



(12) 发明专利

(10) 授权公告号 CN 110392913 B

(45) 授权公告日 2023. 09. 29

(21) 申请号 201880016425.8

(22) 申请日 2018.05.16

(65) 同一申请的已公布的文献号
申请公布号 CN 110392913 A

(43) 申请公布日 2019.10.29

(30) 优先权数据
62/506,805 2017.05.16 US

(85) PCT国际申请进入国家阶段日
2019.09.06

(86) PCT国际申请的申请数据
PCT/US2018/032873 2018.05.16

(87) PCT国际申请的公布数据
W02018/213381 EN 2018.11.22

(73) 专利权人 谷歌有限责任公司
地址 美国加利福尼亚州

(72) 发明人 V.Q.李 R.沙 O.科拉克 D.比奈

T.王

(74) 专利代理机构 北京市柳沈律师事务所
11105
专利代理师 金玉洁

(51) Int.Cl.
G10L 15/30 (2006.01)
G10L 15/22 (2006.01)

(56) 对比文件
US 2002065657 A1,2002.05.30
CN 1581294 A,2005.02.16
CN 101449569 A,2009.06.03
CN 105265005 A,2016.01.20
US 2009204409 A1,2009.08.13
US 5165095 A,1992.11.17
EP 0477688 A2,1992.04.01
US 2015169336 A1,2015.06.18

审查员 贾晶晶

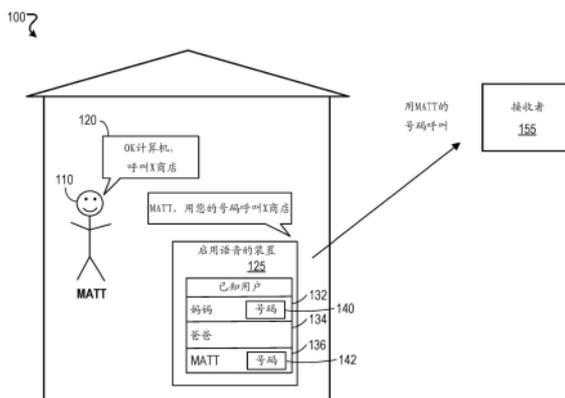
权利要求书3页 说明书24页 附图12页

(54) 发明名称

在共用的启用语音的装置上处理呼叫

(57) 摘要

在一些实现方式中,接收请求语音呼叫的话语,将话语分类为由特定已知用户所说,确定特定已知用户与个人语音号码相关联,以及响应于确定特定已知用户与个人语音号码相关联,用个人语音号码发起语音呼叫。



1. 一种用于自然语音处理的系统,所述系统包括:

一个或多个计算机和存储指令的一个或多个存储装置,所述指令在由一个或多个计算机执行时可操作为使得所述一个或多个计算机进行操作,所述操作包括:

由启用语音的装置接收请求语音呼叫的话语;

由所述启用语音的装置且在发起所述语音呼叫之前,将所述话语分类为由特定已知用户所说;

由所述启用语音的装置且向服务器提供所述话语的表示和所述特定已知用户的指示的二者,所述特定已知用户被分类为已经说出所述话语;

从所述服务器且由所述启用语音的装置接收要呼叫的语音号码和要拨打语音呼叫的指令;

由所述启用语音的装置且在发起所述语音呼叫之前,确定:用于作为所述特定已知用户拨打语音呼叫的语音号码对于分类为已经说出所述话语的特定已知用户而言是否已知,其中用于作为所述特定已知用户拨打语音呼叫的语音号码将所述特定已知用户识别为对所述语音呼叫的接收者的呼叫者;以及

响应于由所述启用语音的装置且在发起所述语音呼叫之前,确定了用于作为所述特定已知用户拨打语音呼叫的语音号码对于分类为已经说出所述话语的特定已知用户而言是已知的,并且响应于接收所述要拨打所述语音呼叫的指令,由所述启用语音的装置通过所述用于作为所述特定已知用户拨打语音呼叫的语音号码而不是通过其他语音号码对所述要呼叫的语音号码发起所述语音呼叫。

2. 根据权利要求1所述的系统,其中将所述话语分类为由特定已知用户所说包括:

确定所述话语中的语音是否与对应于所述特定已知用户的语音相匹配。

3. 根据权利要求1所述的系统,其中将所述话语分类为由特定已知用户所说包括:

确定所述话语的说话者的至少部分的视觉图像是否与对应于所述特定已知用户的视觉信息相匹配。

4. 根据权利要求1所述的系统,其中用于作为所述特定已知用户拨打语音呼叫的所述语音号码还用于向所述特定已知用户拨打呼叫。

5. 根据权利要求1所述的系统,其中确定用于作为所述特定已知用户拨打语音呼叫的语音号码对于分类为已经说出所述话语的特定已知用户而言是否已知包括:

访问所述特定已知用户的账户;

确定所述特定已知用户的账户是否指示电话;以及

确定由所述特定已知用户的账户指示的所述电话通过局部无线连接与所述启用语音的装置连接。

6. 根据权利要求5所述的系统,其中发起所述语音呼叫包括:

通过与所述启用语音的装置连接的所述电话发起所述语音呼叫。

7. 根据权利要求1所述的系统,其中响应于确定用于作为特定已知用户拨打语音呼叫的语音号码对于分类为已经说出所述话语的特定已知用户而言是已知的,并且响应于接收所述要拨打所述语音呼叫的指令包括:

通过互联网协议语音呼叫提供者发起所述语音呼叫。

8. 根据权利要求1所述的系统,其中由所述启用语音的装置确定了用于作为特定已知

用户拨打语音呼叫的语音号码对于分类为已经说出所述话语的特定已知用户而言是已知的包括：

由所述启用语音的装置确定了所述服务器已经向所述启用语音的装置提供所述用于作为所述特定已知用户拨打语音呼叫的语音号码。

9. 根据权利要求1所述的系统, 其中通过所述用于作为所述特定已知用户拨打语音呼叫的语音号码而不是通过其他语音号码对所述要呼叫的语音号码发起所述语音呼叫包括：

识别所述用于作为所述特定已知用户拨打语音呼叫的语音号码; 以及

通过所述用于作为所述特定已知用户拨打语音呼叫的语音号码来呼叫所述要呼叫的语音号码, 所述特定已知用户被分类为已经说出所述话语。

10. 一种用于自然语言处理的计算机实现的方法, 所述方法包括：

由启用语音的装置接收请求语音呼叫的话语；

由所述启用语音的装置和在发起所述语音呼叫之前, 将所述话语分类为由特定已知用户所说；

由所述启用语音的装置且向服务器提供所述话语的表示和所述特定已知用户的指示的二者, 所述特定已知用户的指示被分类为已经说出所述话语；

从所述服务器且由所述启用语音的装置接收要呼叫的语音号码和要拨打语音呼叫的指令；

由所述启用语音的装置且在发起所述语音呼叫之前, 确定: 用于作为所述特定已知用户拨打语音呼叫的语音号码对于分类为已经说出所述话语的特定已知用户而言是否已知, 其中用于作为所述特定已知用户拨打语音呼叫的语音号码将所述特定已知用户识别为对所述语音呼叫的接收者的呼叫者; 以及

响应于由所述启用语音的装置且在发起所述语音呼叫之前, 确定了用于作为所述特定已知用户拨打语音呼叫的语音号码对于分类为已经说出所述话语的特定已知用户而言是已知的, 并且响应于接收所述要拨打所述语音呼叫的指令, 由所述启用语音的装置通过所述用于作为所述特定已知用户拨打语音呼叫的语音号码而不是通过其他语音号码对所述要呼叫的语音号码发起所述语音呼叫。

11. 根据权利要求10所述的方法, 其中将所述话语分类为由特定已知用户所说包括：

确定所述话语中的语音是否与对应于所述特定已知用户的语音相匹配。

12. 根据权利要求10所述的方法, 其中将所述话语分类为由特定已知用户所说包括：

确定所述话语的说话者的至少部分的视觉图像是否与对应于所述特定已知用户的视觉信息相匹配。

13. 根据权利要求10所述的方法, 其中确定用于作为所述特定已知用户拨打语音呼叫的语音号码对于分类为已经说出所述话语的特定已知用户而言是否已知包括：

访问所述特定已知用户的账户信息；

确定所述特定已知用户的账户是否指示电话; 以及

确定由所述特定已知用户的账户指示的所述电话通过局部无线连接与所述启用语音的装置连接。

14. 根据权利要求10所述的方法, 其中发起所述语音呼叫包括：

通过与所述启用语音的装置连接的电话发起所述语音呼叫。

15. 根据权利要求10所述的方法,其中响应于确定用于作为所述特定已知用户拨打语音呼叫的语音号码对于分类为已经说出所述话语的特定已知用户而言是已知的,并且响应于接收所述要拨打所述语音呼叫的指令包括:

通过互联网协议语音呼叫提供者发起所述语音呼叫。

16. 根据权利要求10的所述方法,其中由所述启用语音的装置确定了用于作为特定已知用户拨打语音呼叫的语音号码对于分类为已经说出所述话语的特定已知用户而言是已知的包括:

由所述启用语音的装置确定了所述服务器已经向所述启用语音的装置提供所述用于作为所述特定已知用户拨打语音呼叫的语音号码。

17. 一种非暂时性计算机可读介质,其存储软件,所述软件包括由一个或多个计算机可执行的指令,当一个或多个计算机执行所述指令时使得所述一个或多个计算机进行操作,所述操作包括:

由启用语音的装置接收请求语音呼叫的话语;

由所述启用语音的装置和在发起所述语音呼叫之前将所述话语分类为由特定已知用户所说;

由所述启用语音的装置且向服务器提供所述话语的表示和所述特定已知用户的指示的二者,所述特定已知用户的指示被分类为已经说出所述话语;

从所述服务器且由所述启用语音的装置接收要呼叫的语音号码和要拨打语音呼叫的指令;

由所述启用语音的装置且在发起所述语音呼叫之前,确定:用于作为所述特定已知用户拨打语音呼叫的语音号码对于分类为已经说出所述话语的特定已知用户而言是否已知,其中用于作为所述特定已知用户拨打语音呼叫的语音号码将所述特定已知用户识别为对所述语音呼叫的接收者的呼叫者;以及

响应于由所述启用语音的装置且在发起所述语音呼叫之前,确定了用于作为所述特定已知用户拨打语音呼叫的语音号码对于分类为已经说出所述话语的特定已知用户而言是已知的,并且响应于接收所述要拨打所述语音呼叫的指令,由所述启用语音的装置通过所述用于作为所述特定已知用户拨打语音呼叫的语音号码而不是通过其他语音号码对所述要呼叫的语音号码发起所述语音呼叫。

18. 根据权利要求17所述的介质,其中将所述话语分类为由特定已知用户所说包括:

确定所述话语中的语音是否与对应于所述特定已知用户的语音相匹配。

在共用的启用语音的装置上处理呼叫

[0001] 相关申请的交叉引用

[0002] 本申请要求于2017年5月16日提交的标题为“使用语音控制处理个人电话呼叫”的美国临时专利申请No.62/506,805的优先权,其公开通过引用全文合并于此。

技术领域

[0003] 本说明书总体上涉及自然语音处理。

背景技术

[0004] 启用语音的装置可以响应于来自用户说出的话语而执行动作。例如,用户可以说“OK计算机,今天会下雨吗?”,并且启用语音的装置可听闻地响应“整天都将是晴天”。使用启用语音的装置的益处在于,与启用语音的装置交互可以总体上释放双手。例如,当用户提出问题,启用语音的装置可以提供可听闻的回答而无需用户使用他们的手与任何东西物理交互。然而,常用的启用语音的装置受限于支持的交互类型。

发明内容

[0005] 启用语音的装置可以用于拨打语音呼叫。例如,John Doe可以说“OK计算机,呼叫(555)555-5555”,以使启用语音的装置向电话号码(555)555-5555拨打呼叫。典型地,呼出呼叫与可以用于识别呼叫者的呼叫者号码相关联。例如,当John Doe使用他的电话呼叫(555)555-5555时,接收呼叫的电话可以指示呼叫来自与John Doe的电话相关联的电话号码。

[0006] 与呼叫相关联的呼叫者号码可以是有用的,因为呼叫的接收者可以使用呼叫者号码以决定是否要应答呼叫,而且如果他们需要拨打回叫则也使用呼叫者号码。然而,不同于常规的电话,一些启用语音的装置可能未与可以用于呼叫的呼叫者号码的电话号码相关联。

[0007] 为了在拨打呼叫时提供呼叫者号码,启用语音的装置可以尝试使用说话者的个人语音号码作为呼叫者号码。个人语音号码可以是用于向用户拨打呼叫的号码。例如,当John说“OK计算机,呼叫(555)555-5555”时,启用语音的装置可以使用John Doe的电话的电话号码(555)999-9999作为呼叫者号码。如果启用语音的装置不能确定说话者的个人语音号码,则启用语音的装置可以代替地匿名拨打呼叫,使得呼叫不与可以用于拨打回叫的语音号码相关联。例如,这样的呼叫可以指示“未知号码”或“私人号码”作为呼叫者号码。

[0008] 在一些实例中,如果是向紧急服务呼叫,则可以使用临时号码拨打呼叫,使接收者可以用于回叫启用语音的装置。例如,这样的呼叫可以指示电话号码(555)888-8888,其可以用于在接下来几个小时内向启用语音的装置拨打回叫。

[0009] 附加地或替代地,启用语音的装置可以使用说话者的身份以确定要呼叫的语音号码。例如,当John说“OK计算机,呼叫爸爸”时,启用语音的装置可以识别或以其他方式鉴别John,然后访问John的联系人记录以确定“爸爸”的电话号码。在另一个示例中,当Jane说

“OK计算机,呼叫爸爸”时,启用语音的装置可以由语音识别或其他验证技术将Jane与John区分,并且之后访问Jane的联系人记录以确定“爸爸”的电话号码。在又一个示例中,当访客说“OK计算机,呼叫爸爸”时,启用语音的装置将不会由语音(或其他验证技术)识别访客,并且可以不访问任何用户的联系人记录来确定“爸爸”的电话号码。因此,如在这三个示例中可见,“OK计算机,呼叫爸爸”可以基于说话者的身份具有不同结果。

[0010] 附加地或替代地,在由启用语音的装置拨打语音呼叫期间,启用语音的装置可以对来自用户的话语作出响应。例如,在呼叫期间,启用语音的装置可以对“OK计算机,挂断”、“OK计算机,提高扬声器音量”、“OK计算机,今天天气如何”的命令做出响应。响应于在语音呼叫期间的话语,启用语音的装置可以将至少部分的发声与接收者阻挡开。例如,当用户说“OK计算机,提高扬声器音量”时,启用语音的装置可以提高扬声器音量并且阻挡“提高扬声器音量”,使得接收者只听到“OK计算机”。在另一个示例中,启用语音的装置可以具有向接收者提供音频的等待时间,所以当话语开始于“OK计算机”时可以阻挡整个话语被接收者听见。

[0011] 因此,在一些实现方式中,优点可以是,由多个用户共用的启用语音的装置仍然使用户能够拨打呼叫,并且使得在接收者的电话上呈现的作为呼叫号码的号码是用户的移动计算装置的语音号码。由于人们通常不会接听来自未识别的号码的呼叫,这可以增加使用启用语音的装置拨打的呼叫被应答的可能性。另外,因为正被呼叫的人可能已经基于与用户相关联的语音号码的使用而知道是谁在呼叫,因此呼叫可以更高效。同时,可以提供安全性,在于由于启用语音的装置使用与说话者的语音相匹配的语音号码,因而用户不可以使用启用语音的装置中的任何其他用户的语音号码。

[0012] 在一些实现方式中的另一个优点可以在于,允许对启用语音的装置上的联系人的使用可以使得用户能够更快速地拨打呼叫,因为用户能够快速说出联系人的名字而不是说出语音号码的数字。启用语音的装置还能够消除多个用户之间的联系人的歧义。例如,不同用户的可能具有相同名称“妈妈”的相应的联系人条目,它们与不同电话号码相关联。还可以提供安全性,在于由于启用语音的装置确保了所使用的联系人是与说话者的语音相匹配的联系人,因而用户不可以使用启用语音的装置中的其他用户的联系人。

[0013] 在一些实现方式中的又一个优势可以是,允许在语音呼叫期间对询问进行处理可以使得呼叫能够具有更好的解放双手的体验。例如,响应于自动话务员请求呼叫者以特定号码键入来响应,用户能够虚拟地按下数字。还可以提供安全性,使得在正在处理询问时拨打双向等候,并且一旦解决询问自动结束。附加地,双向等候可以确保对于来自启用语音的虚拟助手的询问的响应不会被来自其他人的声音而模糊不清。例如,在没有双向等候的情况下,其他人可以在输出来自启用语音的虚拟助手的响应的同时说话。

[0014] 在一些方面,在本说明书中所描述的主题可以实施为方法,其可以包含如下动作:接收请求语音呼叫的话语,将话语分类为由特定已知用户所说,确定特定已知用户是否与个人语音号码相关联,以及响应于确定特定已知用户是与个人语音号码相关联的,以个人语音号码发起语音呼叫。

