话者不相依关键词模型存储于关键词模型数据库410中。

[0084] 图6B说明根据本发明的一个实施例的在电子装置110中执行的用于使用说话者相依关键词模型和说话者不相依关键词模型中的至少一者来检测用户指定关键词的方法630的实例的流程图。在640处,电子装置110可将从关键词模型服务器120接收的说话者不相依关键词模型存储于关键词模型数据库410中。在一个实施例中,方法630可前进到650以使用说话者相依和说话者不相依关键词模型两者已将输入声音中的话语的说话者验证为提供样本声音用于产生说话者相依关键词模型的用户。

[0085] 举例来说,电子装置110可通过使用说话者相依关键词模型来确定输入声音的关键词评分,且通过使用说话者不相依关键词模型来确定输入声音的非关键词评分。基于关键词评分和非关键词评分,电子装置110可将输入声音验证为指示用户。举例来说,如果关键词评分与非关键词评分之间的差大于或等于预定置信度值,那么关键词检测单元326可将输入声音验证为指示用户。

[0086] 在另一实施例中,方法可前进到660以确定指示高安全等级的私用关键词是否包含于用户指定关键词中。私用关键词可指示与关键词相关联的功能可由指定用户接入。如果确定用户指定关键词包含一或多个私用关键词(即,在660处为是),那么方法630可前进到670以使用说话者相依关键词模型来从输入声音检测用户指定关键词。另一方面,如果没有私用关键词包含于用户指定关键词中(即,在660处为否),那么方法630前进到662。

[0087] 在662处,电子装置110可确定指示低安全等级的非私用关键词是否包含于用户指定关键词中。非私用关键词可指示与关键词相关联的功能可由任何用户接入。如果确定用户指定关键词不包含非私用关键词(即,在662处为否),那么方法630可前进到670以使用说话者相依关键词模型来从输入声音检测用户指定关键词。另一方面,如果一或多个非私用

关键词包含于用户指定关键词中(即,在662处为是),那么方法630可前进到664。

[0088] 在664处,电子装置110可确定与用户指定关键词相关联的功能是否具有低于预定安全等级的安全等级。举例来说,如果确定相关联功能具有高安全等级(即,在664处为否),那么方法630可前进到670以使用说话者相依关键词模型以从输入声音检测用户指定关键词。另一方面,如果相关联功能具有低安全等级(即,在664处为是),那么方法630可前进到666。

[0089] 在666处,电子装置110可确定用户指定关键词的计数是否小于预定阈值。在这个情况下,用户指定关键词的计数可指示用户指定关键词的被遗漏检测的数目。在一个实施例中,如果说话者相依关键词模型的输入声音的置信度评分大于或等于第二阈值且小于说话者相依关键词模型中的第一阈值,那么电子装置110可确定已遗漏检测用户指定关键词。

[0090] 如果用户指定关键词的计数被确定为小于预定阈值(即,在666处为是),那么方法630可前进到670以使用说话者相依关键词模型以从输入声音检测用户指定关键词。另一方面,如果用户指定关键词的计数被确定为大于预定阈值(即,在666处为是),那么方法630可前进到680以使用说话者不相依关键词模型来检测用户指定关键词。虽然方法630说明为在按那个顺序满足660、662、664和666处的所有四个条件时前进到680,但是方法630也可在满足660、662、664和666处的条件中的至少一者或按任何合适的顺序满足660、662、664和666处的条件中的一些时前进到680。

[0091] 图7说明根据本发明的一个实施例的电子装置110的展示用于将可由用户指定的关键词指定到功能的菜单710、720和730的显示屏幕。最初,可在I/O单元360的显示屏幕上显示菜单710来从功能列表选择功能。如所说明,功能的列表可包含开启TV、展示天气、解锁电子装置110、展示电子装置110的联系人列表等。

[0092] 电子装置110的用户可菜单710选择功能。举例来说,用户可选择功能“展示天气”。作为响应,电子装置110可显示菜单720以允许用户选择用于执行所选功能(即,展示天气)的关键词。

[0093] 用于选择关键词的菜单720可显示器与电子装置110的关键词模型数据库410中的关键词模型相关联的列表。用户可从菜单720选择关键词以将所选关键词指定到功能“展示天气”。菜单720还可提供用于产生新关键词的菜单项,新关键词可由用户指定。举例来说,用户可通过选择菜单项“产生关键词”来激活菜单730。在这个情况下,用户可通过说出关键词5次来产生用户指定关键词(例如,“天气怎么样”)的说话者相依关键词模型。在一个实施例中,电子装置110还可请求用户输入关键词的文本。

[0094] 一旦产生用户指定关键词的说话者相依关键词模型,那么电子装置110可通过将用户指定关键词(例如,“天气怎么样”)和关键词的说话者相依关键词模型映射到所选功能“展示天气”来更新关键词到功能数据库420。在一个实施例中,电子装置110可接收用户指定关键词(例如,“天气怎么样”)的说话者不相依关键词模型且以所接收说话者不相依关键词模型替换说话者相依关键词模型。举例来说,电子装置110可能确定以说话者不相依关键词模型替换说话者相依关键词模型,这是因为用户指定关键词包含非私用关键词“天气”。在所说明实施例中,虽然在选择关键词之前选择功能,但是在一些实施例中可在选择功能之前选择关键词。

[0095] 图8A说明根据本发明的一个实施例的经配置以在从由产生关键词的说话者相依

关键词模型的用户说出的话语检测相关联关键词之后即刻执行展示天气的功能的电子装置110。在所说明实施例中,电子装置110可包含基于来自特定用户(即,用户A)的样本声音所产生的关键词“天气怎么样”的说话者相依关键词模型。电子装置110还可包含从关键词模型服务器120接收的关键词“天气怎么样”的说话者不相依关键词模型。

[0096] 在一些实施例中,电子装置110可确定使用关键词“天气怎么样”的说话者不相依关键词模型而非使用说话者相依关键词。举例来说,电子装置110可确定使用说话者不相依关键词模型,这是因为关键词“天气怎么样”包含非私用关键词“天气”。或者或另外,电子装置110可确定使用说话者不相依关键词模型,这是因为关键词“天气怎么样”不包含私用关键词。

[0097] 在另一实施例中,电子装置110可确定使用说话者不相依关键词模型,这是因为用于展示天气的功能具有低安全等级。在又一个实施例中,电子装置110可确定使用说话者不相依关键词模型,这是因为指示关键词“天气怎么样”的被遗漏检测的数目的计数大于或等于预定数目。在这些情况下,电子装置110可由于说话者不相依特性而高度准确地从不同说话者的话语检测关键词“天气怎么样”。

[0098] 在所说明实施例中,用户A可说出关键词“天气怎么样”以激活用于展示天气的功能。因为电子装置110经配置以基于说话者不相依关键词模型而从输入声音检测关键词“天气怎么样”,所以电子装置110可从用户A检测关键词。因此,子装置110可基于电子装置110的GPS(全球定位系统)信息通过通信网络140接入天气信息,且为用户A将所接收天气信息显示于显示屏幕上。

[0099] 图8B说明根据本发明的一个实施例的经配置以在从由不同用户说出的话语检测相关联关键词之后即刻执行展示天气的功能的电子装置110。在所说明的实施例中,电子装置110可为图8A中的相同电子装置,且可经配置以基于说话者不相依关键词模型而从输入声音检测关键词“天气怎么样”。由于说话者不相依关键词模型的说话者不相依特性,电子装置110可在用户B说出关键词“天气怎么样”时检测关键词。响应于检测到关键词,电子装置110可基于电子装置110的GPS(全球定位系统)信息通过通信网络140接入天气信息,且为用户B将所接收天气信息显示于显示屏幕上。

[0100] 图9A说明根据本发明的一个实施例的经配置以在从由产生关键词的说话者相依关键词模型的用户说出的话语检测相关联关键词之后即刻执行展示联系人列表的功能的电子装置110。在所说明实施例中,电子装置110可包含基于由特定用户(即,用户A)说出的话语所产生的关键词“展示联系人”的说话者相依关键词模型。电子装置110还可包含从关键词模型服务器120接收的关键词“展示联系人”的说话者不相依关键词模型。

