

(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
2020 年 8 月 20 日 (20.08.2020)



(10) 国际公布号
WO 2020/164397 A1

(51) 国际专利分类号:
G10L 15/06 (2013.01) *G10L 21/0216* (2013.01)

(21) 国际申请号: PCT/CN2020/074178

(22) 国际申请日: 2020 年 2 月 3 日 (03.02.2020)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201910111593.1 2019年2月12日 (12.02.2019) CN

(71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [—/CN]; 开曼群岛大开曼资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。

(72) 发明人: 张仕良 (ZHANG, Shiliang); 中国浙江省杭州市余杭区文一西路969号3号楼5楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。雷鸣 (LEI, Ming); 中国浙江省杭州市余杭区文一西路969号3号楼5楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。

(74) 代理人: 北京三友知识产权代理有限公司 (BEIJING SANYOU INTELLECTUAL PROPERTY AGENCY LTD.); 中国北京市金融街 35 号国际企业大厦 A 座 16 层, Beijing 100033 (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, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS,

(54) Title: VOICE RECOGNITION METHOD AND SYSTEM

(54) 发明名称: 一种语音识别方法及系统

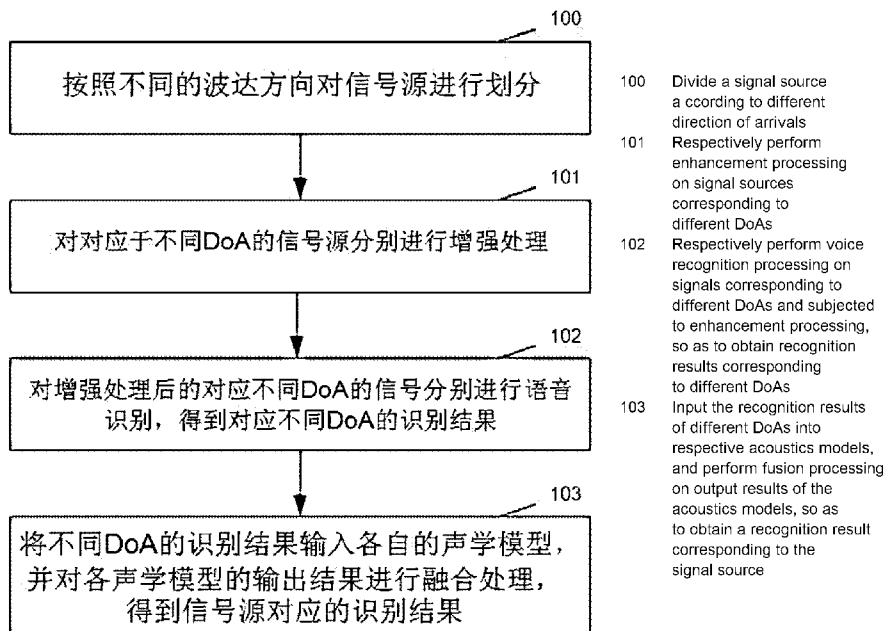


图 1

(57) Abstract: Disclosed are a voice recognition method and system. In the embodiments of the present application, a space is divided into several regions by means of preset DoA angles, so as to divide a signal source into different spatial regions; and signals in different spatial regions are then subjected to enhancement and recognition processing and then are fused to obtain a recognition result of the signal source. The present application does not need to estimate the actual direction of a signal source at each moment, avoiding the problem of recognition inaccuracy caused by estimation of the signal-to-noise ratio and the signal source direction of a signal in a

JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, 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, GH, GM, KE, LR, LS, MW, MZ, NA, RW, 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, 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))。

complex environment, thereby ensuring the accuracy of a voice recognition result.

