

(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 9 月 24 日 (24.09.2020)



(10) 国际公布号
WO 2020/186712 A1

(51) 国际专利分类号:
G06F 3/16 (2006.01)

(21) 国际申请号: PCT/CN2019/106806

(22) 国际申请日: 2019 年 9 月 19 日 (19.09.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201910211472.4 2019年3月20日 (20.03.2019) CN

(71) 申请人: 海信视像科技股份有限公司
(HISENSE VISUAL TECHNOLOGY CO., LTD.) [CN/
CN]; 中国山东省青岛市经济技术开发区前湾
港路218号, Shandong 266555 (CN)。

(72) 发明人: 任晓楠 (REN, Xiaonan); 中国山东省青
岛市经济技术开发区前湾港路218号, Shandong
266555 (CN)。 崔保磊 (CUI, Baolei); 中国山东
省青岛市经济技术开发区前湾港路218号,
Shandong 266555 (CN)。 戴磊 (DAI, Lei); 中国
山东省青岛市经济技术开发区前湾港路
218号, Shandong 266555 (CN)。

(74) 代理人: 北京同达信恒知识产权代理有限公司
(TDP & PARTNERS); 中国北京市西城区裕民路
18号北环中心A座2002, Beijing 100029 (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 APPARATUS, AND TERMINAL

(54) 发明名称: 一种语音识别方法、装置及终端

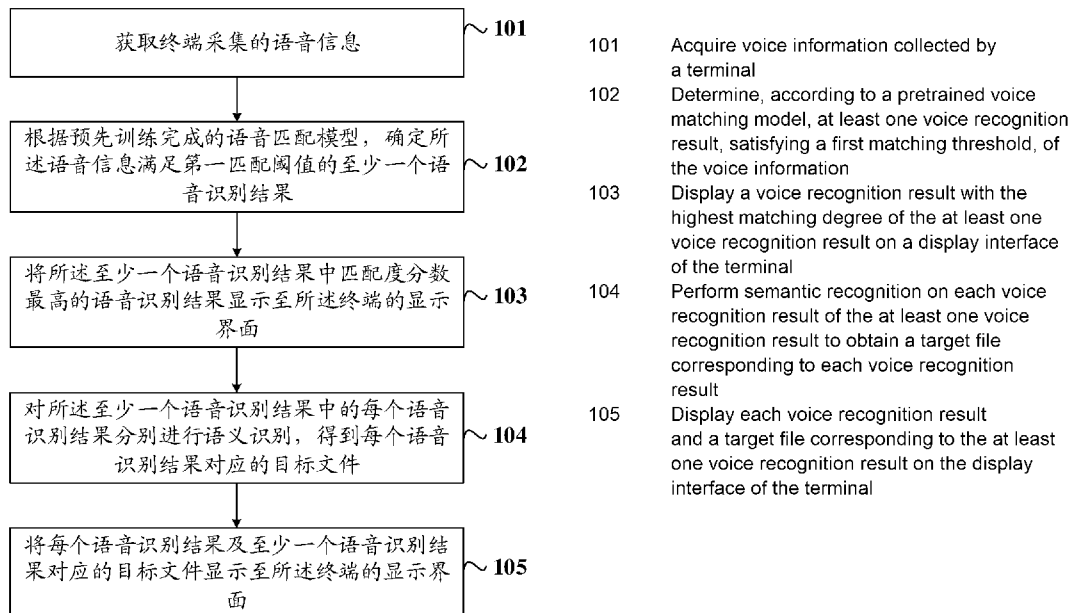


图 1

(57) Abstract: Disclosed are a voice recognition method and apparatus, and a terminal. The method comprises: receiving input voice information; determining, according to a pretrained voice matching model, at least one voice recognition result, satisfying a first matching threshold, of the voice information; determining a voice recognition result with the highest matching degree of the at least one voice recognition result as a target voice recognition result; acquiring a target file corresponding to the target voice recognition result; and displaying each voice recognition result and the target file corresponding to the target voice recognition result on a display interface,

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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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))。

wherein the target voice recognition result is displayed in a first display mode, and the other voice recognition results are displayed in a second display mode.

(57) 摘要: 本公开公开了一种语音识别方法、装置及终端, 所述方法包括: 接收输入的语音信息; 根据预先训练的语音匹配模型, 确定满足第一匹配阈值的所述语音信息的至少一个语音识别结果; 确定所述至少一个语音识别结果中匹配度最高的语音识别结果为目标语音识别结果; 获取所述目标语音识别结果对应的目标文件; 将每个语音识别结果及所述目标语音识别结果对应的目标文件显示至显示界面, 其中所述目标语音识别结果以第一显示方式显示, 其他语音识别结果以第二显示方式显示。

一种语音识别方法、装置及终端

相关申请的交叉引用

本申请要求在2019年03月20日提交中国专利局、申请号为201910211472.4、申请名称为“一种语音识别方法、装置及终端”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。

技术领域

本公开主要涉及智能家居技术领域，尤其涉及一种语音识别方法、装置及终端。

背景技术

目前语音识别产品越来越多，随着技术的进步和普及率的提高，用户对这种交互方式亦逐渐接受和认可。随着语音交互技术及人工智能的不断提高，应用场景从语音助手、智能音箱等方面加速扩围。语音识别产品在使用过程中，通过采集周围环境的声音来进行语义解析并执行用户语音指令操作。

目前，在语音识别技术中，在自适应方面仍需要更大的改进，现实中的语音类型是各种各样的，从声音特征来说可以分为男音、女音和童音，另外，很多人的发音同标准发音有很大的差距，并且还存在一定的同音的情况，导致用户输入语音时得到的识别结果与用户意图不一致，影响了语音识别产品的使用效果。

