

(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 25 日 (25.06.2020)



(10) 国际公布号
WO 2020/125386 A1

(51) 国际专利分类号:
G06K 9/00 (2006.01)

(21) 国际申请号: PCT/CN2019/122313

(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:
201811553986.X 2018 年 12 月 18 日 (18.12.2018) 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 518052 (CN)。

(72) 发明人: 郑子奇 (ZHENG, Ziqi); 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 518052 (CN)。 徐国强 (XU, Guoqiang); 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 518052 (CN)。 邱寒 (QIU, Han); 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 518052 (CN)。

(74) 代理人: 广州华进联合专利商标代理有限公司 (ADVANCE CHINA IP LAW OFFICE); 中国广东省广州市天河区珠江东路 6 号 4501 房 (部位: 自编 01-03 和 08-12 单元) (仅限办公用途), Guangdong 510623 (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,

(54) Title: EXPRESSION RECOGNITION METHOD AND APPARATUS, COMPUTER DEVICE, AND STORAGE MEDIUM

(54) 发明名称: 表情识别方法、装置、计算机设备和存储介质

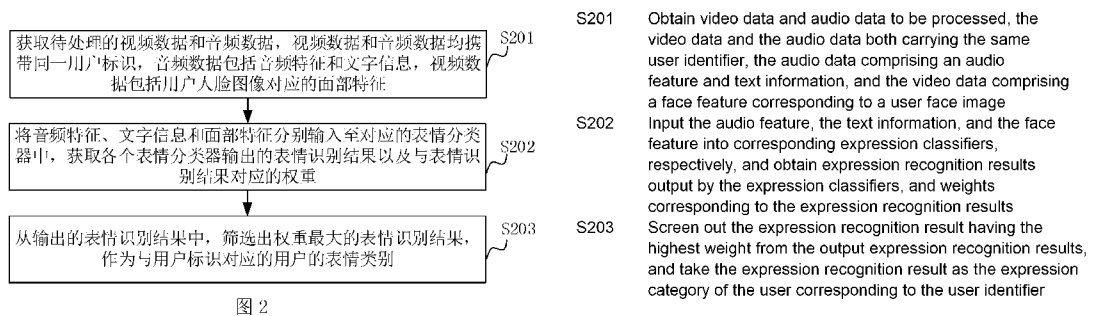


图 2

(57) Abstract: An expression recognition method, comprising: obtaining video data and audio data to be processed, the video data and the audio data both carrying the same user identifier, the audio data comprising an audio feature and text information, and the video data comprising a face feature corresponding to a user face image; inputting the audio feature, the text information, and the face feature into corresponding expression classifiers, respectively, and obtaining expression recognition results output by the expression classifiers, and weights corresponding to the expression recognition results; and screening out the expression recognition result having the highest weight from the output expression recognition results, and taking the expression recognition result as the expression category of the user corresponding to the user identifier.

(57) 摘要: 一种表情识别方法, 包括: 获取待处理的视频数据和音频数据, 所述视频数据和所述音频数据均携带同一用户标识, 所述音频数据包括音频特征和文字信息, 所述视频数据包括用户人脸图像对应的面部特征; 将所述音频特征、文字信息和面部特征分别输入至对应的表情分类器中, 获取各个所述表情分类器输出的表情识别结果以及与所述表情识别结果对应的权重; 从所述输出的表情识别结果中, 筛选出权重最大的表情识别结果, 作为与所述用户标识对应的用户的表情类别。

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, 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))。

表情识别方法、装置、计算机设备和存储介质

本申请要求于 2018 年 12 月 18 日提交中国专利局，申请号为 201811553986.X，申请名称为“表情识别方法、装置、计算机设备和存储介质”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。

5

技术领域

本申请涉及一种表情识别方法、装置、计算机设备和存储介质。

背景技术

10 人类表情蕴含着丰富的情绪，其传递的信息非常重要，故在用户视频面试、面审等既有视频录制也有音频录制的场景中，往往需要对用户表情进行分析，以了解用户真实的情绪。因此，表情识别的准确性显得越来越重要。

然而，传统的表情识别方法主要是基于录制的图像和视频来对用户的表情类型进行判断，因此对于图片的质量和拍摄的场景都有一定的要求。但是，发明人意识到，若图片的质量和拍摄的场景达不到要求，而且在用户面部动作不够丰富时，很容易出现漏判的情况，
15 从而造成表情识别的准确率较低。

发明内容

根据本申请公开的各种实施例，提供一种表情识别方法、装置、计算机设备和存储介
20 质。

一种表情识别方法包括：

获取待处理的视频数据和音频数据，所述视频数据和所述音频数据均携带同一用户标识，所述音频数据包括音频特征和文字信息，所述视频数据包括用户人脸图像对应的面部特征；

25 将所述音频特征、文字信息和面部特征分别输入至对应的表情分类器中，获取各个所述表情分类器输出的表情识别结果以及与所述表情识别结果对应的权重；及

从所述输出的表情识别结果中，筛选出权重最大的表情识别结果，作为与所述用户标识对应的用户的表情类别。

一种表情识别装置包括：

30 数据获取模块，用于获取待处理的视频数据和音频数据，所述视频数据和所述音频数据均携带同一用户标识，所述音频数据包括音频特征和文字信息，所述视频数据包括用户人脸图像对应的面部特征；

表情获取模块，用于将所述音频特征、文字信息和面部特征分别输入至对应的表情分类器中，获取各个所述表情分类器输出的表情识别结果以及与所述表情识别结果对应的权

重；及

表情筛选模块，用于从所述输出的表情识别结果中，筛选出权重最大的表情识别结果，作为与所述用户标识对应的用户的表情类别。

一种计算机设备，包括存储器和一个或多个处理器，所述存储器中储存有计算机可读指令，所述计算机可读指令被所述处理器执行时，使得所述一个或多个处理器执行以下步骤：

获取待处理的视频数据和音频数据，所述视频数据和所述音频数据均携带同一用户标识，所述音频数据包括音频特征和文字信息，所述视频数据包括用户人脸图像对应的面部特征；

将所述音频特征、文字信息和面部特征分别输入至对应的表情分类器中，获取各个所述表情分类器输出的表情识别结果以及与所述表情识别结果对应的权重；及

从所述输出的表情识别结果中，筛选出权重最大的表情识别结果，作为与所述用户标识对应的用户的表情类别。

一个或多个存储有计算机可读指令的非易失性计算机可读存储介质，计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器执行以下步骤：

获取待处理的视频数据和音频数据，所述视频数据和所述音频数据均携带同一用户标识，所述音频数据包括音频特征和文字信息，所述视频数据包括用户人脸图像对应的面部特征；

将所述音频特征、文字信息和面部特征分别输入至对应的表情分类器中，获取各个所述表情分类器输出的表情识别结果以及与所述表情识别结果对应的权重；及

从所述输出的表情识别结果中，筛选出权重最大的表情识别结果，作为与所述用户标识对应的用户的表情类别。

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。本申请的其它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

为了更清楚地说明本申请实施例中的技术方案，下面将对实施例中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其它的附图。

图 1 为根据一个或多个实施例中表情识别方法的应用场景图；

图 2 为根据一个或多个实施例中表情识别方法的流程示意图；

图 3 为根据一个实施例中获取各个表情分类器输出的表情识别结果的步骤的流程示意图；

图 4 为另一个实施例中表情识别方法的流程示意图；

图 5 为根据一个或多个实施例中表情识别装置的框图；

图 6 为根据一个或多个实施例中计算机设备的框图。

具体实施方式

为了使本申请的技术方案及优点更加清楚明白，以下结合附图及实施例，对本申请进行进一步详细说明。应当理解，此处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。

