

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2024年7月4日 (04.07.2024)



(10) 国际公布号
WO 2024/139730 A1

(51) 国际专利分类号:
G10L 21/0208 (2013.01)

(21) 国际申请号: PCT/CN2023/129766

(22) 国际申请日: 2023年11月3日 (03.11.2023)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202211725937.6 2022年12月30日 (30.12.2022) CN

(71) 申请人: 腾讯科技(深圳)有限公司 (TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED) [CN/CN]; 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。

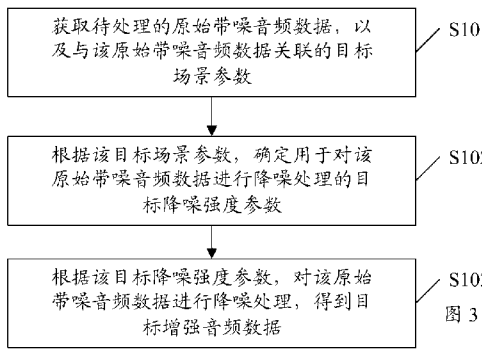
(72) 发明人: 邹欢彬(ZOU, Huanbin); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。李志成(LI, Zhicheng); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。赵军(ZHAO, Jun); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。

(74) 代理人: 北京派特恩知识产权代理有限公司(CHINA PAT INTELLECTUAL PROPERTY OFFICE); 中国北京市海淀区苏州街3号大恒科技大厦南座五层503, Beijing 100080 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI,

(54) Title: AUDIO DATA PROCESSING METHOD AND APPARATUS, AND DEVICE, COMPUTER-READABLE STORAGE MEDIUM AND COMPUTER PROGRAM PRODUCT

(54) 发明名称: 音频数据处理方法、装置、设备、计算机可读存储介质及计算机程序产品



S101 Acquire original noisy audio data to be processed, and a target scenario parameter associated with said original noisy audio data

S102 According to the target scenario parameter, determine a target noise-reduction intensity parameter used for performing noise reduction processing on said original noisy audio data

S103 According to the target noise-reduction intensity parameter, perform noise reduction processing on said original noisy audio data, so as to obtain target enhanced audio data

(57) Abstract: An audio data processing method and apparatus, and a device and a storage medium, which are applied to a cloud server in cloud technology. The method comprises: acquiring original noisy audio data to be processed, and a target scenario parameter associated with said original noisy audio data (S101); according to the target scenario parameter, determining a target noise-reduction intensity parameter used for performing noise reduction processing on said original noisy audio data (S102); and according to the target

GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

noise-reduction intensity parameter, performing noise reduction processing on said original noisy audio data, so as to obtain target enhanced audio data (S103).

(57) 摘要: 一种音频数据处理方法、装置、设备及存储介质, 应用于云技术中的云服务器, 方法包括: 获取待处理的原始带噪音频数据, 以及与原始带噪音频数据关联的目标场景参数(S101), 根据目标场景参数, 确定用于对原始带噪音频数据进行降噪处理的目标降噪强度参数(S102), 根据目标降噪强度参数, 对原始带噪音频数据进行降噪处理, 得到目标增强音频数据(S103)。

音频数据处理方法、装置、设备、计算机可读存储介质及计算机程序产品

相关申请的交叉引用

5 本申请基于申请号为 202211725937.6、申请日为 2022 年 12 月 30 日的中国专利申请提出，并要求以上中国专利申请的优先权，以上中国专利申请的全部内容在此引入本申请作为参考。

技术领域

本申请涉及云技术等领域，尤其涉及一种音频数据处理方法、装置、设备、计算机可读存储介质及计算机程序产品。

10 背景技术

目前，基于互联网的 IP 语音传输（Voice over Internet Protocol, VoIP）通信和蜂窝通信等通信系统已在多种通信场景中普及，例如，网络电话、网络会议、直播等。由于讲话者所处环境复杂且具有多样性，导致采集到的音频数据通常包含噪声数据，因此，需要对带噪音频数据（即包含噪声数据的音频数据）进行去噪声处理，以确保音频数据的质量。目前在对带噪音频数据进行去噪声处理的过程中，需要将噪声数据与纯音频数据（即有效语音数据）完全分离，进而去除噪声；实践中发现，这种去噪声处理方式，会对纯音频数据造成一定程度的损失，导致音频数据的质量不佳。

发明内容

20 本申请实施例提供一种音频数据处理方法、装置、设备、计算机可读存储介质及计算机程序产品，能够避免降噪时对有效音频数据造成损失，从而提高音频数据的质量。

本申请实施例提供一种音频数据处理方法，应用于计算机设备，包括：

获取待处理的原始带噪音频数据，以及与所述原始带噪音频数据关联的目标场景参数；

25 根据所述目标场景参数，确定用于对所述原始带噪音频数据进行降噪处理的目标降噪强度参数；

根据所述目标降噪强度参数，对所述原始带噪音频数据进行降噪处理，得到目标增强音频数据。

本申请实施例提供一种音频数据处理装置，包括：

30 获取模块，用于获取待处理的原始带噪音频数据，以及与所述原始带噪音频数据关联的目标场景参数；

确定模块，用于根据所述目标场景参数，确定用于对所述原始带噪音频数据进行降噪处理的目标降噪强度参数；

处理模块，用于根据所述目标降噪强度参数，对所述原始带噪音频数据进行降噪处理，得到目标增强音频数据。

35 本申请实施例提供了一种计算机设备，包括存储器和处理器，上述存储器存储有计

计算机程序，上述处理器执行上述计算机程序时实现上述的音频数据处理方法的步骤。

本申请实施例一方面提供了一种计算机可读存储介质，其上存储有计算机程序，上述计算机程序被处理器执行时实现上述的音频数据处理方法的步骤。

5 本申请实施例一方面提供了一种计算机程序产品，包括计算机程序，该计算机程序被处理器执行时实现上述的音频数据处理方法的步骤。

在本申请实施例中，通过原始带噪音频数据关联的目标场景参数，自适应地确定用于对原始带噪音频数据进行降噪处理的目标降噪强度参数，基于目标降噪强度参数，定量地降低原始带噪音频数据中的噪声含量。即此处的目标场景参数反映原始带噪音频数据的应用场景和采集场景中的至少一种，目标降噪强度参数反映对原始带噪音频数据中的噪声抑制的力度。换言之，通过原始带噪音频数据的应用场景中针对音频数据的实际需求（和\或原始带噪音频数据的采集场景中的噪声分布情况），定量地降低原始带噪音频数据中的噪声含量，并接受一定程度的噪声残留，不需要完全分离原始带噪音频数据中的噪声数据和音频数据，来彻底抑制噪声，避免在降噪时对有效音频数据造成损失，提高音频数据的质量，提高噪声处理的灵活性。

15 附图说明

为了更清楚地说明本申请实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍，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，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

20 图 1 是本申请提供了一种音频数据处理系统示意图；

图 2 是本申请提供了一种音频数据处理方法的交互场景示意图；

图 3 是本申请提供了一种音频数据处理方法的流程示意图；

图 4 是本申请提供了一种音频数据处理方法的流程示意图；

图 5 是本申请提供了一种目标降噪处理模型的结构示意图；

25 图 6 为是本申请提供了一种带噪音频数据在不同降噪强度参数下的 PESQ 得分情况的示意图；

图 7 为是本申请提供了一种带噪音频数据在不同降噪强度参数下的 SI-SNR 得分情况的示意图；

图 8 是本申请实施例提供了一种音频数据处理装置的结构示意图；

30 图 9 是本申请实施例提供了一种计算机设备的结构示意图。

具体实施方式

下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述，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作出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

35 本申请实施例中主要涉及人工智能云服务，所谓人工智能云服务，一般也被称作是 AIaaS（AI as a Service，中文为“AI 即服务”）。这是目前主流的一种人工智能平台的服务方式，AIaaS 平台会把几类常见的 AI 服务进行拆分，并在云端提供独立或者打包的服务。这种服务模式类似于开了一个 AI 主题商城：所有的开发者都可以通过 API 接口的方式接入使用平台提供的一种或者是多种人工智能服务，部分资深的开发者还可以

40

使用平台提供的 AI 框架和 AI 基础设施来部署和运维自己专属的云人工智能服务。

例如，该人工智能云服务中包括用于对带噪音频数据进行降噪处理的目标降噪处理模型，当需要对原始带噪音频数据进行降噪处理时，计算机设备可以通过 API 接口调用该人工智能云服务中的目标降噪处理模型，将该原始带噪音频数据和目标降噪强度参数输入至该目标降噪处理模型中。通过该目标降噪处理模型基于该目标降噪强度参数，对原始带噪音频数据进行降噪处理，实现定量地降低原始带噪音频数据中的噪声含量，避免降噪时对有效音频数据造成损失，提高音频数据的质量，使音频数据的降噪处理更加智能化。同时，不同计算机设备均可调用该目标降噪处理模型，实现多个计算机设备共享目标降噪处理模型，提高目标降噪处理模型的利用率，不需要计算机设备单独训练得到目标降噪处理模型，降低计算机设备的计算资源开销。

为了便于更清楚理解本申请，首先介绍实现本申请的音频数据处理系统，如图 1 所示，该音频数据处理系统中包括服务器 10 和终端集群，终端集群可以包括一个或者多个终端，这里将不对终端的数量进行限制。如图 1 所示，终端集群可以包括终端 1、终端 2、...、终端 n；可以理解的是，终端 1、终端 2、终端 3、...、终端 n 均可以与服务器 10 进行网络连接，以便于每个终端均可以通过网络连接与服务器 10 之间进行数据交互。

终端中安装有一个或多个目标应用，此处的目标应用可以是指具有语音通信功能的应用，如目标应用包括独立的应用程序、网页应用、宿主应用中的小程序等。终端集群中的任一终端可以作为发送终端，也可以作为接收终端，发送终端可以是指生成原始带噪音频数据并发送原始带噪音频数据的终端，接收终端可以是指接收原始带噪音频数据的终端。例如，终端 1 对应的用户 1 与终端 2 对应的用户 2 进行语音通信时，当用户 1 需要向用户 2 发送音频数据时，这时终端 1 可以称为发送终端，终端 2 可以称为接收终端；同理，当用户 2 需要向用户 1 发送音频数据时，这时终端 2 可以称为发送终端，终端 1 可以称为接收终端。

服务器 10 是指为终端中的目标应用提供后端服务的设备，在一个实施例中，服务器可以用于对发送终端发送的原始带噪音频数据进行降噪处理等等，将降噪处理后的原始带噪音频数据转发至接收终端。在一个实施例中，服务器 10 可以用于将发送终端发送的原始带噪音频数据转发至接收终端，由接收终端对原始带噪音频数据进行降噪处理，得到处理后的原始带噪音频数据。在一个实施例中，服务器可以用于接收发送终端发送的降噪处理后原始带噪音频数据，并将降噪处理后原始带噪音频数据转发至接收终端，即降噪处理后原始带噪音频数据是由发送终端对原始带噪音频数据进行降噪处理得到的。

在一些实施例中，本申请实施例中的原始带噪音频数据可以是指由发送终端的麦克风采集到的音频数据，即原始带噪音频数据是指未进行降噪处理的音频数据，通常原始带噪音频数据包括音频数据和噪声数据。音频数据可以是指对用户有用的数据，如该音频数据可以是指用户语音通信过程中的语音数据，或音频数据可以是指用户录制的音乐作品等等；音频数据可以由人、动物、机器人等所发出的声音进行采集得到的。此处的噪声数据可以是指对用户无意义的的数据，如该噪声数据可以是指环境噪声，例如，在用户语音通信的过程中，除通话双方的语音数据以外的音频数据均为噪声数据。

在一些实施例中，服务器可以是独立的一个物理服务器，也可以是至少两个物理服务器构成的服务器集群或者分布式系统，还可以是提供云服务、云数据库、云计算、云函数、云存储、网络服务、云通信、中间件服务、域名服务、安全服务、内容分发网络 (Content Delivery Network, CDN)、以及大数据和人工智能平台等基础云计算服务的云服务器。终端可以是指车载终端、智能手机、平板电脑、笔记本电脑、桌上型电脑、

智能音箱、有屏音箱、智能手表等等，但并不局限于此。各个终端以及服务器可以通过有线或无线通信方式进行直接或间接地连接，同时，终端以及服务器的数量可以为一个或至少两个，本申请在此不做限制。

图1的音频数据处理系统可应用于语音通信场景、直播场景、音视频录制场景等等，以图1的音频数据系统应用于图2所示的语音通信场景为例进行说明，图2中的终端20a可以为图1中的终端集群中的任一终端，图2中的终端21a可以为图1中的终端集群中除终端20a以外的任一终端，图2中的服务器22a可以为图1中的服务器10。

在终端20a对应的用户1与终端21a对应的用户2进行语音通信的过程中，终端20a可以对用户1的讲话过程进行采集，得到原始带噪音频数据1；该原始带噪音频数据1包括用户1的讲话内容（即语音数据1）以及噪声数据1，噪声数据1反映用户1讲话的过程中的环境噪声，如终端20a发出的啸叫，或其他人的讲话内容等等。终端20a采集到原始带噪音频数据1后，可以将该原始带噪音频数据1发送至服务器22a，服务器22a接收到原始带噪音频数据1后，可以获取该原始带噪音频数据1的目标场景参数1；该目标场景参数1可以用于反映该原始带噪音频数据1的采集场景和应用场景中的至少一种，以该目标场景参数1反映该原始带噪音频数据1的应用场景为例进行说明。该目标场景参数1反映原始带噪音频数据1的应用场景为语音通信场景，服务器22a可以根据应用场景与降噪强度参数之间的对应关系，查询该原始带噪音频数据1的应用场景对应的降噪强度参数，将查询到的降噪强度参数确定为原始带噪音频数据1对应的目标降噪强度参数1。

目标降噪强度参数1反映原始带噪音频数据1需要对噪声数据进行降噪处理的力度，因此，服务器22a可以根据该目标降噪强度参数1对原始带噪音频数据1进行降噪处理，得到目标增强音频数据1，将该目标增强音频数据1发送至终端21a。该目标增强音频数据1中残留了部分噪声数据，避免对原始带噪音频数据1的音频数据与噪声数据完全分离，造成原始带噪音频数据1中的音频数据受损。终端21a接收到目标增强音频数据1后，用户2可以基于目标增强音频数据1感知用户1所处的环境，使语音通信过程更加真实饱满。

同理，在终端20a对应的用户1与终端21a对应的用户2进行语音通信的过程中，终端21a可以对用户2的讲话过程进行采集，得到原始带噪音频数据2；该原始带噪音频数据2包括用户2的讲话内容（即语音数据2）以及噪声数据2，噪声数据2反映用户2讲话的过程中的环境噪声，如终端21a发出的啸叫，或其他人的讲话内容等等。终端21a采集到原始带噪音频数据2后，可以将该原始带噪音频数据2发送至服务器22a，服务器22a接收到原始带噪音频数据2后，可以获取该原始带噪音频数据2的目标场景参数2；该目标场景参数2可以用于反映该原始带噪音频数据2的采集场景和应用场景中的至少一种，以该目标场景参数2反映该原始带噪音频数据2的应用场景为例进行说明。该目标场景参数2反映原始带噪音频数据2的应用场景为语音通信场景，服务器22a可以根据应用场景与降噪强度参数之间的对应关系，查询该原始带噪音频数据2的应用场景对应的降噪强度参数，将查询到的降噪强度参数确定为原始带噪音频数据2对应的目标降噪强度参数2。

目标降噪强度参数2反映需要对原始带噪音频数据2中的噪声数据进行降噪处理的力度，因此，服务器22a可以根据该目标降噪强度参数1对原始带噪音频数据1进行降噪处理，得到目标增强音频数据2，将该目标增强音频数据2发送至终端20a。该目标增强音频数据2中残留了部分噪声数据，避免对原始带噪音频数据2的音频数据与噪声数据完全分离，造成原始带噪音频数据2中的音频数据受损。终端20a接收到目标增强音频数据2后，用户1可以基于目标增强音频数据2感知用户2所处的环境，使语音通

