



(12) 发明专利

(10) 授权公告号 CN 117076702 B

(45) 授权公告日 2023. 12. 15

(21) 申请号 202311184754.2

G06F 16/583 (2019.01)

(22) 申请日 2023.09.14

G06F 16/9536 (2019.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 117076702 A

(56) 对比文件

KR 20210078927 A, 2021.06.29

CN 116431855 A, 2023.07.14

CN 110598037 A, 2019.12.20

US 2022138252 A1, 2022.05.05

CN 110597963 A, 2019.12.20

CN 109034203 A, 2018.12.18

(43) 申请公布日 2023.11.17

(73) 专利权人 荣耀终端有限公司

地址 518040 广东省深圳市福田区香蜜湖
街道东海社区红荔西路8089号深业中
城6号楼A单元3401

审查员 邓力

(72) 发明人 王龙

(74) 专利代理机构 广州三环专利商标代理有限
公司 44202

专利代理师 熊永强

(51) Int. Cl.

G06F 16/532 (2019.01)

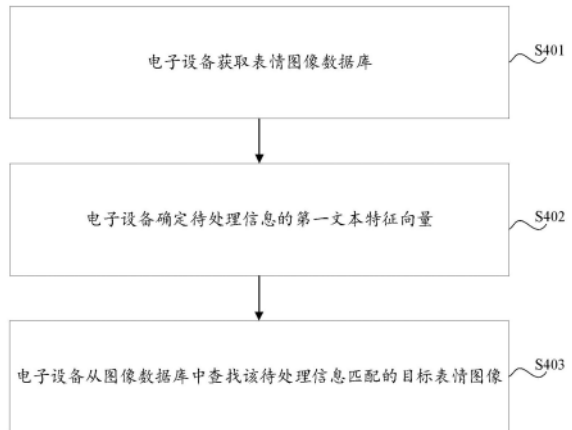
权利要求书2页 说明书22页 附图18页

(54) 发明名称

图像查找方法及电子设备

(57) 摘要

本申请公开了图像查找方法及电子设备,该方法包括:获取表情图像数据库,该表情图像数据库包括多个表情图像、多个表情图像中每个表情图像对应的第一内容特征向量和第一意图特征向量;该第一内容特征向量表示描述表情图像的内容的特征向量,该第一意图特征向量表示描述表情图像的表达意图的特征向量;确定待处理信息的第一文本特征向量;从该图像数据库中查找该待处理信息匹配的目标表情图像,该目标表情图像对应的第一内容特征向量和第一意图特征向量与该第一文本特征向量相匹配。通过本申请描述的方法,能够快速精确地查找到用户需要发送的表情图像,有利于提高线上聊天对话的效率。



1. 一种图像查找方法,其特征在于,所述方法包括:

获取用户在本地收藏的多个表情图像;

调用图像编码器对所述多个表情图像中每个表情图像进行编码处理,得到每个表情图像对应的第一图像特征向量;

调用第一对齐网络对所述第一图像特征向量和第一提示词进行处理,得到每个表情图像对应的第一内容特征向量;所述第一提示词用于提示所述第一对齐网络描述表情图像的内容;

调用第二对齐网络对所述第一图像特征向量和第二提示词进行处理,得到每个表情图像对应的第一意图特征向量;所述第二提示词用于提示所述第二对齐网络描述表情图像的表达意图;

利用所述多个表情图像、所述每个表情图像对应的第一内容特征向量和第一意图特征向量构建表情图像数据库;所述第一内容特征向量表示描述表情图像的内容的特征向量,所述第一意图特征向量表示描述表情图像的表达意图的特征向量;

确定待处理信息的第一文本特征向量;

从所述表情图像数据库中查找所述待处理信息匹配的目标表情图像,所述目标表情图像对应的第一内容特征向量和第一意图特征向量与所述第一文本特征向量相匹配;

其中,从所述表情图像数据库中查找所述待处理信息匹配的目标表情图像,包括:

调用所述第一对齐网络对所述第一文本特征向量进行处理,得到第四内容特征向量;

调用所述第二对齐网络对所述第一文本特征向量进行处理,得到第五意图特征向量;

确定所述表情图像数据库中每个表情图像对应的第一内容特征向量与所述第四内容特征向量之间的第三向量距离,以及确定所述表情图像数据库中每个表情图像对应的第一意图特征向量与所述第五意图特征向量之间的第四向量距离;

基于所述第三向量距离和所述第四向量距离确定所述待处理信息匹配的目标表情图像。

2. 根据权利要求1所述的方法,其特征在于,所述从所述表情图像数据库中查找所述待处理信息匹配的目标表情图像,包括:

确定所述表情图像数据库中每个表情图像对应的第一内容特征向量与所述第一文本特征向量之间的第一向量距离,以及确定所述表情图像数据库中每个表情图像对应的第一意图特征向量与所述第一文本特征向量之间的第二向量距离;

基于所述第一向量距离和所述第二向量距离确定所述待处理信息匹配的目标表情图像。

3. 根据权利要求2所述的方法,其特征在于,所述基于所述第一向量距离和所述第二向量距离确定所述待处理信息匹配的目标表情图像,包括:

对所述第一向量距离和所述第二向量距离进行计算处理,得到每个表情图像对应的第三向量距离;

按照从小到大的顺序排列所述第三向量距离,将前N个第三向量距离对应的表情图像确定为所述待处理信息匹配的目标表情图像;所述N为正整数。

4. 根据权利要求1-3中任一项所述的方法,其特征在于,所述方法还包括:

在第一界面中显示所述待处理信息匹配的目标表情图像,所述第一界面为以下界面中

的任意一项:图像搜索界面、输入法界面、表情图像推荐界面。

5. 根据权利要求1-3中任一项所述的方法,其特征在于,所述方法还包括:

显示表情图像添加界面,所述表情图像添加界面包括表情图像框、自定义标签框和保存选项;

在检测到用户针对所述保存选项的触发操作时,保存所述表情图像框中添加的第一表情图像以及所述自定义标签框中添加的所述第一表情图像的第一标签;

若所述待处理信息与所述第一标签相同,则将所述第一表情图像确定为所述目标表情图像。

6. 一种电子设备,其特征在于,包括:一个或多个处理器,一个或多个存储器;其中,一个或多个存储器与一个或多个处理器耦合,所述一个或多个存储器用于存储计算机程序代码,所述计算机程序代码包括计算机指令,当所述一个或多个处理器在执行所述计算机指令时,使得所述电子设备执行如权利要求1-5中任一项所述的方法。

7. 一种芯片,其特征在于,包括处理器和接口,所述处理器和所述接口耦合;所述接口用于接收或输出信号,所述处理器用于执行代码指令,以使权利要求1-5中任一项所述的方法被执行。

8. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质中存储有计算机程序,所述计算机程序包括程序指令,当所述程序指令在电子设备上运行时,使得所述电子设备执行如权利要求1-5中任一项所述的方法。

9. 一种计算机程序产品,其特征在于,当所述计算机程序产品在计算机上运行时,使得所述计算机执行如权利要求1-5中任一项所述的方法。

图像查找方法及电子设备

技术领域

[0001] 本申请实施例涉及计算机领域,尤其涉及图像查找方法及电子设备。

背景技术

[0002] 随着各种社交应用和新媒体技术的发展,表情包的运用越来越广泛。所谓的表情包也可以称为表情图像,是指一种利用图像来表示感情的方式。当前社交网络中,表情包已成为了一个必不可少的通讯媒介,人们往往更愿意使用表情包来代替文字去传达自己的意图和情感。

[0003] 但是随着人们保存的表情包日益增多,用户查找表情包变得越来越困难且低效,在线上聊天时无法快速精确地找到自己想发送的表情包,从而降低了聊天对话的效率。

发明内容

[0004] 本申请提供了一种图像查找方法及电子设备,能够快速精确地查找到用户需要发送的表情图像,有利于提高线上聊天对话的效率。

[0005] 第一方面,本申请提供了一种图像查找方法,该方法包括:获取表情图像数据库,该表情图像数据库包括多个表情图像、多个表情图像中每个表情图像对应的第一内容特征向量和第一意图特征向量;该第一内容特征向量为描述表情图像的内容的特征向量,该第一意图特征向量表示描述表情图像的表达意图的特征向量;确定待处理信息的第一文本特征向量;从该图像数据库中查找该待处理信息匹配的目标表情图像,该目标表情图像对应的第一内容特征向量和第一意图特征向量与该第一文本特征向量相匹配。

[0006] 基于第一方面所描述的方法,从图像数据库中查找该待处理信息匹配的目标表情图像是通过对表情图像的表达意图和语义内容的多维度快速检索实现的,这样能够更加快速精确地查找到用户需要发送的表情图像,有利于提高线上聊天对话的效率。

[0007] 在一种可能的实现方式中,获取表情图像数据库,包括:获取用户在本地收藏的多个表情图像;调用图像编码器对多个表情图像中每个表情图像进行编码处理,得到每个表情图像对应的第一图像特征向量;调用对齐网络对第一图像特征向量和提示词进行处理,得到每个表情图像对应的第一内容特征向量和第一意图特征向量;该提示词用于提示对齐网络描述表情图像的内容和表达意图;利用多个表情图像、每个表情图像对应的第一内容特征向量和第一意图特征向量构建表情图像数据库。基于该方式,能够从表达意图和语义内容两个维度建立表情图像数据库,有利于提高表情图像的查找效率。

[0008] 在一种可能的实现方式中,该对齐网络包括第一对齐网络和第二对齐网络,该提示词包括第一提示词和第二提示词,调用对齐网络对该第一图像特征向量和提示词进行处理,得到每个表情图像对应的第一内容特征向量和第一意图特征向量,包括:调用该第一对齐网络对第一图像特征向量和第一提示词进行处理,得到每个表情图像对应的第一内容特征向量;该第一提示词用于提示第一对齐网络描述表情图像的内容;调用该第二对齐网络对第一图像特征向量和第二提示词进行处理,得到每个表情图像对应的第一意图特征向

量;该第二提示词用于提示第二对齐网络描述表情图像的表达意图。基于该方式,能够提高每个表情图像对应的第一内容特征向量和第一意图特征向量的准确性。

[0009] 在一种可能的实现方式中,从该图像数据库中查找该待处理信息匹配的目标表情图像,包括:确定该图像数据库中每个表情图像对应的第一内容特征向量与该第一文本特征向量之间的第一向量距离,以及确定该图像数据库中每个表情图像对应的第一意图特征向量与该第一文本特征向量之间的第二向量距离;基于该第一向量距离和该第二向量距离确定该待处理信息匹配的目标表情图像。基于该方式,能够从表达意图和语义内容两个维度查找该待处理信息匹配的目标表情图像,提高表情图像的查找效率。

[0010] 在一种可能的实现方式中,基于该第一向量距离和该第二向量距离确定该待处理信息匹配的目标表情图像,包括:对该第一向量距离和该第二向量距离进行计算处理,得到每个表情图像对应的第三向量距离;按照从小到大的顺序排列第三向量距离,将前N个第三向量距离对应的表情图像确定为该待处理信息匹配的目标表情图像;该N为正整数。基于该方式,能够结合表达意图和语义内容两个维度查找该待处理信息匹配的目标表情图像,提高表情图像的查找效率。

[0011] 在一种可能的实现方式中,从该图像数据库中查找该待处理信息匹配的目标表情图像,包括:调用第一对齐网络对第一文本特征向量进行处理,得到第四内容特征向量;调用第二对齐网络对第一文本特征向量进行处理,得到第五意图特征向量;确定图像数据库中每个表情图像对应的第一内容特征向量与第四内容特征向量之间的第三向量距离,以及确定图像数据库中每个表情图像对应的第一意图特征向量与第五意图特征向量之间的第四向量距离;基于该第三向量距离和该第四向量距离确定待处理信息匹配的目标表情图像。基于该方式,能够从表达意图和语义内容两个维度查找该待处理信息匹配的目标表情图像,提高表情图像的查找效率和准确性。

[0012] 在一种可能的实现方式中,该方法还包括:在第一界面中显示该待处理信息匹配的目标表情图像,该第一界面为以下界面中的任意一项:图像搜索界面、输入法界面、表情图像推荐界面。基于该方式,能够提高图像查找的可视化,提高目标表情图像展示的灵活性。

[0013] 在一种可能的实现方式中,该方法还包括:显示表情图像添加界面,该表情图像添加界面包括表情图像框、自定义标签框和保存选项;在检测到用户针对该保存选项的触发操作时,保存该表情图像框中添加的第一表情图像以及该自定义标签框中添加的第一表情图像的第一标签;若该待处理信息与该第一标签相同,则将该第一表情图像确定为该目标表情图像。基于该方式,能够提高查找目标表情图像的效率和目标表情图像的准确性,有利于提高线上聊天对话的效率。

[0014] 第二方面,本申请提供一种图像查找装置,该装置可以是电子设备,也可以是电子设备中的装置,或者是能够和电子设备匹配使用的装置;其中,该图像查找装置还可以是芯片系统,该图像查找装置可执行第一方面中电子设备所执行的方法。该图像查找装置的功能可以通过硬件实现,也可以通过硬件执行相应的软件实现。该硬件或软件包括一个或多个与上述功能相对应的单元。该单元可以是软件和/或硬件。该图像查找装置执行的操作及有益效果可以参见上述第一方面所述的方法以及有益效果,重复之处不再赘述。

[0015] 第三方面,本申请提供了一种电子设备,包括一个或多个处理器和一个或多个存

存储器。该一个或多个存储器与一个或多个处理器耦合,一个或多个存储器用于存储计算机程序代码,计算机程序代码包括计算机指令,当一个或多个处理器执行计算机指令时,使得电子设备执行上述第一方面任一项可能的实现方式中的图像查找方法。

[0016] 第四方面,本申请提供了一种图像查找装置,该图像查找装置包括用于执行如第一方面中任一项所述方法的功能或单元。

[0017] 第五方面,本申请提供了一种计算机可读存储介质,该计算机可读存储介质中存储有计算机程序,该计算机程序包括程序指令,当该程序指令在电子设备上运行时,使得该电子设备执行上述第一方面任一项可能的实现方式中的图像查找方法。

[0018] 第六方面,本申请提供了一种计算机程序产品,当计算机程序产品在计算机上运行时,使得计算机执行上述第一方面任一项可能的实现方式中的图像查找方法。

附图说明

[0019] 图1A是本申请实施例提供的一种表情图像的示意图;

[0020] 图1B是本申请实施例提供的一种线上聊天对话的示意图;

[0021] 图2是本申请实施例提供的一种电子设备的硬件结构示意图;

[0022] 图3是本申请实施例提供的一种电子设备的软件结构框图;

[0023] 图4是本申请实施例提供的一种图像查找方法的流程示意图;

[0024] 图5是本申请实施例提供的一种电子设备获取表情图像数据库的流程示意图;

[0025] 图6是本申请实施例提供的一种第一对齐网络训练的流程示意图;

[0026] 图7A是本申请实施例提供的一种第一样本表情图像以及对应的第一文本信息的示意图;

[0027] 图7B是本申请实施例提供的一种第一对齐网络的训练过程的示意图;

[0028] 图8是本申请实施例提供的一种第二对齐网络训练的流程示意图;

[0029] 图9A是本申请实施例提供的一种第一样本对话信息的示意图;

[0030] 图9B是本申请实施例提供的一种第二样本对话信息的示意图;

[0031] 图9C是本申请实施例提供的一种第二对齐网络的训练过程的示意图;

[0032] 图10A是本申请实施例提供的一种表情图像搜索界面中待处理信息的示意图;

[0033] 图10B是本申请实施例提供的一种聊天框中待处理信息的示意图;

[0034] 图10C是本申请实施例提供的一种输入法界面中待处理信息的示意图;

[0035] 图11A是本申请实施例提供的一种表情图像搜索界面展示目标表情图像的示意图;

[0036] 图11B是本申请实施例提供的一种表情图像推荐界面展示目标表情图像的示意图;