本申请提供的表情识别方法，可以应用于如图 1 所示的应用环境中。终端 110 通过网络与服务器 120 进行通信。终端 110 中安装了视频录制系统，视频录制系统能够将录制的视频数据和音频数据上传至服务器 120；视频数据和音频数据携带同一用户标识。服务器 120 从音频数据中提取出音频特征和文字信息，从视频数据中提取出用户人脸图像对应的面部特征；将提取出的音频特征、文字信息和面部特征分别输入至对应的表情分类器中，获取各个表情分类器输出的表情识别结果以及与表情识别结果对应的权重；从输出的表情识别结果中，筛选出权重最大的表情识别结果，作为与用户标识对应的用户的表情类别。终端 110 可以但不限于各种个人计算机、笔记本电脑、智能手机和平板电脑，服务器 120 可以用独立的服务器或者是多个服务器组成的服务器集群来实现。

在其中一个实施例中，如图 2 所示，提供了一种表情识别方法，以该方法应用于图 1 中的服务器为例进行说明，包括以下步骤：

步骤 S201，获取待处理的视频数据和音频数据，视频数据和音频数据均携带同一用户标识，音频数据包括音频特征和文字信息，视频数据包括用户人脸图像对应的面部特征。

视频数据和音频数据是指在用户面审、面试等环节中，通过终端中的视频录制系统对用户所录制的数据。用户标识用于标识视频数据和音频数据，便于服务器进行区分；用户标识可以是用户姓名，用户身份证号码等。音频特征是指用于分析用户的表情类别的特征信息，比如音量、频响等；文字信息是指通过语音识别技术，由音频特征中的音频信息转化而来的信息；面部特征是指从视频数据中的人脸图像中提取出来的，用于分析用户的表情类别的特征信息，比如眼部特征，嘴部特征等。用户的表情类别可以分为很多种，比如悲伤、高兴、厌恶等。

终端（比如智能手机）安装了视频录制系统，视频录制系统能够将在用户面审、面试等环节中录制的视频数据和音频数据，作为待处理的视频数据和音频数据，上传至服务器。服务器从待处理的音频数据中提取音频特征和文字信息；从待处理的视频数据中提取用户人脸图像，并从提取出的用户人脸图像提取面部特征，方便后续将提取出的音频特征、文字信息和人脸图像输入对应的表情分类器，以从多个角度对用户的表情类别进行综合判断，进一步提高了表情识别的准确率。

步骤 S202，将音频特征、文字信息和面部特征分别输入至对应的表情分类器中，获取各个表情分类器输出的表情识别结果以及与表情识别结果对应的权重。

表情分类器能够根据输入的信息，输出对应的表情识别结果，比如，基于语音识别的

表情分类器能够根据输入的音频特征，输出用户的表情类别；基于文字识别的表情分类器能够根据输入的文字信息，输出用户的表情类别；基于人脸识别的表情分类器能够根据输入的面部特征，输出用户的表情类别。

权重用于衡量表情识别结果的重要程度，不同的表情识别结果，对应的权重不一样。需要说明的是，同一表情分类器，每次输出的表情识别结果所占有的权重是不同的，具体跟输入的信息有关。

服务器将音频特征、文字信息和面部特征分别输入至对应的表情分类器中，获取各个表情分类器输出的表情识别结果以及与表情识别结果对应的权重，方便从语音、文字和视觉这三个角度对用户的表情类别进行综合判断，能够提供更为多样的表情识别结果，方便从多样的表情识别结果对用户的表情类别进行综合分析，进一步提高了表情识别的准确率，避免了传统方法中仅从视觉角度对用户的表情类别进行判断，导致表情识别的准确率低的缺陷。

步骤 S203，从输出的表情识别结果中，筛选出权重最大的表情识别结果，作为与用户标识对应的用户的表情类别。

服务器从各个表情分类器输出的表情识别结果中，获取权重最大的表情识别结果，作为与用户标识对应的用户的表情类别，实现了从多个角度对用户的表情类别进行综合判断的目的，避免漏判，进一步提高了表情识别的准确率和稳定性。需要说明的是，若各个表情识别结果对应的权重一样，则以基于人脸识别的表情分类器输出的表情识别结果为主。

上述表情识别方法中，服务器获取待处理的视频数据和音频数据，视频数据和音频数据均携带同一用户标识，音频数据包括音频特征和文字信息，视频数据包括用户人脸图像对应的面部特征；将音频特征、文字信息和面部特征分别输入至对应的表情分类器中，获取各个表情分类器输出的表情识别结果以及与表情识别结果对应的权重；从输出的表情识别结果中，筛选出权重最大的表情识别结果，作为与用户标识对应的用户的表情类别；实现了从语音、文字和视觉这三个角度对用户的表情类别进行综合判断的目的，能够提供更为多样的表情识别结果，方便从多样的表情识别结果中对用户的表情类别进行综合分析，避免漏判，进一步提高了表情识别的准确率和稳定性，克服了传统方法中仅从视觉角度对用户的表情类别进行判断，导致表情识别的准确率低的缺陷。

考虑到不同的表情分类器，对应的输入信息不同，为了区分不同的表情分类器，服务器可以将表情分类器分为第一表情分类器、第二表情分类器和第三表情分类器。

在其中一个实施例中，如图 3 所示，将音频特征、文字信息和面部特征分别输入至对应的表情分类器中，获取各个表情分类器输出的表情识别结果以及与表情识别结果对应的权重的步骤具体包括：

步骤 S301，将音频特征输入至第一表情分类器中，获取第一表情分类器的表情识别结果以及与表情识别结果对应的第一权重。

步骤 S302，将文字信息输入至第二表情分类器中，获取第二表情分类器的表情识别

结果以及与表情识别结果对应的第二权重。

步骤 S303，将面部特征输入至第三表情分类器中，获取第三表情分类器的表情识别结果以及与表情识别结果对应的第三权重。

第一表情分类器是基于语音识别的表情分类器，能够根据输入的音频特征，输出与音频特征对应的表情类别，作为用户的表情类别。第二表情分类器是基于文字识别的表情分类器，能够根据输入的文字信息，输出与文字信息对应的表情类型，作为用户的表情类别。第三表情分类器是基于人脸识别的表情分类器，能够根据输入的面部特征，输出与面部特征对应的表情类别，作为用户的表情类别。

服务器将音频特征、文字信息和面部特征分别输入至对应的表情分类器中，以从语音、文字和视觉这三个角度对用户的表情类别进行综合判断，避免漏判，进一步提高了表情识别的准确率；同时，从多个角度对用户的表情类别进行判断，可以提高表情识别的稳定性。

在其中一个实施例中，上述步骤 S301，将音频特征输入至第一表情分类器中，获取第一表情分类器的表情识别结果以及与表情识别结果对应的第一权重，包括：将音频特征输入至第一表情分类器中，第一表情分类器用于从音频特征中提取出目标特征，根据目标特征查询第一数据库，获取与目标特征对应的表情类别，作为表情识别结果，并确定与表情识别结果对应的第一权重；获取第一表情分类器的表情识别结果以及与表情识别结果对应的第一权重。目标特征是指输入的音频特征中，与设定音频特征（比如音调）匹配的特征。第一数据库存储有多个与音频特征对应的表情类别。

服务器基于大数据，预先收集了多个不同的音频特征以及与音频特征对应的表情类别，从音频特征中提取出目标特征；将与音频特征对应的表情类别，作为与从该音频特征中提取的目标特征对应的表情类别，以得到多个与目标特征对应的表情类别；将多个与目标特征对应的表情类别存储至第一数据库中，方便后续通过第一数据库获取与目标特征对应的表情类别。比如，恐惧的情绪中时常会带有尖叫，因此平均音调较高；故通过第一表情分类器，从音频特征中提取出目标特征音调；当识别到音调较高时，可以判断出用户的表情类别为恐惧。通过第一表情分类器，可以获取与输入的音频特征对应的表情类别，结合语音识别技术，通过用户语音中的音频特征来分析用户当前的表情类别，进一步提高了表情识别的准确率，避免漏判而造成表情识别的准确率低缺陷。

为了进一步提高第一表情分类器的表情识别准确率，可以对第一表情分类器进行多次训练。在其中一个实施例中，第一表情分类器通过下述方法得到：获取多个样本音频特征及对应的表情类别；通过待训练的第一表情分类器对样本音频特征进行识别，得到第一表情分类器的表情识别结果；将表情识别结果与对应的实际表情类别进行比较，得到识别误差；当识别误差大于或等于预设第一阈值时，根据识别误差对第一表情分类器进行训练，直到根据训练后的第一表情分类器得到的识别误差小于预设第一阈值，结束训练。

