

(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 6 月 18 日 (18.06.2020)



(10) 国际公布号
WO 2020/119496 A1

(51) 国际专利分类号:
G06K 9/00 (2006.01) **G10L 15/26** (2006.01)

(21) 国际申请号: PCT/CN2019/122329

(22) 国际申请日: 2019 年 12 月 2 日 (02.12.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201811539391.9 2018 年 12 月 14 日 (14.12.2018) CN

(71) 申请人: 深圳壹账通智能科技有限公司(ONE CONNECT SMART TECHNOLOGY CO., LTD. (SHENZHEN)) [CN/CN]; 中国广东省深圳市前海深港合作区前湾一路 1 号 A 栋 201 室, Guangdong 518000 (CN)。

(72) 发明人: 温桂龙 (WEN, Guilong); 中国广东省深圳市前海深港合作区前湾一路 1 号 A 栋 201 室, Guangdong 518000 (CN)。

(74) 代理人: 深圳市世纪恒程知识产权代理事务所 (CENFO INTELLECTUAL PROPERTY AGENCY); 中国广东省深圳市南山区粤海街道高新技术产业园北区松坪山路 3 号奥特讯电力大厦 201, Guangdong 518057 (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, 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,

(54) **Title:** COMMUNICATION METHOD, DEVICE AND EQUIPMENT BASED ON ARTIFICIAL INTELLIGENCE AND READABLE STORAGE MEDIUM

(54) 发明名称: 基于人工智能的通信方法、装置、设备及可读存储介质

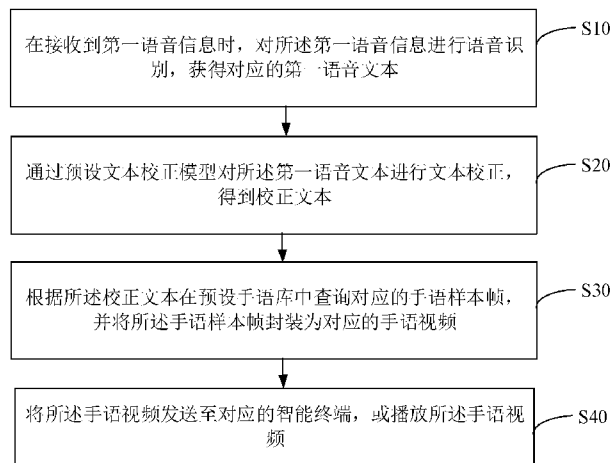


图 2

S10 When receiving first voice information, performing voice recognition to the first voice information to obtain corresponding first voice text

S20 Correcting the first voice text by a preset text correction module to obtain a corrected text

S30 Inquiring a corresponding sign language sample frame in the preset sign language library according to the corrected text, and encapsulating the sign language sample frame into a corresponding sign language video

S40 Sending the sign language video to a corresponding intelligent terminal, or playing the sign language video

(57) **Abstract:** Disclosed are a communication method, device and equipment based on artificial intelligence and a readable storage medium. A translation from voice to sign language or vice versa is realized by artificial intelligence such that deaf-mute users can learn information sent by voice of normal users in social applications, thereby providing convenience for deaf-mute users and improving their use experience; normal users can also conveniently learn the meaning of sign language used by deaf-mute users. Therefore, the method provides convenience both to normal users and to deaf-mute users. Furthermore, during the translation from voice to sign language, the voice text corresponding to voice information is corrected to obtain a corrected text conforming to certain grammar rules or common language habits, the corrected text is then translated into sign language so as to avoid adverse influence to the translation accuracy of sign language video caused by incoherent statement or wrongly written characters, thereby facilitating deaf-mute users to learn the correct meaning of voice information.

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))。

(57) 摘要: 一种基于人工智能的通信方法、设备及可读存储介质, 通过人工智能的方式实现语音与手语之间的转换, 使得聋哑用户也可获知正常用户以语音方式在社交应用发送的信息, 为聋哑用户提供了方便, 提高聋哑用户的使用体验, 而对正常用户而言, 也可方便地获知聋哑用户的手语含义, 为双方提供了方便; 其次, 在语音向手语的转化过程中, 还将对语音信息对应的语音文本进行校正处理, 以得到符合一定语法规则或通用语言习惯的校正文本, 再对该校正文本进行手语转换处理, 从而避免语句不通顺、错别字等情况对手语视频的转换准确度造成不利的影响, 有利于聋哑用户获知语音信息的正确含义。

基于人工智能的通信方法、装置、设备及可读存储介质

本申请要求于2018年12月14日提交中国专利局、申请号为201811539391.9、发明名称为“基于人工智能的通信方法、设备及可读存储介质”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

本申请涉及人工智能领域，尤其涉及一种基于人工智能的通信方法、装置、设备及可读存储介质。

背景技术

随着终端技术的不断发展，社交应用在人们日常生活中的使用越来越频繁，人们可以在智能终端的社交应用中通过发送语音信息的方式来表达自己想要描述的内容，极大的便利了人们之间的沟通。

但是，由于聋哑人语音方面具有一定的缺陷，其通常是以手语与外界进行沟通，因此目前社交应用的语音沟通功能并不利于聋哑人与正常人之间的交流和使用，正常人在不经过系统学习的情况下也无法获取聋哑人的手语所要表达的含义，这就为双方带来了不便。

发明内容

本申请的主要目的在于提供一种基于人工智能的通信方法、设备及可读存储介质，旨在解决聋哑人和正常人无法通过语音功能进行社交通信的技术问题。

为实现上述目的，本申请提供一种基于人工智能的通信方法，所述基于人工智能的通信方法包括：

在接收到第一语音信息时，对所述第一语音信息进行语音识别，获得对应的第一语音文本；

基于预设中文分词规则对所述第一语音文本的原始语句进行词划分，得到对应的语句词序列；

获取所述语句词序列中的目标词对应的混淆词，并通过所述混淆词替换所述语句词序列中的目标词，得到候选序列；

基于预设文法语言模型分别计算所述语句词序列和所述候选序列的文法得分，并根据

所述文法得分在所述语句词序列和候选序列中确定校正序列；

根据所述校正序列获得对应的校正文本；

根据所述校正文本在预设手语库中查询对应的手语样本帧，并将所述手语样本帧封装为对应的手语视频；

将所述手语视频发送至对应的智能终端，或播放所述手语视频。

此外，为实现上述目的，本申请还提供一种基于人工智能的通信方法，所述基于人工智能的通信方法包括：

在接收到手语视频时，对所述各视频帧进行预处理，得到待提取图像；

对所述待提取图像进行实例区域识别，并根据识别结果对所述待提取图像进行图像分割，得到手势区域图像；

对所述手势区域图像进行特征提取，获得对应的手语特征；

根据所述手语特征在预设手语库查询与所述各视频帧对应的手语样本帧，并根据所述手语样本帧确定所述各视频帧的视频帧语义；

根据所述各视频帧的视频帧语义获取所述手语视频的手语视频语义，并根据所述手语视频语义生成对应的第二语音/文本信息；

将所述第二语音/文本信息发送至对应的智能终端，或播放/显示所述第二语音/文本信息。

此外，为实现上述目的，本申请还提供一种基于人工智能的通信装置，所述基于人工智能的通信装置包括：

语音识别模块，用于在接收到第一语音信息时，对所述第一语音信息进行语音识别，获得对应的第一语音文本；

文本校正模块，包括语句划分单元、目标词替换单元、得分计算单元和文本获得单元，其中，所述语句划分单元，用于基于预设中文分词规则对所述第一语音文本的原始语句进行词划分，得到对应的语句词序列；

所述目标词替换单元，用于获取所述语句词序列中的目标词对应的混淆词，并通过所述混淆词替换所述语句词序列中的目标词，得到候选序列；

所述得分计算单元，用于基于预设文法语言模型分别计算所述语句词序列和所述候选序列的文法得分，并根据所述文法得分在所述语句词序列和候选序列中确定校正序列；

所述文本获得单元，用于文本根据所述校正序列获得对应的校正文本；

视频封装模块，用于根据所述校正文本在预设手语库中查询对应的手语样本帧，并将所述手语样本帧封装为对应的手语视频；

视频处理模块，用于将所述手语视频发送至对应的智能终端，或播放所述手语视频。

此外，为实现上述目的，本申请还提供一种基于人工智能的通信装置，所述基于人工智能的通信装置包括：

特征提取模块，包括预处理单元、图像分割单元和特征提取单元，

其中，所述预处理单元，用于对所述各视频帧进行预处理，得到待提取图像；

所述图像分割单元，用于对所述待提取图像进行实例区域识别，并根据识别结果对所述待提取图像进行图像分割，得到手势区域图像；

所述特征提取单元，用于对所述手势区域图像进行特征提取，获得对应的手语特征；

语义确定模块，用于根据所述手语特征在预设手语库查询与所述各视频帧对应的手语样本帧，并根据所述手语样本帧确定所述各视频帧的视频帧语义；

信息生成模块，用于根据所述各视频帧的视频帧语义获取所述手语视频的手语视频语义，并根据所述手语视频语义生成对应的第二语音/文本信息；

信息处理模块，用于将所述第二语音/文本信息发送至对应的智能终端，或播放/显示所述第二语音/文本信息。

此外，为实现上述目的，本申请还提供一种基于人工智能的通信设备，所述基于人工智能的通信设备包括处理器、存储器、以及存储在所述存储器上并可被所述处理器执行的计算机可读指令，其中所述计算机可读指令被所述处理器执行时，实现如上述的基于人工智能的通信方法的步骤。