(57) 摘要: 本申请公开了一种语音识别方法及系统, 本申请实施例通过预先设置的DoA角度, 将空间分成若干个区域, 从而将信号源划分至不同的空间区域; 进而对不同空间区域的信号进行增强和识别处理后融合得到信号源的识别结果。本申请不需要估计每个时刻真实的信号源方向, 避免了在复杂环境下, 由于估计信号的信噪比和信号源方向而带来的识别不准确的问题, 从而保障了语音识别结果的准确性。

一种语音识别方法及系统

本申请要求 2019 年 02 月 12 日递交的申请号为 201910111593.1、发明名称为“一种语音识别方法及系统”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

5 技术领域

本申请涉及但不限于信号处理技术，尤指一种语音识别方法及系统。

背景技术

相关技术中，远场语音识别系统主要包括两个组成部分：一是前端信号增强部分，
10 用于对接收到的多通道带噪语音信号进行处理，从而得到增强后的单通道语音信号。前端信号增强部分通过利用多个通道的语音信号之间的关联性，去除一定的噪声干扰，提升信号的信噪比；另一个是后端常用的语音识别（ASR）部分，用于对前端信号增强部分处理后的单通道语音信号输入一个通用的语音识别系统，以得到最终的语音识别结果。

在复杂环境下，很难估计出信号的信噪比和信号源方向，也就是说，相关技术中的
15 远场语音识别技术很难保障语音识别结果的准确性。

发明内容

本申请提供一种语音识别方法及系统，能够保障语音识别结果的准确性。

本发明实施例提供了一种语音识别方法，包括：

20 按照不同的波达方向 DoA 对信号源进行划分；

对对应于不同 DoA 的信号源分别进行增强处理；

对增强处理后的对应不同 DoA 的信号分别进行语音识别，得到对应不同 DoA 的识别结果；

将不同 DoA 的识别结果输入各自的声学模型，并对各声学模型的输出结果进行融合
25 处理，得到识别结果。

在一种示例性实例中，所述按照不同的波达方向 DoA 对信号源进行划分，包括：

将空间划分为多个区域，从而将信号源划分至不同 DoA 角度形成的区域。

在一种示例性实例中，所述 DoA 角度包括以下至少之一：30 度、60 度、90 度、120
度、150 度。

30 在一种示例性实例中，所述对对应于不同 DoA 的信号源分别进行增强处理，包括：

对所述对应于不同 DoA 的信号源都分别进行基于延时叠加 DAS 的波束形成方法，得到所述增强处理后的信号。

在一种示例性实例中，所述对对应于不同 DoA 的信号源分别进行增强处理，包括：

对所述对应于不同 DoA 的信号源都分别进行 MVDR 的波束形成方法，得到所述增强处理后的信号。

在一种示例性实例中，所述方法之前还包括：根据不同的所述 DoA 对空间进行区域划分；对不同区域内的语音信号进行语音增强处理，得到不同的增强信号样本；利用得到的各样本训练得到对应不同 DoA 的所述声学模型。