比如，当识别误差大于或等于预设第一阈值时，服务器根据识别误差调整第一表情分

类器的参数；根据调整后的第一表情分类器对样本音频特征进行再次识别，获取根据第一表情分类器得到的表情识别结果与对应的实际表情类别之间的识别误差，根据识别误差对第一表情分类器的参数进行再次调整，以对第一表情分类器进行再次训练，直到根据训练后的第一表情分类器得到的识别误差小于预设第一阈值，结束训练。服务器根据识别误差，对第一表情分类器进行多次训练，方便通过第一表情分类器输出更准确的表情识别结果，避免漏判，从而进一步提高了第一表情分类器的表情识别准确率。

在其中一个实施例中，上述步骤 S302，将文字信息输入至第二表情分类器中，获取第二表情分类器的表情识别结果以及与表情识别结果对应的第二权重，包括：将文字信息输入至第二表情分类器中，第二表情分类器用于从文字信息中提取出目标信息，根据目标信息查询第二数据库，获取与目标信息对应的表情类别，作为表情识别结果，并确定与表情识别结果对应的第二权重；获取第二表情分类器的表情识别结果以及与表情识别结果对应的第二权重。目标信息是指从输入的文字信息中，提取出的蕴含情绪的信息，比如开心、愤怒等。第二数据库存储有多个与文字信息对应的表情类别。

服务器基于大数据，预先收集了多个不同用户的语音信息，将语音信息转化成文字信息，从文字信息中提取出目标信息，并确定目标信息对应的表情类别，将多个与目标信息对应的表情类别存储至第二数据库中，方便后续通过第二数据库获取与目标信息对应的表情类别。比如，文字信息“开心”通常代表着快乐，故通过第二表情分类器，从文字信息中提取出目标信息；当识别到目标信息为“开心”时，可以判断出用户的表情类别为快乐。通过第二表情分类器，可以获取与输入的文字信息对应的表情类别，以确定用户当前的表情类别，从而进一步提高了表情识别的准确率，避免漏判而造成表情识别的准确率低的缺陷。

进一步，第二表情分类器还可以从文字信息中提取出目标信息，并从文字信息中确定与目标信息关联的上下文信息；根据目标信息以及与目标信息关联的上下文信息，确定目标信息的实际含义；根据目标信息的实际含义查询第二数据库，获取与目标信息的实际含义对应的表情类别，作为表情识别结果，并确定与表情识别结果对应的第二权重。比如，从文字信息“你说发生这样的事情我能开心吗”中提取目标信息“开心”，并结合“开心”的上下文信息，确定“开心”的实际含义为消极情绪“不开心”。基于此方法，可以获取多个与目标信息对应的表情类别，并将其存储至第二数据库中。结合提取的目标信息的上下文信息，能够进一步减少第二表情分类器的识别误差，从而提高了表情识别的准确率。

在其中一个实施例中，上述步骤 S303，将面部特征输入至第三表情分类器中，获取第三表情分类器的表情识别结果以及与表情识别结果对应的第三权重，包括：将面部特征输入至第三表情分类器中，第三表情分类器用于根据面部特征查询第三数据库，获取与面部特征对应的表情类别，作为表情识别结果，并确定与表情识别结果对应的第三权重；获取第三表情分类器的表情识别结果以及与表情识别结果对应的第三权重。第三数据库存储有多个与面部特征对应的表情类别。

服务器基于大数据，预先收集了多个不同的面部特征以及与面部特征对应的表情类别；将多个与面部特征对应的表情类别存储至第三数据库中，方便后续通过第三数据库获取与面部特征对应的表情类别。比如，人们高兴时的面部特征为嘴角翘起，面颊上抬起皱，眼睑收缩，眼睛尾部会形成“鱼尾纹”；伤心时的面部特征为眯眼，眉毛收紧，嘴角下拉，下巴抬起或收紧。故通过第三表情分类器，当识别到面部特征为嘴角翘起，面颊上抬起皱，眼睑收缩，眼睛尾部会形成“鱼尾纹”，可以判断出用户的表情类别为高兴。通过第三表情分类器，可以获取与输入的面部特征对应的表情类别，结合人脸识别技术，通过用户的面部特征来分析用户当前的表情类别，进一步提高了表情识别的准确率，避免漏判而造成表情识别的准确率低的缺陷。

为了提高表情识别的准确率，服务器还可以从待处理的视频数据中提取多张人脸图像，分别从各张人脸图像中提取面部特征，将面部特征输入第三表情分类器，获取多个表情识别结果及对应的多个第三权重；从多个表情识别结果中筛选出第三权重最大的表情识别结果，作为第三表情分类器最终输出的表情识别结果。从待处理的视频数据中提取多张人脸图像，并通过第三表情分类器进行分析，避免仅仅根据单张人脸图像中的面部特征对用户的表情类别进行判断，导致表情识别的准确率低的缺陷，从而提高了表情识别的准确率。

此外，为了进一步提高表情识别的准确率，服务器还可以对第三表情分类器进行多次训练。在其中一个实施例中，第三表情分类器通过下述方法得到：获取多个样本面部特征及对应的表情类别；通过待训练的第三表情分类器对样本面部特征进行识别，得到第三表情分类器的表情识别结果；获取表情识别结果与对应的实际表情类别之间的相似度；当相似度小于预设第二阈值时，根据相似度对第三表情分类器进行训练，直到根据训练后的第三表情分类器得到的表情识别结果与对应的实际表情类别之间的相似度大于或等于预设第二阈值，结束训练。

比如，当相似度小于预设第二阈值时，服务器根据相似度调整第三表情分类器的参数；根据调整后的第三表情分类器对样本面部特征进行再次识别，获取根据第三表情分类器得到的表情识别结果与对应的实际表情类别之间的相似度，根据相似度对第三表情分类器的参数进行再次调整，以对第三表情分类器进行再次训练，直到根据训练后的第三表情分类器得到的表情识别结果与对应的实际表情类别之间的相似度大于或等于预设第二阈值，结束训练。服务器根据相似度，对第三表情分类器进行多次训练，方便通过第三表情分类器输出更准确的表情识别结果，避免漏判，从而进一步提高了第三表情分类器的表情识别准确率。

在其中一个实施例中，如图4所示，提供了另一种表情识别方法，包括以下步骤：

步骤 S401，获取待处理的视频数据和音频数据，视频数据和音频数据均携带同一用户标识，音频数据包括音频特征和文字信息，视频数据包括用户人脸图像对应的面部特征。

步骤 S402，将音频特征输入至第一表情分类器中，第一表情分类器用于从音频特征

中提取出目标特征，根据目标特征查询第一数据库，获取与目标特征对应的表情类别，作为表情识别结果，并确定与表情识别结果对应的第一权重；获取第一表情分类器的表情识别结果以及与表情识别结果对应的第一权重。

步骤 S403，将文字信息输入至第二表情分类器中，第二表情分类器用于从文字信息中提取出目标信息，根据目标信息查询第二数据库，获取与目标信息对应的表情类别，作为表情识别结果，并确定与表情识别结果对应的第二权重；获取第二表情分类器的表情识别结果以及与表情识别结果对应的第二权重。

步骤 S404，将面部特征输入至第三表情分类器中，第三表情分类器用于根据面部特征查询第三数据库，获取与面部特征对应的表情类别，作为表情识别结果，并确定与表情识别结果对应的第三权重；获取第三表情分类器的表情识别结果以及与表情识别结果对应的第三权重。

步骤 S405，从输出的表情识别结果中，筛选出权重最大的表情识别结果，作为与用户标识对应的用户的表情类别。

上述表情识别方法中，实现了从语音、文字和视觉这三个角度对用户的表情类别进行综合判断的目的，能够提供更为多样的表情识别结果，方便从多样的表情识别结果中对用户的表情类别进行综合分析，避免漏判，进一步提高了表情识别的准确率和稳定性，克服了传统方法中仅从视觉角度对用户的表情类别进行判断，导致表情识别的准确率低的缺陷。

