(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 104981871 B
(45)授权公告日 2018.01.02

(21)申请号 201480007157.5

(22)申请日 2014.02.11

(65)同一申请的已公布的文献号
申请公布号 CN 104981871 A

(43)申请公布日 2015.10.14

(30)优先权数据
13/768,946 2013.02.15 US

(85)PCT国际申请进入国家阶段日
2015.08.03

(86)PCT国际申请的申请数据
PCT/US2014/015852 2014.02.11

(87)PCT国际申请的公布数据
WO2014/126933 EN 2014.08.21

(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 金莱轩 刘尚忆 辛钟元

(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.
G10L 21/038(2006.01)
H04M 9/08(2006.01)

(56)对比文件
US 2006/0190254 A1,2006.08.24,
CN 1503968 A,2004.06.09,
CN 1520590 A,2004.08.11,
CN 102576542 A,2012.07.11,
CN 102473414 A,2012.05.23,
CN 101208972 A,2008.06.25,
US 2008/0300866 A1,2008.12.04,
张勇等.《基于高斯混合模型的语音带宽扩展算法的研究》.《声学学报》.2009,第34卷(第5期),
Peter Jax et al..《On artificial bandwidth extension of telephone speech》.《Signal Processing》.2003,第83卷(第8期),
王媛媛.《语音带宽扩展算法研究》.《中国优秀硕士学位论文全文数据库 信息科技辑》.2010,(第07期),
鲍枫等.《音频带宽扩展技术分析与展望》.《电讯技术》.2011,第51卷(第2期),

审查员 李海龙

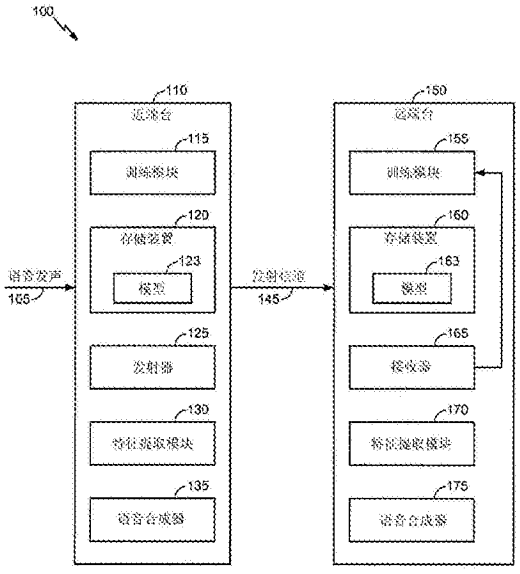
权利要求书3页 说明书11页 附图11页

(54)发明名称

个人化带宽扩展

(57)摘要

本发明提供一种个人化(即,说话者可导出)带宽扩展,其中用于带宽扩展的模型对每一特定用户个人化(例如,定制)。执行训练阶段以产生对用户个人化的带宽扩展模型。所述模型可随后在涉及所述用户的电话呼叫期间在带宽扩展阶段中使用。使用所述个人化带宽扩展模型的所述带宽扩展阶段将当较高频带(例如,宽带)不可用且所述呼叫正在较低频带(例如,窄带)上进行时激活。



CN 104981871 B

1.一种用于从近端台发射到远端台的信号的带宽扩展的方法,其包括:

在所述远端台处接收从所述近端台发射的宽带信号;

在所述远端台处以所述宽带信号的所识别说话者的数据基于由所述近端台发射的所接收宽带信号而产生带宽扩展模型;以及

使用所述带宽扩展模型对包含所识别说话者的窄带呼叫执行带宽扩展,其中提取来自所述窄带呼叫中的语音数据的窄带特征,并使用所述带宽扩展模型将所述窄带特征映射到所述所识别说话者的个人化宽带特征。

2.根据权利要求1所述的方法,其中所述带宽扩展模型是以所述所识别说话者的所述数据个人化。

3.根据权利要求1所述的方法,其中所述带宽扩展模型是基于说话者可导出信息。

4.根据权利要求1所述的方法,其中产生所述带宽扩展模型包括以所述所识别说话者的所述数据训练所述带宽扩展模型。

5.根据权利要求1所述的方法,其中产生所述带宽扩展模型包括识别所述说话者且从所述宽带信号提取宽带特征。

6.根据权利要求1所述的方法,其中对所述窄带呼叫执行带宽扩展包括:

识别所述说话者;以及

从存储装置检索与所述说话者相关联的所述带宽扩展模型。

7.根据权利要求6所述的方法,其进一步包括在执行所述宽带语音合成之后输出宽带语音。

8.根据权利要求1所述的方法,其中对所述窄带呼叫执行带宽扩展是当宽带不可用且所述呼叫正在窄带上进行时激活。

9.一种用于从近端台发射至远端台的信号的带宽扩展的设备,其包括:

用于在所述远端台处接收从所述近端台发射的宽带信号的装置;

用于以所述宽带信号的所识别说话者的数据基于由所述近端台发射的所接收宽带信号而产生带宽扩展模型的装置;以及

用于使用所述带宽扩展模型对包含所识别说话者的窄带呼叫执行带宽扩展的装置,其中提取来自所述窄带呼叫中的语音数据的窄带特征,并使用所述带宽扩展模型将所述窄带特征映射到所述所识别说话者的个人化宽带特征。

10.根据权利要求9所述的设备,其中所述带宽扩展模型是以所述所识别说话者的所述数据个人化。

11.根据权利要求9所述的设备,其中所述带宽扩展模型是基于说话者可导出信息。

12.根据权利要求9所述的设备,其中所述用于产生所述带宽扩展模型的装置包括用于以所述所识别说话者的所述数据训练所述带宽扩展模型的装置。

13.根据权利要求9所述的设备,其中所述用于产生所述带宽扩展模型的装置包括用于识别所述说话者的装置以及用于从所述宽带信号提取宽带特征的装置。

14.根据权利要求9所述的设备,其中所述用于对所述窄带呼叫执行带宽扩展的装置包括:

用于识别所述说话者的装置;以及

用于从存储装置检索与所述说话者相关联的所述带宽扩展模型的装置。

15.根据权利要求14所述的设备，其进一步包括用于在执行所述宽带语音合成之后输出宽带语音的装置。

16.根据权利要求9所述的设备，其中对所述窄带呼叫执行带宽扩展是当宽带不可用且所述呼叫正在窄带上进行时激活。

17.一种非暂时性计算机可读媒体，其包括用于执行从近端台发射到远端台的信号的带宽扩展的指令，其中所述指令致使计算机进行以下操作：

在所述远端台处接收从所述近端台发射的宽带信号；

以所述宽带信号的所识别说话者的数据基于由所述近端台发射的所接收宽带信号而产生带宽扩展模型；以及

使用所述带宽扩展模型对包含所识别说话者的窄带呼叫执行带宽扩展，其中提取来自所述窄带呼叫中的语音数据的窄带特征，并使用所述带宽扩展模型将所述窄带特征映射到所述所识别说话者的个人化宽带特征。

18.根据权利要求17所述的计算机可读媒体，其中所述带宽扩展模型是以所述所识别说话者的所述数据个人化。

19.根据权利要求17所述的计算机可读媒体，其中所述带宽扩展模型是基于说话者可导出信息。

20.根据权利要求17所述的计算机可读媒体，其中所述致使所述计算机产生所述带宽扩展模型的指令包括致使所述计算机以所述所识别说话者的所述数据训练所述带宽扩展模型的指令。

21.根据权利要求17所述的计算机可读媒体，其中所述致使所述计算机产生所述带宽扩展模型的指令包括致使所述计算机识别所述说话者且从所述宽带信号提取宽带特征的指令。

22.根据权利要求17所述的计算机可读媒体，其中所述致使所述计算机对所述窄带呼叫执行带宽扩展的指令包括致使所述计算机进行以下操作的指令：

识别所述说话者；以及

从存储装置检索与所述说话者相关联的所述带宽扩展模型。

23.根据权利要求22所述的计算机可读媒体，其进一步包括致使所述计算机在执行所述宽带语音合成之后输出宽带语音的指令。

24.根据权利要求17所述的计算机可读媒体，其中对所述窄带呼叫执行带宽扩展是当宽带不可用且所述呼叫正在窄带上进行时激活。

25.一种用于从通信的近端台发射至远端台的信号的带宽扩展的系统，其包括：

接收器，其在所述远端台处且适于接收从所述近端台发射的宽带信号；

训练模块，其适于以所述宽带信号的所识别说话者的数据基于由所述近端台发射的所接收宽带信号而产生带宽扩展模型；以及

处理器，其用于使用所述带宽扩展模型对包含所识别说话者的窄带呼叫执行带宽扩展，其中提取来自所述窄带呼叫中的语音数据的窄带特征，并使用所述带宽扩展模型将所述窄带特征映射到所述所识别说话者的个人化宽带特征。

26.根据权利要求25所述的系统，其中所述带宽扩展模型是以所述所识别说话者的所述数据个人化。

27.根据权利要求25所述的系统，其中所述带宽扩展模型是基于说话者可导出信息。

28.根据权利要求25所述的系统，其中所述训练模块适于以所述所识别说话者的所述数据训练所述带宽扩展模型。

29.根据权利要求25所述的系统，其进一步包括特征提取模块，所述特征提取模块适于识别所述说话者且从所述宽带信号提取宽带特征。

30.根据权利要求25所述的系统，其中所述处理器适于识别所述说话者，且从存储装置检索与所述说话者相关联的所述带宽扩展模型。

31.根据权利要求30所述的系统，其中所述处理器适于进一步在执行所述宽带语音合成之后输出宽带语音。

32.根据权利要求25所述的系统，其中对所述窄带呼叫执行带宽扩展是当宽带不可用且所述呼叫正在窄带上进行时激活。

个人化带宽扩展

背景技术

[0001] 信号的带宽扩展是众所周知且可经界定为使其中含有有用内容或效果的信号的频率范围(带宽)扩大的过程。近年来,带宽扩展已经用于经译码语音和音频的高频增强中,例如电话系统中的语音。因此,举例来说,窄带电话语音的质量可通过带宽扩展而改善。

[0002] 一种类型的已知带宽扩展称为盲带宽扩展。盲带宽扩展技术用以再造窄带音频信号的高频带。这些方法增加如电话或无线电通信系统中经由窄频带发射的信号的感知质量。

