(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 111192599 B
(45)授权公告日 2022.11.22

(21)申请号 201811352262.9

(22)申请日 2018.11.14

(65)同一申请的已公布的文献号
申请公布号 CN 111192599 A

(43)申请公布日 2020.05.22

(73)专利权人 中移(杭州)信息技术有限公司
地址 311100 浙江省杭州市余杭区文一西路998号海创园18幢6层
专利权人 中国移动通信集团有限公司

(72)发明人 宋钦梅 方华 袁其政 屈跃强 程宝平

(74)专利代理机构 北京同达信恒知识产权代理有限公司 11291
专利代理师 郭润湘

(51)Int.Cl.
G10L 21/0224(2013.01)
G10L 21/0232(2013.01)
H04M 3/18(2006.01)
H04M 1/19(2006.01)

(56)对比文件
CN 106024002 A,2016.10.12
CN 108346433 A,2018.07.31
CN 107845389 A,2018.03.27
US 2018033449 A1,2018.02.01
CN 105023580 A,2015.11.04

审查员 张岩

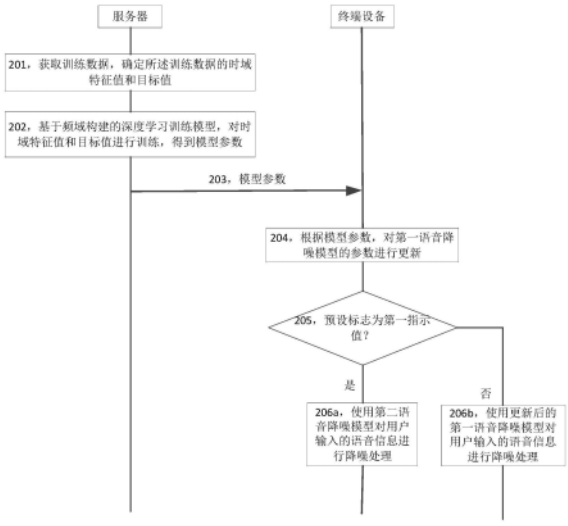
权利要求书2页 说明书9页 附图3页

(54)发明名称

一种降噪方法及装置

(57)摘要

本申请实施例公开了一种降噪方法及装置，其中方法包括：通过基于频域构建的深度学习训练模型对训练数据的时域特征值和目标值进行训练，服务器可以将得到的模型参数发送给终端设备，以使终端设备接收到该模型参数后，对第一语音降噪模型的参数进行更新，并使用更新后的第一语音降噪模型对用户输入的语音信息进行降噪处理。本申请实施例中，通过使用深度学习训练模型训练得到的模型参数对终端设备中的第一语音降噪模型的参数进行更新，可以使得终端设备采用深度学习训练模型训练得到的模型参数对用户输入的语音信息进行降噪处理，从而能够使降噪处理得到的语音信息更加准确，提高用户的体验感。



CN 111192599 B

1.一种降噪方法,其特征在于,所述方法包括:

服务器获取训练数据,所述训练数据包括在第一环境中收集的第一语音信息和在第二环境中收集的第二语音信息,所述第一环境中的噪声小于等于预设阈值,所述第二环境中的噪声大于所述预设阈值;

所述服务器根据所述训练数据,确定所述训练数据的时域特征值和目标值;所述训练数据的时域特征值包括噪声阈值、长时能量值、短时能量值和噪声包络跟踪值,所述噪声阈值用于表征所述训练数据的噪声的幅值范围;所述训练数据的目标值包括所述第一语音信息的语音活动检测值和所述第二语音信息的全带信噪比;

所述服务器基于频域构建的深度学习训练模型,将所述时域特征值输入所述深度学习训练模型的语音活动检测模块得到第一模型参数,将所述第一模型参数输入所述深度学习训练模型的噪声谱估计模块,得到第二模型参数,将所述第一模型参数输入所述深度学习训练模型的谱减模块,得到第三模型参数;

所述服务器将所述第一模型参数、所述第二模型参数和所述第三模型参数共同输入所述谱减模块,在所述谱减模块输出所述目标值时,得到模型参数,并将所述模型参数发送给终端设备,所述模型参数用于所述终端设备对用户输入的语音信息进行降噪处理。

2.一种降噪方法,其特征在于,所述方法包括:

终端设备接收服务器发送的模型参数,所述模型参数为所述服务器基于频域构建的深度学习训练模型,将训练数据的时域特征值输入所述深度学习训练模型,且输出为所述训练数据的目标值时得到的;

所述终端设备根据所述模型参数,对所述终端设备中的第一语音降噪模型的参数进行更新,得到更新后的第一语音降噪模型,所述第一语音降噪模型是对所述深度学习训练模型进行模型训练获取所述模型参数前的模型;

所述终端设备在接收到用户输入的语音信息后,使用所述更新后的第一语音降噪模型对所述语音信息进行降噪处理。

3.根据权利要求2所述的方法,其特征在于,所述终端设备接收服务器发送的模型参数之后,还包括:

所述终端设备将预设标志更新为第一指示值;

所述终端设备得到更新后的第一语音降噪模型之后,还包括:

所述终端设备将所述预设标志更新为第二指示值。

4.根据权利要求3所述的方法,其特征在于,所述终端设备在接收到用户输入的语音信息后,使用所述更新后的第一语音降噪模型对所述语音信息进行更新之前,还包括:

所述终端设备确定所述预设标志为所述第二指示值。

5.根据权利要求3所述的方法,其特征在于,所述方法还包括:

所述终端设备在接收到用户输入的语音信息后,若确定所述预设标志为所述第一指示值,则使用所述终端设备中的第二语音降噪模型对所述语音信息进行降噪处理;所述第二语音降噪模型为所述第一语音降噪模型的备用模型。