[0101] 在一些实施例中,电子装置110可确定使用关键词“展示联系人”的说话者相依关键词模型而非使用说话者不相依关键词。举例来说,电子装置可确定使用说话者相依关键词模型,这是因为关键词“展示联系人”包含私用关键词“联系人”。或者或另外,电子装置110可确定使用说话者相依关键词模型,这是因为关键词“展示联系人”不包含非私用关键词。

[0102] 在另一实施例中,电子装置110可确定使用说话者相依关键词模型,这是因为用于展示电子装置110的联系人列表的功能具有高安全等级。在又一个实施例中,电子装置110可确定使用说话者相依关键词模型,这是因为指示关键词“展示联系人”的被遗漏检测

的数目的计数小于预定数目。在这些情况下,说话者相依关键词模型可从由用户A说出的话语检测关键词“展示联系人”,但由于说话者相依特性而可能不能够从其它用户的话语检测关键词。

[0103] 在所说明实施例中,用户A可说出关键词“展示联系人”以激活用于展示电子装置110的联系人列表的功能。因为电子装置110经配置以基于说话者相依关键词模型而从输入声音检测关键词,所以电子装置110可从来自用户A的语音命令检测关键词“展示联系人”。因此,电子装置110可为用户A在显示屏幕上显示电子装置110的联系人列表。

[0104] 在另一实施例中,电子装置110可使用关键词“展示联系人”的说话者相依和说话者不相依关键词模型两者以将输入声音验证为指示用户A。在这个情况下,电子装置110可基于说话者相依关键词模型而确定输入声音的关键词评分,且基于说话者不相依关键词模型而确定输入声音的非关键词评分。如果可通过从关键词评分减去非关键词评分所计算的置信度评分大于阈值,那么可将输入声音确定为指示用户A。通过使用关键词“展示联系人”的说话者相依和说话者不相依关键词模型两者,电子装置110可更准确地检测由用户A说出的关键词。

[0105] 图9B说明根据本发明的一个实施例的经配置在并不产生说话者相依关键词模型的用户说出相关联关键词时不执行展示联系人列表的功能的电子装置110。在所说明实施例中,电子装置110可为图9A中的相同电子装置,且可经配置以基于由用户A产生的说话者相依关键词模型而从输入声音检测关键词“展示联系人”。由于说话者相依关键词模型的说话者相依特性,电子装置110可从用户A的语音命令检测关键词,但可能不能够从用户B的语音命令检测关键词。因此,甚至当用户B说出关键词“展示联系人”时,电子装置110也可能无法检测关键词并显示尚未检测到的关键词的消息。

[0106] 图10为根据本发明的一些实施例的可实施用于获得由用户指定的关键词的说话者不相依关键词模型的方法和装置的无线通信系统中的示范性电子装置1000的框图。可根据参看图1到9所描述的上述实施例而在电子装置中实施电子装置1000的配置。电子装置1000可为蜂窝式电话、智能电话、平板计算机、膝上型计算机、终端机、手持机、个人数字助理(PDA)、无线调制解调器、无绳电话等。无线通信系统可为码分多址接入(CDMA)系统、移动通信(GSM)系统的广播系统、宽带CDMA(WCDMA)系统、长期演进(LTE)系统、高级LTE系统等。另外,电子装置1000可(例如)使用Wi-Fi直连或蓝牙直接与另一移动装置通信。

[0107] 电子装置1000能够经由接收路径和发射路径提供双向通信。在接收路径上,由基站发射的信号由天线1010接收且被提供到接收器(RCVR) 1014。接收器1014调节且数字化所接收信号且将样本(例如,经调节且经数字化的数字信号)提供到数字区段以供进一步处理。在发射路径上,发射器(TMTR) 1016从数字区段1020接收待发射的数据、处理并调节所述数据且产生经调制信号,所述经调制信号经由天线1010发射到基站。接收器1014和发射器1016可为可支持CDMA、GSM、LTE、LTE高级等的收发器的一部分。

[0108] 数字区段1020包含各种处理、接口和存储器单元,例如调制解调器处理器1022、精简指令集计算机/数字信号处理器(RISC/DSP) 1024、控制器/处理器1026、内部存储器1028、通用音频/视频编码器1032、通用音频解码器1034、图形/显示处理器1036和外部总线接口(EBI) 1038。调制解调器处理器1022可执行供数据发射和接收的处理,例如,编码、调制、解调和解码。RISC/DSP 1024可执行针对电子装置1000的一般和专业化处理。控制器/处理器

1026可执行数字区段1020内的各种处理和接口单元的操作。内部存储器1028可将用于各种单元的数据和/或指令存储在数字区段1020内。

[0109] 通用音频/视频编码器1032可执行对来自音频/视频源1042、麦克风1044、图像传感器1046等的输入信号的编码。通用音频解码器1034可执行对经译码音频数据的解码且可将输出信号提供到扬声器/耳机1048。图形/显示处理器1036可执行对可呈现给显示单元1050的图形、视频、图像和文本的处理。EBI 1038可促进数字区段1020与主存储器1052之间的数据传送。

[0110] 可借由一或多个处理器、DSP、微处理器、RISC等来实施数字区段1020。数字区段1020还可制造于一或多个特殊应用集成电路(ASIC)和/或一些其它类型的集成电路(IC)上。

[0111] 一般来说,本文中所描述的任何装置可表示各种类型的装置,例如无线电话、蜂窝式电话、膝上型计算机、无线多媒体装置、无线通信个人计算机(PC)卡、PDA、外部或内置调制解调器、通过无线通道通信的装置等。装置可具有各种名称,例如接入终端(AT)、接入单元、订户单元、移动台、移动装置、移动单元、移动电话、移动设备、远程站、远程终端、远程单元、用户装置、用户设备、手持式装置等。本文中所描述的任何装置可具有用于存储指令和数据的存储器以及硬件、软件、固件或其组合。

[0112] 图11说明根据本发明的一些实施例的可实施用于识别与由用户指定的关键词相关联的说话者不相依关键词模型的方法和装置的服务器系统1100的框图。服务器系统1100可包含一或多个处理单元(CPU)1102、一或多个网络或其它通信网络接口1110、存储器1112,和用于互连这些组件的一或多个通信总线1114。服务器系统1100还可包含具有显示装置和键盘的用户接口(未展示)。

[0113] 存储器1112可为任何合适的存储器,例如高速随机接入存储器, (例如DRAM、SRAM、DDR RAM或其它随机接入固态存储器装置)。存储器1112可包含或可替代地为非易失性存储器(例如,一或多个磁盘存储器装置、光盘存储装置、快闪存储器装置或其它非易失性固态存储装置)。在一些实施例中,存储器1112可包含从CPU 1102远程定位和/或远程定位于多处中的一或多个存储装置。

[0114] 由存储器1112表示的上述存储器装置中的任一者可存储对应于用于执行和/或实行先前所描述的进程、操作和方法中的任一者的一组指令的任何数目个模组或程序。举例来说,存储器1112可包含操作系统1116,其经配置以存储包含用于处置各种基本系统服务且用于执行硬件相关任务的程序的指令。存储器1112的网络通信模块1110可用于经由一或多个通信网络接口1110(有线或无线)和一或多个通信网络(例如因特网、其它广域网、局域网、城域网等)将服务器系统1100连接到其它计算机。

[0115] 存储器1112还可包含数据库1120,其经配置以存储对于服务器系统1100的操作必需的各种数据和/或信息。举例来说,数据库1120可存储多个关键词的多个说话者不相依关键词模型。在这个配置中,数据库1120还可存储多个私用关键词的列表和多个非私用关键词的列表。

[0116] 应了解,不必将上述所识别模组或程序(即,指令集)实施为单独软件程序、程序或模组,且因此可在各种实施例中组合或以其它方式再配置这些模组的各种子集。另外,存储器1112可存储上文所描述的额外模组和数据结构。