信过程更加真实饱满。

在一些实施例中，请参见图 3，是本申请实施例提供的一种音频数据处理方法的流程图示意图。如图 3 所示，该方法可由图 1 中的终端集群中的任一终端来执行，也可以由图 1 中的服务器来执行，本申请实施例中用于执行该音频数据处理方法的设备可以统称为计算机设备。其中，该方法可以包括如下步骤：

S101、获取待处理的原始带噪音频数据，以及与该原始带噪音频数据关联的目标场景参数。

在一些实施例中，计算机设备可以采集得到待处理的原始带噪音频数据，或，计算机设备可以从其他设备中获取待处理的原始带噪音频数据，然后获取与该原始带噪音频数据关联的目标场景参数。该目标场景参数用于确定该原始带噪音频数据的采集场景和应用场景中的至少一种。

在一个实施例中，计算机设备可以通过传感器对原始带噪音频数据的录制环境进行检测，得到录制环境的环境参数，将该录制环境的环境参数，确定为该原始带噪音频数据的目标场景参数；录制环境的环境参数包括光线、温度、湿度等等中的一种或多种，即目标场景参数包括录制环境的环境参数。该目标场景参数可用于确定原始带噪音频数据的采集场景，例如，如果录制环境的光线为自然光，则表明原始带噪音频数据的采集场景为户外；如果录制环境的光线为人造光，则表明原始带噪音频数据的采集场景为室内。

在一个实施例中，计算机设备可以获取该原始带噪音频数据的采集设备的位置信息，将采集设备的位置信息确定为原始带噪音频数据的采集环境的位置信息，将该采集环境的位置信息确定为该原始带噪音频数据的目标场景参数。该目标场景参数可用于确定原始带噪音频数据的采集场景，例如，如果基于录制环境的位置信息确定录制环境为公园，则表明原始带噪音频数据的采集场景为户外或开放场所；如果基于录制环境的位置信息确定录制环境为办公楼，则表明原始带噪音频数据的采集场景为室内、私密场所等。

在一个实施例中，计算机设备可以获取原始带噪音频数据的录制应用程序对应的程序标识，将该录制应用程序的程序标识，确定为原始带噪音频数据的目标场景参数；录制应用程序可以包括但不限于：语音通话应用程序、会议应用程序、音乐播放应用程序等等，程序标识可以为程序名称、编号等。该目标场景参数可用于确定原始带噪音频数据的应用场景，例如，如果录制应用程序的程序标识指示录制应用程序为语音通话应用程序，则表明该原始带噪音频数据的应用场景为语音通话场景；如果用于录制应用程序的程序标识指示录制应用程序为会议应用程序，则表明该原始带噪音频数据的应用场景为会议应用场景。

在一些实施例中，目标场景参数可以包括原始带噪音频数据的录制环境的环境参数、录制环境的位置信息以及录制应用程序对应的程序标识等中的至少一种或多种。

在一个实施例中，当该目标场景参数用于确定该原始带噪音频数据的采集场景时，计算机设备可以根据采集该原始带噪音频数据的设备的位置信息，确定该原始带噪音频数据关联的采集场景，该采集场景包括室内、户外、私密场所或开放场所等等。在一个实施例中，当该目标场景参数用于确定该原始带噪音频数据的应用场景时，计算机设备可以根据该原始带噪音频数据的所有者的使用指示信息，确定该原始带噪音频数据的应用场景，使用指示信息用于指示该原始带噪音频数据的应用场景，该应用场景可以包括语音通信、直播、音乐作品的播放场景等等。在一个实施例中，当该目标场景参数用于确定该原始带噪音频数据的应用场景和采集场景时，计算机设备可以根据采集该原始带噪音频数据的设备的位置信息，确定该原始带噪音频数据关联的采集场景，根据该原始

带噪音频数据的所有者的使用指示信息，确定该原始带噪音频数据的应用场景。

S102、根据该目标场景参数，确定用于对该原始带噪音频数据进行降噪处理的目标降噪强度参数。

5 在一些实施例中，计算机设备可以根据该目标场景参数，确定用于对该原始带噪音频数据进行降噪处理的目标降噪强度参数；该目标降噪强度参数用于指示原始带噪音频数据中需去除的噪声数据对应的数据量（即含量），也即目标降噪强度参数用于指示对原始带噪音频数据中的噪声数据的降噪力度。例如，假设原始带噪音频数据中的噪声数据的强度为 6dB，该目标降噪强度参数为 5dB，该目标降噪强度参数指示将原始带噪音频数据中的噪声数据的强度（即功率）降低 5dB，降噪处理后的原始带噪音频数据（即目标增强音频数据）中的噪声数据的强度为 1dB。或者，假设原始带噪音频数据的原始信噪比为 10dB，该目标降噪强度参数为 5dB，由于原始带噪音频数据的原始信噪比为：原始带噪音频数据中的音频数据的功率与原始带噪音频数据中的噪声数据的功率之间的比值。因此，将原始带噪音频数据中的噪声数据的强度（即功率）降低 5dB，相当于将原始带噪音频数据中音频数据的信噪比提升 5dB，即降噪处理后的原始带噪音频数据（即目标增强音频数据）的信噪比变更为 $5\text{dB}+6\text{dB}=11\text{dB}$ 。该目标降噪强度参数越大，表明针对原始带噪音频数据的降噪力度越大，原始带噪音频数据中需去除的噪声数据对应数据量越多；该目标降噪强度参数越小，表明针对原始带噪音频数据的降噪力度越小，原始带噪音频数据中需去除的噪声数据对应数据量越少。

20 在一些实施例中，计算机设备可以通过如下三种方式中的任一种方式，确定用于对原始带噪音频数据进行降噪处理的目标降噪强度参数：方式一，如果目标场景参数包括录制应用程序对应的程序标识，计算机设备可以根据录制应用程序对应的程序标识确定该原始带噪音频数据的应用场景，即确定该目标场景参数能够表征该原始带噪音频数据的应用场景，获取在该应用场景下关于音频数据的质量需求等级；该质量需求等级反映该应用场景下针对音频数据的质量要求，即该质量需求等级越高，该应用场景下针对音频数据的质量要求越高；即该质量需求等级越低，该应用场景下针对音频数据的质量要求越低。通常针对该原始带噪音频数据的目标降噪强度参数越大，原始带噪音频数据中需去除的噪声数据对应数据量越多，同时，对原始带噪音频数据中音频数据的损失也越大；针对该原始带噪音频数据的目标降噪强度参数越小，原始带噪音频数据中需去除的噪声数据对应数据量越少，同时，对原始带噪音频数据中音频数据的损失也越小。因此，计算设备可以根据该质量需求等级与降噪强度参数之间的对应关系，查询该原始带噪音频数据对应的质量需求等级对应的降噪强度参数，将查询到的降噪强度参数，确定为用于对该原始带噪音频数据进行降噪处理的目标降噪强度参数。此处质量需求等级与降噪强度参数之间的对应关系可以是根据历史经验得到的，该质量需求等级与目标降噪强度参数之间具有负相关关系。也就是说，质量需求等级越低，目标降噪强度参数越大；质量需求等级越高，目标降噪强度参数越小，避免过度对原始带噪音频数据降噪处理，造成原始带噪音频数据中的音频数据受到损失，提高音频数据的质量。

35 例如，在视频会议场景下用户通常可接受音频数据的质量受到一定程度的损失，不接受该视频会议场景下存在大量噪声数据。因此，计算机设备可以将第一质量等级确定为原始带噪音频数据在视频会议场景下的质量需求等级；将第一降噪强度参数，确定为用于对原始带噪音频数据进行降噪处理的目标降噪强度参数，以便消除视频会议场景下更多的噪声数据，避免噪声数据干扰视频会议。在语音通信场景下用户通常对音频数据的质量要求比较高，接受语音数据场景下存在一定的噪声数据。因此，计算机设备可以将第二质量等级确定为原始带噪音频数据在视频会议场景下的质量需求等级；将第二降噪强度参数，确定为用于对原始带噪音频数据进行降噪处理的目标降噪强度参数，以便

消除视频会议场景下较少的噪声数据，以使用户可以基于残留的噪声数据，感受语音通信双方所在的真实环境，给语音通信双方营造身临其境的氛围。第一质量等级小于第二质量等级，第一降噪强度参数大于第二降噪强度参数。

5 方式二，如果该目标场景参数包括该原始带噪音频数据的录制环境的环境参数和录制环境的位置信息中的至少一种，则计算机设备可以根据该目标场景参数确定原始带噪音频数据的采集场景，即确定该目标场景参数反映该原始带噪音频数据的采集场景，计算设备可以获取在该采集场景中在历史时间段内的历史噪声数据，该历史时间段可以是指近一天内，近一周内，或者，该历史时间段是根据当前时间确定的，如该当前时间段为12月16日的19:20:00~19:30:00，历史时间段可以是指12月15日的19:20:00~19:30:00。
10 由于同一采集场景下在历史时间段内的噪声数据的分布特征，与该当前时间段内的噪声数据的分布特征具有相似性，因此，计算机设备可以根据该历史噪声数据，确定用于对该原始带噪音频数据进行降噪处理的目标降噪强度参数。通过采集场景下的历史噪声数据，确定目标降噪强度参数，避免噪声残留不稳定的问题，即噪声数据忽大忽小，忽有忽无。

15 需要说明的是，上述根据该历史噪声数据，确定用于对该原始带噪音频数据进行降噪处理的目标降噪强度参数，包括：计算机设备可以从该历史噪声数据中，确定该采集场景在该历史时间段内的噪声数据对应噪声类型和噪声变化特征。该噪声类型包括稳态噪声、非稳态噪声以及脉冲噪声等等，稳态噪声是指噪声强度的变化较小（一般不大于3dB），且不随时间有大幅度的变化，如电机、风机及其他电磁噪声，固定转速的摩擦、转动等噪声。非稳态噪声：指噪声强度随时间而有起伏波动（声压变化大于3dB），有的呈周期性噪声，如锤击；有的呈无规律的起伏噪声，如交通噪声。脉冲噪声：指持续时间小于1s的单个或多个突发声组成的噪声，声压级原始水平升至峰值又回至原始水平所需的持续时间短于500ms，其峰值声压级大于40dB。脉冲噪声往往是突发的高强噪声，如爆破、火炮发射等所产生的噪声。该噪声变化特征是指历史噪声数据的强度随
20 时间的变化速度（即变化快慢），也即该噪声变化特征反映历史噪声数据是否稳定。在一些实施例中，计算机设备可以根据该噪声类型和该噪声变化特征，确定用于对该原始带噪音频数据进行降噪处理的目标降噪强度参数，通过采集场景下的历史噪声数据的分布特征（即噪声类型和噪声变化特征），确定目标降噪强度参数，避免噪声残留不稳定的问题。

30 需要说明的是，该采集场景在历史时间段内的历史噪声数据的噪声类型的数量为M时，M为大于或等于1的正整数，计算机设备可以基于M个噪声类型的历史噪声数据分别对应的噪声变化特征，确定用于对原始带噪音频数据进行降噪处理的M个候选降噪强度参数，一个噪声类型的历史噪声数据对应一个候选降噪强度参数。在一些实施例中，计算机设备可以将M个候选降噪强度参数，确定为目标降噪强度参数；或者，计算机设备可以根据M个候选降噪强度参数进行加权平均处理（或算数平均处理），得到目标降噪强度参数。噪声类型的历史噪声数据对应的候选降噪强度参数，可以为随着
35 对应噪声变化特征变化的变量，或者，噪声类型的历史噪声数据对应的候选降噪强度参数，可以为基于对应噪声变化特征变化确定的固定值。这样可以避免无法抑制所有噪声类型的噪声，同时，避免在非稳态噪声情况下，由于噪声变化特征随时间变化而快速变化，从而引起噪声残留不连续的问题，即避免经过降噪处理后的原始带噪音频数据中的
40 噪声数据出现忽大忽小、忽有忽无，导致音频数据的可感知度比较地问题。换言之，通过根据历史噪声数据的噪声类型和噪声变化特征，确定用于对原始带噪音频数据进行降噪处理的目标降噪强度参数，这样有利于对原始带噪音频数据中的所有噪声类型的噪声进行降噪处理（即抑制处理），使降噪处理后的原始带噪音频数据中的噪声残留更加稳

定、平滑，提高降噪处理后的原始带噪音频数据中的音频数据的感知度。

例如，原始带噪音频数据中的第一噪声类型的噪声数据的噪声变化特征指示第一噪声类型的噪声数据的强度位于[5, 10]dB 范围内变化，计算机设备可以根据[5, 10]dB 确定第一噪声类型的噪声数据对应的降噪强度参数，如将 3dB 确定为第一噪声类型的噪声数据对应的降噪强度参数。同理，原始带噪音频数据中的第二噪声类型的噪声数据的噪声变化特征指示第二噪声类型的噪声数据的强度位于[2, 6]dB 范围内变化，计算机设备可以根据[2, 6]dB 确定第二噪声类型的噪声数据对应的降噪强度参数，如将 4dB 确定为第二噪声类型的噪声数据对应的降噪强度参数。然后，可以将第一噪声类型的噪声数据对应的降噪强度参数，以及第二噪声类型的噪声数据对应的降噪强度参数，确定为用于对原始带噪音频数据进行降噪处理的目标降噪强度参数。其中，第一噪声类型的噪声数据对应的降噪强度参数用于对原始带噪音频数据中的第一噪声类型的噪声数据进行降噪处理，第二噪声类型的噪声数据对应的降噪强度参数用于对原始带噪音频数据中的第二噪声类型的噪声数据进行降噪处理。第一噪声数据类型的噪声数据的降噪处理顺序可以位于第二噪声数据类型的噪声数据的降噪处理顺序之前（或之后），第一噪声数据类型的噪声数据的降噪处理顺序与第二噪声数据类型的噪声数据的降噪处理顺序相同。其中，第一噪声类型可以是指稳态噪声，第二噪声类型可以为非稳态噪声。或者，计算机设备可以对第一噪声类型的噪声数据对应的降噪强度参数和第二噪声类型的噪声数据对应的降噪强度参数进行合并处理，得到用于对原始带噪音频数据进行降噪处理的目标降噪强度参数，此处的合并处理可以为求和处理或平均化处理等等。

方式三，如果目标场景参数包括录制应用程度对应的程序标识，以及该原始带噪音频数据的录制环境的环境参数和/或录制环境的位置信息，计算机设备可以根据录制应用程度对应的程序标识，确定原始带噪音频数据的应用场景，根据该原始带噪音频数据的录制环境的环境参数和录制环境的位置信息中的至少一种，确定该原始带噪音频数据的采集场景，即确定该目标场景参数反映该原始带噪音频数据的采集场景和应用场景。计算设备可以获取在该应用场景下关于音频数据的质量需求等级，根据该质量需求等级确定用于对该原始带噪音频数据进行降噪处理的第一降噪强度参数。之后获取在该采集场景中历史时间段内的历史噪声数据，根据该历史噪声数据，确定用于对该原始带噪音频数据进行降噪处理的第二降噪强度参数。上述确定第一降噪强度参数的实现过程可以参考上述方式一，确定第二降噪强度参数的实现过程可以参考上述方式二；然后，对第一降噪强度参数和第二降噪强度参数进行平均化处理，得到用于对原始带噪音频数据进行降噪处理的目标降噪强度参数，或者，计算机设备可以将第一降噪强度参数和第二降噪强度参数，确定为用于对原始带噪音频数据进行降噪处理的目标降噪强度参数，即目标降噪强度参数包括第一降噪强度参数和第二降噪强度参数。通过综合考虑原始带噪音频数据的采集场景和应用场景，确定用于对原始带噪音频数据进行降噪处理的目标降噪强度参数，提高对原始带噪音频数据进行降噪处理的准确度。