发明内容

本公开实施例提供了一种语音识别方法，应用于终端，所述方法包括：
接收输入的语音信息；

根据预先训练的语音匹配模型，确定满足第一匹配阈值的所述语音信息的至少一个语音识别结果；

确定所述至少一个语音识别结果中匹配度最高的语音识别结果为目标语音识别结果；

获取所述目标语音识别结果对应的目标文件；

将每个语音识别结果及所述目标语音识别结果对应的目标文件显示至显示界面，其中所述目标语音识别结果以第一显示方式显示，其他语音识别结果以第二显示方式显示。

在一些实施方式中，所述获取所述目标语音识别结果对应的目标文件，包括：

对所述目标语音识别结果进行语义识别，确定所述目标语音识别结果对应的业务类型；

从资源库中在所述目标语音识别结果对应的业务类型中查找所述目标语音识别结果对应的目标文件。

在一些实施方式中，所述对所述目标语音识别结果进行语义识别，确定所述目标语音识别结果对应的业务类型，包括：

根据预设的词典库，对所述目标语音识别结果进行分词处理，并对所述目标语音识别结果中各分词进行语义识别，确定各分词对应的业务类型；

根据各分词对应的业务类型的权重，确定所述目标语音识别结果对应的业务类型。

在一些实施方式中，所述将每个语音识别结果及所述目标语音识别结果对应的目标文件显示至显示界面，包括：

确定每个语音识别结果的优先级；

在所述终端的显示界面上按照所述优先级排列显示所述每个语音识别结果；

将所述目标语音识别结果对应的目标文件在所述终端的显示界面上显示。

在一些实施方式中，所述将每个语音识别结果及所述目标语音识别结果对应的目标文件显示至显示界面之后，所述方法还包括：

获取用户对所述目标语音识别结果的切换指令；

根据所述切换指令，确定更改后的目标语音识别结果对应的目标文件；

将更改后的所述目标语音识别结果以第一显示方式显示，其他语音识别结果以第二显示方式显示；同时显示更改后的所述目标语音识别结果对应的目标文件。

在一些实施方式中，所述根据预先训练完成的语音匹配模型，确定所述语音信息满足第一匹配阈值的语音识别结果，包括：

将所述语音信息输入至所述语音匹配模型，识别所述语音信息中的拼音序列，组成所有可能的候选字；

针对每个可能的候选字，通过语法规则和统计方法，确定可能的汉字序列及汉字序列的得分；

将得分满足第一匹配阈值的汉字序列作为所述语音识别结果。

本公开实施例提供了一种终端，包括处理器、通信接口、存储器和通信总线；其中，处理器，通信接口，存储器通过通信总线完成相互间的通信；

所述存储器中存储有计算机程序；

所述处理器，配置为运行所述计算机程序以使所述终端实现：

接收输入的语音信息；

根据预先训练的语音匹配模型，确定满足第一匹配阈值的所述语音信息的至少一个语音识别结果；

确定所述至少一个语音识别结果中匹配度最高的语音识别结果为目标语音识别结果；

获取所述目标语音识别结果对应的目标文件；

将每个语音识别结果及所述目标语音识别结果对应的目标文件显示至显示界面，其中所述目标语音识别结果以第一显示方式显示，其他语音识别结果以第二显示方式显示。

在一些实施方式中，所述处理器，配置为通过下述执行所述获取所述目标语音识别结果对应的目标文件：

对所述目标语音识别结果进行语义识别，确定所述目标语音识别结果对

应的业务类型；

从资源库中在所述目标语音识别结果对应的业务类型中查找所述目标语音识别结果对应的目标文件。

在一些实施方式中，所述处理器，配置为通过下述执行所述对所述目标语音识别结果进行语义识别，确定所述目标语音识别结果对应的业务类型：

根据预设的词典库，对所述目标语音识别结果进行分词处理，并对所述目标语音识别结果中各分词进行语义识别，确定各分词对应的业务类型；

根据各分词对应的业务类型的权重，确定所述目标语音识别结果对应的业务类型。

在一些实施方式中，所述处理器，配置为通过下述执行所述将每个语音识别结果及所述目标语音识别结果对应的目标文件显示至显示界面：

确定每个语音识别结果的优先级；

在所述终端的显示界面上按照所述优先级排列显示所述每个语音识别结果；

将所述目标语音识别结果对应的目标文件在所述终端的显示界面上显示。

在一些实施方式中，所述处理器，还配置为运行所述计算机程序以使所述终端实现：

获取用户对所述目标语音识别结果的切换指令；

根据所述切换指令，确定更改后的目标语音识别结果对应的目标文件；

将更改后的所述目标语音识别结果以第一显示方式显示，其他语音识别结果以第二显示方式显示；同时显示更改后的所述目标语音识别结果对应的目标文件。

在一些实施方式中，所述处理器，配置为通过下述执行所述根据预先训练完成的语音匹配模型，确定所述语音信息满足第一匹配阈值的语音识别结果：

将所述语音信息输入至所述语音匹配模型，识别所述语音信息中的拼音序列，组成所有可能的候选字；

针对每个可能的候选字，通过语法规则和统计方法，确定可能的汉字序列及汉字序列的得分；

将得分满足第一匹配阈值的汉字序列作为所述语音识别结果。

本公开实施例提供了一种计算机可读非易失性存储介质，所述计算机可读非易失性存储介质上存储计算机程序，所述计算机程序被处理器执行时实现上述任一项应用于终端的方法的步骤。

附图说明

为了更清楚地说明本公开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公开的一些实施例，对于本领域的普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

图 1 为本公开实施例 1 提供的一种语音识别方法的过程示意图；

图 2 为本公开实施例提供的一种语音匹配模型的示例图；

图 3 为本公开实施例提供的一种语音识别方法的过程示意图；

图 4 为本公开实施例提供的一种语音识别结果显示的过程示意图；

图 4a 为本公开实施例提供的一种语音识别结果显示的示意图；

图 5 为本公开实施例提供的一种语音识别方法的过程示意图；

图 6 为本公开实施例提供的一种语音识别结果显示的过程示意图；

图 7 为本公开实施例提供的一种语音识别结果显示的过程示意图；

图 8 为本公开实施例提供的一种语音识别结果显示的示意图；

图 9 为本公开实施例提供的一种语音识别结果显示的示意图；

图 10 为本公开实施例提供的一种服务器的结构示意图；

图 11 为本公开实施例提供的一种终端的结构示意图。

具体实施方式

下面将结合附图对本公开作进一步地详细描述，显然，所描述的实施例

仅仅是本公开一部分实施例，而不是全部的实施例。基于本公开中的实施例，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，都属于本公开保护的范围。

语音识别是让机器接收、识别和理解语音信号，并将其转换成相应的数字信号。虽然语音识别在许多行业产生了大量的应用，但是要实现真正的人机自然交流还需要很多的工作要做，例如，在自适应方面需要更大的改进，达到不受口音、方言和特定人影响的要求。现实中的语音类型是各种各样的，从声音特征来说可以分为男音、女音和童音，另外，很多人的发音同标准发音有很大的差距，这就需要进行口音和方言的处理。导致用户输入语音时得到的识别结果与用户意图不一致的问题。

用户进行语音输入后，用户需要语音识别的词语中存在较多根据常用词进行音译化后得到的，如：四大名助——>四大名著，陆垚知马莉——>路遥知马力，天气预报——>天气预爆等。具有语音识别功能的终端在进行语音识别处理时，由于多音词以及口音、方言和特定人、连续语音识别中的无意义语气词等影响，导致语音识别的影响因素多样，很可能不能识别出用户想要的结果，导致部分用户的意图可能无法实现，造成误识别的情况，进一步影响了语音识别系统在识别速度、识别效率，降低了用户体验效果。

为了解决现有技术存在的问题，以电视机作为终端的场景为例。本公开实施例提供了一种语音识别方法、装置及终端，所述方法包括：接收输入的语音信息；根据预先训练的语音匹配模型，确定满足第一匹配阈值的所述语音信息的至少一个语音识别结果；确定所述至少一个语音识别结果中匹配度最高的语音识别结果为目标语音识别结果；获取所述目标语音识别结果对应的目标文件；将每个语音识别结果及所述目标语音识别结果对应的目标文件显示至显示界面，其中所述目标语音识别结果以第一显示方式显示，其他语音识别结果以第二显示方式显示。通过对第一显示方式显示匹配度最高的语音识别结果，可以进行快速显示，提高用户使用的便捷度；对至少一个语音识别结果分别进行语义识别，获得更多可能的用户搜索意图，并将每个语音

识别结果及至少一个语音识别结果通过第二显示方式显示至所述终端的显示界面，有效的为用户提供更多的搜索结果，有效提高语音识别结果与用户意图的覆盖率，提高用户通过语音搜索的成功率，提高语音识别产品的使用效果。

本公开实施例提供的所有方案，可以由终端执行，也可以由服务器执行，可以根据需要设置，在此不做限定。如图 1 所示，包括：

步骤 101：接收输入的语音信息；

其中，终端可以通过终端的语音设备获取用户输入的语音信息，也可以通过外接的语音设备获取用户输入的语音信息；具体地，终端中设置有语音识别模块，可以识别语音信息并进行语音信息采集。

另外，终端中设置有通信模块，例如 WIFI 无线通讯模块等，使得该终端能够与服务器连接，可以将采集到的语音信息发送给服务器。当然，也可以全部由终端执行，也可以只发送部分需要服务器处理的语音信息，在此不做限定。

步骤 102：根据预先训练的语音匹配模型，确定满足第一匹配阈值的所述语音信息的至少一个语音识别结果；

在具体实施过程中，语音匹配模型可以设置在终端，也可以设置在服务器上，在此不做限定。若设置于服务器上，则服务器确定满足第一匹配阈值的所述语音信息的至少一个语音识别结果后，将所述至少一个语音识别结果发送至终端。

步骤 103：确定所述至少一个语音识别结果中匹配度最高的语音识别结果为目标语音识别结果；

在具体实施过程中，终端可以根据至少一个语音识别结果的得分，确定匹配度最高的目标语音识别结果，也可以为服务器根据至少一个语音识别结果的得分，确定匹配度最高的目标语音识别结果后，将目标语音识别结果发送至终端。

步骤 104：获取所述目标语音识别结果对应的目标文件；

在具体实施过程中，终端可以根据目标语音识别结果，在本地或网络资源库中搜索所述目标语音识别结果对应的目标文件，也可以为根据目标语音识别结果，在网络资源库中搜索所述目标语音识别结果对应的目标文件，确定目标语音识别结果对应的目标文件后，将目标文件或目标文件的标识信息发送至终端，以使终端确定所述目标语音识别结果对应的目标文件。

步骤 105: 将每个语音识别结果及所述目标语音识别结果对应的目标文件显示至显示界面，其中所述目标语音识别结果以第一显示方式显示，其他语音识别结果以第二显示方式显示。

在具体实施过程中，第二显示方式可以为与第一显示方式相对的显示方式。例如，第一显示方式可以为高亮、带有选中框等显示方式，第二显示方式可以为无高亮显示、无选中框等显示方式。例如，如图 7 所示的目标显示结果为选中框高亮显示的第一显示方式，第二显示方式为无选中框高亮显示的方式。具体的显示方式在此不做限定。

本公开实施例提供的一种语音识别方法，通过对第一显示方式显示匹配度最高的语音识别结果，可以进行快速显示，提高用户使用的便捷度；对至少一个语音识别结果分别进行语义识别，获得更多可能的用户搜索意图，并将每个语音识别结果及至少一个语音识别结果通过第二显示方式显示至所述终端的显示界面，有效的为用户提供更多的搜索结果，有效提高语音识别结果与用户意图的覆盖率，提高用户通过语音搜索的成功率，提高语音识别产品的使用效果。

本公开实施例中，如图 2 所示，提供一种语音识别模型确定语音识别结果的方法，包括：

步骤一、获取用户输入的语音信息，通过声学特征确定所述语音信息的特征声学概率；

具体的，声学特征提取，即从语音信息中提取语音声学特征信息，为了保证识别准确率，该提取部分应该对声学模型的建模单元具有较好的区分性。本公开实施例中的声学特征可以包括：梅尔倒谱系数（MFCC）、线性预测倒