[0117] 图10和11更既定作为对客户端系统和服务器系统的各种特征的功能性描述而非本文中所述的实施例的结构示意图。在实践中,如由所属领域的技术人员认识到,单独展示物件的可组合且一些物件可分离。举例来说,在图11中单独地展示的一些物件可实施于单个服务器上,且单个物件可由一或多个服务器实施。此外,数据库1120可实施于服务器的不同集合上或服务器系统1100的其它组件中。用以实施服务器系统1100的服务器的实际数目和其中的特征分配可在实施方案之间不同。

[0118] 通过各种装置来实施本文中所描述的技术。举例来说,这些技术以硬件、固件、软件或其组合来予以实施。应进一步理解,结合本文中的揭示内容所描述的各种说明性逻辑块、模块、电路和算法步骤可实施为电子硬件、计算机软件,或两者的组合是可能的。为清楚地说明硬件与软件的此可互换性,上文已大体上关于其功能性描述了各种说明性组件、块、模块、电路和步骤。将此功能性实施为硬件还是软件取决于特定应用及施加于整个系统的设计约束。可针对每一特定应用以不同方式来实施所描述功能性,但这些实施方案决策不应被解释为会导致脱离本发明的范围。

[0119] 对于硬件实施方案,用以执行所述技术的处理单元可实施在以下各者内:一或多个ASIC、DSP、数字信号处理装置(DSPD)、可编程逻辑装置(PLD)、现场可编程门阵列(FPGA)、处理器、控制器、微控制器、微处理器、电子装置、经设计以执行本文中所描述的功能的其它电子单元、计算机,或其组合。

[0120] 如果以软件实施,那么先前所描述的方法、操作和程序可存储在计算机可读媒体上的一或多个指令或码上或作为一或多个指令或码通过计算机可读媒体传输。计算机可读媒体包含计算机存储媒体及通信媒体两者,通信媒体包含促进将计算机程序从一处传送到另一处的任何媒体。存储媒体可以是可由计算机接入的任何可用媒体。举例来说且非限制,这些非暂时性计算机可读媒体可包含RAM、ROM、EEPROM、CD-ROM或其它光盘存储装置、磁盘存储装置或其它磁性存储装置,或可用于携载或存储呈指令或数据结构形式的所要程序代码且可由计算机接入的任何其它媒体。并且,任何连接被恰当地称作计算机可读媒体。举例来说,如果使用同轴电缆、光纤电缆、双绞线、数字订户线(DSL)或例如红外线、无线电及微波等无线技术从网站、服务器或其它远程源发射软件,则同轴电缆、光纤电缆、双绞线、DSL或例如红外线、无线电及微波等无线技术包含于媒体的定义中。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、数字多功能光盘(DVD)、软性磁盘和蓝光光盘,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。以上各项的组合也应包含在计算机可读媒体的范围内。

[0121] 提供本发明的前述描述以使所属领域的技术人员能够制造或使用本发明。对本发明的各种修改对于所属领域的技术人员容易显而易见,且本文中界定的一般原理在不脱离本发明的精神或范围的情况下应用于其它变体。因此,本发明并不既定限于本文中所描述的实例,而应符合与本文中所揭示的原理及新颖特征相一致的最广泛范围。

[0122] 虽然已经以特定地针对结构特征和/或方法动作的语言来描述标的物,但应理解,所附权利要求书中所界定的标的物未必限于上文所描述的具体特征或动作。确切地说,上文中所描述的特定特征和动作被揭示为实施权利要求书的实例形式。