[0003] 然而,现有的盲带宽扩展技术使用一般化模型用于执行带宽扩展。此些一般化模型通常在用于任意说话者的带宽扩展中使用,其可导致不准确或不良的用户体验。

发明内容

[0004] 本系统和方法是针对个人化(即,说话者可导出)带宽扩展,其中用于带宽扩展的模型对每一特定用户个人化(例如,定制)。

[0005] 在一实施方案中,执行训练阶段以产生对用户个人化的带宽扩展模型。所述模型可随后在涉及所述用户的电话呼叫期间在带宽扩展阶段中使用。使用所述个人化带宽扩展模型的所述带宽扩展阶段将当较高频带(例如,宽带)不可用且所述呼叫正在较低频带(例如,窄带)上进行时激活。

[0006] 在一实施方案中,在其中可识别用户(例如,通过呼叫者ID)的宽带呼叫期间在远端台(例如,移动电话)处训练模型(对用户个人化,以使得其是基于说话者可导出信息)。

[0007] 在另一实施方案中,在呼叫期间在近端台处(例如,在用户的移动电话处)训练模型(应注意,不管网络条件如何近端都可俘获宽带语音)。

[0008] 在另一实施方案中,以用户的话音或以由移动台的所有者选择的目标话音在移动台处(不在电话呼叫期间)离线训练模型。离线训练涉及当电话不在用于电话呼叫时记录用户的话音或目标话音。

[0009] 提供此发明内容而以简化形式引入下文在具体实施方式中进一步描述的概念选择。本发明内容并不希望识别所主张的标的物的关键特征或基本特征,也并非意图用于限制 所主张的标的物的范围。

附图说明

[0010] 当结合附图阅读时,更好地理解前述发明内容以及说明性实施例的以下详细描述。出于说明实施例的目的,图式中展示实施例的实例构造;然而,所述实施例不限于所揭示的特定方法和工具。在图式中:

[0011] 图1是其中可产生且使用个人化带宽扩展的示范性系统的框图;

[0012] 图2是用于个人化带宽扩展的训练的方法的实施方案的操作流程;

[0013] 图3是执行个人化带宽扩展的方法的实施方案的操作流程;

[0014] 图4是用于个人化带宽扩展的训练的方法的另一实施方案的操作流程;

[0015] 图5是用于个人化带宽扩展的训练的方法的另一实施方案的操作流程；

[0016] 图6是用于个人化带宽扩展的训练的方法的另一实施方案的操作流程；

[0017] 图7是执行用于基于谱包络/激励的个人化带宽扩展的训练的方法的实施方案的操作流程；

[0018] 图8是可与基于谱包络/激励的个人化带宽扩展一起使用的实例映射表的图；

[0019] 图9是可与基于谱包络/激励的个人化带宽扩展一起使用的另一实例映射表的图；

[0020] 图10是执行基于谱包络/激励的个人化带宽扩展的方法的实施方案的操作流程；

[0021] 图11是实例移动台的图；以及

[0022] 图12展示示范性计算环境。

具体实施方式

[0023] 以下参考且并入有图式的详细说明描述且图解说明一或多个具体实施例。并非为了限制而是仅为了举例说明和教示而提供的这些实施例是以充足的细节展示和描述，以使所属领域的技术人员能够实践所主张的内容。因此，出于简洁起见，所述描述可能省略所属领域的技术人员已知的某些信息。

[0024] 图1是其中可(基于说话者可导出信息)产生且用以从所接收窄带(NB)语音信号形成宽带(WB)语音信号的个人化带宽扩展的示范性系统100的框图。系统100可用以产生具有比对应所接收模拟信号的频率范围大的频率范围的模拟信号。因此，信号是否为宽带信号或窄带信号取决于其与另一者的关系。虽然本文所描述的实例和实施方案指代从NB语音信号产生WB语音信号，但本系统和方法并不限于此，因为本文所揭示的系统和方法可用于从语音信号的任何其它集合产生语音信号的任何集合，例如从超宽带 (SWB)到WB或SWB到NB。

[0025] 如图所示，系统100包含可由发射方使用的近端台110，以及可由接收方使用的远端台150。近端台110和远端台150可各自为用于通信的装置，例如移动台或计算装置。相对于图11描述实例移动台800，且相对于图12描述实例计算装置900。

[0026] 近端台110和远端台150可经由一或多个发射信道145通信。所述通信可包括从近端台110到远端台150发射对应于说话者的语音发声105(话音数据)的窄带版本的窄带语音信号。远端台150可使用对说话者训练(即，基于说话者可导出信息)以产生且输出说话者的语音发声105的对应宽带版本的模型163。如在此进一步所描述，取决于实施方案，模型123、163可经在近端台110或远端台150训练、产生且存储。

[0027] 因此，语音发声105和经训练模型123、163或数据(取决于实施方案)可经由一或多个发射通道145发射。更确切地说，发射器125根据(例如)一或多个电信发射标准将说话者的语音发声105(话音数据)转换为原始语音发声105的窄带版本以用于发射。原始语音发声的窄带版本和经训练的模型或数据的发射可大体上以同时方式发生或(例如)可在发射过程期间在单独的时间发生。适合于在此实例中以及在下文阐述的实例中使用的发射信道包含电话网信道、无线蜂窝式网络信道、无线对讲机系统、有线网络或类似物。在此些发射系统中使用的窄带语音信号可限于300Hz-3.4kHz的带宽，其对应于用以使用全球移动通信系统(GSM)网络发射语音信号的带宽。

[0028] 取决于实施方案，近端台110可包括训练模块115、可存储一或多个训练模型123的存储装置120(例如，存储器)、发射器125、特征提取模块130以及语音合成器135。类似地但

不相同地，取决于实施方案，远端台150可包括训练模块155、可存储一或多个训练模型163的存储装置160(例如，存储器)、接收器165、特征提取模块170以及语音合成器175。

[0029] 在一些实施方案中，可不需要或使用近端台110处的语音合成器135。举例来说，在使用仅从近端台110中的WB输入的特征域比较机制的实施方案中，可不产生真实语音且因此可不需要或使用语音合成器135。在此情况下(其使用特征域相似性比较来计算每一现有条目与输入之间的相似性)，将模型和特征(例如NB特征)发送到远端台150以使得WB语音可在远端台150处合成。然而，在使用“按合成分析”机制以通过比较实际WB输入与合成WB输入而学习模型的实施方案中，可使用语音合成器135。此类型的实施方案使用合成语音信号域相似性比较来计算每一现有条目与输入之间的相似性。

[0030] 在近端台110处从近端台110的用户接收语音发声105。语音发声105用以提供个人化带宽扩展，其中用于带宽扩展的模型是对每一特定用户个人化(例如，定制)。

[0031] 在一实施方案中，执行训练阶段以产生对用户个人化的带宽扩展模型(在本文中也被称作“模型”)。可使用训练模块115在近端台110处或使用训练模块155在远端台150处执行训练阶段。由训练模块115产生的个人化带宽扩展模型123可存储在存储装置120中，且由训练模块155产生的模型163可存储在存储装置160中。存储装置120、160可存储一个以上模型(例如，对不同用户个人化)且可呈例如数据库的形式。

[0032] 取决于实施方案，近端台110和/或远端台150可包括特征提取模块(例如，分别为特征提取模块130、170)。特征提取模块从用户的语音提取宽带特征，且这些宽带特征用于产生训练模型。

[0033] 经训练模型123、163可存储在用户的电话(例如，近端台110)中，或发送到其它用户的电话(例如，远端台150)。基于宽带训练数据，可随后在涉及用户的窄带呼叫期间使用模型123、163。使模型在电话呼叫的再生侧(即，在远端台150处)可用。

[0034] 因此，例如模型123或模型163的模型可随后在涉及与所述模型相关联的用户的电话呼叫期间在带宽扩展阶段中使用。电话呼叫可经由一或多个发射信道145在近端台110与远端台150之间发生。确切地说，近端台110的发射器125可经由发射信道145将其已接收的语音发射到远端台的接收器165。使用个人化带宽扩展模型的带宽扩展阶段将在较高频带(例如，宽带)不可用且呼叫正在较低频带(例如，窄带)上进行时激活。在窄带呼叫期间，接收器165使用来自说话者的窄带语音信号连同经训练的模型123、163(先前产生且存储)以产生对应于说话者的语音发声105的宽带版本的宽带语音信号。

[0035] 特征提取模块130、170还可在电话呼叫期间从电话呼叫提取窄带特征，且将这些窄带特征提供到语音合成器135、175。取决于实施方案，近端台110和/或远端台150可包括语音合成器(例如，语音合成器135、175)。语音合成器135、175使用与所提取窄带特征组合的个人化带宽扩展训练模型执行宽带语音合成。可随后输出宽带语音。

[0036] 图2是用于个人化带宽扩展的训练的方法200的实施方案的操作流程。训练阶段用以产生可随后在窄带呼叫中使用的模型。取决于实施方案，如在此进一步所描述，在远端台以传入宽带呼叫、在近端台在宽带呼叫期间或离线地执行训练阶段。在210，例如通过用户将其名称或其它识别符键入或输入台中或通过例如呼叫者ID来识别说话者(在本文中也被称作用户)。

[0037] 在220，说话者发出发声(也被称作语音发声)，其由所述台俘获或另外接收。在

230，从所述发声(例如，从对应于所述发声的话音数据)提取宽带特征。在240执行特征训练(直到已收集足够数据用于恰当训练)以产生训练模型。可收集数据直到满足一或多个预定条件，例如直到所述数据足够不同(满足分集阈值或准则)和/或已收集数据达足够长时间(例 如，取决于实施方案而为若干分钟、小时、天)。此时，在250可存储训练模型，且可在窄带呼叫中使用。

[0038] 预期如果未收集足够数据以产生用于个人化带宽扩展的训练模型，那么可以迄今已收集的个人化数据来调适已针对一般群体训练的带宽扩展模型(即，“一般化模型”)。举例来说，在其中表(例如下文相对于图7-10描述的表)含有用于一般化模型的数据且所述表将经修改为含有个人化数据以用作(或一起使用)个人化训练模型的表的实施方案中，可使用迄今已接收的个人化数据来调适所述表。此允许更平稳地过渡到个人化数据的完整集合已收集(例如，当满足上述一或多个预定条件时)且并入所述表中时(即，当使用已收集的个人化数据的完整集合修改所述表时)。在另一实施方案中，如果未收集足够数据以产生用于个人化带宽扩展的训练模型，那么可使用一般化模型直到已收集足够个人化数据。

[0039] 图3是执行个人化带宽扩展的方法300的实施方案的操作流程。在310，窄带呼叫在近端台与远端台之间发生，且带宽扩展阶段开始。在320(例如，使用呼叫者ID)识别说话者，且在330从存储装置(例如，在再生台侧，例如在远端台)检索或从发射台侧(例如，窄带语音源，例如近端台)接收适当带宽扩展训练模型(与所述说话者相关联)。可采用某种方式的说话者识别来确保当前说话的人是在模型的产生期间已说话的同一个人。如果无法识别说话者，那么可拒绝或终止个人化带宽扩展。