此外，为实现上述目的，本申请还提供一种可读存储介质，所述可读存储介质上存储有计算机可读指令，其中所述计算机可读指令被处理器执行时，实现如上述的基于人工智能的通信方法的步骤。

本申请的一个或多个实施例的细节在下面的附图和描述中提出。本申请的其他特征和优点将从说明书、附图以及权利要求书变得明显。

附图说明

图1为本申请实施例方案中涉及的基于基于人工智能的通信设备的硬件结构示意图；

图2为本申请基于人工智能的通信方法第一实施例的流程示意图。

本申请目的的实现、功能特点及优点将结合实施例，参照附图做进一步说明。

具体实施方式

应当理解，此处所描述的具体实施例仅仅用以解释本申请，并不用于限定本申请。

本申请实施例涉及的基于人工智能的通信方法主要应用于基于人工智能的通信设备，该通信设备可以是个人计算机（personal computer, PC）、笔记本电脑、移动终端（如平板电脑、掌上电脑、手机）等具有数据处理功能的设备。

参照图 1，图 1 为本申请实施例方案中涉及的基于人工智能的通信设备的硬件结构示意图。本申请实施例中，该通信设备可以包括处理器 1001（例如中央处理器 Central Processing Unit, CPU），通信总线 1002，用户接口 1003，网络接口 1004，存储器 1005。其中，通信总线 1002 用于实现这些组件之间的连接通信；用户接口 1003 可以包括显示屏（Display）、输入单元比如键盘（Keyboard）；网络接口 1004 可选的可以包括标准的有线接口、无线接口（如无线保真 Wireless-Fidelity, WI-FI 接口）；存储器 1005 可以是高速随机存取存储器（random access memory, RAM），也可以是稳定的存储器（non-volatile memory），例如磁盘存储器，存储器 1005 可选的还可以是独立于前述处理器 1001 的存储装置。本领域技术人员可以理解，图 1 中示出的硬件结构并不构成对本申请的限定，可以包括比图示更多或更少的部件，或者组合某些部件，或者不同的部件布置。作为一种计算机可读存储介质的存储器 1005 可以包括操作系统、网络通信模块以及计算机可读指令；网络通信模块可用于连接服务器（或基站），与服务器（或基站）进行数据通信；而处理器 1001 可以调用存储器 1005 中存储的计算机可读指令，并执行本申请实施例提供的基于人工智能的通信方法。

本申请实施例提供了一种基于人工智能的通信方法。

参照图 2，图 2 为本申请基于人工智能的通信方法第一实施例的流程示意图。

本实施例中，所述基于人工智能的通信方法包括以下步骤：

步骤 S10，在接收到第一语音信息时，对所述第一语音信息进行语音识别，获得第一语音文本；

本实施例中的基于人工智能的通信方法是由基于人工智能的通信设备实现的，该通信设备可以是个人计算机（personal computer, PC）、笔记本电脑、移动终端（如平板电脑、掌上电脑、手机）等。本实施例中该通信设备以手机为例进行说明；该手机中预先安装有若干社交应用（如即时通讯应用），该手机的用户为一正常人，可称为正常用户；当需要通过该社交应用与某一聋哑用户进行交流时，可以启动该手机中的社交应用，并通过语音的方式进行录入自己的第一语音信息。而对与聋哑用户而言，若将该第一语音信息发送至聋哑用户所用的智能终端（如手机），由于聋哑用户的听力缺陷，可能会出现无法获知该第一语音信息的真实含义；对此，正常用户的手机在接收到该语音信息时，可对该语音信息进行转换，将其转换为对应的手语视频，再将该手语视频发送至聋哑用户的智能终端。

具体的，手机在接收到该第一语音信息时，首先会对该第一语音信息进行分析和识别，提取该语音信息中的语音特征，并根据该语音特征在预设语音库中进行特征查找和匹配，获取到对应的语音文本，并根据该语音文本获取对应的第一文本信息。其中，对第一语音信息的识别和第一文本信息的生成过程，可以通过预先建立的语音识别模型实现，也即先收集一定数量的语音样本和这些语音样本对应的文本样本，并建立初始语音识别模型，然后通过这些语音样本和文本样本对初始语音识别模型进行迭代训练，不断调整模型参数，直至得到最优模型参数（或是满足预设次数的迭代训练次数、或是识别正确率达到一定阈值），从而根据该最优模型参数得到训练好的语音识别模型。当然，在实际中，对于该语音信息向文本信息的转换过程，也可以是通过第三方提供的 API 接口（Application Programming Interface，应用程序编程接口）、SDK（Software Development Kit，软件开发工具包）等实现的，例如手机在获取到第一语音信息时，可以调用第三方提供的 API 接口或 SDK，并通过该 API 接口或 SDK 调用对应应用程序或工具包对第一语音信息进行语音识别，获得对应的第一语音文本。

步骤 S20，通过预设文本校正模型对所述第一语音文本进行文本校正，得到校正文本；

本实施例中，考虑到部分用户由于口语习惯的原因，可能会出现第一语音信息的语句构成与标准语句不完全匹配的情况（句子不通顺），而在语音信息转化为文本信息的过程中，也可能会出现误匹配、误转化，进而导致出现错别字的情况；这些情况会对手语视频的转换准确度造成不利的影响，甚至可能会出现无法识别和转换的情况。对此，本实施例的手机得到在对第一语音文本时，将通过一预设的文本校正模型对该第一语音文本进行文本校正处理，以得到符合一定语法规则（或通用语言习惯）的校正文本，再对该校正文本进行手语转换处理。

具体的，通过预设文本校正模型对第一语音文本进行文本校正，得到校正文本的步骤包括：

基于预设中文分词规则对第一语音文本的原始语句进行词划分，得到对应的语句词序列；

在文本校正过程中，首先需要根据预设中文分词规则对第一语音文本中的原始语句进行词划分，也将一个连续的句子划分成若干个单独的词，这些若干个单独词由于是按照一定的顺序排序，也即划分后得到的是一个语句词序列。在该语句词序列中需要纠正处理的词可称为目标词，当然在实际中该目标词可能具有多个。而对于该预设中文分词规则，可以基于字符串匹配的方式设置，也即预先存储一词典，该词典中包括有若干标准字符串，然后通过标准字符串匹配的方式进行分词；该预设中文分词规则还可以是基于统计分词的

方式设置，也即在收集大量分词样本的基础上，通过统计类机器学习模型学习词语切分的规律（即训练），得到分词模型，再用该分词模型进行分词。当然，还可以是接入第三方的分词工具进行分词（如 SnowNLP、THULAC、jieba 分词等）。

获取所述语句词序列中的目标词对应的混淆词，并通过所述混淆词替换所述语句词序列中的目标词，得到候选序列；

在得到语句词序列时，将确定该语句词序列中的目标词，并获取目标词的混淆词。其中，该混淆词可包括一些发音相似的汉字，如汉语同音字（“字”和“自”）、鼻音（“藏”）和非鼻音（“赞”）、卷舌（“找”和非卷舌“早”等；还可包括形似的汉字，例如汉字的同一字根（“固”和“回”）和类似的五笔输入法（“丐”和“丐”）；此外，还可包括其它类型（如基于统计）的汉字（如“妻”和“西”、“嫂”和“搜”等）。值得说明的是，对于某个目标词的混淆词，可能具有两个以上（此处“以上”包括本数，下同），也即得到可以混淆集。在得到该目标词的混淆词时，会通过混淆词替换语句词序列中的目标词，得到候选序列；例如，原始语句的语句词序列为（李某/的/是/的确/是/一个/问题），目标词为左起的第一个“是”，对应的混淆词为“事”，则对应的候选序列为（李某/的/事/的确/是/一个/问题）。当然，若混淆词有两个以上，则对应的候选序列也为两个以上。

基于预设文法语言模型分别计算所述语句词序列和所述候选序列的文法得分，并根据所述文法得分在所述语句词序列和候选序列中确定校正序列；

在得到语句词序列和候选序列时，即可根据预设文法语言模型分别计算语句词序列和候选序列的文法得分；该文法得分可认为是该序列所对应的语句在整个中文语言中的出现概率，文法得分越高，其出现概率越大。对于该预设文法语言模型，基于 N-Gram 汉语语言模型的思想，可以为二元（假设一个词的出现概率只取决于其前面的一个单词）或三元（假设一个词的概率取决于其前面的两个词）文法语言模型，该模型结合了马尔科夫假设、最大似然估计和条件概率进行设置。

具体的，该模型中包括有预设语料库；该预设语料库中包括有预先收集的若干样本语句，还包括有根据这些样本语句所统计的单词出现次数表、以及多元词相邻次数表；其中单词出现次数表记录有预设语料库中各样本语句的每个词在整个预设语料库中的出现次数，如下表 1 所示：

表 1 单词出现次数表

词	单词出现次数
我	1215
的	3437
事	938

...	...
总计	13738

而多元词相邻次数表，则记录有多个词在整个预设预料库中的出现次数，例如对于二元词相邻次数表，则记录有两个词以一定顺序相邻出现的次数，如下表 2 所示：

表 2 二元词出现的次数

	我	的	事	...
我	8	786	3	...
的	20	0	988	...
事	121	56	228	...
...	...	...	...	...

