



(12) 发明专利

(10) 授权公告号 CN 109726624 B

(45) 授权公告日 2023.07.11

(21) 申请号 201810371277.3

(22) 申请日 2018.04.24

(65) 同一申请的已公布的文献号
申请公布号 CN 109726624 A

(43) 申请公布日 2019.05.07

(30) 优先权数据
15/799,512 2017.10.31 US

(73) 专利权人 百度(美国)有限责任公司
地址 美国加利福尼亚州

(72) 发明人 仲震宇 韦韬 邱喆彬

(74) 专利代理机构 北京英赛嘉华知识产权代理
有限责任公司 11204
专利代理师 马晓亚 王艳春

(51) Int.Cl.

G06F 21/32 (2013.01)

G06V 40/70 (2022.01)

G06V 40/16 (2022.01)

G10L 17/06 (2013.01)

(56) 对比文件

CN 106790260 A, 2017.05.31

US 2011007174 A1, 2011.01.13

刘庆辉;姚鸿勋.基于唇动的说话人识别技术.计算机工程与应用.2006,(第12期),全文.

审查员 赵欣

权利要求书3页 说明书15页 附图6页

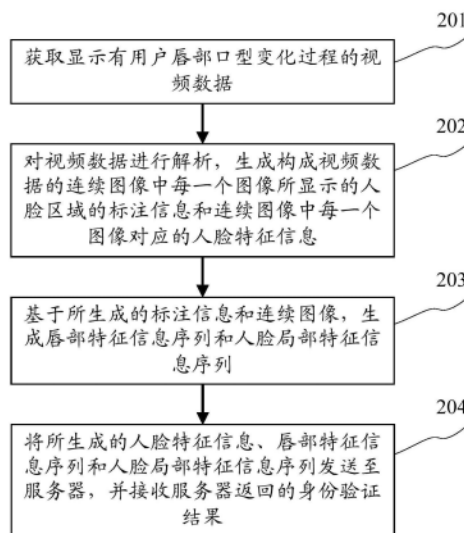
(54) 发明名称

身份认证方法、终端设备和计算机可读存储介质

(57) 摘要

本申请实施例公开了身份认证方法、终端设备和计算机可读存储介质。该方法的一具体实施方式包括:获取显示有用户的唇部口型变化过程的视频数据;对该视频数据进行解析,生成构成该视频数据的连续图像中每一个图像所显示的人脸区域的标注信息和该连续图像中每一个图像对应的人脸特征信息;基于所生成的标注信息和该连续图像,生成唇部特征信息序列和人脸局部特征信息序列;将所生成的人脸特征信息、该唇部特征信息序列和该人脸局部特征信息序列发送至服务器,并接收该服务器返回的身份验证结果。该实施方式提高了身份认证的安全性。

200



1. 一种身份认证方法,所述方法包括:

获取显示有用户的唇部口型变化过程的视频数据;

对所述视频数据进行解析,生成构成所述视频数据的连续图像中每一个图像所显示的人脸区域的标注信息和所述连续图像中每一个图像对应的人脸特征信息,所述标注信息包括所述每一个图像中的所述人脸区域的位置和所述人脸区域中的器官的位置;

基于所生成的标注信息和所述连续图像,生成唇部特征信息序列和人脸局部特征信息序列;

将所生成的人脸特征信息、所述唇部特征信息序列和所述人脸局部特征信息序列发送至服务器,并接收所述服务器返回的身份验证结果。

2. 根据权利要求1所述的身份认证方法,其中,所述获取显示有用户的唇部口型变化过程的视频数据,包括:

显示用于提示用户进行身份认证的第一提示信息,对所述用户的脸部进行拍摄以获取显示有所述用户的唇部口型变化过程的视频数据,并获取所述用户的语音数据。

3. 根据权利要求2所述的身份认证方法,其中,所述第一提示信息包括发音选项;以及所述方法还包括:

响应于确定所述用户对所述发音选项执行了点击操作,对所述语音数据进行解析,生成声纹特征信息;

将所述声纹特征信息发送至所述服务器。

4. 根据权利要求1所述的身份认证方法,其中,

所述基于所生成的标注信息和所述连续图像,生成唇部特征信息序列和人脸局部特征信息序列,包括:

将所生成的标注信息和所述连续图像分别输入至预先训练的神经网络中,得到与所述连续图像相对应的唇部特征信息序列和人脸局部特征信息序列,其中,所述神经网络用于依次从多个图像中提取唇部特征信息以生成唇部特征信息序列和依次从多个图像中提取人脸局部特征信息以生成人脸局部特征信息序列。

5. 根据权利要求1所述的身份认证方法,其中,所述人脸局部特征信息序列包括以下至少一项:面部特征信息序列、眼睛特征信息序列、鼻子特征信息序列、眉毛特征信息序列、下巴特征信息序列、口腔特征信息序列、舌部特征信息序列。

6. 根据权利要求1所述的身份认证方法,其中,所述唇部特征信息序列中的唇部特征信息包括:嘴唇的中心坐标、嘴唇方向、内唇宽度、外唇宽度、外唇上边缘高度、内唇上边缘高度、内唇下边缘高度、外唇下边缘高度、四次曲线中心离坐标原点的偏移量、上唇部四次曲线偏离抛物线的距离、下唇部四次曲线偏离抛物线的距离。

7. 根据权利要求1所述的身份认证方法,其中,在所述获取显示有用户的唇部口型变化过程的视频数据之前,所述方法还包括:

显示用于提示用户进行暗语设置的第二提示信息,并获取显示有所述用户的唇部口型变化过程的目标视频数据;

对所述目标视频数据进行解析,生成构成所述目标视频数据的目标连续图像中的每一个图像所显示的人脸区域的目标标注信息和所述目标连续图像中每一个图像对应的目标人脸特征信息;

基于所生成的目标标注信息和所述目标连续图像,生成目标唇部特征信息序列和目标人脸局部特征信息序列;

将所生成的目标人脸特征信息、所述目标唇部特征信息序列和所述目标人脸局部行为特征信息序列发送至服务器。

8. 根据权利要求7所述的身份认证方法,其中,所述显示用于提示用户进行暗语设置的第二提示信息,并获取显示有所述用户的唇部口型变化过程的目标视频数据,包括:

显示用于提示用户进行暗语设置的第二提示信息,对所述用户的脸部进行拍摄以获取显示有所述用户的唇部口型变化过程的目标视频数据,并获取所述用户的目标声音数据。

9. 根据权利要求8所述的身份认证方法,其中,所述第二提示信息包含发音选项;以及所述方法还包括:

响应于确定所述用户对所述第二提示信息中的发音选项执行了点击操作,对所述目标声音数据进行解析,生成目标声纹特征信息;