6.一种服务器,其特征在于,所述服务器包括:

获取模块,用于获取训练数据,所述训练数据包括在第一环境中收集的第一语音信息和在第二环境中收集的第二语音信息,所述第一环境中的噪声小于等于预设阈值,所述第

二环境中的噪声大于所述预设阈值；

确定模块，用于根据所述训练数据，确定所述训练数据的时域特征值和目标值，所述训练数据的时域特征值包括噪声阈值、长时能量值、短时能量值和噪声包络跟踪值，所述噪声阈值用于表征所述训练数据的噪声的幅值范围；所述训练数据的目标值包括所述第一语音信息的语音活动检测值和所述第二语音信息的全带信噪比；

处理模块，用于基于频域构建的深度学习训练模型，将所述时域特征值输入所述深度学习训练模型的语音活动检测模块得到第一模型参数，将所述第一模型参数输入所述深度学习训练模型的噪声谱估计模块，得到第二模型参数，将所述第一模型参数输入所述深度学习训练模型的谱减模块，得到第三模型参数；

所述服务器将所述第一模型参数、所述第二模型参数和所述第三模型参数共同输入所述谱减模块，在所述谱减模块输出所述目标值时，得到模型参数，并将所述模型参数发送给终端设备，以使所述终端设备使用所述模型参数对用户输入的语音信息进行降噪处理。

7.一种终端设备，其特征在于，所述终端设备包括：

收发模块，用于接收服务器发送的模型参数，所述模型参数为所述服务器基于频域构建的深度学习训练模型，将训练数据的时域特征值输入所述深度学习训练模型，且输出为所述训练数据的目标值时得到的；

更新模块，用于根据所述模型参数，对所述终端设备中的第一语音降噪模型的参数进行更新，得到更新后的第一语音降噪模型，所述第一语音降噪模型是对所述深度学习训练模型进行模型训练获取所述模型参数前的模型；

降噪模块，用于在接收到用户输入的语音信息后，使用所述更新后的第一语音降噪模型对所述语音信息进行降噪处理。

8.根据权利要求7所述的终端设备，其特征在于，在所述收发模块接收到服务器发送的模型参数之后，所述更新模块还用于：将预设标志更新为第一指示值，以及在得到更新后的第一语音降噪模型之后，将所述预设标志更新为第二指示值。

9.根据权利要求8所述的终端设备，其特征在于，所述降噪模块还用于：

确定所述预设标志为所述第二指示值。

10.根据权利要求9所述的终端设备，其特征在于，所述降噪模块还用于：

在接收到用户输入的语音信息后，若确定所述预设标志为所述第一指示值，则使用所述终端设备中的第二语音降噪模型对所述语音信息进行降噪处理；所述第二语音降噪模型为所述第一语音降噪模型的备用模型。

一种降噪方法及装置

技术领域

[0001] 本申请涉及通信技术领域，尤其涉及一种降噪方法及装置。

背景技术

[0002] 现实生活中，用户发出的语音信息中通常含有噪声，比如，环境中的风声、汽车声、机器运转的声音等。在用户使用语音装置进行通话的过程中，这些噪声可能会影响用户的通话质量，使得用户的体验不好。举个例子，用户A和用户B通过终端设备(比如手机)进行通话，若用户A通过手机a发出的语音信息中包含的噪声较大，可能会使得用户B通过手机b无法正常获取到用户A的语音信息，比如获取到的语音信息不够清晰，或者获取不到用户A发出的语音信息。

[0003] 因此，目前亟需一种降噪方法，用以解决因存在噪声而导致用户之间通话质量较低的技术问题。

发明内容

[0004] 本申请实施例提供一种降噪方法及装置，用以解决因存在噪声而导致用户之间通话质量较低的技术问题。

[0005] 本申请实施例提供的一种降噪方法，包括：

[0006] 服务器获取训练数据，所述训练数据包括在第一环境中收集的第一语音信息和在第二环境中收集的第二语音信息，所述第一环境中的噪声小于等于预设阈值，所述第二环境中的噪声大于所述预设阈值；

[0007] 所述服务器根据所述训练数据，确定所述训练数据的时域特征值和目标值；所述训练数据的时域特征值包括噪声阈值、长时能量值、短时能量值和噪声包络跟踪值中的一项或多项；所述训练数据的目标值包括所述第一语音信息的语音活动检测值和/或所述第二语音信息的全带信噪比；

[0008] 所述服务器基于频域构建的深度学习训练模型，对所述时域特征值和所述目标值进行训练，得到模型参数，并将所述模型参数发送给终端设备；所述模型参数用于所述终端设备对用户输入的语音信息进行降噪处理。

[0009] 本申请实施例提供的一种降噪方法，包括：

[0010] 终端设备接收服务器发送的模型参数；

[0011] 所述终端设备根据所述模型参数，对所述终端设备中的第一语音降噪模型的参数进行更新，得到更新后的第一语音降噪模型；

[0012] 所述终端设备在接收到用户输入的语音信息后，使用所述更新后的第一语音降噪模型对所述语音信息进行降噪处理。

[0013] 可选地，所述终端设备接收服务器发送的模型参数之后，还包括：

[0014] 所述终端设备将预设标志更新为第一指示值；

[0015] 所述终端设备得到更新后的第一语音降噪模型之后，还包括：

[0016] 所述终端设备将所述预设标志更新为第二指示值。

[0017] 可选地，所述终端设备在接收到用户输入的语音信息后，使用所述更新后的第一语音降噪模型对所述语音信息进行更新之前，还包括：

[0018] 所述终端设备确定所述预设标志为所述第二指示值。

[0019] 可选地，所述方法还包括：

[0020] 所述终端设备在接收到用户输入的语音信息后，若确定所述预设标志为所述第一指示值，则使用所述终端设备中的第二语音降噪模型对所述语音信息进行降噪处理；所述第二语音降噪模型为所述第一语音降噪模型的备用模型。