[0037] 图11C是本申请实施例提供的一种输入法界面展示目标表情图像的示意图;

[0038] 图12是本申请实施例提供的一种表情图像添加界面的示意图;

[0039] 图13是本申请实施例提供的一种图像查找装置的结构示意图;

[0040] 图14是本申请实施例提供的一种芯片的结构示意图。

具体实施方式

[0041] 下面将结合附图对本申请实施例中的技术方案进行清楚、详尽地描述。其中,在本申请实施例的描述中,除非另有说明,“/”表示或的意思,例如,A/B可以表示A或B;文本中的“和/或”仅仅是一种描述关联对象的关联关系,表示可以存在三种关系,例如,A和/或B,可以表示:单独存在A,同时存在A和B,单独存在B这三种情况,另外,在本申请实施例的描述中,“多个”是指两个或两个以上。

[0042] 以下,术语“第一”、“第二”仅用于描述目的,而不能理解为暗示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此,限定有“第一”、“第二”的特征可以明示或者隐含地包括一个或者更多个该特征,在本申请实施例的描述中,除非另有说明,“多个”的含义是两个或两个以上。

[0043] 本申请以下实施例中的术语“用户界面 (user interface,UI)”,是应用程序或操作系统与用户之间进行交互和信息交换的介质接口,它实现信息的内部形式与用户可以接受形式之间的转换。用户界面是通过java、可扩展标记语言(extensible markup language,XML)等特定计算机语言编写的源代码,界面源代码在终端设备上经过解析,渲染,最终呈现为用户可以识别的内容。用户界面常用的表现形式是图形用户界面(graphic user interface,GUI),是指采用图形方式显示的与计算机操作相关的用户界面。它可以是在终端设备的显示屏中显示的文本、图标、按钮、菜单、选项卡、文本框、对话框、状态栏、导航栏、Widget等可视的界面元素。

[0044] 为了便于理解本申请实施例提供的方案,下面对本申请实施例涉及的相关概念进行介绍:

[0045] 1、人工智能与机器学习

[0046] 在本申请实施例中,涉及人工智能(Artificial Intelligence,AI)技术;所谓的AI是利用数字计算机或者数字计算机控制的机器模拟、延伸和扩展人的智能,感知环境、获取知识并使用知识获得最佳结果的理论、方法、技术及应用系统。具体地,AI技术涉及领域广泛,既有硬件层面的技术也有软件层面的技术;在硬件层面,AI技术一般包括如传感器、专用人工智能芯片、云计算、分布式存储、大数据处理技术、操作/交互系统、机电一体化等技术;在软件层面,AI技术主要包括计算机视觉技术、语音处理技术、自然语言处理技术以及机器学习/深度学习等几大方向。

[0047] 其中,机器学习(Machine Learning,ML)是一门多领域交叉学科,涉及概率论、统计学、逼近论、凸分析、算法复杂度理论等多门学科。专门研究计算机怎样模拟或实现人类的学习行为,以获取新的知识或技能,重新组织已有的知识结构使之不断改善自身的性能。机器学习是人工智能的核心,是使计算机具有智能的根本途径,其应用遍及人工智能的各个领域。机器学习/深度学习通常包括人工神经网络、监督学习、无监督学习等技术。

[0048] 计算机视觉技术(Computer Vision,CV)是一门研究如何使机器“看”的科学,更进一步的说,就是指用摄影机和电脑代替人眼对目标进行识别、追踪和测量等机器视觉,并进一步做图形处理,使电脑处理成为更适合人眼观察或传送给仪器检测的图像。作为一个科学学科,计算机视觉研究相关的理论和技术,试图建立能够从图像或者多维数据中获取信息的人工智能系统。计算机视觉技术通常包括图像处理、图像识别、图像语义理解、图像检索、光学字符识别(Optical Character Recognition,OCR)、视频处理、视频语义理解、视频

内容/行为识别、三维物体重建、3D技术、虚拟现实、增强现实、同步定位与地图构建等技术,还包括常见的人脸识别、指纹识别等生物特征识别技术。

[0049] 自然语言处理(Nature Language processing,NLP)是计算机科学领域与人工智能领域中的一个重要方向。它研究能实现人与计算机之间用自然语言进行有效通信的各种理论和方法。自然语言处理是一门融语言学、计算机科学、数学于一体的科学。因此,这一领域的研究将涉及自然语言,即人们日常使用的语言,所以它与语言学的研究有着密切的联系。自然语言处理技术通常包括文本处理、语义理解、机器翻译、机器人问答、知识图谱等技术。

[0050] 随着人工智能技术的研究和进步,人工智能技术在多个领域展开研究和应用,本申请实施例在实现图像查找方法时可以具体涉及人工智能技术中的机器学习、计算机视觉技术和自然语言处理技术等。

[0051] 2、对齐网络和大语言模型(Large Language Model,LLM)

[0052] 在本申请实施例中,该图像查找方法中会应用到机器学习技术中的对齐网络和大语言模型。所谓的对齐网络是一种将多个特征进行对齐的网络,可以应用于多模态领域,如文本特征对齐网络、图文特征对齐网络、图像特征对齐网络等。所谓的多模态是指多种模态的信息,包括文本、图像、视频、音频等。该图像查找方法主要应用的是图文特征对齐网络,即用于匹配图像特征和文本特征的预训练神经网络模型,例如CLIP(全称是Contrastive Language-Image Pre-Training)模型、BLIP(全称是Bootstrapping Language-Image Pre-training)模型等。其中,CLIP模型是OpenAI在2021年发布的,通过自监督学习的方式从未标注的图像和文本数据中进行预训练,使得模型能够理解图像和文本之间的语义联系,可以用于文本图像检索。BLIP模型是一种多模态 Transformer 模型,提出了一种编码器-解码器混合架构(Multimodal mixture of Encoder-Decoder,MED),MED的特点是很灵活,它既可以作为单模态的编码器,又可以作为基于图像的文本编码器,或者基于图像的文本解码器;BLIP模型由三个视觉语言目标联合训练:图像文本的对比学习、图像文本匹配和图像条件语言建模。

[0053] 所谓的大语言模型是指使用大量文本数据训练的深度学习模型,可以生成自然语言文本或理解语言文本的含义。大语言模型可以处理多种自然语言任务,如文本分类、问答、对话等,是通向人工智能的一条重要途径。常见的大语言模型有LLaMA2(全称是Large Language Model Meta AI 2)模型、ChatGPT(全称是Chat Generative Pre-trained Transformer)模型等。其中,LLaMA2模型是一种用于处理序列问题的神经网络模型,是于2023年2月发布的大型语言模型,它训练了各种模型,这些模型的参数从70亿到650亿不等。ChatGPT模型是人工智能技术驱动的自然语言处理工具,它能够基于在预训练阶段所见的模式和统计规律,来生成回答,还能根据聊天的上下文进行互动,真正像人类一样来聊天交流。

[0054] 3、表情图像

[0055] 所谓的表情图像也可以称为表情包,是指一种利用图像来表示感情的方式。当前社交网络中,表情图像已成为了一个必不可少的通讯媒介,人们往往更愿意使用表情图像来代替文字去传达自己的意图和情感。如图1A中(a)所示,该表情图像可以表达“我很疑惑”;又如图1A中(b)所示,该表情图像可以表达“我很喜欢”。在线上聊天的过程中,用户可

以通过发送表情图像来进行情感或意图的表达。如图1B所示,在聊天对话中,用户A发送“今天不能一起去聚餐了”;用户B发送了表情图像a,并且回复“昨天不是说今天有空吗?”,用户A发送“公司今天临时安排了新任务,有点忙”。其中,表情图像a便可以表达用户B“疑惑”的情感,即表情图像a对应的表达意图是“我很疑惑”。

[0056] 但是随着人们保存的表情图像日益增多,用户查找表情图像变得越来越困难且低效,在线上聊天时无法快速精确地找到自己想发送的表情图像,从而降低了聊天对话的效率。

[0057] 为了能够快速精确地查找到用户需要发送的表情图像,提高线上聊天对话的效率,提高用户体验,本申请提供了一种图像查找方法及电子设备。在具体实现中,上述所提及的图像查找方法可以由电子设备100执行。其中,电子设备100可以是手机、平板电脑、笔记本电脑或具备无线通讯功能的可穿戴电子设备(如智能手表)等,但并不局限于此。电子设备100配置有显示屏,并可以安装有预设应用程序(application,APP),如社交聊天APP。用户可以通过社交聊天APP与其他用户进行线上聊天,在线上聊天过程中可以向其他用户发送表情图像、文字、图片等,在此不做限定。当然,用户也可以保存表情图像至本地,也可以在向其他用户发送表情图像的过程中,搜索查找自己需要的表情图像。

[0058] 下面对电子设备100的硬件结构进行介绍。请参阅图2,图2是本申请实施例提供的电子设备100的硬件结构示意图。

[0059] 电子设备100可以包括处理器110,外部存储器接口120,内部存储器121,通用串行总线(universal serial bus,USB)接口130,充电管理模块140,电源管理模块141,电池142,天线1,天线2,移动通信模块150,无线通信模块160,音频模块170,扬声器170A,受话器170B,麦克风170C,耳机接口170D,传感器模块180,按键190,马达191,指示器192,摄像头193,显示屏194,以及用户标识模块(subscriber identification module,SIM)卡接口195等。其中传感器模块180可以包括压力传感器180A,陀螺仪传感器180B,气压传感器180C,磁传感器180D,加速度传感器180E,距离传感器180F,接近光传感器180G,指纹传感器180H,温度传感器180J,触摸传感器180K,环境光传感器180L,骨传导传感器180M等。

[0060] 可以理解的是,本发明实施例示意的结构并不构成对电子设备100的具体限定。在本申请另一些实施例中,电子设备100可以包括比图示更多或更少的部件,或者组合某些部件,或者拆分某些部件,或者不同的部件布置。图示的部件可以以硬件,软件或软件和硬件的组合实现。

[0061] 处理器110可以包括一个或多个处理单元,例如:处理器110可以包括应用处理器(application processor,AP),调制解调处理器,图形处理器(graphics processing unit,GPU),图像信号处理器(image signal processor,ISP),控制器,存储器,视频编解码器,数字信号处理器(digital signal processor,DSP),基带处理器,和/或神经网络处理器(neural-network processing unit,NPU)等。其中,不同的处理单元可以是独立的器件,也可以集成在一个或多个处理器中。

[0062] 其中,控制器可以是电子设备100的神经中枢和指挥中心。控制器可以根据指令操作码和时序信号,产生操作控制信号,完成取指令和执行指令的控制。

[0063] 处理器110中还可以设置存储器,用于存储指令和数据。在一些实施例中,处理器110中的存储器为高速缓冲存储器。该存储器可以保存处理器110刚用过或循环使用的指令

或数据。如果处理器110需要再次使用该指令或数据,可从所述存储器中直接调用。避免了重复存取,减少了处理器110的等待时间,因而提高了系统的效率。处理器110调用存储器中存储的指令或数据,使电子设备100执行以下方法实施例中电子设备所执行的图像查找方法。

[0064] 在一些实施例中,处理器110可以包括一个或多个接口。接口可以包括集成电路(inter-integrated circuit,I2C)接口,集成电路内置音频(inter-integrated circuit sound,I2S)接口,脉冲编码调制(pulse code modulation,PCM)接口,通用异步收发传输器(universal asynchronous receiver/transmitter,UART)接口,移动产业处理器接口(mobile industry processor interface,MIPI),通用输入输出(general-purpose input/output,GPIO)接口,用户标识模块(subscriber identity module,SIM)接口,和/或通用串行总线(universal serial bus,USB)接口等。

[0065] 充电管理模块140用于从充电器接收充电输入。其中,充电器可以是无线充电器,也可以是有线充电器。

[0066] 电源管理模块141用于连接电池142,充电管理模块140与处理器110。电源管理模块141接收电池142和/或充电管理模块140的输入,为处理器110,内部存储器121,外部存储器,显示屏194,摄像头193,和无线通信模块160等供电在其他一些实施例中,电源管理模块141也可以设置于处理器110中。

[0067] 电子设备100的无线通信功能可以通过天线1,天线2,移动通信模块150,无线通信模块160,调制解调处理器以及基带处理器等实现。

[0068] 天线1和天线2用于发射和接收电磁波信号。电子设备100中的每个天线可用于覆盖单个或多个通信频带。不同的天线还可以复用,以提高天线的利用率。例如:可以将天线1复用为无线局域网的分集天线。在另外一些实施例中,天线可以和调谐开关结合使用。

[0069] 移动通信模块150可以提供应用在电子设备100上的包括2G/3G/4G/5G等无线通信的解决方案。移动通信模块150可以包括至少一个滤波器,开关,功率放大器,低噪声放大器(low noise amplifier,LNA)等。移动通信模块150可以由天线1接收电磁波,并对接收的电磁波进行滤波,放大等处理,传送至调制解调处理器进行解调。移动通信模块150还可以对经调制解调处理器调制后的信号放大,经天线1转为电磁波辐射出去。在一些实施例中,移动通信模块150的至少部分功能模块可以被设置于处理器110中。在一些实施例中,移动通信模块150的至少部分功能模块可以与处理器110的至少部分模块被设置在同一个器件中。

[0070] 调制解调处理器可以包括调制器和解调器。其中,调制器用于将待发送的低频基带信号调制为中高频信号。解调器用于将接收的电磁波信号解调为低频基带信号。随后解调器将解调得到的低频基带信号传送至基带处理器处理。低频基带信号经基带处理器处理后,被传递给应用处理器。

[0071] 无线通信模块160可以提供应用在电子设备100上的包括无线局域网(wireless local area networks,WLAN)(如Wi-Fi网络),蓝牙(blueetooth,BT),BLE广播,全球导航卫星系统(global navigation satellite system,GNSS),调频(frequency modulation,FM),近距离无线通信技术(near field communication,NFC),红外技术(infrared,IR)等无线通信的解决方案。无线通信模块160可以是集成至少一个通信处理模块的一个或多个器件。无线通信模块160经由天线2接收电磁波,将电磁波信号调频以及滤波处理,将处理后

的信号发送到处理器110。无线通信模块160还可以从处理器110接收待发送的信号,对其进行调频,放大,经天线2转为电磁波辐射出去。

[0072] 在一些实施例中,电子设备100的天线1和移动通信模块150耦合,天线2和无线通信模块160耦合,使得电子设备100可以通过无线通信技术与网络以及其他设备通信。

[0073] 电子设备100通过GPU,显示屏194,以及应用处理器等实现显示功能。GPU为图像处理的微处理器,连接显示屏194和应用处理器。GPU用于执行数学和几何计算,用于图形渲染。处理器110可包括一个或多个GPU,其执行程序指令以生成或改变显示信息。

[0074] 显示屏194用于显示图像,视频等。显示屏194包括显示面板。在一些实施例中,电子设备100可以包括1个或N个显示屏194,N为大于1的正整数。

[0075] 电子设备100可以通过ISP,摄像头193,视频编解码器,GPU,显示屏194以及应用处理器等实现拍摄功能。ISP 用于处理摄像头193反馈的数据。摄像头193用于捕获静态图像或视频。数字信号处理器用于处理数字信号,除了可以处理数字图像信号,还可以处理其他数字信号。视频编解码器用于对数字视频压缩或解压缩。电子设备100可以支持一种或多种视频编解码器。

[0076] NPU为神经网络(neural-network ,NN) 计算处理器,通过借鉴生物神经网络结构,例如借鉴人脑神经元之间传递模式,对输入信息快速处理,还可以不断的自学习。

[0077] 外部存储器接口120可以用于连接外部存储卡,例如Micro SD卡,实现扩展电子设备100的存储能力。外部存储卡通过外部存储器接口120与处理器110通信,实现数据存储功能。