应该理解的是，虽然图 2-4 的流程图中的各个步骤按照箭头的指示依次显示，但是这些步骤并不是必然按照箭头指示的顺序依次执行。除非本文中有明确的说明，这些步骤的执行并没有严格的顺序限制，这些步骤可以以其它的顺序执行。而且，图 2-4 中的至少一部分步骤可以包括多个子步骤或者多个阶段，这些子步骤或者阶段并不必然是在同一时刻执行完成，而是可以在不同的时刻执行，这些子步骤或者阶段的执行顺序也不必然是依次进行，而是可以与其它步骤或者其它步骤的子步骤或者阶段的至少一部分轮流或者交替地执行。

在其中一个实施例中，如图 5 所示，提供了一种表情识别装置，包括：数据获取模块 510、表情获取模块 520 和表情筛选模块 530，其中：

数据获取模块 510，用于获取待处理的视频数据和音频数据，视频数据和音频数据均携带同一用户标识，音频数据包括音频特征和文字信息，视频数据包括用户人脸图像对应的面部特征。

表情获取模块 520，用于将音频特征、文字信息和面部特征分别输入至对应的表情分类器中，获取各个表情分类器输出的表情识别结果以及与表情识别结果对应的权重。

表情筛选模块 530，用于从输出的表情识别结果中，筛选出权重最大的表情识别结果，作为与用户标识对应的用户的表情类别。

在其中一个实施例中，表情获取模块还用于将音频特征输入至第一表情分类器中，获取第一表情分类器的表情识别结果以及与表情识别结果对应的第一权重；将文字信息输入

至第二表情分类器中，获取第二表情分类器的表情识别结果以及与表情识别结果对应的第二权重；将面部特征输入至第三表情分类器中，获取第三表情分类器的表情识别结果以及与表情识别结果对应的第三权重。

在其中一个实施例中，表情获取模块还用于将音频特征输入至第一表情分类器中，第一表情分类器用于从音频特征中提取出目标特征，根据目标特征查询第一数据库，获取与目标特征对应的表情类别，作为表情识别结果，并确定与表情识别结果对应的第一权重；获取第一表情分类器的表情识别结果以及与表情识别结果对应的第一权重。

在其中一个实施例中，表情获取模块还用于将文字信息输入至第二表情分类器中，第二表情分类器用于从文字信息中提取出目标信息，根据目标信息查询第二数据库，获取与目标信息对应的表情类别，作为表情识别结果，并确定与表情识别结果对应的第二权重；获取第二表情分类器的表情识别结果以及与表情识别结果对应的第二权重。

在其中一个实施例中，表情获取模块还用于将面部特征输入至第三表情分类器中，第三表情分类器用于根据面部特征查询第三数据库，获取与面部特征对应的表情类别，作为表情识别结果，并确定与表情识别结果对应的第三权重；获取第三表情分类器的表情识别结果以及与表情识别结果对应的第三权重。

在其中一个实施例中，表情识别装置还包括第一训练模块，用于获取多个样本音频特征及对应的表情类别；通过待训练的第一表情分类器对样本音频特征进行识别，得到第一表情分类器的表情识别结果；将表情识别结果与对应的实际表情类别进行比较，得到识别误差；当识别误差大于或等于预设第一阈值时，根据识别误差对第一表情分类器进行训练，直到根据训练后的第一表情分类器得到的识别误差小于预设第一阈值，结束训练。

在其中一个实施例中，表情识别装置还包括第二训练模块，用于获取多个样本面部特征及对应的表情类别；通过待训练的第三表情分类器对样本面部特征进行识别，得到第三表情分类器的表情识别结果；获取表情识别结果与对应的实际表情类别之间的相似度；当相似度小于预设第二阈值时，根据相似度对第三表情分类器进行训练，直到根据训练后的第三表情分类器得到的表情识别结果与对应的实际表情类别之间的相似度大于或等于预设第二阈值，结束训练。

上述各个实施例，表情识别装置实现了从语音、文字和视觉这三个角度对用户的表情类别进行综合判断的目的，能够提供更为多样的表情识别结果，方便从多样的表情识别结果中对用户的表情类别进行综合分析，避免漏判，进一步提高了表情识别的准确率和稳定性，克服了传统方法中仅从视觉角度对用户表情类别进行判断，导致表情识别的准确率低

的缺陷。

关于表情识别装置的具体限定可以参见上文中对于表情识别方法的限定，在此不再赘述。上述表情识别装置中的各个模块可全部或部分通过软件、硬件及其组合来实现。上述各模块可以硬件形式内嵌于或独立于计算机设备中的处理器中，也可以以软件形式存储于计算机设备中的存储器中，以便于处理器调用执行以上各个模块对应的操作。

在其中一个实施例中，提供了一种计算机设备，该计算机设备可以是服务器，其内部结构图可以如图 6 所示。该计算机设备包括通过系统总线连接的处理器、存储器、网络接口和数据库。该计算机设备的处理器用于提供计算和控制能力。该计算机设备的存储器包括非易失性存储介质、内存储器。该非易失性存储介质存储有操作系统、计算机可读指令和数据库。该内存储器为非易失性存储介质中的操作系统和计算机可读指令的运行提供环境。该计算机设备的数据库用于存储表情类别。该计算机设备的网络接口用于与外部的终端通过网络连接通信。该计算机可读指令被处理器执行时以实现一种表情识别方法。

本领域技术人员可以理解，图 6 中示出的结构，仅仅是与本申请方案相关的部分结构的框图，并不构成对本申请方案所应用于其上的计算机设备的限定，具体的计算机设备可以包括比图中所示更多或更少的部件，或者组合某些部件，或者具有不同的部件布置。

一种计算机设备，包括存储器和一个或多个处理器，存储器中储存有计算机可读指令，计算机可读指令被处理器执行时，使得一个或多个处理器执行以下步骤：

获取待处理的视频数据和音频数据，视频数据和音频数据均携带同一用户标识，音频数据包括音频特征和文字信息，视频数据包括用户人脸图像对应的面部特征；

将音频特征、文字信息和面部特征分别输入至对应的表情分类器中，获取各个表情分类器输出的表情识别结果以及与表情识别结果对应的权重；及

从输出的表情识别结果中，筛选出权重最大的表情识别结果，作为与用户标识对应的用户的表情类别。

在其中一个实施例中，处理器执行计算机可读指令时还实现以下步骤：将音频特征输入至第一表情分类器中，获取第一表情分类器的表情识别结果以及与表情识别结果对应的第一权重；将文字信息输入至第二表情分类器中，获取第二表情分类器的表情识别结果以及与表情识别结果对应的第二权重；及将面部特征输入至第三表情分类器中，获取第三表情分类器的表情识别结果以及与表情识别结果对应的第三权重。

在其中一个实施例中，处理器执行计算机可读指令时还实现以下步骤：将音频特征输入至第一表情分类器中，第一表情分类器用于从音频特征中提取出目标特征，根据目标特征查询第一数据库，获取与目标特征对应的表情类别，作为表情识别结果，并确定与表情识别结果对应的第一权重；及获取第一表情分类器的表情识别结果以及与表情识别结果对应的第一权重。

在其中一个实施例中，处理器执行计算机可读指令时还实现以下步骤：将文字信息输入至第二表情分类器中，第二表情分类器用于从文字信息中提取出目标信息，根据目标信息查询第二数据库，获取与目标信息对应的表情类别，作为表情识别结果，并确定与表情识别结果对应的第二权重；及获取第二表情分类器的表情识别结果以及与表情识别结果对应的第二权重。