[0021] 本申请实施例提供的一种服务器，该服务器包括：

[0022] 获取模块，用于获取训练数据，所述训练数据包括在第一环境中收集的第一语音信息和在第二环境中收集的第二语音信息，所述第一环境中的噪声小于等于预设阈值，所述第二环境中的噪声大于所述预设阈值；

[0023] 确定模块，用于根据所述训练数据，确定所述训练数据的时域特征值和目标值，所述训练数据的时域特征值包括噪声阈值、长时能量值、短时能量值和噪声包络跟踪值中的一项或多项；所述训练数据的目标值包括所述第一语音信息的语音活动检测值和/或所述第二语音信息的全带信噪比；

[0024] 处理模块，用于基于频域构建的深度学习训练模型，对所述时域特征值和所述目标值进行训练，得到模型参数，并将所述模型参数发送给终端设备，以使所述终端设备使用所述模型参数对用户输入的语音信息进行降噪处理。

[0025] 本申请实施例提供的一种终端设备，该终端设备包括：

[0026] 收发模块，用于接收服务器发送的模型参数；

[0027] 更新模块，用于根据所述模型参数，对所述终端设备中的第一语音降噪模型的参数进行更新，得到更新后的第一语音降噪模型；

[0028] 降噪模块，用于在接收到用户输入的语音信息后，使用所述更新后的第一语音降噪模型对所述语音信息进行降噪处理。

[0029] 可选地，在所述收发模块接收到服务器发送的模型参数之后，所述更新模块还用于：将预设标志更新为第一指示值，以及在得到更新后的第一语音降噪模型之后，将所述预设标志更新为第二指示值。

[0030] 可选地，所述降噪模块还用于：

[0031] 确定所述预设标志为所述第二指示值。

[0032] 可选地，所述降噪模块还用于：

[0033] 在接收到用户输入的语音信息后，若确定所述预设标志为所述第一指示值，则使用所述终端设备中的第二语音降噪模型对所述语音信息进行降噪处理；所述第二语音降噪模型为所述第一语音降噪模型的备用模型。

[0034] 本申请的上述实施例中，服务器通过将收集到的第一环境(噪声小于等于预设阈值)中的第一语音信息和第二环境(噪声大于预设阈值)中的第二语音信息作为训练数据，并在确定训练数据的时域特征值和目标值后，基于频域构建的深度学习训练模型对时域特征值和目标值进行训练，得到模型参数，以及将模型参数发送给终端设备，使得终端设备可以在接收到服务器发送的模型参数后，对终端设备中的第一语音降噪模型的参数进行更

新，并可以使用更新后的第一语音降噪模型对用户输入的语音信息进行降噪处理。本申请实施例中，通过获取训练数据的时域特征值(可以体现时域特征)和目标值，并采用基于频域构建的深度学习训练模型(可以体现频域特征)对训练数据的时域特征值和目标值进行训练，可以在训练模型的过程中将语音信息的时域特征和频域特征进行结合，进而提升深度学习训练模型的训练性能，加快深度学习训练模型的训练速度；且，由于训练数据可以包括多种环境下的语音信息，并可以多次采用不同的训练数据对深度学习训练模型进行训练，因此可以得到比较准确的模型参数；进一步地，通过使用深度学习训练模型训练得到的模型参数对终端设备中的第一语音降噪模型的参数进行更新，可以使得终端设备采用深度学习训练模型训练得到的模型参数对用户输入的语音信息进行降噪处理，从而能够使降噪处理得到的语音信息更加准确，提高用户的体验感。此外，采用深度学习训练模型对训练数据进行训练获取模型参数的过程和终端设备对用户语音信息的降噪过程可以并行处理，从而能够提高语音降噪的处理速度和处理效率。

附图说明

[0035] 为了更清楚地说明本申请实施例中的技术方案，下面将对实施例描述中所需要使用的附图作简要介绍，显而易见地，下面描述中的附图仅仅是本申请的一些实施例，对于本领域的普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

[0036] 图1为本申请实施例提供的一种可能的系统架构示意图；

[0037] 图2为本申请实施例提供的一种降噪方法对应的流程示意图；

[0038] 图3为本申请实施例中提供的一种服务器的结构示意图；

[0039] 图4为本申请实施例中提供的一种终端设备的结构示意图。

具体实施方式

[0040] 为了使本申请的目的、技术方案和优点更加清楚，下面将结合附图对本申请作进一步地详细描述，显然，所描述的实施例仅仅是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其它实施例，都属于本申请保护的范围。

[0041] 图1为本申请实施例提供的一种可能的系统架构，该系统架构中可以包括网络设备100、一个或多个终端设备(比如，图1中示意出的终端设备101和终端设备102)，终端设备101和终端设备102可以通过网络设备100进行通信。其中，网络设备100可以为基站，比如长期演进(Long Term Evolution，LTE)通信系统中的演进型基站(比如，evolved NodeB)。终端设备可以为手机(mobile phone)、平板电脑(Pad)等，具体不做限定。

[0042] 本申请实施例中，用户A和用户B可以通过终端设备进行通话，其中，用户A使用终端设备101，用户B使用终端设备102，一种可能的场景为，用户A向终端设备101输入语音信息，终端设备101通过网络设备100将用户A输入的语音信息传输给终端设备102，终端设备102在接收到语音信息后，将接收到的语音信息提供给用户B。

[0043] 为了提高通话质量，一种可能的实现方式(为便于描述简称实现方式a)为，采用预设噪声信号对用户发出的语音信息进行降噪处理，即过滤掉用户发出的语音信息中的预设