谱系数 (LPCC)、感知线性预测系数 (PLP) 等。

步骤二、将声学特征提取后的语音信息输入至语音匹配模型中，语音匹配模型包括语言模型和声学模型；

举例来说，在本公开实施例中，所述语音匹配模型的训练过程可以包括：

步骤一、获取样本语音信息，所述样本语音信息中携带有其所属语音的标注信息；

步骤二、将每个样本语音信息输入到语音匹配模型中；

步骤三、根据每个样本语音信息及所述语音匹配模型的输出，对所述语音匹配模型进行训练。

为了方便语音匹配模型的训练，可以收集大量的样本语音信息，该样本语音信息可以为终端采集的，也可以为其他途径获取的；针对样本语音信息，可以该样本语音信息进行标注。

将样本语音信息输入到该语音匹配模型中，对该语音匹配模型进行训练，该模型可以为有动态时间规整技术、隐马尔可夫模型 (Hidden Markov Model, HMM)、人工神经网络、支持向量机等模型。将每个样本语音信息输入到该模型中，根据每个样本语音信息的标注信息及语音匹配模型的输出结果，对该语音匹配模型进行训练。

本公开实施例中，通过对大量样本语音信息进行训练得到语音匹配模型，并通过该语音匹配模型可以对采集的语音信息进行语音识别。

其中，声学模型利用训练语音特征及其对应的标注信息进行有标注的声学模型建模。声学模型构建语音信号中的观测特征和语音建模单元间的映射关系，以此进行音素或音素状态的分类。本公开实施例中，声学模型可以采用 HMM 作为声学模型的建模基础。

其中，语言模型可以采用基于统计学习建模的语音识别框架下的，N-gram 统计语言模型，其包括一个马尔可夫链表示词序列的生成过程，即将产生词序列 W 的概率 $P(W)$ 表示为：

$$p(W) = \prod_{k=1}^N p(w_k | w_1^{k-1}) \approx \prod_{k=1}^N p(w_k | w_{\max[1, k-n+1]}^{k-1})$$

其中, w_k 表示词序列中的第 k 个词, 上式可见, 产生当前词的概率仅与前面 $n-1$ 个词有关;

本公开实施例中, 语言模型的训练和评价指标可以采用语言模型困惑度 (Perplexity, PP), 它的定义是词序列生成概率集合平均的倒数, 即:

$$PP = [p(w_1^N)]^{-\frac{1}{N}} \approx [\prod_{k=1}^N p(w_k | w_{\max[1, k-n+1]}^{k-1})]^{-\frac{1}{N}}$$

从公式中可见, 语言模型对生成词序列的期望困惑度越小, 则该语言模型在给定历史词序列的情况下对产生何种当前词的预测准确度越高, 因此语言模型的训练目标就是最小化训练集语料的困惑度。

在训练过程中, 首先统计训练集语料中出现的各词及相关词组合的概率, 并以此为基础估计语言模型的相关参数。

然而相关词组合的数目是随可能出现的词表规模呈几何级增长的, 统计所有可能出现的情况并不可行, 而且在现实情况下, 训练数据通常是稀疏的, 有些词间组合出现的概率很小甚至是根本没有出现过。针对这些问题, 可以通过降权 (Discounting) 和回溯 (Backing-off) 等方法, 并利用递归神经网络 (Recurrent Neural Network, RNN) 建模语言模型的方法优化语言模型。

步骤三、将语音模型得到的结果输入至解码器解码后, 获得语音识别的可能的文本信息。

结合声学模型计算得到的语音特征声学概率和由语言模型计算出的语言模型概率, 利用解码器通过相关搜索算法分析出最有可能的词序列 W' , 以输出语音信息中可能的文本信息。

在步骤 102 中, 根据预先训练完成的语音匹配模型, 确定所述语音信息满足第一匹配阈值的语音识别结果, 包括:

步骤一、将所述语音信息输入至所述语音匹配模型, 识别所述语音信息中的拼音序列, 组成所有可能的候选字;

为了对每一个音节确认出正确的字符，先根据输入的拼音序列，组成所有可能的字符假设或单音字节、多音节的词假设。如：以输入序列“郑凯的视频”为例，其对应的拼音序列为[zheng4,kai3,de1,shi4,pin2]。如图 3 所示，每条路径为一个可能的识别结果。

步骤二、针对每个可能的候选字，通过语法规则和统计方法，确定可能的汉字序列及汉字序列的得分；

具体的，通过每个待识别音的多个候选字，利用文法规则和统计学原理，得到汉字序列的得分，且纠正一些拼音识别的错误。应用概率统计语言模型在字符序列或词序列中搜索可能正确的路径。本公开实施例中的解码器可以采用动态规划思想的 Viterbi 算法，并通过一定的算法（例如高斯选择算法（Gaussian Selection）、语言模型前看算法（Language Model Look-Ahead）等）进行快速的同步概率计算以及对搜索空间进行裁剪，以此来降低计算复杂度和内存开销，提高实现搜索算法的效率。

步骤三、将得分满足第一匹配阈值的汉字序列作为所述语音识别结果。

具体的，将语言模型匹配的至少一个汉字序列及得分，进行排序。从模板匹配的结果来看，匹配度高的识别结果，识别正确的概率更高。当然也存在由于某个模型中未补充语料导致虽然模型匹配得分较高，但并非正确结果的情况。因此，可以选取匹配得分需要满足第一阈值的语音识别结果进行语义识别。

具体的，得分需要满足第一阈值，即得分大于第一阈值 ρ 的语音识别结果作为可能正确的识别结果。举例来说，识别结果如下：

识别结果	匹配得分
郑恺的视频	0.641
郑凯的视频	0.629
正楷的视频	0.457
正凯德食品	0.231

如上表所示，第一阈值可以为 0.4；则此时的语音识别结果为：“郑恺的视频、郑凯的视频、正楷的视频”。

在一些实施方式中，当所有的识别结果模型匹配分值均小于第一阈值 ρ 时，则取得分最高的识别结果执行步骤 104，进行语义处理。

为提高用户体验，有效显示搜索进程，在步骤 103 之前，还可以将目标语音识别结果，输出到终端的界面并显示。其中，终端的界面可以为采集语音信息的语音助手的客户端的显示界面，也可以是终端的其他界面，在此不做限定。举例来说，可以如图 4a 所示，所述目标语音识别结果为“郑恺的视频”。

如图 4 所示，显示识别结果的流程包括以下步骤：

步骤一、创建界面的布局文件；

其中，所述布局文件包括显示语音识别结果的文本控件。

步骤二、创建界面加载布局文件，初始化文本控件。

步骤三、在终端的显示界面显示语音识别结果，即识别的文本信息。

为了有效地提高识别的准确度和覆盖度，服务器中存储有预设的词典库，该词典库中包括大量的语料数据，具有语义解析功能，在云服务器判断接收到的语音信息后，利用自身的语义解析功能对该语音识别结果进行语义解析处理。具体的，服务器中保存有语义识别模型，该语义识别模型可以识别出语音信息的分词；确定该语音信息中的分词，识别分词的语义，确定每个语义对应的目标文件。当然，若需要检索的词典库较小，为提高解析速率，语义识别可以在终端上完成，在此不做限定。

在步骤 104 中，包括：

步骤一、对目标语音识别结果进行语义识别，确定目标语音识别结果对应的业务类型；

若终端执行语义识别，则可以根据终端上的语义识别模型，输出语音识别结果的分词，解析分词的语义，及语义对应的标注结果，查找该标注结果中是否包含与业务类型相关的业务类型。

若服务器执行语义识别，服务器在接收到终端发送的语音识别结果后，根据服务器上的语义识别模型，输出语音识别结果的分词，解析分词的语义，及语义对应的标注结果，查找该标注结果中是否包含与业务类型相关的业务类型。

步骤二、从资源库中在所述目标语音识别结果对应的业务类型中查找所述目标语音识别结果对应的目标文件。

为进一步提高语义识别的准确度，本公开实施例中，语义识别模型识别具体实施过程，可以包括：

步骤一、根据预设的词典库，对目标语音识别结果进行分词处理，并对目标语音识别结果中各分词进行语义识别，确定各分词对应的业务类型；

其中，预设的词典库可以通过网络爬虫等方法获取语料，以更新分词及对应的业务类型的标注。

步骤二、根据各分词对应的业务类型的权重，确定目标语音识别结果对应的业务类型。

为进一步提高检索效率，针对除目标语音识别结果，超过第一阈值的语音识别结果中的其他语音识别结果也可以与目标语音识别结果同时执行上述操作。当然，也可以在收到用户的切换指令后，再执行上述操作，在此不做限定。