[0015] 在一些实现方式中,将话语分类为由特定已知用户所说包含:确定话语中的语音是否与对应于特定已知用户的语音相匹配。在某些实现方式中,将话语分类为由特定已知用户所说包含:确定说话者的至少部分的视觉图像是否与对应于特定已知用户的视觉信息

相匹配。在一些实现方式中,确定特定已知用户是否与个人语音号码相关联包含:访问特定已知用户的账户信息,并且确定用户的账户信息是否存储特定已知用户的语音号码。

[0016] 在某些实现方式中,确定特定已知用户是否与个人语音号码相关联包含:向服务器提供对特定已知用户的指示和话语的表示,并且从服务器接收特定已知用户的个人语音号码、要呼叫的语音号码和要拨出语音呼叫的指令。在一些实现方式中,确定特定已知用户是否与个人语音号码相关联包含:访问特定已知用户的账户,确定用户的账户是否指示电话,并且确定该电话与启用语音的装置连接。

[0017] 在某些实现方式中,以个人语音号码发起语音呼叫包含,通过与启用语音的装置连接的电话发起语音呼叫。在一些实现方式中,响应于确定特定已知用户与个人语音号码相关联,以个人语音号码发起语音呼叫包含,通过互联网协议语音(Voice over Internet Protocol)呼叫提供者来发起语音呼叫。

[0018] 在一些方面,在本说明书中所描述的主题可以实施为方法,其可以包含如下动作:接收请求语音呼叫的话语,将话语分类为由特定已知用户所说,响应于将话语分类为由特定已知用户所说,基于特定已知用户的联系人确定要呼叫的接收者语音号码,并且向接收者语音号码发起语音呼叫。

[0019] 在一些实现方式中,响应于将话语分类为由特定已知用户所说,获得由特定已知用户创建的联系人条目包含:响应于将话语分类为由特定已知用户所说,确定特定已知用户的联系人条目是可用的,并且响应于确定特定已知用户的联系人条目是可用的,获得由特定已知用户创建的联系人条目。在某些实现方式中,响应于将话语分类为由特定已知用户所说,基于特定已知用户的语音联系人确定要呼叫的接收者语音号码包含:响应于将话语分类为由特定已知用户所说,获得由特定已知用户创建的联系人条目,从联系人条目之中识别特定联系人条目,其中特定联系人条目包含与话语匹配的姓名,以及确定由特定联系人条目指示的语音号码作为接收者语音号码。

[0020] 在一些实现方式中,从联系人条目之中识别特定联系人条目,其中特定联系人条目包含与话语匹配的姓名包含:生成话语的转录,以及确定转录包含姓名。在某些实现方式中,将话语分类为由特定已知用户所说包含:获得由启用语音的装置确定话语中的语音与对应于特定已知用户的语音相匹配的指示。在一些实现方式中,将话语分类为由特定已知用户所说包含:确定话语中的语音是否与对应于特定已知用户的语音相匹配。在某些实现方式中,向接收者语音号码发起语音呼叫包含:向启用语音的装置提供接收者语音号码,以及指令向该接收者语音号码发起语音呼叫。

[0021] 在一些实现方式中,动作包含:接收请求第二语音呼叫的第二话语,将第二话语分类为不是由启用语音的装置的任何已知用户所说,以及响应于将第二话语分类为不是由启用语音的装置的任何已知用户所说,在不访问启用语音的装置的任何已知用户的语音联系人的情况下发起第二语音呼叫。

[0022] 在一些方面,本说明书中所描述的主题可以实施为方法,其包含如下动作:确定在第一方和第三方之间的语音呼叫期间第一方已经对启用语音的虚拟助手说出询问,响应于确定在第一方和第三方之间的语音呼叫期第一方已经对启用语音的虚拟助手说出询问,将第一方和第三方之间的语音呼叫置于等候,确定启用语音的虚拟助手已经解决询问,以及,响应于确定启用语音的虚拟助手已经处理询问,从等候中恢复第一方和第三方之间的

语音呼叫。

[0023] 在一些实现方式中,确定在第一方和第二方之间的语音呼叫期间第一方已经对启用语音的虚拟助手说出询问包含:由启用语音的装置确定在语音呼叫期间由第一方说出热词。在某些实现方式中,将第一方和第二方之间的语音呼叫置于等候包含:向语音呼叫提供者提供指令以将语音呼叫置于等候。在一些实现方式中,将第一方和第二方之间的语音呼叫置于等候包含:将来自麦克风的音频路由到启用语音的虚拟助手而不是语音服务器,并向扬声器路由来自启用语音的虚拟助手的音频,而不是来自语音服务器的音频。

[0024] 在某些实现方式中,确定启用语音的虚拟助手已经解决询问包含:向启用语音的虚拟助手提供询问和在启用语音的装置上正在进行语音呼叫的指示,并且从启用语音的虚拟助手接收对询问的响应和询问被解决的指示。在一些实现方式中,从启用语音的虚拟助手接收对询问的响应和询问被解决的指示包含:接收要输出的作为对询问的响应的音频,以及接收具有指示询问是否被解决的数值的二进制旗标。在某些实现方式中,启用语音的虚拟助手配置为识别对应于询问的命令,确定在语音呼叫期间可以执行该命令,并且响应于确定可以在语音呼叫期间执行该命令,确定响应以指示对命令的应答。

[0025] 在一些实现方式中,启用语音的虚拟助手配置为识别对应于询问的命令,确定在语音呼叫期间无法执行该命令,并且响应于确定在语音呼叫期间无法执行该命令,确定响应以指示无法执行该命令。在某些实现方式中,确定在语音呼叫期间无法执行该命令包含:获得在语音呼叫期间可以正常执行的命令列表,以及确定所识别的命令不在命令列表中。在一些实现方式中,确定在语音呼叫期间无法执行该命令包含:获得在语音呼叫期间无法正常执行的命令列表,以及确定所识别的命令在命令列表中。

[0026] 在某些实现方式中,响应于确定启用语音的虚拟助手已经处理询问,从等候中恢复第一方和第二方之间的语音呼叫包含:向语音呼叫提供者提供指令以从等候中恢复语音呼叫。在一些实现方式中,响应于确定启用语音的虚拟助手已经处理询问,从等候中恢复第一方和第二方之间的语音呼叫包含:将来自麦克风的音频路由到语音服务器而不是启用语音的虚拟助手,并向扬声器路由来自语音服务器的音频,而不是来自启用语音的虚拟助手的音频。在某些实现方式中,响应于确定启用语音的虚拟助手已经处理询问,从等候中恢复第一方和第二方之间的语音呼叫包含:从启用语音的虚拟助手接收产生双音调多频率信号(dual-tone multi-frequency signal)的指令,以及响应于从启用语音的虚拟助手接收产生双音调多频率信号的指令,向语音呼叫提供者提供第二指令以在向语音呼叫提供者提供指令之后产生双音调多频率信号,以从等候中恢复语音呼叫。在一些实现方式中,启用语音的辅助服务器配置为确定询问指示要生成一个或多个双音调多频率信号,以及对应于一个或多个双音调多频率信号的一个或多个号码的命令。该方面和其他方面的其他实现方式包含配置为进行该方法、在计算机存储装置上编码的对应的系统、设备和计算机程序。一个或多个计算机的系统可以因此借助于系统上所安装的软件、固件、硬件及其组合来配置,其在操作中使得系统进行动作。一个或多个计算机程序因此配置为具有当由数据处理设备执行时使得设备进行动作的指令。

[0027] 下面在所附图和描述中提出在本说明书中所描述的主题的一个或多个实现方式的细节。通过说明书、附图和权利要求书,主题的其他特征、方面和潜在优势将变得显而易见。

附图说明

- [0028] 图1A-1D是图示与拨打呼叫的启用语音的装置示例交互的框图。
- [0029] 图2是图示拨打呼叫的过程的示例的流程图。
- [0030] 图3是图示确定要呼叫的语音号码的过程的示例的流程图。
- [0031] 图4是图示在呼叫期间与启用语音的装置交互的示例的框图。
- [0032] 图5是图示与拨打呼叫的启用语音的装置交互的系统的示例的框图。
- [0033] 图6是图示确定呼叫者号码的过程的示例的流程图。
- [0034] 图7是图示确定要呼叫的接收者号码的过程的示例的流程图。
- [0035] 图8是图示语音呼叫期间处理询问的过程的示例的流程图。
- [0036] 图9是计算装置示例的示意图。
- [0037] 在各附图中的相同的参考数字和指定指示相同的元件。

具体实施方式

[0038] 图1A-1D是图示系统100中不同示例性交互的框图。系统100包含启用语音的装置125,其可以由用户110使用以呼叫接收者155,而不需要用户110通过触摸与系统100物理交互。

[0039] 在一些实现方式中,启用语音的装置125可以响应于检测到话语进行动作,话语包含用户说出以称呼(address)启用语音的装置125的预定短语,也称为热词。例如,热词可以是“OK计算机”或一些其他短语,用户必须紧挨在对启用语音的装置125说出任何请求之前将其说出。

[0040] 为了用呼叫者号码拨打呼叫,启用语音的装置125可以将话语分类为由特定已知用户所说,并且用特定已知用户的呼叫者号码拨打呼叫。已知用户可以是注册为系统100的用户的用户,并且访客用户可以是没有注册为系统100的用户的用户。例如,“妈妈”可以注册为启用语音的装置125的已知用户,并且启用语音的装置125稍后可以分类话语是否由已知用户“妈妈”所说。

[0041] 例如,图1A图示启用语音的装置125接收话语“OK计算机,呼叫X商店”,将说话者分类为已知的说话者“Matt”,并且用“Matt”存储的电话号码向X商店拨打呼叫。在另一个示例中,图1B图示启用语音的装置125接收话语“OK计算机,呼叫X商店”,将说话者分类为已知的说话者“爸爸”,并且向X商店拨打匿名的呼叫。在又一个示例中,图1C图示启用语音的装置125接收话语“OK计算机,呼叫X商店”,将说话者分类为访客说话者,并且向X商店拨打匿名的呼叫。

[0042] 在又一个示例中,图1D图示启用语音的装置125接收话语“OK计算机,紧急呼叫”,将说话者分类为访客说话者,并且以临时号码向紧急服务拨打呼叫。临时号码可以是紧急服务可以在至少某一持续时间(例如一小时、两小时、二十四小时等)内用来向启用语音的装置125拨打回叫的语音号码。临时号码可以对说话者是未知的,使得临时号码仅可以由紧急服务用于在紧急期间回叫。

[0043] 更详细地,启用语音的装置125可以包含一个或多个麦克风以及一个或多个扬声器。启用语音的装置125可以使用一个或多个麦克风接收话语,并且通过一个或多个扬声器输出可听闻的响应。

[0044] 启用语音的装置125可以为启用语音的装置125中的每个已知用户存储用户账户信息。例如,启用语音的装置125可以存储已知用户“妈妈”的用户账户信息132的第一集合、已知用户“爸爸”的用户账户信息134的第二集合、已知用户“Matt”的用户账户信息136的第三集合。

[0045] 用户的用户账户信息可以指示语音号码,其可以当用户拨打呼叫时用作呼叫者号码。例如,“妈妈”的用户账户信息132的第一集合可以存储(555) 111-1111的第一电话号码140,“爸爸”的用户账户信息134的第二集合可以是空白的(即没有存储的电话号码),并且“Matt”的用户账户信息136的第三集合可以存储(555) 222-2222的第二电话号码142。在某些实施例中,用户的用户账户信息可以存储多个号码,诸如“家庭”、“工作”、“移动”等。

[0046] 用户的用户账户信息可以指示说话者识别特征,其可以用于识别说话者是否为用户。例如,针对“妈妈”的用户账户信息132的第一集合可以存储美尔频率倒谱系数(mel-frequency cepstral coefficient, MFCC)特征,其可以共同地形成特征向量,该特征向量表示用户“妈妈”先前多次说出热词。

[0047] 在一些实现方式中,用户可以通过移动计算装置上的伴随应用程序注册为已知用户,其中移动计算装置经由本地无线连接与启用语音的装置125通信。例如,用户“妈妈”可以通过她电话上的伴随应用程序登入她的账户中,然后在伴随应用程序中指示她想要注册为启用语音的装置125的已知用户,并且然后多次说出热词到她电话中。

[0048] 作为注册的部分,或者之后,用户可以指示用户是否想要关联语音号码以用作呼叫者号码,以用于用户使用启用语音的装置125拨打的呼叫。例如,用户“妈妈”可以指示她想要由启用语音的装置125拨打她的呼叫,该启用语音的装置125指示呼叫者号码是她的电话的电话号码。在另一个示例中,用户“妈妈”可以指示,当将她的电话(例如通过蓝牙连接)连接到启用语音的装置125时,她想要通过她的电话由启用语音的装置125拨打她的呼叫。

[0049] 启用语音的装置125可以通过多种类型的呼叫提供者拨打呼叫。例如,启用语音的装置125可以具有互联网连接,并使用互联网语音协议(VoIP)拨打呼叫。在另一个示例中,启用语音的装置125可以与蜂窝网络通信并使用蜂窝网络拨打呼叫。在又一个示例中,启用语音的装置125可以与蜂窝(或陆线)电话通信并且通过电话拨打呼叫,因此用户对启用语音的装置125说话并收听,但呼叫是通过电话来建立。

[0050] 在一些实现方式中,用户可以指示要用作呼叫者号码的语音号码,以用于基于选择用户想要使用的呼叫提供者而使用启用语音的装置125拨打的呼叫。例如,“妈妈”可以指示她想要通过例如蜂窝网络提供者的第一呼叫提供者拨打呼叫,对此她还可以使用电话号码(555) 111-1111接收呼叫,并且稍后指示她代替地想要通过例如VoIP提供者的第二呼叫提供者拨打呼叫,对此她可以使用电话号码(555) 111-2222接收呼叫。

[0051] 在一些实现方式中,启用语音的装置125可以基于上下文信息将话语分类为由特定用户所说。上下文信息可以包含一个或多个音频、视觉或其他信息。关于音频信息,启用语音的装置125可以基于已知用户的一个或多个话语的说话者识别特征(例如美尔频率倒谱系数(MFCC)特征,其可以共同地形成特征向量)来分类话语。例如,启用语音的装置125可以存储已知用户的每一个说出“OK计算机”的说话者识别特征。响应于在当前接收的话语中的说话者识别特征充分匹配已知用户“爸爸”说出“OK计算机”的存储的说话者识别特征,启用语音的装置125可以将话语分类为由已知用户“爸爸”说出的。

[0052] 在另一个示例中,启用语音的装置125可以基于话语的完整音频分类话语。例如,启用语音的装置125可以确定在完整接收的话语中的语音是否与对应于已知用户“爸爸”的语音匹配。

[0053] 关于视觉信息,启用语音的装置125可以接收说话者的至少部分的一个或多个图像,并尝试基于一个或多个图像识别说话者。例如,启用语音的装置125可以包含相机,并且确定相机视野内说话者的面部由启用语音的装置125分类为与对应于已知用户“爸爸”的面部相匹配。在其他示例中,启用语音的装置125可以尝试匹配以下中的一个或多个:说话者的指纹、视网膜扫描、面部识别、姿势、另一个装置的共存、或来自另一个装置或软件的元件身份的确认。

[0054] 启用语音的装置125可以是与远程服务器协作拨打呼叫的本地前端装置。例如,当启用语音的装置125接收话语“OK计算机,呼叫X商店”时,启用语音的装置125可以检测何时说话者说出热词“OK计算机”,基于话语“OK计算机”中的说话者识别特征将用户分类为“妈妈”,并且向服务器提供“呼叫X商店”的表示以及说话者是“妈妈”的指示。服务器然后可以转录“呼叫X商店”,确定文本“呼叫X商店”对应于拨打呼叫的动作,X商店的电话号码为(555)999-9999,以及“妈妈”已经指示应该通过她的VoIP账户以(555)111-1111的呼叫者号码拨打她的呼叫。服务器然后可以发送“以VoIP账户(555)111-1111呼叫(555)999-9999”的指令到启用语音的装置125。在其他实现方式中,启用语音的装置125可以独立于远程服务器进行由远程服务器所描述的动作。

[0055] 在一些实现方式中,启用语音的装置125可以基于除了音频信息和视觉信息之外的其他信息来分类话语。具体而言,启用语音的装置125可以基于说话者识别特征和来自验证说话的用户的身份的用户的确认来分类话语。附加地,启用语音的装置125可以基于说话者的至少部分的一个或多个接收的图像和来自验证说话的用户的身份的用户的确认来分类话语。例如,如上文所提及的,启用语音的装置125可以接收来自说话的用户的一个或多个话语。启用语音的装置125可以确定:在一个或多个接收的话语中的说话者识别特征充分匹配已知用户“爸爸”说“OK计算机”的所存储的说话者识别特征。作为响应,启用语音的装置125可以通过提问用户“是爸爸在说话吗?”确认说话用户是“爸爸”的判断。说话者可以通过回答“是”或者“不是”来响应,以便验证启用语音的装置125的确认。如果说话者回答“不是”,则启用语音的装置125可以问附加的问题,诸如“说话者的名字是什么?”,以确定名字是否匹配启用语音的装置125中所存储的已知用户姓名。