在一种示例性实例中，所述将不同 DoA 的识别结果输入各自的声学模型，并对各声学模型的输出结果进行融合处理，得到识别结果，包括：

将对应不同 DoA 的所述识别结果输入各自所述声学模型；对各声学模型的输出结果进行融合，得到所述识别结果。

在一种示例性实例中，所述融合通过基于 ROVER 的融合系统实现。

本申请还提供了一种计算机可读存储介质，存储有计算机可执行指令，所述计算机可执行指令用于执行上述任一项所述的语音识别方法。

本申请又提供了一种用于实现信息分享的装置，包括存储器和处理器，其中，存储器中存储有以下可被处理器执行的指令：用于执行上述任一项所述的语音识别方法的步骤。

本申请还提供了一种音箱，包括存储器和处理器，其中，存储器中存储有以下可被处理器执行的指令：用于执行上述任一项所述的语音识别方法的步骤。

本申请再提供了一种语音识别系统，包括：预处理模块、第一处理模块、第二处理模块、识别模块；其中，

预处理模块，用于按照不同的 DoA 对信号源进行划分；

第一处理模块，用于对对应于不同 DoA 的信号源分别进行增强处理；

第二处理模块，用于对增强处理后的对应不同 DoA 的信号分别进行语音识别，得到对应不同 DoA 的识别结果；

识别模块，用于将不同 DoA 的识别结果输入各自的声学模型，并对各声学模型的输出结果进行融合处理，得到识别结果。

在一种示例性实例中，所述装置还包括：训练模块，用于根据不同的所述 DoA 对空间进行区域划分；对不同区域内的语音信号进行语音增强处理，得到不同的增强信号样

本；利用得到的各样本训练得到对应不同 DoA 的所述声学模型。

本申请包括：按照不同的波达方向 DoA 对信号源进行划分；对对应于不同 DoA 的信号源分别进行增强处理；对增强处理后的对应不同 DoA 的信号分别进行语音识别，得到对应不同 DoA 的识别结果；将不同 DoA 的识别结果输入各自的声学模型，并对各声学模型的输出结果进行融合处理，得到识别结果。本申请实施例通过预先设置的 DoA 角度，将空间分成若干个区域，从而将信号源划分至不同的空间区域；进而对不同空间区域的信号进行增强和识别处理后融合得到信号源的识别结果。本申请不需要估计每个时刻真实的信号源方向，避免了在复杂环境下，由于估计信号的信噪比和信号源方向而带来的识别不准确的问题，从而保障了语音识别结果的准确性。

本发明的其它特征和优点将在随后的说明书中阐述，并且，部分地从说明书中变得显而易见，或者通过实施本发明而了解。本发明的目的和其他优点可通过在说明书、权利要求书以及附图中所特别指出的结构来实现和获得。

附图说明

附图用来提供对本申请技术方案的进一步理解，并且构成说明书的一部分，与本申请的实施例一起用于解释本申请的技术方案，并不构成对本申请技术方案的限制。

图 1 为本申请语音识别方法的流程示意图；

图 2 为本申请一种基于 Delay-and-Sum 波束形成方法的示例图；

图 3 为本申请语音识别系统的组成结构示意图。

具体实施方式

为使本申请的目的、技术方案和优点更加清楚明白，下文中将结合附图对本申请的实施例进行详细说明。需要说明的是，在不冲突的情况下，本申请中的实施例及实施例中的特征可以相互任意组合。

在本申请一个典型的配置中，计算设备包括一个或多个处理器 (CPU)、输入/输出接口、网络接口和内存。

内存可能包括计算机可读介质中的非永久性存储器，随机存取存储器 (RAM) 和/或非易失性内存等形式，如只读存储器 (ROM) 或闪存(flash RAM)。内存是计算机可读介质的示例。

计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或

技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括，但不限于相变内存 (PRAM)、静态随机存取存储器 (SRAM)、动态随机存取存储器 (DRAM)、其他类型的随机存取存储器 (RAM)、只读存储器 (ROM)、电可擦除可编程只读存储器 (EEPROM)、快闪记忆体或其他内存技术、
5 只读光盘只读存储器 (CD-ROM)、数字多功能光盘 (DVD) 或其他光学存储、磁盒式磁带，磁带磁盘存储或其他磁性存储设备或任何其他非传输介质，可用于存储可以被计算设备访问的信息。按照本文中的界定，计算机可读介质不包括非暂存电脑可读媒体 (transitory media)，如调制的数据信号和载波。

在附图的流程图示出的步骤可以在诸如一组计算机可执行指令的计算机系统中执行。并且，虽然在流程图中示出了逻辑顺序，但是在某些情况下，可以以不同于此处的顺序执行所示出或描述的步骤。
10

图 1 为本申请语音识别方法的流程示意图，如图 1 所示，包括：

步骤 100：按照不同的波达方向 (DoA, Direction Of Arrival) 对信号源进行划分。

声音到达麦克风阵列不同麦克风的延迟，通过这个延迟可以计算出目标声源即步骤
15 100 中的信号源可能在空间的某个角度内即某个 DOA 角度内。本申请发明人发现，当不能准确估计 DoA 时，可以将空间划分成不同的方向，然后假设目标声源在这个方向。

在一种示例性实例中，按照预先设置的至少一个 DoA 角度，比如 30 度、60 度、90 度、120 度、150 度等，将空间划分为多个区域，从而假设信号源出现在这些 DoA 角度区域内，也就是说，将信号源划分至不同 DoA 角度形成的区域。这里需要说明的是，信号源是移动的，所以不同时刻可能处于不同的 DoA 角度区域内，但是肯定会处于某个 DoA 角度区域内。
20