如图 5 所示，具体的，可以包括：

步骤一、对所述至少一个语音识别结果中的每个语音识别结果进行语义识别，确定所述语音识别结果对应的业务类型；

具体地，将语音识别结果 1 输入语义识别模型，如果语义识别模型输出的结果中包含的业务类型 1，则认为该语音识别结果 1 中包含业务类型 1，需要在业务类型 1 对应的应用程序中执行后续处理。

举例来说，语音识别结果 1 为“郑恺的视频”，语义识别模型输出的分词结果为：郑恺,的,视频。其中，视频的业务类型为视频类型，则语音识别结果的业务类型为视频类型。

在一些实施方式中，业务类型，还可以根据分词的属性确定。例如，语音识别结果 2 为“天气预报”，语义识别模型确定的分词结果为：天气，预报；“天气”具有天气属性（weatherKeys），则确定业务类型为天气查询类型。

为进一步提高语义识别的准确度，本公开实施例中，语义识别模型识别具体实施过程，可以包括：

步骤一、根据预设的词典库，对所述语音识别结果进行分词处理，并对所述语音识别结果中各分词进行语义识别，确定各分词对应的业务类型；

其中，预设的词典库可以通过网络爬虫等方法获取语料，以更新分词及对应的业务类型的标注。

步骤二、根据各分词对应的业务类型的权重，确定所述语音识别结果对应的业务类型。

在一些实施方式中，所述业务类型的权重为根据所述业务类型在所述终端中的优先级、所述预设的词典库中的分词所来源的资料库的优先级或所述终端用户的用户偏好中的至少一项确定的。

举例来说，语音识别结果 3 为“正楷的视频”，语义识别模型确定的分词结果为：正楷,的,视频。视频的业务类型为视频类型；正楷的业务类型为教育类型；若确定“视频”对应的视频类型的权重大于“正楷”对应的教育类型的权重，则确定语音识别结果 3 的业务类型为视频类型。若确定“视频”对应的视频类型的权重与“正楷”对应的教育类型的权重相同，也可以将语音识别结果 3 对应的业务类型确定为教育类型和视频类型。

再比如，语音识别结果 4 为“天气预爆”，语义识别模型确定的分词结果为：天气预爆；根据预设的词典库确定天气预爆为一电影，其对应的业务类型包括视频类型、歌曲类型等；则根据“天气预爆”对应的视频类型的权重，及“天气预爆”对应的歌曲类型的权重，确定语音识别结果 4 的业务类型。

在步骤二中，从资源库中在至少一个语音识别结果对应的业务类型中查找至少一个语音识别结果对应的目标文件。

结合上述举例，对于语音识别结果 1，可以从资源库中的视频类型中搜索

郑恺的目标文件。对于语音识别结果 2，可以从资源库中的天气查询业务中搜索天气预报的目标文件。对于语音识别结果 3，可以从资源库中的视频类型或教育类型或教育视频类型中搜索正楷的目标文件。对于语音识别结果 4，可以从资源库中的视频类型或歌曲类型中搜索天气预爆的目标文件。

在步骤 105 中，具体包括：

步骤一、确定所述至少一个语音识别结果中每个语音识别结果的优先级；

具体的，结合语义分析用户界面（User Interface，UI）以键盘制表定位键（Tabulator key，TAB）形式展示搜索结果，展示结果的排序主要是根据热搜排名进行 TAB 排序。

步骤二、在所述终端的显示界面上按照所述优先级排列显示每个语音识别结果；

其优先级可以基于用户大数据分析、得分及用户偏好等方式确定，在此不做限定。

步骤三、将每个语音识别结果及所述目标语音识别结果对应的目标文件显示至显示界面。

在具体实施过程中，如图 6 所示，包括：

语义识别模块将语音识别结果对应的 TAB 数据和目标文件转换为 JSON 数据，传输至终端的显示模块；

终端的显示模块获取所述 JSON 数据后，解析对应的语音识别结果和目标文件；

根据解析结果显示每个语音识别结果及对应的目标文件。

结合上述举例，若确定排序结果为：郑恺>郑凯>正楷，则显示结果可以如图 7 所示。

在一些实施方式中，对于语义分析无法确定的业务类型或无法确定对应的目标文件时，则不将该语音识别结果显示至终端。如“正楷的视频”语义无法理解或资源库中搜不到“正楷”书法相关内容，则不将该语音识别结果显示至终端。

结合上述举例，若确定排序结果为：天气预报>天气预爆，则显示结果可以如图 8 和图 9 所示。

进一步的，若用户想要切换目标语音识别结果，则可以进行语音识别结果的切换，具体包括：

获取用户对所述目标语音识别结果的切换指令；

根据所述切换指令，确定更改后的目标语音识别结果对应的目标文件；

将更改后的所述目标语音识别结果以第一显示方式显示，其他语音识别结果以第二显示方式显示；同时显示更改后的所述目标语音识别结果对应的目标文件。

具体的，确定更改后的目标语音识别结果对应的目标文件可以参考上述实施例，在此不再赘述。

为了进一步提高识别的语音识别的准确率，在本公开实施例中，所述方法还包括：

获取用户对所述语音识别结果或目标文件的操作指令；

增加所述操作指令对应的语音识别结果或目标文件的匹配度，以更新用户偏好。

例如，用户在显示界面中选择“天气预爆”，则在用户的用户偏好中，记录“天气预爆”，并增加“天气预爆”的匹配度。

为了进一步提高识别的语音识别的准确率，本公开实施例，还提供在一些实施方式中，包括：

判断所述语音信息是否包括对所述终端进行控制的第一控制指令；

如果所述用户语音信息为对所述终端进行控制的第一控制指令，则在所述终端中执行所述第一控制指令。

在一些实施方式中，如果语音信息中还包含操作类型的分词，说明终端有必要根据该语音信息进行相应操作。此时可以直接向终端发送根据所述语音信息进行处理的分词。例如，打开，观看，播放等操作类型的分词。

在一些实施方式中，对语音信息的语义中是否包含针对终端设置的目标

控制指令进行判断，如果是，则在所述终端中执行所述第一控制指令。

例如，识别的语音识别结果为“打开郑凯的视频”，则可以确定第一控制指令为打开。

在一些实施方式中，若确定“郑凯的视频”的目标文件唯一，则可以直接执行打开“郑凯的视频”的目标文件。

在一些实施方式中，若确定“郑凯的视频”的目标文件有多个，可以先显示多个目标文件，在获取用户的操作指令后，执行打开的控制指令。

本公开实施例，通过在语音匹配模型识别所述语音信息将匹配得分最高的识别结果展示给用户，同时将满足第一匹配阈值的至少一个语音识别结果分别进行语义识别，结合语义处理结果，可以通过 UI 交互展示不同的业务搜索结果给用户，更好地理解用户意图，与现有技术中的语音识别方法相比，通过多次语义分析请求，实现同音词名称业务的搜索和展现，用户可根据意向选择想要的结果。

在上述各实施例的基础上，本公开实施例还提供了一种服务器 1100，如图 10 所示，包括：处理器 1101、通信接口 1102、存储器 1103 和通信总线 1104，其中，处理器 1101，通信接口 1102，存储器 1103 通过通信总线 1104 完成相互间的通信；

所述存储器 1103 中存储有计算机程序；

所述处理器 1101，配置为运行所述计算机程序以使服务器 1100 实现：

接收终端发送的语音信息；根据预先训练完成的语音匹配模型，确定所述语音信息满足第一匹配阈值的至少一个语音识别结果；确定所述至少一个语音识别结果中匹配度最高的语音识别结果为目标语音识别结果；获取所述目标语音识别结果对应的目标文件；将每个语音识别结果及所述目标语音识别结果对应的目标文件显示至终端的显示界面，其中所述目标语音识别结果以第一显示方式显示，其他语音识别结果以第二显示方式显示。

在一些实施方式中，所述处理器 1101，配置为通过下述执行所述获取所述目标语音识别结果对应的目标文件：

对所述目标语音识别结果进行语义识别，确定所述目标语音识别结果对应的业务类型；从资源库中在所述目标语音识别结果对应的业务类型中查找所述目标语音识别结果对应的目标文件。

在一些实施方式中，所述处理器 1101，配置为通过下述执行所述对所述目标语音识别结果进行语义识别，确定所述目标语音识别结果对应的业务类型：

根据预设的词典库，对所述目标语音识别结果进行分词处理，并对所述目标语音识别结果中各分词进行语义识别，确定各分词对应的业务类型；根据各分词对应的业务类型的权重，确定所述目标语音识别结果对应的业务类型。

在一些实施方式中，所述处理器 1101，配置为通过下述执行所述将每个语音识别结果及所述目标语音识别结果对应的目标文件显示至显示界面：

确定每个语音识别结果的优先级，并将每个语音识别结果的优先级发送给终端，以使终端在显示界面上按照所述优先级排列显示所述每个语音识别结果，并将所述目标语音识别结果对应的目标文件在所述终端的显示界面上显示。