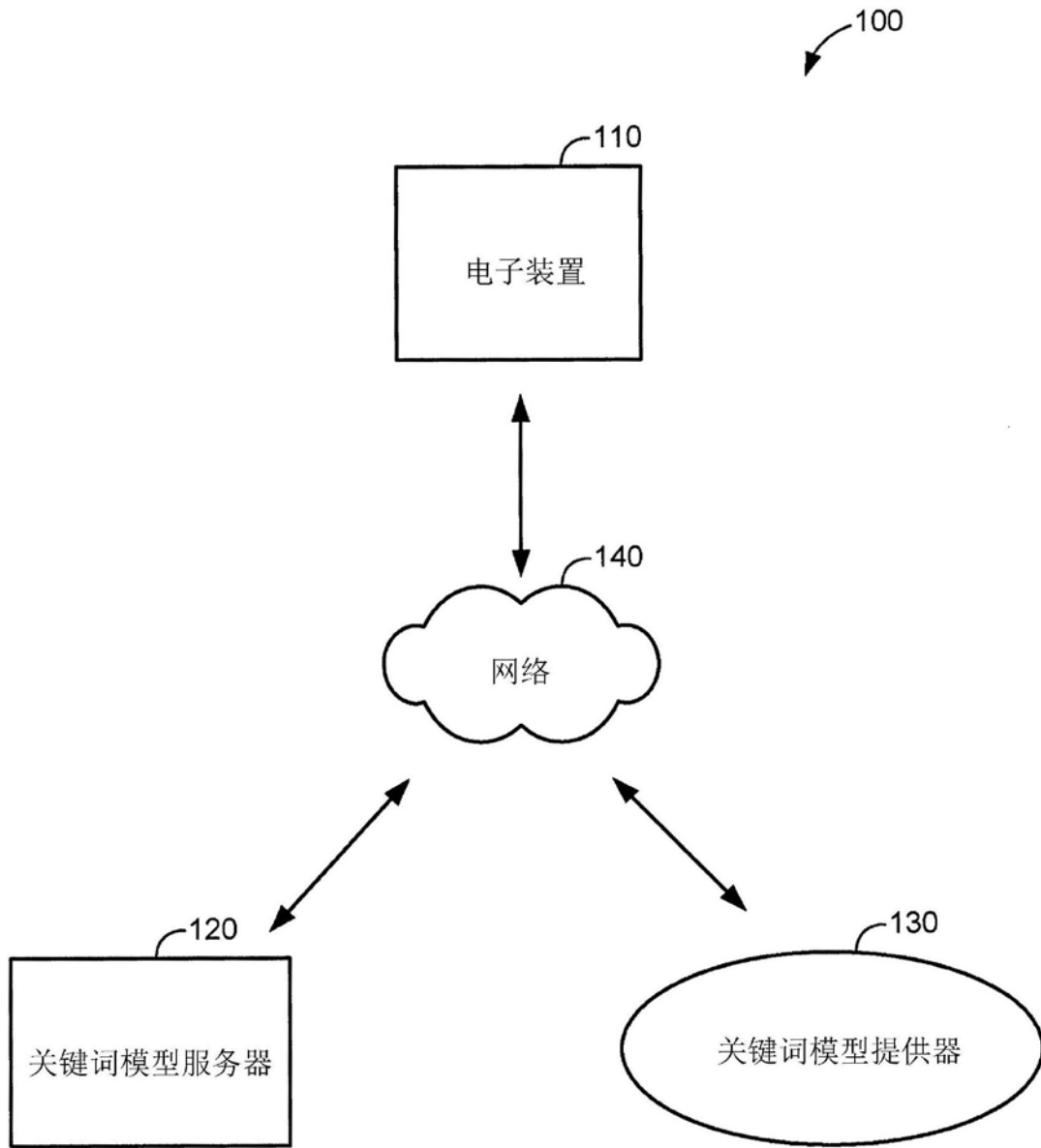


图1



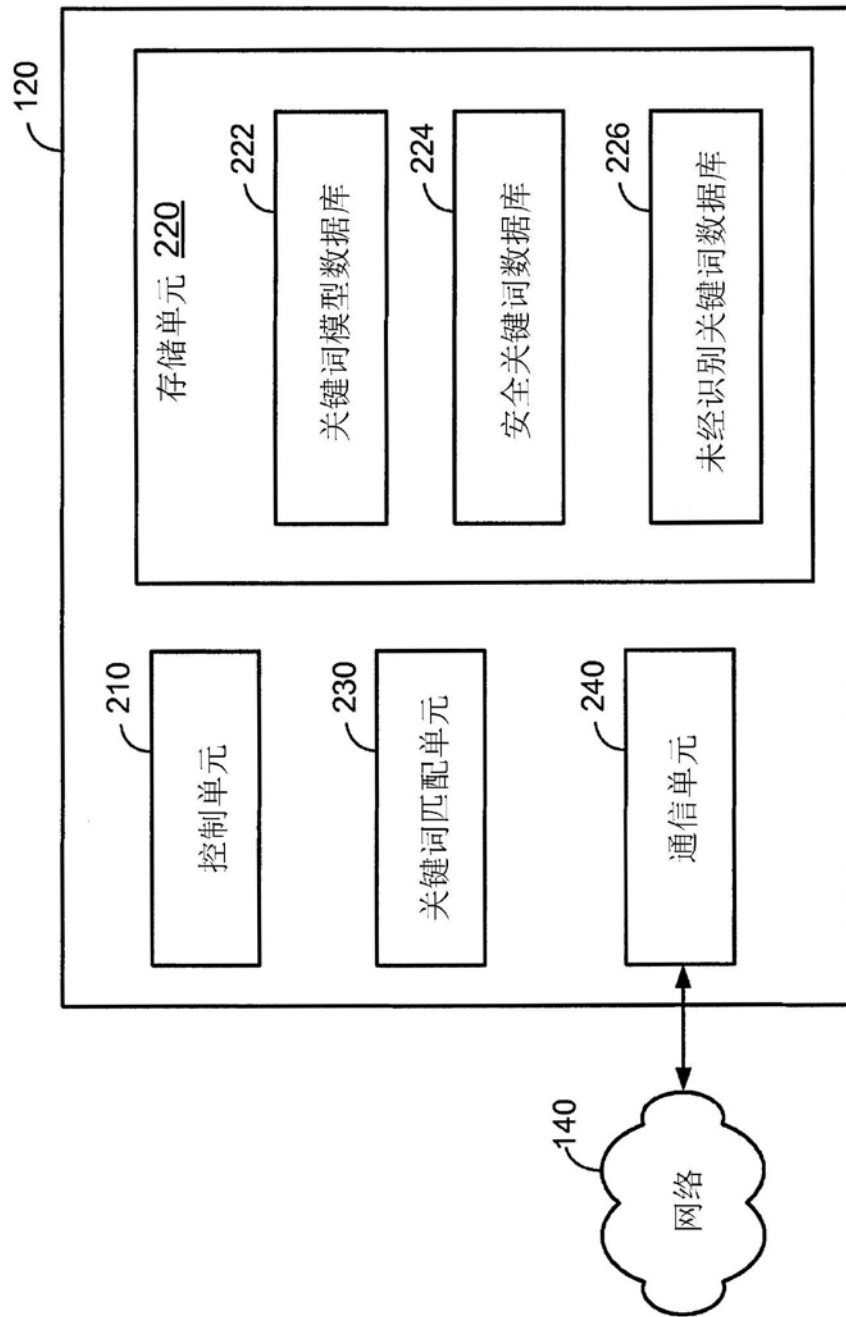


图2

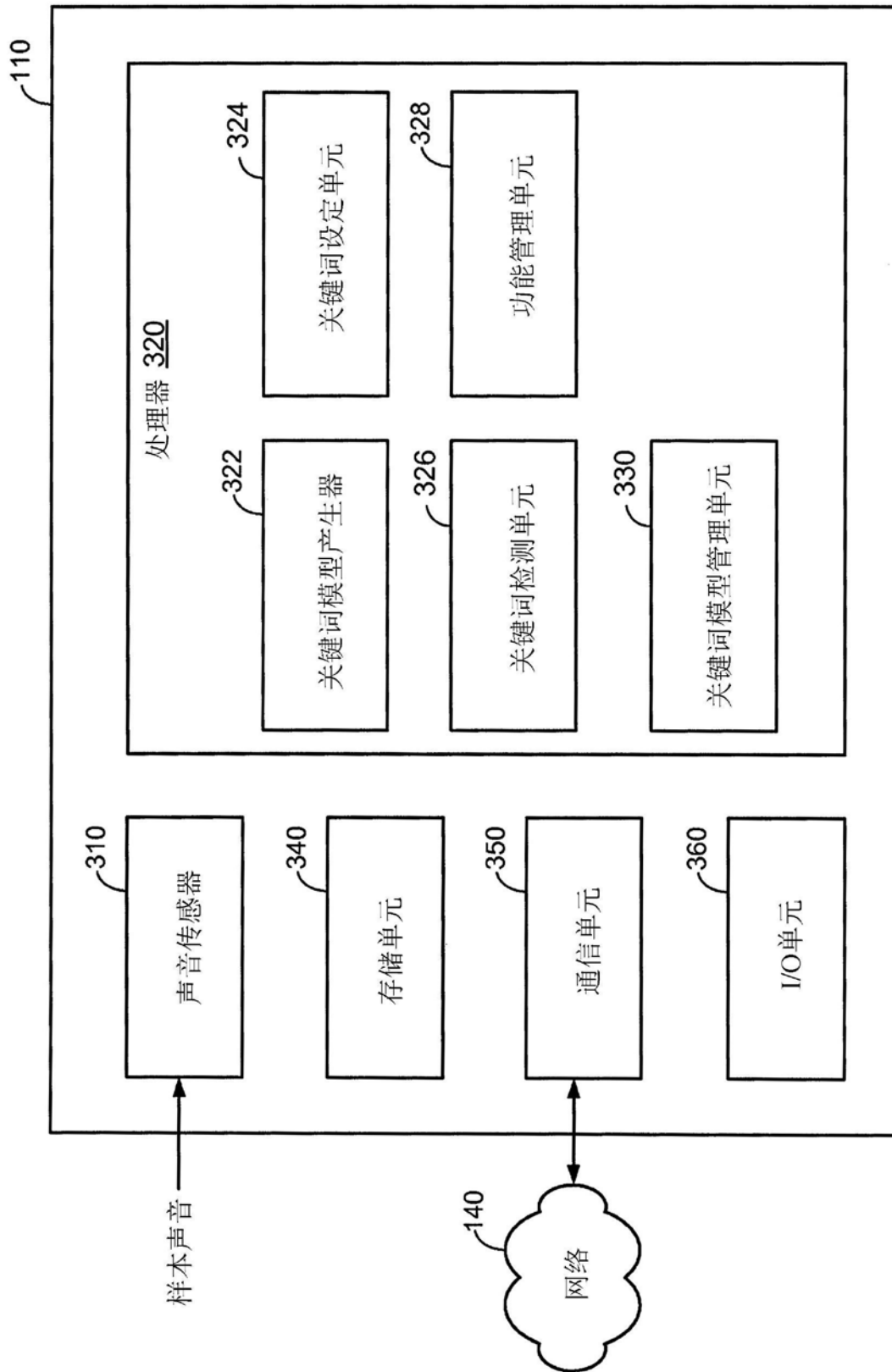