[0056] 图2是图示拨打呼叫的过程200的示例的流程图。可以由诸如图1A-1D的系统100的一个或多个计算系统进行过程200的操作。过程200包含接收话语(210)。例如,启用语音的装置125可以接收“OK计算机,呼叫(555)999-9999”的话语。

[0057] 过程200包含确定是否呼叫紧急服务(212)。例如,启用语音的装置125可以确定呼叫号码不是呼叫紧急服务,因为(555)999-9999不与任何紧急服务相关联。在另一个示例中,启用语音的装置125可以确定呼叫号码“911”是呼叫紧急服务,因为911与紧急服务相关联。

[0058] 如果过程200确定呼叫的是紧急服务,则过程200包含用临时号码(214)发起呼叫。例如,启用语音的装置125可以请求呼叫提供者生成可以使用二十四小时向启用语音的装置回叫的电话号码,并且然后发起呼叫紧急服务,显示临时号码作为呼叫者号码。

[0059] 如果过程200确定呼叫的不是紧急服务,则过程200包含确定话语的说话者是否为已知用户(216)。例如,启用语音的装置125可以响应于将说话者分类为已知用户“Matt”,确定“OK计算机,呼叫(555)999-9999”的说话者是已知用户。在另一个示例中,启用语音的装置125可以响应于将说话者分类为已知用户“爸爸”,确定说话者是已知用户。在又一个示例中,启用语音的装置125可以响应于将说话者分类为访客用户,确定说话者不是已知用户。

[0060] 在一些实现方式中,确定话语的说话者是否为已知用户包含确定话语中的语音是否与对应于特定已知用户的语音相匹配。例如,启用语音的装置125可以确定说话者说“OK计算机”的方式与已知用户“Matt”如何说“OK计算机”相匹配,并且作为响应,将说话者分类为已知用户“Matt”。在另一个示例中,启用语音的装置125可以确定说话者说“OK计算机”的方式与已知用户“爸爸”如何说“OK计算机”相匹配,并且作为响应,将说话者分类为已知用户“爸爸”。附加地或替代地,确定话语的说话者是否为已知用户包含确定说话者的至少部分的视觉图像是否与对应于特定已知用户的视觉信息相匹配。

[0061] 如果过程200确定话语的说话者是已知用户,则过程200包含确定已知用户是否与个人语音号码相关联(218)。例如,启用语音的装置125可以确定已知用户“Matt”具有账户信息,该账户信息指示已知用户想要在通过启用语音的装置125拨打呼叫时使用的呼叫提供者,并且作为响应,确定已知用户关联个人电话号码。在另一个示例中,启用语音的装置125可以确定已知用户“爸爸”没有账户信息,该账户信息指示已知用户想要在通过启用语音的装置125拨打呼叫时使用的呼叫提供者,并且作为响应,确定已知用户没有关联个人电话号码。

[0062] 如果过程200确定已知用户与个人语音号码相关联,则过程200包含用个人语音号码发起呼叫(220)。例如,启用语音的装置125可以联系由“Matt”的账户信息所指示的呼叫提供者,并请求为“Matt”向电话号码(555)999-9999拨打呼叫。

[0063] 回到218,如果过程200确定已知用户没有与个人语音号码相关联,则过程包含发起匿名呼叫(222)。例如,启用语音的装置125可以请求呼叫提供者向(555)999-9999拨打匿名呼叫。

[0064] 回到216,如果过程200确定话语的说话者不是已知用户,则过程200包含发起匿名呼叫(222),如上所述的222。

[0065] 尽管确定是否呼叫紧急服务(212)首先在过程200中示出,但是过程200可以是不同的。例如,过程200可以代替地首先如上所述在(216)中确定说话者是已知用户,然后如上所述在(218)中确定已知用户与个人语音号码相关联,并且接下来如上所述在(212)中确定呼叫是紧急服务,并且然后使用已知用户的个人语音号码。

[0066] 向紧急答复者提供已知用户的个人语音号码,而不是启用语音的装置125的临时号码的一个原因,在于无论已知用户是否在启用语音的装置125附近紧急答复者都可以联系已知用户。

[0067] 图3是图示确定要呼叫的语音号码的过程300的示例的流程图。可以由诸如图1A-1D的系统100的一个或多个计算系统进行过程300的操作。

[0068] 过程300包含接收请求呼叫的话语(310)。例如,启用语音的装置125可以接收用户110的话语,以请求诸如“OK计算机,呼叫奶奶”的话语。

[0069] 过程300包含确定话语的说话者是否为已知用户(312)。例如,启用语音的装置125

可以将说话者分类为已知用户“妈妈”。

[0070] 如果过程300确定话语的说话者是已知用户,则过程300包含确定已知用户的个人联系人是否可用(314)。例如,启用语音的装置125可以基于确定启用语音的装置125已经访问已知用户“妈妈”的联系人记录,而确定已知用户“妈妈”的个人联系人是可用的。已知用户的个人联系人可以是指为已知用户创建的电话联系人条目。例如,已知用户可以通过如下创建已知用户的电话联系人条目:打开创建新的电话联系人条目的界面,键入电话号码“(123)456-7890”和联系人姓名“John Doe”,并且然后选择以创建标记有“John Doe”的姓名和指示电话号码“(123)456-7890”的电话条目。可以由已知用户的全部个人联系人来形成已知用户的联系人列表。例如,已知用户的联系人列表可以包含“John Doe”的联系人条目以及由已知用户创建的其他联系人条目。

[0071] 如果过程300确定个人联系人对已知用户是可用的,则过程300包含使用个人联系人确定与接收者相关联的号码(316)。例如,启用语音的装置125从已知用户“妈妈”的联系人记录中对接收者“奶奶”扫描个人联系人列表,并取回与“奶奶”相关联的号码。

[0072] 回到314,如果过程300代替地确定已知用户的个人联系人不可用,则过程300包含在没有与已知用户相关联的个人联系人的情况下确定接收者号码(318)。例如,启用语音的装置125可以在互联网搜索接收者号码。在该实例中,启用语音的装置125可以在互联网搜索使用地理位置服务可能在已知用户的附近的对应于“奶奶”的多个接收者号码,而不能够识别出接收者号码,并且向已知用户提供语音消息以说明“没有找到联系人号码”。如果没有找到接收者号码,则启用语音的装置125可以促使说话者说出要呼叫的语音号码然后呼叫该号码。

[0073] 回到312,如果过程300代替地确定话语的说话者不是已知用户,则过程300包含如上所述的在没有个人联系人的情况下确定接收者号码(318)。

[0074] 图4是图示在呼叫期间与启用语音的装置交互的示例的框图。4图示了阶段(A)到(C)的各种操作,其可以以指示的顺序中或另一个顺序进行。

[0075] 在一些实现方式中,启用语音的装置125可以响应于检测包含诸如热词的预定短语的话语进行动作,用户在呼叫期间说出该预定短语以称呼启用语音的装置125。例如,图4图示了启用语音的装置125接收发生“OK计算机,呼叫X商店”,将说话者分类为已知的说话者“Matt”,并且用“Matt”的所存储的电话号码向X商店拨打呼叫。附加地,说话者“Matt”在接收者未接听呼叫期间向启用语音的装置125通信命令。响应于电话呼叫期间的命令,启用语音的装置125可以将至少部分的话语与接收者阻挡开。

[0076] 在阶段(A)期间,启用语音的装置125接收话语120“OK计算机,呼叫X商店”。响应于接收话语120,启用语音的装置125使用前面提及的方法之一将说话者分类为已知说话者“Matt”,并且向“Matt”返回陈述“用您的号码呼叫X商店”的响应。对用户110的响应指示:启用语音的装置125通过分类说话者、采取与命令相关联的动作、并且使用与“Matt”相关联的号码而理解了话语。在阶段(B)期间,启用语音的装置125向例如X商店的接收者155发起呼叫。例如启用语音的装置125在用户110和接收者155之间发起电话呼叫。启用语音的装置125使用可以由接收者155用于回叫用户110的用户110的号码来呼叫接收者155。接收者155通过说出“你好?”应答电话呼叫。作为响应,用户110经由启用语音的装置125对接收者155说“嘿,商店开门吗?”接收者155应答道“是的,晚上10点关门。”

[0077] 在阶段(B)期间,启用语音的装置125在与接收者155的电话呼叫期间从来自用户110的命令中检测热词。例如,启用语音的装置125获得来自用户110的命令,其陈述“OK计算机,几点了”。响应于电话呼叫期间所接收的话语,启用语音的装置125发送用户110说出热词“OK计算机”,但是阻止热词之后的命令,使得接收者155听到“OK计算机”而没有听到“几点了”。启用语音的装置125仅对用户110答复道“下午9点”,使得接收者155没有听到该答复。替代地,等待时间量可以引入到通信中,以允许启用语音的装置125在作为呼叫的部分同样对接收者进行广播之前检测热词。以这种方法,不仅与热词相关联的指令而且热词本身可以被阻止作为呼叫的部分对接收者传递。

[0078] 在一些实现方式中,启用语音的装置125可以通过在检测用户110说出热词之后在用户110和接收者155之间拨打双向等候,防止接收者155听到用户110和启用语音的装置125之间的通信。在双向等候期间,接收者155和用户110无法彼此听到。例如,响应于接收“OK计算机,几点了”的话语,启用语音的装置125可以就在“OK计算机”之后、“几点了”之前发起双向等候,使得X商店处的接收者155只听到“OK计算机”。

[0079] 一旦启用语音的装置125确定来自用户的命令已经被解决,则启用语音的装置125可以结束双向等候。例如,启用语音的装置125可以确定“下午9点”的答复回答用户的问题“几点了”,并且作为响应,结束双向等候。在另一个示例中,响应于用户110说“OK计算机,设定下午七点的闹钟”,启用语音的装置125可以答复“您想要在哪天设定下午七点的闹钟”并且继续双向等候用户110提供日期。在其他实施例中,用户110可以请求启用语音的装置125将呼叫置于等候,例如通过陈述“OK计算机,将呼叫置于等候”。启用语音的装置125可以继续等候呼叫直到用户请求结束等候,例如通过陈述“OK计算机,恢复呼叫”。

[0080] 在一些实现方式中,启用语音的装置125可以阻止与用户110进行长时间交互的命令。例如,启用语音的装置125可以阻止关于如下的特征:播放诸如音乐、新闻、博客的媒体,播放每日简报,第三方对话动作,进行附加的电话呼叫,以及玩诸如益智问答(trivia)的游戏。启用语音的装置125可以在阻止这些特征时提供错误,例如输出“抱歉,呼叫期间音乐无法播放”,或者忽略与这些任务之一相关联的任何命令并继续电话呼叫。

[0081] 在阶段(C)期间,启用语音的装置125在与X商店处的接收者155电话呼叫期间从来自用户110的另一个命令中检测热词。例如,启用语音的装置125获得来自用户110的命令,其陈述“OK计算机,挂断”。响应于电话呼叫期间所接收的话语,启用语音的装置125对用户110响应“呼叫结束”或非语言声音提示。附加地,启用语音的装置125不将响应“呼叫结束”或非语言声音提示发送到X商店处的接收者155。图5是图示与拨打呼叫的启用语音的装置交互的系统500的示例的框图。系统500包含启用语音的装置125、辅助服务器502、联系人数据库504、语音服务器506、客户端装置510、网络508以及通信链路512和514。

[0082] 在一些实现方式中,启用语音的装置125可以包含一个或多个计算机,并且可以包含在多个地理位置上分布的计算机。启用语音的装置125与一个或多个客户端装置510、辅助服务器502和语音服务器506通信。

[0083] 在一些实现方式中,辅助服务器502和语音服务器506可以各自包含一个或多个计算机,并且可以包含在多个地理位置上分布的计算机。辅助服务器502与启用语音的装置125和联系人数据库504通信。语音服务器506与启用语音的装置125和诸如X商店的一个或多个接收者通信。

[0084] 客户端装置510可以是例如桌上型计算机、膝上型计算机、平板计算机、可穿戴计算机、蜂窝电话、智能电话、音乐播放器、电子书阅读器、导航系统或任何其他合适的计算设备。网络508可以是有线的或无线的或者两者的组合，并且可以包含互联网。

[0085] 在一些实现方式中，启用语音的装置125可以使用诸如蓝牙、WiFi的短程通信协议或其他短程通信协议通过通信链路512连接到客户端装置510。例如，启用语音的装置125可以配对或连接多达7个不同的客户端装置510，每个客户端装置510具有相关联的通信链路512。在一些实现方式中，启用语音的装置125可以在任何给定的时间路由来自客户端装置510之一的音频。

[0086] 在一些实现方式中，启用语音的装置125可以从用户110接收话语“OK计算机，呼叫X商店”120。启用语音的装置125还可以将说话者(用户110)分类为已知的说话者“Matt”。例如，启用语音的装置125可以将包含在与“Matt”相关联的用户账户信息中的说话者识别特征与所接收的由用户110说出的热词进行比较。启用语音的装置125可以响应于比较来确定用户110是“Matt”。在一些实现方式中，启用语音的装置125可以然后将作为询问的话语的音频表示发送到辅助服务器502以进一步处理。

[0087] 在一些实现方式中，当用户110请求拨打呼叫时，启用语音的装置125可以停止各种事件。例如，一旦用户说“OK计算机，呼叫X商店”，启用语音的装置125可以停止播放音乐或闹钟。为了在用户110请求拨打呼叫时停止各种事件，启用语音的装置125可以存储在用户请求拨打电话时应当停止的事件的特定类型，并且响应于检测到用户正在拨打呼叫，结束那些所存储的特定类型的事件。例如，启用语音的装置125可以存储在用户拨打呼叫时要停止的播放音乐和闹钟的事件，并且响应于检测到用户正在拨打呼叫，结束任何播放音乐和闹钟的事件但是继续其他事件。

[0088] 在一些实现方式中，启用语音的装置125可以请求用户110在拨打呼叫前禁用任何事件。例如，启用语音的装置125可能当前正在播放音乐或者由于闹钟或定时器引起的铃声。启用语音的装置125可以不允许用户110做出任何呼叫直到用户110解除音乐或者由于闹钟或定时器引起的铃声。在一些实现方式中，用户110可以分别通过说“OK计算机，关闭音乐”或“OK计算机，关闭闹钟”来禁用音乐或者由于警报或计时器引起的铃声。在其他实现方式中，用户110可以通过点击启用语音的装置125上的交互按钮来禁用音乐或者由于警报或计时器引起的铃声。例如，启用语音的装置125可以存储需要用户交互的特定事件，以在用户拨打呼叫请求时禁用。响应于检测到用户拨打呼叫的请求和至少一个特定事件正在发生，启用语音的装置125可以向用户陈述说“请在呼叫前禁用事件”的警告消息，并且忽略拨打呼叫的请求。一旦用户通过发送语音命令到启用语音的装置125或者点击启用语音的装置125上的交互按钮来命令启用语音的装置125禁用特定事件，那么用户可以请求启用语音的装置125拨打呼叫。

[0089] 在一些实现方式中，响应于接收来自用户110的拨打呼叫的命令，启用语音的装置125可以警告用户110即将发生的闹钟。例如，用户110可能设定要在下午6:30时在启用语音的装置125上响铃的闹钟。用户110可能在下午6:29对启用语音的装置125说出话语“OK计算机，呼叫X商店”。响应于接收话语，启用语音的装置125可以向用户输出说“请在拨打电话呼叫前禁用闹钟”或者“设定为下午6:30的闹钟还有一分钟，您想要在我拨打呼叫前禁用该闹钟吗？”随后，用户110可以在用启用语音的装置125拨打电话呼叫前禁用闹钟或让闹钟经

过。

[0090] 在一些实现方式中,启用语音的装置125可以基于确定闹钟是否设定要在拨打电话呼叫的预定时间长度(例如,一分钟、五分钟、十五分钟或其他一些时间长度)内响起,警告用户110即将发生的闹钟。例如,启用语音的装置125可以在下午6:29接收要拨打呼叫的请求,确定下午6:29在五分钟内设定了下午6:30的闹钟,并且响应于确定下午6:29在五分钟内设定了闹钟,向用户110提供即将发生的闹钟的警告。

[0091] 在一些实现方式中,辅助服务器502获得请求516。例如,启用语音的装置125可以发送数据,该数据包含指示从用户110接收的话语的音频表示的搜索请求。数据可以指示所识别的已知说话者“Matt”,话语的音频表示“OK计算机,呼叫X商店”120,与启用语音的装置125相关联的唯一ID,以及与所识别的已知说话者“Matt”相关联的个人结果位。与启用语音的装置125相关联的唯一ID指示辅助服务器502向哪里发送响应。例如,唯一ID可以是与启用语音的装置125相关联的IP地址、URL或MAC地址。

[0092] 在一些实现方式中,辅助服务器502处理所获得的请求516。具体而言,辅助服务器502解析所获得的请求516,以确定与话语相关联的命令。例如,辅助服务器502可以通过将话语的音频表示转换为话语的文本表示来处理所获得的请求516。响应于转换,辅助服务器502解析热词后面的命令的文本表示“呼叫X商店”。在一些实现方式中,辅助服务器502确定与文本命令相关联的动作。例如,辅助服务器502通过将文本动作“呼叫”与存储文本动作进行比较来确定来自所获得的请求516的动作是“呼叫X商店”。