将所述目标声纹特征信息发送至所述服务器。

10. 根据权利要求7所述的身份认证方法,其中,所述目标标注信息包括人脸区域的关键点的位置和人脸区域的位置;以及

所述基于所生成的目标标注信息和所述目标连续图像,生成目标唇部特征信息序列和目标人脸局部特征信息序列,包括:

将所生成的目标标注信息和所述目标连续图像分别输入至预先训练的神经网络中,得到与所述目标连续图像相对应的目标唇部特征信息序列和目标人脸局部特征信息序列。

11. 一种身份认证方法,所述方法包括:

接收客户端发送的身份认证请求,其中,所述身份认证请求包含用户的人脸特征信息、唇部特征信息序列、人脸局部特征信息序列,所述人脸特征信息、所述唇部特征信息序列和所述人脸局部特征信息序列为所述客户端基于显示有所述用户唇部口型变化过程的视频数据的解析而生成的,其中,所述唇部特征信息序列和所述人脸局部特征信息序列基于构成所述视频数据的连续图像中每一个图像所显示的人脸区域的标注信息而生成,所述标注信息包括所述每一个图像中的所述人脸区域的位置和所述人脸区域中的器官的位置;

提取预置的、所述用户的目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列;

基于所述人脸特征信息、所述唇部特征信息序列、所述人脸局部特征信息序列与所述目标人脸特征信息、所述目标唇部特征信息序列、所述目标人脸局部特征信息序列的匹配结果,生成用于表征所述用户是否通过身份认证的身份认证结果;

向所述客户端发送所述身份认证结果。

12. 根据权利要求11所述的身份认证方法,其中,所述基于所述人脸特征信息、所述唇部特征信息序列和所述人脸局部特征信息序列与所述目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列的匹配结果,生成用于表征所述用户是否通过身份认证的身份认证结果,包括:

在以下三个条件同时满足的情况下,生成用于表征所述用户通过身份认证的身份认证结果:所述人脸特征信息与所述目标人脸特征信息相匹配;所述唇部特征信息序列与所述目标唇部特征信息序列相匹配;所述人脸局部特征信息序列与所述目标人脸局部特征信息

序列相匹配。

13. 根据权利要求11所述的身份认证方法,其中,所述身份认证请求还包括所述用户的声纹特征信息;以及

所述基于所述人脸特征信息、所述唇部特征信息序列和所述人脸局部特征信息序列与所述目标人脸特征信息、所述目标唇部特征信息序列和所述目标人脸局部特征信息序列的匹配结果,生成用于表征所述用户是否通过身份认证的身份认证结果,包括:

提取预置的、所述用户的目标声纹特征信息;

在以下四个条件同时满足的情况下,生成用于表征所述用户通过身份认证的身份认证结果:所述人脸特征信息与所述目标人脸特征信息相匹配;所述唇部特征信息序列与所述目标唇部特征信息序列相匹配;所述人脸局部特征信息序列与所述目标人脸局部特征信息序列相匹配;所述声纹特征信息与所述目标声纹特征信息相匹配。

14. 根据权利要求11所述的身份认证方法,其中,在所述接收客户端发送的身份认证请求之前,所述方法包括:

接收客户端发送的身份设定请求,其中,所述身份设定请求包含用户的目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列,所述目标人脸特征信息、所述目标唇部特征信息序列和所述目标人脸局部特征信息序列为所述客户端基于显示有所述用户唇部口型变化过程的目标视频数据的解析而生成的;

将所述目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列进行存储。

15. 根据权利要求14所述的身份认证方法,其中,所述身份设定请求还包括所述用户的目标声纹特征信息,其中,所述目标声纹特征信息是所述客户端基于对所述用户的语音数据的解析而生成的;以及

所述将所述目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列进行存储,包括:

将所述目标声纹特征信息、所述目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列进行存储。

16. 一种终端设备,包括:

一个或多个处理器;

存储装置,用于存储一个或多个程序,

当所述一个或多个程序被所述一个或多个处理器执行,使得所述一个或多个处理器实现如权利要求1所述的方法。

17. 一种计算机可读存储介质,其上存储有计算机程序,该程序被处理器执行时实现如权利要求1所述的方法。

身份认证方法、终端设备和计算机可读存储介质

技术领域

[0001] 本申请实施例涉及计算机技术领域,具体涉及互联网技术领域,尤其涉及身份认证方法、终端设备和计算机可读存储介质。

背景技术

[0002] 随着全球经济与信息技术的发展以及全球互联网时代的到来,越来越多的领域需要可靠的身份认证。在信息化的背景下,如何准确地鉴别一个人的身份、保证信息安全,是信息化时代的一个重要挑战。

[0003] 现有的方法通常是利用口令、密钥或人脸识别技术等进行身份认证,然而,口令、密钥容易被用户忘记,人脸识别技术所依赖的人脸特征容易被照片或视频等伪造。

发明内容

[0004] 本申请实施例的目的在于提出一种身份认证方法、终端设备和计算机可读存储介质。

[0005] 第一方面,本申请实施例提供了一种身份认证方法,该方法包括:获取显示有用户的唇部口型变化过程的视频数据;对视频数据进行解析,生成构成视频数据的连续图像中每一个图像所显示的人脸区域的标注信息和连续图像中每一个图像对应的人脸特征信息;基于所生成的标注信息和连续图像,生成唇部特征信息序列和人脸局部特征信息序列;将所生成的人脸特征信息、唇部特征信息序列和人脸局部特征信息序列发送至服务器,并接收服务器返回的身份验证结果。

[0006] 在一些实施例中,获取显示有用户的唇部口型变化过程的视频数据,包括:显示用于提示用户进行身份认证的第一提示信息,对用户的脸部进行拍摄以获取显示有用户的唇部口型变化过程的视频数据,并获取用户的语音数据。

[0007] 在一些实施例中,第一提示信息包括发音选项;该方法还包括:响应于确定用户对发音选项执行了点击操作,对语音数据进行解析,生成声纹特征信息;将声纹特征信息发送至服务器。

[0008] 在一些实施例中,标注信息包括人脸区域的关键点的位置和人脸区域的位置;以及基于所生成的标注信息和连续图像,生成唇部特征信息序列和人脸局部特征信息序列,包括:将所生成的标注信息和连续图像分别输入至预先训练的神经网络中,得到与连续图像相对应的唇部特征信息序列和人脸局部特征信息序列,其中,神经网络用于依次从多个图像中提取唇部特征信息以生成唇部特征信息序列和依次从多个图像中提取人脸局部特征信息以生成人脸局部特征信息序列。