在一些实施方式中，所述处理器 1101，还配置为运行所述计算机程序以使所述服务器 1100 实现：

通过通信接口 1102 接收终端发送的用户对所述目标语音识别结果的切换指令；根据所述切换指令，确定更改后的目标语音识别结果对应的目标文件并发送给终端，以使终端将更改后的所述目标语音识别结果以第一显示方式显示，其他语音识别结果以第二显示方式显示，同时显示更改后的所述目标语音识别结果对应的目标文件。

在一些实施方式中，所述处理器 1101，配置为通过下述执行所述根据预先训练完成的语音匹配模型，确定所述语音信息满足第一匹配阈值的语音识别结果：

将所述语音信息输入至所述语音匹配模型，识别所述语音信息中的拼音

序列，组成所有可能的候选字；针对每个可能的候选字，通过语法规则和统计方法，确定可能的汉字序列及汉字序列的得分；将得分满足第一匹配阈值的汉字序列作为所述语音识别结果。

上述服务器提到的通信总线可以是外设部件互连标准（Peripheral Component Interconnect, PCI）总线或扩展工业标准结构（Extended Industry Standard Architecture, EISA）总线等。该通信总线可以分为地址总线、数据总线、控制总线等。为便于表示，图中仅用一条粗线表示，但并不表示仅有一根总线或一种类型的总线。

通信接口 1102 用于上述服务器与其他设备之间的通信。

存储器可以包括随机存取存储器（Random Access Memory, RAM），也可以包括非易失性存储器（Non-Volatile Memory, NVM），例如至少一个磁盘存储器。可选地，存储器还可以是至少一个位于远离前述处理器的存储装置。

上述处理器可以是通用处理器，包括中央处理器、网络处理器（Network Processor, NP）等；还可以是数字指令处理器（Digital Signal Processing, DSP）、专用集成电路、现场可编程门阵列或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。

在上述各实施例的基础上，本公开实施例还提供了一种计算机存储可读非易失性存储介质，所述计算机可读非易失性存储介质内存储有可由服务器执行的计算机程序，当所述程序在所述服务器上运行时，使得所述服务器执行时实现上述实施例中的任一种方法。

上述计算机可读非易失性存储介质可以是服务器中的处理器能够存取的任何可用介质或数据存储设备，包括但不限于磁性存储器如软盘、硬盘、磁带、磁光盘（MO）等、光学存储器如光盘（CD）、数字通用光盘（DVD）、蓝光光盘（BD）、高清通用光盘（HVD）等、以及半导体存储器如只读存储器（ROM）、可擦除可编程只读存储器（EPROM）、带电可擦除可编程只读存储器（EEPROM）、非易失性存储器（NAND FLASH）、固态硬盘（SSD）等。

在上述各实施例的基础上，本公开实施例还提供了一种终端 1200，如图

11 所示, 包括: 处理器 1201、通信接口 1202、存储器 1203 和通信总线 1204, 其中, 处理器 1201, 通信接口 1202, 存储器 1203 通过通信总线 1204 完成相互间的通信;

所述存储器 1203 中存储有计算机程序;

所述处理器 1201, 配置为运行所述计算机程序以使终端 1200 实现:

接收输入的语音信息; 根据预先训练的语音匹配模型, 确定满足第一匹配阈值的所述语音信息的至少一个语音识别结果; 确定所述至少一个语音识别结果中匹配度最高的语音识别结果为目标语音识别结果; 获取所述目标语音识别结果对应的目标文件; 将每个语音识别结果及所述目标语音识别结果对应的目标文件显示至显示界面, 其中所述目标语音识别结果以第一显示方式显示, 其他语音识别结果以第二显示方式显示。

在一些实施方式中, 所述处理器 1201, 配置为通过下述执行所述获取所述目标语音识别结果对应的目标文件:

对所述目标语音识别结果进行语义识别, 确定所述目标语音识别结果对应的业务类型; 从资源库中在所述目标语音识别结果对应的业务类型中查找所述目标语音识别结果对应的目标文件。

在一些实施方式中, 所述处理器 1201, 配置为通过下述执行所述对所述目标语音识别结果进行语义识别, 确定所述目标语音识别结果对应的业务类型:

根据预设的词典库, 对所述目标语音识别结果进行分词处理, 并对所述目标语音识别结果中各分词进行语义识别, 确定各分词对应的业务类型; 根据各分词对应的业务类型的权重, 确定所述目标语音识别结果对应的业务类型。

在一些实施方式中, 所述处理器 1201, 配置为通过下述执行所述将每个语音识别结果及所述目标语音识别结果对应的目标文件显示至显示界面:

确定每个语音识别结果的优先级; 在所述终端的显示界面上按照所述优先级排列显示所述每个语音识别结果; 将所述目标语音识别结果对应的目标

文件在所述终端的显示界面上显示。

在一些实施方式中，所述处理器 1201，还配置为运行所述计算机程序以使所述终端实现：

获取用户对所述目标语音识别结果的切换指令；根据所述切换指令，确定更改后的目标语音识别结果对应的目标文件；将更改后的所述目标语音识别结果以第一显示方式显示，其他语音识别结果以第二显示方式显示；同时显示更改后的所述目标语音识别结果对应的目标文件。其中，切换指令为通过通信接口 1202 获取的用户对所述目标语音识别结果的切换指令。

在一些实施方式中，所述处理器 1201，配置为通过下述执行所述根据预先训练完成的语音匹配模型，确定所述语音信息满足第一匹配阈值的语音识别结果：

将所述语音信息输入至所述语音匹配模型，识别所述语音信息中的拼音序列，组成所有可能的候选字；针对每个可能的候选字，通过语法规则和统计方法，确定可能的汉字序列及汉字序列的得分；将得分满足第一匹配阈值的汉字序列作为所述语音识别结果。

上述终端提到的通信总线可以是外设部件互连标准（Peripheral Component Interconnect, PCI）总线或扩展工业标准结构（Extended Industry Standard Architecture, EISA）总线等。该通信总线可以分为地址总线、数据总线、控制总线等。为便于表示，图中仅用一条粗线表示，但并不表示仅有一根总线或一种类型的总线。

通信接口 1202 用于上述终端与其他设备之间的通信。

存储器可以包括 RAM，也可以包括 NVM，例如至少一个磁盘存储器。可选地，存储器还可以是至少一个位于远离前述处理器的存储装置。

上述处理器可以是通用处理器，包括中央处理器、NP 等；还可以是 DSP、专用集成电路、现场可编程门阵列或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。

在上述各实施例的基础上，本公开实施例还提供了一种计算机存储可读

非易失性存储介质，所述计算机可读非易失性存储介质内存储有可由终端执行的计算机程序，当所述程序在所述终端上运行时，使得所述终端执行时实现上述实施例中的任一种方法。

上述计算机可读非易失性存储介质可以是终端中的处理器能够存取的任何可用介质或数据存储设备，包括但不限于磁性存储器如软盘、硬盘、磁带、磁光盘（MO）等、光学存储器如 CD、DVD、BD、HVD 等、以及半导体存储器如 ROM、EPROM、EEPROM、非易失性存储器（NAND FLASH）、固态硬盘（SSD）等。

对于系统/装置实施例而言，由于其基本相似于方法实施例，所以描述的比较简单，相关之处参见方法实施例的部分说明即可。

需要说明的是，在本文中，诸如第一和第二等之类的关系术语仅仅用来将一个实体或者一个操作与另一个实体或者另一个操作区分开来，而不一定要求或者暗示这些实体或者操作之间存在任何这种实际的关系或者顺序。

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、系统、或计算机程序产品。因此，本申请可采用完全硬件实施例、完全应用实施例、或结合应用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。

本公开是参照根据本公开实施例的方法、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。

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。

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。

尽管已描述了本公开的优选实施例，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，所附权利要求要求意欲解释为包括优选实施例以及落入本公开范围的所有变更和修改。