[0040] 在340，提取来自电话呼叫中的语音发声(例如，来自话音数据)的窄带特征且使用所述模型映射到说话者的个人化宽带特征。在350使用与所提取窄带特征组合的个人化带宽扩展训练模型执行宽带语音合成。在360输出宽带语音。

[0041] 图4是用于个人化带宽扩展的训练的方法的另一实施方案的操作流程。在此实施方案中，在其中可识别(例如，通过呼叫者ID)用户的宽带呼叫期间在远端台150处训练(对用户个人化)模型。

[0042] 在410，在来自近端台的宽带呼叫期间在远端台处接收说话者的话音(例如，语音发声)。在415，确定说话者的话音或语音发声的声学条件是否具有足够高的质量以用于训练。所述质量可基于例如信噪比(SNR)或其它基于噪声的统计数据和/或测量值。因此，举例来说，如果SNR高于预定阈值，那么所述发声的质量足够高以在个人化带宽扩展训练模型的产生中使用。如果不是，那么处理返回到410且继续直到声学条件可接受。

[0043] 如果在415确定声学条件可接受，那么在420从语音发声提取窄带和宽带特征，且在430通过例如说话者、远端台的用户或呼叫者ID 425识别说话者。如果任何带宽扩展特征已先前产生且存储，那么在435可从存储装置(例如本地数据库)检索关于说话者的 带宽扩展(BWE)特征，因为这是迭代过程。

[0044] 在440，使用从420提取的特征和从435检索的带宽扩展特征执行特征训练。此特征训练可包含更新特征空间且采用用于说话者的窄带宽带映射。可使用用于确定相似性测量和匹配的任何技术，例如上述特征域相似性比较、上述合成语音信号域相似性比较以及任何距离度量，包含但不限于具有可允许时间调整(在时域或频域中)的最小二乘拟合、基于特征的方法(例如使用LPC/LPCC、MFCC或音频指纹)或基于较高阶的方法(例如交叉累积量、

经验库尔贝克-莱伯尔发散(Kullback-Leibler Divergence),或板仓-西都(Itakura-Saito)距离)。

[0045] 在一实施方案中,可从初始化特征空间(例如,用于一般群体的通用码簿)朝向个人化特征空间(例如,用于说话者的个人化码簿)执行更新。可以当前特征向量与最近的现有码簿条目之间的经加权和替换码簿条目。如果太多的存储器消耗变成问题,那么在一实施方案中通过减轻从通用特征空间的更新条件而可对某些说话者给出存储器消耗的优先级。或者或另外,可存储仅不同特征空间或仅从通用特征空间的差量或改变。一般化特征空间可预先存储在所述台处或者发送或另外提供到所述台用于后续使用。

[0046] 预期即使不具有特征空间的更新,个人化带宽扩展仍可使用个人化窄带宽带映射模型来获得。

[0047] 在450,可以训练的完成百分比(例如,相对于预定阈值量的特征和/或在训练模型的完成之前发生的映射)来存储(例如,在本地数据库中)用于说话者的带宽扩展特征空间和映射。在一实施方案中,可通过或基于测量用于以恰当VAD(话音活动检测)训练以排除非语音部分且测量构造多少新窄带宽带映射的计数而获得模型训练完成百分比。

[0048] 在一实施方案中,指示符可提供于地址簿中对用户可存取,以指示宽带以某个训练完成百分比而可用。地址簿可包括(例如)存储且列出各种用户的数据库以及指示宽带是否可用于每一用户和/或训练完成百分比的指示符。用户可决定是否基于用户接口以训练完成百分比使用对宽带的带宽扩展。或者或另外,可设定默认值以使得每当训练完成百分比到达某一水平(例如,99%)时,那么当宽带不可用时或例如当一些译码资源将保存或保藏(例如,在呼叫期间的不良网络条件或繁重数据传送)时自动切换到对宽带的带宽扩展。

[0049] 在460,确定是否存在所存储的足够带宽扩展数据(例如,相对于预定阈值)以当宽带连接不可用时在呼叫中使用。如果不是,那么处理在410继续以搜集更多发声且产生更多带宽扩展数据。然而,如果已产生且存储足够带宽扩展数据,那么在470处处理停止且所得模型准备好在窄带话音通信中使用。

[0050] 取决于实施方案,个人化带宽扩展可在任何若干情况中使用,例如(1)当宽带在整个呼叫中不可用时;(2)当宽带对于呼叫的一部分可用但对于呼叫的另一部分不可用且当宽带不可用时个人化带宽扩展变成激活时;以及(3)当宽带始终可用或对于呼叫的某一部分可用时,且即使当宽带可用时在个人化带宽扩展模型变成足够良好且运营商公司决定切换到个人化带宽扩展模式的情况下个人化带宽扩展变成激活。

[0051] 图5是用于个人化带宽扩展的训练的方法500的另一实施方案的操作流程。在此实施方案中,在呼叫期间在近端台110(例如,在用户的移动电话)训练模型。应注意不管网络条件如何,近端台110都可俘获宽带语音。

[0052] 在510,电话呼叫发生(即,在线呼叫),其中说话者在所述呼叫上且对所述说话者的台发声(提供语音发声)。在515,确定说话者的话音或语音发声的声学条件是否具有足够高的质量以用于训练。如上所述,SNR技术可用以确定所述声学条件是否可接受。如果不是,那么处理返回到510且继续直到声学条件可接受。

[0053] 如果在515确定声学条件可接受,那么在520经由例如用户输入525由说话者识别说话者。在530,确定所识别说话者是否为经登记用户(例如,通过将说话者的识别与存储装置中与说话者的台相关联的登记用户的列表)。确定所识别说话者是否为经登记用户用以

防止对除了作为所述台的所有者(例如,移动电话的所有者)的经登记用户之外的其它人的不希望的训练。如果说话者不是经登记用户,那么处理返回到510。然而,如果确定所识别说话者是经登记用户,那么在540从语音发声提取窄带和宽带特征,且如果任何带宽扩展特征已经先前产生且所存储,那么在545可从存储装置(例如本地数据库)检索关于说话者的带宽扩展特征,因为这是迭代过程。

[0054] 在550,使用从540提取的特征和从545检索的带宽扩展特征执行特征训练。此特征训练可包含更新特征空间且调适用于说话者的窄带-宽带映射,类似于上文相对于方法400所描述。在555,可以训练的完成百分比(例如,相对于在训练模型的完成之前发生的特征和/或映射的量的预定阈值)(例如,在本地数据库中)存储用于说话者的带宽扩展特征空间和映射。

[0055] 在560,确定是否存在所存储的足够带宽扩展数据(例如,相对于预定阈值)以当宽带连接不可用时在呼叫中使用。如果不是,那么处理在510继续以搜集更多发声且产生更多带宽扩展数据。然而,如果已产生且存储足够带宽扩展数据,那么处理停止且所得模型准备好在窄带话音通信中使用。

[0056] 在570处可将模型发送到远端台以在窄带话音通信中使用。每当用户到用户数据连接变成可用的且存在与先前数据的差量时,可在背景中进行此数据发射。或者,类似于邮件或短消息推送中,每当模型的所有者想要更新她/他的模型时,可向在其电话地址簿中具有联系人的用户通知可用的用户的模型。或者,在呼叫的开始中,如果经训练模型可用且由于最后同步而为新的或经修改的,那么可将其发射。

[0057] 图6是用于个人化带宽扩展的训练的方法600的另一实施方案的操作流程。在此实施方案中,以用户的话音或以由移动台的所有者选定的目标话音在移动台处(例如,在近端台110或远端台150或具有适当能力的任何移动台处)离线(例如,不在电话呼叫期间)训练模型。离线训练涉及当电话不在用于电话呼叫时记录用户的话音或目标话音。

[0058] 在610,虽然台是离线的(即,电话呼叫不在进行),那么说话者对所述台发声以将语音发声提供到说话者的台。在615,确定说话者的话音或语音发声的声学条件是否具有足够高的质量以用于训练。如果不是,那么处理返回到610且继续直到声学条件可接受。

[0059] 如果在615确定声学条件可接受,那么在625经由用户输入630(例如,来自与台相关联的预定电话簿的识别的用户选择)由说话者识别说话者。

[0060] 在620从语音发声提取窄带和宽带特征,且如果任何带宽扩展特征已经先前产生且存储,那么在635可从存储装置(例如本地数据库)检索识别的说话者,因为这是迭代过程。

[0061] 在640,使用从620提取的特征和从635检索的带宽扩展特征执行特征训练。此特征训练可包含更新特征空间且调适用于说话者的窄带宽带映射。在650,类似于上文相对于方法400所描述,可以训练的完成百分比(例如,相对于在训练模型完成之前发生的特征和/或映射的量的预定阈值)(例如,在本地数据库中)存储用于说话者的带宽扩展特征空间和映射。

[0062] 在660,确定是否存在所存储的足够带宽扩展数据(例如,相对于预定阈值)以当宽带连接不可用时在呼叫中使用。如果不是,那么处理在610继续以搜集更多发声且产生更多带宽扩展数据。然而,如果已产生且存储足够带宽扩展数据,那么在670处处理停止且所得

模型准备好在窄带话音通信中使用。

[0063] 图7是当宽带源可用时执行用于基于谱包络/激励的个人化带宽扩展的训练的方法700的实施方案的操作流程。在705，产生来自语音源(例如，语音发声)的语音且提供到台。在710执行宽带(WB)特征提取且在720执行窄带(NB)特征提取。在715将WB线谱频率(LSF)和NB LSF提供到WB LSF码簿更新，且构造从NB LSF码映射到WB LSF码的映射表。图8中展示实例映射表740，例如在715产生的一个映射表。图8中连同对应WB LSF 747一起提供NB LSF索引745。以一般化LSF(例如，关于人的一般群体) 初始化WB码簿。对于NB LSF码i，基于当前帧的WB LSF而更新相关联WB LSF码簿条目。

[0064] 返回到图7，在725将WB激励连同NB LSF和NB激励一起提供到WB激励码簿更新，且构造从NB到WB激励的码簿映射表。在730确定是否停止训练(例如，基于是否已获得足够训练数据)。如果是，那么训练在735停止；否则，处理在705继续，其中在训练中使用额外语音发声。图9中展示实例映射表750，例如在725产生的一个映射表。连同对应激励索引757一起提供LSF索引755。以一般化激励(即，从一般群体导出的激励)初始化码簿。对于NB LSF码i和NB激励码j，更新相关联WB激励码簿条目。