在复杂环境下，很难估计信号的信噪比和信号源方向，因此，本申请实施例中，并不需要估计每个时刻真实的信号源方向，而是通过预先设置的 DoA 角度，将空间分成若干个区域，从而假设信号源出现在这些 DoA 角度区域内。通过假设信号源总是会处于其中某一个 DoA 角度范围内，使得后续可以针对每个区域的信号源分别进行信号增强处理。
25

步骤 101：对对应于不同 DoA 的信号源分别进行增强处理。

在一种示例性实例中，增强处理可以包括：

对对应于不同 DoA 的信号源都分别进行基于延时叠加 (DAS, Delay-and-Sum) 的波束形成方法，得到增强处理后的信号。图 2 为本申请一种基于 Delay-and-Sum 波束形
30

成方法的示例，具体实现可以参见相关技术，这里仅仅是举例说明，并不用于限定本申请的保护范围。

在一种示例性实例中，增强处理可以包括：

对对应于不同 DoA 的信号源都分别进行 MVDR 的波束形成方法，得到增强处理后
5 的信号。其中，MVDR（Minimum Variance Distortionless Response）是 Capon 于 1967 年提出的一种自适应的空间波数谱估计算法。

步骤 102：对增强处理后的对应不同 DoA 的信号分别进行语音识别，得到对应不同 DoA 的识别结果。

在一种示例性实例中，语音识别可以包括如 ASR 系统。

10 本申请中，由于对对应不同 DoA 的信号都进行了波束形成，因此，经过语音识别如 ASR 系统后会得到若干个对应不同 DoA 的识别结果。

步骤 103：将不同 DoA 的识别结果输入各自的声学模型，并对各声学模型的输出结果进行融合处理，得到信号源对应的识别结果。

本步骤之前还包括：根据不同的 DoA 对空间进行区域划分，然后对不同区域的语音
15 信号进行语音增强处理，得到不同的增强信号样本，利用得到的各样本训练得到对应不同 DoA 的声学模型。训练的方法很多，可以采用相关技术来实现，具体实现并不用于限定本申请的保护范围。

在一种示例性实例中，本步骤可以包括：将对应不同 DoA 的识别结果输入各自训练好的声学模型，然后再将各声学模型的输出结果采用如基于 ROVER 的融合系统进行融
20 合，得到最终信号源对应的识别结果。

在一种示例性实例中，融合可以通过一种基于识别结果投票的错误降低（ROVER，Recognizer Output Voting Error Reduction）方法的融合系统来实现。

本申请实施例通过预先设置的 DoA 角度，将空间分成若干个区域，从而将信号源划分至不同的空间区域；进而对不同空间区域的信号进行增强和识别处理后融合得到信号
25 源的识别结果。本申请不需要估计每个时刻真实的信号源方向，避免了在复杂环境下，由于估计信号的信噪比和信号源方向而带来的识别不准确的问题，从而保障了语音识别结果的准确性。

本申请还提供一种计算机可读存储介质，存储有计算机可执行指令，所述计算机可执行指令用于执行上述任一项的语音识别方法。

30 本申请再提供一种实现信息分享的装置，包括存储器和处理器，其中，存储器中存

储有上述任一项的语音识别方法的步骤。

本申请还提供一种音箱，包括存储器和处理器，其中，存储器中存储有以下可被处理器执行的指令：用于执行上述任一项所述的语音识别方法的步骤。

图 3 为本申请语音识别系统的组成结构示意图，如图 3 所示，至少包括：预处理模

5 块、第一处理模块、第二处理模块、识别模块；其中，

预处理模块，用于按照不同的 DoA 对信号源进行划分；

第一处理模块，用于对对应于不同 DoA 的信号源分别进行增强处理；

第二处理模块，用于对增强处理后的对应不同 DoA 的信号分别进行语音识别，得到识别结果；

10 识别模块，用于将不同 DoA 的识别结果输入各自的声学模型，并对各声学模型的输出结果进行融合处理，得到识别结果。

在一种示例性实例中，预处理模块具体用于：

按照预先设置的至少一个 DoA 角度，比如 30 度、60 度、90 度、120 度、150 度等，将空间划分为多个区域，从而假设信号源出现在这些 DoA 角度区域内，也就是说，将信