[0009] 在一些实施例中,人脸局部特征信息序列包括以下至少一项:面部特征信息序列、眼睛特征信息序列、鼻子特征信息序列、眉毛特征信息序列、下巴特征信息序列、口腔特征信息序列、舌部特征信息序列。

[0010] 在一些实施例中,唇部特征信息序列中的唇部特征信息包括:嘴唇的中心坐标、嘴

唇方向、内唇宽度、外唇宽度、外唇上边缘高度、内唇上边缘高度、内唇下边缘高度、外唇下边缘高度、四次曲线中心离坐标原点的偏移量、上唇部四次曲线偏离抛物线的距离、下唇部四次曲线偏离抛物线的距离。

[0011] 在一些实施例中,在获取显示有用户的唇部口型变化过程的视频数据之前,该方法还包括:显示用于提示用户进行暗语设置的第二提示信息,并获取显示有用户的唇部口型变化过程的目标视频数据;对目标视频数据进行解析,生成构成目标视频数据的目标连续图像中的每一个图像所显示的人脸区域的目标标注信息和目标连续图像中每一个图像对应的目标人脸特征信息;基于所生成的目标标注信息和目标连续图像,生成目标唇部特征信息序列和目标人脸局部特征信息序列;将所生成的目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列发送至服务器。

[0012] 在一些实施例中,显示用于提示用户进行暗语设置的第二提示信息,并获取显示有用户的唇部口型变化过程的目标视频数据,包括:显示用于提示用户进行暗语设置的第二提示信息,对用户的脸部进行拍摄以获取显示有用户的唇部口型变化过程的目标视频数据,并获取用户的目标声音数据。

[0013] 在一些实施例中,第二提示信息包含发音选项;该方法还包括:响应于确定用户对第二提示信息中的发音选项执行了点击操作,对目标声音数据进行解析,生成目标声纹特征信息;将目标声纹特征信息发送至服务器。

[0014] 在一些实施例中,目标标注信息包括人脸区域的关键点的位置和人脸区域的位置;以及基于所生成的目标标注信息和目标连续图像,生成目标唇部特征信息序列和目标人脸局部特征信息序列,包括:将所生成的目标标注信息和目标连续图像分别输入至预先训练的神经网络中,得到与目标连续图像相对应的目标唇部特征信息序列和目标人脸局部特征信息序列。

[0015] 第二方面,本申请实施例提供了一种身份认证方法,该方法包括:接收客户端发送的身份认证请求,其中,身份认证请求包含用户的人脸特征信息、唇部特征信息序列、人脸局部特征信息序列,人脸特征信息、唇部特征信息序列和人脸局部特征信息序列为客户端基于显示有用户唇部口型变化过程的视频数据的解析而生成的;提取预置的、用户的目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列;基于人脸特征信息、唇部特征信息序列、人脸局部特征信息序列与目标人脸特征信息、目标唇部特征信息序列、目标人脸局部特征信息序列的匹配结果,生成用于表征用户是否通过身份认证的身份认证结果;向客户端发送身份认证结果。

[0016] 在一些实施例中,基于人脸特征信息、唇部特征信息序列和人脸局部特征信息序列与目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列的匹配结果,生成用于表征用户是否通过身份认证的身份认证结果,包括:在以下三个条件同时满足的情况下,生成用于表征用户通过身份认证的身份认证结果:人脸特征信息与目标人脸特征信息相匹配;唇部特征信息序列与目标唇部特征信息序列相匹配;人脸局部特征信息序列与目标人脸局部特征信息序列相匹配。

[0017] 在一些实施例中,身份认证请求还包括用户的声纹特征信息;以及基于人脸特征信息、唇部特征信息序列和人脸局部特征信息序列与目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列的匹配结果,生成用于表征用户是否通过身份认证的

身份认证结果,包括:提取预置的、用户的目标声纹特征信息;在以下四个条件同时满足的情况下,生成用于表征用户通过身份认证的身份认证结果:人脸特征信息与目标人脸特征信息相匹配;唇部特征信息序列与目标唇部特征信息序列相匹配;人脸局部特征信息序列与目标人脸局部特征信息序列相匹配;声纹特征信息与目标声纹特征信息相匹配。

[0018] 在一些实施例中,在接收客户端发送的身份认证请求之前,该方法包括:接收客户端发送的身份设定请求,其中,身份设定请求包含用户的目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列,目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列为客户端基于显示有用户唇部口型变化过程的目标视频数据的解析而生成的;将目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列进行存储。

[0019] 在一些实施例中,身份设定请求还包括用户的目标声纹特征信息,其中,目标声纹特征信息是客户端基于对用户的语音数据的解析而生成的;以及将目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列进行存储,包括:将目标声纹特征信息、目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列进行存储。

[0020] 第三方面,本申请实施例提供了一种终端设备,包括:一个或多个处理器;存储装置,用于存储一个或多个程序,当一个或多个程序被一个或多个处理器执行,使得一个或多个处理器实现如上述第一方面所提供的身份认证方法中任一实施例的方法。

[0021] 第四方面,本申请实施例提供了一种计算机可读存储介质,其上存储有计算机程序,该程序被处理器执行时实现如上述第一方面所提供的身份认证方法中任一实施例的方法。

[0022] 本申请实施例提供的身份认证方法,通过对所获取的显示有用户唇部口型变化过程的视频数据进行解析,以便生成构成视频数据的连续图像中每一个图像所显示的人脸区域的标注信息和连续图像中每一个图像对应的人脸特征信息,而后基于所生成的标注信息和连续图像,生成唇部特征信息序列和人脸局部特征信息序列,最后将所生成的人脸特征信息、唇部特征信息序列和人脸局部特征信息序列发送至服务器,并接收服务器返回的身份验证结果,从而可以结合人脸的面部识别、用户使用唇语时的嘴型动作以及用户使用唇语时的脸部局部行为的识别进行用户身份认证,由于用户使用唇语时的嘴型动作以及脸部行为是用户自身独有的生物特征,攻击者不易效仿,因此,提高了身份认证的安全性。

附图说明

[0023] 通过阅读参照以下附图所作的对非限制性实施例所作的详细描述,本申请的其它特征、目的和优点将会变得更明显:

[0024] 图1是本申请可以应用于其中的示例性系统架构图;

[0025] 图2是根据本申请的身份认证方法的一个实施例的流程图;

[0026] 图3是根据本申请的身份认证方法的一个应用场景的示意图;

[0027] 图4是根据本申请的身份认证方法的又一个实施例的流程图;

[0028] 图5是根据本申请的身份认证方法的再一个实施例的流程图;

[0029] 图6是根据本申请的身份认证装置的一个实施例的结构示意图;

[0030] 图7是根据本申请的身份认证装置的又一个实施例的结构示意图;