图3

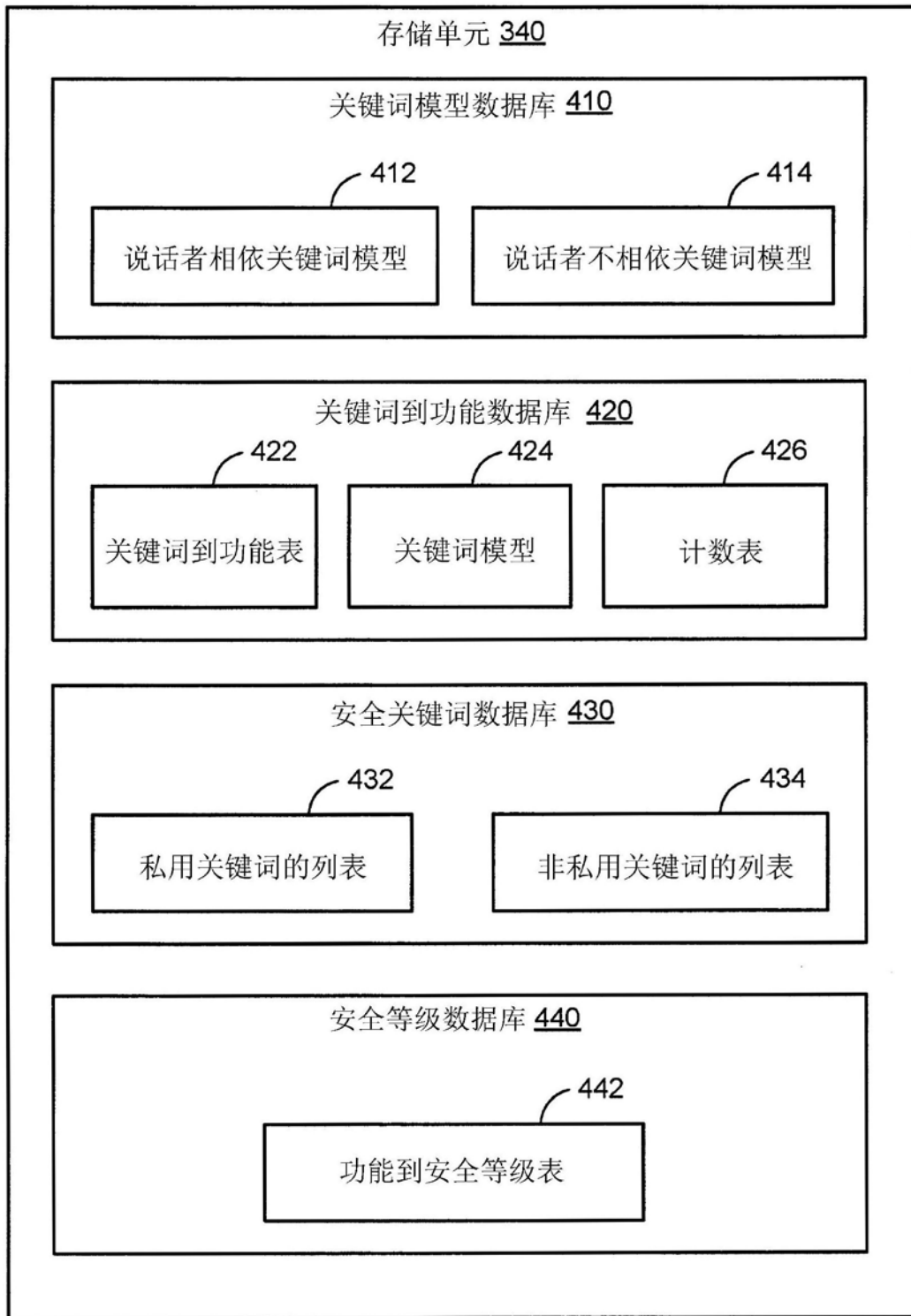


图4

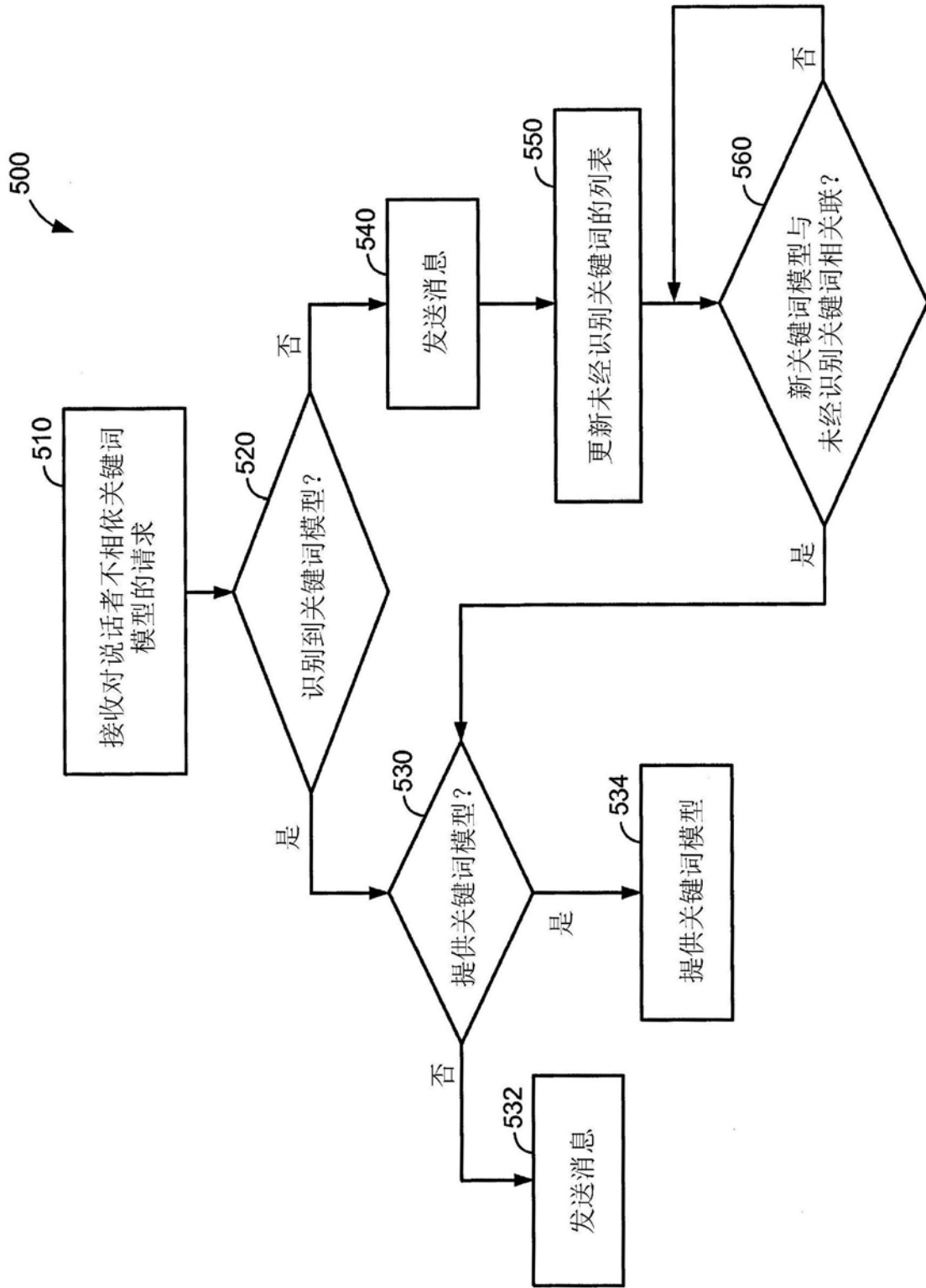


图5

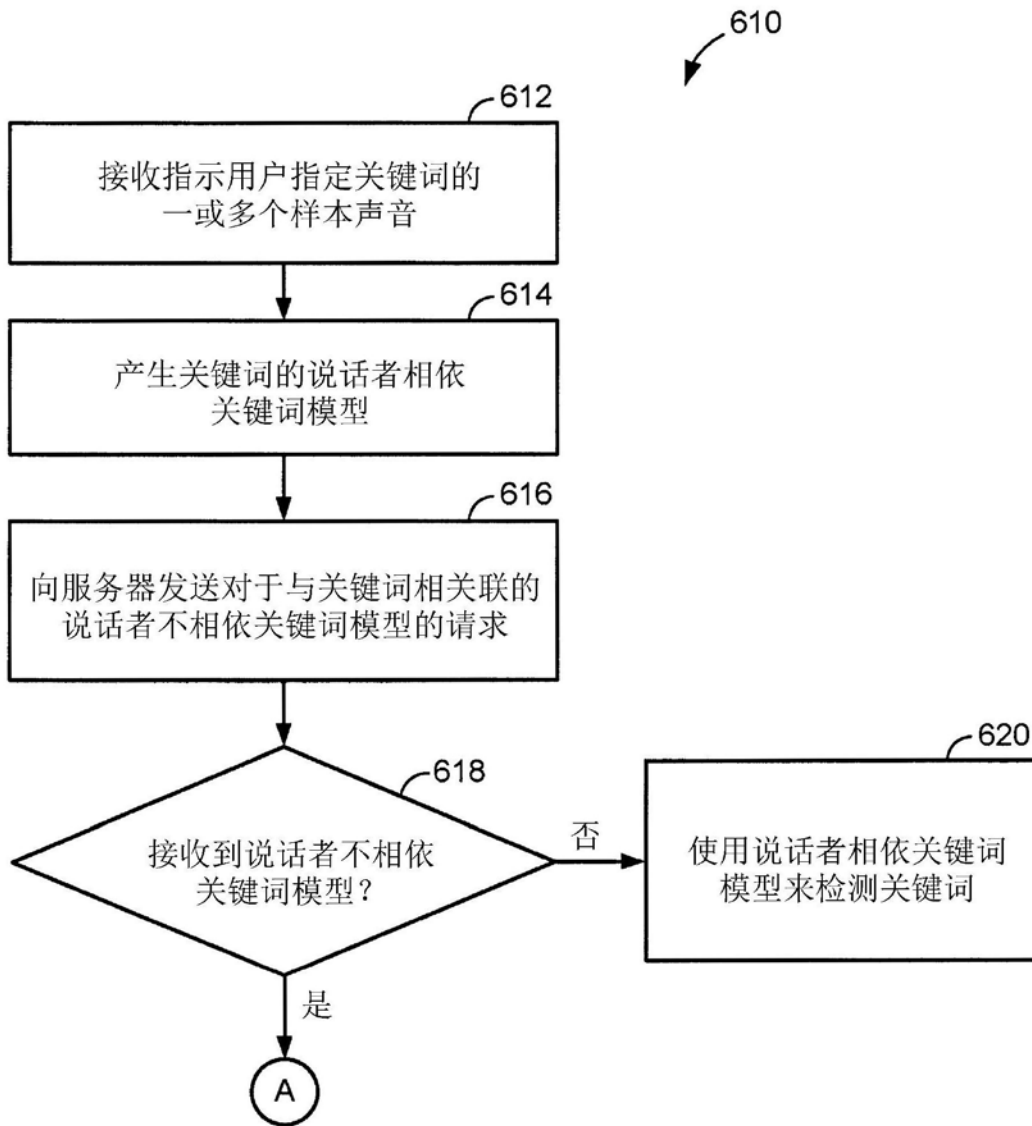