15 号源划分至不同 DoA 角度形成的区域。

在一种示例性实例中，第一处理模块具体用于：

对对应于不同 DoA 的信号源都分别进行基于 DAS 的波束形成方法，得到增强后的信号；

或者，对对应于不同 DoA 的信号源都分别进行 MVDR 的波束形成方法，得到增强

20 后的信号。其中，MVDR 是 Capon 于 1967 年提出的一种自适应的空间波数谱估计算法。

在一种示例性实例中，第二处理模块可以是 ASR 系统。

在一种示例性实例中，识别模块具体用于：

将对应不同 DoA 的识别结果输入训练好的各自的声学模型，将各声学模型的识别结果采用如基于 ROVER 的融合系统进行融合，得到信号源对应的识别结果。

25 本申请语音识别装置还包括：训练模块，用于根据不同的所述 DoA 对空间进行区域划分；对不同区域内的语音信号进行语音增强处理，得到不同的增强信号样本；利用得到的各样本训练得到对应不同 DoA 的所述声学模型。

需要说明的是，本申请语音识别系统中的各模块可以单独设置在不同的实体设备中，也可以合理组合后设置在多个实体设备中，还可以是都设置在同一实体设备中。

30 虽然本申请所揭露的实施方式如上，但所述的内容仅为便于理解本申请而采用的实

施方式，并非用以限定本申请。任何本申请所属领域内的技术人员，在不脱离本申请所揭露的精神和范围的前提下，可以在实施的形式及细节上进行任何的修改与变化，但本申请的专利保护范围，仍须以所附的权利要求书所界定的范围为准。

权 利 要 求 书

1. 一种语音识别方法, 包括:

按照不同的波达方向 DoA 对信号源进行划分;

对对应于不同 DoA 的信号源分别进行增强处理;

5 对增强处理后的对应不同 DoA 的信号分别进行语音识别, 得到对应不同 DoA 的识别结果;

将不同 DoA 的识别结果输入各自的声学模型, 并对各声学模型的输出结果进行融合处理, 得到识别结果。

2. 根据权利要求 1 所述的语音识别方法, 其中, 所述按照不同的波达方向 DoA 对信号源进行划分, 包括:

将空间划分为多个区域, 从而将信号源划分至不同 DoA 角度形成的区域。

3. 根据权利要求 2 所述的语音识别方法, 其中, 所述 DoA 角度包括以下至少之一: 30 度、60 度、90 度、120 度、150 度。

4. 根据权利要求 1 所述的语音识别方法, 其中, 所述对对应于不同 DoA 的信号源分别进行增强处理, 包括:

对所述对应于不同 DoA 的信号源都分别进行基于延时叠加 DAS 的波束形成方法, 得到所述增强处理后的信号。

5. 根据权利要求 1 所述的语音识别方法, 其中, 所述对对应于不同 DoA 的信号源分别进行增强处理, 包括:

20 对所述对应于不同 DoA 的信号源都分别进行 MVDR 的波束形成方法, 得到所述增强处理后的信号。

6. 根据权利要求 1 所述的语音识别方法, 所述方法之前还包括: 根据不同的所述 DoA 对空间进行区域划分; 对不同区域内的语音信号进行语音增强处理, 得到不同的增强信号样本; 利用得到的各样本训练得到对应不同 DoA 的所述声学模型。

25 7. 根据权利要求 6 所述的语音识别方法, 其中, 所述将不同 DoA 的识别结果输入各自的声学模型, 并对各声学模型的输出结果进行融合处理, 得到识别结果, 包括:

将对应不同 DoA 的所述识别结果输入各自所述声学模型; 对各声学模型的输出结果进行融合, 得到所述识别结果。

8. 根据权利要求 6 所述的语音识别方法, 其中, 所述融合通过基于 ROVER 的融合系统实现。

9. 一种计算机可读存储介质，存储有计算机可执行指令，所述计算机可执行指令用于执行权利要求 1~权利要求 8 任一项所述的语音识别方法。

10. 一种用于实现信息分享的装置，包括存储器和处理器，其中，存储器中存储有以下可被处理器执行的指令：用于执行权利要求 1~权利要求 8 任一项所述的语音识别方法的步骤。

11. 一种音箱，包括存储器和处理器，其中，存储器中存储有以下可被处理器执行的指令：用于执行权利要求 1~权利要求 8 任一项所述的语音识别方法的步骤。

12. 一种语音识别系统，包括：预处理模块、第一处理模块、第二处理模块、识别模块；其中，

10 预处理模块，用于按照不同的 DoA 对信号源进行划分；

第一处理模块，用于对对应于不同 DoA 的信号源分别进行增强处理；

第二处理模块，用于对增强处理后的对应不同 DoA 的信号分别进行语音识别，得到对应不同 DoA 的识别结果；

15 识别模块，用于将不同 DoA 的识别结果输入各自的声学模型，并对各声学模型的输出结果进行融合处理，得到识别结果。

13. 根据权利要求 12 所述的语音识别系统，所述系统还包括：训练模块，用于根据不同的所述 DoA 对空间进行区域划分；对不同区域内的语音信号进行语音增强处理，得到不同的增强信号样本；利用得到的各样本训练得到对应不同 DoA 的所述声学模型。

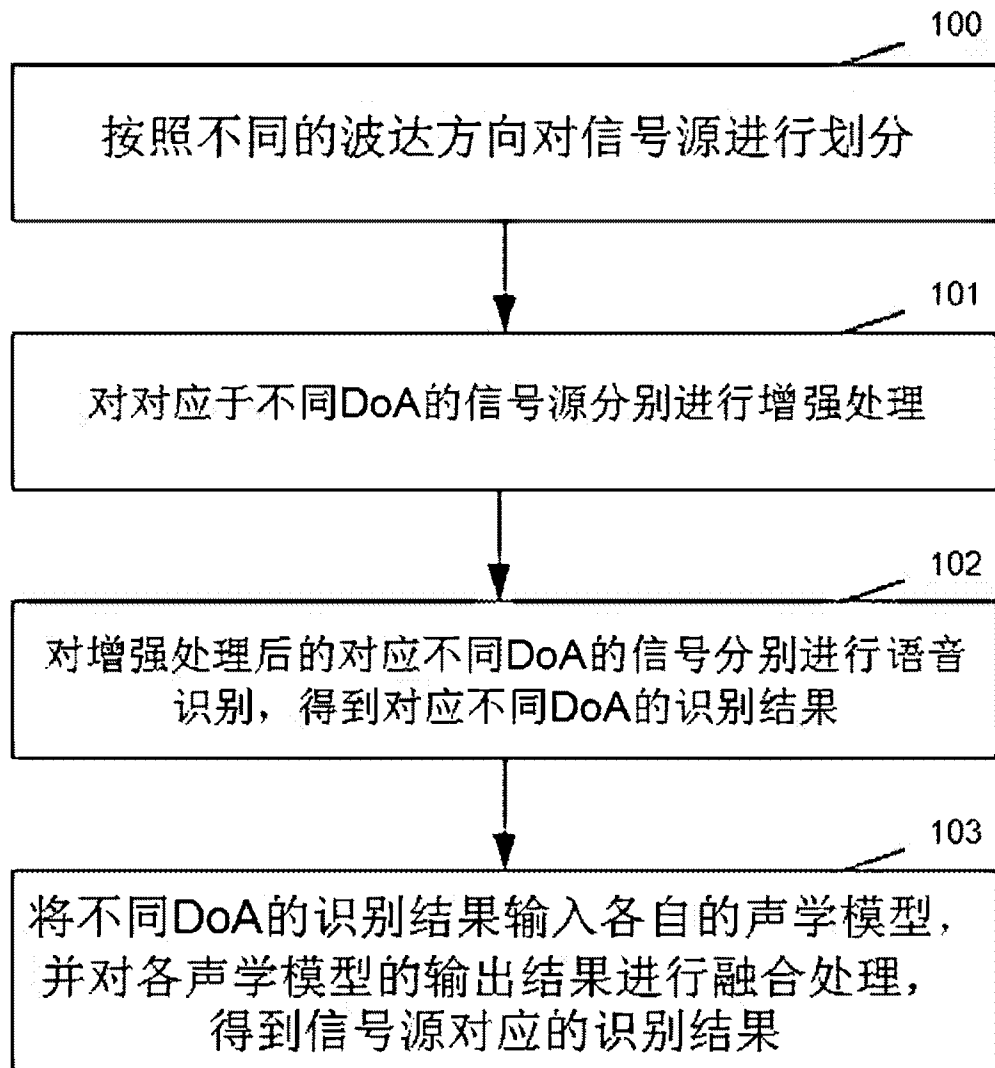
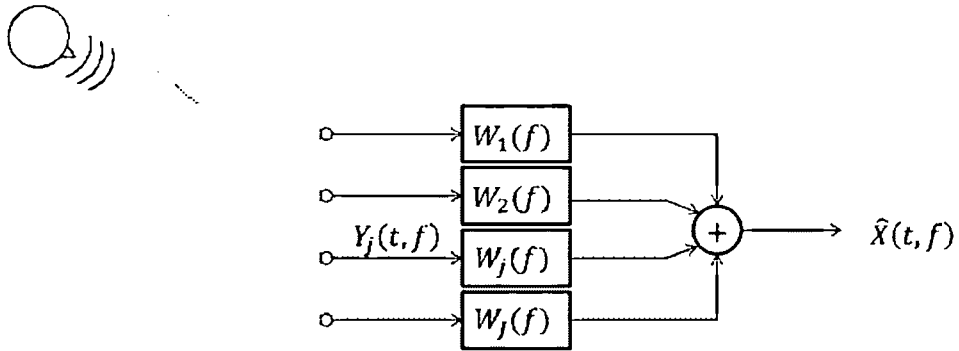