在其中一个实施例中，处理器执行计算机可读指令时还实现以下步骤：将面部特征输入至第三表情分类器中，第三表情分类器用于根据面部特征查询第三数据库，获取与面部

特征对应的表情类别，作为表情识别结果，并确定与表情识别结果对应的第三权重；及获取第三表情分类器的表情识别结果以及与表情识别结果对应的第三权重。

在其中一个实施例中，处理器执行计算机可读指令时还实现以下步骤：获取多个样本音频特征及对应的表情类别；通过待训练的第一表情分类器对样本音频特征进行识别，得到第一表情分类器的表情识别结果；将表情识别结果与对应的实际表情类别进行比较，得到识别误差；及当识别误差大于或等于预设第一阈值时，根据识别误差对第一表情分类器进行训练，直到根据训练后的第一表情分类器得到的识别误差小于预设第一阈值，结束训练。

在其中一个实施例中，处理器执行计算机可读指令时还实现以下步骤：获取多个样本面部特征及对应的表情类别；通过待训练的第三表情分类器对样本面部特征进行识别，得到第三表情分类器的表情识别结果；获取表情识别结果与对应的实际表情类别之间的相似度；及当相似度小于预设第二阈值时，根据相似度对第三表情分类器进行训练，直到根据训练后的第三表情分类器得到的表情识别结果与对应的实际表情类别之间的相似度大于或等于预设第二阈值，结束训练。

上述各个实施例，计算机设备通过处理器上运行的计算机可读指令，实现了从语音、文字和视觉这三个角度对用户的表情类别进行综合判断的目的，能够提供更为多样的表情识别结果，方便从多样的表情识别结果中对用户的表情类别进行综合分析，避免漏判，进一步提高了表情识别的准确率和稳定性，克服了传统方法中仅从视觉角度对用户表情类别进行判断，导致表情识别的准确率低缺陷。

一个或多个存储有计算机可读指令的非易失性存储介质，计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器执行以下步骤：

获取待处理的视频数据和音频数据，视频数据和音频数据均携带同一用户标识，音频数据包括音频特征和文字信息，视频数据包括用户人脸图像对应的面部特征；

将音频特征、文字信息和面部特征分别输入至对应的表情分类器中，获取各个表情分类器输出的表情识别结果以及与表情识别结果对应的权重；及

从输出的表情识别结果中，筛选出权重最大的表情识别结果，作为与用户标识对应的用户的表情类别。

在其中一个实施例中，计算机可读指令被处理器执行时还实现以下步骤：将音频特征输入至第一表情分类器中，获取第一表情分类器的表情识别结果以及与表情识别结果对应的第一权重；将文字信息输入至第二表情分类器中，获取第二表情分类器的表情识别结果以及与表情识别结果对应的第二权重；及将面部特征输入至第三表情分类器中，获取第三表情分类器的表情识别结果以及与表情识别结果对应的第三权重。

在其中一个实施例中，计算机可读指令被处理器执行时还实现以下步骤：将音频特征输入至第一表情分类器中，第一表情分类器用于从音频特征中提取出目标特征，根据目标特征查询第一数据库，获取与目标特征对应的表情类别，作为表情识别结果，并确定与表

情识别结果对应的第一权重；及获取第一表情分类器的表情识别结果以及与表情识别结果对应的第一权重。

在其中一个实施例中，计算机可读指令被处理器执行时还实现以下步骤：将文字信息输入至第二表情分类器中，第二表情分类器用于从文字信息中提取出目标信息，根据目标信息查询第二数据库，获取与目标信息对应的表情类别，作为表情识别结果，并确定与表情识别结果对应的第二权重；及获取第二表情分类器的表情识别结果以及与表情识别结果对应的第二权重。

在其中一个实施例中，计算机可读指令被处理器执行时还实现以下步骤：将面部特征输入至第三表情分类器中，第三表情分类器用于根据面部特征查询第三数据库，获取与面部特征对应的表情类别，作为表情识别结果，并确定与表情识别结果对应的第三权重；及获取第三表情分类器的表情识别结果以及与表情识别结果对应的第三权重。

在其中一个实施例中，计算机可读指令被处理器执行时还实现以下步骤：获取多个样本音频特征及对应的表情类别；通过待训练的第一表情分类器对样本音频特征进行识别，得到第一表情分类器的表情识别结果；将表情识别结果与对应的实际表情类别进行比较，得到识别误差；及当识别误差大于或等于预设第一阈值时，根据识别误差对第一表情分类器进行训练，直到根据训练后的第一表情分类器得到的识别误差小于预设第一阈值，结束训练。

在其中一个实施例中，计算机可读指令被处理器执行时还实现以下步骤：获取多个样本面部特征及对应的表情类别；通过待训练的第三表情分类器对样本面部特征进行识别，得到第三表情分类器的表情识别结果；获取表情识别结果与对应的实际表情类别之间的相似度；及当相似度小于预设第二阈值时，根据相似度对第三表情分类器进行训练，直到根据训练后的第三表情分类器得到的表情识别结果与对应的实际表情类别之间的相似度大于或等于预设第二阈值，结束训练。

上述各个实施例，计算机可读存储介质通过其存储的计算机可读指令，实现了从语音、文字和视觉这三个角度对用户的表情类别进行综合判断的目的，能够提供更为多样的表情识别结果，方便从多样的表情识别结果中对用户的表情类别进行综合分析，避免漏判，进一步提高了表情识别的准确率和稳定性，克服了传统方法表情识别的准确率低缺陷。

本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，是可以通过计算机可读指令来指令相关的硬件来完成，该计算机可读指令可存储于一非易失性计算机可读存储介质中，该计算机可读指令在执行时，可包括如上述各方法的实施例的流程。其中，本申请所提供的各实施例中所使用的对存储器、存储、数据库或其它介质的任何引用，均可包括非易失性和/或易失性存储器。非易失性存储器可包括只读存储器（ROM）、可编程 ROM（PROM）、电可编程 ROM（EPROM）、电可擦除可编程 ROM（EEPROM）或闪存。易失性存储器可包括随机存取存储器（RAM）或者外部高速缓冲存储器。作为说明而非局限，RAM 以多种形式可得，诸如静态 RAM（SRAM）、动态 RAM（DRAM）、

同步 DRAM(SDRAM)、双数据率 SDRAM(DDRSDRAM)、增强型 SDRAM(ESDRAM)、同步链路(Synchlink) DRAM(SLDRAM)、存储器总线(Rambus)直接 RAM(RDRAM)、直接存储器总线动态 RAM(DRDRAM)、以及存储器总线动态 RAM(RDRAM)等。

5 以上实施例的各技术特征可以进行任意的组合,为使描述简洁,未对上述实施例中的各个技术特征所有可能的组合都进行描述,然而,只要这些技术特征的组合不存在矛盾,都应当认为是本说明书记载的范围。

以上所述实施例仅表达了本申请的几种实施方式,其描述较为具体和详细,但并不能因此而理解为对发明专利范围的限制。应当指出的是,对于本领域的普通技术人员来说,在不脱离本申请构思的前提下,还可以做出若干变形和改进,这些都属于本申请的保护范围。因此,本申请专利的保护范围应以所附权利要求为准。

权利要求书

1、一种表情识别方法，包括：

获取待处理的视频数据和音频数据，所述视频数据和所述音频数据均携带同一用户标识，所述音频数据包括音频特征和文字信息，所述视频数据包括用户人脸图像对应的面部特征；

5 将所述音频特征、文字信息和面部特征分别输入至对应的表情分类器中，获取各个所述表情分类器输出的表情识别结果以及与所述表情识别结果对应的权重；及

从所述输出的表情识别结果中，筛选出权重最大的表情识别结果，作为与所述用户标识对应的用户的表情类别。

2、根据权利要求1所述的方法，其特征在于，所述表情分类器包括第一表情分类器、
10 第二表情分类器和第三表情分类器；

所述将所述音频特征、文字信息和面部特征分别输入至对应的表情分类器中，获取各个所述表情分类器输出的表情识别结果以及与所述表情识别结果对应的权重，包括：

将所述音频特征输入至第一表情分类器中，获取第一表情分类器的表情识别结果以及
与所述表情识别结果对应的第一权重；

15 将所述文字信息输入至第二表情分类器中，获取第二表情分类器的表情识别结果以及
与所述表情识别结果对应的第二权重；及

将所述面部特征输入至第三表情分类器中，获取第三表情分类器的表情识别结果以及
与所述表情识别结果对应的第三权重。

3、根据权利要求2所述的方法，其特征在于，所述将所述音频特征输入至第一表情
20 分类器中，获取第一表情分类器的表情识别结果以及与所述表情识别结果对应的第一权重，
包括：