[0093] 附加地,辅助服务器502通过访问联系人数据库504解决接收者“X商店”的号码。在一些实现方式中,辅助服务器502访问联系人数据库504以取回与已知用户相关联的联系人。联系人数据库504通过由与联系人相关联的已知用户姓名来索引联系人而存储联系人。例如,联系人数据库504包含“Matt”的条目,其进一步包含与“Matt”相关联的个人联系人。个人联系人包含姓名和相关联的号码,诸如“妈妈”-(555) 111-1111、“爸爸”-(555) 222-2222、和“X商店”-(555) 333-3333。

[0094] 附加地,当在所获得的请求516中接收到的个人结果位被启用时,辅助服务器502可以仅解决接收者的号码。如果个人结果位未被启用、或为“0”,则辅助服务器502在动作消息518中发送标识符以指示启用语音的装置125对用户110中继陈述“请允许计算机访问个人联系人”的消息。如果个人结果位被启用、或为“1”,则辅助服务器502访问所识别的已知说话者的个人联系人的联系人数据库504。在一些实现方式中,辅助服务器502在所识别的已知说话者的个人联系人中取回与接收者相关联的号码。在该示例中,辅助服务器502取回X商店的号码(555) 333-3333。在其他实现方式中,接收者的号码可以包含在热词之后的命令的文本表示中。例如,命令可以包含“OK计算机,呼叫555-333-3333”。

[0095] 在一些实现方式中,辅助服务器502可以在联系人数据库504中识别所获得的请求516中未在所识别的已知说话者的个人联系人中找到的接收者。例如,辅助服务器502可以确定来自所获得的请求516的热词之后的命令的文本表示包含“呼叫奶奶”。然而,来自与“Matt”相关联的联系人数据库504中的个人联系人不包含“奶奶”的条目。反之,联系人包含“妈妈”、“爸爸”和“X商店”。为了解决接收者“奶奶”的号码,辅助服务器502可以搜索其他数据库和/或互联网以找到“奶奶”的号码。

[0096] 在搜索其他数据库和/或互联网时,辅助服务器502可以在知识图中搜索。例如,辅

助服务器502可能无法将“X公司客户服务”与用户的个人联系人中的任何记录相匹配,然后在知识图上搜索具有姓名“X公司客户服务”的实体,并且识别该实体在知识图中所存储的电话号码。

[0097] 在一些实现方式中,命令可以包含呼叫地理上接近启用语音的装置125的商业。辅助服务器502可以在互联网上搜索与启用语音的装置125的最接近商业相关联的语音号码。然而,如果辅助服务器502没有找到与请求的接收者相关联的号码,则辅助服务器502可以在动作消息518中发送标识符以指示启用语音的装置125对用户110中继陈述“没有找到联系人”的消息。例如,如果在个人联系人记录或知识图中未能找到“X商店”的电话号码,则辅助服务器502可以在地图数据库中搜索具有“X商店”的名称的附近的本地商业。

[0098] 在一些实现方式中,辅助服务器502可以确定包含在命令中的号码可以是不支持的语音号码。例如,号码可以仅包含7位数,诸如123-4567。作为响应,辅助服务器502可以在动作消息518中发送标识符,以指示启用语音的装置125对用户110中继陈述“电话号码不支持”的消息。

[0099] 响应于确定与接收者相关联的联系人号码,辅助服务器502生成去往启用语音的装置125的动作消息518。具体地,动作消息518可以包含联系人号码和触发呼叫的动作。例如,动作消息518可以包含555-333-3333的“X商店”的电话号码以及指令启用语音的装置125立即呼叫“X商店”的动作。在一些实现方式中,辅助服务器502可以基于命令的上下文在动作消息518中包含要使用的呼出号码。例如,如果命令包含呼叫紧急服务,则辅助服务器502可以在动作消息518中包含接收者155可以用于在特定时间段回叫启用语音的装置125的号码。例如,电话号码(555)888-8888可以用于未来几个小时拨打对启用语音的装置125的回叫。

[0100] 在一些实现方式中,启用语音的装置125从辅助服务器502获得动作消息518。响应于获得动作消息518,启用语音的装置125对动作消息518采取行动。例如,动作消息指示启用语音的装置125使用所指示的电话号码555-333-3333呼叫“X商店”。

[0101] 在一些实现方式中,启用语音的装置125可以基于用户110的偏好使用语音服务器506或相关联的客户端装置510来呼叫由辅助服务器502指定的接收者。具体地,用户110的偏好可以被存储在启用语音的装置125中。例如,启用语音的装置125可以确定用户110的偏好是使用语音服务器506或IP语音(VoIP),以进行任何呼出呼叫。如此,启用语音的装置125将指示发送到语音服务器506以呼叫接收者。在一些实现方式中,语音服务器506可以对呼出呼叫使用相关联的号码。在一些实现方式中,启用语音的装置125可以使得用户能够从多个不同VoIP提供者中选择使用VoIP提供者,并且然后在用户发起将来的呼叫时使用该VoIP提供者。

[0102] 在一些实现方式中,响应于确定用户110在启用语音的装置125附近,启用语音的装置125可以将与语音服务器506相关联的号码用于呼叫紧急服务。例如,响应于确定将客户端装置510之一连接到启用语音的装置125,启用语音的装置125可以使用与语音服务器506相关联的号码呼叫紧急服务。通过确保客户端装置510和启用语音的装置125之间的连接,启用语音的装置125可以确保用户110在启用语音的装置125附近。

[0103] 替代地,启用语音的装置125可以确定用户110的次要偏好为将现有的客户端装置510用于向接收者拨打呼出呼叫。如果启用语音的装置125确定用户110的次要偏好是使用

相关联的客户端装置510来呼叫接收者,则启用语音的装置125将验证去往客户端装置510的通信链路512。例如,启用语音的装置125可以验证去往客户端装置510的蓝牙连接。如果启用语音的装置125不能创建去往客户端装置510的蓝牙连接,则启用语音的装置125可以对用户110中继陈述“请确保您的蓝牙连接激活”的消息。一旦蓝牙连接被建立,启用语音的装置125发送指示到客户端装置510以呼叫接收者。在其他实施例中,如果启用语音的装置125不能通过任何短程通信协议发现客户端装置510,则启用语音的装置125可以使用去往接收者的语音服务器506以私人号码向接收者拨打电话呼叫。

[0104] 在一些实现方式中,响应于连接到接收者电话,启用语音的装置125可以播放可听闻的声音供用户听到。例如,如果接收者电话可用于应答,则启用语音的装置125可以播放可听闻的响铃音调。在另一个示例中,如果接收者电话不可用于应答,则启用语音的装置125可以播放忙音信号音调。在另一个示例中,如果接收者电话号码是无效的,则启用语音的装置125可以向用户提供语音消息,诸如“电话号码不支持”。在其他实施例中,在尝试连接对接收者电话的呼叫期间,用户110可以点击启用语音的装置125上的交互按钮以断开对接收者电话的呼叫。

[0105] 在一些实现方式中,启用语音的装置125可以重拨由用户110拨打的最近呼叫。例如,用户110可以说“OK计算机,重拨”而不用说号码,并且启用语音的装置125将重拨被呼叫的最后一个接收者号码。

[0106] 在一些实现方式中,为了启用语音的装置125重拨最近呼叫,启用语音的装置125在每次呼叫之后在存储器中存储与最近呼叫相关联的设定。存储器中与最近呼叫相关联的设定包含要拨打呼叫的用户、用于做出呼叫的号码、和接收者的号码。

[0107] 在一些实现方式中,启用语音的装置125可以接收双音调多频率(DTMF)音调以导航交互语音响应系统。例如,用户110可以说“OK计算机,按N”,其中N是*键、#键或者0和9之间的数字键。作为响应,启用语音的装置125可以在检测“OK计算机”之后拨打双向等候,生成发送到接收者155的数字N的拨号音,并且结束双向等候。

[0108] 在一些实现方式中,启用语音的装置125可以向用户110提供状态灯。例如,状态灯可以是指示启用语音的装置125的状态的LED灯。状态灯可以改变改变颜色、闪烁持续时间或亮度,以指示呼叫连接中、已连接的呼叫、呼叫结束、接收来自用户的语音命令,以及向用户110提供消息。

[0109] 在一些实现方式中,用户110可以用特定语音命令结束呼叫。例如,用户110可以说“OK计算机,停止呼叫”,“OK计算机,挂断”,或者“OK计算机,断开呼叫”。在一些实现方式中,接收者可以结束电话呼叫。在结束呼叫之后,启用语音的装置125可以播放可听闻的忙音,并且将启用语音的装置125返回到先前的状态然后连接电话呼叫。例如,将启用语音的装置125返回到先前的状态可以包含在当发起呼叫时媒体停止的时间点处,继续播放诸如歌曲的媒体。

[0110] 在一些实现方式中,启用语音的装置125可以指示何时接收到呼入呼叫。例如,启用语音的装置125可以闪烁LED、可听闻地输出响铃噪声、或可听闻地输出“呼入呼叫”,以指示启用语音的装置125正在接收呼叫。作为响应,用户110可以对于呼入呼叫采取动作。例如,用户110可以通过说出如下之一来应答呼叫:“OK计算机,接通”,“OK计算机,应答”,“OK计算机,接受”,或者“OK计算机,是”,作为几个示例。

[0111] 在另一个示例中,用户110可以说出如下之一来拒绝呼叫并且断开对连接的尝试:“OK计算机,不”,“OK计算机,拒绝”,或者“OK计算机,挂断”,作为几个示例。

[0112] 在一些实现方式中,启用语音的装置125可以仅通过临时号码接受呼入呼叫。具体地,启用语音的装置125可以仅当接收来自对临时号码的呼入呼叫时响铃,该临时号码用于拨打对紧急服务的向外呼叫。例如,启用语音的装置125可以使用号码(555) 555-5555作为呼出呼叫的临时号码以拨打紧急服务,并且可以仅接收对号码(555) 555-5555的呼入呼叫。

[0113] 在一些实现方式中,用户110可以将另一装置上的呼入呼叫转移到启用语音的装置125以用作说话者电话。用户110可以在呼叫响铃时或呼叫期间转移呼叫。例如,用户110可以说“OK计算机,将呼叫从我的电话转移到你”。在一些实现方式中,启用语音的装置125可以使用短程通信协议与其他装置通信以转移电话呼叫。例如,启用语音的装置125可以例如使用蓝牙或WiFi来连接到其他装置,以指令其他装置将当前电话呼叫路由到启用语音的装置125的扬声器。

[0114] 在一些实现方式中,用户110可以将呼叫从启用语音的装置125转移到客户端装置510。具体地,用户110可以在呼叫响铃时或呼叫期间转移呼叫。如果使用诸如蓝牙的短程通信协议中的至少一个将客户端装置510连接到启用语音的装置125,则可以进行转移呼叫。例如,用户110可以说“OK计算机,转移呼叫至我的电话”。附加地,用户110可以将呼叫从一个启用语音的装置125转移到位于分开的房间中的另一个启用语音的装置125。例如,用户110可以说“OK计算机,转移呼叫至卧室计算机”。如果客户端装置510或其他启用语音的装置125没有开机或连接到启用语音的装置125,则启用语音的装置125可以陈述“请将装置开机以建立连接”。

[0115] 图6是图示确定呼叫者号码的过程600的示例的流程图。可以由诸如系统500的一个或多个计算系统进行过程600的操作。

[0116] 过程600包含接收请求语音呼叫的话语(610)。例如,启用语音的装置125可以在用户说“OK计算机,呼叫(123) 456-7890”时接收话语,并且启用语音的装置125中的麦克风然后生成对应于话语的音频数据。在一些实现方式中,语音呼叫可以指代仅包含音频的呼叫。在其他实现方式中,语音呼叫可以指代不仅包含音频的呼叫,例如包含音频和视频的视频会议呼叫。

[0117] 过程600包含将话语分类为由特定已知用户所说(620)。例如,启用语音的装置125可以将话语“OK计算机,呼叫(123) 456-7890”分类为由特定已知用户“Matt”所说。在另一个示例中,启用语音的装置125可以将话语“OK计算机,呼叫(123) 456-7890”分类为由对启用语音的装置未知的用户所说。

[0118] 将话语分类为由特定已知用户所说可以包含确定话语中的语音是否与对应于特定已知用户的语音相匹配。例如,如先前所描述的,启用语音的装置125可以存储对应于先前说出热词“OK计算机”的已知用户“Matt”的MFCC,从刚收到的话语中的热词“OK计算机”中确定MFCC,然后确定来自话语的MFCC与为已知用户“Matt”所存储的MFCC相匹配,并且作为响应,将话语分类为已知用户“Matt”所说。在另一个示例中,启用语音的装置125可以存储对应于先前说出热词“OK计算机”的已知用户“Matt”的MFCC,从刚收到的话语中的热词“OK计算机”中确定MFCC,然后确定来自话语的MFCC与为已知用户“Matt”所存储的MFCC不匹配,并且作为响应,不将话语分类为已知用户“Matt”所说。

[0119] 将话语分类为由特定已知用户所说可以包含确定说话者的至少一部分的视觉图像是否与对应于特定已知用户的视觉信息相匹配。例如,如先前在上文所描述的,启用语音的装置125可以包含相机,获得由相机捕获的说话者的面部的图像,确定图像中说话者的面部与描述已知用户“Matt”的面部的信息相匹配,并且响应于该确定,将说话者分类为已知用户“Matt”。在另一个示例中,启用语音的装置125可以包含相机,获得由相机捕获的说话者的面部的图像,确定图像中说话者的面部与描述已知用户“Matt”的面部的信息不匹配,并且响应于该确定,将说话者分类为不是已知用户“Matt”。在一些实现方式中,视觉图像和语音可以组合地考虑,以将话语分类为是否由特定已知用户所说。

[0120] 过程600包含确定特定已知用户是否与个人语音号码相关联(630)。例如,启用语音的装置125可以确定已知用户“Matt”与(555) 222-2222的个人电话号码相关联。在另一个示例中,启用语音的装置125可以确定特定已知用户“爸爸”不与个人号码相关联。

[0121] 确定特定已知用户是否与个人语音号码相关联可以包含,访问特定已知用户的账户信息,以及确定用户的账户信息是否存储特定已知用户的语音号码。例如,启用语音的装置125可以访问已知用户“Matt”在启用语音的装置125上存储的账户信息,确定该账户信息包含(555) 222-2222的个人电话号码,并且作为响应,确定已知用户“Matt”与个人号码相关联。在另一个示例中,启用语音的装置125可以访问已知用户“爸爸”在启用语音的装置125上存储的账户信息,确定该账户信息不包含个人电话号码,并且作为响应,确定已知用户“爸爸”不与个人号码相关联。

[0122] 附加地或替代地,确定特定已知用户是否与个人语音号码相关联可以包含,向服务器提供特定已知用户的指示和话语的表示,以及从服务器接收特定已知用户的个人语音号码、要呼叫的语音号码和要拨出语音呼叫的指令。例如,在一些实现方式中,启用语音的装置125可以不存储个人电话号码,而辅助服务器502可以存储个人电话号码。因此,启用语音的装置125可以为辅助服务器502提供话语“OK计算机,呼叫(123) 456-7890”的音频表示以及说话者是已知用户“Matt”的指示。然后,辅助服务器502可以转录话语,从转录中的“呼叫”确定话语是在请求发起呼叫,从转录中确定“(123) 456-7890”是要呼叫的号码,响应于确定话语是在请求呼叫,访问已知用户“Matt”的所存储的帐户信息,确定已知用户“Matt”的存储的帐户包含个人语音号码(555) 222-2222,并且作为响应,向启用语音的装置125提供指令以拨打对号码(123) 456-7890的呼叫,显示(555) 222-2222作为发起呼叫的电话号码。

[0123] 确定特定已知用户是否与个人语音号码相关联可以包含,访问特定已知用户的账户,确定用户的账户是否指示电话,以及确定电话连接启用语音的装置。例如,在启用语音的装置125将话语分类为已经由已知用户“Matt”所说之后,启用语音的装置125可以访问存储的帐户信息以确定特定电话是否被指示为关联已知用户“Matt”,响应于确定帐户指示特定电话,确定特定电话是否已连接(例如通过 Bluetooth®),并且响应于确定特定电话已连接,通过特定电话发起电话呼叫。

[0124] 过程600包含用个人语音号码发起语音呼叫(640)。例如,启用语音的装置125可以向语音服务器506提供指令以用个人号码“(555) 222-2222”发起对“(123) 456-7890”的呼叫。在一些实现方式中,用个人语音号码发起电话呼叫可以包含,通过VoIP呼叫提供者发起电话呼叫。例如,语音服务器506可以是VoIP提供者并且启用语音的装置125可以请求语音