[0031] 图8是适于用来实现本申请实施例的终端设备或服务器的计算机系统的结构示意图。

具体实施方式

[0032] 下面结合附图和实施例对本申请作进一步的详细说明。可以理解的是,此处所描述的具体实施例仅仅用于解释相关发明,而非对该发明的限定。另外还需要说明的是,为了便于描述,附图中仅示出了与有关发明相关的部分。

[0033] 需要说明的是,在不冲突的情况下,本申请中的实施例及实施例中的特征可以相互组合。下面将参考附图并结合实施例来详细说明本申请。

[0034] 图1示出了可以应用本申请的身份认证方法的示例性系统架构100。

[0035] 如图1所示,系统架构100可以包括终端设备101、102、103,网络104和服务器105。网络104用以在终端设备101、102、103和服务器105之间提供通信链路的介质。网络104可以包括各种连接类型,例如有线、无线通信链路或者光纤电缆等等。

[0036] 用户可以使用终端设备101、102、103通过网络104与服务器105交互,以接收或发送消息等。终端设备101、102、103上可以安装有各种通讯客户端应用,例如网页浏览器应用、购物类应用、搜索类应用、即时通信工具、邮箱客户端、社交平台软件等。

[0037] 终端设备101、102、103可以是具有摄像装置和显示屏幕、并且支持网络传输的各种电子设备,包括但不限于智能手机、平板电脑、电子书阅读器、MP3播放器(Moving Picture Experts Group Audio Layer III,动态影像专家压缩标准音频层面3)、MP4(Moving Picture Experts Group Audio Layer IV,动态影像专家压缩标准音频层面4)播放器、膝上型便携计算机和台式计算机等等。

[0038] 服务器105可以是提供各种服务的服务器,例如用于对使用终端设备101、102、103的用户进行身份认证的后台服务器。后台服务器可以对接收到的信息(例如人脸特征信息、上述唇部特征信息序列、上述人脸局部特征信息序列等)进行分析、匹配等处理,并将处理结果(例如身份认证结果)反馈给终端设备。此处,服务器105不限于物理服务器,还可以是云服务器。

[0039] 需要说明的是,本申请实施例所提供的身份认证方法一般由服务器105执行。

[0040] 应该理解,图1中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需要,可以具有任意数目的终端设备、网络和服务器。

[0041] 继续参考图2,示出了根据本申请的身份认证方法的一个实施例的流程200。该身份认证方法,包括以下步骤:

[0042] 步骤201,获取显示有用户唇部口型变化过程的视频数据。

[0043] 在本实施例中,身份验证方法运行于其上的电子设备(例如图1所示的终端设备101、102、103)可以安装有摄像装置,例如摄像头。上述电子设备可以开启摄像装置进行视频采集,以获取显示有用户唇部口型变化过程的视频数据。此处,上述视频数据是指连续的图像序列,其实质是由一组组连续的图像构成的。

[0044] 实践中,上述电子设备在视频采集过程中可以实时将所采集的每一帧图像进行显示。用户可以基于上述电子设备所显示的影像调整头部或上述电子设备的位置,以使上述电子设备所采集的视频显示有上述用户的脸部区域。此外,在视频采集过程中,上述用户可

以以发音或者不发音(唇语)的方式说出预先设定的暗语(例如一个短语、任意一句话、或几个发音等),以使上述电子设备所采集的视频数据显示有上述用户唇部的口型变化过程。

[0045] 步骤202,对视频数据进行解析,生成构成视频数据的连续图像中每一个图像所显示的人脸区域的标注信息和连续图像中每一个图像对应的人脸特征信息。

[0046] 在本实施例中,上述电子设备可以对上述视频数据进行解析,生成构成上述视频数据的连续图像中每一个图像所显示的人脸区域的标注信息和上述连续图像中每一个图像对应的人脸特征信息。其中,标注信息可以包括图像中人脸区域中的各个器官(例如眼睛、眉毛、鼻子、嘴巴、下巴等)位于人脸区域的位置(可以用坐标值进行表示)、人脸区域位于图像的位置等信息。需要说明的是,上述人脸特征信息可以是用于对图像中的人脸特征进行表征的信息,其中,上述人脸特征可以是与人脸相关的各种基本要素。例如,上述基本要素可以包括但不限于人脸中的各器官位于图像的人脸区域中的位置、人脸中各个器官之间的结构关系,以及多个点之间的欧氏距离、曲率和角度等。实践中,上述人脸特征信息可以用向量或矩阵表示。

[0047] 需要指出的是,上述电子设备可以利用各种人脸检测方法对上述连续图像进行检测,生成人脸区域的标注信息和人脸特征信息。作为示例,上述电子设备可以将上述连续图像输入至预先训练的、用于进行人脸检测的卷积神经网络(Convolutional Neural Network, CNN)中,得到上述卷积神经网络输出的、与上述连续图像中的每一个图像所显示的人脸区域的标注信息和该图像对应的人脸特征信息。需要说明的是,上述卷积神经网络可以是上述电子设备利用深度学习和训练样本对现有的深度卷积神经网络(例如DenseBox、VGGNet、ResNet、SegNet等)进行有监督训练而得到的。其中,上述训练样本可以包括大量的图像和每一个图像的标注,例如,用于指示是否是人脸区域的标注、用于指示人脸区域的位置的标注(例如坐标值等)、用于指示是否是人脸中各个器官的标注和用于指示各个器官位于人脸区域的位置的标注等。此处,标注信息和人脸特征信息可以用向量或矩阵的形式表示。

[0048] 实践中,卷积神经网络是一种前馈神经网络,它的人工神经元可以响应一部分覆盖范围内的周围单元,对于图像处理有出色表现,因而,可以利用卷积神经网络进行人脸特征的提取和人脸检测。卷积神经网络可以包含至少一个卷积层和至少一个池化层,其中,卷积层可以用于提取图像特征,池化层可以用于对输入的信息进行降采样(downsample)。需要说明的是,利用卷积神经网络进行人脸检测和特征提取是目前广泛研究和应用的公知技术,在此不再赘述。

[0049] 步骤203,基于所生成的标注信息和连续图像,生成唇部特征信息序列和人脸局部特征信息序列。

[0050] 在本实施例中,上述电子设备可以基于所生成的标注信息和上述连续图像,生成唇部特征信息序列和人脸局部特征信息序列。具体地,可以首先对上述连续图像的每一个图像对应的标注信息进行提取、计算、汇总等处理;而后,将处理后的标注信息确定为该图像的唇部特征信息和人脸局部特征信息;最后,将所确定的、与各个图像对应的唇部特征信息依次汇总成唇部特征信息序列,并将所确定的、与各个图像对应的人脸局部特征信息依次汇总成人脸局部特征信息序列。作为示例,由于所生成的标注信息中可以包括图像中人脸区域中的各个器官位于人脸区域的位置和人脸区域位于图像的位置等信息,因此,上述