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

权 利 要 求

1、一种语音识别方法，所述方法包括：

接收输入的语音信息；

根据预先训练的语音匹配模型，确定满足第一匹配阈值的所述语音信息的至少一个语音识别结果；

确定所述至少一个语音识别结果中匹配度最高的语音识别结果为目标语音识别结果；

获取所述目标语音识别结果对应的目标文件；

将每个语音识别结果及所述目标语音识别结果对应的目标文件显示至显示界面，其中所述目标语音识别结果以第一显示方式显示，其他语音识别结果以第二显示方式显示。

2、如权利要求 1 所述的方法，所述获取所述目标语音识别结果对应的目标文件，包括：

对所述目标语音识别结果进行语义识别，确定所述目标语音识别结果对应的业务类型；

从资源库中在所述目标语音识别结果对应的业务类型中查找所述目标语音识别结果对应的目标文件。

3、如权利要求 2 所述的方法，所述对所述目标语音识别结果进行语义识别，确定所述目标语音识别结果对应的业务类型，包括：

根据预设的词典库，对所述目标语音识别结果进行分词处理，并对所述目标语音识别结果中各分词进行语义识别，确定各分词对应的业务类型；

根据各分词对应的业务类型的权重，确定所述目标语音识别结果对应的业务类型。

4、如权利要求 1 所述的方法，所述将每个语音识别结果及所述目标语音识别结果对应的目标文件显示至显示界面，包括：

确定每个语音识别结果的优先级；

在所述终端的显示界面上按照所述优先级排列显示所述每个语音识别结果；

将所述目标语音识别结果对应的目标文件在所述终端的显示界面上显示。

5、如权利要求4所述的方法，所述将每个语音识别结果及所述目标语音识别结果对应的目标文件显示至显示界面之后，所述方法还包括：

获取用户对所述目标语音识别结果的切换指令；

根据所述切换指令，确定更改后的目标语音识别结果对应的目标文件；

将更改后的所述目标语音识别结果以第一显示方式显示，其他语音识别结果以第二显示方式显示；同时显示更改后的所述目标语音识别结果对应的目标文件。

6、如权利要求1至5任一项所述的方法，所述根据预先训练完成的语音匹配模型，确定所述语音信息满足第一匹配阈值的语音识别结果，包括：

将所述语音信息输入至所述语音匹配模型，识别所述语音信息中的拼音序列，组成所有可能的候选字；

针对每个可能的候选字，通过语法规则和统计方法，确定可能的汉字序列及汉字序列的得分；

将得分满足第一匹配阈值的汉字序列作为所述语音识别结果。

7、一种终端，包括存储器和处理器，其中，

所述存储器，存储有计算机程序；

所述处理器，与所述存储器通信，配置为运行所述计算机程序以使所述终端实现：

接收输入的语音信息；

根据预先训练的语音匹配模型，确定满足第一匹配阈值的所述语音信息的至少一个语音识别结果；

确定所述至少一个语音识别结果中匹配度最高的语音识别结果为目标语音识别结果；

获取所述目标语音识别结果对应的目标文件；

将每个语音识别结果及所述目标语音识别结果对应的目标文件显示至显示界面，其中所述目标语音识别结果以第一显示方式显示，其他语音识别结果以第二显示方式显示。

8、如权利要求 7 所述的终端，所述处理器，配置为通过下述执行所述获取所述目标语音识别结果对应的目标文件：

对所述目标语音识别结果进行语义识别，确定所述目标语音识别结果对应的业务类型；

从资源库中在所述目标语音识别结果对应的业务类型中查找所述目标语音识别结果对应的目标文件。

9、如权利要求 8 所述的终端，所述处理器，配置为通过下述执行所述对所述目标语音识别结果进行语义识别，确定所述目标语音识别结果对应的业务类型：

根据预设的词典库，对所述目标语音识别结果进行分词处理，并对所述目标语音识别结果中各分词进行语义识别，确定各分词对应的业务类型；

根据各分词对应的业务类型的权重，确定所述目标语音识别结果对应的业务类型。

10、如权利要求 7 所述的终端，所述处理器，配置为通过下述执行所述将每个语音识别结果及所述目标语音识别结果对应的目标文件显示至显示界面：

确定每个语音识别结果的优先级；

在所述终端的显示界面上按照所述优先级排列显示所述每个语音识别结果；

将所述目标语音识别结果对应的目标文件在所述终端的显示界面上显示。

11、如权利要求 10 所述的终端，所述处理器，还配置为运行所述计算机程序以使所述终端实现：

获取用户对所述目标语音识别结果的切换指令；

根据所述切换指令，确定更改后的目标语音识别结果对应的目标文件；

将更改后的所述目标语音识别结果以第一显示方式显示，其他语音识别结果以第二显示方式显示；同时显示更改后的所述目标语音识别结果对应的目标文件。

12、如权利要求 7 至 11 任一项所述的终端，所述处理器，配置为通过下述执行所述根据预先训练完成的语音匹配模型，确定所述语音信息满足第一匹配阈值的语音识别结果：