[0078] 内部存储器121可以用于存储计算机可执行程序代码,所述可执行程序代码包括指令。处理器110通过运行存储在内部存储器121的指令,从而执行电子设备100的各种功能应用以及数据处理。内部存储器121可以包括存储程序区和存储数据区。其中,存储程序区可存储操作系统,至少一个功能所需的应用程序(比如声音播放功能)等。存储数据区可存储电子设备100使用过程中所创建的数据(比如音频数据)等。此外,内部存储器121可以包括高速随机存取存储器,还可以包括非易失性存储器,例如闪存器件等。

[0079] 电子设备100可以通过音频模块170,扬声器170A,受话器170B,麦克风170C,耳机接口170D,以及应用处理器等实现音频功能。例如音乐播放,录音等。

[0080] 音频模块170用于将数字音频信息转换成模拟音频信号输出,也用于将模拟音频输入转换为数字音频信号。音频模块170还可以用于对音频信号编码和解码。在一些实施例中,音频模块170可以设置于处理器110中,或将音频模块170的部分功能模块设置于处理器110中。

[0081] 扬声器170A,也称“喇叭”,用于将音频电信号转换为声音信号。受话器170B,也称“听筒”,用于将音频电信号转换成声音信号。麦克风170C,也称“话筒”,“传声器”,用于将声音信号转换为电信号。耳机接口170D用于连接有线耳机。压力传感器180A用于感受压力信号,可以将压力信号转换成电信号。在一些实施例中,压力传感器180A可以设置于显示屏194。陀螺仪传感器180B可以用于确定电子设备100的运动姿态。气压传感器180C用于测量气压。磁传感器180D包括霍尔传感器。加速度传感器180E可检测电子设备100在各个方向上(一般为三轴)加速度的大小。距离传感器180F,用于测量距离。接近光传感器180G可以包括例如发光二极管(LED)和光检测器。环境光传感器180L用于感知环境光亮度。指纹传感器

180H用于采集指纹。温度传感器180J用于检测温度。触摸传感器180K,也称“触控面板”。触摸传感器180K可以设置于显示屏194,由触摸传感器180K与显示屏194组成触摸屏,也称“触控屏”。触摸传感器180K用于检测作用于其上或附近的触摸操作。骨传导传感器180M可以获取振动信号。按键190包括开机键,音量键等。马达191可以产生振动提示。指示器192可以是指示灯,可以用于指示充电状态,电量变化,也可以用于指示消息,未接来电,通知等。SIM卡接口195用于连接SIM卡。

[0082] 电子设备100的软件系统可以采用分层架构,事件驱动架构,微核架构,微服务架构,或云架构。本发明实施例以分层架构的Android系统为例,示例性说明电子设备100的软件结构。图3是本申请实施例的电子设备100的软件结构框图。分层架构将软件分成若干层,每一层都有清晰的角色和分工。层与层之间通过软件接口通信。在一些实施例中,将Android系统分为四层,从上至下分别为应用程序层,应用程序框架层,安卓运行时(Android runtime)和系统库,以及内核层。

[0083] 应用程序层可以包括一系列应用程序包。如图3所示,应用程序包可以包括聊天APP、相机,图库,日历,通话, WLAN,蓝牙,音乐,视频,短信息等应用程序。

[0084] 应用程序框架层为应用程序层的应用程序提供应用编程接口(application programming interface,API)和编程框架。应用程序框架层包括一些预先定义的函数。如图3所示,应用程序框架层可以包括窗口管理器,内容提供者,视图系统,电话管理器,资源管理器,通知管理等。

[0085] 窗口管理器用于管理窗口程序。窗口管理器可以获取显示屏大小,判断是否有状态栏,锁定屏幕,截取屏幕等。

[0086] 内容提供者用来存放和获取数据,并使这些数据可以被应用程序访问。所述数据可以包括视频,图像,音频,拨打和接听的电话,浏览历史和书签,电话簿等。

[0087] 视图系统包括可视控件,例如显示文字的控件,显示图片的控件等。视图系统可用于构建应用程序。显示界面可以由一个或多个视图组成的。例如,包括短信通知图标的显示界面,可以包括显示文字的视图以及显示图片的视图。

[0088] 电话管理器用于提供电子设备100的通信功能。例如通话状态的管理(包括接通,挂断等)。

[0089] 资源管理器为应用程序提供各种资源,比如本地化字符串,图标,图片,布局文件,视频文件等等。

[0090] 通知管理器使应用程序可以在状态栏中显示通知信息,可以用于传达告知类型的消息,可以短暂停留后自动消失,无需用户交互。比如通知管理器被用于告知下载完成,消息提醒等。通知管理器还可以是以图表或者滚动条文本形式出现在系统顶部状态栏的通知,例如后台运行的应用程序的通知,还可以是以对话框形式出现在屏幕上的通知。例如在状态栏提示文本信息,发出提示音,电子设备振动,指示灯闪烁等。

[0091] Android runtime包括核心库和虚拟机。Android runtime负责安卓系统的调度和管理。

[0092] 核心库包含两部分:一部分是java语言需要调用的功能函数,另一部分是安卓的核心库。

[0093] 应用程序层和应用程序框架层运行在虚拟机中。虚拟机将应用程序层和应用程序

框架层的java文件执行为二进制文件。虚拟机用于执行对象生命周期的管理,堆栈管理,线程管理,安全和异常的管理,以及垃圾回收等功能。

[0094] 系统库可以包括多个功能模块。例如:表面管理器(surface manager),媒体库(media libraries),三维图形处理库(例如:OpenGL ES),2D图形引擎(例如:SGL)等。

[0095] 表面管理器用于对显示子系统进行管理,并且为多个应用程序提供了2D和3D图层的融合。

[0096] 媒体库支持多种常用的音频,视频格式回放和录制,以及静态图像文件等。媒体库可以支持多种音视频编码格式。

[0097] 三维图形处理库用于实现三维图形绘图,图像渲染,合成,和图层处理等。

[0098] 2D图形引擎是2D绘图的绘图引擎。

[0099] 内核层是硬件和软件之间的层。内核层至少包含显示驱动,摄像头驱动,音频驱动,传感器驱动。

[0100] 下面进一步对本申请提供的图像查找方法进行详细描述。请参见图4,图4是本申请实施例提供的一种图像查找方法的流程示意图。如图4所示,该图像查找方法包括如下步骤S401~S403,图4所示的方法执行主体包括电子设备(如图2中的电子设备100)。或者,图4所示的方法执行主体包括电子设备中的芯片。图4以电子设备为执行主体为例进行说明。其中:

[0101] S401、电子设备获取表情图像数据库。

[0102] 在本申请实施例中,该表情图像数据库包括多个表情图像、多个表情图像中每个表情图像对应的第一内容特征向量和第一意图特征向量。该第一内容特征向量表示描述表情图像的内容的特征向量,该第一意图特征向量表示描述表情图像的表达意图的特征向量。例如,该表情图像数据库包括3个表情图像,即表情图像a、表情图像b和表情图像c;该表情图像数据库还包括每个表情图像对应的第一内容特征向量和第一意图特征向量,即表情图像a对应的第一内容特征向量和第一意图特征向量、表情图像b对应的第一内容特征向量和第一意图特征向量、表情图像c对应的第一内容特征向量和第一意图特征向量。电子设备首先需要获取到表情图像数据库,以便于后续可以在表情图像数据库中搜寻需要的表情图像。

[0103] 针对表情图像的内容和表达意图,以图1A中(a)所示的表情图像为例,该表情图像的内容是“一只白色的小狗,侧坐在地上,眼睛望着前方,小狗的头上有问号的标记”;该表情图像的表达意图是“我很疑惑”。

[0104] 其中,这里的表情图像可以是用户在本本地收藏或添加的表情图像(如用户在本本地收藏或添加的自定义表情图像),也可以是用户在云端商城下载的表情图像,在此不做限定。也就是说,该表情图像数据库既可以支持对云端商城下载的表情图像的搜寻,也可以支持在本本地收藏的表情图像的搜寻。本申请实施例是以该多个表情图像是用户在本本地收藏的表情图像为例进行说明。

[0105] 需要说明的是,本申请实施例所涉及到用户、用户在本本地收藏的表情图像等相关的数据,均是经过用户授权后所获取到的。并且,当本申请实施例运用到具体产品或技术中时,所涉及使用到的数据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。

[0106] 在一种可能的实现方式中,电子设备获取表情图像数据库时,如图5所示,具体的实现方式可以包括以下步骤s11~s14。基于该方式,能够从表达意图和语义内容两个维度建立表情图像数据库,有利于提高表情图像的查找效率。

[0107] s11、电子设备获取用户在本地收藏的多个表情图像。

[0108] 在具体实现中,用户在本地收藏的多个表情图像可以是用户在线上聊天过程中收藏的其他用户发送的自定义表情图像,也可以是用户自己在本地制作的自定义表情图像,也可以是用户自己在本地添加的自定义表情图像,在此不做限定。

[0109] s12、电子设备调用图像编码器对多个表情图像中每个表情图像进行编码处理,得到每个表情图像对应的第一图像特征向量。

[0110] 在具体实现中,图像编码器可以称为Image Encoder,可以是已经训练好的图像编码器,用于将图像转化为特征向量表示,例如该图像编码器可以是Transformer模型中的编码器,也可以是其他模型中的编码器,在此不做限定。电子设备可以调用图像编码器对每个表情图像进行编码处理,得到每个表情图像对应的第一图像特征向量。

[0111] s13、电子设备调用对齐网络对第一图像特征向量和提示词进行处理,得到每个表情图像对应的第一内容特征向量和第一意图特征向量。

[0112] 在具体实现中,该提示词用于提示对齐网络描述表情图像的内容和表达意图,例如该提示词为“详细描述表情图像的内容和表达意图”。针对每个表情图像,电子设备需要将该表情图像对应的第一图像特征向量和该提示词一起输入到对齐网络中处理,该对齐网络输出该表情图像对应的第一内容特征向量和第一意图特征向量。其中,该对齐网络可以是一个既提取表情图像的第一内容特征向量,又提取表情图像的第一意图特征向量的对齐网络;也可以由两个独立的对齐网络直接组成的对齐网络,例如该对齐网络可以包括第一对齐网络和第二对齐网络,第一对齐网络用于提取表情图像的第一内容特征向量,第二对齐网络用于提取表情图像的第一意图特征向量。

[0113] 可选地,以该对齐网络是由两个独立的对齐网络组成的为例,该对齐网络包括第一对齐网络和第二对齐网络,该提示词包括第一提示词和第二提示词。电子设备调用对齐网络对第一图像特征向量和提示词进行处理,得到每个表情图像对应的第一内容特征向量和第一意图特征向量时,具体的实现方式可以包括以下步骤1和步骤2。其中,步骤1和步骤2的执行顺序不做限定。基于该方式,能够提高每个表情图像对应的第一内容特征向量和第一意图特征向量的准确性。

[0114] 步骤1、电子设备调用第一对齐网络对第一图像特征向量和第一提示词进行处理,得到每个表情图像对应的第一内容特征向量。

[0115] 在具体实现中,该第一提示词用于提示第一对齐网络描述表情图像的内容,例如该第一提示词为“详细描述表情图像的内容”。针对每个表情图像,电子设备需要将该表情图像对应的第一图像特征向量和第一提示词一起输入到第一对齐网络中处理,该第一对齐网络输出该表情图像对应的第一内容特征向量。

[0116] 可选地,该第一对齐网络可以是其他计算机设备训练好后发送给电子设备使用的。这里的计算机设备可以是服务器或其他设备,服务器可以是独立的物理服务器,或者由多台服务器组成的服务器集群,或者是一个云端计算中心(如云服务器)。其中,所谓的云服务器是一种简单高效、安全可靠计算服务,其管理方式要比物理服务器更简单高效;可以提

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。可以理解为,第一对齐网络是离线在云端训练好的,不会在边缘端进行训练。基于该方式,有利于减少对电子设备硬件资源的依赖,减少电子设备的负担。

[0117] 以计算机设备是服务器为例,如图6所示,具体可以包括以下步骤A~步骤H。其中,步骤B~D和步骤E~F的执行顺序不做限定,可以先执行步骤B~D,再执行步骤E~F;也可以先执行步骤E~F,再执行步骤B~D,也可以同时执行步骤B~D和步骤E~F。

[0118] 步骤A、服务器获取第一样本表情图像以及对应的第一文本信息。

[0119] 在具体实现中,第一样本表情图像以及对应的第一文本信息可以组成一个第一训练样本,服务器需要获取多个这样的第一训练样本对模型进行训练。例如,该第一样本表情图像为图7A中(a)所示;该第一样本表情图像对应的第一文本信息为“一只白色的小狗,侧坐在地上,眼睛望着前方,小狗的头上有问号的标记”,如图7A中(b)所示。

[0120] 步骤B、服务器调用图像编码器对第一样本表情图像进行编辑处理,得到第二图像特征向量。

[0121] 在具体实现中,服务器也可以将第一样本表情图像输入到已经训练好的图像编码器中进行编码处理,可以得到第二图像特征向量。

[0122] 步骤C、服务器调用第一初始对齐网络对第二图像特征向量和第一提示词进行处理,得到第二内容特征向量。

[0123] 在具体实现中,这里的第一初始对齐网络可以是神经网络模型,如CLIP模型、BLIP模型等,在此不做限定。服务器可以将第二图像特征向量和第一提示词输入到第一初始对齐网络中进行处理,让第一初始对齐网络描述一下表情图像的内容,得到第二内容特征向量。

[0124] 步骤D、服务器调用大语言模型对第二内容特征向量和第一提示词进行处理,得到第三内容特征向量。

[0125] 在具体实现中,这里的大语言模型是已经使用大量文本数据训练好的深度学习模型,如LLaMA2模型、ChatGPT模型等。服务器可以将第二内容特征向量和第一提示词输入到大语言模型中进行处理,让大语言模型描述一下表情图像的内容,得到第三内容特征向量。

[0126] 步骤E、服务器调用文本编码器对第一文本信息进行编码处理,得到第二文本特征向量。

[0127] 在具体实现中,文本编码器可以称为Text Encoder,可以是已经训练好的文本编码器,用于将文本信息转化为特征向量表示,例如该文本编码器可以是Transformer模型中的编码器,也可以是BERT模型中的编码器,也可以是其他模型中的编码器,在此不做限定。服务器可以将第一文本信息输入到文本编码器中进行编码处理,得到第二文本特征向量。

[0128] 步骤F、服务器调用大语言模型对第二文本特征向量和第三提示词进行处理,得到第三文本特征向量。

[0129] 在具体实现中,该第三提示词用于提示大语言模型解释文本信息,例如该第三提示词为“重新解释文本信息的描述”。服务器可以将第二文本特征向量和第三提示词输入到大语言模型中进行处理,让大语言模型重新解释一下第一文本信息,得到第三文本特征向量。

[0130] 步骤G、服务器基于第三内容特征向量和第三文本特征向量更新第一初始对齐网络的模型参数,得到第一对齐网络。

[0131] 在具体实现中,服务器利用第三内容特征向量和第三文本特征向量计算第一模型损失值(Loss),这里的第一模型损失值可以是第三内容特征向量与第三文本特征向量之间的相似度,例如余弦相似度(Cosine Similarity),即计算两个向量之间夹角的余弦值,用来衡量两个向量在方向上的相似程度。服务器可以按照减少该第一模型损失值的方向,调整该第一初始对齐网络的模型参数;当训练达到次数后,完成训练,得到第一对齐网络。具体调整该第一初始对齐网络的模型参数的方式也可以采用随机梯度下降方法、自适应梯度算法等,在此不做限定。

[0132] 步骤H、服务器向电子设备发送第一对齐网络。相应地,电子设备接收来自服务器的第一对齐网络。

[0133] 在具体实现中,服务器训练好第一对齐网络后,便可以将该第一对齐网络发送给电子设备,以供电子设备使用。该第一对齐网络可以用于提取表情图像的第一内容特征向量。