在一些实施例中，当目标降噪强度参数包括上述第一降噪强度参数和第二降噪强度参数时，第一降噪强度参数的对应的处理顺序位于第二降噪强度参数对应的处理顺序之前，即计算机设备可以先采用第一降噪强度参数对原始带噪音频数据进行降噪处理，得到第一候选增强音频数据，然后，采用第二降噪强度参数对第一候选增强音频数据进行降噪处理，得到目标增强音频数据。第一降噪强度参数的对应的处理顺序可以位于第二降噪强度参数对应的处理顺序之后，即计算机设备可以先采用第二降噪强度参数对原始带噪音频数据进行降噪处理，得到第二候选增强音频数据，然后，采用第一降噪强度参数对第二候选增强音频数据进行降噪处理，得到目标增强音频数据。或者，第一降噪强度参数的对应的处理顺序与第二降噪强度参数对应的处理顺序相同，即计算机设备可以

同时采用第一降噪强度参数和第二降噪强度参数对原始带噪音频数据进行降噪处理，得到目标增强音频数据。

S103、根据该目标降噪强度参数，对该原始带噪音频数据进行降噪处理，得到目标增强音频数据。

5 在一些实施例中，计算机设备可以根据该目标降噪强度参数，对该原始带噪音频数据进行降噪处理，得到该目标增强音频数据，即该目标增强音频数据为降噪处理后的原始带噪音频数据，该目标增强音频数据中的噪声数据的强度低于原始带噪音频数据中的噪声数据的强度，且，该目标增强音频数据中的噪声数据的稳定度大于原始带噪音频数据中的噪声数据的稳定度，即目标增强音频数据中的噪声数据更加稳定、平滑，有利于
10 用户感知目标增强音频数据中的音频数据（即语音数据）。

在一些实施例中，通过原始带噪音频数据关联的目标场景参数，自适应地确定用于对原始带噪音频数据进行降噪处理的目标降噪强度参数，基于目标降噪强度参数，定量地降低原始带噪音频数据中的噪声含量。即此处的目标场景参数反映原始带噪音频数据的应用场景和采集场景中的至少一种，目标降噪强度参数反映对原始带噪音频数据中的
15 噪声抑制的力度。换言之，通过原始带噪音频数据的应用场景中针对音频数据的实际需求（和/或原始带噪音频数据的采集场景中的噪声分布情况），定量地降低原始带噪音频数据中的噪声含量，并接受一定程度的噪声残留，不需要完全分离原始带噪音频数据中的噪声数据和音频数据，来彻底抑制噪声，避免降噪时对有效音频数据造成损失，提高音频数据的质量，提高噪声处理的灵活性。

20 在一些实施例中，请参见图4，是本申请实施例提供的一种音频数据处理方法的流程示意图。如图4所示，该方法可由图1中的终端集群中的任一终端来执行，也可以由图1中的服务器来执行，本申请实施例中用于执行该音频数据处理方法的设备可以统称为计算机设备。其中，该方法可以包括如下步骤：

25 本申请实施例中，步骤S201至S205为对初始降噪处理模型进行优化训练，得到目标降噪处理模型的过程，步骤S206至S208为通过目标降噪处理模型，基于目标降噪强度参数对原始带噪音频数据进行降噪处理的过程。

S201、获取样本音频数据以及样本噪声数据，根据该样本音频数据和该样本噪声数据生成样本带噪音频数据。

30 本申请实施例中，计算机设备可以获取语音数据集和噪声数据集，语音数据集中包括多个样本音频数据（即纯语音数据），噪声数据集包括多个样本噪声数据（即纯噪声数据）。然后将该语音数据集的样本音频数据与噪声数据集中的样本噪声数据进行组合，得到多个样本带噪音频数据。

35 例如，假设样本音频数据为 s_n ，样本噪声数据为 d_n ，样本带噪音频数据为 x_n ，则样本带噪音频数据可以采用如下公式（1）表示：

$$x_n = s_n + d_n \quad (1);$$

S202、获取用于对该样本带噪音频数据进行降噪处理的样本降噪强度参数。

40 本申请实施例中，计算机设备可以随机生成用于对样本带噪音频数据进行降噪处理的样本降噪强度参数；或者，计算机设备可以根据样本带噪音频数据中的样本噪声数据的噪声类型和噪声变化特征，生成用于对样本带噪音频数据进行降噪处理的样本降噪强度参数；或者，计算机设备可以根据样本带噪音频数据中的样本音频数据的应用场景，生成用于对样本带噪音频数据进行降噪处理的样本降噪强度参数。计算机设备生成用于对该样本带噪音频数据进行降噪处理的样本降噪强度参数的实现过程，可以参考上述生成用于对原始带噪音频数据进行降噪处理的目标降噪强度参数的实现过程。

45 在一些实施例中，上述根据样本带噪音频数据中的样本噪声数据的噪声类型和噪声

变化特征，生成用于对样本带噪音频数据进行降噪处理的样本降噪强度参数，包括：样本带噪音频数据中的第一噪声类型的噪声数据的噪声变化特征指示第一噪声类型的噪声数据的强度位于[5, 10]dB 范围内变化，计算机设备可以根据[5, 10]dB 确定第一噪声类型的噪声数据对应的降噪强度参数，如将 7.5dB 确定为第一噪声类型的噪声数据对应的降噪强度参数。同理，样本带噪音频数据中的第二噪声类型的噪声数据的噪声变化特征指示第二噪声类型的噪声数据的强度位于[2, 6]dB 范围内变化，计算机设备可以根据[2, 6]dB 确定第二噪声类型的噪声数据对应的降噪强度参数，如将 3dB 确定为第二噪声类型的噪声数据对应的降噪强度参数。然后，可以将第一噪声类型的噪声数据对应的降噪强度参数，以及第二噪声类型的噪声数据对应的降噪强度参数，确定为用于对样本带噪音频数据进行降噪处理的样本降噪强度参数。其中，第一噪声类型的噪声数据对应的降噪强度参数用于对样本带噪音频数据中的第一噪声类型的噪声数据进行降噪处理，第二噪声类型的噪声数据对应的降噪强度参数用于对样本带噪音频数据中的第二噪声类型的噪声数据进行降噪处理。如第一噪声类型可以是指稳态噪声，第二噪声类型可以为非稳态噪声。或者，计算机设备可以对第一噪声类型的噪声数据对应的降噪强度参数和第二噪声类型的噪声数据对应的降噪强度参数进行合并处理，得到用于对样本带噪音频数据进行降噪处理的样本降噪强度参数，此处的合并处理可以为求和处理或平均化处理等等。在一些实施例中，上述根据样本带噪音频数据中的样本音频数据的应用场景，生成用于对样本带噪音频数据进行降噪处理的样本降噪强度参数，包括：计算机设备可以获取在样本音频数据的应用场景下关于音频数据的质量需求等级；该质量需求等级反映该应用场景下针对音频数据的质量要求，即该质量需求等级越高，该应用场景下针对音频数据的质量要求越高；即该质量需求等级越低，该应用场景下针对音频数据的质量要求越低。在一些实施例中，根据质量需求等级确定用于对样本带噪音频数据进行降噪处理的样本降噪强度参数，如质量需求等级越低，样本降噪强度参数越大；质量需求等级越高，样本降噪强度参数越小，避免过度对样本带噪音频数据降噪处理，造成样本带噪音频数据中的音频数据受到损失，提高音频数据的质量。

S203、根据该样本降噪强度参数、该样本音频数据以及样本噪声数据生成标注语音增强数据。

本申请实施例中，该样本降噪强度参数用于将样本带噪音频数据中的样本噪声数据进行抑制处理，因此，计算机设备可以根据该样本降噪强度参数，该样本音频数据以及样本噪声数据生成标注语音增强数据。

在一些实施例中，上述步骤 S203 包括：计算机设备可以根据该样本降噪强度参数生成降噪因子，将降噪因子与样本噪声数据之间的乘积，确定为对该样本噪声数据进行降噪处理得到的处理后的样本噪声数据。该降噪因子可以为小于 1 的正数，当该样本降噪强度参数为小于 1 的正数，该降噪因子可以为样本降噪强度参数；当该样本降噪强度参数为大于 1 的正数，该降噪因子可以为对该样本降噪强度参数进行归一化处理得到的，如样本降噪强度参数为 δ_{snr} ，降噪因子可以为 $10^{-\frac{\delta_{snr}}{20}}$ 。在一些实施例中，计算机设备可以将该处理后的样本噪声数据与该样本音频数据进行组合（即求和处理），得到标注语音增强数据。

例如，假设标注语音增强数据为 y_n ，则该标注语音增强数据可以采用如下公式 (2) 表示：

$$y_n = s_n + d_n \cdot 10^{-\frac{\delta_{snr1}}{20}} \quad (2) ;$$

其中，公式 (2) 中的 δ_{snr1} 为样本降噪强度参数，该标注语音增强数据为初始降噪处理模型的优化训练的目标，根据公式 (2) 可知，初始降噪处理模型的优化训练的目标为：在降低样本带噪音频数据中的样本音频数据的损失的情况下，按照样本降噪强度

参数，对样本带噪音频数据中的样本噪声数据进行抑制处理。

S204、通过初始降噪处理模型基于该样本降噪强度参数，对该样本带噪音频数据进行降噪处理，得到预测语音增强数据。

5 在一些实施例中，计算机设备可以将该样本降噪增强度参数和样本带噪音频数据输入至初始降噪处理模型中，通过该初始降噪处理模型，基于该样本降噪强度参数，对该样本带噪音频数据进行降噪处理，得到预测语音增强数据。

需要说明的是，初始降噪处理模型基于该样本降噪强度参数，对该样本带噪音频数据进行降噪处理，得到预测语音增强数据的实现过程，可以参考目标降噪处理模型基于目标降噪强度参数，对原始带噪音频数据进行降噪处理，得到目标增强音频数据的实现过程。

10 在一些实施例中，初始降噪处理模型可以为深度神经网络、卷积神经网络、长短时间记忆网络等中的一种。

S205、根据该预测语音增强数据和该标注语音增强数据，对该初始降噪处理模型进行优化训练，得到目标降噪处理模型。

15 在一些实施例中，如果该预测语音增强数据与该标注语音增强数据之间的差异比较小，表明初始降噪处理模型的降噪处理准确度比较高；如果该预测语音增强数据与该标注语音增强数据之间的差异比较大，表明初始降噪处理模型的降噪处理准确度比较低。也就是说，预测语音增强数据和标注语音增强数据可用于衡量初始降噪处理模型的降噪处理准确度，因此，计算机设备可以根据该预测语音增强数据和标注语音增强数据，对该初始降噪处理模型进行优化训练，得到目标降噪处理模型，提高目标降噪处理模型的降噪处理准确度。

25 在一些实施例中，上述步骤 S205 包括：计算设备可以获取该初始降噪处理模型的误差函数，将该预测语音增强数据和标注语音增强数据代入至误差函数中，得到该初始降噪处理模型的降噪处理误差。初始降噪处理模型的误差函数可以为均方误差函数或交叉熵函数等等，该降噪处理误差用于衡量初始降噪处理模型的降噪处理准确度，即降噪处理误差越大，表明初始降噪处理模型的降噪处理准确度越低；降噪处理误差越小，表明初始降噪处理模型的降噪处理准确度越高。之后计算机设备可以检测预测语音增强数据中所包含的噪声数据的强度的噪声变化特征，根据预测语音增强数据中所包含的噪声数据的强度的噪声变化特征，确定该预测语音增强数据中所包含的噪声数据的稳定度。此处该预测语音增强数据中所包含的噪声数据的稳定度：用于反映预测语音增强数据中残留的噪声数据的稳定度，也用于反映初始降噪处理模型的降噪处理稳定度。然后，可以根据该降噪处理误差和该稳定度，对该初始降噪处理模型的模型参数进行调整，得到该目标降噪处理模型，以便可以提高目标降噪处理模型的降噪处理准确度以及降噪处理的稳定度。

35 在一些实施例中，上述根据该降噪处理误差和该稳定度，对该初始降噪处理模型的模型参数进行调整，得到该目标降噪处理模型，包括：计算机设备可以根据该降噪处理误差确定该初始降噪处理模型的收敛状态；该初始降噪处理模型的收敛状态用于反映该初始降噪处理模型的降噪处理误差是否达到最小值，收敛状态包括已收敛状态或未收敛状态。通常该降噪处理误差小于误差阈值时，计算机设备可以确定该初始降噪处理模型的收敛状态为已收敛状态，即该初始降噪处理模型的降噪处理误差为最小值；如果该降噪处理误差大于或等于误差阈值时，计算机设备可以确定该初始降噪处理模型的收敛状态为未收敛状态，即该初始降噪处理模型的降噪处理误差大于最小值。因此，若该初始降噪处理模型的收敛状态为已收敛状态，且该稳定度大于或等于稳定度阈值，表明该初始降噪处理模型的降噪处理误差达到最小值，或该初始降噪处理模型的降噪处理稳定度

40

比较高。这时，不需要对初始降噪处理模型的模型参数进行调整，计算机设备可以将该初始降噪处理模型确定为目标降噪处理模型。此处的该稳定度阈值可以手动设置的，或者，该稳定度阈值可以是指根据该样本噪声数据的采集场景或样本音频数据的应用场景确定的。同理，若该初始降噪处理模型的收敛状态为未收敛状态，或该稳定度小于稳定度阈值，表明该初始降噪处理模型的降噪处理误差未达到最小值，或该初始降噪处理模型的降噪处理稳定度比较差。这时，计算机设备可以根据该降噪处理误差，对该初始降噪处理模型的模型参数进行调整；直到调整后的初始降噪处理模型的收敛状态为已收敛状态，且对应的稳定度大于或等于该稳定度阈值，将该调整后的初始降噪处理模型，确定为该目标降噪处理模型。通过根据稳定度和收敛状态，对初始降噪处理模型的模型参数进行调整，有利于训练得到降噪处理准确度高，且降噪处理稳定度高的目标降噪处理模型。

S206、获取待处理的原始带噪音频数据，以及与该原始带噪音频数据关联的目标场景参数。

S207、根据该目标场景参数，确定用于对该原始带噪音频数据进行降噪处理的目标降噪强度参数。

需要说明的是，本申请实施例中对步骤 S206 的解释说明可以参考上述对步骤 S101 的解释说明，本申请实施例中对步骤 S207 的解释说明可以参考上述对步骤 S102 的解释说明。

S208、通过目标降噪处理模型根据该目标降噪强度参数，对该原始带噪音频数据进行降噪处理，得到目标增强音频数据。

在一些实施例中，该目标降噪处理模型可以包括特征提取网络以及语音解析网络以及语音生成网络，上述步骤 S208 可以包括：计算机设备可以通过目标降噪处理模型的特征提取网络，提取该原始带噪音频数据的频域信号，该原始带噪音频数据的频域信号反映原始带噪音频数据的频域特征，如该原始带噪音频数据的频域信号反映原始带噪音频数据的频率与信号强度之间的变化特征。然后计算机设备可以通过该目标降噪处理模型的语音解析网络，对该原始带噪音频数据的频域信号进行解析，得到该原始带噪音频数据的余弦变换掩码；该余弦变换掩码用于反映该原始带噪音频数据中的音频数据的占比，即该余弦变换掩码用于反映该原始带噪音频数据中的音频数据在原始带噪音频数据中的占比。然后，可以通过该目标降噪处理模型的语音生成网络，根据该原始带噪音频数据的余弦变换掩码、该原始带噪音频数据的频域信号以及该目标降噪强度参数，生成目标增强音频数据。如，计算机设备可以对该目标降噪强度参数进行指数运算，得到降噪因子，如 $10^{-\frac{\delta_{snr}2}{20}}$ ，获取 1 与余弦变换掩码之间的差值，获取该差值与降噪因子之间的乘积，获取该乘积与余弦变换掩码之间的和，得到降噪值，将该降噪值与该原始带噪音频数据的频域信号的乘积，确定频域增强音频数据，对该频域音频数据进行时域变换，得到目标增强音频数据。通过目标降噪处理模型，对原始带噪音频数据进行降噪处理，可避免对原始带噪音频数据中的音频数据造成损失的问题，以及避免目标增强音频数据中的噪声残留不稳定的问题，提高目标增强音频数据中的噪声残留的稳定度和平滑度，提高目标增强音频数据中的音频数据的感知度。