对于上述表 2，各数值所表示的是（行，列）顺序的单词相邻出现的此时，例如对于上述第二行第三列的数值“786”，所表示的是“我”和“的”两个词以“我/的”形式在预设预料库中出现次数为 786 次。对于三元词相邻次数表，则记录有三个词以一定顺序相邻出现的次数，以此类推。

在计算语句词序列和候选序列的文法得分时，首先分别需要统计语句词序列和候选序列中各单词在预设语料库中的单词出现次数；对于该单词出现次数的统计，可根据各序列所包括的单词在预设预料库中的单词出现次数表查询得到。其次，还将统计语句词序列中各单词与前位历史词在语料库中的语句词相邻次数，以及候选序列中各单词与前位历史词在预设语料库中的候选词相邻次数；值得说明的是，对于不同的文法语义模型，所需要统计的相邻词的词数可能不同；例如，对于语句词序列（李某/的/是/的确/是/一个/问题），当预设文法语言模型为三元文法语言模型时，所统计的是二元词相邻次数，即需要统计“李某/的”、“的/是”、“是/的确”、“的确/是”、“是/一个”、“一个/问题”的语句词相邻次数；而当预设文法语言模型为三元文法语言模型时，所统计的包括二元词相邻次数和三元词相邻此时，即需要统计“李某/的”、“李某/的/是”、“的/是/的确”、“是/的确/是”、“的确/是/一个”、“是/一个/问题”的语句词相邻次数；以此类推。在得到各单词出现次数、语句词相邻次数和候选词相邻次数时，即可分别计算各序列的文法得分；在计算时，即将语句词序列对应的单词出现次数和语句词相邻次数代入至预设得分公式计算语句词序列的文法得分，并将候选

序列对应的单词出现次数和候选词相邻次数代入预设得分公式计算所述候选序列的文法得分。

可选地，当预设文法语言模型为二元文法语言模型，该预设得分公式为：

$$P = k1 * \frac{C(w_1)}{C_0} * \prod_{i=2}^m \frac{C(w_{i-1}, w_i)}{C(w_{i-1})}$$

其中， P 为语句词序列或候选序列的文法得分；

$k1$ 为第一预设得分系数， $k1 > 0$ ；

m 为语句词序列或候选序列的单词数， $m \geq 2$ ；

C_0 为预设语料库中的单词总数， $C_0 > 0$ ；

$C(w_1)$ 为语句词序列或候选序列中第 1 位单词在预设语料库中的单词出现次数；

$C(w_{i-1})$ 为语句词序列或候选序列中第 $i-1$ 位单词在预设语料库中的单词出现次数， $2 \leq i \leq m$ ；

$C(w_{i-1}, w_i)$ 为语句词序列中第 i 位单词与前一位历史词在预设语料库中的语句词相邻次数，或为候选序列中第 i 位单词与前一位历史词在预设语料库中的候选词相邻次数。

可选地，当预设文法语言模型为三元文法语言模型，该预设得分公式为：

$$P = k2 * \frac{C(w_1)}{C_0} * \frac{C(w_1, w_2)}{C(w_1)} * \prod_{i=3}^m \frac{C(w_{i-2}, w_{i-1}, w_i)}{C(w_{i-2}, w_{i-1})}$$

其中， P 为语句词序列或候选序列的文法得分；

$k2$ 为第二预设得分系数， $k2 > 0$ ；

m 为语句词序列或候选序列的单词数， $m \geq 3$ ；

C_0 为预设语料库中的单词总数， $C_0 > 0$ ；

$C(w_1)$ 为语句词序列或候选序列中第 1 位单词在预设语料库中的单词出现次数；

$C(w_1, w_2)$ 为语句词序列中第 1 位单词与第 2 位单词在预设语料库中的语句词相邻次数，或为候选序列中第 1 位单词与第 2 位单词在预设语料库中的候选词相邻次数；

$C(w_{i-2}, w_{i-1})$ 为语句词序列中第 i 位单词的前一位历史词与前二位历史词在预设语料库中的语句词相邻次数，或为候选序列中第 i 位单词的前一位历史词与前二位历史词在预设语料库中的候选词相邻次数；

$C(w_{i-2}, w_{i-1}, w_i)$ 为语句词序列中第 i 位单词与前一位历史词、前二位历史词在预设语料库中的语句词相邻次数，或为候选序列中第 i 位单词与前一位历史词、前二位历史词在预设语料库中的候选词相邻次数。

在计算得到语句词序列和候选序列的文法得分时，该文法得分可认为是该序列所对应

的语句在整个中文语言中的出现概率，语法得分越高，其出现概率越大，因此可将语法得分最高的序列所对应的序列作为校正序列。当然，在实际中，可能是语句词序列即为校正序列，也可能是候选序列为校正序列。在确定校正序列时，即可根据该校正序列获得对应的校正文本。

步骤 S30，根据所述校正文本在预设手语库中查询对应的手语样本帧，并将所述手语样本帧封装为对应的手语视频；

本实施例中，在得到校正文本时，即可根据该校正文本进行手语视频的转换。具体的，手机可对该校正文本进行语义识别，确定该校正文本的文本语义；当然该文本语义可能是以多段的方式记录的。在得到各段文本语义时，服务器将根据该文本语义在预设手语库中查找与之对应的手语样本帧，然后将各段手语样本帧进行拼接和封装，得到对应的手语视频；其中对于该预设手语库，预先存储有若干语义对应的标准手语样本视频，这些手语样本视频可以是预先录制并存储的。对于该预设手语库，可以是存储在社交应用的服务器或数据库，由手机根据校正文本在服务器中查询得到，从而节约手机的存储空间；当然该预设手语库也可以是存储在手机本地，以方便离线使用。而对于封装的手语视频，则可以根据实际情况（如手机播放器的支持格式或网络条件）选择不同的视频格式（如 mp4、mkv、flv 等）和清晰度（如 480P、720P 等）等。

步骤 S40，将所述手语视频发送至对应的智能终端，或播放所述手语视频。

本实施例中，手机根据校正文本得到对应的手语视频时，即可将该手语视频发送至聋哑用户所用的智能终端（如手机），以使得聋哑用户通过该智能终端看到该手语视频，并获知正常用户所表达的含义。当然，手机在得到该手语视频时，也可以是直接播放该手语视频，以供正常用户查看，并根据正常用户的操作对该手语视频执行发送、存储或是删除操作。

进一步的，本实施例中的语音信息转化手语视频，是以“正常用户要向聋哑用户发送语音消息前进行转换”为例进行说明；而在实际中，还可以是正常用户通过手机直接向聋哑用户的智能终端发送语音信息，当聋哑用户的智能终端接收到该语音信息时，可根据聋哑用户的操作将该语音信息转化为手语视频，其过程与上述实施例的各步骤类似，此处不再赘述，此时，聋哑用户的智能终端也可认为是本实施例中的基于人工智能的通信设备。

再进一步的，本实施例中的语音信息转化为手语视频的过程中，是先将语音信息转化为文本信息；而若正常用户是直接以文本录入的方式录入文本信息，则手机也可直接对该文本进行校正处理，并进一步对校正后的文本执行手语视频的转化，其过程与上述类似，此处不再赘述。

值得说明的是，本实施例中的语音信息向手语视频转化功能，可以是社交应用中的某个功能，也可以是以 API 接口或 SDK 的方式集成在手机等终端中，以供终端中的应用进行调用。

本实施例中，在接收到第一语音信息时，对所述第一语音信息进行语音识别，获得对应的第一语音文本；通过预设文本校正模型对所述第一语音文本进行文本校正，得到校正文本；根据所述校正文本在预设手语库中查询对应的手语样本帧，并将所述手语样本帧封装为对应的手语视频；将所述手语视频发送至对应的智能终端，或播放所述手语视频。通过以上方式，本实施例通过人工智能的方式实现语音向手语的转换，使得聋哑用户也可获知正常用户以语音方式在社交应用发送的信息，为聋哑用户提供了方便，提高聋哑用户的使用体验；其次，在语音转化过程中，还将对语音信息对应的语音文本进行校正处理，以得到符合一定语法规则（或通用语言习惯）的校正文本，再对该校正文本进行手语转换处理，从而避免语句不通顺、错别字等情况对手语视频的转换准确度造成不利的影响，有利于聋哑用户获知语音信息的正确含义。