[0134] 总的来说,针对第一对齐网络的训练过程,如图7B所示,服务器获取多个第一训练样本,每个第一训练样本包括第一样本表情图像以及对应的第一文本信息。针对每个第一训练样本,服务器将获取的第一样本表情图像输入到已经训练好的图像编码器中进行编码处理,可以得到第二图像特征向量;将第二图像特征向量和第一提示词输入到第一初始对齐网络中进行处理,让第一初始对齐网络描述一下表情图像的内容,得到第二内容特征向量;将第二内容特征向量和第一提示词(即①和②)输入到大语言模型中进行处理,让大语言模型描述一下表情图像的内容,得到第三内容特征向量。另外,服务器也会将获取的第一样本表情图像对应的第一文本信息输入到文本编码器中进行编码处理,得到第二文本特征向量;将第二文本特征向量和第三提示词(即③和④)输入到大语言模型中进行处理,让大语言模型重新解释一下第一文本信息,得到第三文本特征向量。最后,服务器利用第三内容特征向量和第三文本特征向量计算第一模型损失值,按照减少该第一模型损失值的方向,调整该第一初始对齐网络的模型参数;当训练达到次数后,完成训练,得到第一对齐网络。

[0135] 当然,如果电子设备有足够大的处理能力,也可以自己训练第一对齐网络,具体的训练过程也可以参考上述步骤A~步骤G,在此不做限定。

[0136] 步骤2、电子设备调用第二对齐网络对第一图像特征向量和第二提示词进行处理,得到每个表情图像对应的第一意图特征向量。

[0137] 在具体实现中,该第二提示词用于提示所述第二对齐网络描述表情图像的表达意图,例如该第一提示词为“详细描述表情图像的表达意图”。针对每个表情图像,电子设备需要将该表情图像对应的第一图像特征向量和第二提示词一起输入到第二对齐网络中处理,该第二对齐网络输出该表情图像对应的第一意图特征向量。

[0138] 可选地,该第二对齐网络可以是其他计算机设备训练好后发送给电子设备使用的。这里的计算机设备也可以是服务器或其他设备,服务器可以是独立的物理服务器,或者由多台服务器组成的服务器集群,或者是一个云端计算中心(如云服务器)。可以理解为,第二对齐网络也是离线在云端训练好的,不会在边缘端进行训练。基于该方式,有利于减少对电子设备硬件资源的依赖,减少电子设备的负担。

[0139] 以计算机设备是服务器为例,如图8所示,具体可以包括以下步骤a~步骤i。其中,步骤b~e和步骤f~g的执行顺序不做限定,可以先执行步骤b~e,再执行步骤f~g;也可以先执行步骤f~g,再执行步骤b~e,也可以同时执行步骤b~e和步骤f~g。

[0140] 步骤a、服务器获取第一样本对话信息。

[0141] 在具体实现中,该第一样本对话信息包括对话文本信息和第二样本表情图像。如图9A所示,该第一样本对话信息中的对话文本信息可以是:“今天不能一起去聚餐了”“昨天不是说今天有空吗?”“公司今天临时安排了新任务,有点忙”;该第二样本表情图像是在第一样本对话信息中用户发送的表情图像。

[0142] 步骤b、服务器基于第一样本对话信息确定第二样本表情图像对应的第二文本信息。

[0143] 可选地,服务器基于第一样本对话信息确定第二样本表情图像对应的第二文本信息时,具体的实现方式可以是:调用大语言模型对第一样本对话信息进行处理,得到第二样本对话信息;从第二样本对话信息中确定第二样本表情图像对应的第二文本信息。

[0144] 在具体实现中,服务器可以将第一样本对话信息输入到大语言模型中进行处理,让大语言模型根据对话上下文,将第一样本对话信息中的所有表情图像替换为文本信息,从而输出第二样本对话信息,如图9B所示。此时,便可以从第二样本对话信息中找到第二样本表情图像对应的第二文本信息,即在图9B中,第二样本表情图像对应的第二文本信息为“我很疑惑”。其中,可以将第一样本对话信息和第二样本对话信息组成一个第二训练样本,服务器需要获取多个这样的第二训练样本对模型进行训练。

[0145] 步骤c、服务器调用图像编码器对第二样本表情图像进行编辑处理,得到第三图像特征向量。

[0146] 在具体实现中,服务器也可以将第二样本表情图像输入到已经训练好的图像编码器中进行编码处理,可以得到第三图像特征向量。

[0147] 步骤d、服务器调用第二初始对齐网络对第三图像特征向量和第二提示词进行处理,得到第二意图特征向量。

[0148] 在具体实现中,这里的第二初始对齐网络也可以是神经网络模型,如CLIP模型、BLIP模型等,在此不做限定。服务器可以将第三图像特征向量和第二提示词输入到第二初始对齐网络中进行处理,让第二初始对齐网络描述一下表情图像的表达意图,得到第二意图特征向量。

[0149] 步骤e、服务器调用大语言模型对第二意图特征向量、第二提示词和对话文本信息中提取的上下文信息进行处理,得到第三意图特征向量。

[0150] 在具体实现中,这里的大语言模型也可以是已经使用大量文本数据训练好的深度学习模型,如LLaMA2模型、ChatGPT模型等。服务器可以将第二意图特征向量、第二提示词以及对话文本信息中提取的上下文信息输入到大语言模型中进行处理,让大语言模型描述一下表情图像的表达意图,得到第三意图特征向量。其中,服务器可以采用训练好的视觉模型提取对话文本信息中的上下文信息,这里的视觉模型可以是BERT模型、GPT模型、ELMO模型等,在此不做限定。

[0151] 步骤f、服务器调用文本编码器对第二文本信息进行编码处理,得到第四文本特征向量。

[0152] 在具体实现中,该文本编码器也可以是Transformer模型中的编码器,也可以是BERT模型中的编码器,也可以是其他模型中的编码器,在此不做限定。服务器可以将第二文本信息输入到文本编码器中进行编码处理,得到第四文本特征向量。

[0153] 步骤g、服务器调用大语言模型对第四文本特征向量、第四提示词和上下文信息进行处理,得到第四意图特征向量。

[0154] 在具体实现中,第四提示词用于提示大语言模型描述文本信息的表达意图,例如该第四提示词可以是“详细描述文本信息的表达意图”。服务器可以将第四文本特征向量、第四提示词以及上述提取的上下文信息,输入到大语言模型中进行处理,让大语言模型描述一下第二文本信息的表达意图,得到第四意图特征向量。

[0155] 步骤h、服务器基于第三意图特征向量和第四意图特征向量更新第二初始对齐网络的模型参数,得到第二对齐网络。

[0156] 在具体实现中,服务器利用第三意图特征向量和第四意图特征向量计算第二模型损失值,这里的第二模型损失值可以是第三意图特征向量与第四意图特征向量之间的相似度,例如余弦相似度,即计算两个向量之间夹角的余弦值,用来衡量两个向量在方向上的相似程度。服务器可以按照减少该第二模型损失值的方向,调整该第二初始对齐网络的模型参数;当训练达到次数后,完成训练,得到第二对齐网络。具体调整该第二初始对齐网络的模型参数的方式也可以采用随机梯度下降方法、自适应梯度算法等,在此不做限定。

[0157] 步骤i、服务器向电子设备发送第二对齐网络。相应地,电子设备接收来自服务器的第二对齐网络。

[0158] 在具体实现中,服务器训练好第二对齐网络后,便可以将该第二对齐网络发送给电子设备,以供电子设备使用。该第二对齐网络可以用于提取表情图像的第一意图特征向量。

[0159] 总的来说,针对第二对齐网络的训练过程,如图9C所示,服务器获取到第一样本对话信息,该第一样本对话信息包括对话文本信息和第二样本表情图像;将第一样本对话信息输入到大语言模型中处理,输出第二样本对话信息;从第二样本对话信息中确定第二样本表情图像对应的第二文本信息。服务器将第二样本表情图像输入到已经训练好的图像编码器中进行编码处理,得到第三图像特征向量。将第三图像特征向量和第二提示词输入到第二初始对齐网络中进行处理,让第二初始对齐网络描述一下表情图像的表达意图,得到第二意图特征向量;将第二意图特征向量、第二提示词以及对话文本信息中提取的上下文信息(即①、②和③)输入到大语言模型中进行处理,让大语言模型描述一下表情图像的表达意图,得到第三意图特征向量。服务器将第二文本信息输入到文本编码器中进行编码处理,得到第四文本特征向量;将第四文本特征向量、第四提示词以及上述提取的上下文信息(即③、④和⑤)输入到大语言模型中进行处理,让大语言模型描述一下第二文本信息的表达意图,得到第四意图特征向量。最后,服务器利用第三意图特征向量和第四意图特征向量计算第二模型损失值,按照减少该第二模型损失值的方向,调整该第二初始对齐网络的模型参数;当训练达到次数后,完成训练,得到第二对齐网络。

[0160] 当然,如果电子设备有足够大的处理能力,也可以自己训练第二对齐网络,具体的训练过程也可以参考上述步骤a~步骤h,在此不做限定。

[0161] s14、电子设备利用多个表情图像、每个表情图像对应的第一内容特征向量和第一

意图特征向量构建表情图像数据库。

[0162] 在具体实现中,将多个表情图像、每个表情图像对应的第一内容特征向量和第一意图特征向量一起组成一个表情图像数据库,以便后续用于表情图像的查找。

[0163] S402、电子设备确定待处理信息的第一文本特征向量。

[0164] 在本申请实施例中,这里的待处理信息可以是用户输入或搜索的文本信息(如词语、语句等),这里的文本信息可以是需要查找的表情图像所表达的情绪意图,也可以是对需要查找的表情图像的内容描述。也就是说,该图像查找方法既可以支持针对表情图像的表达意图的搜索,也可以支持针对表情图像的语义内容的搜索。

[0165] 例如,该待处理信息可以是用户在表情图像搜索框中输入的文字。示例性地,如图10A所示,用户在表情图像搜索界面的表情图像搜索框中输入“疑惑”,这里的“疑惑”便是待处理信息;也可以在表情图像搜索界面的表情图像搜索框中输入“一只小狗或小猫,头上有问号”,这里的“一只小狗或小猫,头上有问号”便是待处理信息。

[0166] 又例如,该待处理信息可以是用户在聊天框中输入的未发送的文字。示例性地,如图10B所示,用户在聊天框中编辑输入的未发送的文字为“疑惑”,这里的“疑惑”便是待处理信息。

[0167] 又例如,该待处理信息可以是用户在输入法界面中的表情图像搜索栏中输入的文字。示例性地,如图10C中(a)所示,用户在输入法界面中,点击表情按钮,显示表情界面;如图10C中(b)所示,该表情界面中包括Emoji选项、颜文字选项、表情图选项、斗图选项,用户可以点击表情图选项进入表情图像搜索界面(可以认为是本地TAB页);如图10C中(c)所示,该表情图像搜索界面包括表情图像搜索界栏,用户可以在表情图像搜索栏中输入需要搜索的文字,如“疑惑”,这里的“疑惑”便是待处理信息;当然,也可以在表情图像搜索栏中输入需要搜索的语句,如“一只小狗或小猫,头上有问号”,这里的“一只小狗或小猫,头上有问号”便是待处理信息。

[0168] 当然,该待处理信息也可以是其他方式输入的信息,在此不做限定。电子设备需要将这些待处理信息的第一文本特征向量提取出来,以便后续处理。

[0169] 在一种可能的实现方式中,电子设备确定待处理信息的第一文本特征向量时,具体的实现方式可以是:调用文本编码器对待处理信息进行处理,得到待处理信息的第一文本特征向量。也就是说,电子设备也可以将待处理信息输入到已经训练好的文本编码器中处理,得到待处理信息的第一文本特征向量。这里的文本编码器可以是上述提及的文本编码器,也可以是其他文本编码器,在此不做限定。

[0170] S403、电子设备从图像数据库中查找该待处理信息匹配的目标表情图像。

[0171] 在本申请实施例中,该目标表情图像对应的第一内容特征向量和第一意图特征向量与第一文本特征向量相匹配。可以理解为,在图像数据库中包括多个表情图像以及每个表情图像对应的第一内容特征向量和第一意图特征向量,电子设备需要在图像数据库中查找到与第一文本特征向量相匹配的第一内容特征向量和第一意图特征向量,即目标表情图像对应的第一内容特征向量和第一意图特征向量,此时目标表情图像即为该待处理信息匹配的表情图像。其中,目标表情图像可以是一个或多个。

[0172] 可选地,电子设备可以记录该待处理信息匹配的目标表情图像,这样后续获取到相同的待处理信息时,便可以展示匹配的目标表情图像。基于该方式,有利于提高查找

表情图像的效率。另外,如果在第一预设时间段内,用户未使用过第一界面中展示的第一目标表情图像,电子设备可以在第一界面中不展示该第一目标表情图像,也可以重新查找该待处理信息匹配的其他表情图像。

[0173] 在一种可能的实现方式中,电子设备从图像数据库中查找该待处理信息匹配的目标表情图像时,具体的实现方式可以是:确定图像数据库中每个表情图像对应的第一内容特征向量与第一文本特征向量之间的第一向量距离,以及确定图像数据库中每个表情图像对应的第一意图特征向量与第一文本特征向量之间的第二向量距离;基于第一向量距离和第二向量距离确定待处理信息匹配的目标表情图像。基于该方式,能够从表达意图和语义内容两个维度查找该待处理信息匹配的目标表情图像,提高表情图像的查找效率。这里的向量距离可以是指向量之间的欧式距离,下面对基于第一向量距离和第二向量距离确定待处理信息匹配的目标表情图像的不同方式进行详细描述。

[0174] 方式一:按照从小到大的顺序排列每个表情图像对应的第一向量距离,将前X个第一向量距离对应的表情图像确定为该待处理信息匹配的目标表情图像;按照从小到大的顺序排列每个表情图像对应的第二向量距离,将前Y个第二向量距离对应的表情图像确定为该待处理信息匹配的目标表情图像。其中,X和Y为正整数。

[0175] 例如,假设X为1,Y为1;图像数据库包括表情图像a、表情图像b、表情图像c。表情图像a对应的第一向量距离为0.3,表情图像b对应的第一向量距离为0.5,表情图像c对应的第一向量距离为0.7。将这些第一向量距离从小到大排列,最小的第一向量距离(即排列在第一个的第一向量距离)为0.3,因此可以将表情图像a作为待处理信息匹配的目标表情图像。表情图像a对应的第二向量距离为0.2,表情图像b对应的第二向量距离为0.5,表情图像c对应的第二向量距离为0.6。将这些第二向量距离从小到大排列,最小的第二向量距离(即排列在第一个的第二向量距离)为0.2,因此可以将表情图像a作为待处理信息匹配的目标表情图像。此时,只有表情图像a作为待处理信息匹配的目标表情图像。

[0176] 又例如,假设X为1,Y为1;图像数据库包括表情图像a、表情图像b、表情图像c。表情图像a对应的第一向量距离为0.3,表情图像b对应的第一向量距离为0.5,表情图像c对应的第一向量距离为0.7。将这些第一向量距离从小到大排列,最小的第一向量距离(即排列在第一个的第一向量距离)为0.3,因此可以将表情图像a作为待处理信息匹配的目标表情图像。表情图像a对应的第二向量距离为0.4,表情图像b对应的第二向量距离为0.3,表情图像c对应的第二向量距离为0.6。将这些第二向量距离从小到大排列,最小的第二向量距离(即排列在第一个的第二向量距离)为0.3,因此可以将表情图像b作为待处理信息匹配的目标表情图像。此时,表情图像a和表情图像b都可以作为待处理信息匹配的目标表情图像。

[0177] 又例如,假设X为2,Y为2;图像数据库包括表情图像a、表情图像b、表情图像c。表情图像a对应的第一向量距离为0.3,表情图像b对应的第一向量距离为0.5,表情图像c对应的第一向量距离为0.7。将这些第一向量距离从小到大排列,排列在前2个的第一向量距离为0.3和0.5,因此可以将表情图像a和表情图像b作为待处理信息匹配的目标表情图像。表情图像a对应的第二向量距离为0.2,表情图像b对应的第二向量距离为0.5,表情图像c对应的第二向量距离为0.6。将这些第二向量距离从小到大排列,排列在前2个的第二向量距离为0.2和0.5,因此也是将表情图像a和表情图像b作为待处理信息匹配的目标表情图像。此时,表情图像a和表情图像b作为待处理信息匹配的目标表情图像。