在一些实施例中，上述通过该目标降噪处理模型的语音解析网络，对该原始带噪音频数据的频域信号进行解析，得到该原始带噪音频数据的余弦变换掩码，包括：计算机设备可以通过该语音解析网络中的编码层，按照第一语音特征提取模式对该原始带噪音频数据的频域信号进行语音特征提取，得到第一关键语音特征，按照第二语音特征提取模式对该第一关键语音特征进行语音特征提取，得到第二关键语音特征，按照第三语音特征提取模式对该第一关键语音特征和该第二关键语音特征进行语音特征提取，得到第

三关键语音特征。对该第一关键语音特征、该第二关键语音特征以及该第三关键语音特征进行解析，得到该原始带噪音频数据的余弦变换掩码。通过根据不同语音特征提取模式，提取原始带噪音频数据中的语音数据（即音频数据），避免原始带噪音频数据中的语音特征丢失，造成语音数据损失的问题。

5 需要说明的是，由于用户的声带振动产生一般低于 500Hz 的基音及其谐波信号，因此，计算机设备可以根据原始带噪音频数据的频率分布特征提取关键语音特征。通常语音数据的频谱往往会在基频（基音）和谐波信号中出现峰值，而噪声数据的频谱则相对平坦，因此，计算机设备可以根据原始带噪音频数据的频谱平坦度，提取关键语音特征。同时，噪声数据的频谱比语音数据的频谱更稳定，即噪声数据的频谱的整体波形形状在
10 任何给定阶段都倾向于保持相同，因此，可以通过原始带噪音频数据的频谱模板差异来区分噪声数据和语音数据，即计算机设备可以根据原始带噪音频数据的频谱模板差异，来提取关键语音特征。本申请实施例中的第一语音特征提取模式、第二语音特征提取模式以及第三语音提取模式分别为从不同角度提取关键语音特征的方式，如，第一语音特征提取模式、第二语音特征提取模式以及第三语音提取模式分别为上述基于频率分布特征的提取模式、基于频谱平坦度的提取模式以及频谱模板差异的提取模式中的一种，
15 第一语音特征提取模式、第二语音特征提取模式以及第三语音提取模式可以不相同，也可以存在至少两个提取模式相同。

在一些实施例中，上述对该第一关键语音特征、该第二关键语音特征以及该第三关键语音特征进行解析，得到该原始带噪音频数据的余弦变换掩码，包括：计算机设备可以
20 通过该语音解析网络中的时序解析层，对该第三关键语音特征进行解析，得到该原始带噪音频数据的时序信息，该原始带噪音频数据的时序信息反映原始带噪音频数据中的关键语音特征与时间之间的关系。在一些实施例中，计算机设备可以通过该语音解析网络中的解码层，根据该时序信息、该第一关键语音特征、该第二关键语音特征以及该第三关键语音特征进行解析，得到该原始带噪音频数据的余弦变换掩码。

25 在一些实施例中，上述通过该目标降噪处理模型的语音生成网络，根据该原始带噪音频数据的余弦变换掩码、该原始带噪音频数据的频域信号以及该目标降噪强度参数，生成目标增强音频数据，包括：计算机设备可以通过该目标降噪处理模型的语音生成网络，根据该原始带噪音频数据的频域信号，确定该原始带噪音频数据的原始信噪比，根据该原始信噪比以及该目标降噪强度参数，生成该原始带噪音频数据降噪后的增强信噪比。例如，假设目标降噪强度参数为 δ_{snr2} ，目标降噪强度参数 δ_{snr2} 的单位为 dB，代表的物理意义是原始带噪音频数据需要提升的信噪比。原始带噪音频数据的原始信噪比为 λ ，则该原始带噪音频数据降噪后的增强信噪比可以为 $\lambda + \delta_{snr2}$ 。之后计算机设备可以根据该增强信噪比、该原始带噪音频数据的余弦变换掩码以及该原始带噪音频数据的频域信号，生成目标增强音频数据。通过定量地抑制原始带噪音频数据中的噪声数据，
30 避免原始带噪音频数据的音频数据损失，提高目标增强音频数据中的噪声残留的稳定度和平滑度，提高目标增强音频数据中的音频数据的感知度。

例如，如图 5 所示，该目标降噪处理模型包括特征提取网络 501、语音解析网络 502 以及语音生成网络 503，该特征提取网络用于将时域的原始带噪音频数据进行频域转换，得到原始带噪音频数据的频域信号。在一些实施例中，该特征提取网络首先对原始带噪
40 音频数据 x_n 进行重采样操作，将各种采样率类型的原始带噪音频数据重采样至 48kHz。重采样完成后，对重采样后的原始带噪音频数据进行分帧加窗处理，如可以按照帧长 1024、帧移 512，将重采样后的原始带噪音频数据分割成多个带噪音频数据片段，采用汉明窗分别对多个带噪音频数据片段进行调制。分帧加窗处理结束后，对调制后的多个带噪音频数据片段进行离散余弦变换（Discrete cosine transform, DCT）操作，得到原始

带噪音频数据的频域信号 X_k 。对原始带噪音频数据进行分帧加窗与余弦变换操作结合又可称之为短时余弦变换 (Short-time discrete cosine transform, SDCT)。语音解析网络 502 用于提前原始带噪音频数据的余弦变换掩码, 该语音解析网络可以为深度学习网络模块, 深度学习网络模块包括编码层 5021、时序解析层 5022 以及解码层 5023, 编码层 5021 可以由多个二维卷积构成, 每一个二维卷积的卷积核大小 kernel size 为 (5,2), 这代表频域视野为 5, 时域视野为 2, 每一帧信号特征 (即当前的带噪音频数据片段对应的频域信号) 的分析处理会参考前一帧信号 (即前一个带噪音频数据片段对应的频域信号)。而二维卷积的步长 stride 为 (2,1), 这能够让频域信号数目逐层减半, 时域帧数不变, 起到了降维减小计算量的作用。如图 5 所示, 以该编码层 5021 包括三个二维卷积为例, 分别为二维卷积 1、二维卷积 2 以及二维卷积 3。二维卷积 1、二维卷积 2 以及二维卷积 3 分别提取原始带噪音频数据的第一关键语音特征、第二关键语音特征以及第三关键语音特征。解码层 5023 部分主要由转置二维卷积 (ConvTranspose2d) 为内核的 DecTConv2d 组成, 图 5 以解码层包括三个转置二维卷积为例, 分别为转置二维卷积 1、转置二维卷积 2 以及转置二维卷积 3, 每一层的 DecTConv2d 参数都与对应的二维卷积相同, 实现了信号维度的还原。在编码层与解码层之间, 采用了时序解析层 5022, 该时序解析层 5022 可以由门控循环单元 (Gated Recurrent Units, GRU) 堆叠构成的循环神经网络模块 RNNs, RNNs 的作用主要提取并分析音频信号帧间时序信息。所以深度学习网络模块的工作流程是, 编码层接受来自特征提取网络的原始带噪音频数据的频域信号, 然后经过二维卷积逐层提取高维度特征 (即第一关键语音特征、第二关键语音特征以及第三关键语音特征)。对应的输出通过跳连接方式给到转置二维卷积。RNNs 接受来自最后一层二维卷积 3 输出的第三关键语音特征, 进行时序信息提取分析, 并且将输入给到解码层。而解码层接受来自 RNNs 和编码层的输出, 经过逐层维度升高处理, 最终得到余弦变换掩码 \tilde{m}_k 。

在一些实施例中, 上述根据该增强信噪比以及该频域信号, 生成目标增强音频数据, 包括: 计算机设备可以根据该增强信噪比、该原始带噪音频数据的余弦变换掩码对该原始带噪音频数据的频域信号进行降噪处理, 得到频域增强音频数据, 对该频域增强音频数据进行变换, 得到时域增强音频数据, 将该时域增强音频数据, 确定为目标增强音频数据。

例如, 假设该原始带噪音频数据的频域信号为 X_k , 原始带噪音频数据中的音频数据的频域信号为 Y_k , 原始带噪音频数据中的噪声数据的频域信号为 D_k , 该原始带噪音频数据的频域信号可以采用如下公式 (3) 表示:

$$X_k = Y_k + D_k \quad (3);$$

其中, 公式 (3) 中的 k 为原始带噪音频数据的第 k 个采样点, k 为大于 1 的正整数, 基于公式 (3), 原始带噪音频数据的原始信噪比可以采用如下公式 (4) 表示:

$$\lambda = 10 \log_{10} \left(\frac{Y_k^2}{D_k^2} \right) \quad (4);$$

假设频域增强音频数据为 \hat{X}_k , 频域增强音频数据中的音频数据的频域信号为 \hat{Y}_k , 频域增强音频数据中的噪声数据的频域信号为 \hat{D}_k , 该频域增强音频数据的频域信号可以采用如下公式 (5) 表示:

$$\hat{X}_k = \hat{Y}_k + \hat{D}_k \quad (5);$$

在一些实施例中, 基于公式 (5), 该原始带噪音频数据降噪后的增强信噪比可以采用如下公式 (6) 表示:

$$\lambda + \delta_{snr2} = 10 \log_{10} \left(\frac{\hat{Y}_k^2}{\hat{D}_k^2} \right) \quad (6);$$

由于原始带噪音频数据的余弦变换掩码反映原始带噪音频数据中的音频数据的占

比, 因此, 该原始带噪音频数据的频域信号 X_k 与频域增强音频数据中的音频数据的频域信号 \hat{Y}_k 之间的关系可以采用如下公式 (7) 表示:

$$\hat{Y}_k = X_k \cdot \hat{m}_k \quad (7);$$

其中, 公式 (7) 中的 \hat{m}_k 为原始带噪音频数据的余弦变换掩码, 基于公式 (4)、公式 (6) 和公式 (7), 频域增强音频数据中的噪声数据的频域信号 \hat{D}_k 可以采用如下公式 (8) 表示:

$$(8);$$

在一些实施例中, 基于公式 (7) 和公式 (8), 上述公式 (5) 可以变换为如下公式 (9):

$$\hat{X}_k = X_k \cdot \left[\hat{m}_k + (1 - \hat{m}_k) \cdot 10^{-\frac{\delta_{snr2}}{20}} \right] \quad (9);$$

然后, 计算机设备对上述公式 (9) 进行时域变换, 得到目标增强音频数据。

在一些实施例中, 本申请实施例中通过引入目标降噪强度参数, 定量地控制算法对原始带噪音频数据的噪声处理力度。针对原始带噪音频数据不同的应用场景和/或采集场景, 可以灵活配置该目标降噪强度参数, 提升了本申请对不同场景的适配性, 提升了本申请的泛化性。本申请可以做到覆盖大多数语音数据应用场景和实际需求, 减小算法开发难度和系统复杂性。由于申请中采用新的模型训练模式来满足可控降噪力度的需求, 不以纯净语音当做目标增强语音, 而是按照一定信噪比 (样本降噪强度参数) 来混合语音信号 (即样本音频数据) 和噪声信号 (即样本噪声数据) 得到目标增强语音 (即标注语音增强数据), 这在一定程度上规避了常规语音增强降噪算法容易出现的语音损失问题和噪声残留不连续问题。

接下来给出本申请在不同降噪强度参数下的降噪效果表现, 按照信噪比范围 [-10,30]dB 生成了一批测试数据 (即带噪音频数据), 将降噪强度参数 δ_{snr} 分别设置为 5dB, 10dB, 20dB, 40dB, 选取可感知语音质量参数 (Perceptual evaluation of speech quality, PESQ) 和尺度不变性信噪比参数 (Scale-Invariant Source-to-Noise Ratio, SI-SNR) 两项常用的语音增强降噪质量评价指标作为降噪效果参考指标。图 6 为带噪音频数据在不同降噪强度参数下的 PESQ 得分情况, 图 6 的横坐标表示带噪音频数据的原始信噪比, 纵坐标表示带噪音频数据在根据降噪强度参数进行降噪处理后的 PESQ 得分, 其中每个原始信噪比对应 5 个矩形, 同一原始信噪比下, 从左到右的第一个矩形的长度表示带噪音频数据不进行降噪处理的 PESQ 得分, 第二个矩形到第五个矩形的长度分别表示带噪音频数据在根据降噪强度参数 5dB, 10dB, 20dB, 40dB 进行降噪处理后的 PESQ 得分。从图 6 可知, 带噪音频数据根据降噪强度参数处理后的 PESQ 得分高于带噪音频数据不进行降噪处理的 PESQ 得分, 这种情况在带噪音频数据的原始信噪比大于 4dB 时特别明显。同时, 在同一原始信噪比下, 降噪强度参数越大, 带噪音频数据根据降噪强度参数处理后的 PESQ 得分越高; 降噪强度参数越小, 带噪音频数据根据降噪强度参数处理后的 PESQ 得分越低。

图 7 为带噪音频数据在不同降噪强度参数下的 SI-SNR 得分情况, 图 7 的横坐标表示带噪音频数据的原始信噪比, 纵坐标表示带噪音频数据在根据降噪强度参数进行降噪处理后的 SI-SNR 得分, 其中每个原始信噪比对应 5 个矩形, 同一原始信噪比下, 从左到右的第一个矩形的长度表示带噪音频数据不进行去噪处理的 SI-SNR 得分, 第二个矩形到第五个矩形的长度分别表示带噪音频数据在根据降噪强度参数 5dB, 10dB, 20dB, 40dB 进行降噪处理后的 SI-SNR 得分。从图 7 可知, 带噪音频数据根据降噪强度参数处理后的 SI-SNR 得分高于带噪音频数据不进行降噪处理的 SI-SNR 得分, 这种情况在带噪音频数据的原始信噪比大于 4dB 时特别明显。同时, 在同一原始信噪比下, 降噪强度参数越大, 带噪音频数据根据降噪强度参数处理后的 SI-SNR 得分越高; 降噪强度参数

越小，带噪音频数据根据降噪强度参数处理后的 SI-SNR 得分越低。

在本申请实施例中，通过原始带噪音频数据关联的目标场景参数，自适应地确定用于对原始带噪音频数据进行降噪处理的目标降噪强度参数，基于目标降噪强度参数，定量地降低原始带噪音频数据中的噪声含量。即此处的目标场景参数反映原始带噪音频数据的应用场景和采集场景中的至少一种，目标降噪强度参数反映对原始带噪音频数据中的噪声抑制的力度。换言之，通过原始带噪音频数据的应用场景中针对音频数据的实际需求（和/或原始带噪音频数据的采集场景中的噪声分布情况），定量地降低原始带噪音频数据中的噪声含量，并接受一定程度的噪声残留，不需要完全分离原始带噪音频数据中的噪声数据和音频数据，来彻底抑制噪声，避免降噪时对有效音频数据造成损失，提高音频数据的质量，提高噪声处理的灵活性。

请参见图 8，是本申请实施例提供的一种音频数据处理装置的结构示意图。上述音频数据处理装置可以是运行于网络设备中的一个计算机程序（包括程序代码），例如该音频数据处理装置为一个应用软件；该装置可以用于执行本申请实施例提供的方法中的相应步骤。如图 8 所示，该音频数据处理装置可以包括：