将所述语音信息输入至所述语音匹配模型，识别所述语音信息中的拼音序列，组成所有可能的候选字；

针对每个可能的候选字，通过语法规则和统计方法，确定可能的汉字序列及汉字序列的得分；

将得分满足第一匹配阈值的汉字序列作为所述语音识别结果。

13、一种计算机可读非易失性存储介质，所述可读非易失性存储介质上存储计算机程序，所述计算机程序被处理器执行时实现如权利要求 1 至 6 中任一项所述的方法。

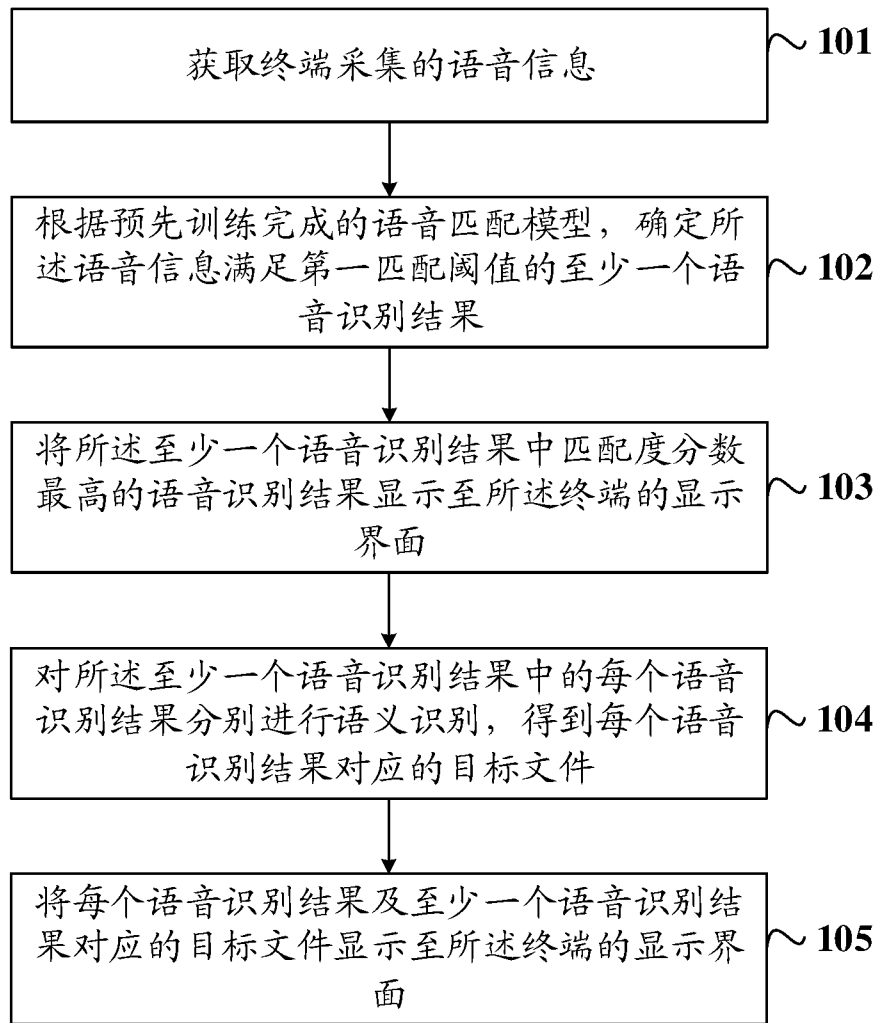


图 1

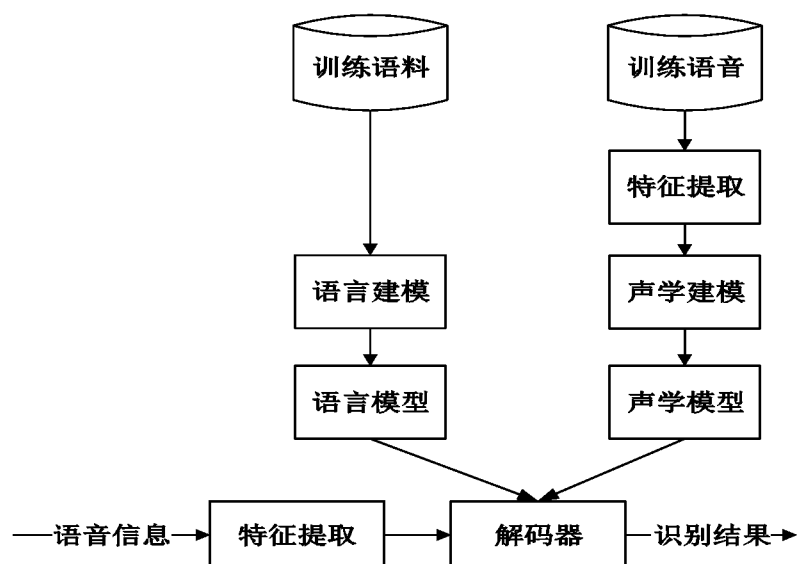


图 2

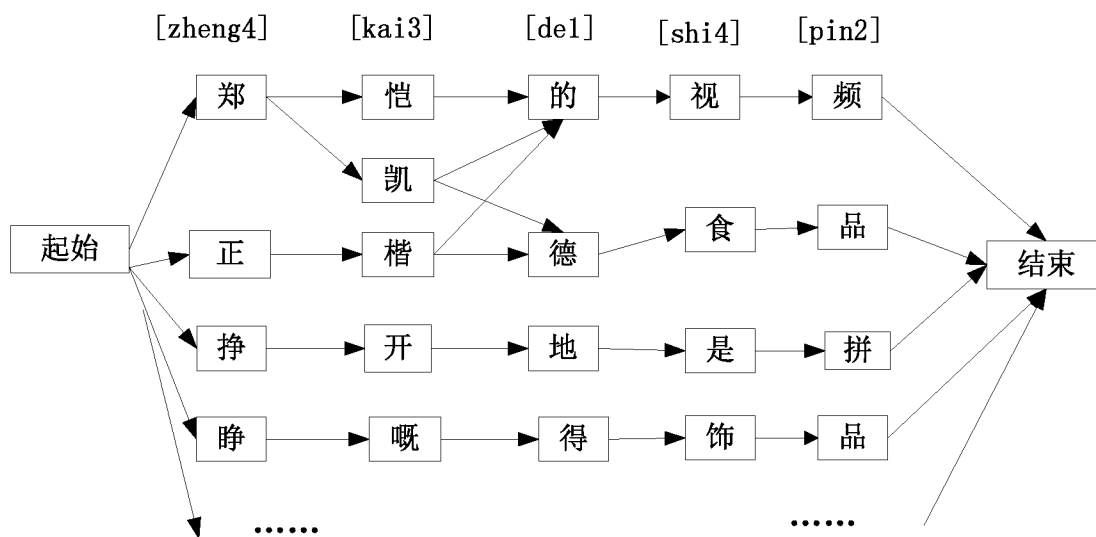


图 3

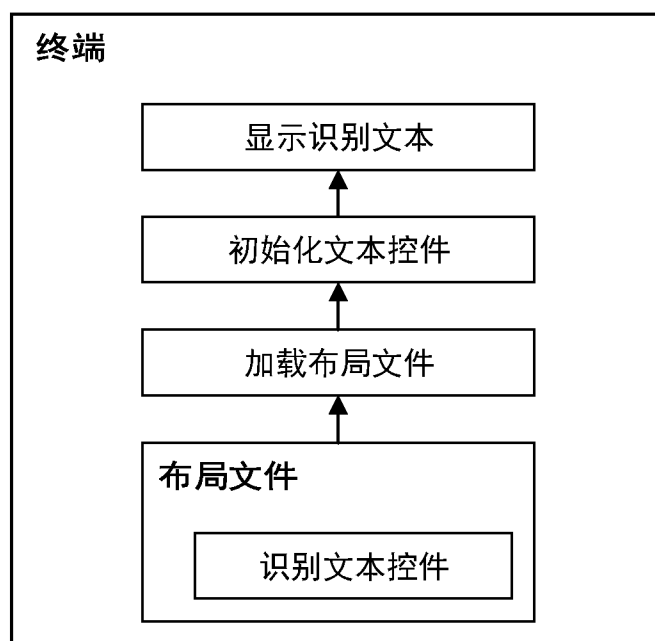


图 4



图 4a

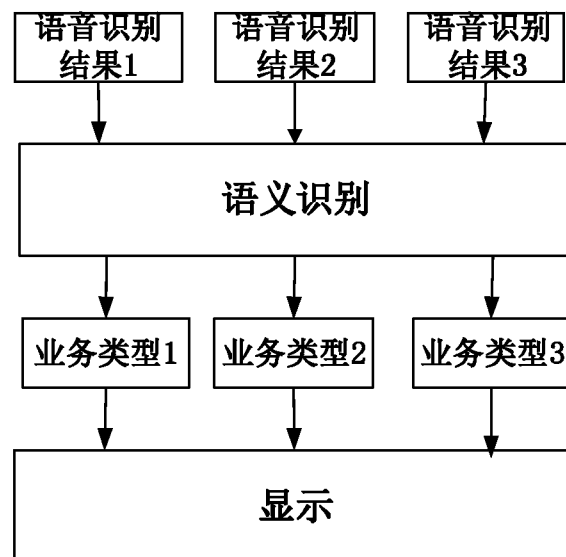


图 5

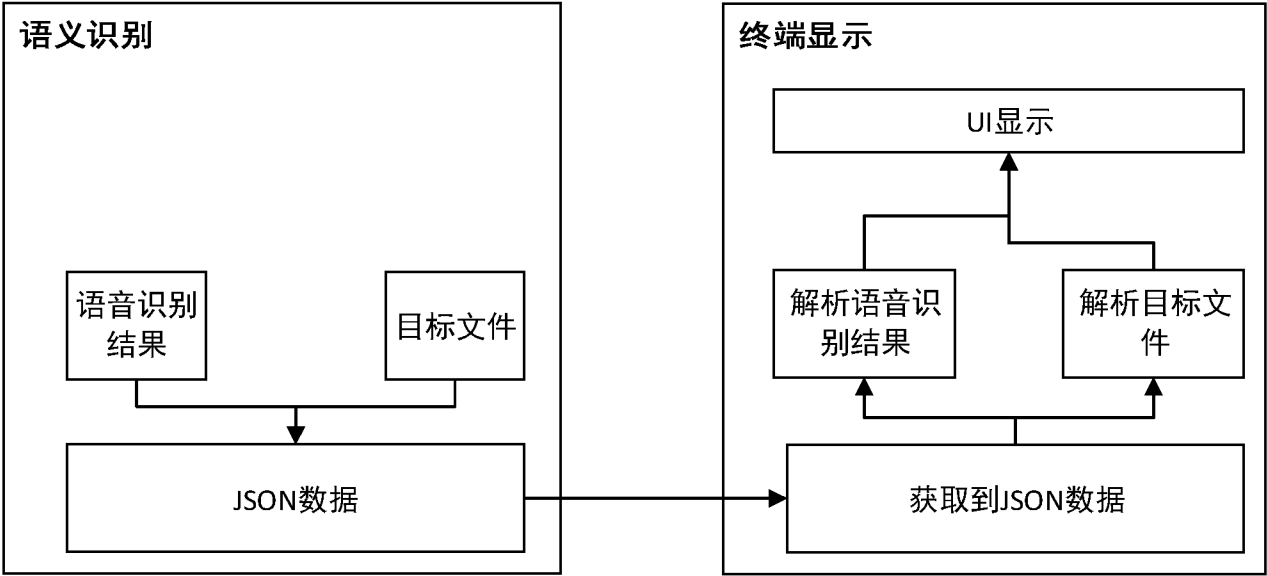


图 6



图 7



图 8



图 9

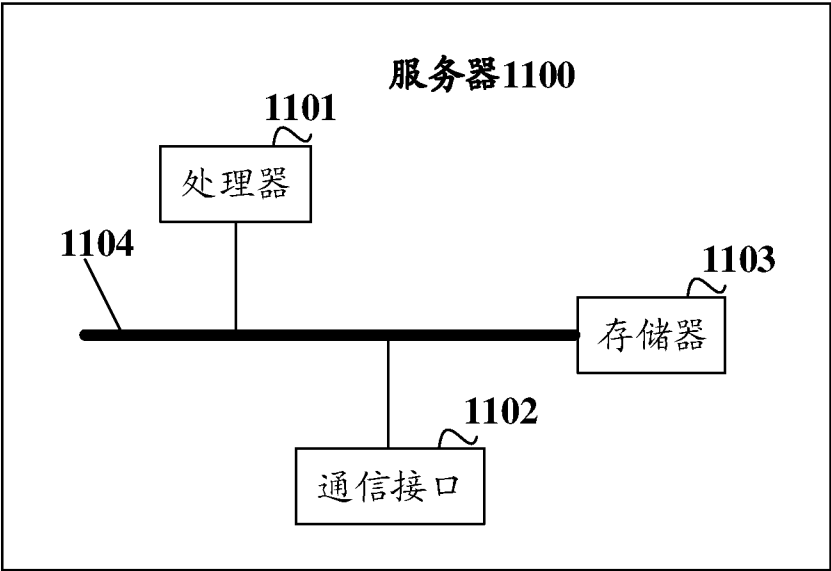


图 10

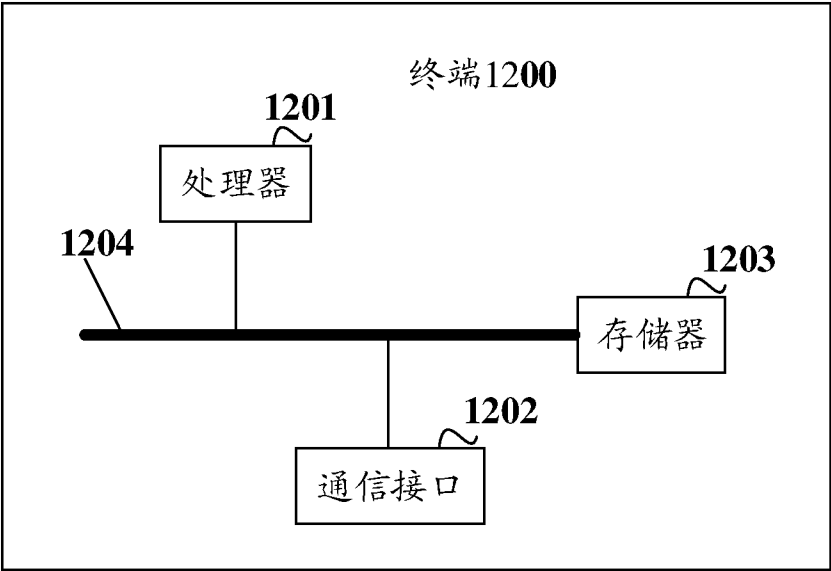


图 11

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/106806

A. CLASSIFICATION OF SUBJECT MATTER

G06F 3/16(2006.01)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)