服务器506发起呼叫。在另一个示例中,启用语音的装置125可以向确定已连接启用语音的装置提供要与已知用户“Matt”相关联的电话拨打呼叫的指令。

[0125] 图7是图示确定要呼叫的接收者号码的过程示例的流程图。可以由诸如系统500的一个或多个计算系统进行过程600的操作。

[0126] 过程700包含接收请求语音呼叫的话语(710)。例如,辅助服务器502可以从启用语音的装置125接收“呼叫奶奶”的表示以及话语由启用语音的装置125确定为已经由已知用户“Matt”所说的指示。指示可以包含字母数字值或者与字母数字值相关联的二进制值,字母数字值从其他用户的账户中唯一识别Matt的账户,与字母数字值相关联的二进制值指示话语的说话者与由字母数字值识别的账户相关联。

[0127] 过程包含将话语分类为由特定已知用户所说(720)。例如,辅助服务器502可以将话语分类为如已经由已知用户“Matt”所说。将话语分类为由特定已知用户所说可以包含获得指示:话语中的语音由启用语音的装置确定为与对应于特定已知用户的语音相匹配。例如,辅助服务器502可以确定启用语音的装置125已提供值“854978”,该值将已知用户“Matt”的账户唯一识别为与话语“呼叫奶奶”的说话者相匹配,并且作为响应,将话语分类为如已经由已知用户“Matt”所说。

[0128] 附加地或替代地,将话语分类为由特定已知用户所说可以包含,确定话语中的语音是否与对应于特定已知用户的语音相匹配。例如,辅助服务器502可以从话语的音频表示生成MFCC,确定来自话语的MFCC是否匹配已知用户“Matt”的存储的MFCC,并且响应于MFCC匹配,将话语分类为已经由已知用户“Matt”所说。

[0129] 过程700包含响应于将话语分类为由特定已知用户所说,基于特定已知用户的联系人确定要呼叫的接收者语音号码(730)。例如,响应于将“呼叫奶奶”分类为由已知用户“Matt”所说,辅助服务器502可以基于已知用户“Matt”所存储的电话联系人确定要呼叫的接收者号码为“(987)654-3210”。在另一个示例中,响应于将“呼叫奶奶”分类为由已知用户“爸爸”所说,辅助服务器502可以基于已知用户“爸爸”所存储的电话联系人确定要呼叫的接收者号码为“(876)543-2109”。

[0130] 获得由特定已知用户创建的联系人条目可以包含,响应于将话语分类为由特定已知用户所说,确定特定已知用户的联系人条目是可用的,并且响应于确定特定已知用户的联系人条目是可用的,获得由特定已知用户创建的联系人条目。例如,响应于将话语分类为由已知用户“Matt”所说,辅助服务器502可以确定已知用户“Matt”的电话联系人条目,并且作为响应,访问已知用户“Matt”的电话联系人条目。

[0131] 确定特定已知用户的联系人条目是可用的可以包含,确定特定已知用户先前是否指示特定已知用户想要个性化的结果。例如,辅助服务器502可以从启用语音的装置125与话语一起接收个性化结果位,确定个性化结果位设定为指示已知用户“Matt”想要个性化结果的值,并且作为响应,确定已知用户“Matt”的电话联系人条目是可用的。在另一个示例中,辅助服务器502可以从启用语音的装置125与话语一起接收个性化结果位,确定个性化结果位设定为指示已知用户“爸爸”不想要个性化结果的值,并且作为响应,确定已知用户“爸爸”的电话联系人条目不是可用的。

[0132] 响应于将话语分类为由特定已知用户所说,基于特定已知用户的联系人确定要呼叫的接收者语音号码可以包含:响应于将话语分类为由特定已知用户所说,获得由特定已

知用户创建的联系人条目,从联系人条目之中识别包含与话语匹配的姓名的特定联系人条目,并且确定由特定联系人条目指示的语音号码作为接收者语音号码。例如,响应于将话语“呼叫奶奶”分类为由已知用户“Matt”所说,辅助服务器502可以获得由已知用户“Matt”创建的电话联系人条目,识别电话联系人条目之一被命名为与话语“奶奶”相匹配的“奶奶”,且具有号码“(987)654-3210”,并且确定接收者电话号码是号码“(987)654-3210”。

[0133] 从联系人条目之中包含与话语匹配的姓名的特定联系人条目可以包含,生成话语的转录并确定该转录包含姓名。例如,辅助服务器502可以生成话语“呼叫奶奶”的转录,确定来自转录的“奶奶”与确定已知用户“Matt”的电话联系人条目的“奶奶”的姓名相一致,并且作为响应,识别命名为“奶奶”的联系人条目。

[0134] 过程700包含向接收者语音号码发起语音呼叫(740)。例如,辅助服务器502可以对从命名为“奶奶”的已知用户的电话联系人条目中获得的接收者电话号码“(987)654-3210”发起呼叫。对接收者语音号码发起语音呼叫可以包含,向启用语音的装置提供接收者语音号码以及要向接收者语音号码发起语音呼叫的指示。例如,辅助服务器502可以向启用语音的装置125提供指令,以用号码(555)222-2222发起对号码(987)654-3210的呼叫。

[0135] 在一些实现方式中,过程700可以包含,接收请求第二语音呼叫的第二话语,将第二话语分类为不是由启用语音的装置125的任何已知用户所说,并且响应于将第二话语分类为不是由启用语音的装置的任何已知用户所说,在不访问启用语音的装置的任何已知用户的联系人的情况下发起第二语音呼叫。例如,辅助服务器502可以接收“呼叫X商店”的第二话语,将第二话语分类为不是由启用语音的装置125的任何已知用户所说并确定话语中的“X商店”不是电话号码,并且响应于将第二话语分类为不是由启用语音的装置的任何已知用户所说以及话语中的“X商店”不是电话号码,在地图数据库上搜索具有“X商店”名称的附近本地商业,识别出一个具有“X商店”名称和“(765)432-1098”的电话号码的附近本地商业,并且在不访问启用语音的装置的任何已知用户的电话联系人发起对(765)432-1098的第二电话呼叫。

[0136] 图8是图示语音呼叫期间处理询问的过程示例的流程图。可以由诸如系统500的一个或多个计算系统进行过程800的操作。

[0137] 过程800包含确定在第一方和第三方之间的语音呼叫期间第一方已经对启用语音的虚拟助手说出询问(810)。例如,启用语音的装置125可以确定在用户和其他人之间的电话呼叫期间用户已经对辅助服务器502说出询问。确定在第一方和第三方之间的电话呼叫期间第一方已经对启用语音的虚拟助手说出询问可以包含,由启用语音的装置确定在电话呼叫期间由第一方说出热词。例如,启用语音的装置125可以确定在通过启用语音的装置125进行呼叫时已经说出热词“OK计算机”。当启用语音的装置125的麦克风和扬声器正在用于为其他人接听来自用户的语音并输出其他人的语音到用户时,呼叫可以被认为通过启用语音的装置125正在进行。

[0138] 过程800包含,响应于确定在第一方和第三方之间的电话呼叫期间第一方已经对启用语音的虚拟助手说出询问,将第一方和第三方之间的语音呼叫置于等候(810)。例如,响应于确定在第一方和第三方之间的电话呼叫期间第一方已经对启用语音的虚拟助手说出询问“OK计算机,我的下一个预约是什么?”,启用语音的装置125可以将电话呼叫置于双向等候。语音呼叫可以被置于双向等候,使得其他人无法听到来自用户对启用语音的虚拟

助手的询问,并且不能听到来自启用语音的虚拟助手对询问的响应。

[0139] 过程800包含将语音呼叫置于等候(820)。例如,启用语音的装置125可以将电话呼叫置于双向等候。将第一方和第二方之间的语音呼叫置于等候可以包含,向语音呼叫提供者提供指令以将语音呼叫置于等候。例如,启用语音的装置125可以指令语音服务器506以将正在进行的呼叫置于等候。附加地或替代地,将第一方和第二方的语音呼叫置于等候可以包含,将音频从麦克风路由到启用语音的虚拟助手而不是语音服务器,并向扬声器路由来自启用语音的虚拟助手的音频,而不是来自语音服务器的音频。例如,启用语音的装置125可以将音频从启用语音的装置125中的麦克风路由到辅助服务器502而不是语音服务器506,并且向启用语音的装置125的扬声器路由来自辅助服务器502的音频,而不是来自语音服务器506的音频。

[0140] 过程800包含确定启用语音的虚拟助手已经解决询问(830)。例如,启用语音的装置125可以确定辅助服务器502已经解决询问“OK计算机,我的下一个预约是什么”。确定启用语音的虚拟助手已经解决询问包含,向启用语音的虚拟助手提供询问和在启用语音的装置上正在进行语音呼叫的指示,并且从启用语音的虚拟助手接收对询问的响应和询问被解决的指示。例如,启用语音的装置125提供询问“OK计算机,我的下一个预约是什么”的表示和“正在进行呼叫=真”的指示,并且作为响应,接收合成语音“您的下一个预约是下午3:30的‘茶歇’”的表示作为对询问的响应和“询问被解决=真”的指示。

[0141] 在一些实现方式中,启用语音的虚拟助手可以配置为识别对应于询问的命令,确定在语音呼叫期间可以执行命令,并且响应于确定可以在语音呼叫期间执行命令,确定响应以指示对命令的应答。例如,辅助服务器502可以接收话语“OK计算机,我的下一个预约是什么”的表示,从话语的表示中识别来自转录的“识别下一个预约”的命令,确定在电话呼叫期间可以执行命令“识别下一个预约”,并且响应于确定在电话呼叫期间可以执行命令,确定响应以指示“您的下一个预约是下午3:30的‘茶歇’”的应答。

[0142] 在一些实现方式中,启用语音的虚拟助手可以配置为识别对应于询问的命令,确定在语音呼叫期间无法执行该命令,并且响应于确定在语音呼叫期间无法执行该命令,确定响应以指示无法执行该命令。例如,辅助服务器502可以接收话语“OK计算机,播放一些音乐”的表示,从话语的表示中识别来自转录的“播放音乐”的命令,确定在电话呼叫期间无法执行该命令“播放音乐”,并且响应于确定在电话呼叫期间无法执行该命令,确定响应以指示“抱歉,我不能在呼叫期间播放音乐”的应答。

[0143] 在一些实现方式中,确定在语音呼叫期间无法执行该命令包含,获得在语音呼叫期间可以正常执行的命令列表,以及确定所识别的命令不在命令列表中。例如,辅助服务器502可以获得包含“识别下一个预约”而不包含“播放音乐”的可执行的命令列表,确定在列表中没有识别“播放音乐”,并且作为响应,确定在电话呼叫期间不能正常执行“播放音乐”的命令。

[0144] 在一些实现方式中,确定在语音呼叫期间无法执行该命令包含,获得在语音呼叫期间无法正常执行的命令列表,并且确定所识别的命令在命令列表中。例如,辅助服务器502可以获得包含“播放音乐”而不包含“识别下一个预约”的不能执行的命令列表,确定在列表中识别“播放音乐”,并且作为响应,确定在电话呼叫期间无法正常执行“播放音乐”的命令。

[0145] 过程800包含,响应于确定启用语音的虚拟助手已经处理询问,从等候中恢复第一方和第二方之间的语音呼叫(840)。例如,启用语音的装置125可以恢复电话呼叫。响应于确定启用语音的虚拟助手已经处理询问,从等候中恢复第一方和第二方之间的语音呼叫可以包含,向语音呼叫提供者提供指令以从等候中恢复语音呼叫。例如,启用语音的装置125可以向语音服务器506提供指令以从等候中恢复语音呼叫。

[0146] 附加地或替代地,响应于确定启用语音的虚拟助手已经处理询问,从等候中恢复第一方和第二方之间的语音呼叫包含,将来自麦克风的音频路由到语音服务器而不是启用语音的虚拟助手,并向扬声器路由来自语音服务器的音频,而不是来自启用语音的虚拟助手的音频。例如,启用语音的装置125可以将来自麦克风的音频路由到语音服务器506而不是辅助服务器502,并向扬声器路由来自语音服务器506的音频,而不是来自辅助服务器502的音频。

[0147] 在一些实现方式中,响应于确定启用语音的虚拟助手已经处理询问,从等候中恢复第一方和第二方之间的语音呼叫包含,从启用语音的虚拟助手接收指令以产生双音调多频率信号,并且响应于从启用语音的虚拟助手接收产生双音调多频率信号的指令,向语音呼叫提供者提供第二指令以在向语音呼叫提供者提供指令之后产生双音调多频率信号来从等候中恢复语音呼叫。例如,启用语音的装置125可以接收“生成1的DTMF”的指令,并且作为响应,指令语音服务器506已生成表示按“1”键的DTMF。

[0148] 在一些实现方式中,启用语音的辅助服务器配置为确定询问指示要生成一个或多个双音调多频率信号以及对应于一个或多个双音调多频率信号的一个或多个号码的命令。例如,辅助服务器502可以接收话语“OK计算机,按1”的表示,根据转录确定“按1”指示要生成转录中由“1”表示的的数字的DTMF信号,并且作为响应,向启用语音的装置125提供指令以指令启用语音的装置125生成“1”的DTMF。附加地或替代地,在一些实现方式中,启用语音的装置125可以生成DTMF。例如,启用语音的装置125可以从辅助服务器502接收指令以生成“1”的DTMF,并且作为响应,产生“1”的DTMF音调并发送这些音调到语音服务器506。

[0149] 除了如上描述,用户可以配备有控制,以允许用户关于以下做出选择:本文中所描述的系统、程序或特征是否可以并且何时可以启用对用户信息(例如关于用户社交网络、社交行为或活动、职业、用户的偏好、或者用户的当前位置的信息)的收集,以及用户是否从服务器发送内容或通信。附加地,某些数据可以在其被存储或使用前以一个或多个方式来处理,使得个人可识别信息被移除。例如,可以处理用户的身份,使得无法确定用户的个人可识别信息,或者在获得的位置信息的情况下,可以将用户的地理位置普遍化(诸如城市、邮政编码或州级别),使得不能确定用户的特定位置。因此,用户可以控制收集关于用户的什么信息、如何使用信息以及向用户提供什么信息。

[0150] 可以使用系统100的不同配置,其中启用语音的装置125、辅助服务器502、和语音服务器506的功能可以组合、进一步分离、分散或互换。例如,并非包含要转录的对辅助服务器502的询问中话语的音频表示,启用语音的装置125可以转录话语并且包含对辅助服务器502的询问中的转录。

[0151] 图9示出了计算装置900和移动计算装置950的示例,其可以用于实现本文所描述的技术。计算装置900旨在表现数字计算机的各种形式,诸如笔记本计算机、台式机、工作站、个人数字助理、服务器、叶片服务器、大型主机和其他合适的计算机。移动计算装置950

旨在表示移动装置的各种形式,诸如个人数字助理、蜂窝电话、智能电话和其他类似的计算装置。在此所示的组件、它们的连接和关系以及它们的功能仅为示例性的,并不意味限制。

[0152] 计算装置900包含处理器902、存储器904、存储装置906、连接到存储器904的高速接口908和多个高速扩展端口910、以及连接到低速扩展端口914和存储装置906的低速接口912。处理器902、存储器904、存储装置906、高速接口908、高速扩展端口910、以及低速接口912使用各种总线互连并且可以安装在公共的母板上或以其他适合的方式安装。处理器902可以处理在计算装置900内执行的指令,指令包含存储在存储器904中或在存储装置906上的指令,以在外部输入/输出装置(诸如耦接到高速接口908的显示器916)上显示图形用户界面(GUI)的图形信息。在其他实现方式中,可以连同多个存储器和多种存储器一起,适当地使用多个处理器和/或多个总线。同样,可以连接多个计算装置,每个装置提供必要操作的部分(例如,作为服务器库、一组叶片服务器或者多处理器系统)。

[0153] 存储器904在计算装置900内存储信息。在一些实现方式中,存储器904是(多个)易失性存储器单元。在一些实现方式中,存储器904是(多个)非易失性存储器单元。存储器904还可以是另一种形式的计算机可读介质,诸如磁盘或光盘。

[0154] 存储装置906能够为计算装置900提供大容量存储。在一些实现方式中,存储装置906可以是或者含有计算机可读介质(诸如软盘装置、硬盘装置、光盘装置或者磁带装置、闪存或者其他类似的固态存储器装置、或者装置阵列),包含存储区域网络中的装置或者其他配置。指令可以被存储在信息载体中。指令在由一个或多个处理装置(例如处理器902)执行时进行一个或多个方法,诸如上文所描述的那些方法。指令还可以由一个或多个存储装置,诸如计算机可读介质或机器可读介质(例如存储器904、存储装置906或处理器902上的存储器)。

[0155] 高速接口908管理计算装置900的带宽密集的操作,而低速接口912管理较不带宽密集的操作。功能的此类分配仅是示例性的。在一些实现方式中,高速接口908耦接到存储器904、显示器916(例如,通过图形处理器或者加速器)和高速扩展端口910,其可以接受各种扩展卡(未示出)。在实现方式中,低速接口912耦接到存储装置906和低速扩展端口914。低速扩展端口914(其可以包含各种通信端口(例如,USB、蓝牙、以太网、无线以太网))可以例如通过网络适配器耦接到一个或多个输入/输出装置,诸如键盘、指向装置、扫描仪或者诸如交换机或路由器的网络装置。