获取模块 801，配置为获取待处理的原始带噪音频数据，以及与所述原始带噪音频数据关联的目标场景参数；确定模块 802，配置为根据所述目标场景参数，确定用于对所述原始带噪音频数据进行降噪处理的目标降噪强度参数；处理模块 803，配置为根据所述目标降噪强度参数，对所述原始带噪音频数据进行降噪处理，得到目标增强音频数据。

在一些实施例中，确定模块 802 包括获取单元 81a 和确定单元 82a；获取单元 81a，配置为若所述目标场景参数用于确定所述原始带噪音频数据的应用场景，则获取在所述应用场景下关于音频数据的质量需求等级；确定单元 82a，配置为根据所述质量需求等级，确定用于对所述原始带噪音频数据进行降噪处理的目标降噪强度参数。

获取单元 81a，配置为若所述目标场景参数用于确定所述原始带噪音频数据的采集场景，获取在所述采集场景中历史时间段内的历史噪声数据；确定单元 82a，配置为根据所述历史噪声数据，确定用于对所述原始带噪音频数据进行降噪处理的目标降噪强度参数。

在一些实施例中，确定单元 82a 根据所述历史噪声数据，确定用于对所述原始带噪音频数据进行降噪处理的目标降噪强度参数，包括：从所述历史噪声数据中，确定所述采集场景在所述历史时间段内的噪声数据对应噪声类型和噪声变化特征；根据所述噪声类型和所述噪声变化特征，确定用于对所述原始带噪音频数据进行降噪处理的目标降噪强度参数。

在一些实施例中，采集场景在所述历史时间段内的噪声数据对应 M 个噪声类型，确定单元 82a 根据所述噪声类型和所述噪声变化特征，确定用于对所述原始带噪音频数据进行降噪处理的目标降噪强度参数，包括：基于所述 M 个噪声类型分别对应的噪声变化特征，确定用于对原始带噪音频数据进行降噪处理的 M 个候选降噪强度参数；将所述 M 个候选降噪强度参数确定为目标降噪强度参数；或者，对所述 M 个候选噪声强度参数进行均值计算，得到目标降噪强度参数。

处理模块 803 包括提取单元 83a、解析单元 84a 以及生成单元 85a；提取单元 83a，配置为通过目标降噪处理模型的特征提取网络，提取所述原始带噪音频数据的频域信号；解析单元 84a，配置为通过所述目标降噪处理模型的语音解析网络，对所述原始带噪音频数据的频域信号进行解析，得到所述原始带噪音频数据的余弦变换掩码；所述余弦变换掩码用于反映所述原始带噪音频数据中的音频数据的占比；生成单元 85a，配置为通过所述目标降噪处理模型的语音生成网络，根据所述原始带噪音频数据的余弦变换

掩码、所述原始带噪音频数据的频域信号以及所述目标降噪强度参数，生成目标增强音频数据。

5 在一些实施例中，解析单元 84a 通过所述目标降噪处理模型的语音解析网络，对所述原始带噪音频数据的频域信号进行解析，得到所述原始带噪音频数据的余弦变换掩码，包括：通过所述语音解析网络中的编码层，按照第一语音特征提取模式对所述原始带噪音频数据的频域信号进行语音特征提取，得到第一关键语音特征；按照第二语音特征提取模式对所述第一关键语音特征进行语音特征提取，得到第二关键语音特征；按照第三语音特征提取模式对所述第一关键语音特征和所述第二关键语音特征进行语音特征提取，得到第三关键语音特征；对所述第一关键语音特征、所述第二关键语音特征以及所述第三关键语音特征进行解析，得到所述原始带噪音频数据的余弦变换掩码。

10 在一些实施例中，解析单元 84a 对所述第一关键语音特征、所述第二关键语音特征以及所述第三关键语音特征进行解析，得到所述原始带噪音频数据的余弦变换掩码，包括：通过所述语音解析网络中的时序解析层，对所述第三关键语音特征进行解析，得到所述原始带噪音频数据的时序信息；通过所述语音解析网络中的解码层，根据所述时序信息、所述第一关键语音特征、所述第二关键语音特征以及所述第三关键语音特征进行解析，得到所述原始带噪音频数据的余弦变换掩码。

15 在一些实施例中，生成单元 85a 通过所述目标降噪处理模型的语音生成网络，根据所述原始带噪音频数据的余弦变换掩码、所述原始带噪音频数据的频域信号以及所述目标降噪强度参数，生成目标增强音频数据，包括：通过所述目标降噪处理模型的语音生成网络，根据所述原始带噪音频数据的频域信号，确定所述原始带噪音频数据的原始信噪比；根据所述原始信噪比以及所述目标降噪强度参数，生成所述原始带噪音频数据降噪后的增强信噪比；根据所述增强信噪比、所述原始带噪音频数据的余弦变换掩码以及所述原始带噪音频数据的频域信号，生成目标增强音频数据。

20 在一些实施例中，生成单元 85a 根据所述增强信噪比、所述原始带噪音频数据的余弦变换掩码以及所述原始带噪音频数据的频域信号，生成目标增强音频数据，包括：根据所述增强信噪比、所述原始带噪音频数据的余弦变换掩码对所述原始带噪音频数据的频域信号进行降噪处理，得到频域增强音频数据；对所述频域增强音频数据进行变换，得到时域增强音频数据，将所述时域增强音频数据，确定为目标增强音频数据。

25 获取模块 801，还配置为获取样本音频数据以及样本噪声数据，根据所述样本音频数据和所述样本噪声数据生成样本带噪音频数据；获取用于对所述样本带噪音频数据进行降噪处理的样本降噪强度参数；生成模块 804，配置为根据所述样本降噪强度参数、所述样本音频数据以及样本噪声数据生成标注语音增强数据；处理模块 803，配置为通过初始降噪处理模型基于所述样本降噪强度参数，对所述样本带噪音频数据进行降噪处理，得到预测语音增强数据；训练模块 805，配置为根据所述预测语音增强数据和所述标注语音增强数据，对所述初始降噪处理模型进行优化训练，得到所述目标降噪处理模型。

30 在一些实施例中，训练模块 805 根据所述预测语音增强数据和所述标注语音增强数据，对所述初始降噪处理模型进行优化训练，得到所述目标降噪处理模型，包括：根据所述预测语音增强数据和所述标注语音增强数据，确定所述初始降噪处理模型的降噪处理误差；根据所述预测语音增强数据确定所述预测语音增强数据中所包含的噪声数据的稳定度；根据所述降噪处理误差和所述稳定度，对所述初始降噪处理模型的模型参数进行调整，得到所述目标降噪处理模型。

40 在一些实施例中，训练模块 805 根据所述降噪处理误差和所述稳定度，对所述初始降噪处理模型的模型参数进行调整，得到所述目标降噪处理模型，包括：根据所述降噪

处理误差确定所述初始降噪处理模型的收敛状态；若所述初始降噪处理模型的收敛状态为未收敛状态，或所述稳定度小于稳定度阈值，则根据所述降噪处理误差，对所述初始降噪处理模型的模型参数进行调整；直到调整后的初始降噪处理模型的收敛状态为已收敛状态，且对应的稳定度大于或等于所述稳定度阈值，将所述调整后的初始降噪处理模型，确定为所述目标降噪处理模型。

在一些实施例中，生成模块 804 根据所述样本降噪强度参数、所述样本音频数据以及样本噪声数据生成标注语音增强数据，包括：根据所述样本降噪强度参数对所述样本噪声数据进行降噪处理，得到处理后的样本噪声数据；将所述处理后的样本噪声数据与所述样本音频数据进行组合，得到标注语音增强数据。

根据本申请的一个实施例，前文所示的音频数据处理方法所涉及的步骤可由图 8 所示的音频数据处理装置中的各个模块来执行。例如，图 3 中所示的步骤 S101 可由图 8 中的获取模块 801 来执行，图 3 中所示的步骤 S102 可由图 8 中的确定模块 802 来执行；图 3 中所示的步骤 S103 可由图 8 中的处理模块 803 来执行。

根据本申请的一个实施例，图 8 所示的音频数据处理装置中的各个模块可以分别或全部合并为一个或若干个单元来构成，或者其中的某个（些）单元还可以再拆分为功能上更小的至少两个子单元，可以实现同样的操作，而不影响本申请的实施例的技术效果的实现。上述模块是基于逻辑功能划分的，在实际应用中，一个模块的功能也可以由至少两个单元来实现，或者至少两个模块的功能由一个单元实现。在本申请的其它实施例中，音频数据处理装置也可以包括其它单元，在实际应用中，这些功能也可以由其它单元协助实现，并且可以由至少两个单元协作实现。

根据本申请的一个实施例，可以通过在包括中央处理单元（CPU）、随机存取存储介质（RAM）、只读存储介质（ROM）等处理组件和存储组件的例如计算机的通用计算机设备上运行能够执行如前文所示的相应方法所涉及的各步骤的计算机程序（包括程序代码），来构造如图 8 中所示的音频数据处理装置，以及来实现本申请实施例的音频数据处理方法。上述计算机程序可以记载于例如计算机可读记录介质上，并通过计算机可读记录介质装载于上述计算设备中，并在其中运行。

在一些实施例中，通过原始带噪音频数据关联的目标场景参数，自适应地确定用于对原始带噪音频数据进行降噪处理的目标降噪强度参数，基于目标降噪强度参数，定量地降低原始带噪音频数据中的噪声含量。即此处的目标场景参数反映原始带噪音频数据的应用场景和采集场景中的至少一种，目标降噪强度参数反映对原始带噪音频数据中的噪声抑制的力度。换言之，通过原始带噪音频数据的应用场景中针对音频数据的实际需求（和/或原始带噪音频数据的采集场景中的噪声分布情况），定量地降低原始带噪音频数据中的噪声含量，并接受一定程度的噪声残留，不需要完全分离原始带噪音频数据中的噪声数据和音频数据，来彻底抑制噪声，避免降噪时对有效音频数据造成损失，提高音频数据的质量，提高噪声处理的灵活性。

可以理解的是，在本申请实施例中，涉及到原始带噪音频数据、目标增加音频数据等相关的数据，当本申请实施例运用到具体产品或技术中时，需要获得用户许可或者同意，且相关数据的收集、使用和处理需要遵守相关国家和地区的相关法律法规和标准。

请参见图 9，是本申请实施例提供的一种计算机设备的结构示意图。如图 9 所示，上述计算机设备 1000 可以为上述方法中的第一设备，可以是指终端或服务器，包括：处理器 1001，网络接口 1004 和存储器 1005，此外，上述计算机设备 1000 还可以包括：用户接口 1003，和至少一条通信总线 1002。其中，通信总线 1002 用于实现这些组件之间的连接通信。其中，在一些实施例中，用户接口 1003 可以包括显示屏（DiSPlay）、键盘（Keyboard），可选用户接口 1003 还可以包括标准的有线接口、无线接口。网络

接口 1004 可选的可以包括标准的有线接口、无线接口（如 WI-FI 接口）。存储器 1005 可以是高速 RAM 存储器，也可以是非易失性的存储器（non-volatile MeMory），例如至少一个磁盘存储器。存储器 1005 可选的还可以是至少一个在远离前述处理器 1001 的存储装置。如图 9 所示，作为一种计算机可读存储介质的存储器 1005 中可以包括操作
5 系统、网络通信模块、用户接口模块以及计算机应用程序。

在图 9 所示的计算机设备 1000 中，网络接口 1004 可提供网络通讯功能；而用户接口 1003 主要用于提供输入的接口；而处理器 1001 可以用于调用存储器 1005 中存储的计算机应用程序，以实现：获取待处理的原始带噪音频数据，以及与所述原始带噪音频数据关联的目标场景参数；根据所述目标场景参数，确定用于对所述原始带噪音频数据
10 进行降噪处理的目标降噪强度参数；根据所述目标降噪强度参数，对所述原始带噪音频数据进行降噪处理，得到目标增强音频数据。

在一些实施例中，处理器 1001 可以用于调用存储器 1005 中存储的计算机应用程序，以实现根据所述目标场景参数，确定用于对所述原始带噪音频数据进行降噪处理的目标降噪强度参数，包括：若所述目标场景参数反映所述原始带噪音频数据的应用场景，则
15 获取在所述应用场景下关于音频数据的质量需求等级；根据所述质量需求等级，确定用于对所述原始带噪音频数据进行降噪处理的目标降噪强度参数。

在一些实施例中，处理器 1001 可以用于调用存储器 1005 中存储的计算机应用程序，以实现根据所述目标场景参数，确定用于对所述原始带噪音频数据进行降噪处理的目标降噪强度参数，包括：若所述目标场景参数反映所述原始带噪音频数据的采集场景，则
20 获取在所述采集场景中历史时间段内的历史噪声数据；根据所述历史噪声数据，确定用于对所述原始带噪音频数据进行降噪处理的目标降噪强度参数。

在一些实施例中，处理器 1001 可以用于调用存储器 1005 中存储的计算机应用程序，以实现根据所述历史噪声数据，确定用于对所述原始带噪音频数据进行降噪处理的目标降噪强度参数，包括：从所述历史噪声数据中，统计所述采集场景在所述历史时间段内的
25 噪声数据对应噪声类型和噪声变化特征；根据所述噪声类型和所述噪声变化特征，确定用于对所述原始带噪音频数据进行降噪处理的目标降噪强度参数。

在一些实施例中，处理器 1001 可以用于调用存储器 1005 中存储的计算机应用程序，以实现根据所述目标降噪强度参数，对所述原始带噪音频数据进行降噪处理，得到目标增强音频数据，包括：通过目标降噪处理模型的特征提取网络，提取所述原始带噪音频
30 数据的频域信号；通过所述目标降噪处理模型的语音解析网络，对所述原始带噪音频数据的频域信号进行解析，得到所述原始带噪音频数据的余弦变换掩码；所述余弦变换掩码用于反映所述原始带噪音频数据中的音频数据的占比；通过所述目标降噪处理模型的语音生成网络，根据所述原始带噪音频数据的余弦变换掩码、所述原始带噪音频数据的频域信号以及所述目标降噪强度参数，生成目标增强音频数据。

在一些实施例中，处理器 1001 可以用于调用存储器 1005 中存储的计算机应用程序，以实现通过所述目标降噪处理模型的语音解析网络，对所述原始带噪音频数据的频域信号进行解析，得到所述原始带噪音频数据的余弦变换掩码，包括：通过所述语音解析网络中的编码层，按照第一语音特征提取模式对所述原始带噪音频数据的频域信号进行语音
35 特征提取，得到第一关键语音特征；按照第二语音特征提取模式对所述第一关键语音特征进行语音特征提取，得到第二关键语音特征；按照第三语音特征提取模式对所述第一关键语音特征和所述第二关键语音特征进行语音特征提取，得到第三关键语音特征；
40 对所述第一关键语音特征、所述第二关键语音特征以及所述第三关键语音特征进行解析，得到所述原始带噪音频数据的余弦变换掩码。

在一些实施例中，处理器 1001 可以用于调用存储器 1005 中存储的计算机应用程序，

以实现所述第一关键语音特征、所述第二关键语音特征以及所述第三关键语音特征进行解析，得到所述原始带噪音频数据的余弦变换掩码，包括：通过所述语音解析网络中的时序解析层，对所述第三关键语音特征进行解析，得到所述原始带噪音频数据的时序信息；通过所述语音解析网络中的解码层，根据所述时序信息、所述第一关键语音特征、
5 所述第二关键语音特征以及所述第三关键语音特征进行解析，得到所述原始带噪音频数据的余弦变换掩码。

在一些实施例中，处理器 1001 可以用于调用存储器 1005 中存储的计算机应用程序，以实现通过所述目标降噪处理模型的语音生成网络，根据所述原始带噪音频数据的余弦变换掩码、所述原始带噪音频数据的频域信号以及所述目标降噪强度参数，生成目标增强
10 增强音频数据，包括：通过所述目标降噪处理模型的语音生成网络，根据所述原始带噪音频数据的频域信号，确定所述原始带噪音频数据的原始信噪比；根据所述原始信噪比以及所述目标降噪强度参数，生成所述原始带噪音频数据降噪后的增强信噪比；根据所述增强信噪比、所述原始带噪音频数据的余弦变换掩码以及所述原始带噪音频数据的频域信号，生成目标增强音频数据。