噪声信号，得到过滤后的语音信息。采用这种语音降噪方法，可以对用户发出的语音信息中的预设噪声信号进行过滤。然而，该种方法所采用的预设噪声信号是固定的，也就是说，终端设备可能会采用该预设噪声信号对所有后续用户发出的语音信息进行降噪处理，而不会对终端设备中的模型参数进行更新。在某种情况下，采用预设噪声信号对用户在多种不同环境下的语音信息进行降噪处理得到的语音信息可能会不准确。

[0044] 基于此，本申请实施例提供一种语音降噪方法，用以解决因存在噪声而导致用户之间通话质量较低的技术问题。

[0045] 图2为本申请实施例提供的一种降噪方法对应的流程示意图，该方法包括：

[0046] 步骤201，服务器获取训练数据，并确定训练数据的时域特征值和目标值。

[0047] 此处，训练数据可以包括在第一环境中收集的第一语音信息和在第二环境中收集的第二语音信息，第一环境中的噪声可以小于等于预设阈值，也就是说，第一环境中的噪声较小，第二环境中的噪声可以大于预设阈值，也就是说，第一环境中的噪声较大。其中，预设阈值可以是由本领域技术人员根据经验进行设置得到的，或者也可以是通过一次或多次实验确定的，本申请实施例对此不作限定。

[0048] 本申请实施例中，训练数据的获取方式可以有多种，在一种可能的实现方式中，可以由测试人员携带语音采集装置分别在第一环境和第二环境中收集第一语音信息和第二语音信息。举个例子，以预设阈值为-30分贝为例，若应用场景为预设楼层的某个房间，则可以将语音采集装置采集到的测试人员在该房间噪声小于等于-30分贝(或者没有噪声)的第一位置发出的语音信息作为第一语音信息，且可以将语音采集装置采集到的测试人员在该房间噪声大于-30分贝的第二位置发出的相同的语音信息作为第二语音信息。比如，用户在室内远离窗户的A位置对语音采集装置说“测试室内环境的噪声”(为便于描述，简称语音1)，并在窗户旁边的位置B对语音采集装置说“测试室内环境的噪声”(为便于描述，简称语音2)，则第一语音信息可以为语音1，第二语音信息可以为语音2。

[0049] 本申请实施例中，语音信息作为一种信号，一般可以具有时域特征和频域特征。具体地说，语音信息的时域特征可以由语音信息的幅值维度和语音信息的时间维度来描述，比如，可以表示为幅值随时间变化的函数；通过语音信息的时域特征，可以获取到语音信息在某个时间段内的能量值、语音信息在某个时间点的幅值等信息。语音信息的频域特征可以由语音信息的幅值维度和语音信息的频率维度来描述，比如，可以表示为幅值随频率变化的函数。语音信息的时域特征与频域特征之间可以进行转换，在一个示例中，可以通过傅里叶变换将随时间变化的语音信息拆解为多个具有不同频率的语音信息。

[0050] 本申请实施例中，服务器可以获取语音采集装置采集的训练数据，并确定训练数据的时域特征值，训练数据的时域特征值可以用于指示训练数据在时间域上的特征，在一种可能的实现方式中，训练数据的时域特征值可以包括噪声阈值、长时能量值、短时能量值和噪声包络跟踪值中的一项或多项。可以理解的，训练数据的时域特征值还可以包括其他的信息，本申请实施例对此不作限定。

[0051] 具体实施中，噪声阈值可以用于指示噪声的幅值范围，在一个示例中，可以预先设置一个初始噪声阈值(比如，-40分贝)，此时，训练数据中幅值小于-40分贝的语音信息可以认定为噪声，幅值大于-40分贝的语音信息可以认定为用户的语音。具体实施中，不同的环境可以对应有不同的初始噪声阈值，比如，办公环境对应的初始噪声阈值可以为-40分贝，

商场对应的初始噪声阈值可以为-30分贝。本申请实施例中，每一个环境对应的初始噪声阈值可以由本领域技术人员根据经验进行设置，或者也可以由本领域技术人员通过实验进行确定，对此不作限定。

[0052] 长时能量值和短时能量值可以用于指示语音信息在预设时间段内的能量信息。下面以第一语音信息为例描述获取第一语音信息的长时能量值和短时能量值的实现过程，获取第二语音信息的长时能量值和短时能量值的过程可以参照第一语音信息来处理。

[0053] 本申请实施例中，可以预先根据第一语音信息的总时长对第一语音信息进行分帧操作，得到多帧语音信息。其中，多帧语音信息中任意两帧语音信息的时长可以相同，或者也可以不同。具体实施中，将第一语音信息划分为多帧语音信息的方式可以有多种，在一种可能的实现方式中，可以预先设置一个标准帧时长(比如，10ms)，从而根据标准帧时长将第一语音信息划分为多帧语音信息。在一个示例中(为便于描述，简称示例一)，第一语音信息的总时长为205ms，则根据标准帧时长划分后，可以得到21帧语音信息，其中第1帧语音信息至第20帧语音信息的时长和标准帧时长相同，均为10ms，第21帧语音信息的时长为5ms。

[0054] 进一步地，可以设置长时能量值为A帧(比如，A为3)语音信息的总能量的平均值，并可以设置短时能量值为最近B帧(比如，B为1)语音信息的能量值。基于示例一中的多帧语音信息和每帧语音信息的时长，则长时能量值可以为30ms时长的语音信息的总能量的平均值(若包括第21帧语音信息，则可以为25ms时长的语音信息的总能量的平均值)；相应地，短时能量值可以为第21帧语音信息的能量值。需要说明的是，本申请实施例中，A和B的数值可以由本领域技术人员根据实际情况进行调整，对此不作限定。