[0065] 图10是仅当窄带源可用时执行基于谱包络/激励的个人化带宽扩展的方法760的实施方案的操作流程。在765，将窄带语音提供到台，例如远端台150。在770，提取窄带特征，从而导致NB激励码和NB LSF码。

[0066] 在775，使用NB激励码和NB LSF码执行到WB激励的映射。还在780执行从NB LSF到WB LSF的映射。使用所述映射，在785合成且在790输出WB语音。

[0067] 应注意，NB与WB之间可能存在非唯一映射问题。为了解决此问题，可使用类似于用于语音辨识的语言模型(称为N元语法)的技术。在一实施方案中，查看过去(或在准许延迟的情况下可能的未来)帧以构造映射。举例来说，如果构造2元语法，啊么映射变成NB_{i-1},NB{i}-WB{i}而不是NB_{i}-WB{i}。

[0068] 在用于谱包络(例如，LPC(线性预测系数)、LSF、梅尔频率倒谱系数(MFCC)等)的训练阶段的实施方案中，当宽带信号可用时，找到用于窄带语音的谱包络(SE)码簿条目和对应宽带信号的SE以构造例如码簿映射或HMM/GMM(隐式马尔可夫模型/高斯混合模型)。也可以应用保持较低频带SE和/或防止太近的SE的任何方法。

[0069] 在一实施方案中，对于码簿映射，以用于一般化群体的BWE的表初始化WB SE表。当观测到对应NB SE时替换/调适表中的每一条目。可维持且分析经更新条目的数目和/或每条目的更新的数目以确定是否已获得足够的训练数据。

[0070] 在用于谱包络的个人化带宽扩展阶段的实施方案中，给定NB SE，通过码簿映射(表查找或滤波)或从概率模型(GMM/HMM)确定对应WB SE。

[0071] 关于用于激励实施方案的训练阶段，给定NB激励码簿(FCB(固定码簿)或FCB/ACB(自适应码簿))和NB SE，本系统和方法可构造从NB SE和激励到WB激励的映射/概率模型。举例来说，可能构造具有N*M个条目的WB激励表，其中N是NB激励码簿中的条目的数目且M是NB SE码簿中的条目的数目。以用于一般化群体的带宽扩展的条目初始化此表。当观测到对应NB SE/激励组合时替换/调适所述表中的每一条 目。可维持且分析经更新条目的数目和/或每条目的更新的数目以确定是否已获得足够的训练数据。

[0072] 对于带宽扩展阶段，可靠性量度(准则)可用以确定是否使用个人化带宽扩展且可

基于表中的经更新条目的数目和/或在时间上的训练数据长度。如果例如高于预定量或阈值，那么可使用个人化带宽扩展。

[0073] 此外，在一些实施方案中，可使用多个等级的可靠性量度。举例来说，(1)如果可靠性量度极高，那么甚至当网络资源足够用于WB(以减少网络容量)时可使用个人化带宽扩展；(2)如果可靠性量度处于中等范围内，那么当所述网络无法支持WB时可使用个人化带宽扩展；以及(3)如果可靠性量度是低的，那么可使用窄带通信或一般化带宽扩展。

[0074] 除非另有指示，否则对具有特定特征的设备的操作的任何揭示内容还明确地希望揭示具有类似特征的方法(且反过来也一样)，且对根据特定配置的设备的操作的任何揭示内容还明确地希望揭示根据类似配置的方法(且反过来也一样)。

[0075] 如本文所使用，术语“确定”(及其语法变体)在极其宽广意义上使用。术语“确定”涵盖许多种类的动作，且因此“确定”可包含计算、估计、处理、导出、调查、查找(例如，在表格、数据库或另一数据结构中查找)、查实等。又，“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)等等。并且，“确定”可包含解析、选择、挑选、建立等等。

[0076] 术语“信号处理”(及其语法变体)可指代信号的处理和解译。所关注的信号可包含声音、图像和许多其它信号。此些信号的处理可包含存储和重构、信息从噪声的分离、压缩以及特征提取。术语“数字信号处理”可指代数字表示中的信号的研究以及这些信号的处理方法。数字信号处理是例如移动台、非移动台和因特网等许多通信技术的元素。用于数字信号处理的算法可使用专用计算机而执行，所述专用计算机可利用称为数字信号处理器(有时简称为DSP)的专用微处理器。

[0077] 结合本文所揭示的实施例描述的方法、过程或算法的步骤可直接以硬件、以由处理器执行的软件模块或以所述两个的组合实施。方法或过程中的各种步骤或动作可以展示的次序执行，或可以另一次序执行。另外，可省略一或多个过程或方法步骤，或可将一或多个过程或方法步骤添加到所述方法和过程。可在方法和过程的开始、末尾或介入的现有元素中添加额外步骤、框或动作。

[0078] 图11展示无线通信系统中的实例移动台800的设计的框图。移动台800可为智能电话、蜂窝式电话、终端、手持机、PDA、无线调制解调器、无绳电话等。所述无线通信系统可为码分多址(CDMA)系统、GSM系统等。

[0079] 移动台800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812接收且提供到接收器(RCVR)814。接收器814调节且数字化所接收的信号且将样本提供到数字区段820用于进一步处理。在发射路径上，发射器(TMTR)816从数字区段820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812发射到基站。接收器814和发射器816可为可支持CDMA、GSM等的收发器的部分。

[0080] 数字区段820包含各种处理、接口和存储器单元，例如调制解调器处理器822、精简指令集计算机/数字信号处理器(RISC/DSP)824、控制器/处理器826、内部存储器828、一般化音频编码器832、一般化音频解码器834、图形/显示处理器836，以及外部总线接口(EBI)838。调制解调器处理器822可执行用于数据发射和接收的处理，例如编码、调制、解调和解码。RISC/DSP 824可执行用于移动台800的一般且专用处理。控制器/处理器826可引导数字区段820内的各种处理和接口单元的操作。内部存储器828可存储用于数字区段820内的各种单元的数据和/或指令。

[0081] 一般化音频编码器832可执行用于来自音频源842、麦克风843等的输入信号的编码。一般化音频解码器834可执行用于经译码音频数据的解码且可将输出信号提供到扬声器/头戴式耳机844。图形/显示处理器836可执行用于可呈现给显示单元846的图形、视频、图像和文本的处理。EBI 838可促进数字区段820与主存储器848之间的数据传送。

[0082] 数字区段820可用一或多个处理器、DSP、微处理器、RISC等来实施。数字区段820还可制造于一或多个专用集成电路(ASIC)和/或一些其它类型的集成电路(IC)上。

[0083] 图12展示其中可实施实例实施方案和方面的示范性计算环境。计算系统环境仅是合适的计算环境的一个实例，且不希望暗示对于使用范围或功能性的任何限制。

[0084] 可使用正由计算机执行的例如程序模块等计算机可执行指令。总地来说，程序模块包含例程、程序、对象、组件、数据结构等，其执行特定任务或实施特定抽象数据类型。可使用分布式计算环境，其中任务由远程处理装置执行，所述远程处理装置经由通信网络或其它数据发射媒体链接。在分布式计算环境中，程序模块和其它数据可位于包含存储器存储装置的本地和远程计算机存储媒体两者中。

[0085] 参考图12，用于实施本文描述的方面的示范性系统包含计算装置，例如计算装置900。在其最基础配置中，计算装置900通常包含至少一个处理单元902和存储器904。取决于计算装置的确切配置和类型，存储器904可为易失性(例如，随机存取存储器(RAM))、非易失性(例如，只读存储器(ROM)、快闪存储器等)，或两者的某一组合。此 最基础配置在图12中由虚线906说明。

[0086] 计算装置900可具有额外特征和/或功能性。举例来说，计算装置900可包含额外存储装置(可装卸式和/或非可装卸式)，其包含但不限于磁盘或光盘或磁带。此额外存储装置在图12中由可装卸式存储装置908和非可装卸式存储装置910说明。

[0087] 计算装置900通常包含多种计算机可读媒体。计算机可读媒体可为可由装置900存取且包含易失性和非易失性媒体以及可装卸式和非可装卸式媒体的任何可用媒体。计算机存储媒体包含在任何方法或技术中实施用于存储例如计算机可读指令、数据结构、程序模块或其它数据等信息的易失性和非易失性以及可装卸式和非可装卸式媒体。存储器904、可装卸式存储装置908以及非可装卸式存储装置910全部是计算机存储媒体的实例。计算机存储媒体包含(但不限于)RAM、ROM、电可擦除程序只读存储器(EEPROM)、快闪存储器或其它存储器技术、CD-ROM、数字多功能光盘(DVD)或其它光学存储装置、盒式磁带、磁带、磁盘存储器装置或其它磁性存储装置，或可用于存储所要信息且可由计算装置700存取的任何其它媒体。任何此类计算机存储媒体都可以是装置900的一部分。

[0088] 计算装置900可含有允许装置与其它装置通信的通信连接912。计算装置900还可具有输入装置914，例如键盘、鼠标、笔、话音输入装置、触摸输入装置等。还可包含例如显示器、扬声器、打印机等输出装置916。所有这些装置在此项技术中是众所周知的且此处不需要详细论述。

[0089] 一般来说，本文所描述的任何装置可表示各种类型装置，例如无线或有线电话、蜂窝式电话、膝上型计算机、无线多媒体装置、无线通信PC卡、PDA、外部或内部调制解调器、通过无线或有线信道通信的装置等等。装置可具有各种名称，例如接入终端(AT)、存取单元、订户单元、移动台、移动装置、移动单元、移动电话、移动装置、远程站、远程终端、远程单元、用户装置、用户设备、手持式装置、非移动台、非移动装置、端点等等。本文所述的任何装置

均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。

[0090] 可通过各种装置来实施本文中所描述的技术。举例来说，这些技术可以硬件、固件、软件或其组合来实施。所属领域的技术人员将进一步了解，结合本文中的揭示内容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。为清楚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而描述了各种说明性组件、块、模块、电路及步骤。此功能性是实施为硬件还是软件取决于特定应用和强加于整个系统的设计约束。所属领域的技术人员可针对每一特定应用以不同方式实 施所描述的功能性，但此类实施决策不应被解译为引起对本发明的范围的偏离。

[0091] 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。

[0092] 因此，结合本文中的揭示内容描述的各种说明性逻辑块、模块和电路可以用通用处理器、DSP、ASIC、FPGA或经设计以执行本文所述的功能的其它可编程逻辑装置、离散门或晶体管逻辑、离散硬件组件或其任何组合来实施或执行。通用处理器可以是微处理器，但在替代方案中，处理器可以是任何处理器、控制器、微控制器或状态机。处理器还可实施为计算装置的组合，例如，DSP与微处理器的组合、多个微处理器、结合DSP核心的一或多个微处理器，或任何其它此类配置。