[0178] 又例如,假设X为2,Y为2;图像数据库包括表情图像a、表情图像b、表情图像c。表情图像a对应的第一向量距离为0.3,表情图像b对应的第一向量距离为0.5,表情图像c对应的第一向量距离为0.7。将这些第一向量距离从小到大排列,排列在前2个的第一向量距离为0.3和0.5,因此可以将表情图像a和表情图像b作为待处理信息匹配的目标表情图像。表情图像a对应的第二向量距离为0.5,表情图像b对应的第二向量距离为0.3,表情图像c对应的第二向量距离为0.4。将这些第二向量距离从小到大排列,排列在前2个的第二向量距离为0.3和0.4,因此可以将表情图像b和表情图像c作为待处理信息匹配的目标表情图像。此时,表情图像a、表情图像b、表情图像c都可以作为待处理信息匹配的目标表情图像。

[0179] 方式二:对第一向量距离和第二向量距离进行计算处理,得到每个表情图像对应的第三向量距离;按照从小到大的顺序排列第三向量距离,将前N个第三向量距离对应的表情图像确定为该待处理信息匹配的目标表情图像;该N为正整数。

[0180] 例如,假设N为1,图像数据库包括表情图像a、表情图像b、表情图像c;第一向量距离对应的权重为0.4,第二向量距离对应的权重为0.6。表情图像a对应的第一向量距离为0.3,表情图像b对应的第一向量距离为0.5,表情图像c对应的第一向量距离为0.7。表情图像a对应的第二向量距离为0.2,表情图像b对应的第二向量距离为0.5,表情图像c对应的第二向量距离为0.6。因此,计算得到:表情图像a对应第三向量距离为0.24,表情图像b对应第三向量距离为0.5,表情图像c对应第三向量距离为0.64。将这些第三向量距离从小到大排列,最小的第三向量距离(即排列在第一个的第三向量距离)为0.24,因此可以将表情图像a作为待处理信息匹配的目标表情图像。

[0181] 又例如,假设N为2,图像数据库包括表情图像a、表情图像b、表情图像c;第一向量距离对应的权重为0.4,第二向量距离对应的权重为0.6。表情图像a对应的第一向量距离为0.3,表情图像b对应的第一向量距离为0.5,表情图像c对应的第一向量距离为0.7。表情图像a对应的第二向量距离为0.2,表情图像b对应的第二向量距离为0.5,表情图像c对应的第二向量距离为0.6。因此,计算得到:表情图像a对应第三向量距离为0.24,表情图像b对应第三向量距离为0.5,表情图像c对应第三向量距离为0.64。将这些第三向量距离从小到大排列,排列在前2个的第三向量距离为0.24和0.5,因此可以将表情图像a和表情图像b作为待处理信息匹配的目标表情图像。

[0182] 当然,电子设备基于第一向量距离和第二向量距离确定待处理信息匹配的目标表情图像的方式,也可以采用其他策略,在此不做限定。

[0183] 在一种可能的实现方式中,电子设备从图像数据库中查找该待处理信息匹配的目标表情图像时,具体的实现方式可以是:调用第一对齐网络对第一文本特征向量进行处理,得到第四内容特征向量;调用第二对齐网络对第一文本特征向量进行处理,得到第五意图特征向量;确定图像数据库中每个表情图像对应的第一内容特征向量与第四内容特征向量之间的第三向量距离,以及确定图像数据库中每个表情图像对应的第一意图特征向量与第五意图特征向量之间的第四向量距离;基于第三向量距离和第四向量距离确定待处理信息匹配的目标表情图像。可以理解为,电子设备也需要利用第一对齐网络提取第一文本信息的内容特征向量(即第四内容特征向量),利用第二对齐网络提取第一文本信息的意图特征向量(第五意图特征向量),针对性地从语义内容和表达意图两个维度查找与第一文本信息匹配的目标表情图像,提高目标表情图像的准确性。另外,基于第三向量距离和第四向量距

离确定待处理信息匹配的目标表情图像的具体实现方式可以参考上述基于第一向量距离和第二向量距离确定待处理信息匹配的目标表情图像的具体实现方式,在此不做赘述。

[0184] 在一种可能的实现方式中,该方法还包括:在第一界面中显示待处理信息匹配的目标表情图像,该第一界面为以下界面中的任意一项:图像搜索界面、输入法界面、表情图像推荐界面。也就是说,电子设备在表情图像数据库中查找到待处理信息匹配的目标表情图像后,可以在第一界面中进行展示,以供用户进行选择。基于该方式,能够提高图像查找的可视化,提高目标表情图像展示的灵活性。

[0185] 如图10A所示,该表情图像搜索界面包括表情图像搜索框和搜索按钮,假设用户在表情图像搜索框中输入了“疑惑”,当用户点击了搜索按钮时,电子设备便执行上述描述的图像查找方法,在表情图像数据库中查找到与“疑惑”相匹配的目标表情图像后,便将这些目标表情图像展示在表情图像搜索界面,如图11A所示,以供用户进行选择。

[0186] 如图10B所示,假设用户在聊天框中编辑文字时输入了“疑惑”,该“疑惑”是用户未发送的文字,此时电子设备可以执行上述描述的图像查找方法,在表情图像数据库中查找到与“疑惑”相匹配的目标表情图像,然后弹出表情图像推荐界面,在该表情图像推荐界面中展示这些目标表情图像,如图11B所示,以供用户进行选择,有利于提高对话效率。

[0187] 如图10C所示,假设用户在输入法界面中,点击表情按钮,显示表情界面;进一步点击显示表情界面中的表情图选项进入表情图像搜索界面(可以认为是本地TAB页);用户可以在表情图像搜索界面的表情图像搜索栏中输入需要搜索的文字,即“疑惑”;当用户点击了搜索按钮时,电子设备便执行上述描述的图像查找方法,在表情图像数据库中查找到与“疑惑”相匹配的目标表情图像后,便将这些目标表情图像展示在输入法界面,如图11C所示,以供用户进行选择。

[0188] 当然,针对目标表情图像的展示顺序,电子设备可以根据用户的历史习惯或历史行为进行展示,例如按照在预设时间段内用户历史使用频率从多到少的顺序进行展示,也可以按照随机顺序进行展示,也可以按照用户添加表情图像的时间顺序(如从近到远)进行展示,在此不做限定。

[0189] 在一种可能的实现方式中,该方法还包括:显示表情图像添加界面,该表情图像添加界面包括表情图像框、自定义标签框和保存选项;在检测到用户针对该保存选项的触发操作时,保存该表情图像框中添加的第一表情图像以及该自定义标签框中添加的第一表情图像的第一标签;若该待处理信息与该第一标签相同,则将该第一表情图像确定为该目标表情图像。可以理解为,用户可以预先对自己收藏的表情图像添加标签并保存,当用户需要查找表情图像时,电子设备除了应用上述步骤S403查找到目标表情图像外,如果用户输入的待处理信息与用户之前保存的第一标签相同,可以将该第一标签对应的第一表情图像作为目标表情图像,可以一起展示给用户进行选择。基于该方式,能够提高查找目标表情图像的效率和目标表情图像的准确性,有利于提高线上聊天对话的效率。

[0190] 如图12所示,该表情图像添加界面包括表情图像框、自定义标签框和保存选项。用户可以在表情图像框中添加第一表情图像,在自定义标签框中添加第一表情图像的第一标签。当用户点击了保存选项时,电子设备保存该表情图像框中添加的第一表情图像以及该自定义标签框中添加的第一表情图像的第一标签。在用户查找表情图像的过程中,如果用户输入的待处理信息与用户曾经保存的第一标签相同时,电子设备也可以直接将该第一

表情图像作为目标表情图像。

[0191] 可见,基于上述所描述的方法,电子设备获取表情图像数据库;然后确定出待处理信息的第一文本特征向量,这里的待处理信息可以认为是用户输入的文本信息、用户搜索的文本信息等;最后可以从图像数据库中查找该待处理信息匹配的目标表情图像,并且展示该目标表情图像供用户选择。其中,该目标表情图像对应的第一内容特征向量和第一意图特征向量与第一文本特征向量相匹配;可以理解为,从图像数据库中查找该待处理信息匹配的目标表情图像是通过表情图像的情绪意图和语义内容的多维度快速检索实现的,这样能够更加快速精确地查找到用户需要发送的表情图像,有利于提高线上聊天对话的效率。

[0192] 请参见图13,图13示出了本申请实施例的一种图像查找装置1300的结构示意图。图13所示的图像查找装置可以电子设备,也可以是电子设备中的装置,或者是能够和电子设备匹配使用的装置。图13所示的图像查找装置可以包括获取单元1301、确定单元1302和查找单元1303。其中:

[0193] 获取单元1301,用于获取表情图像数据库,该表情图像数据库包括多个表情图像、多个表情图像中每个表情图像对应的第一内容特征向量和第一意图特征向量;该第一内容特征向量表示描述表情图像的内容的特征向量,该第一意图特征向量表示描述表情图像的表达意图的特征向量;

[0194] 确定单元1302,用于确定待处理信息的第一文本特征向量;

[0195] 查找单元1303,用于从该图像数据库中查找该待处理信息匹配的目标表情图像,该目标表情图像对应的第一内容特征向量和第一意图特征向量与该第一文本特征向量相匹配。

[0196] 在一种可能的实现方式中,获取单元1301,在获取表情图像数据库时,具体用于:获取用户在本地收藏的多个表情图像;调用图像编码器对多个表情图像中每个表情图像进行编码处理,得到每个表情图像对应的第一图像特征向量;调用对齐网络对第一图像特征向量和提示词进行处理,得到每个表情图像对应的第一内容特征向量和第一意图特征向量;该提示词用于提示对齐网络描述表情图像的内容和表达意图;利用多个表情图像、每个表情图像对应的第一内容特征向量和第一意图特征向量构建表情图像数据库。

[0197] 在一种可能的实现方式中,该对齐网络包括第一对齐网络和第二对齐网络,该提示词包括第一提示词和第二提示词,获取单元1301,在调用对齐网络对该第一图像特征向量和提示词进行处理,得到每个表情图像对应的第一内容特征向量和第一意图特征向量时,具体用于:调用该第一对齐网络对第一图像特征向量和第一提示词进行处理,得到每个表情图像对应的第一内容特征向量;该第一提示词用于提示第一对齐网络描述表情图像的内容;调用该第二对齐网络对第一图像特征向量和第二提示词进行处理,得到每个表情图像对应的第一意图特征向量;该第二提示词用于提示第二对齐网络描述表情图像的表达意图。

[0198] 在一种可能的实现方式中,查找单元1303,在从该图像数据库中查找该待处理信息匹配的目标表情图像时,具体用于:确定该图像数据库中每个表情图像对应的第一内容特征向量与该第一文本特征向量之间的第一向量距离,以及确定该图像数据库中每个表情图像对应的第一意图特征向量与该第一文本特征向量之间的第二向量距离;基于该第一向

量距离和该第二向量距离确定该待处理信息匹配的目标表情图像。

[0199] 在一种可能的实现方式中,查找单元1303,在基于该第一向量距离和该第二向量距离确定该待处理信息匹配的目标表情图像时,具体用于:对该第一向量距离和该第二向量距离进行计算处理,得到每个表情图像对应的第三向量距离;按照从小到大的顺序排列第三向量距离,将前N个第三向量距离对应的表情图像确定为该待处理信息匹配的目标表情图像;该N为正整数。

[0200] 在一种可能的实现方式中,查找单元1303,在从该图像数据库中查找该待处理信息匹配的目标表情图像时,具体用于:调用第一对齐网络对第一文本特征向量进行处理,得到第四内容特征向量;调用第二对齐网络对第一文本特征向量进行处理,得到第五意图特征向量;确定图像数据库中每个表情图像对应的第一内容特征向量与第四内容特征向量之间的第三向量距离,以及确定图像数据库中每个表情图像对应的第一意图特征向量与第五意图特征向量之间的第四向量距离;基于该第三向量距离和该第四向量距离确定待处理信息匹配的目标表情图像。

[0201] 在一种可能的实现方式中,该装置还包括显示单元,该显示单元,用于在第一界面中显示该待处理信息匹配的目标表情图像,该第一界面为以下界面中的任意一项:图像搜索界面、输入法界面、表情图像推荐界面。

[0202] 在一种可能的实现方式中,该显示单元,还用于:显示表情图像添加界面,该表情图像添加界面包括表情图像框、自定义标签框和保存选项;该装置还包括处理单元,该处理单元,用于在检测到用户针对该保存选项的触发操作时,保存该表情图像框中添加的第一表情图像以及该自定义标签框中添加的该第一表情图像的第一标签;若该待处理信息与该第一标签相同,则将该第一表情图像确定为该目标表情图像。

[0203] 对于图像查找装置可以是芯片或芯片系统的情况,可参见图14所示的芯片的结构示意图。图14所示的芯片1400包括处理器1401、接口1402。可选的,还可包括存储器1403。其中,处理器1401的数量可以是一个或多个,接口1402的数量可以是多个。

[0204] 对于芯片用于实现本申请实施例中电子设备的情况:

[0205] 所述接口1402,用于接收或输出信号;

[0206] 所述处理器1401,用于执行电子设备的数据处理操作。

[0207] 可以理解的是,本申请实施例中的一些可选的特征,在某些场景下,可以不依赖于其他特征,比如其当前所基于的方案,而独立实施,解决相应的技术问题,达到相应的效果,也可以在某些场景下,依据需求与其他特征进行结合。相应地,本申请实施例中给出的图像查找装置也可以相应的实现这些特征或功能,在此不予赘述。

[0208] 应理解,本申请实施例中的处理器可以是一种集成电路芯片,具有信号的处理能力。在实现过程中,上述方法实施例的各步骤可以通过处理器中的硬件的集成逻辑电路或者软件形式的指令完成。上述的处理器可以是通用处理器、数字信号处理器(digital signal processor,DSP)、专用集成电路(application specific integrated circuit,ASIC)、现场可编程门阵列(field programmable gate array,FPGA)或者其它可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件。

[0209] 可以理解,本申请实施例中的存储器可以是易失性存储器或非易失性存储器,或可包括易失性和非易失性存储器两者。其中,非易失性存储器可以是只读存储器(read-

only memory,ROM)、可编程只读存储器(programmable ROM,PROM)、可擦除可编程只读存储器(erasable PROM,EPROM)、电可擦除可编程只读存储器(electrically EPROM,EEPROM)或闪存。易失性存储器可以是随机存取存储器(random access memory,RAM),其用作外部高速缓存。通过示例性但不是限制性说明,许多形式的RAM可用,例如静态随机存取存储器(static RAM,SRAM)、动态随机存取存储器(dynamic RAM,DRAM)、同步动态随机存取存储器(synchronous DRAM,SDRAM)、双倍数据速率同步动态随机存取存储器(double data rate SDRAM,DDR SDRAM)、增强型同步动态随机存取存储器(enhanced SDRAM,ESDRAM)、同步连接动态随机存取存储器(synchlink DRAM,SLDRAM)和直接内存总线随机存取存储器(direct rambus RAM,DR RAM)。应注意,本文描述的系统和方法的存储器旨在包括但不限于这些和任意其它适合类型的存储器。

[0210] 本申请还提供了一种计算机可读存储介质,该计算机可读存储介质中存储有计算机程序,该计算机程序包括程序指令,当该程序指令在图像查找装置上运行时,实现上述任一方法实施例的功能。