[0055] 噪声包络跟踪值可以用于对噪声的幅值进行估计。一般来说，噪声可能会比语音具有更宽的时间特性，因此，可以通过跟踪第二语音信息中每帧语音信息对应的最小幅值来获取噪声幅值的估计值。比如，可以预先提取第二语音信息中每帧语音信息对应的最小值，并据此绘制噪声包络跟踪图，并可以根据噪声包络跟踪图和预设指标(比如，快降慢升的原则)，计算得到估计的噪声信息(比如，长时能量值对应的噪声包络值、短时能量值对应的噪声包络值等)。

[0056] 需要说明的是，本申请实施例中，噪声阈值的大小还可以根据噪声包络跟踪值进行调整。在一个示例中，可以预设一个增量(比如，2分贝)，若计算得到的噪声包络跟踪值大于初始噪声阈值(比如，-40分贝)，则可以按照预设增量依次增大噪声阈值。举个例子，若第一次计算得到的噪声包络跟踪值为-35分贝，则可以将噪声阈值调整为-33分贝，并可以基于-33分贝的噪声阈值对噪声包络跟踪值进行计算。

[0057] 本申请实施例中，服务器还可以确定训练数据的目标值，训练数据的目标值可以包括第一语音信息的语音活动检测值和/或第二语音信息的全带信噪比。其中，第一语音信息的语音活动检测值可以用于指示检测到的是语音还是噪声，在一个示例中，可以预设第一指示值(1)和第二指示值(0)，比如，“1”可以用于指示当前检测到的是语音，“0”可以用于指示当前检测到的是噪声，或者，“0”可以用于指示当前检测到的是语音，“1”可以用于指示当前检测到的是噪声，本申请实施例对此不做限定。第二语音信息的全带信噪比可以用于指示语音和噪声的对应关系，在一个示例中，第二语音信息的全带信噪比可以为第二语音信息的平均功率和第二语音信息相对应的噪声的平均功率的比值。

[0058] 步骤202，服务器基于频域构建的深度学习训练模型，对时域特征值和目标值进行

训练,得到模型参数。

[0059] 此处,深度学习训练模型的构建方法可以有多种,在一种可能的实现方式中,可以基于keras构建深度学习训练模型。具体地说,Keras是一个基于Theano的高度模块化的神经网络库,比如,Keras可以基于Torch并可以采用Python语言进行编写,且Keras可以支持图形处理器(Graphics Processing Unit,GPU)和中央处理器(Central Processing Unit,CPU)。

[0060] 本申请实施例中,深度学习训练模型可以包括语音活动检测模块、噪声谱估计模块和谱减模块。其中,语音活动检测模块可以通过检测第一语音信息和第二语音信息,并根据检测到的第一语音信息和第二语音信息的活动标志(比如,幅值范围等)来区分语音和幅值。噪声谱估计模块可以用于对第一语音信息和第二语音信息进行计算,并可以根据计算得到的结果对噪声的频谱特性进行估计。谱减模块可以用于根据语音活动检测模块和噪声谱估计模块得到的计算结果,确定增益值,该增益值可以用于对语音信息中的噪声进行抑制。

[0061] 具体实施中,服务器可以将训练数据的时域特征值作为深度学习训练模型的输入信息,并可以将训练数据的目标值作为深度学习训练模型的输出信息,进而控制深度学习训练模型根据输入信息和输出信息进行模型训练,得到模型参数。在一个示例中,服务器通过将时域特征值输入语音活动检测模块,可以得到第一模型参数;进一步地,服务器通过将第一模型参数输入噪声谱估计模块,可以得到第二模型参数,同时,服务器通过将第一模型参数输入谱减模块,可以得到第三模型参数;最后,服务器通过将第一模型参数、第二模型参数和第三模型参数共同输入谱减模块,可以得到训练数据经过深度学习训练模型训练后的模型参数。需要说明的是,本申请实施例中,深度学习训练模型中的各个模块可以为通过Keras构建的功能模块,也就是说,语音活动检测模块、噪声谱估计模块和谱减模块仅为对确定模型参数的过程进行描述而引入的,具体实施中,还可以包括其它模块,具体不做限定。且,各个功能模块的名称也可以为其它可能的名称,具体不做限定。

[0062] 步骤203,将模型参数发送给终端设备。

[0063] 此处,服务器可以通过与终端设备进行通信,将模型参数发送给终端设备,其中,服务器与终端设备进行通信的方式可以有多种,在一个示例中,服务器可以通过无线方式与终端设备进行通信;在又一个示例中,服务器也可以通过有线(比如,光纤、网线等)与终端设备进行通信,本申请实施例对此不作限定。

[0064] 步骤204,终端设备接收服务器发送的模型参数,并使用该模型参数对终端设备中的第一语音降噪模型的参数进行更新。

[0065] 本申请实施例中,终端设备中可以预先设置有第一语音降噪模型和第二语音降噪模型,第二语音降噪模型可以为第一语音降噪模型的备用模型。其中,第一语音降噪模块和第二语音降噪模型在初始状态(比如,在终端设备出厂时,或者终端设备被初始化后)的参数可以为未对基于Kears搭建的深度语音学习训练模型进行模型训练前的参数。

[0066] 具体实施中,终端设备中可以设置有一个预设标志,该预设标志可以用于指示终端设备中的第一语音降噪模型是否处于更新状态。在一个示例中,终端设备在接收到服务器发送的模型参数,并在使用该模型参数对终端设备中的第一语音降噪模型的参数进行更新之前,终端设备可以将预设标志更新为第一指示值,第一指示值可以用于指示终端设备