[0156] 如图所示,可以采用若干不同形式来实现计算装置900。例如,它可以实现为标准服务器920,或者此类服务器的多次组合。此外,它可以在诸如笔记本计算机922的个人计算机中实现。它还可以实现为机架服务器系统924的部分。替代地,来自计算装置900的组件可以与诸如移动计算装置950的移动装置(未示出)中的其他组件结合。此类装置中的每一个可以含有计算装置900、移动计算装置950中的一个或多个,并且整个系统可以由相互通信的多个计算装置构成。

[0157] 移动计算装置950包含处理器952、存储器964、诸如显示器954的输入/输出装置、通信接口966和收发器968以及其他组件。移动计算装置950还可以配备有存储装置(诸如微驱动器或者其他装置),以提供附加的存储。处理器952、存储器964、显示器954、通信接口966和收发器968中的每一个使用各种总线互连,并且可以在公共的母板上或以其他合适的方式安装若干组件。

[0158] 处理器952可以执行移动计算装置950内的指令,包含存储在存储器964中的指令。处理器952可以实现为芯片的芯片组,该芯片包含单独的和多个模拟和数字处理器。例如,可以提供处理器952,以用于移动计算装置950的其他组件(诸如用户界面的控制、由移动计算装置950运行的应用以及由移动计算装置950运行的无线通信)的协调。

[0159] 处理器952可以通过耦接到显示器954的控制接口958和显示接口956与用户通信。例如,显示器954可以是TFT LCD(薄膜晶体管液晶显示器)或OLED(有机发光二极管)显示器,或者其他适当的显示技术。显示接口956可以包括驱动显示器954的适当的电路,以向用户呈现图形和其他信息。控制接口958可以从用户接收命令,并且将其转换以提交给处理器952。附加地,可以提供外部接口962以与处理器952通信,以便使移动计算装置950可能与其他装置进行近域通信。例如,外部接口962可以在一些实现方式中提供有线通信,或者在其它实施方式中提供无线通信,并且还可以使用多个接口。

[0160] 存储器964在移动计算装置950内存储信息。存储器964可以实现为计算机可读介质或多个介质、易失性存储器单元或多个单元、或者非易失性存储器单元或多个单元中的一个或多个。扩展存储器974也可以通过扩展接口972来提供并连接到移动计算装置950,其中扩展接口可以包含例如SIMM(单列直插式存储器模块)卡接口。扩展存储器974可以为移动计算装置950提供额外的存储空间,或者还可以存储移动计算装置950的应用或者其他信息。具体而言,扩展存储器974可以包含要执行的或者补充如上所描述的过程的指令,并且还可以包含安全信息。因此,例如,可以提供扩展存储器974作为移动计算装置950的安全模块,并且可以用指令来编程,该指令许可移动计算装置950的安全使用。此外,可以经由SIMM卡提供安全应用连同附加信息,诸如以非可攻击的方式将识别信息放置在SIMM卡上。

[0161] 存储器可以包含例如闪存和/或NVRAM存储器(非易失性随机存取存储器),如下所描述。在一些实现方式中,指令存储在信息载体中,其中指令在由一个或多个处理装置(例如处理器952)执行时进行一个或多个方法,诸如上文所描述的那些方法。指令还可以由一个或多个存储装置,诸如一个或多个计算机可读介质或机器可读介质(例如存储器964、扩展存储器974或处理器952上的存储器)。在一些实现方式中,可以在传播的信号(例如在收发器968或外部接口962)中接收指令。

[0162] 移动计算装置950可以通过通信接口966进行无线通信,通信接口966可以在必要时包含数字信号处理电路。通信接口966可以提供各种模式或协议下的通信,诸如GSM语音呼叫(全球移动通信系统)、SMS(短消息服务)、EMS(增强消息服务)、或MMS消息(多媒体消息服务)、CDMA(码分多址)、TDMA(时分多址)、PDC(个人数字蜂窝)、WCDMA(宽带码分多址)、CDMA2000或GPRS(通用分组无线服务)等。此类通信可以例如通过使用射频的收发器968来发生。附加地,可以发生短程通信,诸如使用蓝牙、Wi-Fi或其他此类收发器(未示出)。附加地,GPS(全球定位系统)接收器模块970可以向移动计算装置950提供附加的导航和位置相关的无线数据,其可以由在移动计算装置950上运行的应用适当地使用。

[0163] 移动计算装置950还可以使用音频编解码器960来可听闻地通信,其可以接收由用户所说的信息并且将其转换为可用的数字信息。音频编解码器960同样可以为用户产生可听闻的声音,诸如通过例如在移动计算装置950的听筒中的扬声器。此类声音可以包含来自语音电话呼叫的声音,可以包含记录的声音(例如语音消息、音乐文件等),并且还可以包含在移动计算装置950上操作的应用所生成的声音。

[0164] 如图所示,可以采用若干不同形式来实现移动计算装置950。例如,它可以实现为蜂窝电话980。它还可以实现为智能电话982、个人数字助理或其他类似的移动装置的部分。

[0165] 本文中所描述的系统和技术各种实现方式可以实现为数字电子电路、集成电路、专门设计的ASIC、计算机硬件、固件、软件和/或其组合。这些各种实现方式可以包含在可编程系统上可执行和/或可解释的一个或多个计算机程序中的实现方式,该可编程系统包含至少一个可编程处理器,其可以是专用或通用的,耦接以从存储系统、至少一个输入设备和至少一个输出设备接收数据和指令,并且向存储系统、至少一个输入设备和至少一个输出设备传输数据和指令。

[0166] 这些计算机程序(还已知为程序、软件、软件应用或代码)包含可编程处理器的机器指令,并且可以实现为高级过程语言和/或面向对象的编程语言、和/或汇编/机器语言。程序可以被存储在保存其他程序或数据的部分文件中,例如在标记语言文档中存储的一个或多个脚本、在专用于讨论中的程序的单个文件中、或在多个协同文件中,例如存储一个或多个模块、子程序或部分代码的文件。计算机程序可以部署为在一个计算机上或者在多个计算机上执行,该多个计算机位于一个站点处或者分布跨越多个站点并由通信网络互连。

[0167] 如本文所使用的,术语“机器可读介质”、“计算机可读介质”指代任何计算机程序产品、设备和/或装置(例如磁盘、光盘、存储器、可编程逻辑器件(PLD)),其用于可编程处理器提供机器指令和/或数据,包含接收机器指令作为机器可读信号的机器可读介质。术语“机器可读信号”指代用于向可编程处理器提供机器指令和/或数据的任何信号。

[0168] 为了提供与用户的交互,在此所描述的系统和技术可以实现在具有用于向用户显示信息的显示装置(例如CRT(阴极射线管)或LCD(液晶显示器)监控器)以及键盘和指向装置(例如鼠标或轨迹球)的计算机上,用户可以通过该指向装置向计算机提供输入。其他类型的装置也可以用于提供与用户的交互;例如向用户所提供的反馈可以是任何形式的传感反馈,例如视觉反馈、听觉反馈或者触觉反馈;并且来自用户的输入可以用包含声音、语音或触觉输入的任何形式来接收。

[0169] 可以在计算系统中实现在此所描述的系统和技术,该计算系统包含后端组件(例如作为数据服务器),或者包含中间件组件(例如应用服务器),或者包含前端组件(例如具有图形用户界面或网络浏览器的客户端计算机,用户可以通过该图形界面或Web浏览器与在此所描述的系统和技术实现方式互),或者这样的后端、中间件或前端组件的任何组合交互。系统的组件可以通过例如通信网络的数字数据通信的任何形式或者介质互连。通信网络的示例包含局域网(“LAN”)、广域网(“WAN”)和互联网。

[0170] 计算系统可以包括客户端和服务端。客户端和服务端总体上彼此远离,并且典型地通过通信网络交互。客户端和服务端的关系借助于在相应的计算机上运行并彼此之间具有客户端-服务端关系的计算机程序而出现。

[0171] 除了如上描述,用户可以配备有控制,以允许用户关于以下做出选择:本文中所描述的系统、程序或特征是否可以并且何时可以启用对用户信息(例如关于用户社交网络、社交动作或活动、职业、用户的偏好、或者用户的当前位置的信息)的收集,以及用户是否从服务器发送内容或通信。附加地,某些数据可以在其被存储或使用前以一个或多个方式来处理,使得个人可识别的信息被移除。

[0172] 例如,在一些实施例中,可以处理用户的身份使得无法确定用户的个人可识别的

信息,或者在获得的位置信息的情况下可以将用户的地理位置普遍化(诸如城市、邮政编码或州级别),使得不能确定用户的特定位置。因此,用户可以控制收集关于用户的什么信息、如何使用信息以及向用户提供什么信息。

[0173] 已经描述了若干实施例。然而,应当理解的是,可以在不脱离本公开的精神和范围的情况下做出各种修改。例如,可以在重新排序、添加或移除步骤的情况下使用如上所示的各种形式的流程。同样,尽管已经描述了系统和方法的若干应用,但是应当理解的是许多其他应用是预期的。因此,其他实施例在所附权利要求的范围内。

[0174] 已经描述了主题的特定实施例。其他实施例在所附权利要求的范围内。例如,权利要求中所述的动作可以以不同的顺序进行,并且仍然实现期望的结果。作为一个示例,所附附图中所描绘的步骤不是必须按所示出的特定顺序或先后顺序,以实现期望的结果。在一些情况下,多任务和并行处理可以是有利的。