[0211] 本申请还提供了一种计算机程序产品,当计算机程序产品在计算机上运行时,使得计算机实现上述任一方法实施例的功能。

[0212] 上述实施例中,可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。当使用软件实现时,可以全部或部分地以计算机程序产品的形式实现。所述计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行所述计算机指令时,全部或部分地产生按照本申请实施例所述的流程或功能。所述计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。所述计算机指令可以存储在计算机可读存储介质中,或者从一个计算机可读存储介质向另一个计算机可读存储介质传输,例如,所述计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线(例如同轴电缆、光纤、数字用户线(digital subscriber line,DSL))或无线(例如红外、无线、微波等)方式向另一个网站站点、计算机、服务器或数据中心进行传输。所述计算机可读存储介质可以是计算机能够存取的任何可用介质或者是包含一个或多个可用介质集成的服务器、数据中心等数据存储设备。所述可用介质可以是磁性介质(例如,软盘、硬盘、磁带)、光介质(例如,高密度数字视频光盘(digital video disc,DVD))、或者半导体介质(例如,固态硬盘(solid state disk,SSD))等。

[0213] 以上所述,仅为本申请的具体实施方式,但本申请的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本申请揭露的技术范围内,可轻易想到变化或替换,都应涵盖在本申请的保护范围之内。因此,本申请的保护范围应以权利要求的保护范围为准。