基于图 2 所示实施例，提出本申请基于人工智能的通信方法第二实施例。本实施例中，所述基于人工智能的通信方法包括以下步骤：

步骤 S50，在接收到手语视频时，基于预设特征提取算法对所述手语视频的各视频帧进行特征提取，获得所述各视频帧的手语特征；

目前，由于聋哑人语音方面具有一定的缺陷，因此通常是以手语与外界进行沟通，在这种情况下聋哑人往往无法直接使用社交应用的语音功能，但对正常人而言在不经系统学习的情况下也无法获取聋哑人的手语所要表达的含义，这就为双方带来了不便。对此，本实施例中提出一种基于人工智能的通信方法，通过人工智能的方式实现手语视频向语音/文本进行转换，使得正常人利用该方法也可通过社交应用与聋哑人进行沟通，为双方提供了方便。

本实施例中的基于人工智能的通信方法是由基于人工智能的通信设备实现的，该通信设备可以是个人计算机（personal computer，PC）、笔记本电脑、移动终端（如平板电脑、掌上电脑、手机）等。本实施例中该通信设备以手机为例进行说明；该手机中预先安装有若干社交应用（如即时通讯应用），该手机的用户为一聋哑人，可称为聋哑用户；当聋哑用户需要通过该社交应用与某一正常用户进行交流时，可以启动该手机中的社交应用，并通过手机录像的方式录入自己的手语视频。手机在接收到该手语视频时，首先基于预设特征提取算法对手语视频的各视频帧进行特征提取，获取到各视频帧的手语特征，例如手势形状、各手指的指向、手部动作的运动轨迹等。对于该特征提取的过程所使用的预设特征

提取算法，可以是依靠机器学习得到的；也即先收集一定数量的手语图像样本，并建立初始特征提取模型，然后通过这些手语图像样本对初始特征提取模型进行迭代训练，不断调整模型参数，直至得到最优模型参数（或是满足预设次数的迭代训练次数、或是识别正确率达到一定阈值），从而根据该最优模型参数得到训练好的特征提取模型，该特征提取模型中的模型逻辑即为预设特征提取算法。当然，在实际中，该预设特征提取算法也可以根据静态规则设置的。

进一步，对于手语而言，其含义主要是靠手势以及手部动作运动轨迹等进行表达，因此在进行特征提取时，也主要是针对手语视频帧中的手部分图像进行。对此，为了提高特征提取的效率，可先对各视频帧进行一定的分割处理，再对分割后的图像进行特征提取。具体的，基于预设特征提取算法对所述手语视频的各视频帧进行特征提取，获得所述各视频帧的手语特征的步骤包括：

对所述各视频帧进行预处理，得到待提取图像；

手机在得到手语视频时，首先可对手语视频的各视频帧进行预处理，得到待提取图像，从而为后续的图像分割处理提供方便，例如图像二值化、降噪等操作。其中，二值化是指将视频帧转化为只含两种黑白颜色的二值图像，这是由于彩色图像所含信息量巨大，对其进行后续处理时计算量也会随之增大，花费的时间也会边长，因此为提高处理和识别的效率，可先对彩色图进行二值化处理；去噪则是通过滤波器或是其它方式解决数字图像由于噪声干扰而导致图像质量下降的问题。当然预处理的内容还可以根据实际情况添加或替换为其它的处理操作。

对所述待提取图像进行实例区域识别，并根据识别结果对所述待提取图像进行图像分割，得到手势区域图像；

在得到待提取图像时，将对该待提取图像进行实例区域识别，识别出其中的手势区域；对于该实例区域识别的过程，可以通过基于待测物的图像中每个像素点的像素值进行，例如通过各像素点的像素值以及临近相似点的像素值确定出像素突变点，进而得出各区域轮廓，然后对区域轮廓进行对象识别，确定各区域轮廓的含义，从而识别出手势区域轮廓；然后即可根据该轮廓进行图像分割，得到手势区域图像。

对所述手势区域图像进行特征提取，获得对应的手语特征。

在得到手势区域图像时，即可对该手势区域图像进行特征提取操作，获得对应的手语特征，从而无需对视频帧中的非手势区域图像进行相关检测和特征提取，降低了运算量，提高了特征提取的效率。

步骤 S60，根据所述手语特征在预设手语库查询与所述各视频帧对应的手语样本帧，

并根据所述手语样本帧确定所述各视频帧的视频帧语义；

本实施例中，手机在提取得到手语特征时，即可根据该手语特征在预设手语库中查询与各视频帧对应的手语样本帧；其中该预设手语库中包括有标准手语样本帧、该手语样本帧的样本语义以及该手语样本帧的样本特征；而该手语样本帧的样本特征，也是通过上述步骤 S50 的预设特征提取算法对手语样本帧提取得到。在查询与各视频帧对应的手语样本帧时，即将视频帧的手语特征与各手语样本帧的样本特征进行比较，当某一样本特征与该手语特征的相似度达到一定阈值时，即可认为该样本特征所属的手语样本帧与该视频帧对应；而对于该相似度，则可以通过距离的方式进行表征（如欧式距离等）。值得说明的是，若存在两个以上的样本特征与某一手语特征的相似度均达到一定阈值，则可将相似度最高的样本特征所属的手语样本帧作为与该视频帧对应的手语样本帧。在查询到与视频帧对应的手语样本帧时，手机获取该手语样本帧的样本含义，并根据该样本含义确定该视频帧的视频帧语音。

步骤 S70，根据所述各视频帧的视频帧语义获取所述手语视频的手语视频语义，并根据所述手语视频语义生成对应的第二语音/文本信息；

本实施例中，在确定各视频帧的视频帧语义时，即可将各视频帧语义进行拼接，从而获得整段手语视频的手语视频语义，并根据该手语视频语义生成对应的第二语音/文本信息。当然，对于是生成语音信息还是文本信息，可以根据聋哑用户的选择进行确定，又或是在手机系统中进行预先设置，又或者是两者均进行生成）。

步骤 S80，将所述第二语音/文本信息发送至对应的智能终端，或播放/显示所述第二语音/文本信息。

本实施例中，手机得到对应的第二语音/文本信息时，即可将该第二语音/文本信息发送至正常用户所用的智能终端（如手机），以使得正常用户通过该智能终端看到该第二语音/文本信息，并获知聋哑用户所表达的含义。当然，手机在得到该第二语音/文本信息时，也可以是播放第二语音信息/显示第二文本信息，并根据聋哑用户的操作对该第二语音/文本信息执行发送、存储或是删除操作。

进一步的，本实施例中的手语视频转化语音/文本信息，是以“聋哑用户要向正常用户发送手语视频前进行转换”为例进行说明；而在实际中，还可以是聋哑用户通过手机直接向正常用户的智能终端发送手语视频，当正常用户的智能终端接收到该手语视频时，可根据正常用户的操作将该手语视频转化为语音信息/文本信息，其过程与上述实施例的各步骤类似，此处不再赘述，此时，正常用户的智能终端也可认为是本实施例中的基于人工智能的通信设备。再进一步，本实施例中的手语视频向语音/文本信息转化功能，可以是社交应用

中的某个功能，也可以是以 API 接口或 SDK 的方式集成在手机等终端中，以供终端中的应用进行调用。

本实施例中，在接收到手语视频时，基于预设特征提取算法对所述手语视频的各视频帧进行特征提取，获得所述各视频帧的手语特征；根据所述手语特征在预设手语库查询与所述各视频帧对应的手语样本帧，并根据所述手语样本帧确定所述各视频帧的视频帧语义；根据所述各视频帧的视频帧语义获取所述手语视频的手语视频语义，并根据所述手语视频语义生成对应的第二语音/文本信息；将所述第二语音/文本信息发送至对应的智能终端，或播放/显示所述第二语音/文本信息。通过以上方式，本实施例通过人工智能的方式实现手语视频向语音/文本的转换，使得正常用户可以方便地获知聋哑用户的手语含义，而聋哑用户也可以以手语表达的方式与正常用户进行社交通信，为双方提供了方便，提高了双方的使用体验。

值得说明的是，对于上述第一实施例的语音向手语的转换功能和第二实施例的手语向语音/文字的转换功能，可以是同时集成在同一基于人工智能的通信设备中，根据用户的选择和设置确定是否启动其中的某一功能，从而为用户的使用提供方便。

此外，本申请实施例还提供一种基于人工智能的通信装置，所述基于人工智能的通信装置包括：

语音识别模块，用于在接收到第一语音信息时，对所述第一语音信息进行语音识别，获得对应的第一语音文本；

文本校正模块，用于通过预设文本校正模型对所述第一语音文本进行文本校正，得到校正文本；

视频封装模块，用于根据所述校正文本在预设手语库中查询对应的手语样本帧，并将所述手语样本帧封装为对应的手语视频；

视频处理模块，用于将所述手语视频发送至对应的智能终端，或播放所述手语视频。

其中，上述基于人工智能的通信装置各虚拟功能模块存储于图 1 所示基于人工智能的通信设备的存储器 1005 中，用于实现计算机可读指令的所有功能；各模块被处理器 1001 执行时，可实现语音向手语之间的转换和发送的功能。