在一些实施例中，处理器 1001 可以用于调用存储器 1005 中存储的计算机应用程序，以实现根据所述增强信噪比、所述原始带噪音频数据的余弦变换掩码以及所述原始带噪
15 音频数据的频域信号，生成目标增强音频数据，包括：根据所述增强信噪比、所述原始带噪音频数据的余弦变换掩码对所述原始带噪音频数据的频域信号进行降噪处理，得到频域增强音频数据；对所述频域增强音频数据进行变换，得到时域增强音频数据，将所述时域增强音频数据，确定为目标增强音频数据。

在一些实施例中，处理器 1001 可以用于调用存储器 1005 中存储的计算机应用程序，以实现：获取样本音频数据以及样本噪声数据，根据所述样本音频数据和所述样本噪声数据生成样本带噪音频数据；获取用于对所述样本带噪音频数据进行降噪处理的样本降
20 噪强度参数；根据所述样本降噪强度参数、所述样本音频数据以及样本噪声数据生成标注语音增强数据；通过初始降噪处理模型基于所述样本降噪强度参数，对所述样本带噪音频数据进行降噪处理，得到预测语音增强数据；根据所述预测语音增强数据和所述标注语音增强数据，对所述初始降噪处理模型进行优化训练，得到所述目标降噪处理模型。

在一些实施例中，处理器 1001 可以用于调用存储器 1005 中存储的计算机应用程序，以实现根据所述预测语音增强数据和所述标注语音增强数据，对所述初始降噪处理模型
25 进行优化训练，得到所述目标降噪处理模型，包括：根据所述预测语音增强数据和所述标注语音增强数据，确定所述初始降噪处理模型的降噪处理误差；根据所述预测语音增强数据确定所述预测语音增强数据中所包含的噪声数据的稳定度；根据所述降噪处理误差和所述稳定度，对所述初始降噪处理模型的模型参数进行调整，得到所述目标降噪处理模型。

在一些实施例中，处理器 1001 可以用于调用存储器 1005 中存储的计算机应用程序，以实现根据所述降噪处理误差和所述稳定度，对所述初始降噪处理模型的模型参数进行
30 调整，得到所述目标降噪处理模型，包括：根据所述降噪处理误差确定所述初始降噪处理模型的收敛状态；若所述初始降噪处理模型的收敛状态为未收敛状态，或所述稳定度小于稳定度阈值，则根据所述降噪处理误差，对所述初始降噪处理模型的模型参数进行调整；直到调整后的初始降噪处理模型的收敛状态为已收敛状态，且对应的稳定度大于
40 或等于所述稳定度阈值，将所述调整后的初始降噪处理模型，确定为所述目标降噪处理模型。

在一些实施例中，处理器 1001 可以用于调用存储器 1005 中存储的计算机应用程序，以实现根据所述样本降噪强度参数、所述样本音频数据以及样本噪声数据生成标注语音

增强数据,包括:根据所述样本降噪强度参数对所述样本噪声数据进行降噪处理,得到处理后的样本噪声数据;将所述处理后的样本噪声数据与所述样本音频数据进行组合,得到标注语音增强数据。

5 在一些实施例中,通过原始带噪音频数据关联的目标场景参数,自适应地确定用于对原始带噪音频数据进行降噪处理的目标降噪强度参数,基于目标降噪强度参数,定量地降低原始带噪音频数据中的噪声含量。即此处的目标场景参数反映原始带噪音频数据的应用场景和采集场景中的至少一种,目标降噪强度参数反映对原始带噪音频数据中的噪声抑制的力度。换言之,通过原始带噪音频数据的应用场景中针对音频数据的实际需求(和/或原始带噪音频数据的采集场景中的噪声分布情况),定量地降低原始带噪音频数据中的噪声含量,并接受一定程度的噪声残留,不需要完全分离原始带噪音频数据中的噪声数据和音频数据,来彻底抑制噪声,避免降噪时对有效音频数据造成损失,提高音频数据的质量,提高噪声处理的灵活性。

10 应当理解,本申请实施例中所描述的计算机设备可执行前文所对应实施例中对上述音频数据处理方法的描述,也可执行前文所对应实施例中对上述音频数据处理装置的描述。

15 此外,这里需要指出的是:本申请实施例还提供了一种计算机可读存储介质,且上述计算机可读存储介质中存储有前文提及的音频数据处理装置所执行的计算机程序,且上述计算机程序包括程序指令,当上述处理器执行上述程序指令时,能够执行前文对应实施例中对上述音频数据处理方法的描述。另外,对采用相同方法的有益效果描述,也不再进行重复说明。对于本申请所涉及的计算机可读存储介质实施例中未披露的技术细节,请参照本申请方法实施例的描述。

20 作为示例,上述程序指令可被部署在一个计算机设备上执行,或者被部署在一个地点的至少两个计算机设备上执行,又或者,在分布在至少两个地点且通过通信网络互连的至少两个计算机设备上执行,分布在至少两个地点且通过通信网络互连的至少两个计算机设备可以组成区块链网络。

25 上述计算机可读存储介质可以是前述任一实施例提供的音频数据处理装置或者上述计算机设备的中部存储单元,例如计算机设备的硬盘或中存。该计算机可读存储介质也可以是该计算机设备的外部存储设备,例如该计算机设备上配备的插接式硬盘,智能存储卡(SMART MEDIA CARD, SMC),安全数字(Secure digital, SD)卡,闪存卡(flash card)等。在一些实施例中,该计算机可读存储介质还可以既包括该计算机设备的中部存储单元也包括外部存储设备。该计算机可读存储介质用于存储该计算机程序以及该计算机设备所需的其他程序和数据。该计算机可读存储介质还可以用于暂时地存储已经输出或者将要输出的数据。

30 本申请实施例的说明书和权利要求书及附图中的术语“第一”、“第二”等是用于区别不同媒体内容,而非用于描述特定顺序。此外,术语“包括”以及它们任何变形,意图在于覆盖不排他的包含。例如包含了一系列步骤或单元的过程、方法、装置、产品或设备没有局限于已列出的步骤或模块,而是可选地还包括没有列出的步骤或模块,或可选地还包括对于这些过程、方法、装置、产品或设备固有的其他步骤单元。

35 在一些实施例中,本申请以上实施例如果需要用到用户信息等,需要获得用户许可或者同意,需要遵守相关国家和地区的相关法律法规。

40 本申请实施例还提供了一种计算机程序产品,包括计算机程序/指令,上述计算机程序/指令被处理器执行时实现前文对应实施例中对上述音频数据处理方法、解码方法的描述。另外,对采用相同方法的有益效果描述不再重复说明。对于本申请所涉及的计算机程序产品的实施例中未披露的技术细节,请参照本申请方法实施例的描述。

本领域普通技术人员可以意识到，结合本文中所公开的实施例描述的各示例的单元及算法步骤，能够以电子硬件、计算机软件或者二者的结合来实现，为了清楚地说明硬件和软件的可互换性，在上述说明中已经按照功能一般性地描述了各示例的组成及步骤。这些功能究竟以硬件还是软件方式来执行，取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能，但是这种实现不应认为超出本申请的范围。

本申请实施例提供的方法及相关装置是参照本申请实施例提供的方法流程图和/或结构示意图来描述的，可由计算机程序指令实现方法流程图和/或结构示意图的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。这些计算机程序指令可提供到通用计算机、专用计算机、嵌入式处理机或其他可编程网络连接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网络连接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/或结构示意图一个方框或多个方框中指定的功能的装置。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网络连接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品，该指令装置实现在流程图一个流程或多个流程和/或结构示意图一个方框或多个方框中指定的功能。这些计算机程序指令也可装载到计算机或其他可编程网络连接设备上，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/或结构示意图一个方框或多个方框中指定的功能的步骤。以上所揭露的仅为本申请较佳实施例而已，当然不能以此来限定本申请之权利范围，因此依本申请权利要求所作的等同变化，仍属本申请所涵盖的范围。

权利要求书

1、一种音频数据处理方法，应用于计算机设备，包括：

获取待处理的原始带噪音频数据，以及与所述原始带噪音频数据关联的目标场景参数：

5 根据所述目标场景参数，确定用于对所述原始带噪音频数据进行降噪处理的目标降噪强度参数；

根据所述目标降噪强度参数，对所述原始带噪音频数据进行降噪处理，得到目标增强音频数据。

2、如权利要求 1 所述的方法，其中，所述根据所述目标场景参数，确定用于对所述原始带噪音频数据进行降噪处理的目标降噪强度参数，包括：

10 若所述目标场景参数用于确定所述原始带噪音频数据的应用场景，获取在所述应用场景下关于音频数据的质量需求等级；

根据所述质量需求等级，确定用于对所述原始带噪音频数据进行降噪处理的目标降噪强度参数。

3、如权利要求 1 所述的方法，其中，所述根据所述目标场景参数，确定用于对所述原始带噪音频数据进行降噪处理的目标降噪强度参数，包括：

15 若所述目标场景参数用于确定所述原始带噪音频数据的采集场景，获取在所述采集场景中历史时间段内的历史噪声数据；

20 根据所述历史噪声数据，确定用于对所述原始带噪音频数据进行降噪处理的目标降噪强度参数。

4、如权利要求 3 所述的方法，其中，所述根据所述历史噪声数据，确定用于对所述原始带噪音频数据进行降噪处理的目标降噪强度参数，包括：

从所述历史噪声数据中，确定所述采集场景在所述历史时间段内的噪声数据对应噪声类型和噪声变化特征；

25 根据所述噪声类型和所述噪声变化特征，确定用于对所述原始带噪音频数据进行降噪处理的目标降噪强度参数。

5、根据权利要求 4 中所述的方法，其中，所述采集场景在所述历史时间段内的噪声数据对应 M 个噪声类型，所述根据所述噪声类型和所述噪声变化特征，确定用于对所述原始带噪音频数据进行降噪处理的目标降噪强度参数，包括：

30 基于所述 M 个噪声类型分别对应的噪声变化特征，确定用于对原始带噪音频数据进行降噪处理的 M 个候选降噪强度参数；

将所述 M 个候选降噪强度参数确定为目标降噪强度参数；或者，

对所述 M 个候选噪声强度参数进行均值计算，得到目标降噪强度参数。

6、如权利要求 1 所述的方法，其中，所述根据所述目标降噪强度参数，对所述原始带噪音频数据进行降噪处理，得到目标增强音频数据，包括：

35 获取目标降噪处理模型，所述目标降噪处理模型包括特征提取网络、语音解析网络和语音生成网络；

通过所述特征提取网络，提取所述原始带噪音频数据的频域信号；

40 通过所述语音解析网络，对所述原始带噪音频数据的频域信号进行解析，得到所述原始带噪音频数据的余弦变换掩码；所述余弦变换掩码用于反映所述原始带噪音频数据中的音频数据的占比；

通过所述语音生成网络，根据所述原始带噪音频数据的余弦变换掩码、所述原始带噪音频数据的频域信号以及所述目标降噪强度参数，生成目标增强音频数据。

7、如权利要求 6 所述的方法，其中，所述通过所述语音解析网络，对所述原始带噪音频数据的频域信号进行解析，得到所述原始带噪音频数据的余弦变换掩码，包括：

通过所述语音解析网络中的编码层，按照第一语音特征提取模式对所述原始带噪音频数据的频域信号进行语音特征提取，得到第一关键语音特征；

5 按照第二语音特征提取模式对所述第一关键语音特征进行语音特征提取，得到第二关键语音特征；

按照第三语音特征提取模式对所述第一关键语音特征和所述第二关键语音特征进行语音特征提取，得到第三关键语音特征；

10 对所述第一关键语音特征、所述第二关键语音特征以及所述第三关键语音特征进行解析，得到所述原始带噪音频数据的余弦变换掩码。

8、如权利要求 7 所述的方法，其中，所述对所述第一关键语音特征、所述第二关键语音特征以及所述第三关键语音特征进行解析，得到所述原始带噪音频数据的余弦变换掩码，包括：

15 通过所述语音解析网络中的时序解析层，对所述第三关键语音特征进行解析，得到所述原始带噪音频数据的时序信息；

通过所述语音解析网络中的解码层，根据所述时序信息、所述第一关键语音特征、所述第二关键语音特征以及所述第三关键语音特征进行解析，得到所述原始带噪音频数据的余弦变换掩码。

20 9、如权利要求 6 所述的方法，其中，所述通过所述语音生成网络，根据所述原始带噪音频数据的余弦变换掩码、所述原始带噪音频数据的频域信号以及所述目标降噪强度参数，生成目标增强音频数据，包括：

通过所述语音生成网络，根据所述原始带噪音频数据的频域信号，确定所述原始带噪音频数据的原始信噪比；

25 根据所述原始信噪比以及所述目标降噪强度参数，生成所述原始带噪音频数据降噪后的增强信噪比；

根据所述增强信噪比、所述原始带噪音频数据的余弦变换掩码以及所述原始带噪音频数据的频域信号，生成目标增强音频数据。

30 10、如权利要求 7 所述的方法，其中，所述根据所述增强信噪比、所述原始带噪音频数据的余弦变换掩码以及所述原始带噪音频数据的频域信号，生成目标增强音频数据，包括：

根据所述增强信噪比、所述原始带噪音频数据的余弦变换掩码对所述原始带噪音频数据的频域信号进行降噪处理，得到频域增强音频数据；

对所述频域增强音频数据进行变换，得到时域增强音频数据，将所述时域增强音频数据，确定为目标增强音频数据。

35 11、如权利要求 6 所述的方法，其中，所述方法还包括：

获取样本音频数据以及样本噪声数据，根据所述样本音频数据和所述样本噪声数据生成样本带噪音频数据；

获取用于对所述样本带噪音频数据进行降噪处理的样本降噪强度参数；

40 根据所述样本降噪强度参数、所述样本音频数据以及样本噪声数据生成标注语音增强数据；

通过初始降噪处理模型基于所述样本降噪强度参数，对所述样本带噪音频数据进行降噪处理，得到预测语音增强数据；

根据所述预测语音增强数据和所述标注语音增强数据，对所述初始降噪处理模型进行优化训练，得到所述目标降噪处理模型。

12、如权利要求 11 所述的方法，其中，所述根据所述预测语音增强数据和所述标注语音增强数据，对所述初始降噪处理模型进行优化训练，得到所述目标降噪处理模型，包括：