中的第一语音降噪模型处于更新状态。进一步地，终端设备可以在得到更新后的第一语音降噪模型之后，将预设标志更新为第二指示值，第二指示值可以用于指示终端设备中的第一语音降噪模型处于未更新状态，或者可以用于指示终端设备中的第一语音降噪模型已更新完成。本申请实施例中，预设标志可以通过一个比特位来表示，比如，第一指示值可以为“0”，第二指示值可以为“1”；或者，第一指示值可以为“1”，第二指示值可以为“0”，具体不做限定。

[0067] 本申请实施例中，若终端设备检测到第一语音降噪模型由于某些原因(比如，终端设备的某些硬件损坏或者第一语音降噪模型的更新算法出错等)无法进行降噪处理，则终端设备也可以将预设标志更新为第一指示值。

[0068] 需要说明的是，本申请实施例中，终端设备可以在得到更新后的第一语音降噪模型后，将第二语音降噪模块更新为第一语音降噪模块，比如，可以将第二语音降噪模块的参数更新为第一语音降噪模块的参数。

[0069] 步骤205，终端设备在接收到用户输入的语音信息后，若确定预设标志为第一指示值，则可以执行步骤206a；若确定预设标志为第二指示值，则可以执行步骤206b；若确定预设标志为第二指示值，则可以执行步骤206b。

[0070] 步骤206a，终端设备使用第二语音降噪模型对语音信息进行降噪处理。

[0071] 此处，终端设备若确定预设标志为第一指示值，说明终端设备中的第一语音降噪模型处于更新状态，或者第一语音降噪模型无法进行降噪处理，此时，终端设备可以控制终端设备中的第二语音降噪模型对用户输入的语音信息进行降噪处理。

[0072] 具体地说，第二语音降噪模型的参数可以为终端设备接收到的服务器上一次进行模型训练所得到的模型参数(或者在初始状态下可以为传统语音降噪模型的模型参数)，因此，终端设备可以采用上一次得到的模型参数(或者传统语音降噪模型的模型参数)对用户输入的语音信息进行降噪处理。

[0073] 步骤206a，终端设备使用更新后的第一语音降噪模型对语音信息进行降噪处理。

[0074] 此处，终端设备确定预设标志为第二指示值，说明终端设备中的第一语音降噪模型已更新完成，此时，终端设备可以控制更新后的第一语音降噪模型对用户输入的语音信息进行降噪处理。也就是说，终端设备可以采用服务器进行模型训练得到的最新的模型参数对用户输入的语音信息进行降噪处理。

[0075] 本申请的上述实施例中，通过获取训练数据的时域特征值(可以体现时域特征)和目标值，并采用基于频域构建的深度学习训练模型(可以体现频域特征)对训练数据的时域特征值和目标值进行训练，可以在训练模型的过程中将语音信息的时域特征和频域特征进行结合，进而提升深度学习训练模型的训练性能，加快深度学习训练模型的训练速度；且，由于训练数据可以包括多种环境下的语音信息，并可以多次采用不同的训练数据对深度学习训练模型进行训练，因此可以得到比较准确的模型参数；进一步地，通过使用深度学习训练模型训练得到的模型参数对终端设备中的第一语音降噪模型的参数进行更新，可以使得终端设备采用深度学习训练模型训练得到的模型参数对用户输入的语音信息进行降噪处理，从而能够使降噪处理得到的语音信息更加准确，提高用户的体验感。此外，采用深度学习训练模型对训练数据进行训练获取模型参数的过程和终端设备对用户语音信息的降噪过程可以并行处理，从而能够提高语音降噪的处理速度和处理效率。

[0076] 针对上述方法流程，本申请实施例还提供一种降噪装置，该装置的具体内容可以参照上述方法实施。

[0077] 图3为本申请实施例提供的一种服务器的结构示意图，包括：

[0078] 获取模块301，用于获取训练数据，所述训练数据包括在第一环境中收集的第一语音信息和在第二环境中收集的第二语音信息，所述第一环境中的噪声小于等于预设阈值，所述第二环境中的噪声大于所述预设阈值；

[0079] 确定模块302，用于根据所述训练数据，确定所述训练数据的时域特征值和目标值，所述训练数据的时域特征值包括噪声阈值、长时能量值、短时能量值和噪声包络跟踪值中的一项或多项；所述训练数据的目标值包括所述第一语音信息的语音活动检测值和/或所述第二语音信息的全带信噪比；

[0080] 处理模块303，用于基于频域构建的深度学习训练模型，对所述时域特征值和所述目标值进行训练，得到模型参数，并将所述模型参数发送给终端设备，以使所述终端设备使用所述模型参数对用户输入的语音信息进行降噪处理。

[0081] 图4为本申请实施例提供的一种终端设备的结构示意图，包括：

[0082] 收发模块401，用于接收服务器发送的模型参数；

[0083] 更新模块402，用于根据所述模型参数，对所述终端设备中的第一语音降噪模型的参数进行更新，得到更新后的第一语音降噪模型；

[0084] 降噪模块403，用于在接收到用户输入的语音信息后，使用所述更新后的第一语音降噪模型对所述语音信息进行降噪处理。

[0085] 可选地，在所述收发模块401接收到服务器发送的模型参数之后，所述更新模块402还用于：将预设标志更新为第一指示值，以及在得到更新后的第一语音降噪模型之后，将所述预设标志更新为第二指示值。

[0086] 可选地，所述降噪模块403还用于：

[0087] 确定所述预设标志为所述第二指示值。

[0088] 可选地，所述降噪模块403还用于：

[0089] 在接收到用户输入的语音信息后，若确定所述预设标志为所述第一指示值，则使用所述终端设备中的第二语音降噪模型对所述语音信息进行降噪处理；所述第二语音降噪模型为所述第一语音降噪模型的备用模型。