电子设备可以提取上述连续图像中的每一个图像的标注信息中的、与唇部相关的标注信息(例如嘴角、唇峰等位置的坐标),将所提取标注信息作为唇部特征信息进行计算、汇总(例如计算唇部宽度、高度等),生成唇部特征信息序列。同时,上述电子设备可以提取上述连续图像中的每一个图像的标注信息中的、与人脸局部(例如眼睛、鼻子、面部、眉毛、下巴、口腔、舌部中的一项或多项)相关的标注信息,对所提取的标注信息进行计算、汇总,生成人脸局部特征信息序列。实践中,人脸局部特征信息序列可以是一个或多个,若为多个,则每个人脸局部特征信息序列可以对应一个人脸局部。

[0051] 需要说明的是,上述唇部特征信息可以是用于对唇部特征进行表征的信息,其中,唇部特征可以是与人脸区域中的唇部区域相关的各种基本要素。例如,可以包括但不限于用户唇部多个位置(例如嘴角、唇峰等)的坐标值、唇部的尺寸(例如唇部宽度、高度、厚度等)等信息。唇部特征信息序列可以是由上述连续图像中的各个图像中对应的唇部特征信息所依次构成的序列。实践中,唇部特征信息可以用向量或矩阵表示。

[0052] 需要指出的是,上述人脸局部特征信息是用于对人脸局部特征进行表征的信息,其中,上述人脸局部特征可以是与人脸中的局部(例如眼睛、鼻子、下巴、眉毛、口腔、舌部、脸部等)相关的各种基本要素。例如,可以包括但不限于人脸中的某个局部位于图像的人脸区域中的位置、人脸中某个局部的关键点的欧氏距离、曲率和角度等。上述人脸局部特征信息序列可以是由上述连续图像中的各个图像中对应的人脸局部特征信息所依次构成的序列。实践中,上述人脸局部特征信息可以用向量或矩阵表示。

[0053] 步骤204,将所生成的人脸特征信息、唇部特征信息序列和人脸局部特征信息序列发送至服务器,并接收服务器返回的身份验证结果。

[0054] 在本实施例中,上述电子设备可以将所生成的人脸特征信息、唇部特征信息序列和人脸局部特征信息序列发送至服务器(例如图1所示的服务器105),并接收服务器返回的身份验证结果。

[0055] 实践中,上述服务器中可以存储有预置的、上述用户的目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列。上述目标人脸特征信息、上述目标唇部特征信息序列和上述目标人脸局部特征信息序列可以是上述电子设备基于显示有用户唇部口型变化过程的目标视频数据的解析而生成的,其中,上述目标视频数据可以是所述电子设备在该用户进行暗语设置的过程中所获取的、显示有该用户的唇部口型变化过程的视频数据。此处,暗语设置是在该身份认证方法之前预先完成的。上述服务器可以基于上述人脸特征信息、上述唇部特征信息序列和上述人脸局部特征信息序列与上述目标人脸特征信息、上述目标唇部特征信息序列和上述目标人脸局部特征信息序列的匹配结果,生成用于表征上述用户是否通过身份认证的身份认证结果,并向上述电子设备发送上述身份认证结果。具体地,上述服务器可以在以下三个条件同时满足的情况下,生成用于表征上述用户通过身份认证的身份认证结果:上述人脸特征信息与上述目标人脸特征信息相匹配;上述唇部特征信息序列与上述目标唇部特征信息序列相匹配;上述人脸局部特征信息序列与上述目标人脸局部特征信息序列相匹配。

[0056] 需要说明的是,上述服务器可以预先接收并存储上述电子设备发送的身份设定请求,其中,上述身份设定请求可以包含上述用户的目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列,上述目标人脸特征信息、上述目标唇部特征信息序列和

上述目标人脸局部特征信息序列可以为上述电子设备基于显示有上述用户唇部口型变化过程的目标视频数据的解析而生成的。

[0057] 上述服务器可以利用各种匹配方式进行上述人脸特征信息、上述唇部特征信息序列、上述人脸局部特征信息序列与上述目标人脸特征信息、上述目标唇部特征信息序列、上述目标人脸局部特征信息序列的匹配。作为示例，上述服务器中可以存储有分别用于进行人脸识别、唇部行为识别和人脸局部行为识别的三个神经网络。上述服务器可以分别将上述人脸特征信息、上述唇部特征信息序列和上述人脸局部特征信息序列输入至相应的神经网络中，并得到上述三个神经网络分别输出的人脸识别结果、唇部行为识别结果、人脸局部行为识别结果。需要说明的是，上述三个神经网络可以是利用机器学习方法，并基于大量的训练样本进行有监督训练而得到的。用于进行人脸识别的神经网络所使用的训练样本中可以包含上述目标人脸特征信息；用于进行唇部行为识别的神经网络所使用的训练样本中可以包含上述目标唇部特征信息序列；用于进行人脸局部行为识别的神经网络所使用的训练样本中可以包含上述目标人脸局部特征信息序列。需要说明的是，进行人脸识别、唇部行为识别和人脸局部行为识别的神经网络的数量不限于三个，还可以是一个、两个或多个。

[0058] 继续参见图3，图3是根据本实施例的身份认证方法的应用场景的一个示意图。在图3的应用场景中，终端设备301首先获取了显示有用户的唇部口型变化过程的视频数据；而后，终端设备301对上述视频数据进行解析，生成构成上述视频数据的连续图像中每一个图像所显示的人脸区域的标注信息和上述连续图像中每一个图像对应的人脸特征信息302；之后，终端设备301基于所生成的标注信息和上述连续图像，生成唇部特征信息序列303和人脸局部特征信息序列304；然后，终端设备301将所生成的人脸特征信息302、上述唇部特征信息序列303和上述人脸局部特征信息序列304发送至服务器305；最后，终端设备301接收上述服务器305返回的身份验证结果306。

[0059] 本申请的上述实施例提供的方法，通过对所获取的显示有用户唇部口型变化过程的视频数据进行解析，以便生成构成视频数据的连续图像中每一个图像所显示的人脸区域的标注信息和连续图像中每一个图像对应的人脸特征信息，而后基于所生成的标注信息和连续图像，生成唇部特征信息序列和人脸局部特征信息序列，最后将所生成的人脸特征信息、唇部特征信息序列和人脸局部特征信息序列发送至服务器，并接收服务器返回的身份验证结果，从而可以结合人脸的面部识别、用户使用唇语时的嘴型动作以及用户使用唇语时的脸部局部行为的识别进行用户身份认证，即将用户所拥有的（即人脸特征信息）、所知道的（即预先设定的暗语）和用户的行为特征（即用户在进行暗语描述时的唇部特征信息序列和人脸局部特征信息序列）等多方面因素有效结合。即使用户预先设定的暗语被攻击者效仿，但由于用户使用唇语时的嘴型以及脸部行为是用户自身独有的生物特征，攻击者不易效仿，因此，提高了身份认证的安全性。