将所述音频特征输入至第一表情分类器中，所述第一表情分类器用于从所述音频特征
中提取出目标特征，根据所述目标特征查询第一数据库，获取与所述目标特征对应的表情
类别，作为表情识别结果，并确定与所述表情识别结果对应的第一权重；及

25 获取所述第一表情分类器的表情识别结果以及与所述表情识别结果对应的第一权重。

4、根据权利要求2所述的方法，其特征在于，所述将所述文字信息输入至第二表情
分类器中，获取第二表情分类器的表情识别结果以及与所述表情识别结果对应的第二权重，
包括：

30 将所述文字信息输入至第二表情分类器中，所述第二表情分类器用于从所述文字信息
中提取出目标信息，根据所述目标信息查询第二数据库，获取与所述目标信息对应的表情
类别，作为表情识别结果，并确定与所述表情识别结果对应的第二权重；及

获取所述第二表情分类器的表情识别结果以及与所述表情识别结果对应的第二权重。

5、根据权利要求2所述的方法，其特征在于，所述将所述面部特征输入至第三表情
分类器中，获取第三表情分类器的表情识别结果以及与所述表情识别结果对应的第三权重，

包括：

将所述面部特征输入至第三表情分类器中，所述第三表情分类器用于根据所述面部特征查询第三数据库，获取与所述面部特征对应的表情类别，作为表情识别结果，并确定与所述表情识别结果对应的第三权重；及

5 获取所述第三表情分类器的表情识别结果以及与所述表情识别结果对应的第三权重。

6、根据权利要求 1 至 5 任意一项所述的方法，其特征在于，所述第一表情分类器通过下述方法得到：

获取多个样本音频特征及对应的表情类别；

10 通过待训练的第一表情分类器对所述样本音频特征进行识别，得到第一表情分类器的表情识别结果；

将所述表情识别结果与对应的实际表情类别进行比较，得到识别误差；及

当所述识别误差大于或等于预设第一阈值时，根据所述识别误差对所述第一表情分类器进行训练，直到根据训练后的第一表情分类器得到的识别误差小于所述预设第一阈值，结束训练。

15 7、根据权利要求 6 所述的方法，其特征在于，所述第三表情分类器通过下述方法得到：

获取多个样本面部特征及对应的表情类别；

通过待训练的第三表情分类器对所述样本面部特征进行识别，得到第三表情分类器的表情识别结果；

20 获取所述表情识别结果与对应的实际表情类别之间的相似度；及

当所述相似度小于预设第二阈值时，根据所述相似度对所述第三表情分类器进行训练，直到根据训练后的第三表情分类器得到的表情识别结果与对应的实际表情类别之间的相似度大于或等于所述预设第二阈值，结束训练。

8、一种表情识别装置，包括：

25 数据获取模块，用于获取待处理的视频数据和音频数据，所述视频数据和所述音频数据均携带同一用户标识，所述音频数据包括音频特征和文字信息，所述视频数据包括用户人脸图像对应的面部特征；

表情获取模块，用于将所述音频特征、文字信息和面部特征分别输入至对应的表情分类器中，获取各个所述表情分类器输出的表情识别结果以及与所述表情识别结果对应的权重；及

30 表情筛选模块，用于从所述输出的表情识别结果中，筛选出权重最大的表情识别结果，作为与所述用户标识对应的用户的表情类别。

9、根据权利要求 8 所述的装置，其特征在于，所述表情分类器包括第一表情分类器、第二表情分类器和第三表情分类器；

35 所述表情获取模块还用于将所述音频特征输入至第一表情分类器中，获取第一表情分

类器的表情识别结果以及与所述表情识别结果对应的第一权重；将所述文字信息输入至第二表情分类器中，获取第二表情分类器的表情识别结果以及与所述表情识别结果对应的第二权重；将所述面部特征输入至第三表情分类器中，获取第三表情分类器的表情识别结果以及与所述表情识别结果对应的第三权重。

5 10、根据权利要求9所述的装置，其特征在于，所述表情获取模块还用于将所述音频特征输入至第一表情分类器中，所述第一表情分类器用于从所述音频特征中提取出目标特征，根据所述目标特征查询第一数据库，获取与所述目标特征对应的表情类别，作为表情识别结果，并确定与所述表情识别结果对应的第一权重；获取所述第一表情分类器的表情识别结果以及与所述表情识别结果对应的第一权重。

10 11、一种计算机设备，包括存储器及一个或多个处理器，所述存储器中储存有计算机可读指令，所述计算机可读指令被所述一个或多个处理器执行时，使得所述一个或多个处理器执行以下步骤：

获取待处理的视频数据和音频数据，所述视频数据和所述音频数据均携带同一用户标识，所述音频数据包括音频特征和文字信息，所述视频数据包括用户人脸图像对应的面部特征；

15 将所述音频特征、文字信息和面部特征分别输入至对应的表情分类器中，获取各个所述表情分类器输出的表情识别结果以及与所述表情识别结果对应的权重；及

从所述输出的表情识别结果中，筛选出权重最大的表情识别结果，作为与所述用户标识对应的用户的表情类别。

20 12、根据权利要求11所述的计算机设备，其特征在于，所述表情分类器包括第一表情分类器、第二表情分类器和第三表情分类器；

所述处理器执行所述计算机可读指令时还执行以下步骤：

将所述音频特征输入至第一表情分类器中，获取第一表情分类器的表情识别结果以及与所述表情识别结果对应的第一权重；

25 将所述文字信息输入至第二表情分类器中，获取第二表情分类器的表情识别结果以及与所述表情识别结果对应的第二权重；及

将所述面部特征输入至第三表情分类器中，获取第三表情分类器的表情识别结果以及与所述表情识别结果对应的第三权重。

30 13、根据权利要求12所述的计算机设备，其特征在于，所述处理器执行所述计算机可读指令时还执行以下步骤：

将所述音频特征输入至第一表情分类器中，所述第一表情分类器用于从所述音频特征中提取出目标特征，根据所述目标特征查询第一数据库，获取与所述目标特征对应的表情类别，作为表情识别结果，并确定与所述表情识别结果对应的第一权重；及

获取所述第一表情分类器的表情识别结果以及与所述表情识别结果对应的第一权重。

35 14、根据权利要求12所述的计算机设备，其特征在于，所述处理器执行所述计算机

可读指令时还执行以下步骤:

将所述文字信息输入至第二表情分类器中,所述第二表情分类器用于从所述文字信息中提取出目标信息,根据所述目标信息查询第二数据库,获取与所述目标信息对应的表情类别,作为表情识别结果,并确定与所述表情识别结果对应的第二权重;及

5 获取所述第二表情分类器的表情识别结果以及与所述表情识别结果对应的第二权重。

15、根据权利要求 12 所述的计算机设备,其特征在于,所述处理器执行所述计算机可读指令时还执行以下步骤:

10 将所述面部特征输入至第三表情分类器中,所述第三表情分类器用于根据所述面部特征查询第三数据库,获取与所述面部特征对应的表情类别,作为表情识别结果,并确定与所述表情识别结果对应的第三权重;及

获取所述第三表情分类器的表情识别结果以及与所述表情识别结果对应的第三权重。

16、一个或多个存储有计算机可读指令的非易失性计算机可读存储介质,所述计算机可读指令被一个或多个处理器执行时,使得所述一个或多个处理器执行以下步骤:

15 获取待处理的视频数据和音频数据,所述视频数据和所述音频数据均携带同一用户标识,所述音频数据包括音频特征和文字信息,所述视频数据包括用户人脸图像对应的面部特征;