[0093] 对于固件和/或软件实施方案，所述技术可实施为在计算机可读媒体上的指令，所述计算机可读媒体例如随机存取RAM、ROM、非易失性RAM、可编程ROM、EEPROM、快闪存储器、压缩光盘(CD)、磁性或光学数据存储装置，或类似物。所述指令可由一或多个处理器执行且可致使处理器执行本文所描述的功能性的某些方面。

[0094] 如果以软件实施，则可将功能作为一或多个指令或代码而存储在计算机可读媒体上或经由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与通信媒体两者，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通用或专用计算机存取的任何可用媒体。借助于实例而非限制，这些计算机可读媒体可包括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，或任何其它可以用于运载或存储指令或数据结构的形式的期望程序代码装置并且可以由通用或专用计算机或通用或专用处理器存取的媒体。而且，任何连接被适当地称为计算机可读媒体。举例来说，如果使用同轴缆线、光纤缆线、双绞线、数字订户线(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发射软件，则所述同轴缆线、光纤缆线、双绞线、DSL或无线技术(例如，红外线、无线电和微波)包含在媒体的定义中。如本文所使用，磁盘和光盘包含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，而光盘用激光以光学方式再现数据。上文的组合也应包含在计算机可读媒体的范围内。

[0095] 软件模块可驻留在RAM存储器、快闪存储器、ROM存储器、EPROM存储器、EEPROM存储器、寄存器、硬盘、可卸除式磁盘、CD-ROM，或所属领域中已知的任何其它形式的存储媒体中。示例性存储媒体耦合到处理器，使得处理器可从存储媒体读取 信息和将信息写入到存储媒体。在替代方案中，存储媒体可与处理器成一体式。处理器及存储媒体可驻留在ASIC

中。ASIC可驻留于用户终端中。在替代方案中，处理器和存储媒体可作为离散组件驻留在用户终端中。

[0096] 虽然示范性实施方案可能参考利用在一或多个独立计算机系统的上下文中的当前揭示的标的物的方面，所述标的物不限于此，而是可结合任何计算环境(例如网络或分布式计算环境)来实施。再者，当前揭示的标的物的方面可在多个处理芯片或装置中或跨越多个处理芯片或装置而实施，且可类似地跨越多个装置实现存储。此些装置可能包含例如PC、网络服务器和手持式装置。

[0097] 虽然已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特定特征或动作。事实上，揭示上文所描述的特定特征和动作作为实施权利要求书的实例形式。

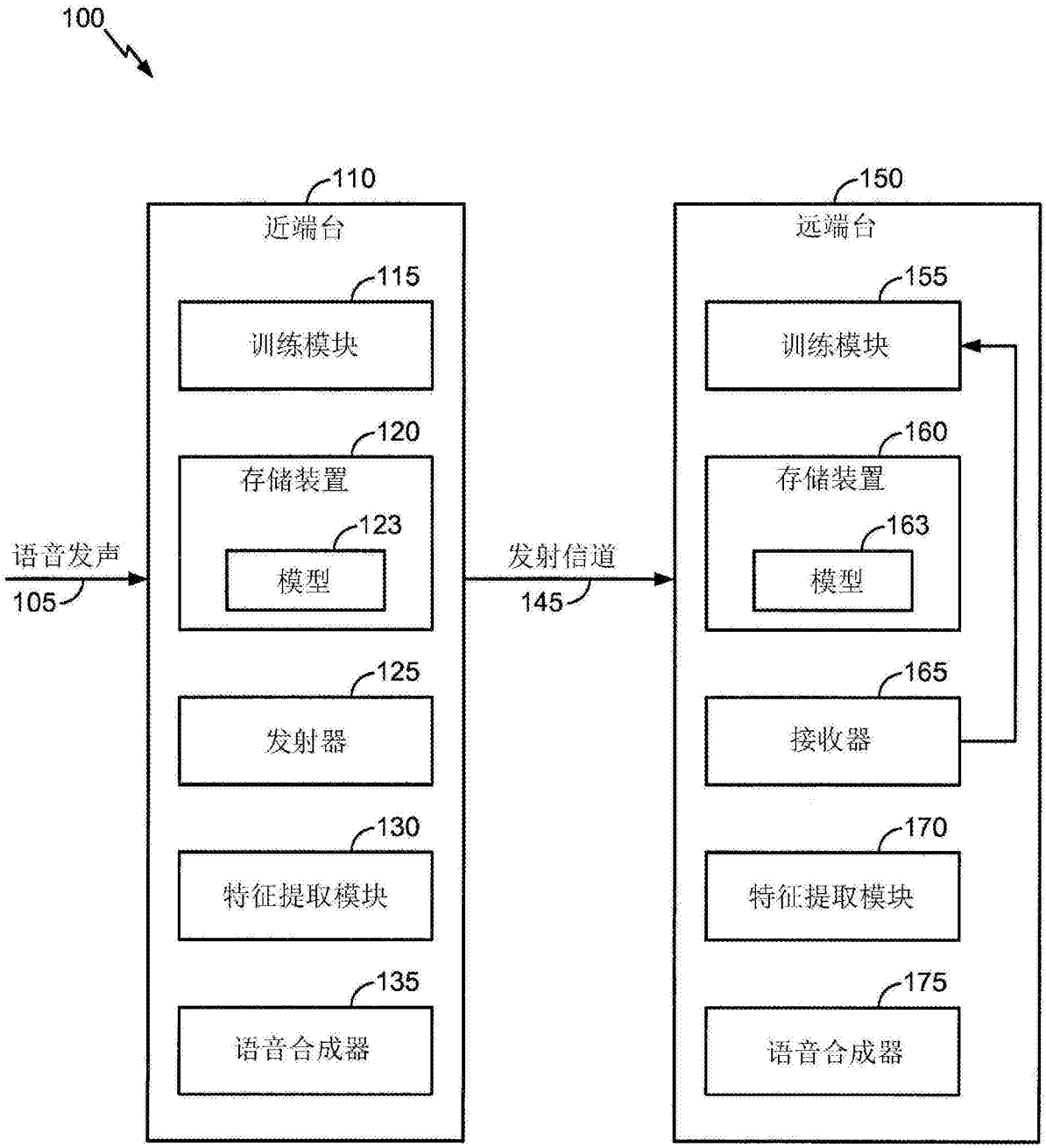

100
110
近端台
115
训练模块
120
存储装置
123
模型
125
发射器
130
特征提取模块
135
语音合成器
语音发声
105
发射信道
145
150
远端台
155
训练模块
160
存储装置
163
模型
165
接收器
170
特征提取模块
175
语音合成器

图1

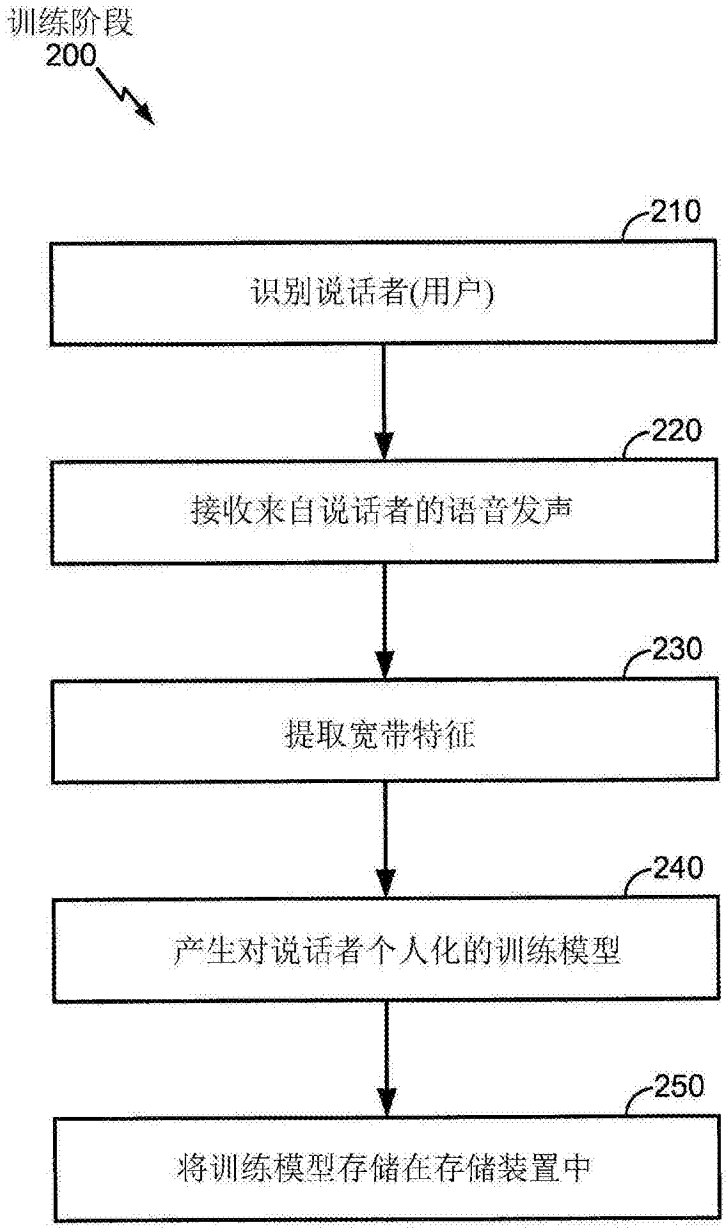

训练阶段
200
210
识别说话者(用户)
220
接收来自说话者的语音发声
230
提取宽带特征
240
产生对说话者个人化的训练模型
250
将训练模型存储在存储装置中