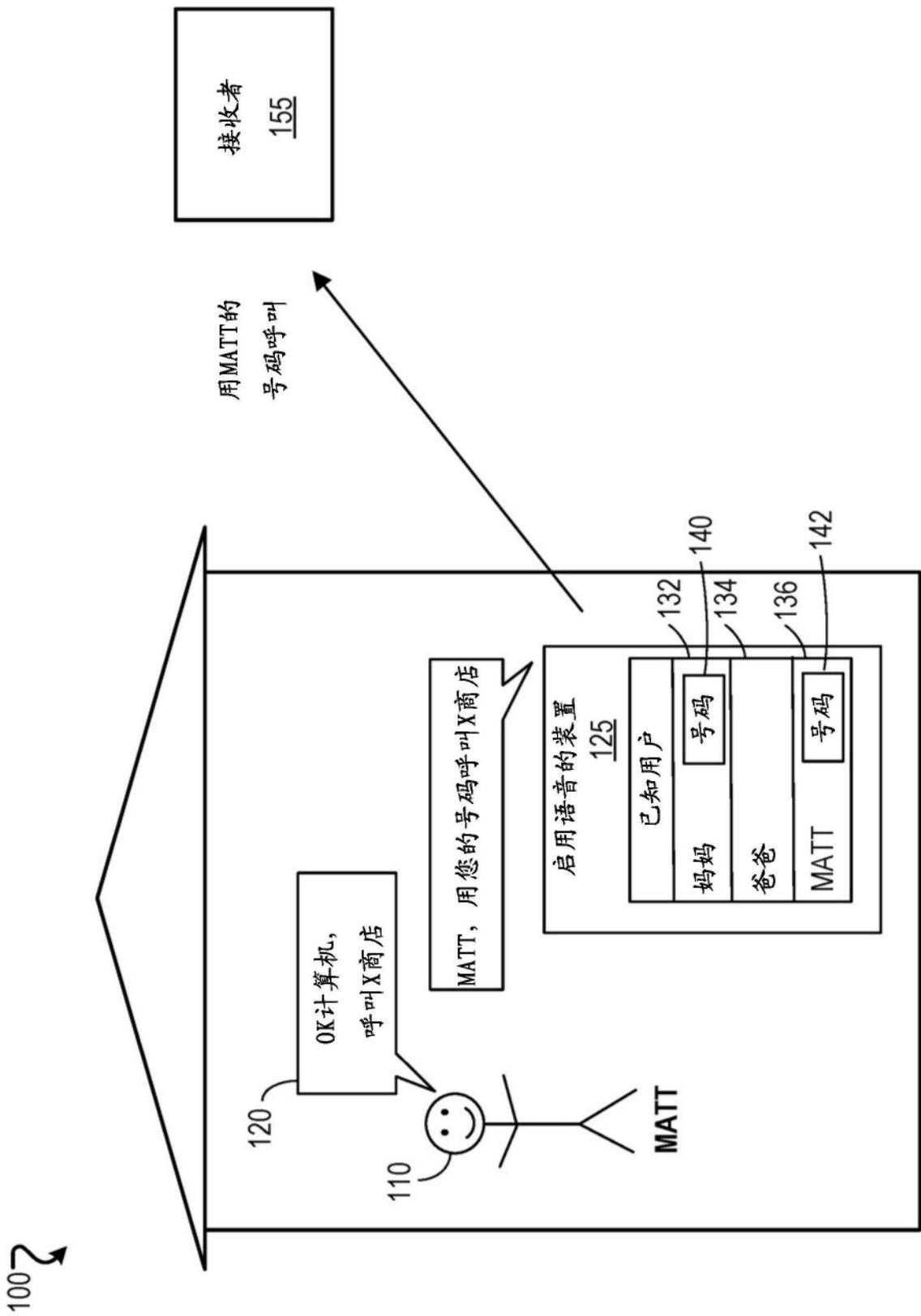


图1A

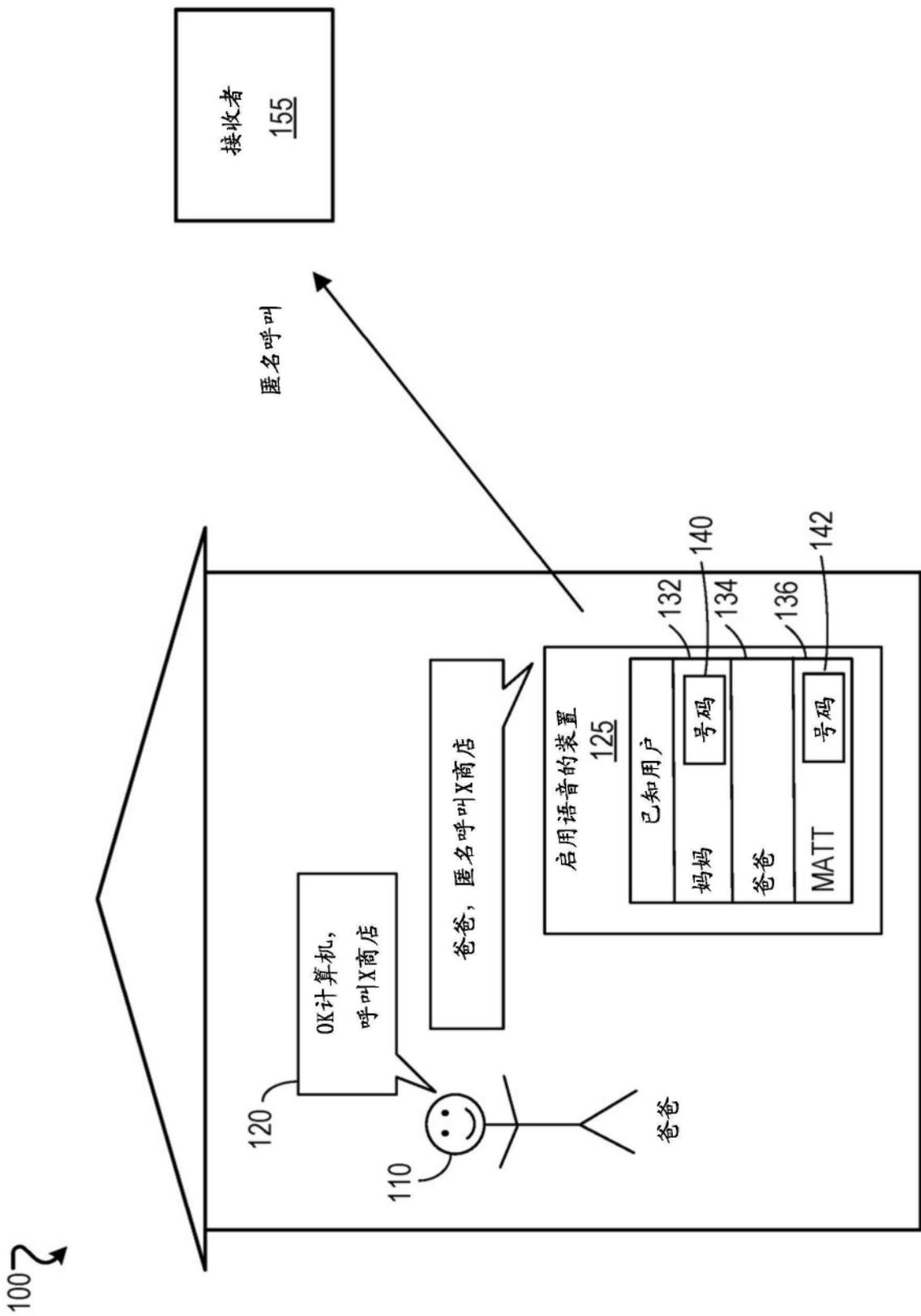


图1B

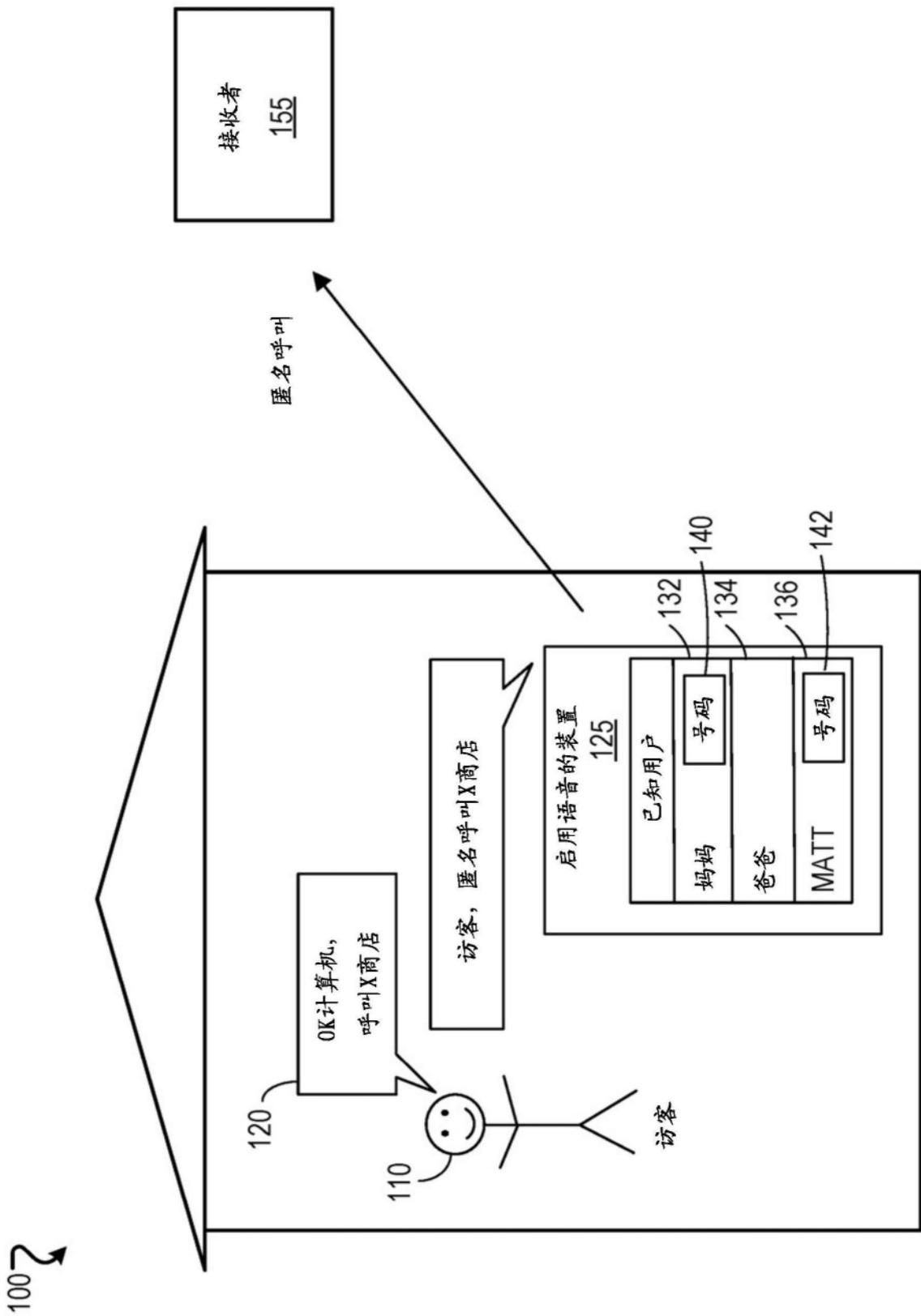


图1C

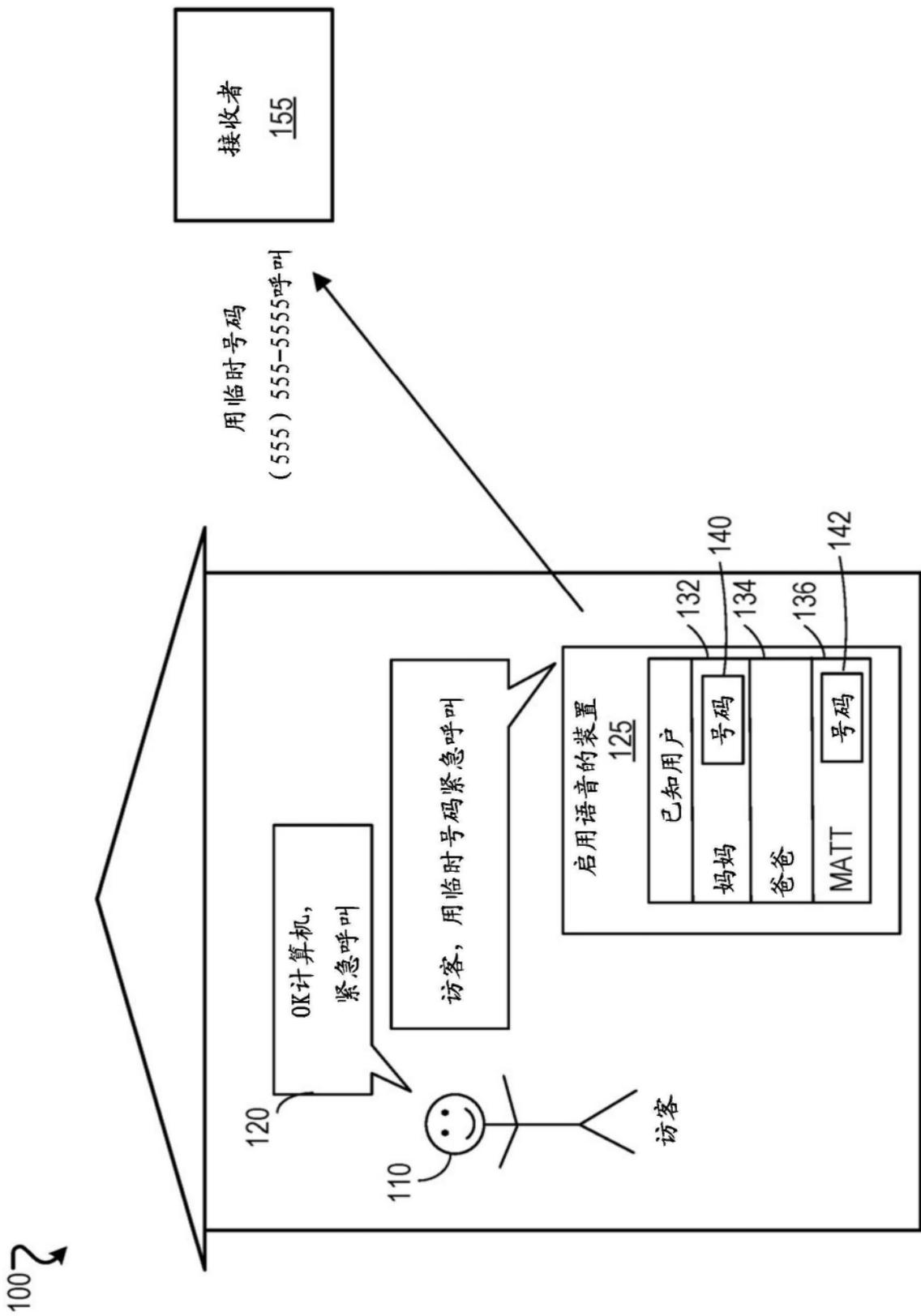


图1D

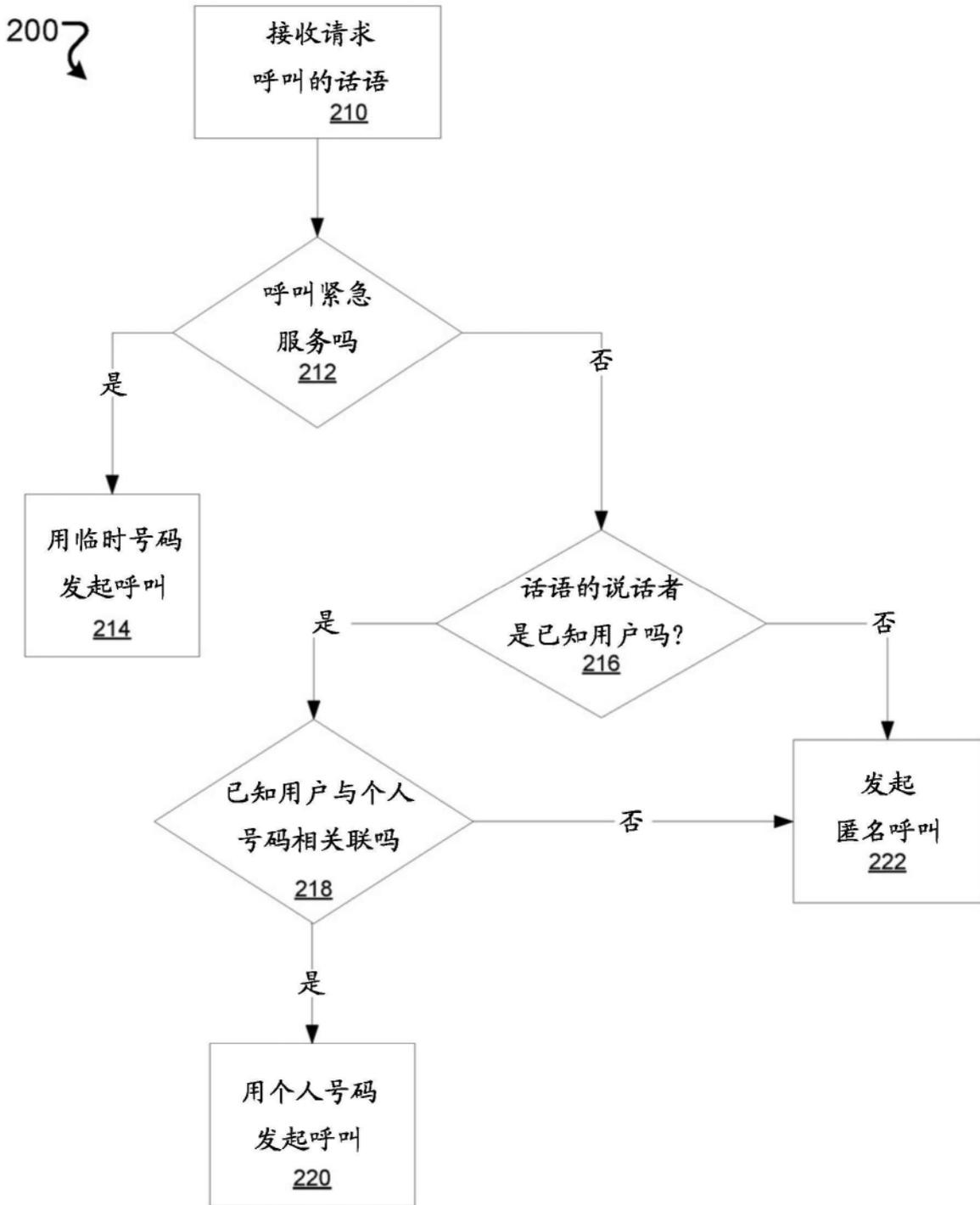


图2

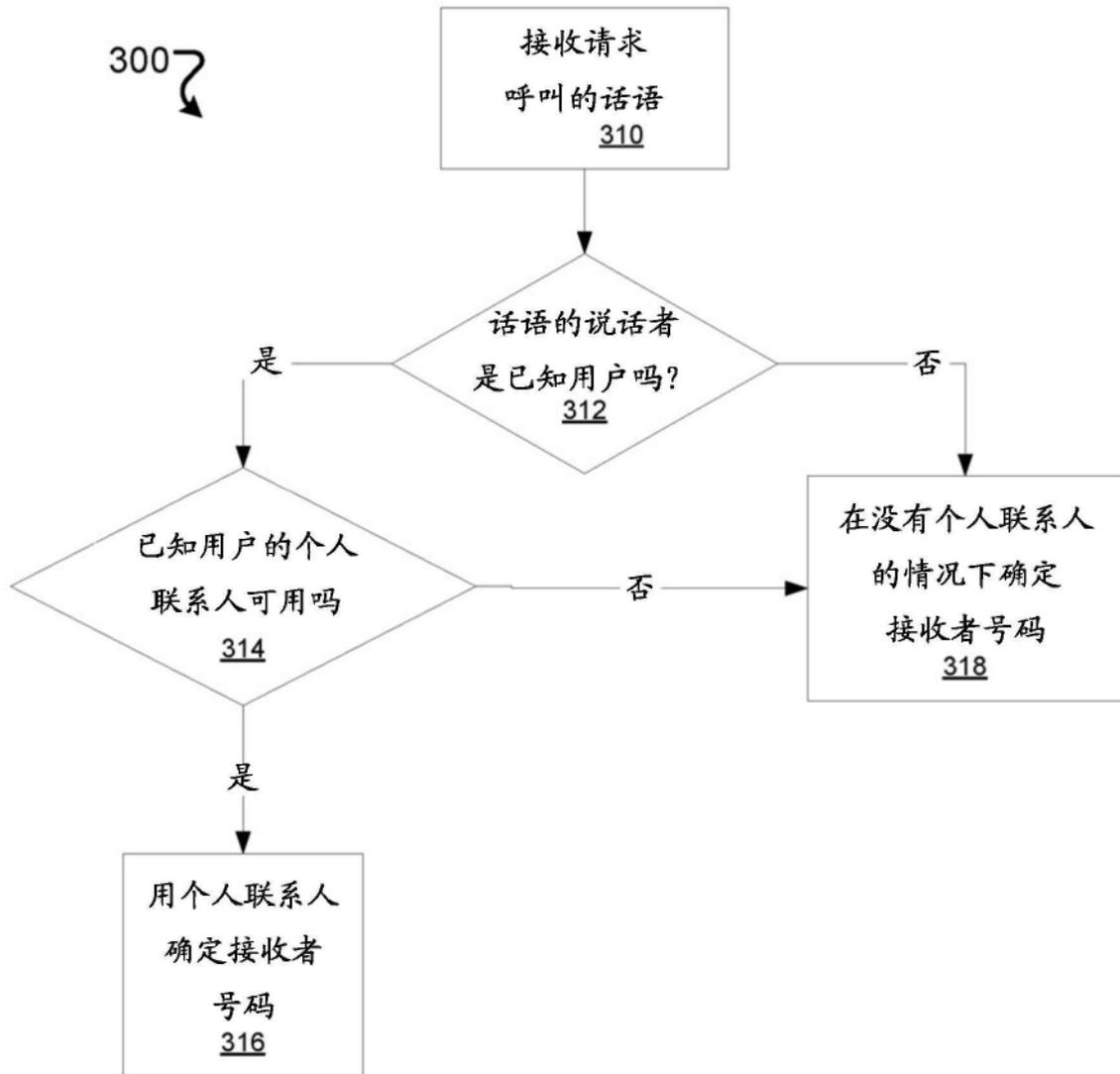