图6A

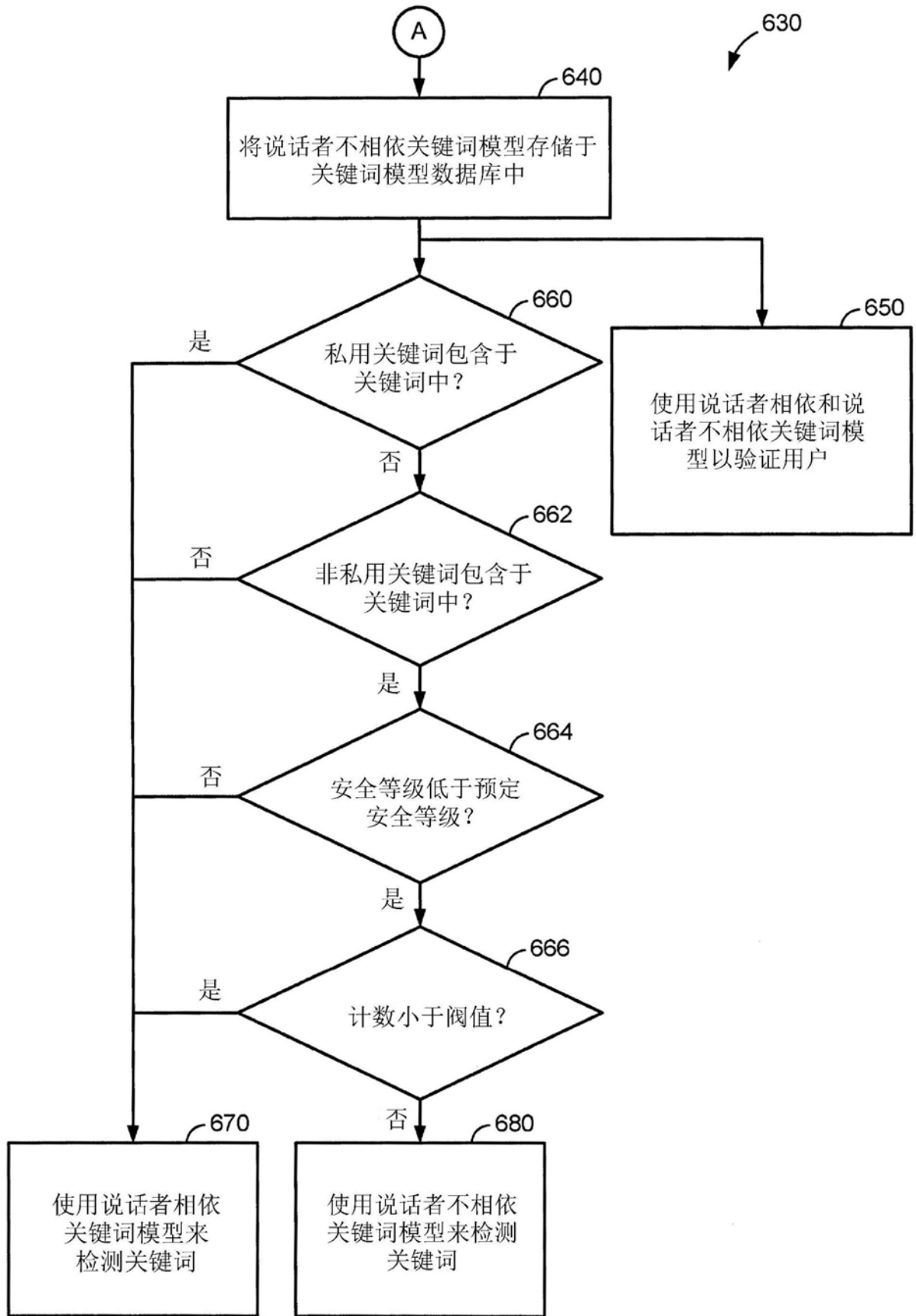


图6B

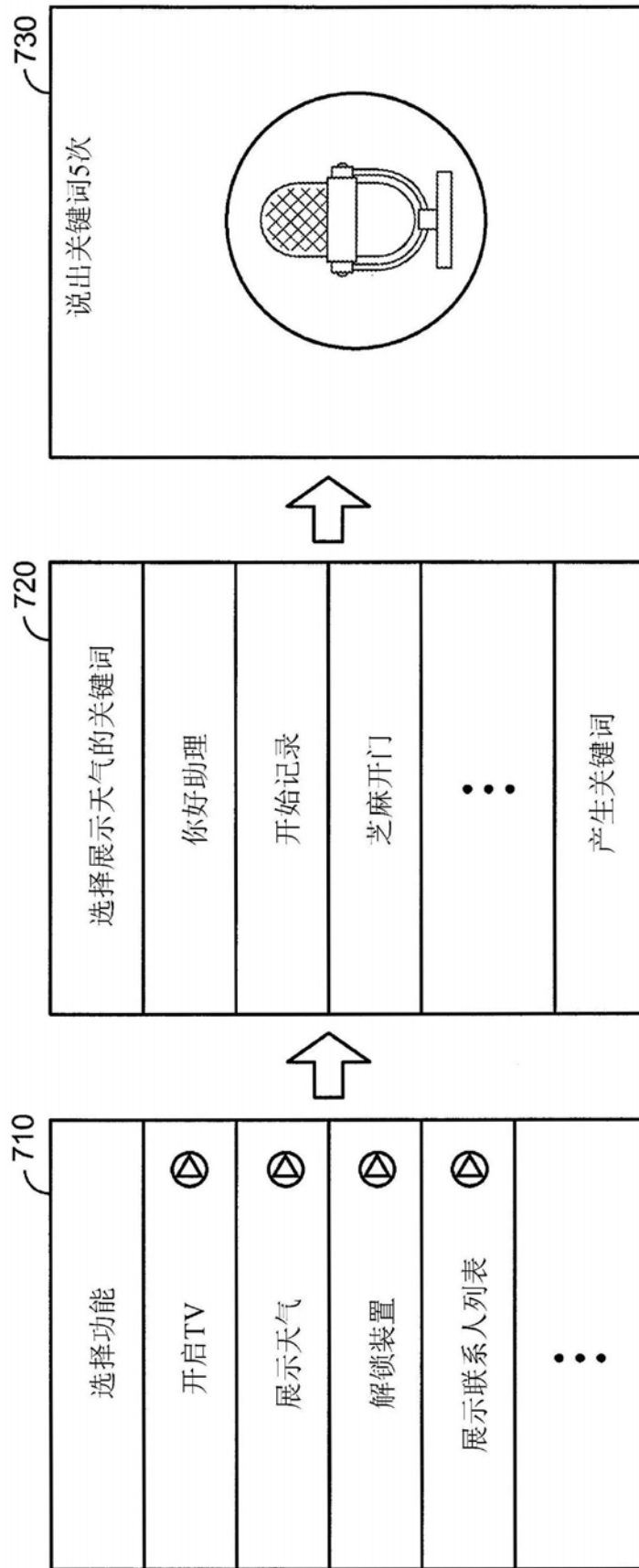


图7

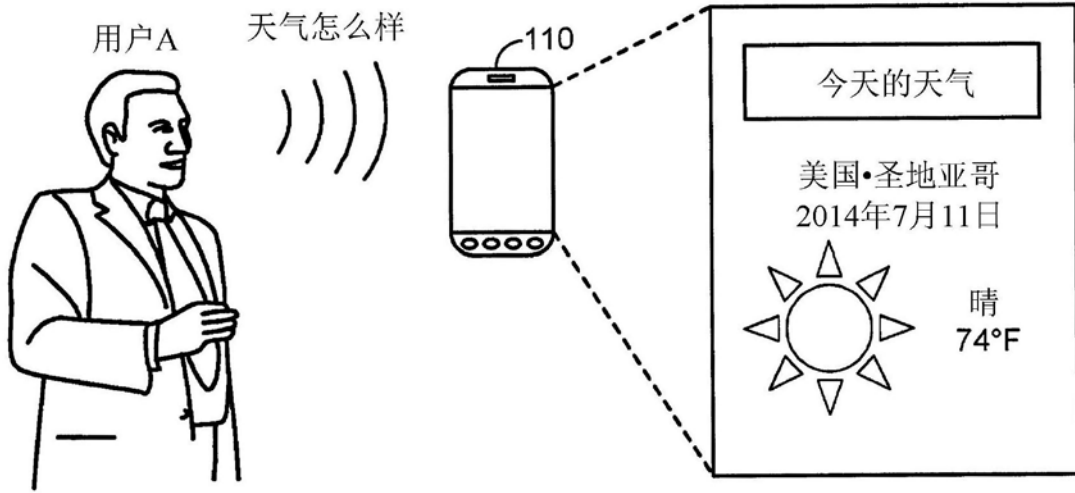