根据所述预测语音增强数据和所述标注语音增强数据，确定所述初始降噪处理模型的降噪处理误差；

根据所述预测语音增强数据确定所述预测语音增强数据中所包含的噪声数据的稳定度；

根据所述降噪处理误差和所述稳定度，对所述初始降噪处理模型的模型参数进行调整，得到所述目标降噪处理模型。

13、如权利要求 12 所述的方法，其中，所述根据所述降噪处理误差和所述稳定度，对所述初始降噪处理模型的模型参数进行调整，得到所述目标降噪处理模型，包括：

根据所述降噪处理误差确定所述初始降噪处理模型的收敛状态；

若所述初始降噪处理模型的收敛状态为未收敛状态，或所述稳定度小于稳定度阈值，则根据所述降噪处理误差，对所述初始降噪处理模型的模型参数进行调整；

直到调整后的初始降噪处理模型的收敛状态为已收敛状态，且对应的稳定度大于或等于所述稳定度阈值，将所述调整后的初始降噪处理模型，确定为所述目标降噪处理模型。

14、如权利要求 11 所述的方法，其中，所述根据所述样本降噪强度参数、所述样本音频数据以及样本噪声数据生成标注语音增强数据，包括：

根据所述样本降噪强度参数对所述样本噪声数据进行降噪处理，得到处理后的样本噪声数据；

将所述处理后的样本噪声数据与所述样本音频数据进行组合，得到标注语音增强数据。

15、一种音频数据处理装置，包括：

获取模块，配置为获取待处理的原始带噪音频数据，以及与所述原始带噪音频数据关联的目标场景参数；

确定模块，配置为根据所述目标场景参数，确定用于对所述原始带噪音频数据进行降噪处理的目标降噪强度参数；

处理模块，配置为根据所述目标降噪强度参数，对所述原始带噪音频数据进行降噪处理，得到目标增强音频数据。

16、一种计算机设备，包括存储器和处理器，所述存储器存储有计算机程序，所述处理器执行所述计算机程序时实现权利要求 1 至 14 中任一项所述的方法的步骤。

17、一种计算机可读存储介质，其上存储有计算机程序，所述计算机程序被处理器执行时实现权利要求 1 至 14 中任一项所述的音频数据处理方法的步骤。

18、一种计算机程序产品，包括计算机程序，该计算机程序被处理器执行时实现权利要求 1 至 14 中任一项所述的音频数据处理方法的步骤。

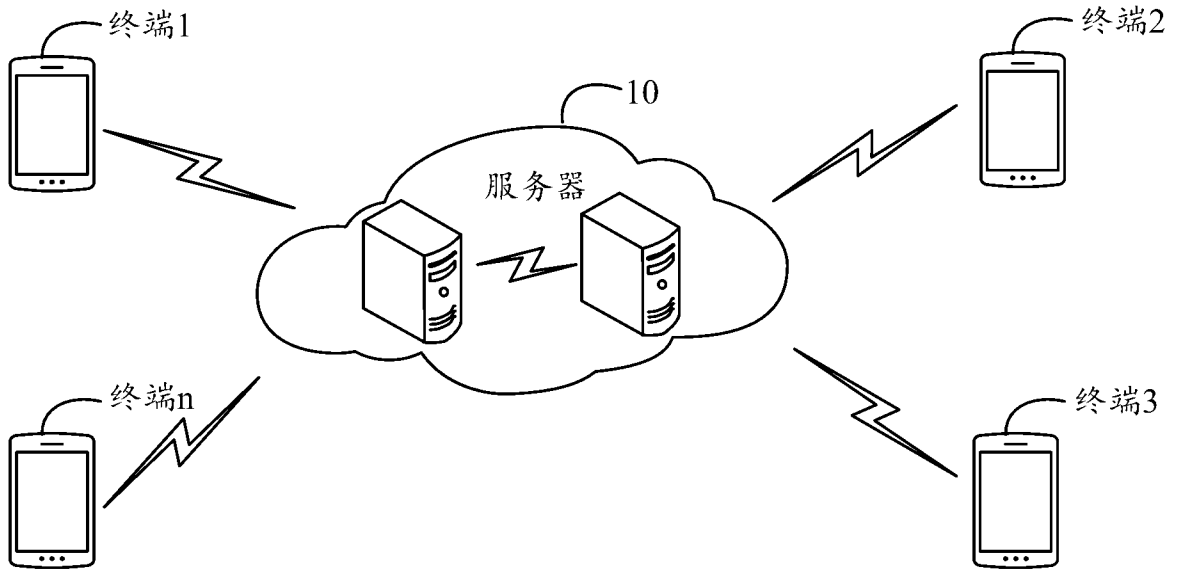


图 1

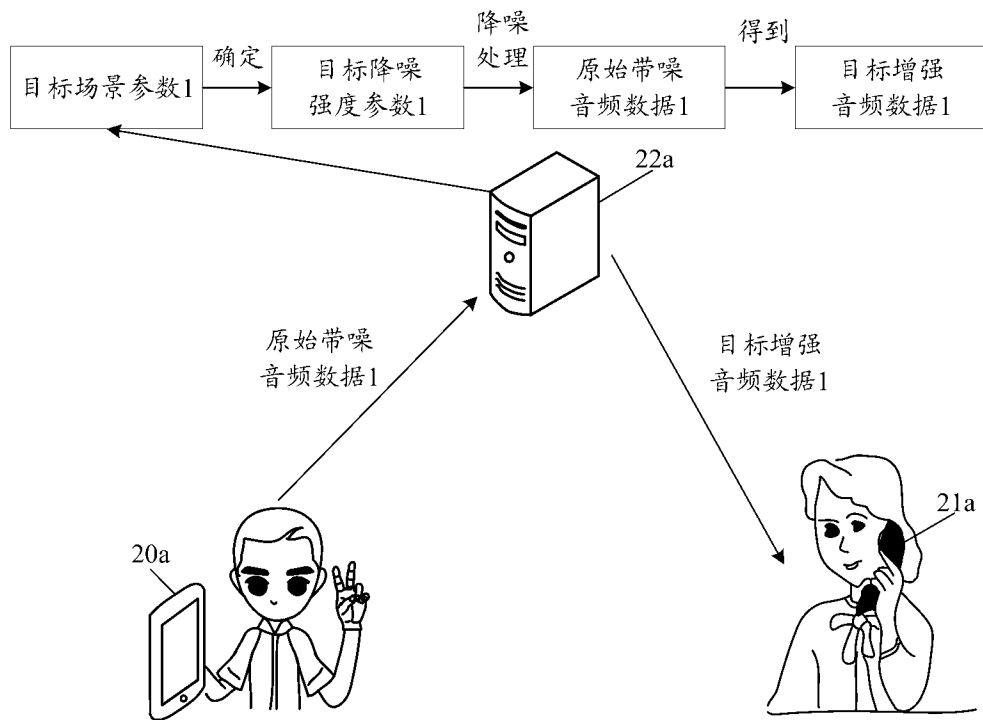


图 2

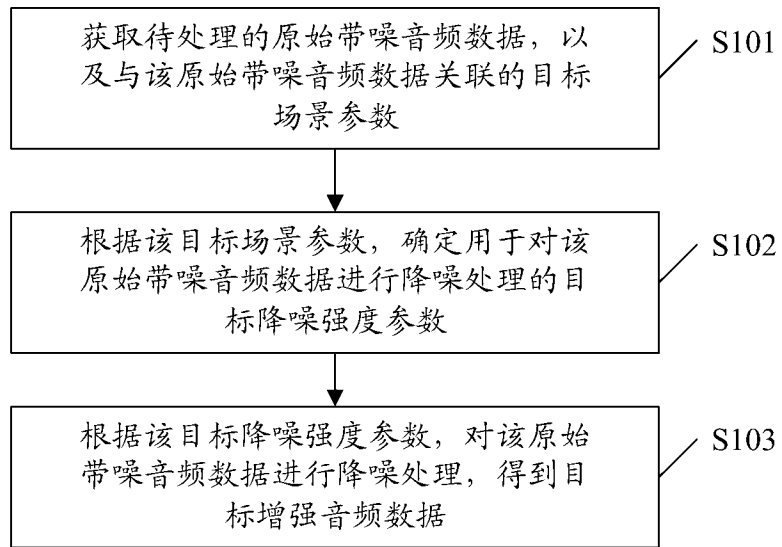


图 3

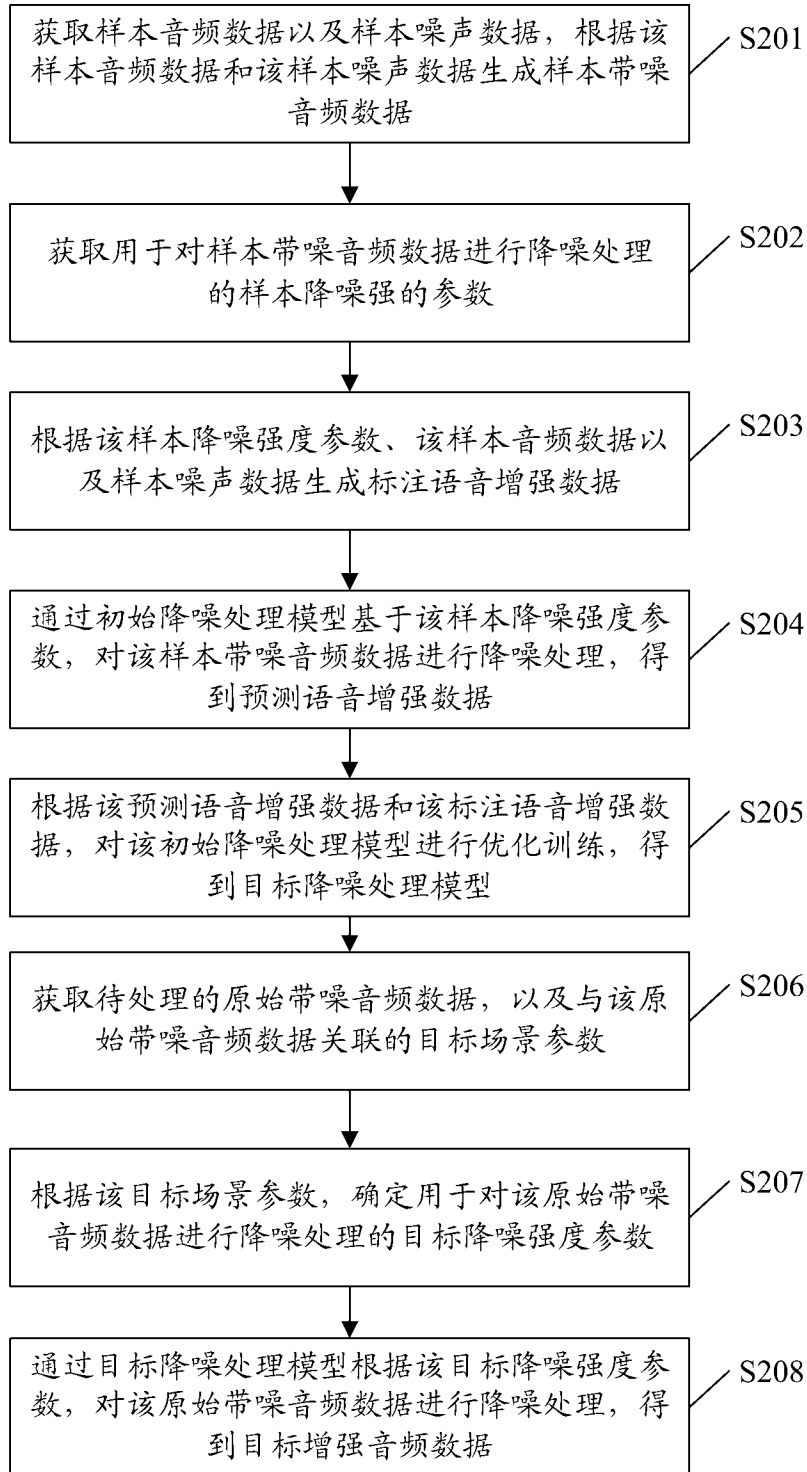


图 4

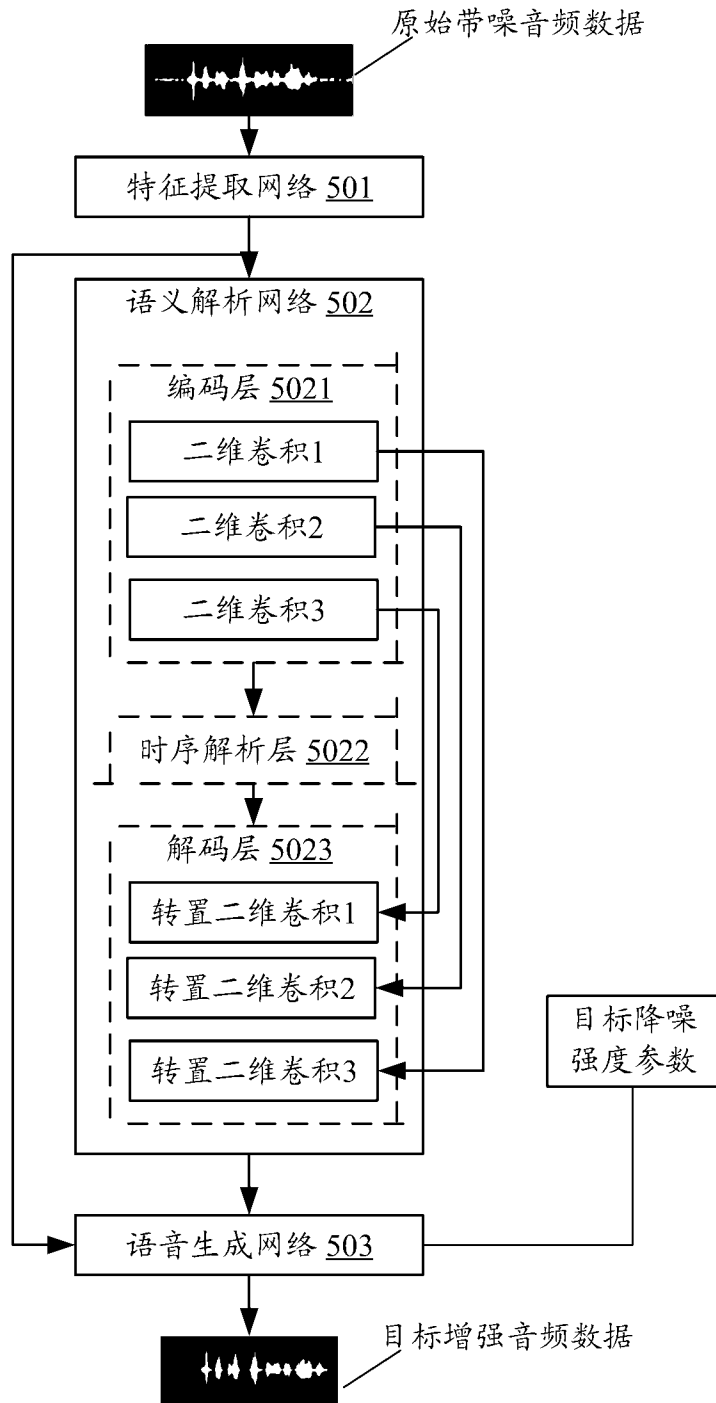


图 5

不同降噪强度参数下PESQ得分情况

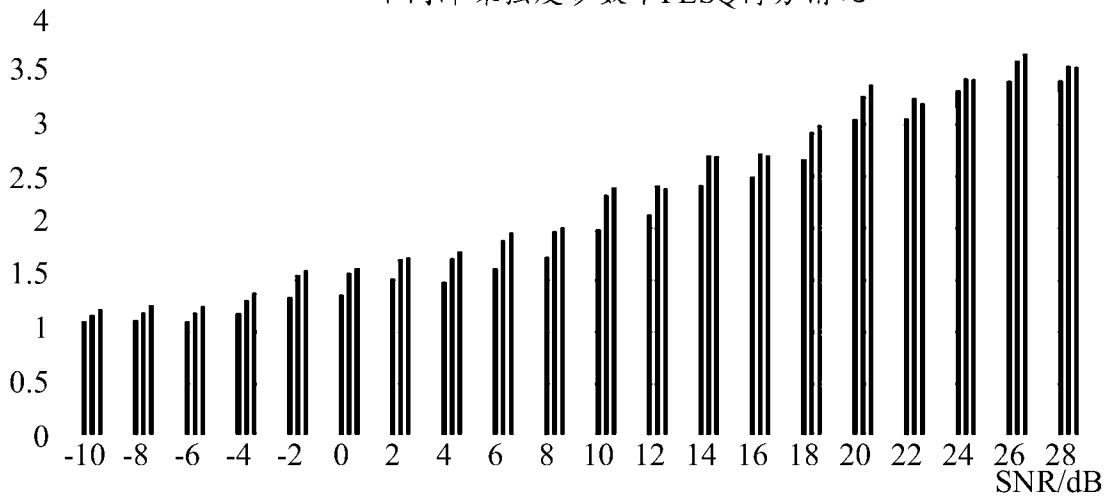


图 6

不同降噪强度参数下SI-SNR得分情况

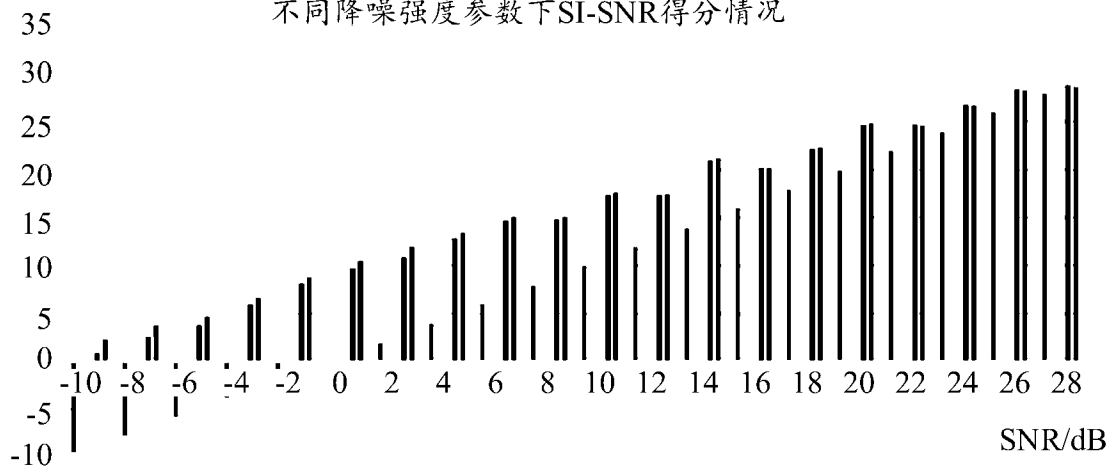


图 7

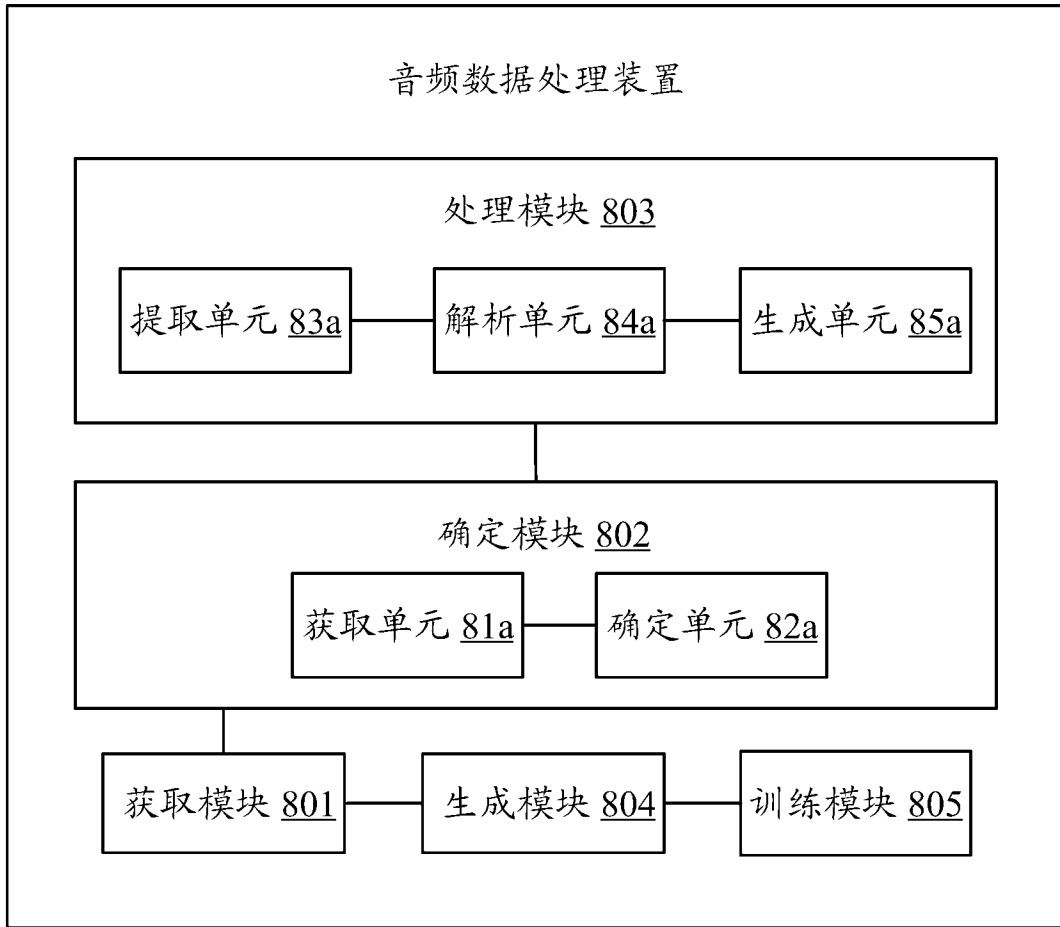


图 8

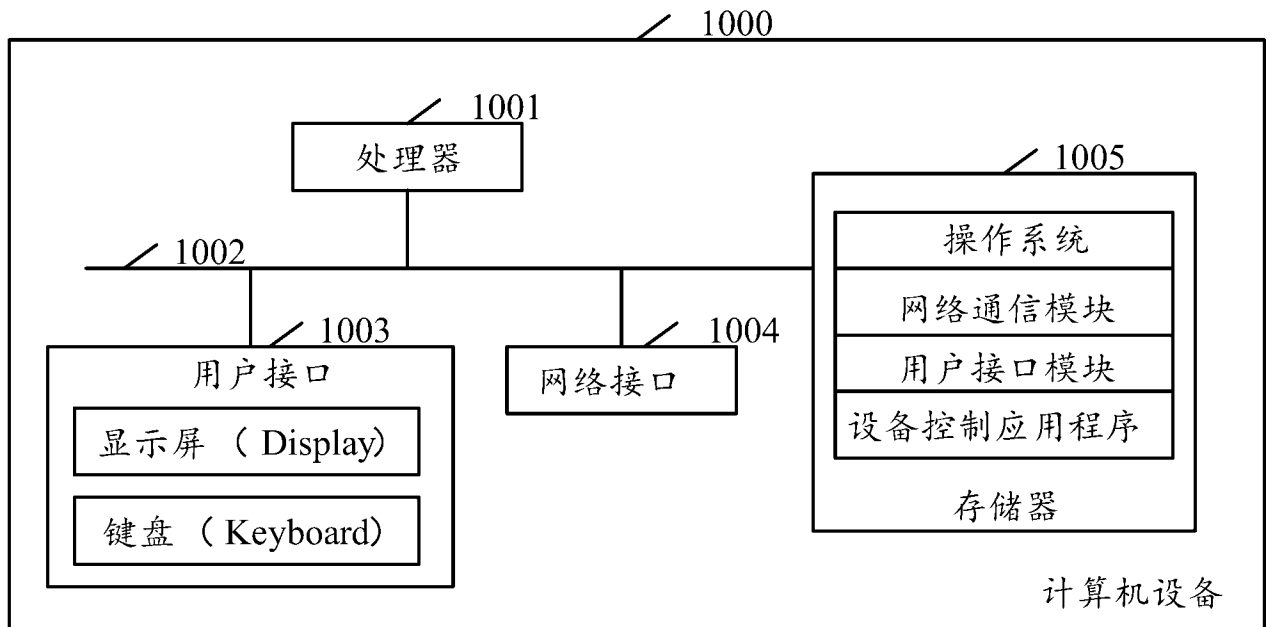


图 9

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2023/129766

A. CLASSIFICATION OF SUBJECT MATTER G10L 21/0208(2013.01)i According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC:G10L Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) DWPI, CNTXT, WPABS, ENTXT, CNKI: 噪声, 噪音, 带噪, 降噪, 去噪, 除噪, 场景, 环境, 周围, 强度, 含量, 质量, 需求, 力度, 历史, 变化, 类型, 候选, 生成网络, 提取网络, 掩码, 增强, 收敛, noise, noising, noise reduction, de-noising, scene, environment, surroundings, intensity, content, quality, demand, strength, history, variation, type, candidate, generative network, extractive network, mask, enhancement, convergence		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 113395539 A (BEIJING BYTEDANCE NETWORK TECHNOLOGY CO., LTD.) 14 September 2021 (2021-09-14) description, paragraphs [0032]-[0080]	1-3, 15-18
A	CN 110197670 A (VOLKSWAGEN-MOBVOI (BEIJING) INFORMATION TECHNOLOGY CO., LTD.) 03 September 2019 (2019-09-03) entire document	1-18
A	CN 111785288 A (BEIJING DIDI INFINITY TECHNOLOGY AND DEVELOPMENT CO., LTD.) 16 October 2020 (2020-10-16) entire document	1-18
A	CN 113362845 A (APOLLO ZHILIAN (BEIJING) TECHNOLOGY CO., LTD.) 07 September 2021 (2021-09-07) entire document	1-18
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "D" document cited by the applicant in the international application "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 12 December 2023		Date of mailing of the international search report 20 December 2023
Name and mailing address of the ISA/CN China National Intellectual Property Administration (ISA/ CN) China No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088		Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2023/129766

C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 113539283 A (TENCENT TECHNOLOGY (SHENZHEN) CO., LTD.) 22 October 2021 (2021-10-22) entire document	1-18
A	DE 102021203815 A1 (ROBERT BOSCH GMBH) 20 October 2022 (2022-10-20) entire document	1-18
A	US 2021074282 A1 (MASSACHUSETTS INSTITUTE OF TECHNOLOGY) 11 March 2021 (2021-03-11) entire document	1-18

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No. PCT/CN2023/129766

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 113395539 A	14 September 2021	None	
CN 110197670 A	03 September 2019	None	
CN 111785288 A	16 October 2020	None	
CN 113362845 A	07 September 2021	None	
CN 113539283 A	22 October 2021	None	
DE 102021203815 A1	20 October 2022	None	
US 2021074282 A1	11 March 2021	US 11227586 B2	18 January 2022

<p>A. 主题的分类</p> <p>G10L 21/0208(2013.01)i</p> <p>按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类</p>																									