图 1



- Output of beamformer

$$\hat{X}(t, f) = \sum_j w_j(f) Y_j(t, f)$$

- Matrix notations

$$\hat{X}(t, f) = \mathbf{w}_f^H \mathbf{y}_{t,f}$$

$$\mathbf{w}_f = [W_1(f), \dots, W_j(f)]^T \quad \mathbf{y}_{t,f} = [Y_1(t, f), \dots, Y_j(t, f)]^T$$

图 2

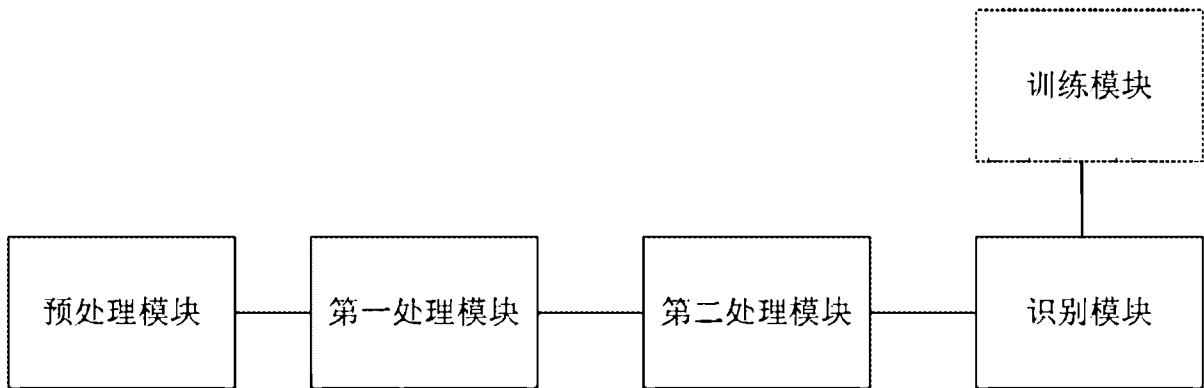


图 3

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/074178

A. CLASSIFICATION OF SUBJECT MATTER

G10L 15/06(2013.01)i; G10L 21/0216(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)