[0090] 从上述内容可以看出：本申请的上述实施例中，服务器通过将收集到的第一环境(噪声小于等于预设阈值)中的第一语音信息和第二环境(噪声大于预设阈值)中的第二语音信息作为训练数据，并在确定训练数据的时域特征值和目标值后，基于频域构建的深度学习训练模型对时域特征值和目标值进行训练，得到模型参数，以及将模型参数发送给终端设备，使得终端设备可以在接收到服务器发送的模型参数后，对终端设备中的第一语音降噪模型的参数进行更新，并可以使用更新后的第一语音降噪模型对用户输入的语音信息进行降噪处理。本申请实施例中，通过获取训练数据的时域特征值(可以体现时域特征)和目标值，并采用基于频域构建的深度学习训练模型(可以体现频域特征)对训练数据的时域特征值和目标值进行训练，可以在训练模型的过程中将语音信息的时域特征和频域特征进行结合，进而提升深度学习训练模型的训练性能，加快深度学习训练模型的训练速度；且，由于训练数据可以包括多种环境下的语音信息，并可以多次采用不同的训练数据对深度学

习训练模型进行训练，因此可以得到比较准确的模型参数；进一步地，通过使用深度学习训练模型训练得到的模型参数对终端设备中的第一语音降噪模型的参数进行更新，可以使得终端设备采用深度学习训练模型训练得到的模型参数对用户输入的语音信息进行降噪处理，从而能够使降噪处理得到的语音信息更加准确，提高用户的体验感。此外，采用深度学习训练模型对训练数据进行训练获取模型参数的过程和终端设备对用户语音信息的降噪过程可以并行处理，从而能够提高语音降噪的处理速度和处理效率。

[0091] 本领域内的技术人员应明白，本申请的实施例可提供为方法、或计算机程序产品。因此，本申请可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且，本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0092] 本申请是参照根据本申请实施例的方法、设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器，使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0093] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

[0094] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上，使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

[0095] 尽管已描述了本申请的优选实施例，但本领域内的技术人员一旦得知了基本创造性概念，则可对这些实施例作出另外的变更和修改。所以，所附权利要求意欲解释为包括优选实施例以及落入本申请范围的所有变更和修改。

[0096] 显然，本领域的技术人员可以对本申请进行各种改动和变型而不脱离本申请的精神和范围。这样，倘若本申请的这些修改和变型属于本申请权利要求及其等同技术的范围之内，则本申请也意图包含这些改动和变型在内。