将所述音频特征、文字信息和面部特征分别输入至对应的表情分类器中,获取各个所述表情分类器输出的表情识别结果以及与所述表情识别结果对应的权重;及

20 从所述输出的表情识别结果中,筛选出权重最大的表情识别结果,作为与所述用户标识对应的用户的表情类别。

17、根据权利要求 16 所述的存储介质,其特征在于,所述表情分类器包括第一表情分类器、第二表情分类器和第三表情分类器;

所述计算机可读指令被所述处理器执行时还执行以下步骤:

25 将所述音频特征输入至第一表情分类器中,获取第一表情分类器的表情识别结果以及与所述表情识别结果对应的第一权重;

将所述文字信息输入至第二表情分类器中,获取第二表情分类器的表情识别结果以及与所述表情识别结果对应的第二权重;及

将所述面部特征输入至第三表情分类器中,获取第三表情分类器的表情识别结果以及与所述表情识别结果对应的第三权重。

30 18、根据权利要求 17 所述的存储介质,其特征在于,所述计算机可读指令被所述处理器执行时还执行以下步骤:

将所述音频特征输入至第一表情分类器中,所述第一表情分类器用于从所述音频特征中提取出目标特征,根据所述目标特征查询第一数据库,获取与所述目标特征对应的表情类别,作为表情识别结果,并确定与所述表情识别结果对应的第一权重;及

35 获取所述第一表情分类器的表情识别结果以及与所述表情识别结果对应的第一权重。

19、根据权利要求 17 所述的存储介质，其特征在于，所述计算机可读指令被所述处理器执行时还执行以下步骤：

将所述文字信息输入至第二表情分类器中，所述第二表情分类器用于从所述文字信息中提取出目标信息，根据所述目标信息查询第二数据库，获取与所述目标信息对应的表情类别，作为表情识别结果，并确定与所述表情识别结果对应的第二权重；及

获取所述第二表情分类器的表情识别结果以及与所述表情识别结果对应的第二权重。

20、根据权利要求 17 所述的存储介质，其特征在于，所述计算机可读指令被所述处理器执行时还执行以下步骤：

将所述面部特征输入至第三表情分类器中，所述第三表情分类器用于根据所述面部特征查询第三数据库，获取与所述面部特征对应的表情类别，作为表情识别结果，并确定与所述表情识别结果对应的第三权重；及