图8A

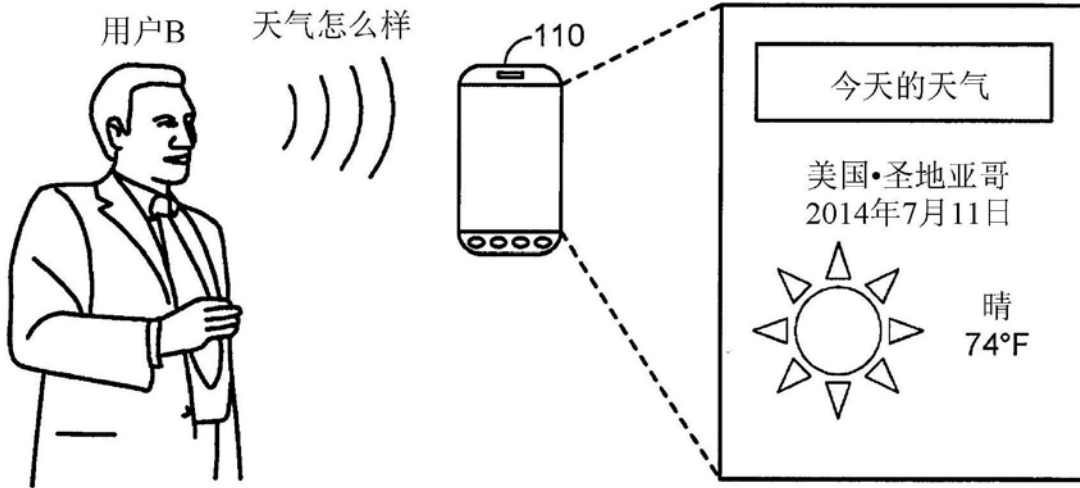


图8B



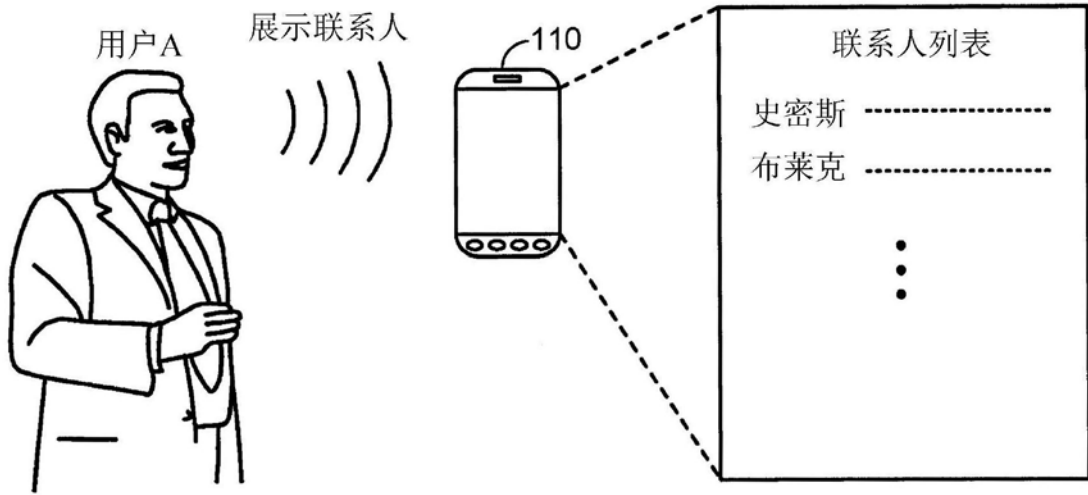


图9A