图2

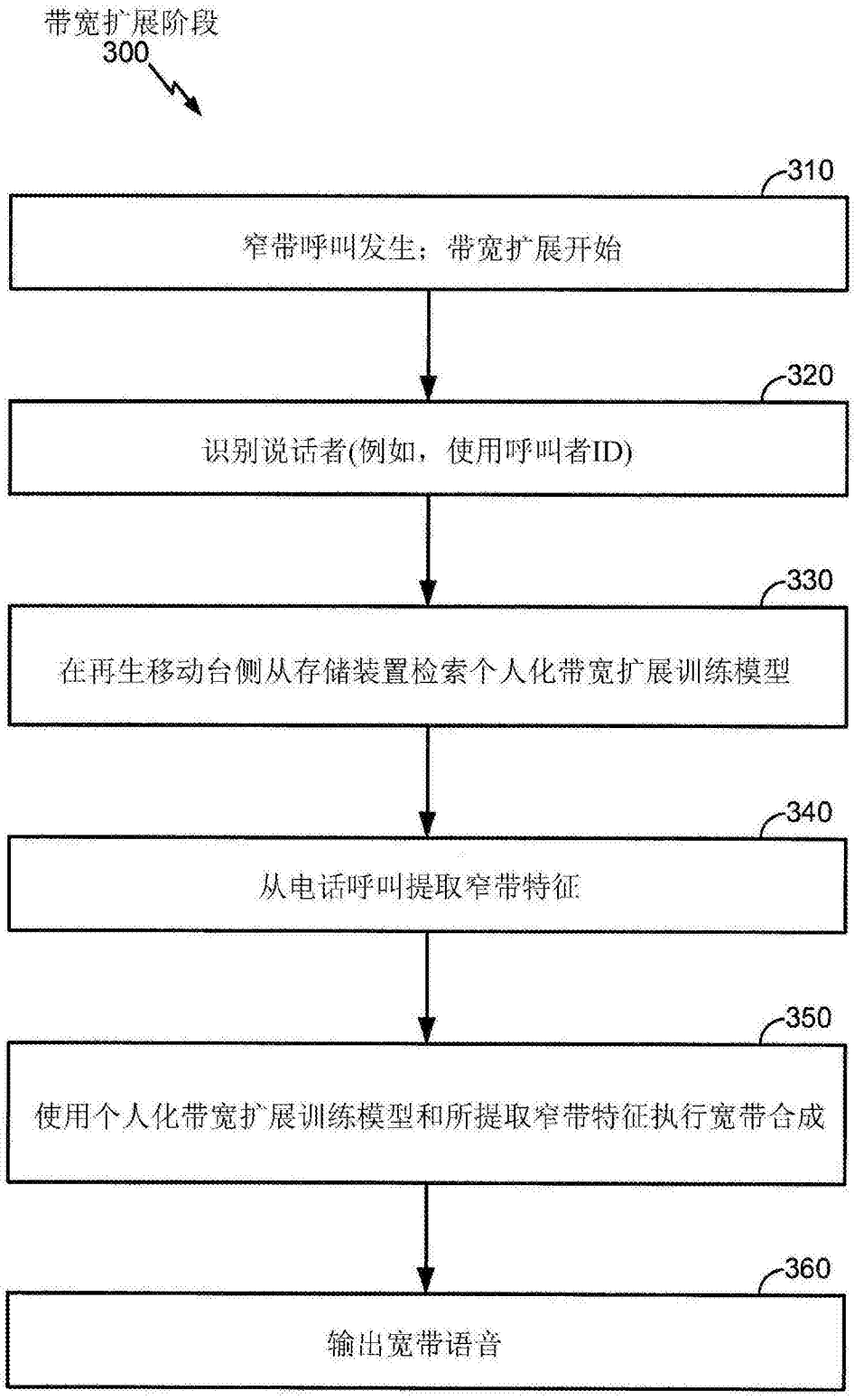

带宽扩展阶段
300
310
窄带呼叫发生；带宽扩展开始
320
识别说话者(例如，使用呼叫者ID)
330
在再生移动台侧从存储装置检索个人化带宽扩展训练模型
340
从电话呼叫提取窄带特征
350
使用个人化带宽扩展训练模型和所提取窄带特征执行宽带合成
360
输出宽带语音

图3

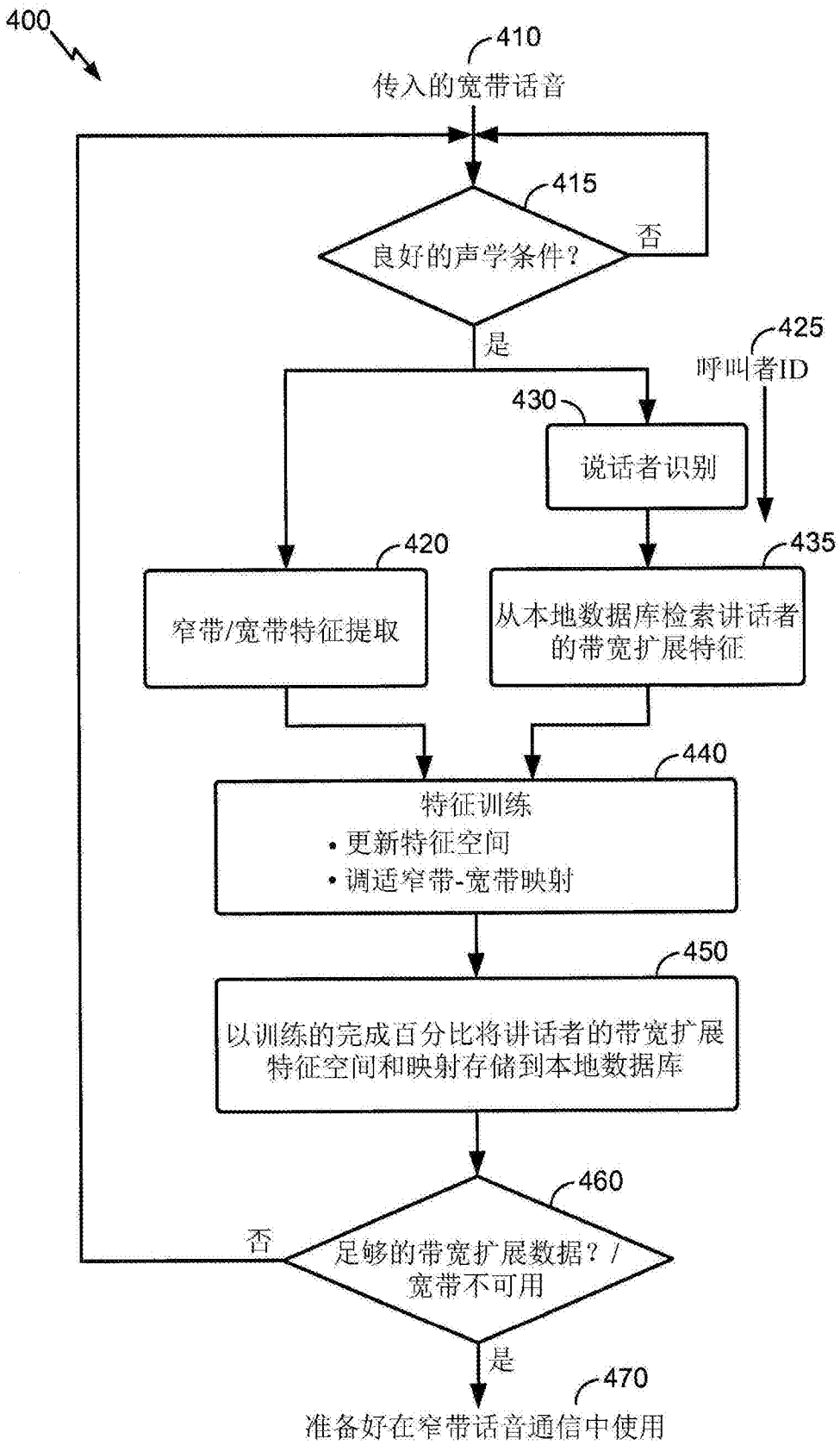

400
410
传入的宽带话音
415
良好的声学条件？
否
是
425
呼叫者ID
430
说话者识别
420
窄带/宽带特征提取
435
从本地数据库检索讲话者的带宽扩展特征
440
特征训练
• 更新特征空间
• 调适窄带-宽带映射
450
以训练的完成百分比将讲话者的带宽扩展特征空间和映射存储到本地数据库
460
否
足够的带宽扩展数据？/宽带不可用
是
470
准备好在窄带话音通信中使用

图4

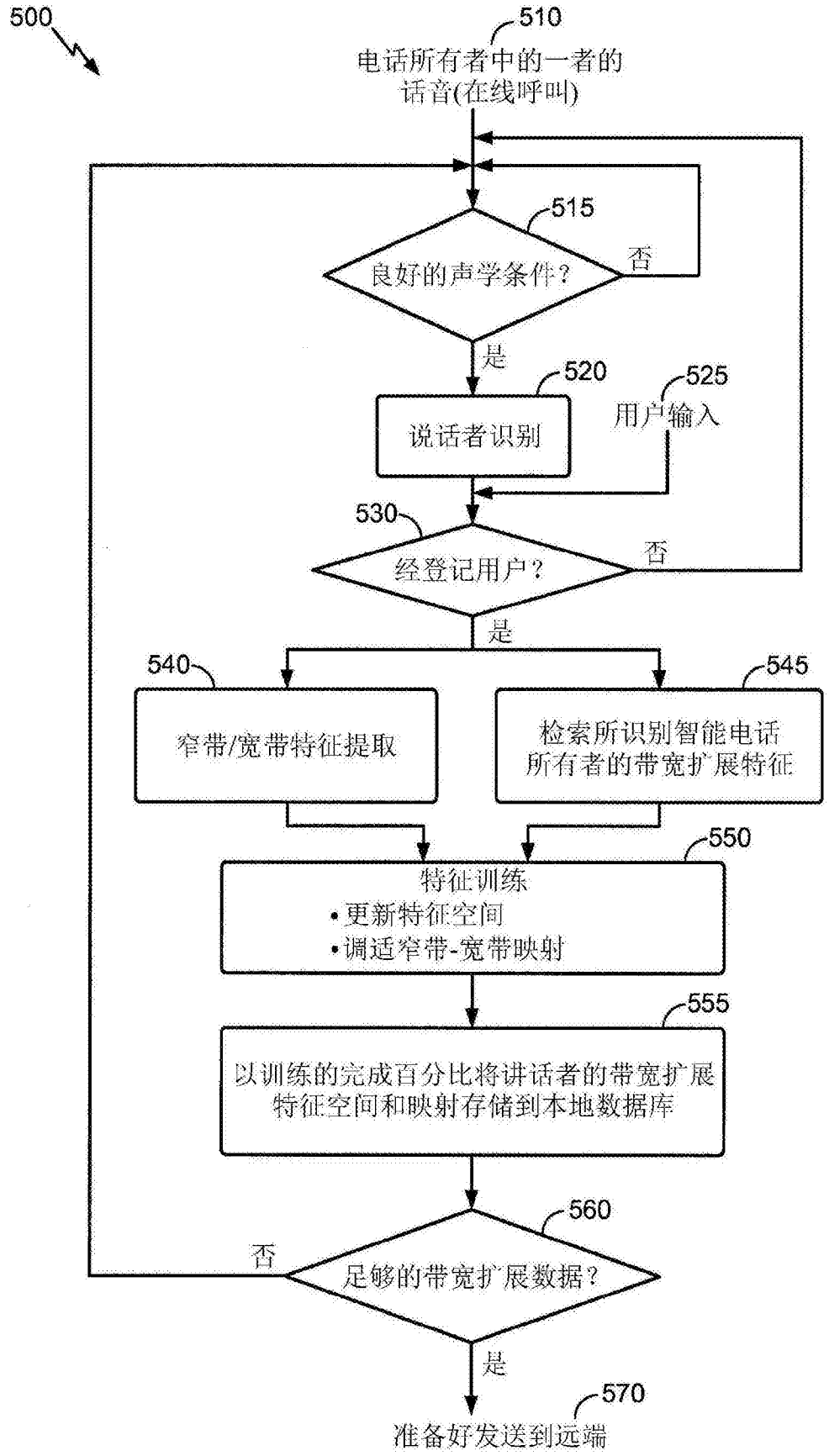

500
510
电话所有者中的一者的话音(在线呼叫)
515
良好的声学条件？
否
是
520
说话者识别
525
用户输入
530
经登记用户？
否
是
540
窄带/宽带特征提取
545
检索所识别智能电话所有者的带宽扩展特征
550
特征训练
•更新特征空间
•调适窄带-宽带映射
555
以训练的完成百分比将讲话者的带宽扩展特征空间和映射存储到本地数据库
560
足够的带宽扩展数据？
否
是
570
准备好发送到远端