G06F; 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)

CNPAT, CNKI, WPI, EPODOC: 语音, 识别, 匹配, 阈值, 门限, 度, 分数, 分值, 打分, 最, 第一, 文件, 曲目, 影像, 视频, 音乐, 显示方式, 高亮, 突出, 醒目, 反色, 粗体, 加粗, 字号, ASR, speech recogn+, match+, threshold?, most, highest, best, file, video, music, media, movie?, highlight, font, bold

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 109976702 A (QINGDAO HISENSE ELECTRIC CO., LTD.) 05 July 2019 (2019-07-05) description, paragraphs [0063]-[0221], and claims 1-10	1-13
X	CN 101557432 A (LG ELECTRONICS INC.) 14 October 2009 (2009-10-14) description, page 7, paragraphs 3 and 4, page 13, paragraph 2 to page 20, paragraph 1, page 23, paragraph 2 to page 24, paragraph 3, page 28, paragraph 4 to page 30, paragraph 2, figures 1-19, and claims 1-9	1-13
X	CN 101557651 A (LG ELECTRONICS INC.) 14 October 2009 (2009-10-14) description, page 6, paragraphs 6 and 7 and page 10, the last paragraph but one to page 19, paragraph 1, figures 1-9, and claims 1-10	1-13
X	JP 1021254 A (TOSHIBA CORPORATION) 23 January 1998 (1998-01-23) description, paragraphs [0007]-[0054], and figures 1-8	1-13
A	CN 106356056 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 25 January 2017 (2017-01-25) 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

19 November 2019

Date of mailing of the international search report

27 November 2019

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/CN2019/106806

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	109976702	A	05 July 2019	None			
CN	101557432	A	14 October 2009	KR	20090107364	A	13 October 2009
				US	2009253463	A1	08 October 2009
				RU	2008126780	A	10 January 2010
				KR	20090107366	A	13 October 2009
				KR	20090107852	A	14 October 2009
				JP	2009253970	A	29 October 2009
				EP	2109295	A1	14 October 2009
				US	2015126252	A1	07 May 2015
				US	2017019515	A1	19 January 2017
				US	2017257470	A1	07 September 2017
CN	101557651	A	14 October 2009	US	2009254351	A1	08 October 2009
				JP	2009252238	A	29 October 2009
				EP	2109298	A1	14 October 2009
				RU	2008126782	A	10 January 2010
				KR	20090107365	A	13 October 2009
				US	2012130712	A1	24 May 2012
JP	1021254	A	23 January 1998	None			
CN	106356056	A	25 January 2017	None			

国际检索报告

国际申请号

PCT/CN2019/106806

A. 主题的分类 G06F 3/16 (2006.01) 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) G06F; 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)) CNPAT, CNKI, WPI, EPDOC: 语音, 识别, 匹配, 阈值, 门限, 度, 分数, 分值, 打分, 最, 第一, 文件, 曲目, 影像, 视频, 音乐, 显示方式, 高亮, 突出, 醒目, 反色, 粗体, 加粗, 字号, ASR, speech recogn+, match+, threshold?, most, highest, best, file, video, music, media, movie?, highlight, font, bold																				
C. 相关文件 <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>PX</td> <td>CN 109976702 A (青岛海信电器股份有限公司) 2019年 7月 5日 (2019 - 07 - 05) 说明书第[0063]-[0221]段、权利要求1-10</td> <td>1-13</td> </tr> <tr> <td>X</td> <td>CN 101557432 A (LG电子株式会社) 2009年 10月 14日 (2009 - 10 - 14) 说明书第7页第3-4段, 第13页第2段-第20页第1段, 第23页第2段-第24页第3段, 第28页第4段-第30页第2段、图1-19、权利要求1-9</td> <td>1-13</td> </tr> <tr> <td>X</td> <td>CN 101557651 A (LG电子株式会社) 2009年 10月 14日 (2009 - 10 - 14) 说明书第6页第6-7段, 第10页倒数第2段-第19页第1段、图1-9、权利要求1-10</td> <td>1-13</td> </tr> <tr> <td>X</td> <td>JP 1021254 A (TOSHIBA CORP.) 1998年 1月 23日 (1998 - 01 - 23) 说明书第[0007]-[0054]段、图1-8</td> <td>1-13</td> </tr> <tr> <td>A</td> <td>CN 106356056 A (腾讯科技深圳有限公司) 2017年 1月 25日 (2017 - 01 - 25) 全文</td> <td>1-13</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 109976702 A (青岛海信电器股份有限公司) 2019年 7月 5日 (2019 - 07 - 05) 说明书第[0063]-[0221]段、权利要求1-10	1-13	X	CN 101557432 A (LG电子株式会社) 2009年 10月 14日 (2009 - 10 - 14) 说明书第7页第3-4段, 第13页第2段-第20页第1段, 第23页第2段-第24页第3段, 第28页第4段-第30页第2段、图1-19、权利要求1-9	1-13	X	CN 101557651 A (LG电子株式会社) 2009年 10月 14日 (2009 - 10 - 14) 说明书第6页第6-7段, 第10页倒数第2段-第19页第1段、图1-9、权利要求1-10	1-13	X	JP 1021254 A (TOSHIBA CORP.) 1998年 1月 23日 (1998 - 01 - 23) 说明书第[0007]-[0054]段、图1-8	1-13	A	CN 106356056 A (腾讯科技深圳有限公司) 2017年 1月 25日 (2017 - 01 - 25) 全文	1-13
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																		
PX	CN 109976702 A (青岛海信电器股份有限公司) 2019年 7月 5日 (2019 - 07 - 05) 说明书第[0063]-[0221]段、权利要求1-10	1-13																		
X	CN 101557432 A (LG电子株式会社) 2009年 10月 14日 (2009 - 10 - 14) 说明书第7页第3-4段, 第13页第2段-第20页第1段, 第23页第2段-第24页第3段, 第28页第4段-第30页第2段、图1-19、权利要求1-9	1-13																		
X	CN 101557651 A (LG电子株式会社) 2009年 10月 14日 (2009 - 10 - 14) 说明书第6页第6-7段, 第10页倒数第2段-第19页第1段、图1-9、权利要求1-10	1-13																		
X	JP 1021254 A (TOSHIBA CORP.) 1998年 1月 23日 (1998 - 01 - 23) 说明书第[0007]-[0054]段、图1-8	1-13																		
A	CN 106356056 A (腾讯科技深圳有限公司) 2017年 1月 25日 (2017 - 01 - 25) 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 2019年 11月 19日		国际检索报告邮寄日期 2019年 11月 27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 李文 电话号码 86-(10)-53961682																		

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/106806

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	109976702	A	2019年 7月 5日	无			
CN	101557432	A	2009年 10月 14日	KR	20090107364	A	2009年 10月 13日
				US	2009253463	A1	2009年 10月 8日
				RU	2008126780	A	2010年 1月 10日
				KR	20090107366	A	2009年 10月 13日
				KR	20090107852	A	2009年 10月 14日
				JP	2009253970	A	2009年 10月 29日
				EP	2109295	A1	2009年 10月 14日
				US	2015126252	A1	2015年 5月 7日
				US	2017019515	A1	2017年 1月 19日
				US	2017257470	A1	2017年 9月 7日
CN	101557651	A	2009年 10月 14日	US	2009254351	A1	2009年 10月 8日
				JP	2009252238	A	2009年 10月 29日
				EP	2109298	A1	2009年 10月 14日
				RU	2008126782	A	2010年 1月 10日
				KR	20090107365	A	2009年 10月 13日
				US	2012130712	A1	2012年 5月 24日
JP	1021254	A	1998年 1月 23日	无			
CN	106356056	A	2017年 1月 25日	无			