[0060] 进一步参考图4，其示出了身份认证方法的又一个实施例的流程400。该身份认证方法的流程400，包括以下步骤：

[0061] 步骤401，显示用于提示用户进行身份认证的第一提示信息，对用户的脸部进行拍摄以获取显示有用户的唇部口型变化过程的视频数据，并获取用户的声音数据。

[0062] 在本实施例中，身份验证方法运行于其上的电子设备（例如图1所示的终端设备101、102、103）可以在显示屏幕中显示用于提示用户进行身份认证的第一提示信息。此处，

上述第一提示信息可以提示用户通过对预先设定的暗语进行描述的方式进行身份验证。上述电子设备可以安装有摄像装置,也可以安装有麦克风等装置。上述电子设备可以开启摄像装置和麦克风进行视频数据的采集和声音数据的采集。在此过程中,用户可以调整头部或上述电子设备的位置、方向等,以使上述电子设备对用户的脸部区域进行拍摄,进而获取显示有用户的唇部的口型变化过程的视频数据。

[0063] 步骤402,响应于确定用户对发音选项执行了点击操作,对声音数据进行解析,生成声纹特征信息。

[0064] 在本实施例中,上述第一提示信息可以包含发音选项,上述发音选项可以用于提示用户以发出声音的方式描述预先设定的暗语。响应于确定上述用户对上述发音选项执行了点击操作,上述电子设备可以利用各种声纹特征提取方法对上述声音数据进行解析,生成声纹特征信息。作为示例,上述电子设备可以将上述声音数据输入至预先训练的用于进行声纹特征提取的神经网络模型(例如现有的用于进行声纹特征提取的DNN-ivector模型或者现有的用于进行声纹特征提取的端对端深度神经网络模型),得到上述神经网络模型输出的声纹特征信息。上述声纹特征信息可以是用于对声纹特征进行表征的信息,其中,上述声纹特征可以是与声纹相关的各种基本要素。例如,可以包括音调、响度、频率、音色等。实践中,上述声纹特征信息可以用向量或矩阵表示。此处,利用神经网络模型提取声纹特征信息的方法是是目前广泛研究和应用的公知技术,在此不再赘述。

[0065] 实践中,用户可以在进行暗语描述之前或之后点击上述发音选项。若确定用户点击了上述发音选项,则上述电子设备可以对上述声音数据进行解析生成声纹特征信息;若确定用户未点击上述发音选项,则上述电子设备可以不执行对声音数据解析的步骤。

[0066] 在本实施例的一些可选的实现方式中,上述第一提示信息也可以包含不发音选项,上述不发音选项可以用于提示用户以不发音的方式描述暗语(即使用唇语进行描述)。若确定用户点击了上述不发音选项,则上述电子设备可以不执行对声音数据解析的步骤。

[0067] 在本实施例的一些可选的实现方式中,若上述电子设备在暗语描述之前点击了上述不发音选项,则上述电子设备可以不进行声音数据的获取,直接对上述用户的脸部进行拍摄,获取显示有上述用户的唇部口型变化过程的视频数据。

[0068] 步骤403,对视频数据进行解析,生成构成视频数据的连续图像中每一个图像所显示的人脸区域的标注信息和连续图像中每一个图像对应的人脸特征信息。

[0069] 在本实施例中,上述电子设备可以对上述视频数据进行解析,生成构成上述视频数据的连续图像中每一个图像所显示的人脸区域的标注信息和上述连续图像中每一个图像对应的人脸特征信息。其中,标注信息可以包括人脸区域的关键点的位置和人脸区域的位置。上述关键点可以是预先指定的、人脸中具有较强语义信息的点(例如眼角、嘴角、鼻翼、眉尖、眉峰、轮廓中的点等)。实践中,人脸关键点的数量可以是72个,也可以是其他预设的数值,在此不做限定。需要说明的是,标注信息还可以包括但不限于关键点的类型(例如眼角、嘴角、鼻翼、眉尖、眉峰、轮廓等)、是关键点的概率等信息。

[0070] 在本实施例中,上述电子设备可以将上述连续图像输入至预先训练的、用于进行人脸检测的卷积神经网络中,使上述卷积神经网络输出与上述连续图像中的每一个图像所显示的人脸区域的标注信息和该图像对应的人脸特征信息。需要说明的是,上述步骤403与上述步骤202的操作基本相同,在此不再赘述。

[0071] 步骤404,基于所生成的标注信息和连续图像,生成唇部特征信息序列和人脸局部特征信息序列。

[0072] 在本实施例中,上述电子设备可以基于所生成的标注信息和连续图像,生成唇部特征信息序列和人脸局部特征信息序列。需要说明的是,上述唇部特征信息序列中的唇部特征信息可以包括但不限于:嘴唇的中心坐标、嘴唇方向、内唇宽度、外唇宽度、外唇上边缘高度、内唇上边缘高度、内唇下边缘高度、外唇下边缘高度、四次曲线中心离坐标原点的偏移量、上唇部四次曲线偏离抛物线的距离、下唇部四次曲线偏离抛物线的距离。需要指出的是,上述人脸局部特征信息序列可以包括以下至少一项:面部特征信息序列、眼睛特征信息序列、鼻子特征信息序列、眉毛特征信息序列、下巴特征信息序列、口腔特征信息序列、舌部特征信息序列。此处,每个部位特征信息序列可以是用于对该部位特征进行表征的信息。作为示例,眉毛特征序列可以是用于对眉毛特征进行表征的信息,其中,眉毛特征可以是与眉毛相关的各种基本要素,例如眉峰位置、眉头位置、眉毛长度、眉毛宽度、眉毛高度等等。