<p>B. 检索领域</p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>IPC:G10L</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))</p> <p>DWPI, CNTXT, WPABS, ENTXT, CNKI: 噪声, 噪音, 带噪, 降噪, 去噪, 除噪, 场景, 环境, 周围, 强度, 含量, 质量, 需求, 力度, 历史, 变化, 类型, 候选, 生成网络, 提取网络, 掩码, 增强, 收敛, noise, noising, noise reduction, de-noising, scene, environment, surroundings, intensity, content, quality, demand, strength, history, variation, type, candidate, generative network, extractive network, mask, enhancement, convergence</p>																									
<p>C. 相关文件</p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>CN 113395539 A (北京字节跳动网络技术有限公司) 2021年9月14日 (2021 - 09 - 14) 说明书[0032]-[0080]段</td> <td>1-3, 15-18</td> </tr> <tr> <td>A</td> <td>CN 110197670 A (大众问问(北京)信息科技有限公司) 2019年9月3日 (2019 - 09 - 03) 全文</td> <td>1-18</td> </tr> <tr> <td>A</td> <td>CN 111785288 A (北京嘀嘀无限科技发展有限公司) 2020年10月16日 (2020 - 10 - 16) 全文</td> <td>1-18</td> </tr> <tr> <td>A</td> <td>CN 113362845 A (阿波罗智联(北京)科技有限公司) 2021年9月7日 (2021 - 09 - 07) 全文</td> <td>1-18</td> </tr> <tr> <td>A</td> <td>CN 113539283 A (腾讯科技(深圳)有限公司) 2021年10月22日 (2021 - 10 - 22) 全文</td> <td>1-18</td> </tr> <tr> <td>A</td> <td>DE 102021203815 A1 (ROBERT BOSCH G.M.B.H.) 2022年10月20日 (2022 - 10 - 20) 全文</td> <td>1-18</td> </tr> </tbody> </table> <p><input checked="" type="checkbox"/> 其余文件在C栏的续页中列出。 <input checked="" type="checkbox"/> 见同族专利附件。</p> <table border="0"> <tr> <td> <p>* 引用文件的具体类型:</p> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“D” 申请人在国际申请中引证的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“P” 公布日先于国际申请日但迟于所要求的优先权日的文件</p> </td> <td> <p>“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件</p> <p>“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性</p> <p>“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性</p> <p>“&” 同族专利的文件</p> </td> </tr> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	X	CN 113395539 A (北京字节跳动网络技术有限公司) 2021年9月14日 (2021 - 09 - 14) 说明书[0032]-[0080]段	1-3, 15-18	A	CN 110197670 A (大众问问(北京)信息科技有限公司) 2019年9月3日 (2019 - 09 - 03) 全文	1-18	A	CN 111785288 A (北京嘀嘀无限科技发展有限公司) 2020年10月16日 (2020 - 10 - 16) 全文	1-18	A	CN 113362845 A (阿波罗智联(北京)科技有限公司) 2021年9月7日 (2021 - 09 - 07) 全文	1-18	A	CN 113539283 A (腾讯科技(深圳)有限公司) 2021年10月22日 (2021 - 10 - 22) 全文	1-18	A	DE 102021203815 A1 (ROBERT BOSCH G.M.B.H.) 2022年10月20日 (2022 - 10 - 20) 全文	1-18	<p>* 引用文件的具体类型:</p> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“D” 申请人在国际申请中引证的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“P” 公布日先于国际申请日但迟于所要求的优先权日的文件</p>	<p>“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件</p> <p>“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性</p> <p>“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性</p> <p>“&” 同族专利的文件</p>
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																							
X	CN 113395539 A (北京字节跳动网络技术有限公司) 2021年9月14日 (2021 - 09 - 14) 说明书[0032]-[0080]段	1-3, 15-18																							
A	CN 110197670 A (大众问问(北京)信息科技有限公司) 2019年9月3日 (2019 - 09 - 03) 全文	1-18																							
A	CN 111785288 A (北京嘀嘀无限科技发展有限公司) 2020年10月16日 (2020 - 10 - 16) 全文	1-18																							
A	CN 113362845 A (阿波罗智联(北京)科技有限公司) 2021年9月7日 (2021 - 09 - 07) 全文	1-18																							
A	CN 113539283 A (腾讯科技(深圳)有限公司) 2021年10月22日 (2021 - 10 - 22) 全文	1-18																							
A	DE 102021203815 A1 (ROBERT BOSCH G.M.B.H.) 2022年10月20日 (2022 - 10 - 20) 全文	1-18																							
<p>* 引用文件的具体类型:</p> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“D” 申请人在国际申请中引证的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“P” 公布日先于国际申请日但迟于所要求的优先权日的文件</p>	<p>“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件</p> <p>“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性</p> <p>“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性</p> <p>“&” 同族专利的文件</p>																								
<p>国际检索实际完成的日期</p> <p>2023年12月12日</p>	<p>国际检索报告邮寄日期</p> <p>2023年12月20日</p>																								
<p>ISA/CN的名称和邮寄地址</p> <p>中国国家知识产权局 中国北京市海淀区蓟门桥西土城路6号 100088</p>	<p>授权官员</p> <p>王玥</p> <p>电话号码 (+86) 010-53962545</p>																								

C. 相关文件		
类型*	引用文件，必要时，指明相关段落	相关的权利要求
A	US 2021074282 A1 (MASSACHUSETTS INSTITUTE TECHNOLOGY) 2021年3月11日 (2021 - 03 - 11) 全文	1-18

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2023/129766

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	113395539	A	2021年9月14日	无	
CN	110197670	A	2019年9月3日	无	
CN	111785288	A	2020年10月16日	无	
CN	113362845	A	2021年9月7日	无	
CN	113539283	A	2021年10月22日	无	
DE	102021203815	A1	2022年10月20日	无	
US	2021074282	A1	2021年3月11日	US	11227586 B2 2022年1月18日