图5

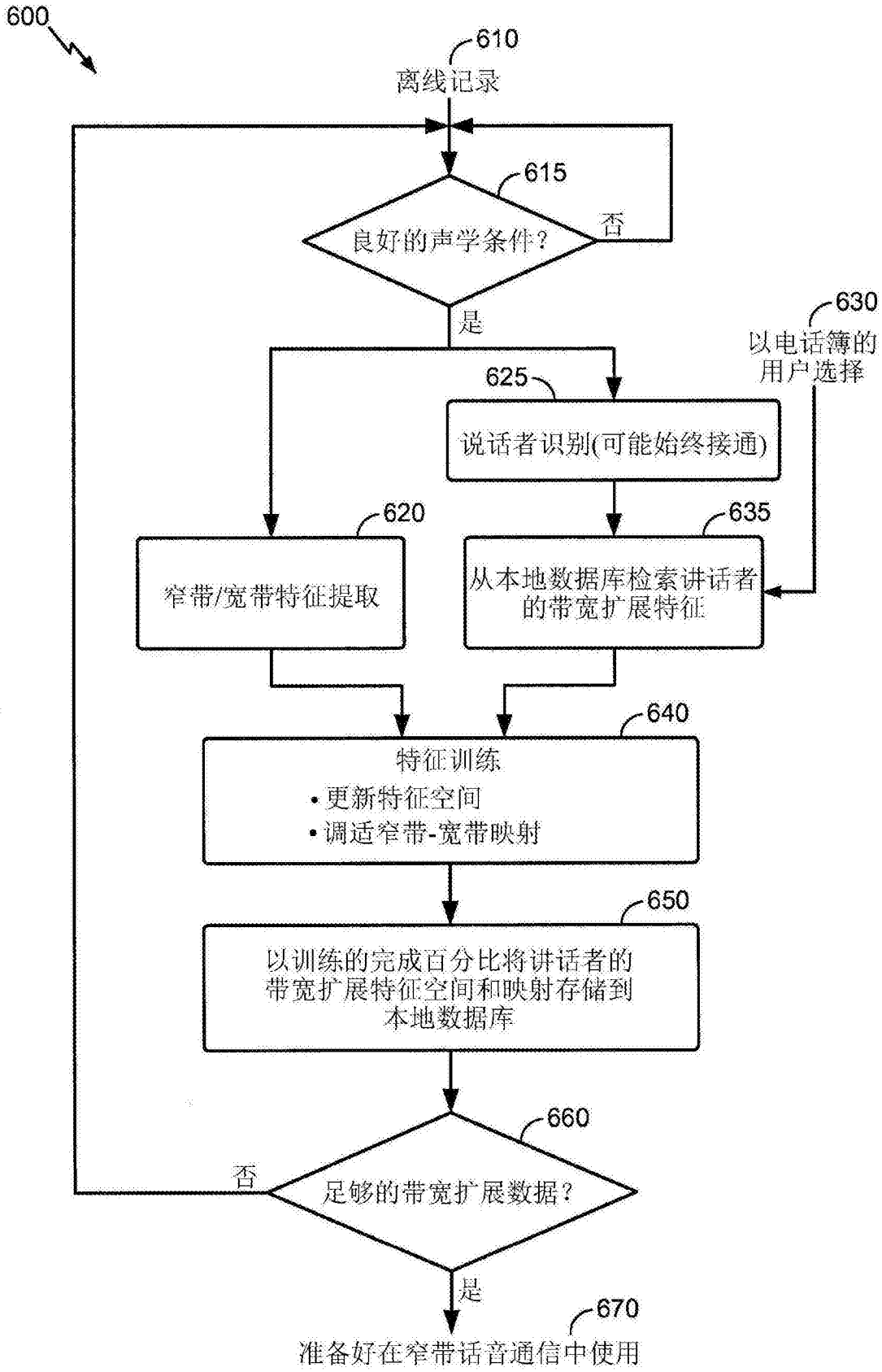

600
610
离线记录
615
良好的声学条件？
否
是
630
以电话簿的
用户选择
625
说话者识别(可能始终接通)
620
窄带/宽带特征提取
635
从本地数据库检索讲话者
的带宽扩展特征
640
特征训练
• 更新特征空间
• 调适窄带-宽带映射
650
以训练的完成百分比将讲话者的
带宽扩展特征空间和映射存储到
本地数据库
660
否
足够的带宽扩展数据？
是
670
准备好在窄带话音通信中使用

图6

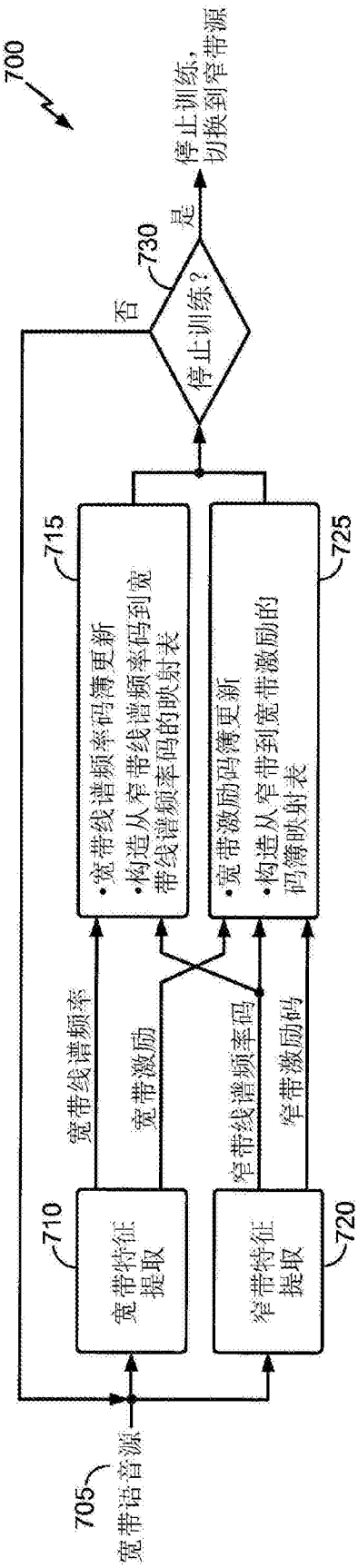

700
停止训练，
切换到窄带源
是
730
停止训练？
否
715
•宽带线谱频率码簿更新
•构造从窄带线谱频率码到宽带线谱频率码的映射表
725
•宽带激励码簿更新
•构造从窄带到宽带激励的码簿映射表
宽带线谱频率
宽带激励
窄带线谱频率码
窄带激励码
710
宽带特征提取
720
窄带特征提取
705
宽带语音源