[0073] 在本实施例中,上述电子设备可以存储有预先训练的神经网络。其中,上述神经网络可以用于依次从多个图像中提取唇部特征信息以生成唇部特征信息序列和依次从多个图像中提取人脸局部特征信息以生成人脸局部特征信息序列。上述电子设备可以将所生成的标注信息和上述连续图像分别输入至上述神经网络中,得到与上述连续图像相对应的唇部特征信息序列和人脸局部特征信息序列。实践中,上述神经网络可以是使用各种现有的、可进行时序特征序列提取的网络结构或模型进行有监督训练而得到的。作为示例,上述网络结构或模型可以是深度神经网络(Deep Neural Network,DNN)与隐马尔可夫模型(Hidden Markov Model,HMM)相结合的网络结构、卷积神经网络与长短期记忆网络(Long Short-Term Memory,LSTM)相结合的网络结构、循环神经网络(Recurrent neural Network,RNN)与CTC(Connectionist Temporal Classification)模型相结合的网络结构、循环神经网络与注意力模型(Attention Model,AM)相结合的网络结构,也可以是卷积神经网络、长短期记忆网络与深度神经网络三者相结合的网络结构。

[0074] 需要说明的是,上述预先训练的神经网络的数量可以是一个或多个。作为示例,上述预先训练的神经网络的数量可以是一个,该神经网络可以同时提取唇部特征信息和人脸局部特征信息,并生成唇部特征信息序列和人脸局部特征信息序列。作为又一示例,上述预先训练的神经网络的数量可以是两个,分别为预先训练的第一神经网络和第二神经网络。其中,上述第一神经网络可以用于依次从多个图像中提取唇部特征信息以生成唇部特征信息序列,上述第二神经网络用于依次从多个图像中提取人脸局部特征信息以生成人脸局部特征信息序列。上述电子设备可以将所生成的标注信息和上述连续图像输入至预先训练的第一神经网络和第二神经网络中,分别得到与上述连续图像相对应的唇部特征信息序列和人脸局部特征信息序列。此处,上述第一神经网络与上述第二神经网络可以是使用相同的或不同可进行时序特征序列提取的网络结构或模型进行有监督训练得到的。作为再一示例,上述预先训练的神经网络的数量可以不小于三个,其中一个神经网络可以用于提取唇部特征信息,其余各个神经网络可以用于提取不同的人脸局部特征信息(例如眼睛特征信息、下巴特征信息等)。

[0075] 步骤405,将声纹特征信息发送至服务器,并将所生成的人脸特征信息、唇部特征信息序列和人脸局部特征信息序列发送至服务器,接收服务器返回的身份验证结果。

[0076] 在本实施例中,上述电子设备可以将所生成的声纹特征信息、人脸特征信息、唇部特征信息序列和人脸局部特征信息序列发送至服务器(例如图1所示的服务器105),并接收服务器返回的身份验证结果。

[0077] 实践中,上述服务器中可以预先存储有上述用户的目标声纹特征信息、目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列。上述电子设备可以在以下四个条件同时满足的情况下,生成用于表征上述用户通过身份认证的身份认证结果:上述人脸特征信息与上述目标人脸特征信息相匹配;上述唇部特征信息序列与上述目标唇部特征信息序列相匹配;上述人脸局部特征信息序列与上述目标人脸局部特征信息序列相匹配;上述声纹特征信息与上述目标声纹特征信息相匹配。

[0078] 在本实施例的一些可选的实现方式中,在获取显示有用户的唇部口型变化过程的视频数据之前,上述电子设备还可以执行如下步骤:

[0079] 首先,显示用于提示用户进行暗语设置的第二提示信息,并获取显示有上述用户的唇部口型变化过程的目标视频数据。可选的,上述电子设备在显示上述第二提示信息后,可以对上述用户的脸部进行拍摄以获取显示有上述用户的唇部口型变化过程的目标视频数据,还可以获取在视频录制过程中所产生的上述用户的目标声音数据。此处,上述第二提示信息还可以包括发音选项。上述电子设备在确定上述用户对上述发音选项执行了点击操作后,可以对上述目标声音数据进行解析,生成目标声纹特征信息。需要说明的是,目标声纹特征信息与步骤402中的声纹特征信息的生成方法基本相同,在此不再赘述。

[0080] 而后,对上述目标视频数据进行解析,生成构成上述目标视频数据的目标连续图像中的每一个图像所显示的人脸区域的目标标注信息和上述目标连续图像中每一个图像对应的目标人脸特征信息。此处,目标标注信息和目标人脸特征信息的生成方法与上述步骤403中的标注信息、人脸特征信息的生成方法基本相同,在此不再赘述。

[0081] 之后,基于所生成的目标标注信息和上述目标连续图像,生成目标唇部特征信息序列和目标人脸局部特征信息序列。其中,上述目标标注信息可以包括人脸区域的关键点的位置和人脸区域的位置。上述电子设备可以将所生成的目标标注信息和上述目标连续图像分别输入至预先训练的神经网络中,得到与上述目标连续图像相对应的目标唇部特征信息序列和目标人脸局部特征信息序列。此处,目标唇部特征信息序列和目标人脸局部特征信息序列的生成方法与步骤404中的唇部特征信息序列和人脸局部特征信息序列的生成方法基本相同,在此不再赘述。

[0082] 最后,将所生成的目标人脸特征信息、上述目标唇部特征信息序列和上述目标人脸局部行为特征信息序列发送至服务器。需要说明的是,若已生成目标声纹特征信息,则可以同时将上述目标声纹特征信息发送至上述服务器。

[0083] 需要指出的是,上述服务器可以利用各种匹配方式进行上述声纹特征信息、上述人脸特征信息、上述唇部特征信息序列、上述人脸局部特征信息序列与上述目标声纹特征信息、上述目标人脸特征信息、上述目标唇部特征信息序列、上述目标人脸局部特征信息序列的匹配。作为示例,上述服务器中可以分别存储有用于进行声纹识别、人脸识别、唇部行为识别和人脸局部行为识别的四个神经网络。上述服务器可以分别将上述声纹特征信息、上述人脸特征信息、上述唇部特征信息序列和上述人脸局部特征信息序列输入至相应的神经网络中,并得到上述四个神经网络分别输出的声纹识别结果、人脸识别结果、唇部行为识

别结果、人脸局部行为识别结果。需要说明的是,上述四个神经网络可以是利用机器学习方法,并基于大量的训练样本进行有监督训练而得到的。用于进行声纹识别的神经网络所使用的训练样本中可以包含上述目标声纹信息;用于进行人脸识别的神经网络所使用的训练样本中可以包含上述目标人脸特征信息;用于进行唇部行为识别的神经网络所使用的训练样本中可以包含上述目标唇部特征信息序列;用于进行人脸局部行为识别的神经网络所使用的训练样本中可以包含上述目标人脸局部特征信息序列。需要说明的是,进行声纹识别、人脸识别、唇部行为识别和人脸局部行为识别的神经网络的数量不限于四个,还可以是一个、两个或多个。