进一步的，所述文本校正模块包括：

语句划分单元，用于基于预设中文分词规则对所述第一语音文本的原始语句进行词划分，得到对应的语句词序列；

目标词替换单元，用于获取所述语句词序列中的目标词对应的混淆词，并通过所述混淆词替换所述语句词序列中的目标词，得到候选序列；

得分计算单元,用于基于预设文法语言模型分别计算所述语句词序列和所述候选序列的文法得分,并根据所述文法得分在所述语句词序列和候选序列中确定校正序列;

文本获得单元,用于文本根据所述校正序列获得对应的校正文本。

进一步的,所述得分计算单元包括:

第一统计子单元,用于统计所述语句词序列和所述候选序列中各单词在预设语料库中的单词出现次数;

第二统计子单元,用于统计所述语句词序列中各单词与前位历史词在预设语料库中的语句词相邻次数,并统计所述候选序列中各单词与前位历史词在所述预设语料库中的候选词相邻次数;

得分计算子单元,用于将所述单词出现次数和所述语句词相邻次数代入至预设得分公式计算所述语句词序列的文法得分,并将所述单词出现次数和所述候选词相邻次数代入至预设得分公式计算所述候选序列的文法得分。

进一步的,所述预设文法语言模型为二元文法语言模型,所述预设得分公式为:

$$P = k1 * \frac{C(w_1)}{C_0} * \prod_{i=2}^m \frac{C(w_{i-1}, w_i)}{C(w_{i-1})}$$

其中, P 为所述语句词序列或所述候选序列的文法得分;

$k1$ 为第一预设得分系数, $k1 > 0$;

m 为所述语句词序列或所述候选序列的单词数, $m \geq 2$;

C_0 为所述预设语料库中的单词总数, $C_0 > 0$;

$C(w_1)$ 为所述语句词序列或所述候选序列中第 1 位单词在所述预设语料库中的单词出现次数;

$C(w_{i-1})$ 为所述语句词序列或所述候选序列中第 $i-1$ 位单词在所述预设语料库中的单词出现次数, $2 \leq i \leq m$;

$C(w_{i-1}, w_i)$ 为所述语句词序列中第 i 位单词与前一位历史词在所述预设语料库中的语句词相邻次数, 或为所述候选序列中第 i 位单词与前一位历史词在所述预设语料库中的候选词相邻次数。

进一步的,所述预设文法语言模型为三元文法语言模型,所述预设得分公式为:

$$P = k2 * \frac{C(w_1)}{C_0} * \frac{C(w_1, w_2)}{C(w_1)} * \prod_{i=3}^m \frac{C(w_{i-2}, w_{i-1}, w_i)}{C(w_{i-2}, w_{i-1})}$$

其中, P 为所述语句词序列或所述候选序列的文法得分;

$k2$ 为第二预设得分系数, $k2 > 0$;

m 为所述语句词序列或所述候选序列的单词数, $m \geq 3$;

C_0 为所述预设语料库中的单词总数, $C_0 > 0$;

$C(w_1)$ 为所述语句词序列或所述候选序列中第 1 位单词在所述预设语料库中的单词出现次数;

$C(w_1, w_2)$ 为所述语句词序列中第 1 位单词与第 2 位单词在所述预设语料库中的语句词相邻次数, 或为所述候选序列中第 1 位单词与第 2 位单词在所述预设语料库中的候选词相邻次数;

$C(w_{i-2}, w_{i-1})$ 为所述语句词序列中第 i 位单词的前一位历史词与前二位历史词在所述预设语料库中的语句词相邻次数, 或为所述候选序列中第 i 位单词的前一位历史词与前二位历史词在所述预设语料库中的候选词相邻次数;

$C(w_{i-2}, w_{i-1}, w_i)$ 为所述语句词序列中第 i 位单词与前一位历史词、前二位历史词在所述预设语料库中的语句词相邻次数, 或为所述候选序列中第 i 位单词与前一位历史词、前二位历史词在所述预设语料库中的候选词相邻次数。

其中, 上述基于人工智能的通信装置中各个模块的功能实现与上述基于人工智能的通信方法第一实施例中各步骤相对应, 其功能和实现过程在此处不再一一赘述。

此外, 本申请实施例还提供一种基于人工智能的通信装置, 所述基于人工智能的通信装置包括:

特征提取模块, 在接收到手语视频时, 基于预设特征提取算法对所述手语视频的各视频帧进行特征提取, 获得所述各视频帧的手语特征;

语义确定模块, 用于根据所述手语特征在预设手语库查询与所述各视频帧对应的手语样本帧, 并根据所述手语样本帧确定所述各视频帧的视频帧语义;

信息生成模块, 用于根据所述各视频帧的视频帧语义获取所述手语视频的手语视频语义, 并根据所述手语视频语义生成对应的第二语音/文本信息;

信息处理模块, 用于将所述第二语音/文本信息发送至对应的智能终端, 或播放/显示所述第二语音/文本信息。

其中, 上述基于人工智能的通信装置各虚拟功能模块也可存储于图 1 所示基于人工智能的通信设备的存储器 1005 中, 用于实现计算机可读指令的所有功能; 各模块被处理器 1001 执行时, 可实现手语向语音/文本之间的转换和发送的功能。

进一步的, 所述特征提取模块包括:

预处理单元, 用于对所述各视频帧进行预处理, 得到待提取图像;

图像分割单元, 用于对所述待提取图像进行实例区域识别, 并根据识别结果对所述待

提取图像进行图像分割，得到手势区域图像；

特征提取单元，用于对所述手势区域图像进行特征提取，获得对应的手语特征。

其中，上述基于人工智能的通信装置中各个模块的功能实现与上述基于人工智能的通信方法第二实施例中各步骤相对应，其功能和实现过程在此处不再一一赘述。

此外，本申请实施例还提供一种可读存储介质，所述计算机可读存储介质可以为非易失性可读存储介质。本申请可读存储介质上存储有计算机可读指令，其中所述计算机可读指令被处理器执行时，实现如上述的基于人工智能的通信方法第一实施例或第二实施例的步骤。

其中，计算机可读指令被执行时所实现的方法可参照上述基于人工智能的通信方法各实施例，此处不再赘述。

以上仅为本申请的优选实施例，并非因此限制本申请的专利范围，凡是利用本申请说明书及附图内容所作的等效结构或等效流程变换，或直接或间接运用在其他相关的技术领域，均同理包括在本申请的专利保护范围内。

权 利 要 求 书

1、一种基于人工智能的通信方法，其中，所述基于人工智能的通信方法包括：

在接收到第一语音信息时，对所述第一语音信息进行语音识别，获得对应的第一语音文本；

基于预设中文分词规则对所述第一语音文本的原始语句进行词划分，得到对应的语句词序列；

获取所述语句词序列中的目标词对应的混淆词，并通过所述混淆词替换所述语句词序列中的目标词，得到候选序列；

基于预设文法语言模型分别计算所述语句词序列和所述候选序列的文法得分，并根据所述文法得分在所述语句词序列和候选序列中确定校正序列；

根据所述校正序列获得对应的校正文本；

根据所述校正文本在预设手语库中查询对应的手语样本帧，并将所述手语样本帧封装为对应的手语视频；

将所述手语视频发送至对应的智能终端，或播放所述手语视频。

2、如权利要求1所述的基于人工智能的通信方法，其中，所述基于预设文法语言模型分别计算所述语句词序列和所述候选序列的文法得分的步骤包括：

统计所述语句词序列和所述候选序列中各单词在预设语料库中的单词出现次数；

统计所述语句词序列中各单词与前位历史词在预设语料库中的语句词相邻次数，并统计所述候选序列中各单词与前位历史词在所述预设语料库中的候选词相邻次数；

将所述单词出现次数和所述语句词相邻次数代入至预设得分公式计算所述语句词序列的文法得分，并将所述单词出现次数和所述候选词相邻次数代入至预设得分公式计算所述候选序列的文法得分。

3、如权利要求2所述的基于人工智能的通信方法，其中，所述预设文法语言模型为二元文法语言模型，所述预设得分公式为：

$$P = k1 * \frac{C(w_1)}{C_0} * \prod_{i=2}^m \frac{C(w_{i-1}, w_i)}{C(w_{i-1})}$$

其中， P 为所述语句词序列或所述候选序列的文法得分；

$k1$ 为第一预设得分系数， $k1 > 0$ ；

m 为所述语句词序列或所述候选序列的单词数， $m \geq 2$ ；

C_0 为所述预设语料库中的单词总数， $C_0 > 0$ ；

$C(w_1)$ 为所述语句词序列或所述候选序列中第 1 位单词在所述预设语料库中的单词出现次数；

$C(w_{i-1})$ 为所述语句词序列或所述候选序列中第 $i-1$ 位单词在所述预设语料库中的单词出现次数， $2 \leq i \leq m$ ；

$C(w_{i-1}, w_i)$ 为所述语句词序列中第 i 位单词与前一位历史词在所述预设语料库中的语句词相邻次数，或为所述候选序列中第 i 位单词与前一位历史词在所述预设语料库中的候选词相邻次数。