图7

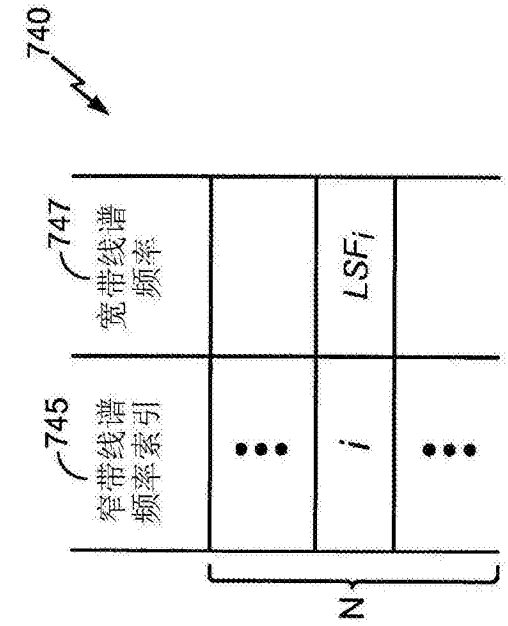

740
747
宽带线谱频率
745
窄带线谱频率索引
LSF_i
i
N

图8

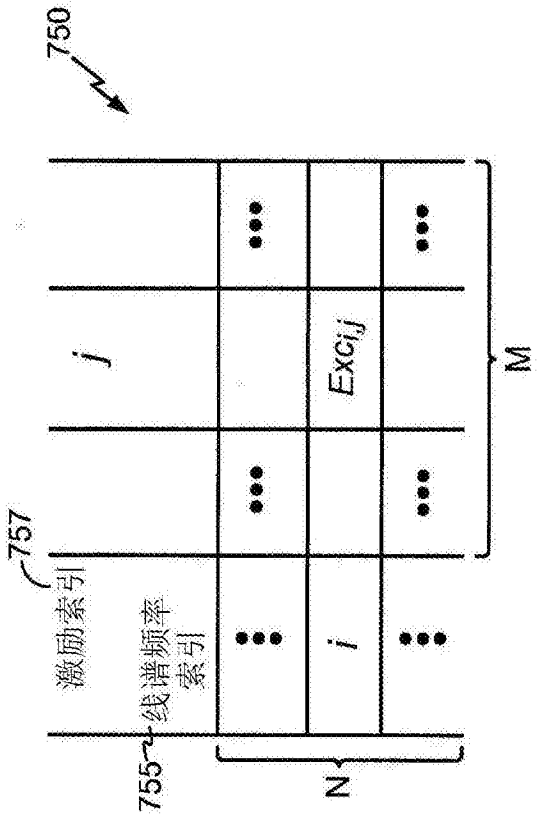

750
757
激励索引
线谱频率索引
755
j
Exc_{ij}
i
M
N

图9

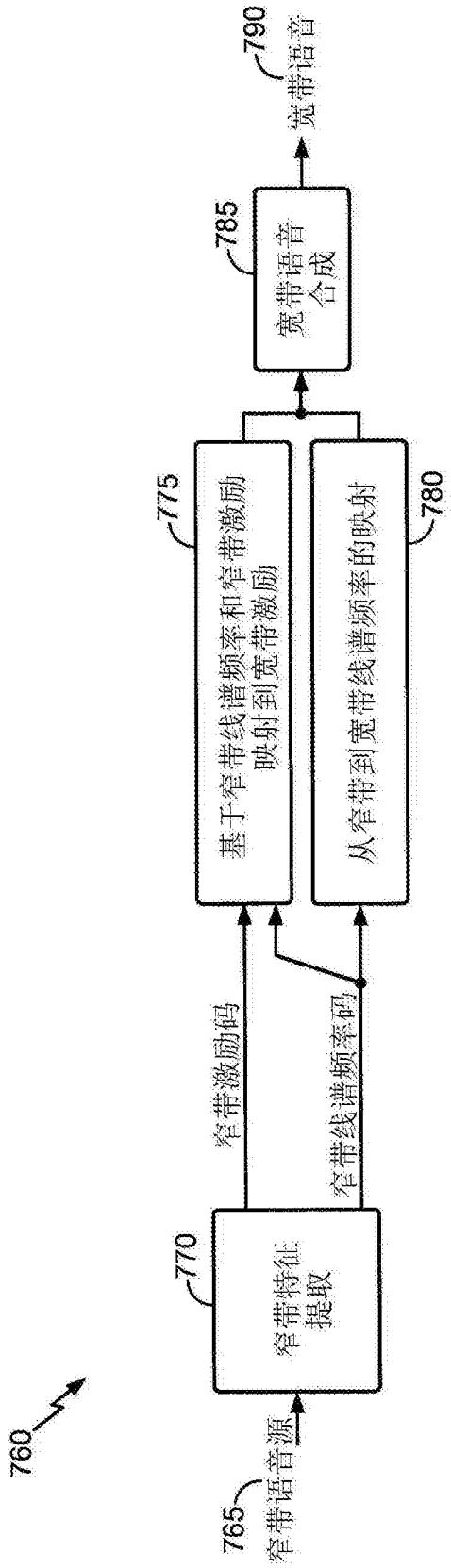

760
窄带语音源
765
窄带特征提取
770
窄带激励码
窄带线谱频率码
基于窄带线谱频率和窄带激励映射到宽带激励
775
从窄带到宽带线谱频率的映射
780
宽带语音合成
785
宽带语音
790

图10

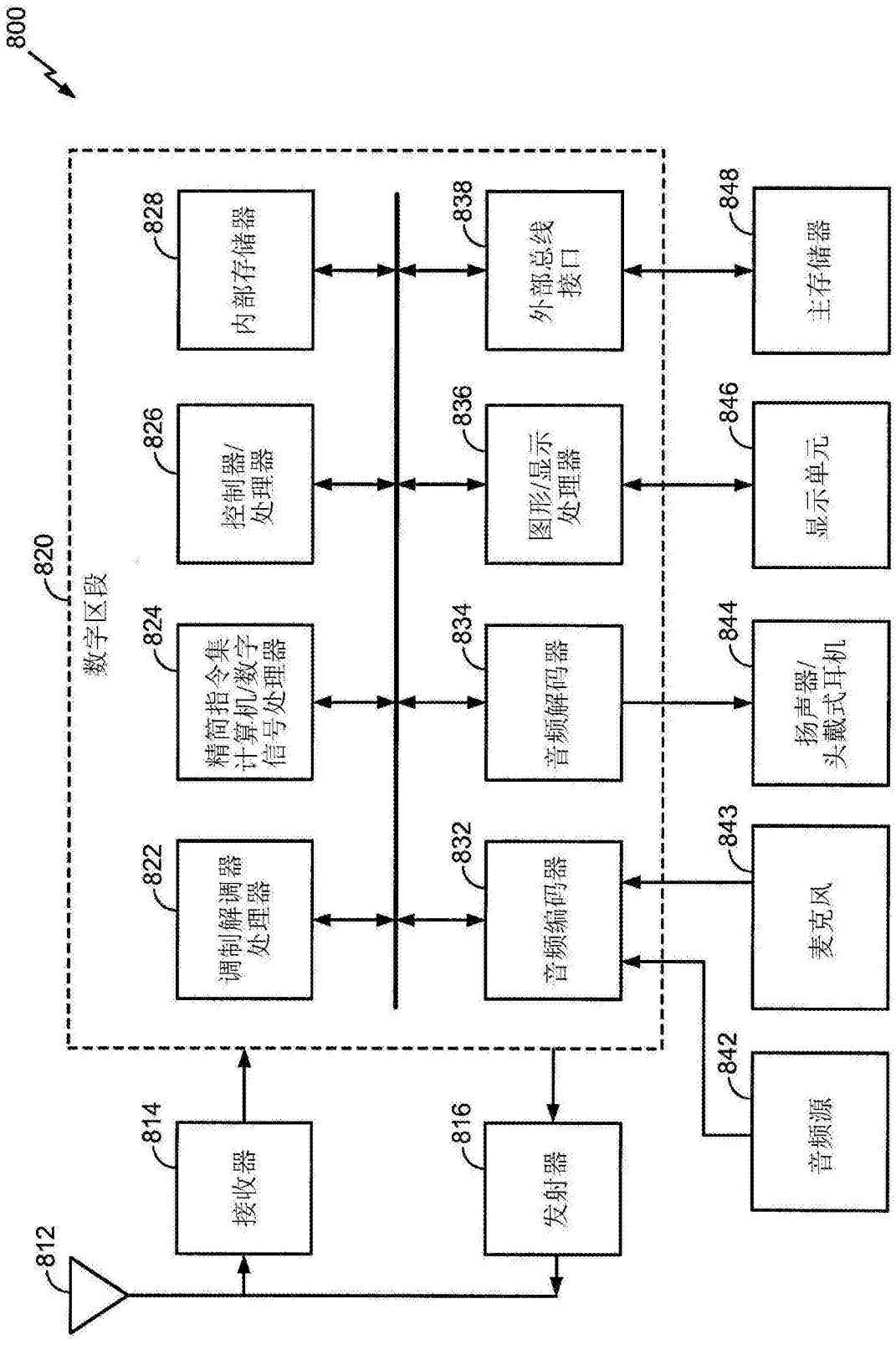

800
820
数字区段
828
内部存储器
826
控制器/处理器
824
精简指令集计算机/数字信号处理器
822
调制解调器处理器
838
外部总线接口
836
图形/显示处理器
834
音频解码器
832
音频编码器
848
主存储器
846
显示单元
844
扬声器/头戴式耳机
843
麦克风
842
音频源
814
接收器
816
发射器
812

图11

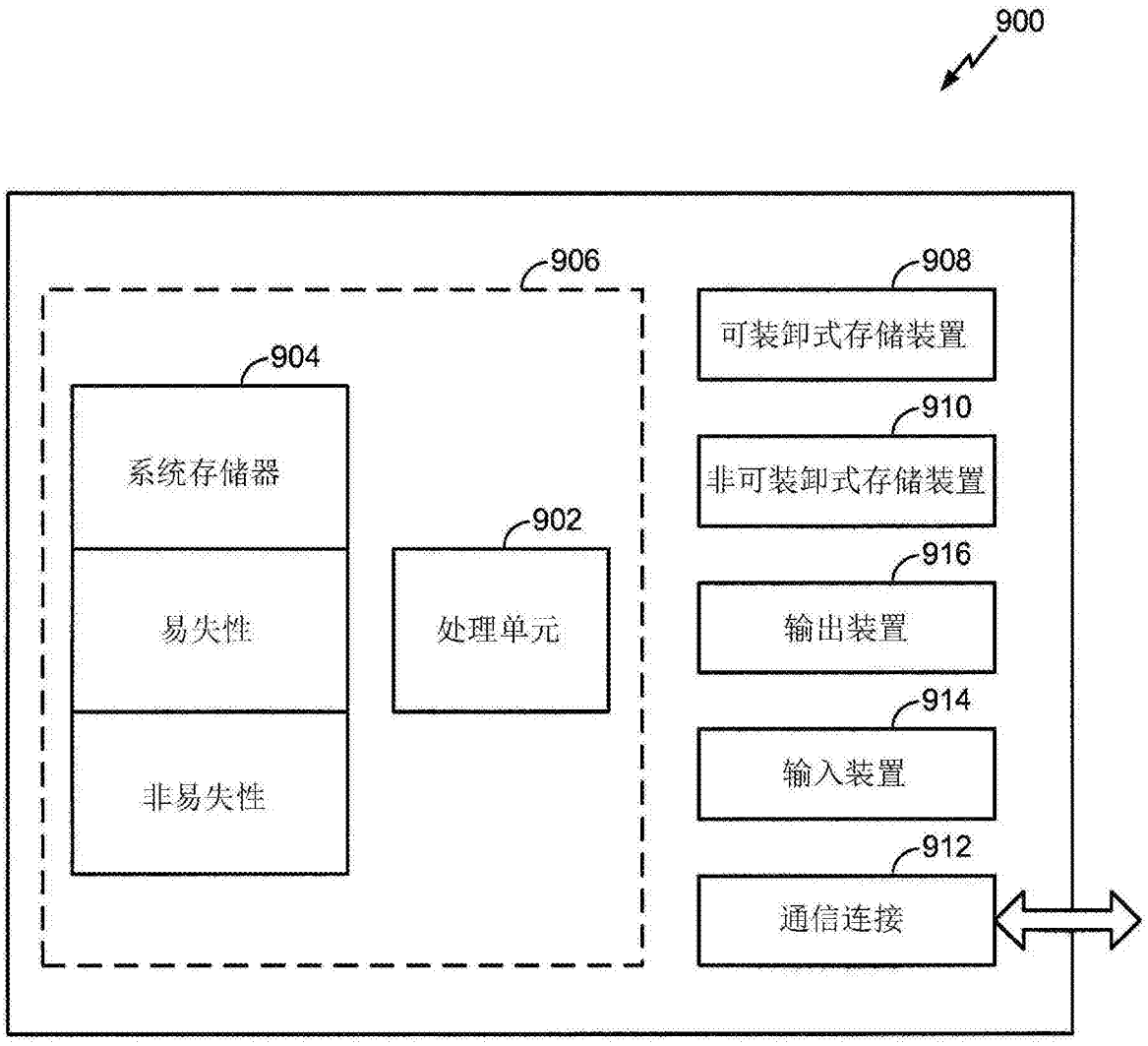

900
906
908
可装卸式存储装置
904
系统存储器
910
非可装卸式存储装置
902
处理单元
916
输出装置
易失性
914
输入装置
非易失性
912
通信连接

图12