图3

400 ↷

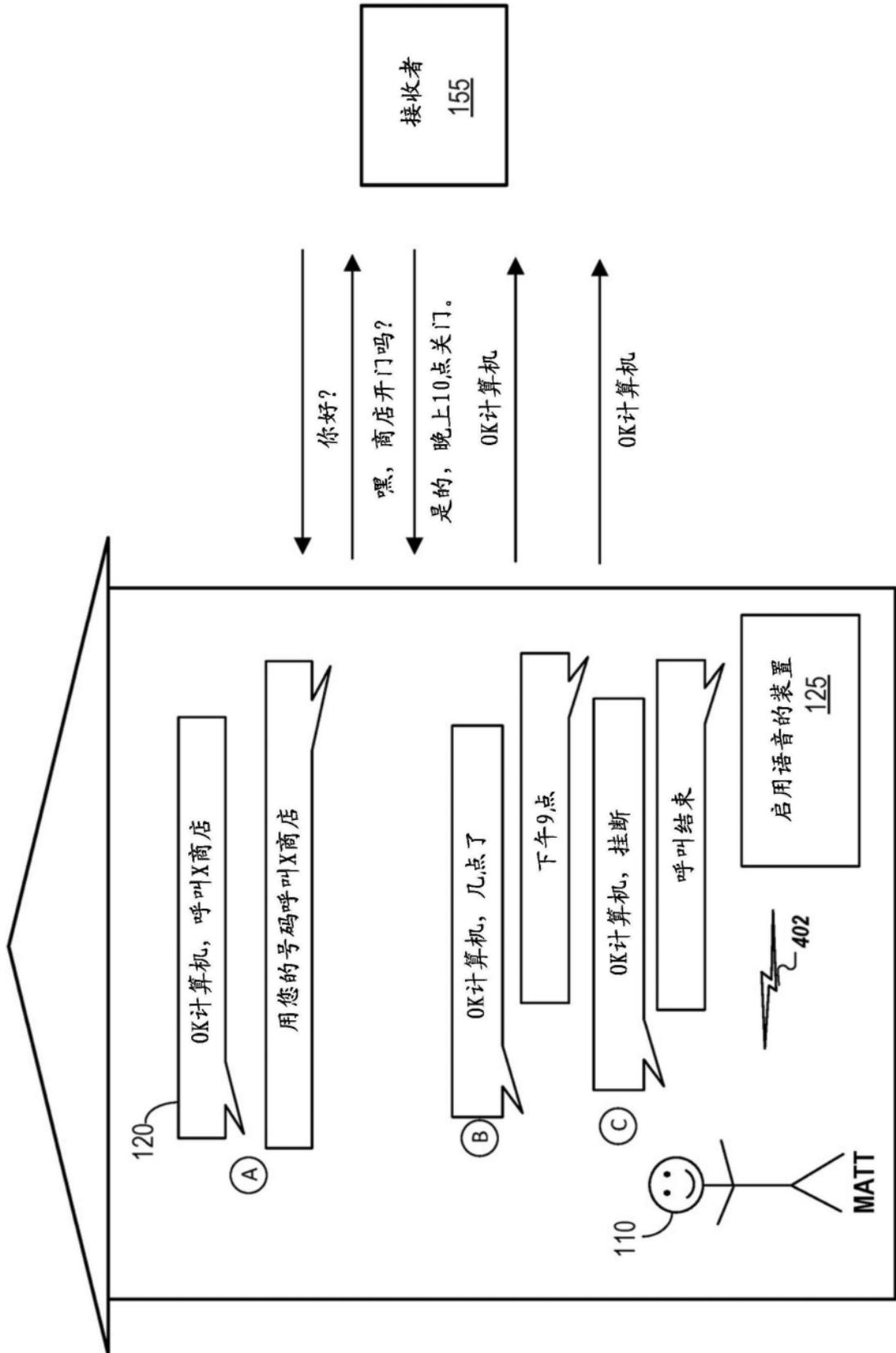


图4

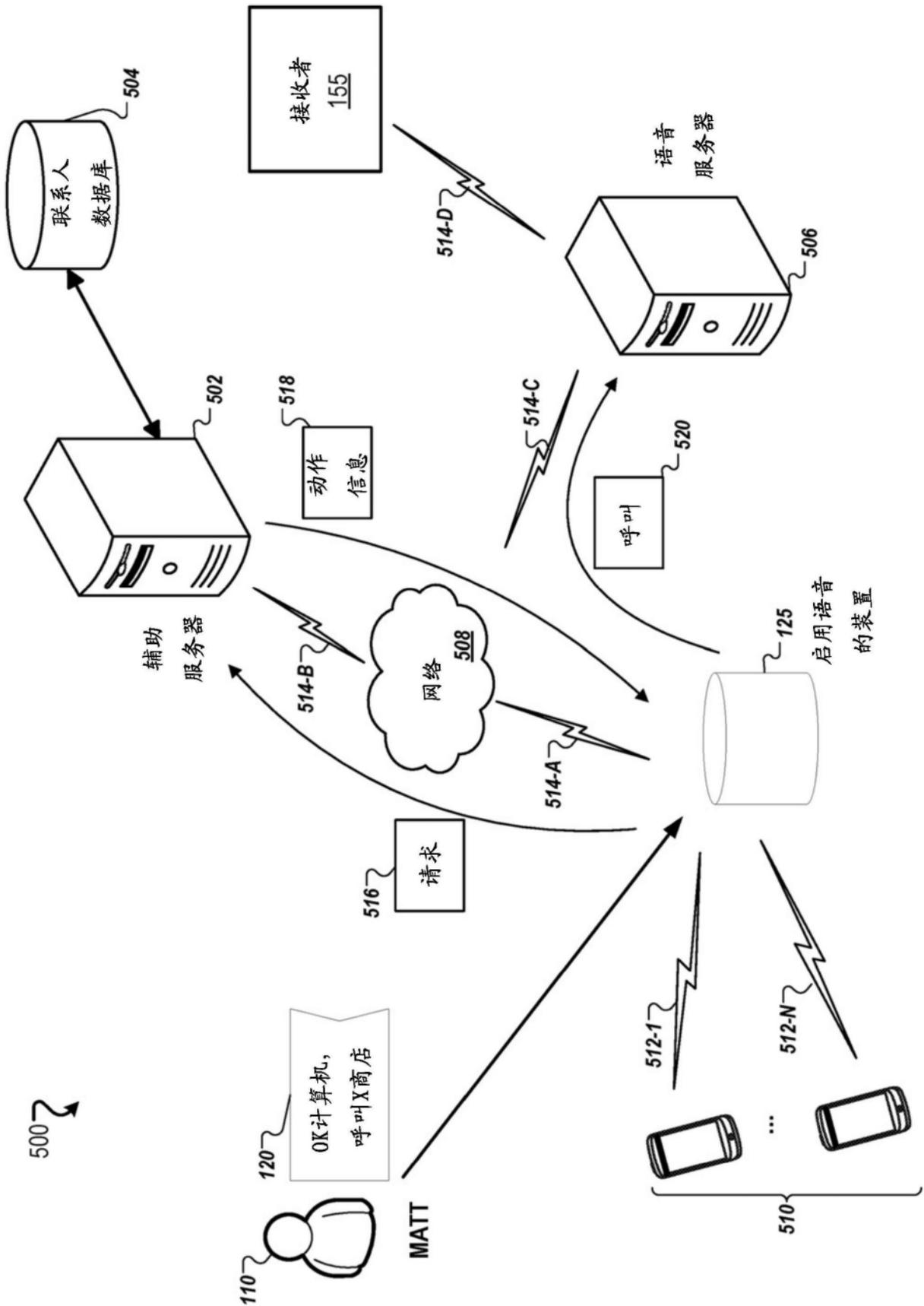


图5

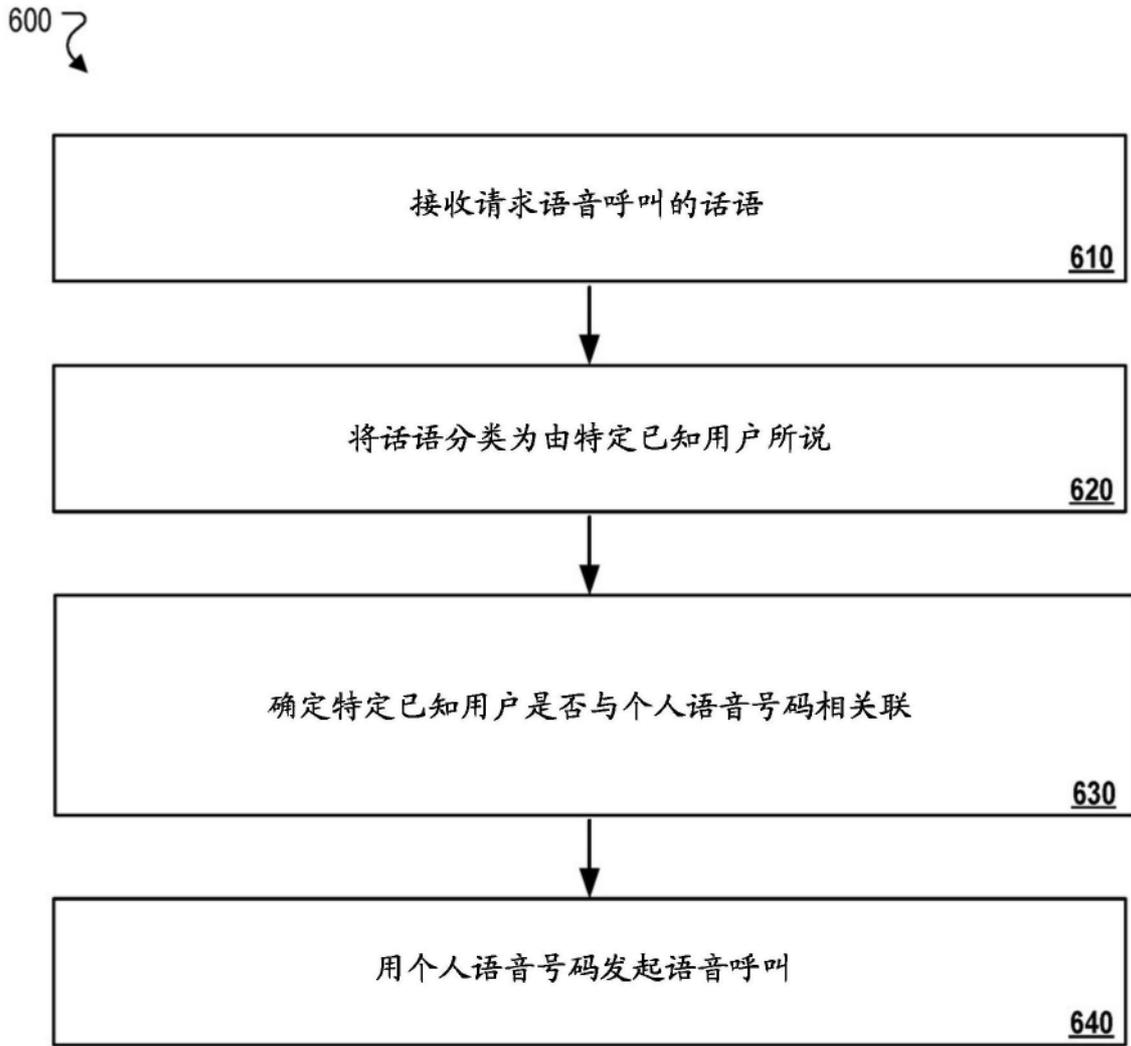


图6

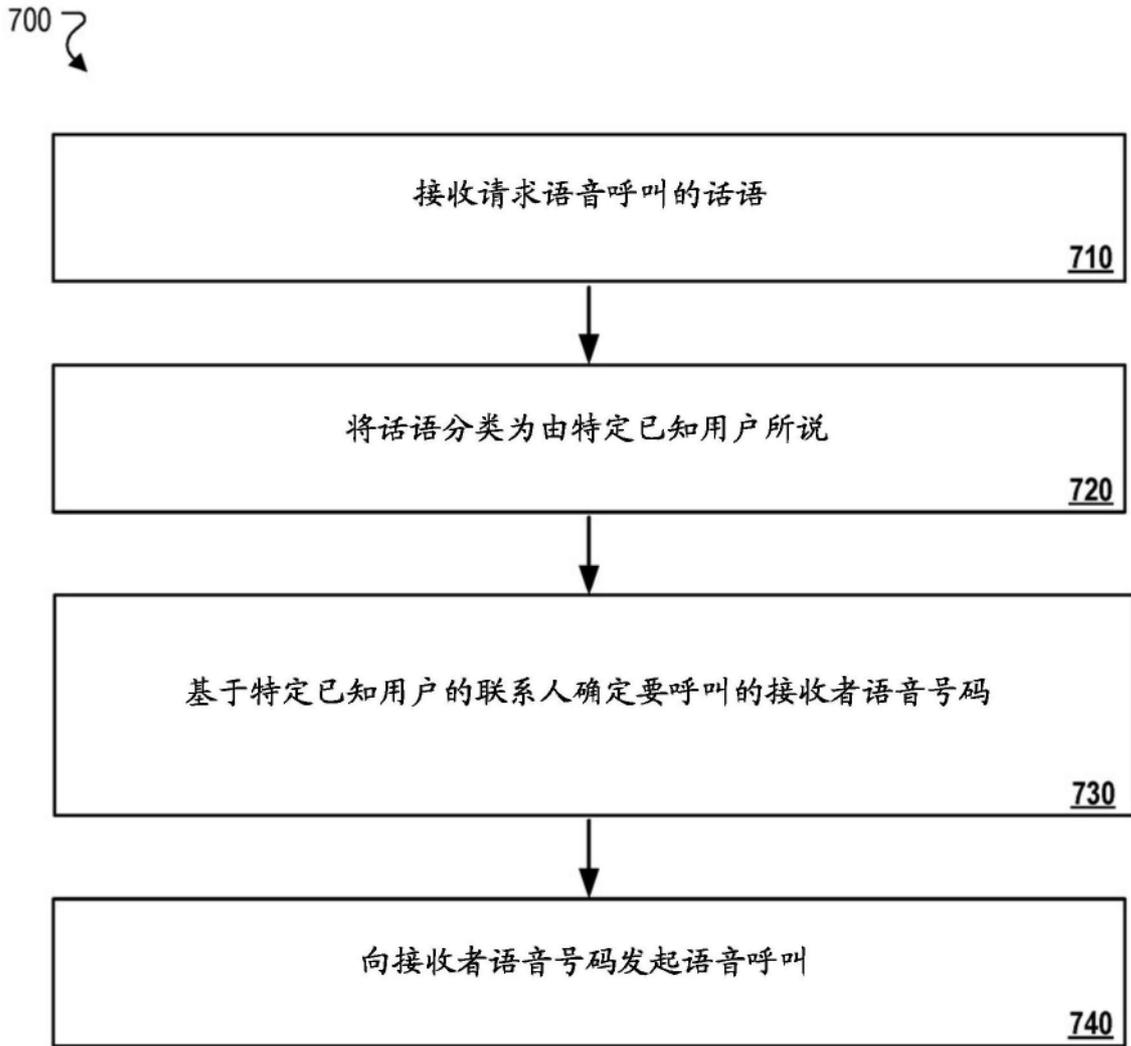


图7

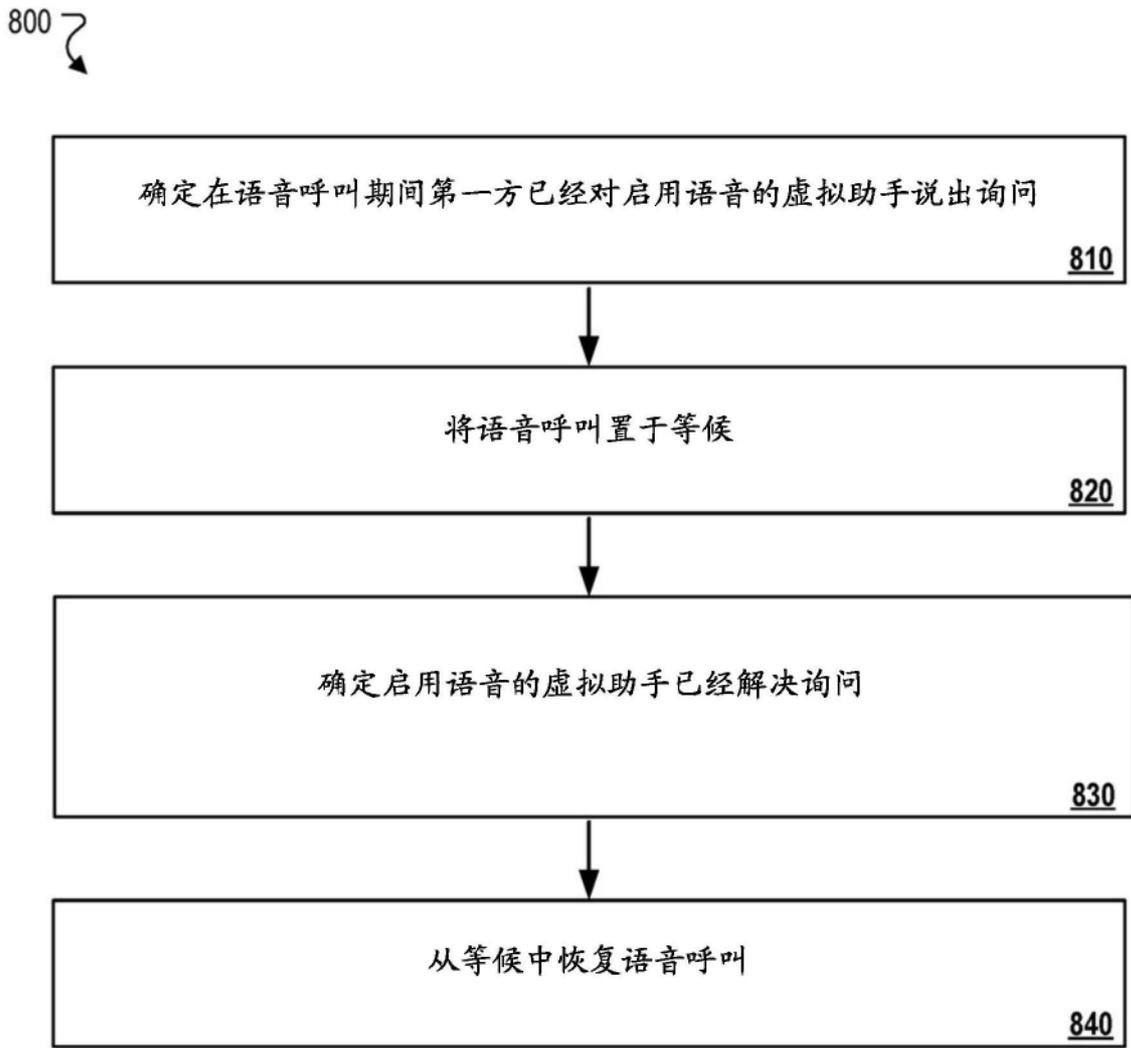


图8