4、如权利要求 2 所述的基于人工智能的通信方法，其中，所述预设文法语言模型为三元文法语言模型，所述预设得分公式为：

$$P = k2 * \frac{C(w_1)}{C_0} * \frac{C(w_1, w_2)}{C(w_1)} * \prod_{i=3}^m \frac{C(w_{i-2}, w_{i-1}, w_i)}{C(w_{i-2}, w_{i-1})}$$

其中， P 为所述语句词序列或所述候选序列的文法得分；

$k2$ 为第二预设得分系数， $k2 > 0$ ；

m 为所述语句词序列或所述候选序列的单词数， $m \geq 3$ ；

C_0 为所述预设语料库中的单词总数， $C_0 > 0$ ；

$C(w_1)$ 为所述语句词序列或所述候选序列中第 1 位单词在所述预设语料库中的单词出现次数；

$C(w_1, w_2)$ 为所述语句词序列中第 1 位单词与第 2 位单词在所述预设语料库中的语句词相邻次数，或为所述候选序列中第 1 位单词与第 2 位单词在所述预设语料库中的候选词相邻次数；

$C(w_{i-2}, w_{i-1})$ 为所述语句词序列中第 i 位单词的前一位历史词与前二位历史词在所述预设语料库中的语句词相邻次数，或为所述候选序列中第 i 位单词的前一位历史词与前二位历史词在所述预设语料库中的候选词相邻次数；

$C(w_{i-2}, w_{i-1}, w_i)$ 为所述语句词序列中第 i 位单词与前一位历史词、前二位历史词在所述预设语料库中的语句词相邻次数，或为所述候选序列中第 i 位单词与前一位历史词、前二位历史词在所述预设语料库中的候选词相邻次数。

5、一种基于人工智能的通信方法，其中，所述基于人工智能的通信方法包括：

在接收到手语视频时，对所述各视频帧进行预处理，得到待提取图像；

对所述待提取图像进行实例区域识别，并根据识别结果对所述待提取图像进行图像分割，得到手势区域图像；

对所述手势区域图像进行特征提取，获得对应的手语特征；

根据所述手语特征在预设手语库查询与所述各视频帧对应的手语样本帧,并根据所述手语样本帧确定所述各视频帧的视频帧语义;

根据所述各视频帧的视频帧语义获取所述手语视频的手语视频语义,并根据所述手语视频语义生成对应的第二语音/文本信息;

将所述第二语音/文本信息发送至对应的智能终端,或播放/显示所述第二语音/文本信息。

6、一种基于人工智能的通信装置,其中,所述基于人工智能的通信装置包括:

语音识别模块,用于在接收到第一语音信息时,对所述第一语音信息进行语音识别,获得对应的第一语音文本;

文本校正模块,包括语句划分单元、目标词替换单元、得分计算单元和文本获得单元,其中,所述语句划分单元,用于基于预设中文分词规则对所述第一语音文本的原始语句进行词划分,得到对应的语句词序列;

所述目标词替换单元,用于获取所述语句词序列中的目标词对应的混淆词,并通过所述混淆词替换所述语句词序列中的目标词,得到候选序列;

所述得分计算单元,用于基于预设文法语言模型分别计算所述语句词序列和所述候选序列的文法得分,并根据所述文法得分在所述语句词序列和候选序列中确定校正序列;

所述文本获得单元,用于文本根据所述校正序列获得对应的校正文本;

视频封装模块,用于根据所述校正文本在预设手语库中查询对应的手语样本帧,并将所述手语样本帧封装为对应的手语视频;

视频处理模块,用于将所述手语视频发送至对应的智能终端,或播放所述手语视频。

7、如权利要求6所述的基于人工智能的通信装置,其中,所述得分计算单元包括:

第一统计子单元,用于统计所述语句词序列和所述候选序列中各单词在预设语料库中的单词出现次数;

第二统计子单元,用于统计所述语句词序列中各单词与前位历史词在预设语料库中的语句词相邻次数,并统计所述候选序列中各单词与前位历史词在所述预设语料库中的候选词相邻次数;

得分计算子单元,用于将所述单词出现次数和所述语句词相邻次数代入至预设得分公式计算所述语句词序列的文法得分,并将所述单词出现次数和所述候选词相邻次数代入至预设得分公式计算所述候选序列的文法得分。

8、如权利要求7所述的基于人工智能的通信装置,其中,所述预设文法语言模型为二元文法语言模型,所述预设得分公式为:

$$P = k1 * \frac{C(w_1)}{C_0} * \prod_{i=2}^m \frac{C(w_{i-1}, w_i)}{C(w_{i-1})}$$

其中, P 为所述语句词序列或所述候选序列的文法得分;

$k1$ 为第一预设得分系数, $k1 > 0$;

m 为所述语句词序列或所述候选序列的单词数, $m \geq 2$;

C_0 为所述预设语料库中的单词总数, $C_0 > 0$;

$C(w_1)$ 为所述语句词序列或所述候选序列中第 1 位单词在所述预设语料库中的单词出现次数;

$C(w_{i-1})$ 为所述语句词序列或所述候选序列中第 $i-1$ 位单词在所述预设语料库中的单词出现次数, $2 \leq i \leq m$;

$C(w_{i-1}, w_i)$ 为所述语句词序列中第 i 位单词与前一位历史词在所述预设语料库中的语句词相邻次数, 或为所述候选序列中第 i 位单词与前一位历史词在所述预设语料库中的候选词相邻次数。

9、如权利要求 7 所述的基于人工智能的通信装置, 其中, 所述预设文法语言模型为三元文法语言模型, 所述预设得分公式为:

$$P = k2 * \frac{C(w_1)}{C_0} * \frac{C(w_1, w_2)}{C(w_1)} * \prod_{i=3}^m \frac{C(w_{i-2}, w_{i-1}, w_i)}{C(w_{i-2}, w_{i-1})}$$

其中, P 为所述语句词序列或所述候选序列的文法得分;

$k2$ 为第二预设得分系数, $k2 > 0$;

m 为所述语句词序列或所述候选序列的单词数, $m \geq 3$;

C_0 为所述预设语料库中的单词总数, $C_0 > 0$;

$C(w_1)$ 为所述语句词序列或所述候选序列中第 1 位单词在所述预设语料库中的单词出现次数;

$C(w_1, w_2)$ 为所述语句词序列中第 1 位单词与第 2 位单词在所述预设语料库中的语句词相邻次数, 或为所述候选序列中第 1 位单词与第 2 位单词在所述预设语料库中的候选词相邻次数;

$C(w_{i-2}, w_{i-1})$ 为所述语句词序列中第 i 位单词的前一位历史词与前二位历史词在所述预设语料库中的语句词相邻次数, 或为所述候选序列中第 i 位单词的前一位历史词与前二位历史词在所述预设语料库中的候选词相邻次数;

$C(w_{i-2}, w_{i-1}, w_i)$ 为所述语句词序列中第 i 位单词与前一位历史词、前二位历史词在所述预设语料库中的语句词相邻次数, 或为所述候选序列中第 i 位单词与前一位历史词、

前二位历史词在所述预设语料库中的候选词相邻次数。

10、一种基于人工智能的通信装置，其中，所述基于人工智能的通信装置包括：

特征提取模块，包括预处理单元、图像分割单元和特征提取单元，

其中，所述预处理单元，用于对所述各视频帧进行预处理，得到待提取图像；

所述图像分割单元，用于对所述待提取图像进行实例区域识别，并根据识别结果对所述待提取图像进行图像分割，得到手势区域图像；

所述特征提取单元，用于对所述手势区域图像进行特征提取，获得对应的手语特征；

语义确定模块，用于根据所述手语特征在预设手语库查询与所述各视频帧对应的手语样本帧，并根据所述手语样本帧确定所述各视频帧的视频帧语义；

信息生成模块，用于根据所述各视频帧的视频帧语义获取所述手语视频的手语视频语义，并根据所述手语视频语义生成对应的第二语音/文本信息；

信息处理模块，用于将所述第二语音/文本信息发送至对应的智能终端，或播放/显示所述第二语音/文本信息。

11、一种基于人工智能的通信设备，其中，所述基于人工智能的通信设备包括处理器、存储器、以及存储在所述存储器上并可被所述处理器执行的计算机可读指令，其中所述计算机可读指令被所述处理器执行时，实现以下步骤：