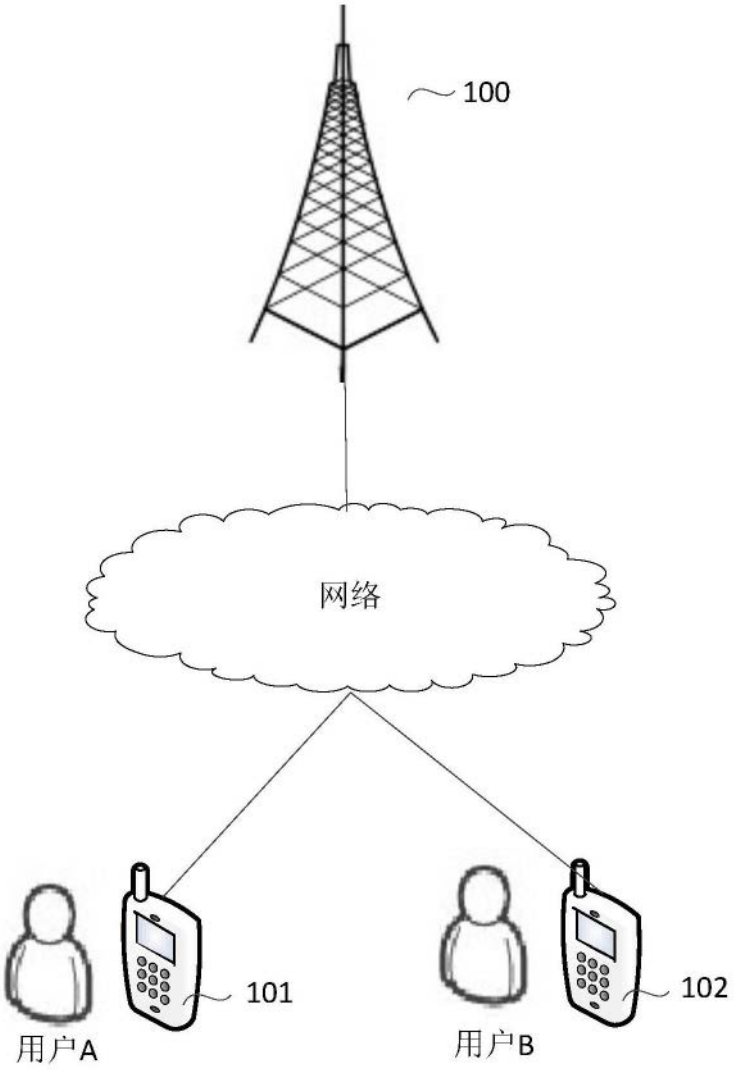


图1

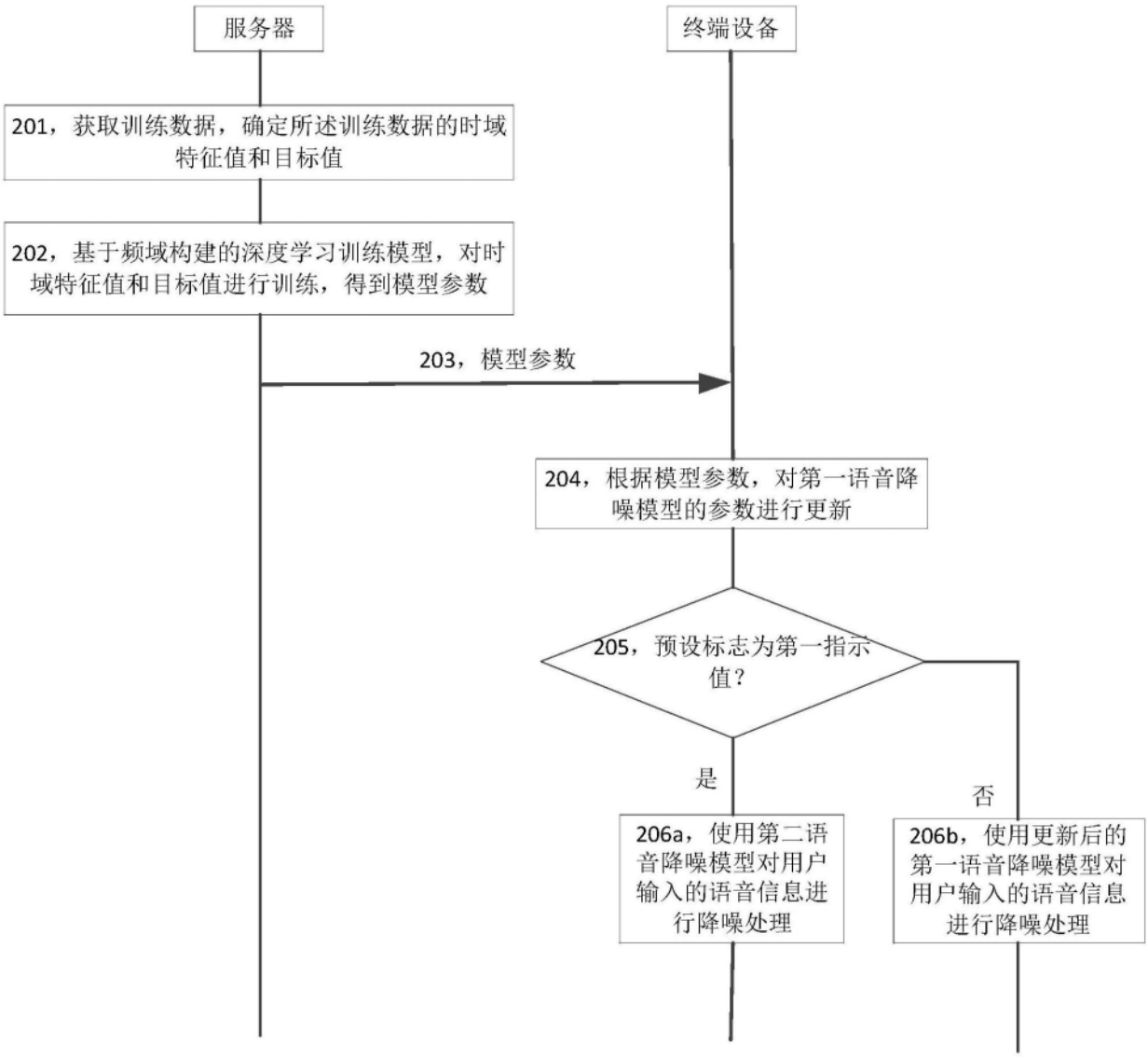

服务器
终端设备
201，获取训练数据，确定所述训练数据的时域特征值和目标值
202，基于频域构建的深度学习训练模型，对时域特征值和目标值进行训练，得到模型参数
203，模型参数
204，根据模型参数，对第一语音降噪模型的参数进行更新
205，预设标志为第一指示值？
是
否
206a，使用第二语音降噪模型对用户输入的语音信息进行降噪处理
206b，使用更新后的第一语音降噪模型对用户输入的语音信息进行降噪处理

图2

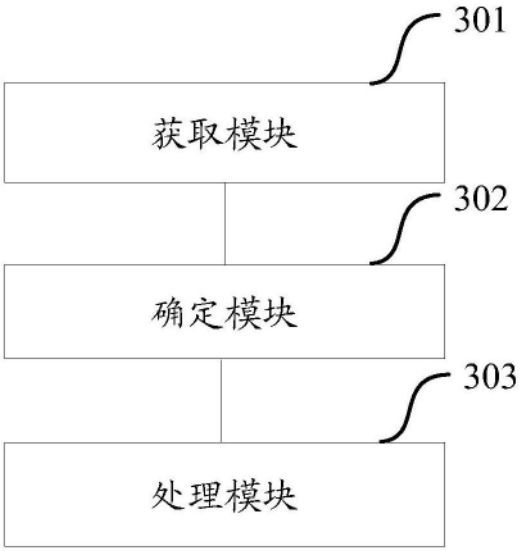


图3

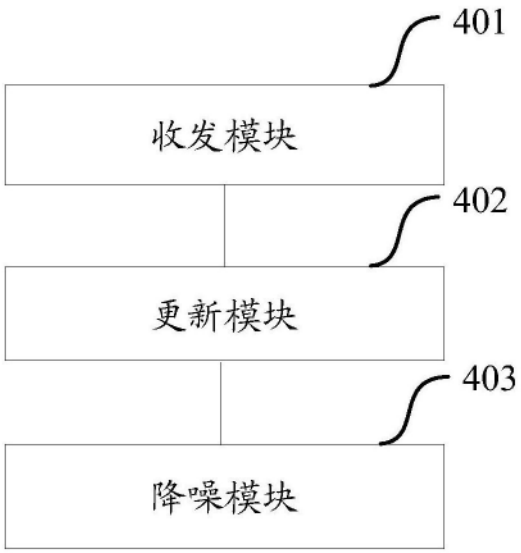


图4