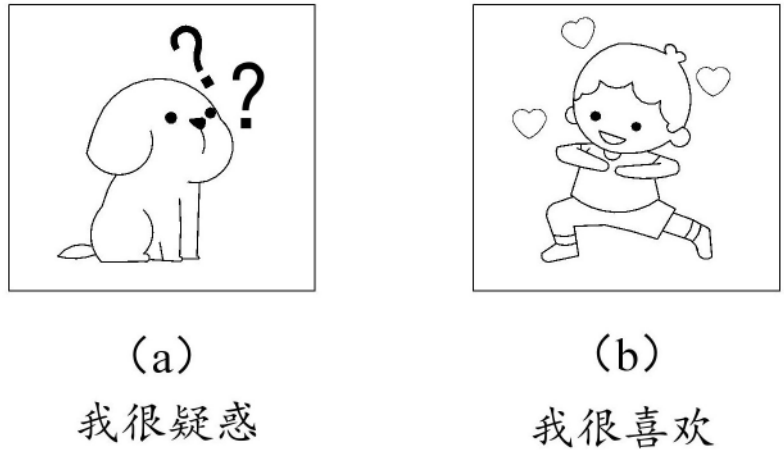


图1A

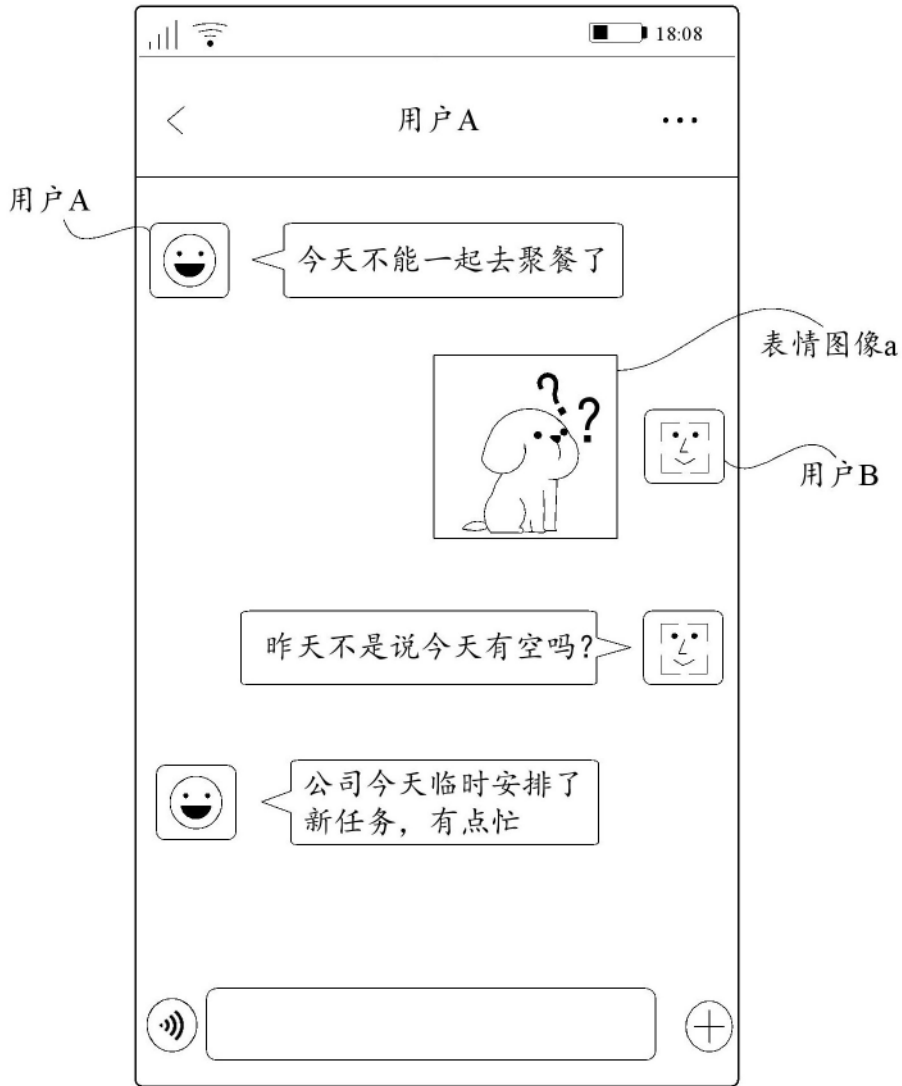


图1B

电子设备100

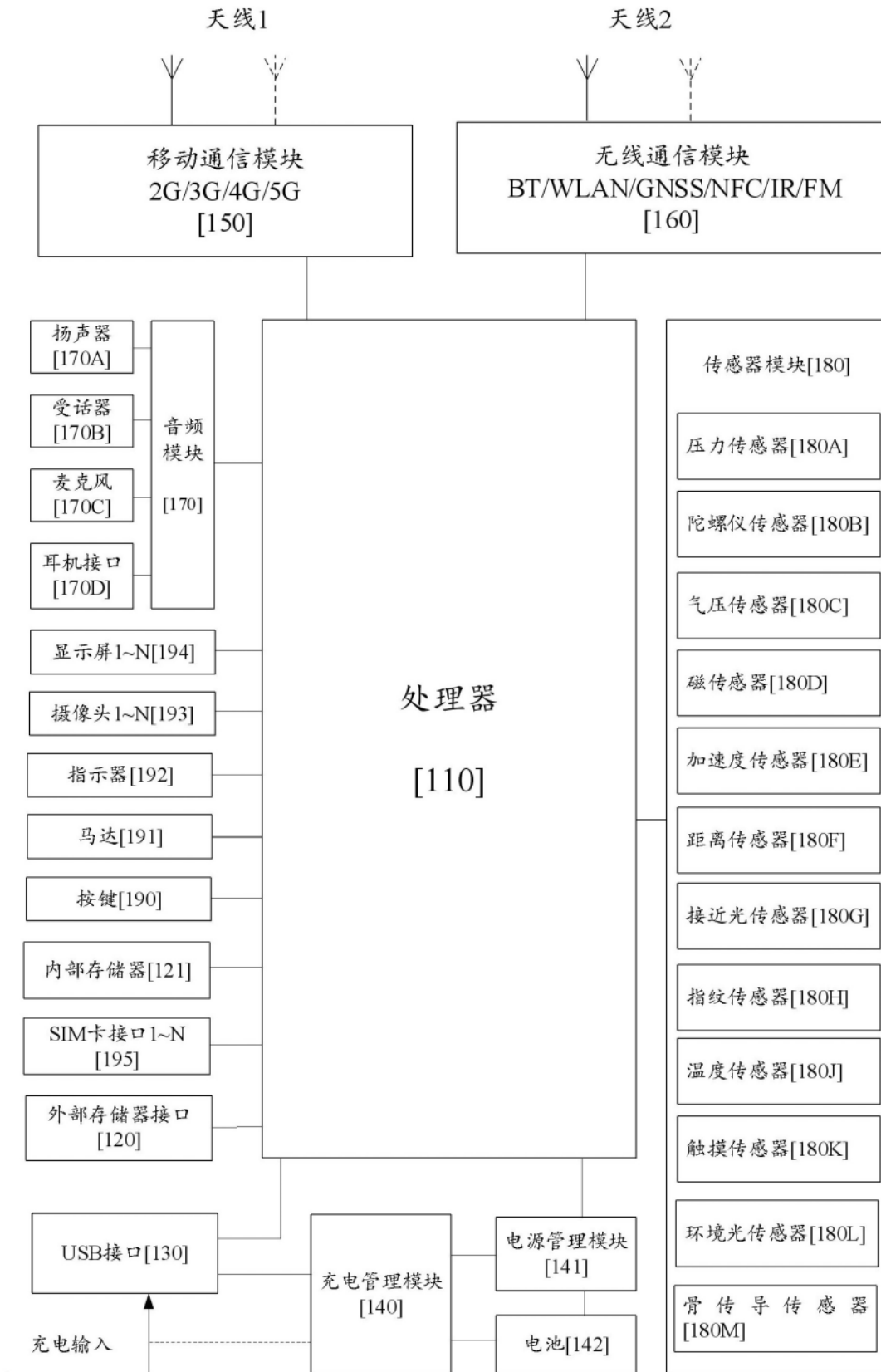


图2

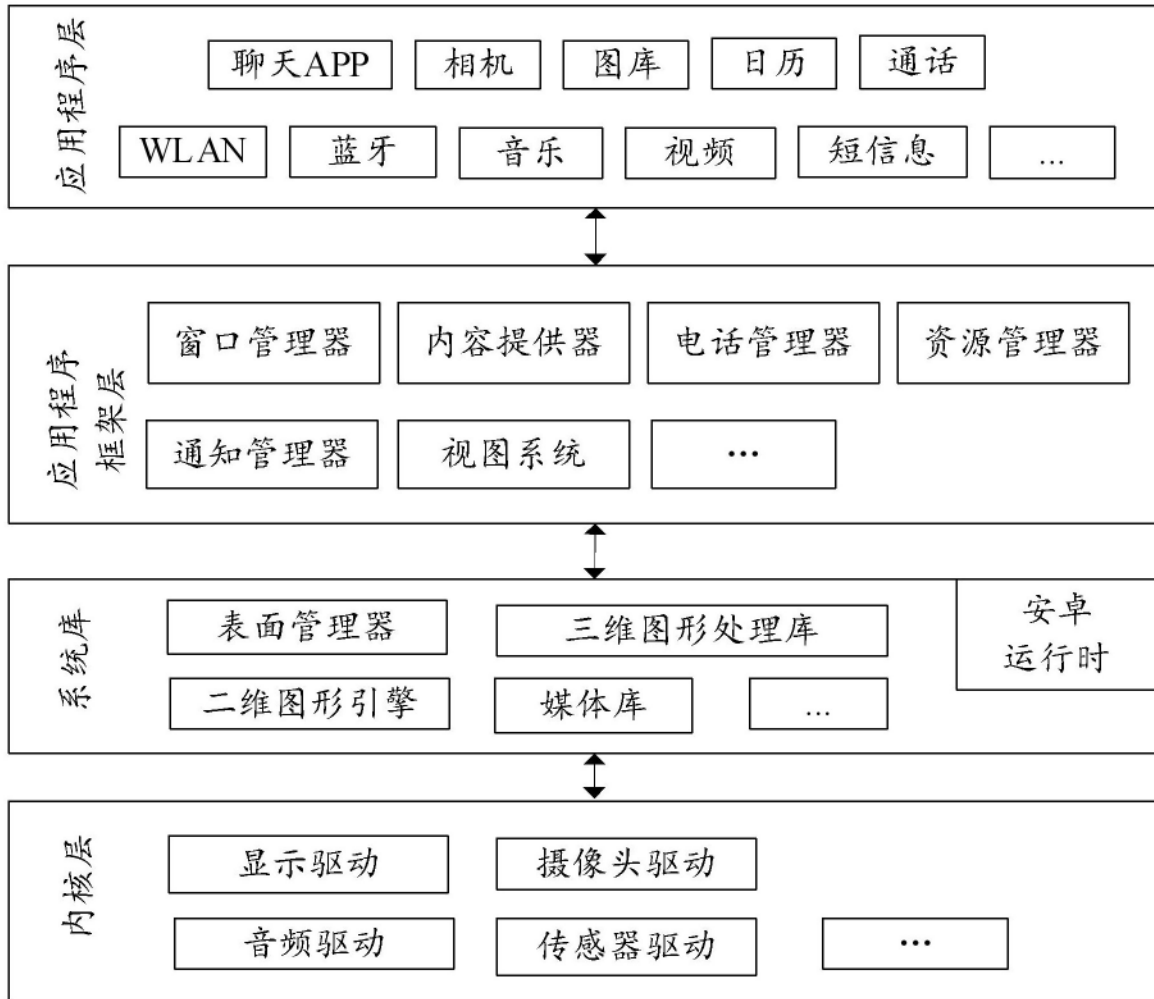


图3

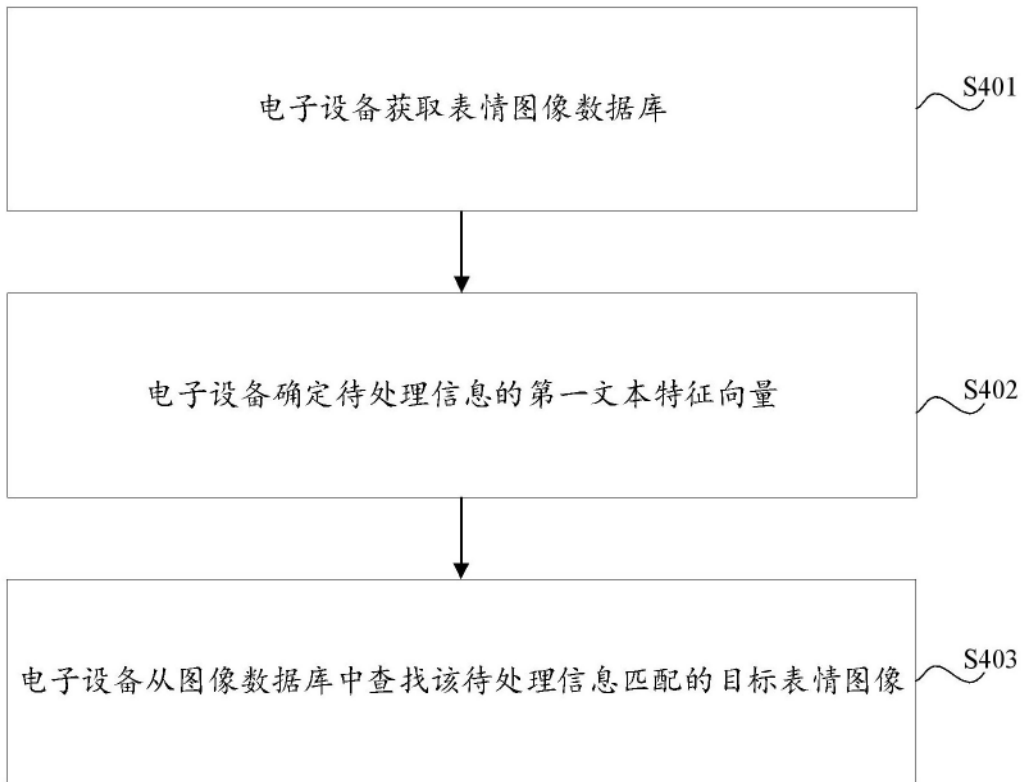


图4

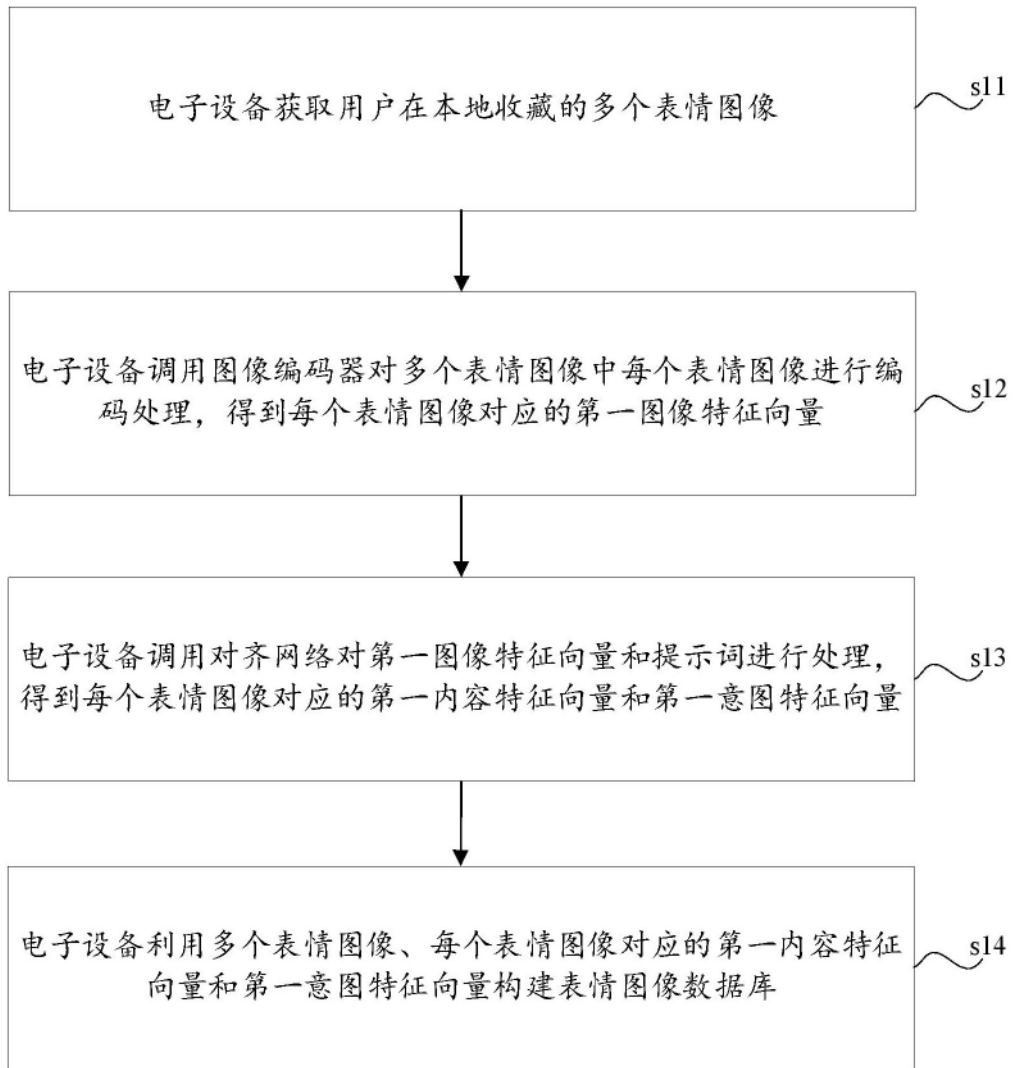
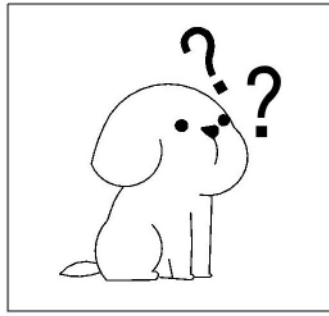


图5



图6



(a)

一只白色的小狗，侧坐在地上，眼睛望着前方，小狗的头上有问号的标记

(b)