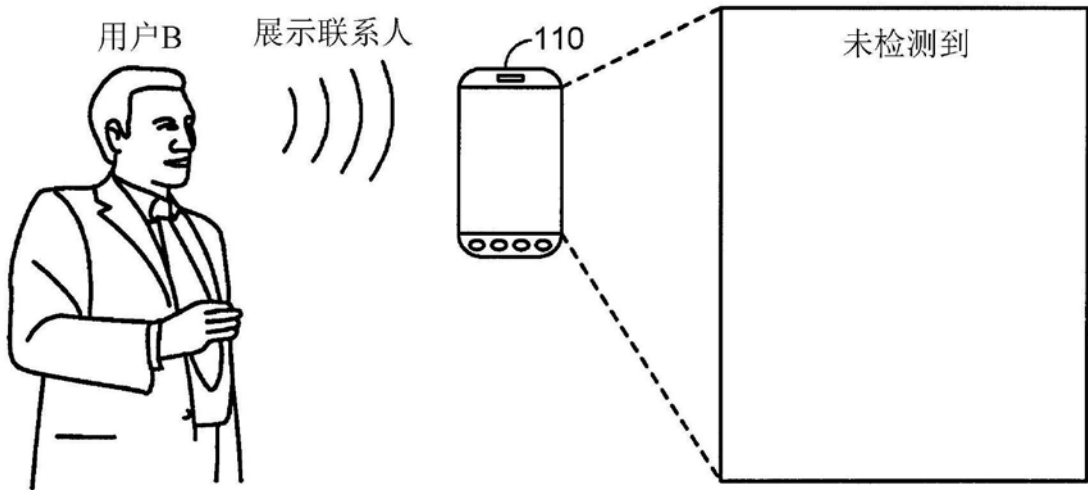


图9B

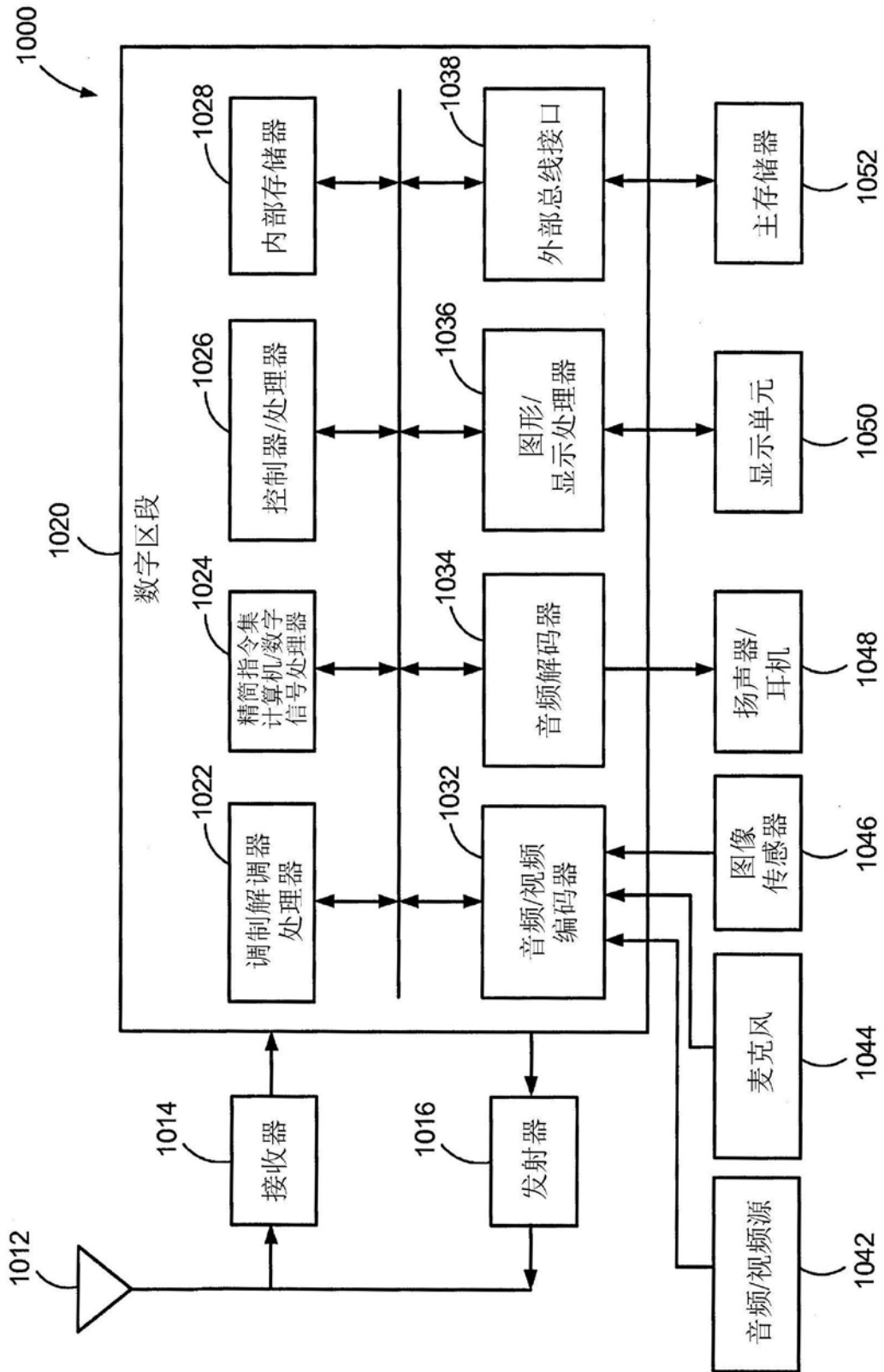


图10

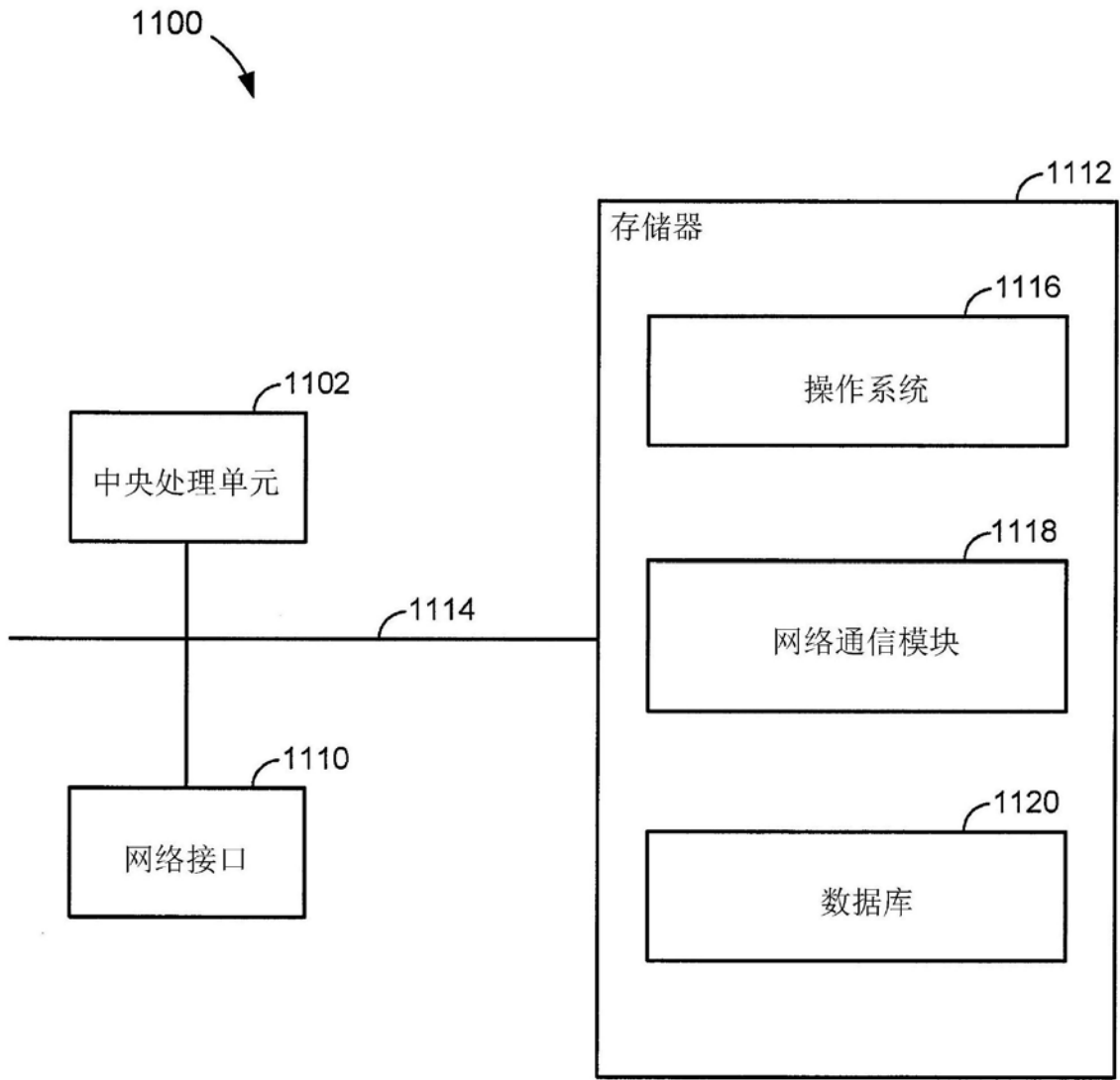


图11