在接收到第一语音信息时，对所述第一语音信息进行语音识别，获得对应的第一语音文本；

基于预设中文分词规则对所述第一语音文本的原始语句进行词划分，得到对应的语句词序列；

获取所述语句词序列中的目标词对应的混淆词，并通过所述混淆词替换所述语句词序列中的目标词，得到候选序列；

基于预设文法语言模型分别计算所述语句词序列和所述候选序列的文法得分，并根据所述文法得分在所述语句词序列和候选序列中确定校正序列；

根据所述校正序列获得对应的校正文本；

根据所述校正文本在预设手语库中查询对应的手语样本帧，并将所述手语样本帧封装为对应的手语视频；

将所述手语视频发送至对应的智能终端，或播放所述手语视频。

12、如权利要求 11 所述的基于人工智能的通信设备，其中，所述基于预设文法语言模型分别计算所述语句词序列和所述候选序列的文法得分的步骤包括：

统计所述语句词序列和所述候选序列中各单词在预设语料库中的单词出现次数；

统计所述语句词序列中各单词与前位历史词在预设语料库中的语句词相邻次数,并统计所述候选序列中各单词与前位历史词在所述预设语料库中的候选词相邻次数;

将所述单词出现次数和所述语句词相邻次数代入至预设得分公式计算所述语句词序列的文法得分,并将所述单词出现次数和所述候选词相邻次数代入至预设得分公式计算所述候选序列的文法得分。

13、如权利要求 12 所述的基于人工智能的通信设备,其中,所述预设文法语言模型为二元文法语言模型,所述预设得分公式为:

$$P = k1 * \frac{C(w_1)}{C_0} * \prod_{i=2}^m \frac{C(w_{i-1}, w_i)}{C(w_{i-1})}$$

其中, P 为所述语句词序列或所述候选序列的文法得分;

$k1$ 为第一预设得分系数, $k1 > 0$;

m 为所述语句词序列或所述候选序列的单词数, $m \geq 2$;

C_0 为所述预设语料库中的单词总数, $C_0 > 0$;

$C(w_1)$ 为所述语句词序列或所述候选序列中第 1 位单词在所述预设语料库中的单词出现次数;

$C(w_{i-1})$ 为所述语句词序列或所述候选序列中第 $i-1$ 位单词在所述预设语料库中的单词出现次数, $2 \leq i \leq m$;

$C(w_{i-1}, w_i)$ 为所述语句词序列中第 i 位单词与前一位历史词在所述预设语料库中的语句词相邻次数,或为所述候选序列中第 i 位单词与前一位历史词在所述预设语料库中的候选词相邻次数。

14、如权利要求 12 所述的基于人工智能的通信设备,其中,所述预设文法语言模型为三元文法语言模型,所述预设得分公式为:

$$P = k2 * \frac{C(w_1)}{C_0} * \frac{C(w_1, w_2)}{C(w_1)} * \prod_{i=3}^m \frac{C(w_{i-2}, w_{i-1}, w_i)}{C(w_{i-2}, w_{i-1})}$$

其中, P 为所述语句词序列或所述候选序列的文法得分;

$k2$ 为第二预设得分系数, $k2 > 0$;

m 为所述语句词序列或所述候选序列的单词数, $m \geq 3$;

C_0 为所述预设语料库中的单词总数, $C_0 > 0$;

$C(w_1)$ 为所述语句词序列或所述候选序列中第 1 位单词在所述预设语料库中的单词出现次数;

$C(w_1, w_2)$ 为所述语句词序列中第 1 位单词与第 2 位单词在所述预设语料库中的语

句词相邻次数，或为所述候选序列中第1位单词与第2位单词在所述预设语料库中的候选词相邻次数；

$C(w_{i-2}, w_{i-1})$ 为所述语句词序列中第 i 位单词的前一位历史词与前二位历史词在所述预设语料库中的语句词相邻次数，或为所述候选序列中第 i 位单词的前一位历史词与前二位历史词在所述预设语料库中的候选词相邻次数；

$C(w_{i-2}, w_{i-1}, w_i)$ 为所述语句词序列中第 i 位单词与前一位历史词、前二位历史词在所述预设语料库中的语句词相邻次数，或为所述候选序列中第 i 位单词与前一位历史词、前二位历史词在所述预设语料库中的候选词相邻次数。

15、一种基于人工智能的通信设备，其中，所述基于人工智能的通信设备包括处理器、存储器、以及存储在所述存储器上并可被所述处理器执行的计算机可读指令，其中所述计算机可读指令被所述处理器执行时，实现以下步骤：

在接收到手语视频时，对所述各视频帧进行预处理，得到待提取图像；

对所述待提取图像进行实例区域识别，并根据识别结果对所述待提取图像进行图像分割，得到手势区域图像；

对所述手势区域图像进行特征提取，获得对应的手语特征；

根据所述手语特征在预设手语库查询与所述各视频帧对应的手语样本帧，并根据所述手语样本帧确定所述各视频帧的视频帧语义；

根据所述各视频帧的视频帧语义获取所述手语视频的手语视频语义，并根据所述手语视频语义生成对应的第二语音/文本信息；

将所述第二语音/文本信息发送至对应的智能终端，或播放/显示所述第二语音/文本信息。

16、一种可读存储介质，其中，所述可读存储介质上存储有计算机可读指令，其中所述计算机可读指令被处理器执行时，实现以下步骤：

在接收到第一语音信息时，对所述第一语音信息进行语音识别，获得对应的第一语音文本；

基于预设中文分词规则对所述第一语音文本的原始语句进行词划分，得到对应的语句词序列；

获取所述语句词序列中的目标词对应的混淆词，并通过所述混淆词替换所述语句词序列中的目标词，得到候选序列；

基于预设文法语言模型分别计算所述语句词序列和所述候选序列的文法得分，并根据所述文法得分在所述语句词序列和候选序列中确定校正序列；

根据所述校正序列获得对应的校正文本；

根据所述校正文本在预设手语库中查询对应的手语样本帧，并将所述手语样本帧封装为对应的手语视频；

将所述手语视频发送至对应的智能终端，或播放所述手语视频。

17、如权利要求 16 所述的可读存储介质，其中，所述基于预设文法语言模型分别计算所述语句词序列和所述候选序列的文法得分的步骤包括：

统计所述语句词序列和所述候选序列中各单词在预设语料库中的单词出现次数；

统计所述语句词序列中各单词与前位历史词在预设语料库中的语句词相邻次数，并统计所述候选序列中各单词与前位历史词在所述预设语料库中的候选词相邻次数；

将所述单词出现次数和所述语句词相邻次数代入至预设得分公式计算所述语句词序列的文法得分，并将所述单词出现次数和所述候选词相邻次数代入至预设得分公式计算所述候选序列的文法得分。

18、如权利要求 17 所述的可读存储介质，其中，所述预设文法语言模型为二元文法语言模型，所述预设得分公式为：

$$P = k1 * \frac{C(w_1)}{C_0} * \prod_{i=2}^m \frac{C(w_{i-1}, w_i)}{C(w_{i-1})}$$

其中， P 为所述语句词序列或所述候选序列的文法得分；

$k1$ 为第一预设得分系数， $k1 > 0$ ；

m 为所述语句词序列或所述候选序列的单词数， $m \geq 2$ ；

C_0 为所述预设语料库中的单词总数， $C_0 > 0$ ；

$C(w_1)$ 为所述语句词序列或所述候选序列中第 1 位单词在所述预设语料库中的单词出现次数；

$C(w_{i-1})$ 为所述语句词序列或所述候选序列中第 $i-1$ 位单词在所述预设语料库中的单词出现次数， $2 \leq i \leq m$ ；

$C(w_{i-1}, w_i)$ 为所述语句词序列中第 i 位单词与前一位历史词在所述预设语料库中的语句词相邻次数，或为所述候选序列中第 i 位单词与前一位历史词在所述预设语料库中的候选词相邻次数。

19、如权利要求 17 所述的可读存储介质，其中，所述预设文法语言模型为三元文法语言模型，所述预设得分公式为：

$$P = k2 * \frac{C(w_1)}{C_0} * \frac{C(w_1, w_2)}{C(w_1)} * \prod_{i=3}^m \frac{C(w_{i-2}, w_{i-1}, w_i)}{C(w_{i-2}, w_{i-1})}$$

其中， P 为所述语句词序列或所述候选序列的文法得分；

k_2 为第二预设得分系数, $k_2 > 0$;

m 为所述语句词序列或所述候选序列的单词数, $m \geq 3$;

C_0 为所述预设语料库中的单词总数, $C_0 > 0$;

$C(w_1)$ 为所述语句词序列或所述候选序列中第 1 位单词在所述预设语料库中的单词出现次数;

$C(w_1, w_2)$ 为所述语句词序列中第 1 位单词与第 2 位单词在所述预设语料库中的语句词相邻次数, 或为所述候选序列中第 1 位单词与第 2 位单词在所述预设语料库中的候选词相邻次数;

$C(w_{i-2}, w_{i-1})$ 为所述语句词序列中第 i 位单词的前一位历史词与前二位历史词在所述预设语料库中的语句词相邻次数, 或为所述候选序列中第 i 位单词的前一位历史词与前二位历史词在所述预设语料库中的候选词相邻次数;