图7A

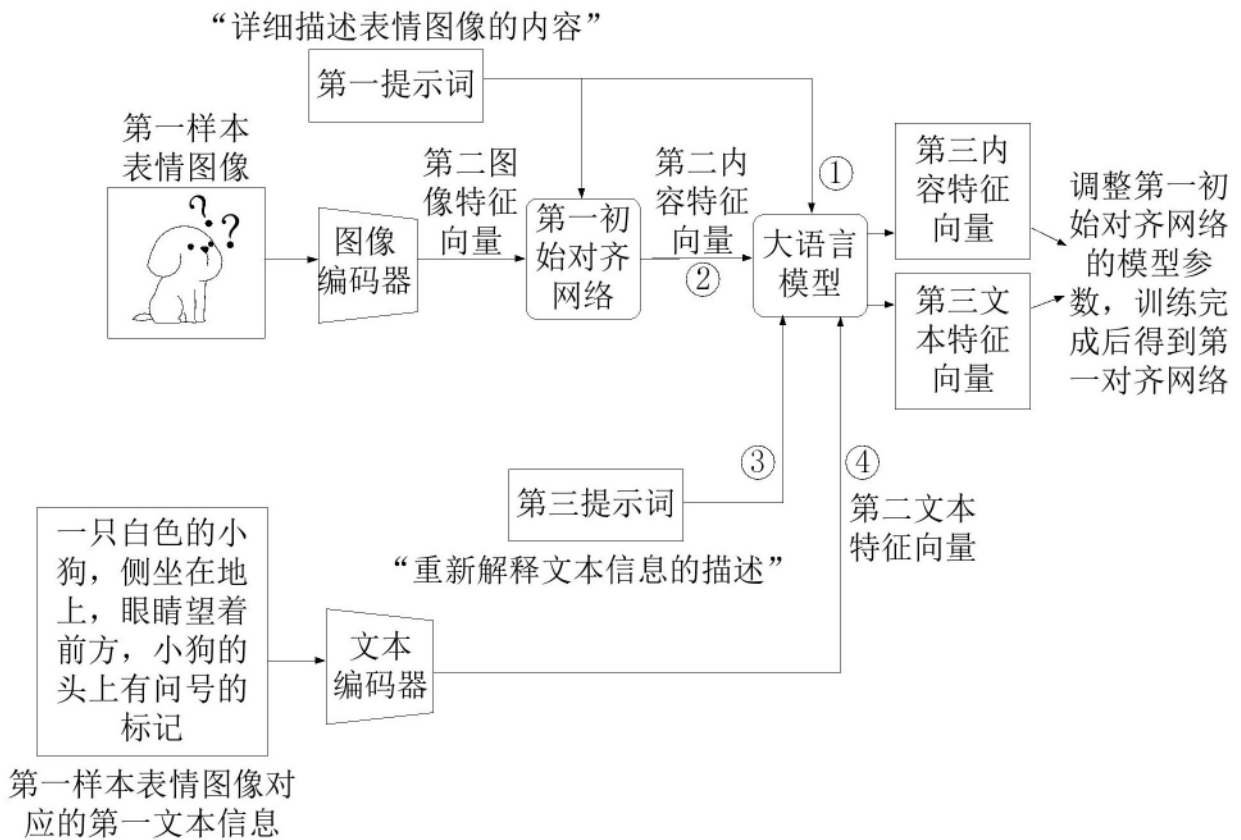


图7B



图8

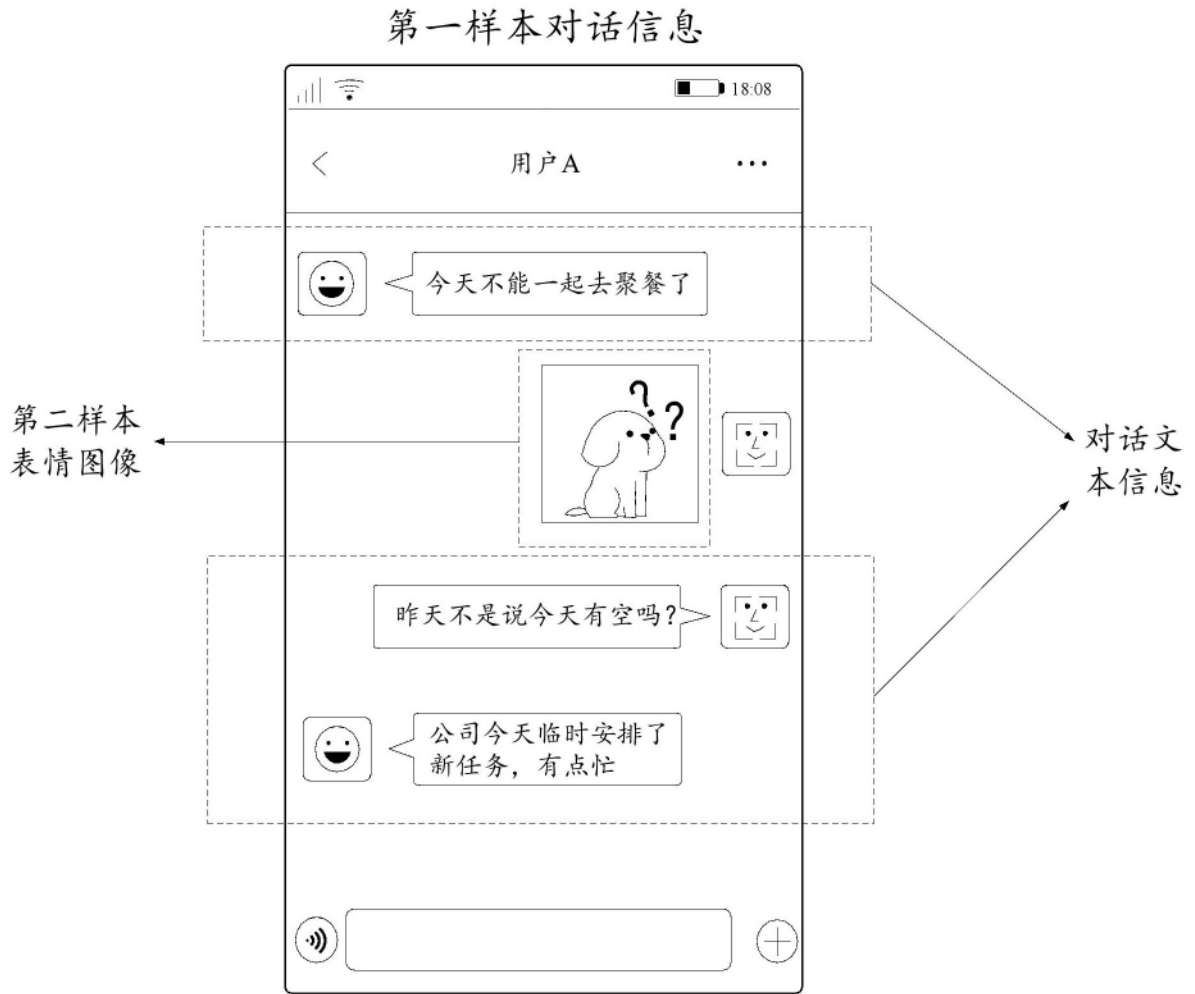


图9A

第二样本对话信息

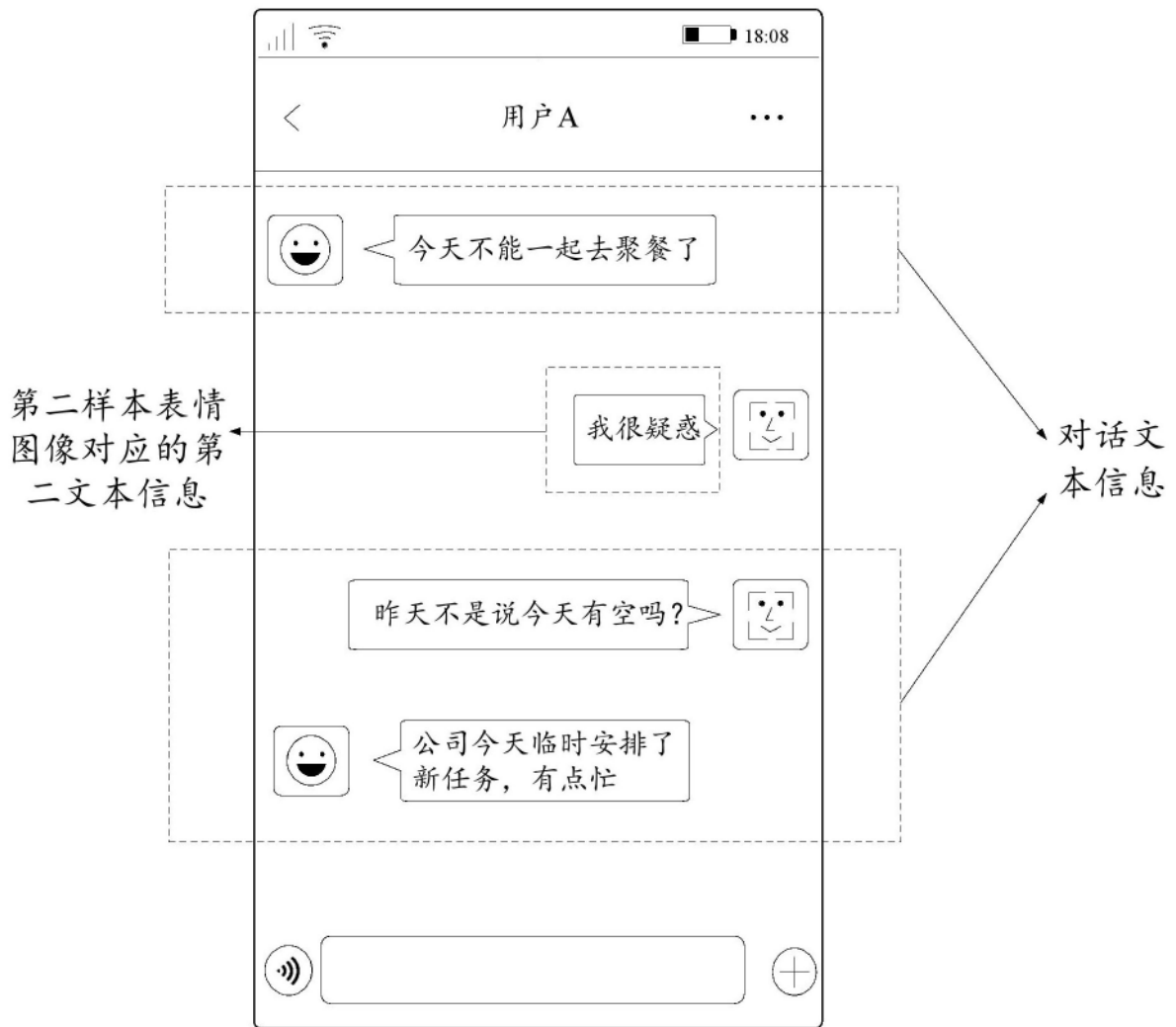


图9B

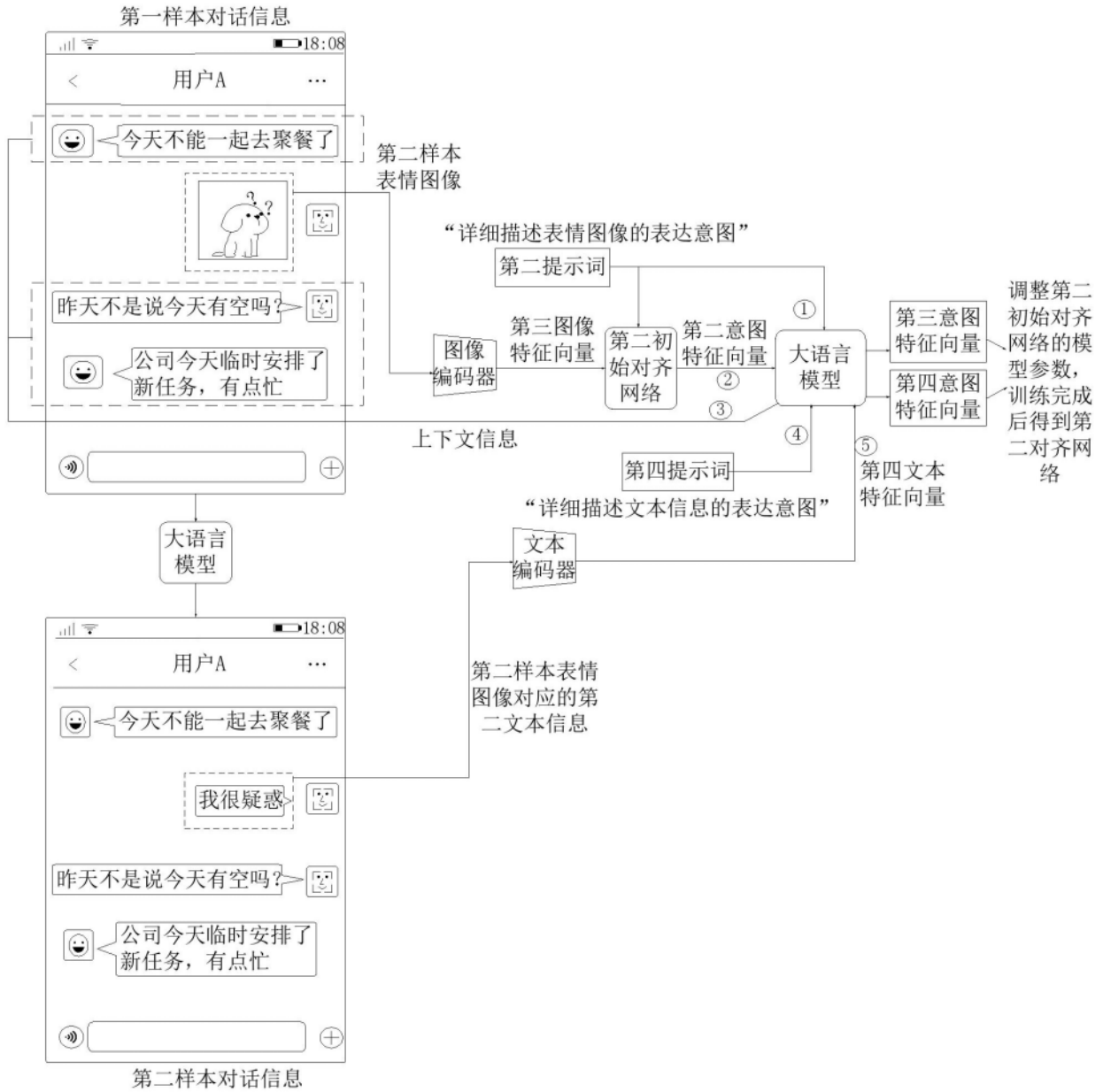


图9C

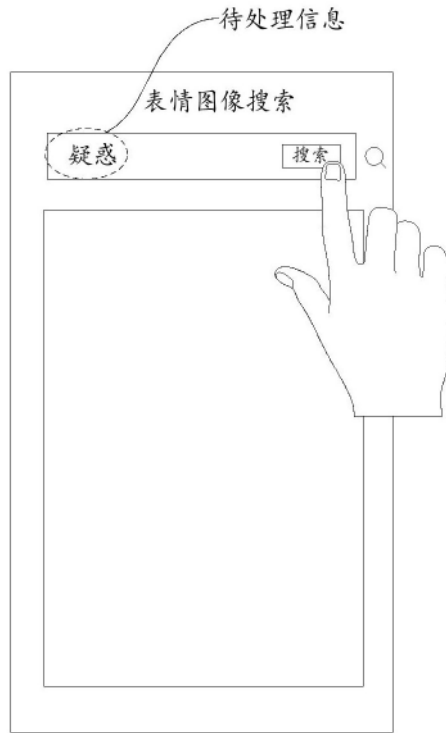


图10A

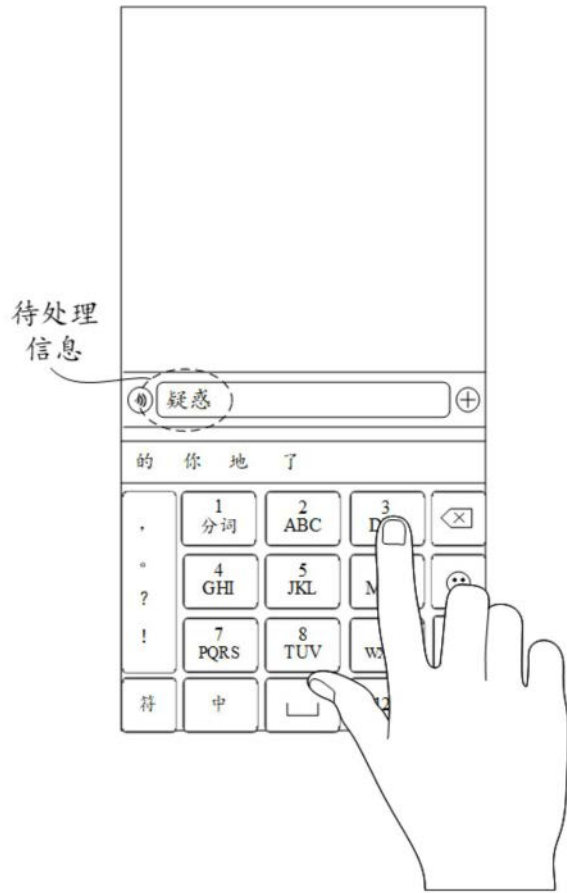


图10B

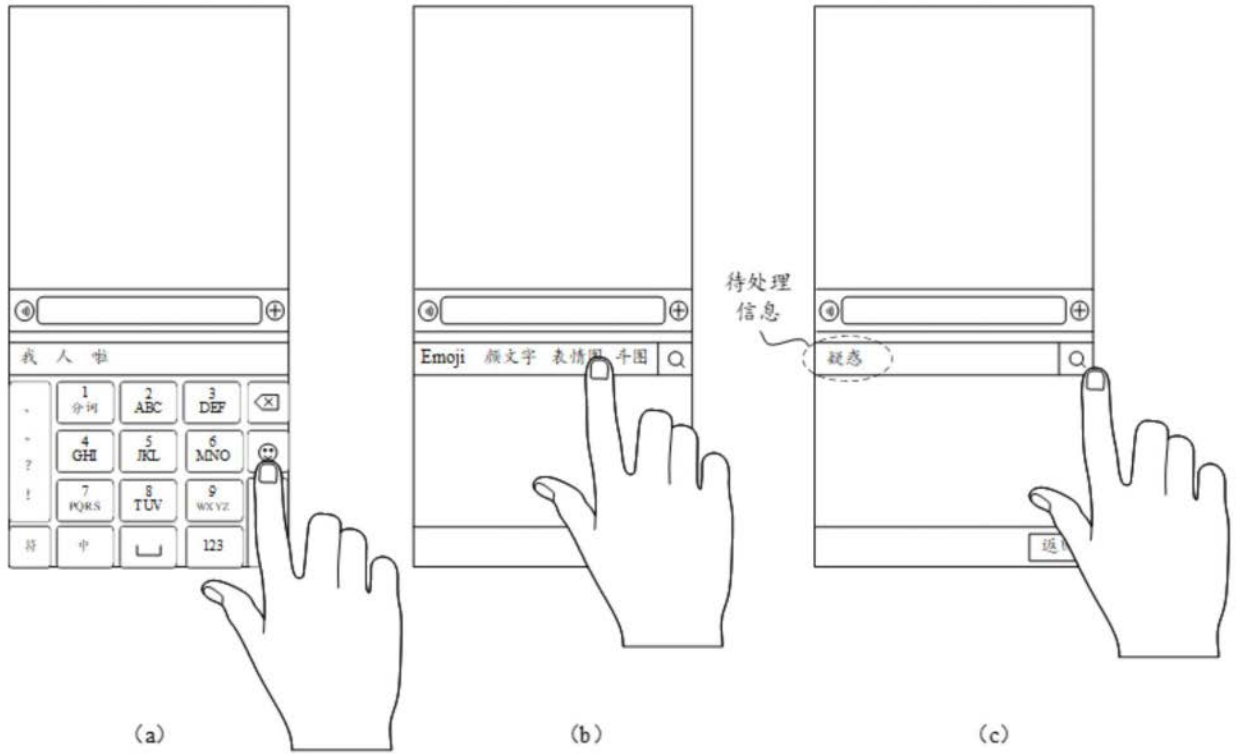


图10C

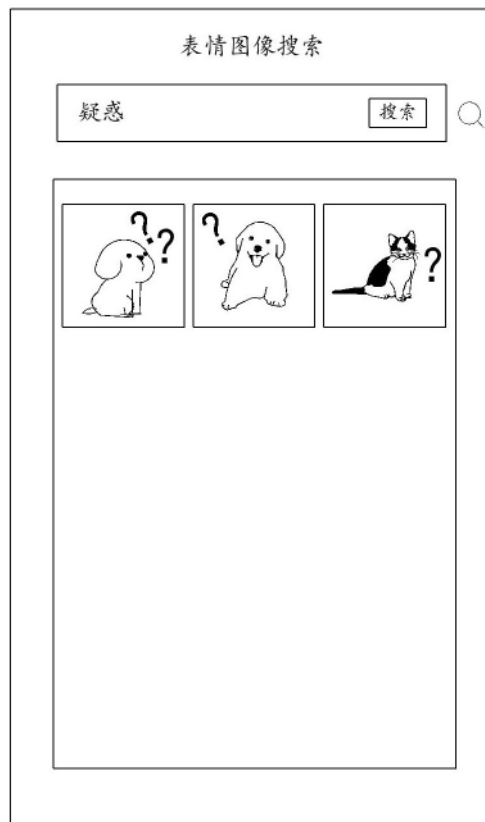


图11A



图11B

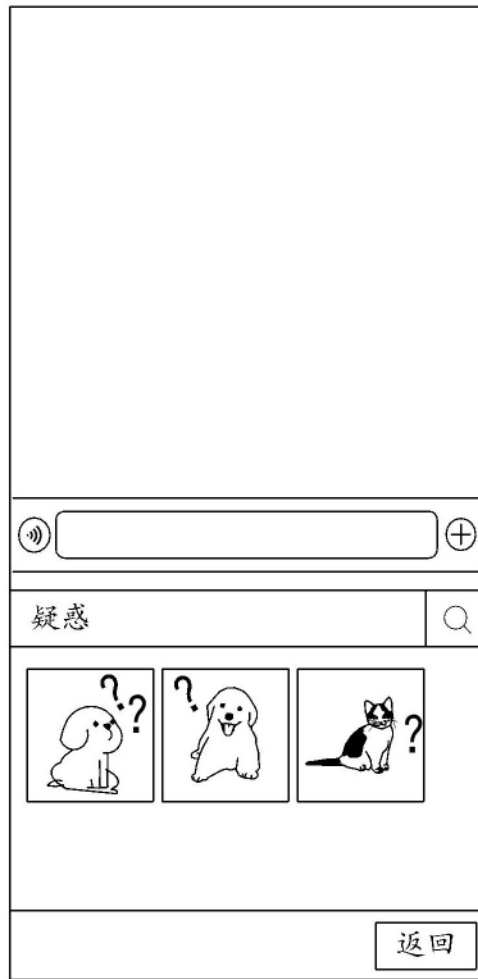


图11C

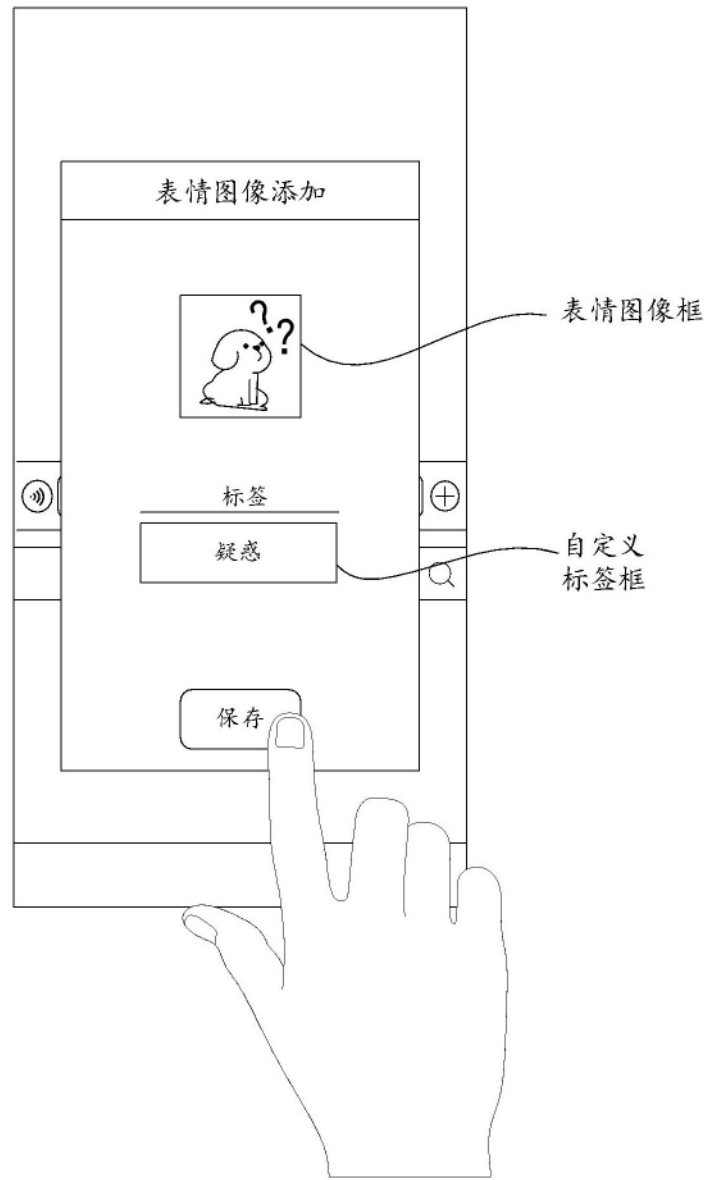


图12

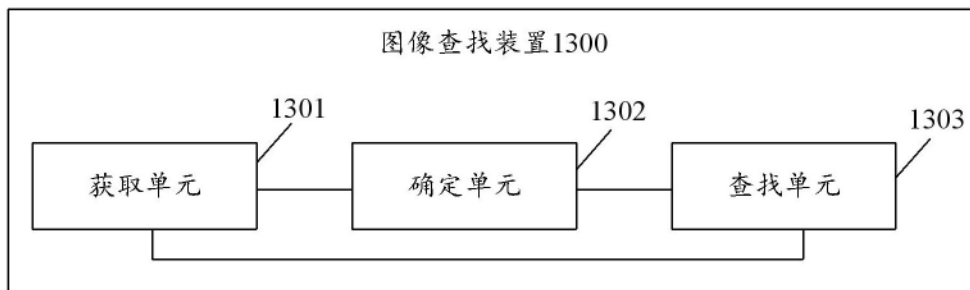


图13

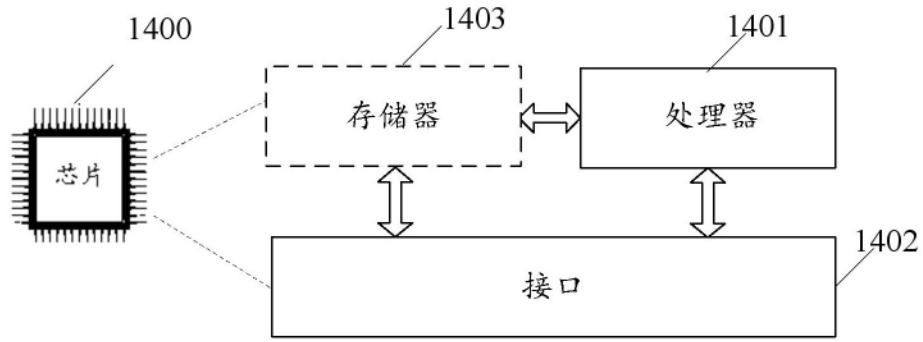


图14