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)

CNABS; CNTXT; CNKI; 万方; 超星; VEN; USTXT; EPTXT; WOTXT; IEEE: 区域, 方向, 空间, 划分, 方位, 语音, 语义, 远程, 远场, 信号源, 声源, 麦克风, 话筒, 增强, 加强, 角度, 波束, 模型, 加权, 融合, 权重, DOA, direction of arrival, divid+, far field, model, direction, area, voice, sound, angle

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 102271299 A (SONY CORPORATION) 07 December 2011 (2011-12-07) description, paragraphs [0030]-[0049]	1-5, 9-12
Y	CN 109272989 A (BEIJING JINGDONG SHANGKE INFORMATION TECHNOLOGY CO., LTD. et al.) 25 January 2019 (2019-01-25) description, paragraphs [0005]-[0038]	1-5, 9-12
A	CN 107742522 A (IFLYTEK CO., LTD.) 27 February 2018 (2018-02-27) entire document	1-13



Further documents are listed in the continuation of Box C.



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

“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

31 March 2020

Date of mailing of the international search report

21 April 2020

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/CN)

No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088

China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2020/074178

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	102271299	A	07 December 2011	KR	20110132245	A	07 December 2011
				JP	5593852	B2	24 September 2014
				US	8976978	B2	10 March 2015
				EP	2392938	B1	13 March 2013
				JP	2011254189	A	15 December 2011
				CN	102271299	B	15 July 2015
				US	2011293107	A1	01 December 2011
				EP	2392938	A1	07 December 2011
CN	109272989	A	25 January 2019	None			
CN	107742522	A	27 February 2018	WO	2019080553	A1	02 May 2019

国际检索报告

国际申请号

PCT/CN2020/074178

A. 主题的分类

G10L 15/06(2013.01)i; G10L 21/0216(2013.01)i

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

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G10L

包含在检索领域中的除最低限度文献以外的检索文献

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))

CNABS;CNTXT;CNKI;万方;超星;VEN;USTXT;EPTXT;WOTXT;IEEE:区域, 方向, 空间, 划分, 方位, 语音, 语义, 远程, 远场, 信号源, 声源, 麦克风, 话筒, 增强, 加强, 角度, 波束, 模型, 加权, 融合, 权重, DOA, direction of arrival, divid+, far field, model, direction, area, voice, sound, angle

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
Y	CN 102271299 A (索尼公司) 2011年 12月 7日 (2011 - 12 - 07) 说明书第[0030]-[0049]段	1-5、9-12
Y	CN 109272989 A (北京京东尚科信息技术有限公司 等) 2019年 1月 25日 (2019 - 01 - 25) 说明书第[0005]-[0038]段	1-5、9-12
A	CN 107742522 A (科大讯飞股份有限公司) 2018年 2月 27日 (2018 - 02 - 27) 全文	1-13

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“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)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“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

“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

“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

“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

“&” 同族专利的文件

国际检索实际完成的日期

2020年 3月 31日

国际检索报告邮寄日期

2020年 4月 21日

ISA/CN的名称和邮寄地址

中国国家知识产权局(ISA/CN)

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

授权官员

周群

电话号码 86-(0512)-88997268

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/074178

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	102271299	A	2011年 12月 7日	KR	20110132245	A	2011年 12月 7日
				JP	5593852	B2	2014年 9月 24日
				US	8976978	B2	2015年 3月 10日
				EP	2392938	B1	2013年 3月 13日
				JP	2011254189	A	2011年 12月 15日
				CN	102271299	B	2015年 7月 15日
				US	2011293107	A1	2011年 12月 1日
				EP	2392938	A1	2011年 12月 7日
CN	109272989	A	2019年 1月 25日	无			
CN	107742522	A	2018年 2月 27日	WO	2019080553	A1	2019年 5月 2日