获取所述第三表情分类器的表情识别结果以及与所述表情识别结果对应的第三权重。

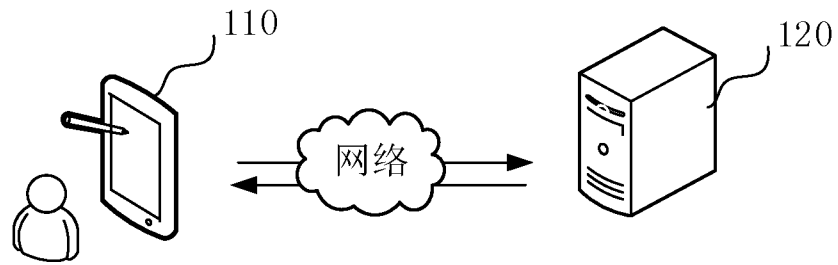


图 1

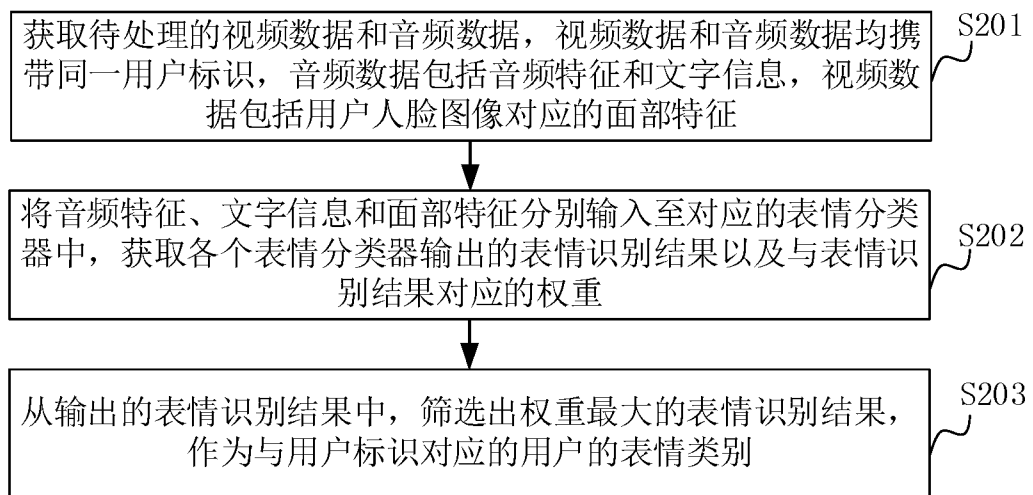


图 2

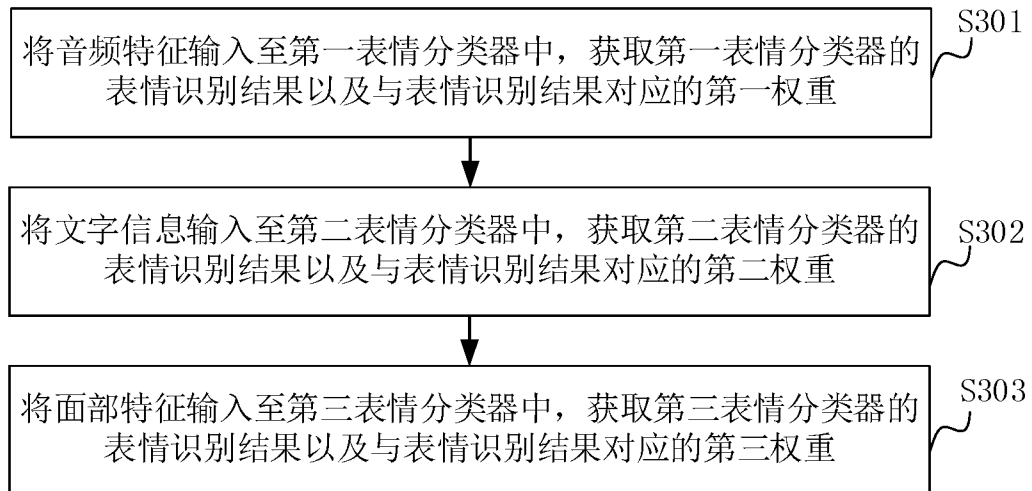


图 3

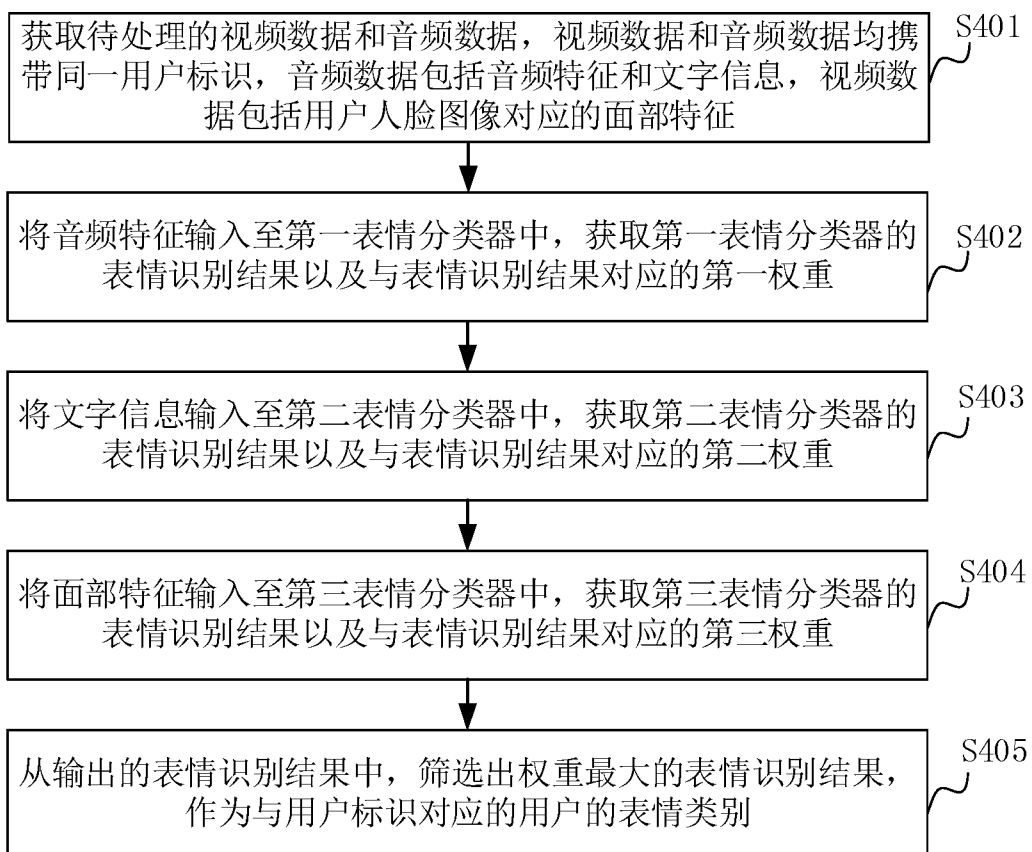


图 4

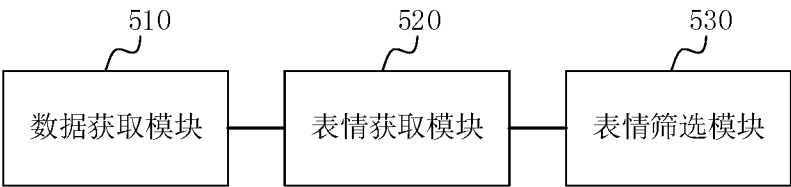


图 5

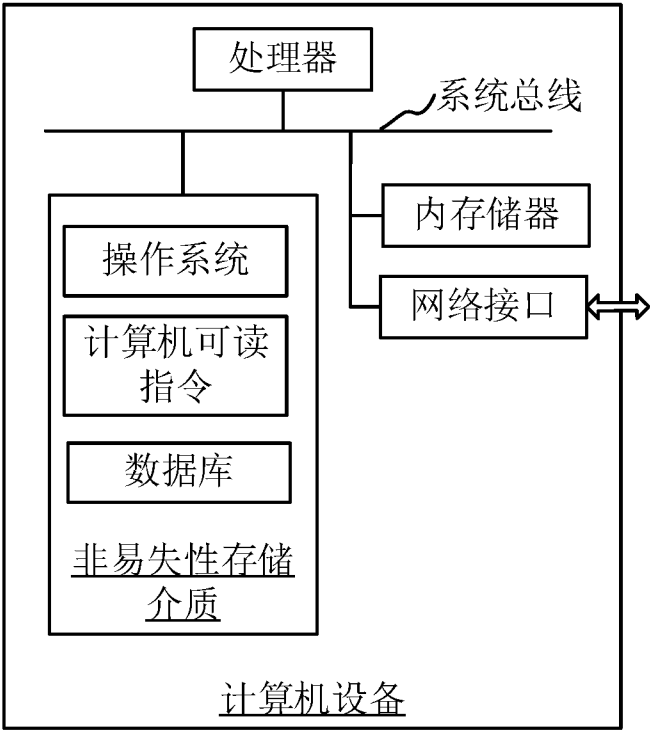


图 6

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/122313

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/00(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

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, CNTXT, CNKI, VEN: 表情, 识别, 视频, 音频, 声音, 数据, 标识, 特征, 文字, 文本, 人脸, 面, 分类, 分类器, 权重, 阈值, expression, recognition, video, audio, voice, data, identification, feature, text, fac+, classif+, weight, threshold

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 109829363 A (ONE CONNECT SMART TECHNOLOGY CO., LTD. (SHENZHEN)) 31 May 2019 (2019-05-31) description, paragraphs 4-128	1-20
X	CN 108805089 A (NANJING XINKTECH INFORMATION TECHNOLOGY CO., LTD.) 13 November 2018 (2018-11-13) description, paragraphs 9-39	1-20
A	CN 108268838 A (FUJIAN INSTITUTE OF RESEARCH ON THE STRUCTURE OF MATTER, CHINESE ACADEMY OF SCIENCES) 10 July 2018 (2018-07-10) entire document	1-20
A	CN 107256392 A (NANJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS) 17 October 2017 (2017-10-17) entire document	1-20
A	CN 107818785 A (PING AN PUHUI ENTERPRISE MANAGEMENT CO., LTD.) 20 March 2018 (2018-03-20) 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

21 February 2020

Date of mailing of the international search report

05 March 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/CN2019/122313

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	109829363	A	31 May 2019	None	
CN	108805089	A	13 November 2018	None	
CN	108268838	A	10 July 2018	None	
CN	107256392	A	17 October 2017	None	
CN	107818785	A	20 March 2018	None	

国际检索报告

国际申请号

PCT/CN2019/122313

A. 主题的分类 G06K 9/00 (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 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用)) CNABS, CNTXT, CNKI, VEN, 表情, 识别, 视频, 音频, 声音, 数据, 标识, 特征, 文字, 文本, 人脸, 面, 分类, 分类器, 权重, 阈值, expression, recognition, video, audio, voice, data, identification, feature, text, face, classif+, weight, threshold																				
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 109829363 A (深圳壹账通智能科技有限公司) 2019年 5月 31日 (2019 - 05 - 31) 说明书第4-128段</td> <td>1-20</td> </tr> <tr> <td>X</td> <td>CN 108805089 A (南京云思创智信息科技有限公司) 2018年 11月 13日 (2018 - 11 - 13) 说明书第9-39段</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 108268838 A (中国科学院福建物质结构研究所) 2018年 7月 10日 (2018 - 07 - 10) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 107256392 A (南京邮电大学) 2017年 10月 17日 (2017 - 10 - 17) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 107818785 A (平安普惠企业管理有限公司) 2018年 3月 20日 (2018 - 03 - 20) 全文</td> <td>1-20</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 109829363 A (深圳壹账通智能科技有限公司) 2019年 5月 31日 (2019 - 05 - 31) 说明书第4-128段	1-20	X	CN 108805089 A (南京云思创智信息科技有限公司) 2018年 11月 13日 (2018 - 11 - 13) 说明书第9-39段	1-20	A	CN 108268838 A (中国科学院福建物质结构研究所) 2018年 7月 10日 (2018 - 07 - 10) 全文	1-20	A	CN 107256392 A (南京邮电大学) 2017年 10月 17日 (2017 - 10 - 17) 全文	1-20	A	CN 107818785 A (平安普惠企业管理有限公司) 2018年 3月 20日 (2018 - 03 - 20) 全文	1-20
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																		
PX	CN 109829363 A (深圳壹账通智能科技有限公司) 2019年 5月 31日 (2019 - 05 - 31) 说明书第4-128段	1-20																		
X	CN 108805089 A (南京云思创智信息科技有限公司) 2018年 11月 13日 (2018 - 11 - 13) 说明书第9-39段	1-20																		
A	CN 108268838 A (中国科学院福建物质结构研究所) 2018年 7月 10日 (2018 - 07 - 10) 全文	1-20																		
A	CN 107256392 A (南京邮电大学) 2017年 10月 17日 (2017 - 10 - 17) 全文	1-20																		
A	CN 107818785 A (平安普惠企业管理有限公司) 2018年 3月 20日 (2018 - 03 - 20) 全文	1-20																		
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																				
<table border="0"> <tr> <td style="vertical-align: top;"> * 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 </td> <td style="vertical-align: top;"> “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 </td> </tr> </table>			* 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																
* 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 21日		国际检索报告邮寄日期 2020年 3月 5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		受权官员 史江峰 电话号码 62089926																		

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/122313

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	109829363	A	2019年 5月 31日	无	
CN	108805089	A	2018年 11月 13日	无	
CN	108268838	A	2018年 7月 10日	无	
CN	107256392	A	2017年 10月 17日	无	
CN	107818785	A	2018年 3月 20日	无	