$C(w_{i-2}, w_{i-1}, w_i)$ 为所述语句词序列中第 i 位单词与前一位历史词、前二位历史词在所述预设语料库中的语句词相邻次数, 或为所述候选序列中第 i 位单词与前一位历史词、前二位历史词在所述预设语料库中的候选词相邻次数。

20、一种可读存储介质, 其中, 所述可读存储介质上存储有计算机可读指令, 其中所述计算机可读指令被处理器执行时, 实现以下步骤:

在接收到手语视频时, 对所述各视频帧进行预处理, 得到待提取图像;

对所述待提取图像进行实例区域识别, 并根据识别结果对所述待提取图像进行图像分割, 得到手势区域图像;

对所述手势区域图像进行特征提取, 获得对应的手语特征;

根据所述手语特征在预设手语库查询与所述各视频帧对应的手语样本帧, 并根据所述手语样本帧确定所述各视频帧的视频帧语义;

根据所述各视频帧的视频帧语义获取所述手语视频的手语视频语义, 并根据所述手语视频语义生成对应的第二语音/文本信息;

将所述第二语音/文本信息发送至对应的智能终端, 或播放/显示所述第二语音/文本信息。

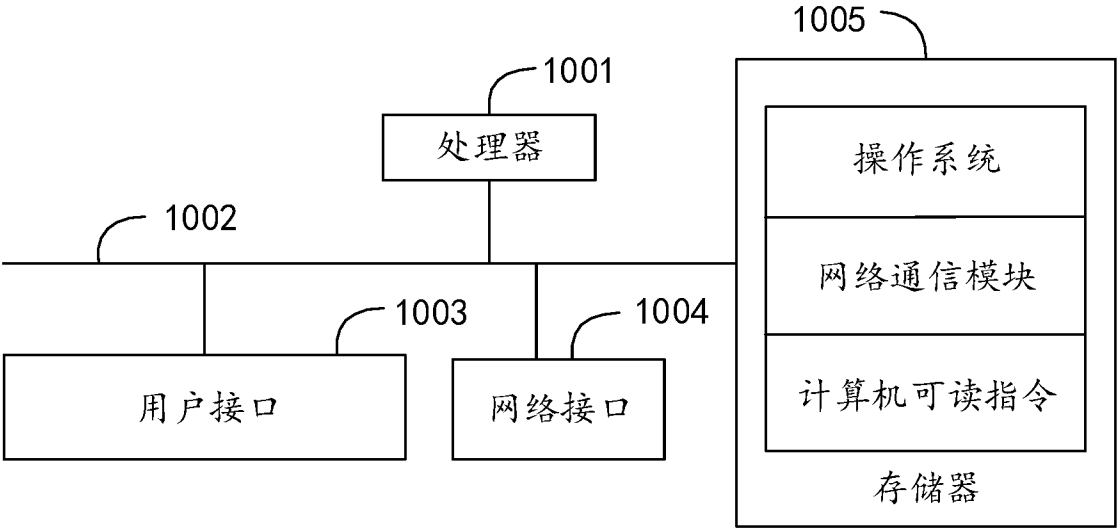


图 1

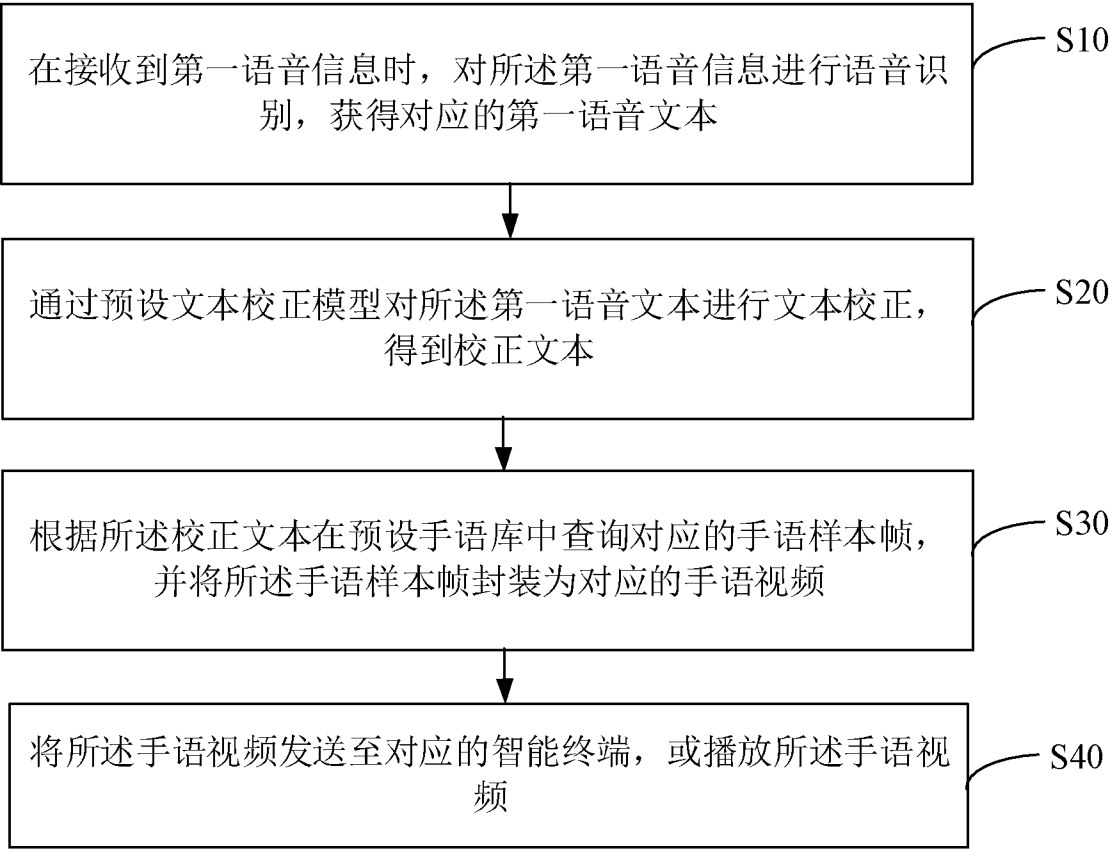


图 2

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/122329

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/00(2006.01)i; G10L 15/26(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)

G06K; 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, WPI, EPODOC, CNKI, IEEE, GOOGLE: 语音, 手语, 聋哑, 校正, 词, 序列, speech, voice, sign, language, deaf, dumb, mute, check, correct, word, order, sequence

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 109740447 A (ONECONNECT FINANCIAL TECHNOLOGY CO., PTE. LTD.) 10 May 2019 (2019-05-10) description, paragraphs [0029]-[0099]	1-20
X	CN 102708866 A (WUHAN UNIVERSITY) 03 October 2012 (2012-10-03) description, paragraph [0023]	5, 10, 15, 20
A	CN 102708866 A (WUHAN UNIVERSITY) 03 October 2012 (2012-10-03) entire document	1-4, 6-9, 11-14, 16-19
A	CN 108615009 A (SHANDONG NORMAL UNIVERSITY) 02 October 2018 (2018-10-02) entire document	1-20
A	US 2005033578 A1 (ZUCKERMAN, M.) 10 February 2005 (2005-02-10) entire document	1-20



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

17 February 2020

Date of mailing of the international search report

25 February 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/CN2019/122329

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	109740447	A	10 May 2019	None	
CN	102708866	A	03 October 2012	None	
CN	108615009	A	02 October 2018	None	
US	2005033578	A1	10 February 2005	None	

国际检索报告

国际申请号

PCT/CN2019/122329

A. 主题的分类

G06K 9/00 (2006.01) i; G10L 15/26 (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)

G06K; 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, WPI, EPDOC, CNKI, IEEE, GOOGLE: 语音, 手语, 聋哑, 校正, 词, 序列, speech, voice, sign, language, deaf, dumb, mute, check, correct, word, order, sequence

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
PX	CN 109740447 A (深圳壹账通智能科技有限公司) 2019年 5月 10日 (2019 - 05 - 10) 说明书第[0029]-[0099]段	1-20
X	CN 102708866 A (武汉大学) 2012年 10月 3日 (2012 - 10 - 03) 说明书第[0023]段	5, 10, 15, 20
A	CN 102708866 A (武汉大学) 2012年 10月 3日 (2012 - 10 - 03) 全文	1-4, 6-9, 11-14, 16-19
A	CN 108615009 A (山东师范大学) 2018年 10月 2日 (2018 - 10 - 02) 全文	1-20
A	US 2005033578 A1 (ZUCKERMAN, Mara) 2005年 2月 10日 (2005 - 02 - 10) 全文	1-20

☐ 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 2月 17日

国际检索报告邮寄日期

2020年 2月 25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

孙韬敏

电话号码 86-(10)-53961420

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/122329

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	109740447	A	2019年 5月 10日	无	
CN	102708866	A	2012年 10月 3日	无	
CN	108615009	A	2018年 10月 2日	无	
US	2005033578	A1	2005年 2月 10日	无	