[0084] 需要说明的是,用户在暗语设置的过程中,若未选择发音选项,则在身份认证的过程中通常不进行发音选项的点击操作。

[0085] 从图4中可以看出,与图2对应的实施例相比,本实施例中的身份认证方法的流程400突出了对声纹特征信息的提取步骤和结合声纹特征信息进行身份认证的步骤。由此,本实施例描述的方案可以结合更多的生物特征进行身份识别,且声纹特征信息不易被效仿,从而进一步提高了身份认证的安全性。

[0086] 继续参考图5,示出了根据本申请的身份认证方法的一个实施例的流程500。该身份认证方法,包括以下步骤:

[0087] 步骤501,接收客户端发送的身份认证请求。

[0088] 在本实施例中,身份认证方法运行于其上的电子设备(例如图1所示的服务器105)可以通过有线连接方式或无线连接方式接收客户端(例如图1所示的终端设备101、102、103)发送的身份认证请求。其中,上述身份认证请求包含用户的人脸特征信息、唇部特征信息序列、人脸局部特征信息序列,其中,上述人脸特征信息、上述唇部特征信息序列和上述人脸局部特征信息序列为客户端基于显示有用户唇部口型变化过程的视频数据的解析而生成的。需要指出的是,上述无线连接方式可以包括但不限于3G/4G连接、WiFi连接、蓝牙连接、WiMAX连接、Zigbee连接、UWB(ultra wideband)连接、以及其他现在已知或将来开发的无线连接方式。

[0089] 步骤502,提取预置的、用户的目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列。

[0090] 在本实施例中,上述电子设备可以提取预置的、上述用户的目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列。实践中,上述电子设备中可以预先存储大量的用户的目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列。上述电子设备在接收到身份认证请求后,可以直接从所存储的信息中检索。

[0091] 步骤503,基于人脸特征信息、唇部特征信息序列、人脸局部特征信息序列与目标人脸特征信息、目标唇部特征信息序列、目标人脸局部特征信息序列的匹配结果,生成用于表征用户是否通过身份认证的身份认证结果。

[0092] 在本实施例中,上述电子设备可以基于上述人脸特征信息、上述唇部特征信息序列、上述人脸局部特征信息序列与上述目标人脸特征信息、上述目标唇部特征信息序列、上述目标人脸局部特征信息序列的匹配结果,生成用于表征用户是否通过身份认证的身份认证结果。

[0093] 此处,上述电子设备可以利用各种匹配方式进行上述人脸特征信息、上述唇部特

征信息序列、上述人脸局部特征信息序列与上述目标人脸特征信息、上述目标唇部特征信息序列、上述目标人脸局部特征信息序列的匹配。作为示例，上述电子设备中可以分别存储有用于进行人脸识别、唇部行为识别和人脸局部行为识别的三个神经网络。上述电子设备可以分别将上述人脸特征信息、上述唇部特征信息序列、上述人脸局部特征信息序列输入至相应的神经网络中，并得到上述三个神经网络分别输出的人脸识别结果、唇部行为识别结果、人脸局部行为识别结果。需要指出的是，上述人脸识别结果、唇部行为识别结果、人脸局部行为识别结果可以为三个数值。若上述人脸识别结果大于技术人员基于大量数据统计而预先设定的第一阈值，则可以确定上述人脸特征信息与上述目标人脸特征信息相匹配。若上述唇部行为识别结果大于技术人员基于大量数据统计而预先设定的第二阈值，则可以确定上述唇部特征信息序列与上述目标唇部特征信息序列相匹配。若上述人脸局部行为识别结果大于技术人员基于大量数据统计而预先设定的第三阈值，则可以确定上述人脸局部特征信息序列与上述目标人脸局部特征信息序列相匹配。

[0094] 需要说明的是，上述三个神经网络可以是利用机器学习方法，并基于大量的训练样本进行有监督训练而得到的。用于进行人脸识别的神经网络所使用的训练样本中可以包含上述目标人脸特征信息；用于进行唇部行为识别的神经网络所使用的训练样本中可以包含上述目标唇部特征信息序列；用于进行人脸局部行为识别的神经网络所使用的训练样本中可以包含上述目标人脸局部特征信息序列。需要说明的是，进行人脸识别、唇部行为识别和人脸局部行为识别的神经网络的数量不限于三个，还可以是一个、两个或多个。

[0095] 在本实施例的一些可选的实现方式中，上述电子设备可以在以下三个条件同时满足的情况下，生成用于表征上述用户通过身份认证的身份认证结果：上述人脸特征信息与上述目标人脸特征信息相匹配；上述唇部特征信息序列与上述目标唇部特征信息序列相匹配；上述人脸局部特征信息序列与上述目标人脸局部特征信息序列相匹配。

[0096] 在本实施例的一些可选的实现方式中，上述身份认证请求还可以包括上述用户的声纹特征信息。上述电子设备还可以提取预置的、上述用户的目标声纹特征信息；并在以下四个条件同时满足的情况下，生成用于表征上述用户通过身份认证的身份认证结果：上述人脸特征信息与上述目标人脸特征信息相匹配；上述唇部特征信息序列与上述目标唇部特征信息序列相匹配；上述人脸局部特征信息序列与上述目标人脸局部特征信息序列相匹配；上述声纹特征信息与上述目标声纹特征信息相匹配。

[0097] 在本实施例的一些可选的实现方式中，在上述接收客户端发送的身份认证请求之前，上述电子设备还可以接收客户端发送的身份设定请求，其中，上述身份设定请求可以包含上述用户的目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列，上述目标人脸特征信息、上述目标唇部特征信息序列和上述目标人脸局部特征信息序列为上述客户端基于显示有上述用户唇部口型变化过程的目标视频数据的解析而生成的。而后，上述电子设备可以将上述目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列进行存储。

[0098] 在本实施例的一些可选的实现方式中，上述身份设定请求还包括上述用户的目标声纹特征信息，其中，上述目标声纹特征信息是上述客户端基于对上述用户的语音数据的解析而生成的。上述电子设备可以将上述目标声纹特征信息、上述目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列进行存储。

[0099] 步骤504,向客户端发送身份认证结果。

[0100] 在本实施例中,上述电子设备可以向上述客户端发送上述身份认证结果。

[0101] 本申请的上述实施例提供的方法,通过将客户端发送的身份认证请求中的人脸特征信息、唇部特征信息序列、人脸局部特征信息序列与预置的上述用户的目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列进行分别匹配,而后基于匹配结果生成用于表征上述用户是否通过身份认证的身份认证结果,最后向上述客户端发送上述身份认证结果,从而可以结合人脸的面部识别、用户使用唇语时的嘴型动作以及用户使用唇语时的脸部局部行为的识别进行用户身份认证,由于用户使用唇语时的嘴型动作以及脸部行为是用户自身独有的生物特征,攻击者不易效仿,因此,提高了身份认证的安全性。

[0102] 进一步参考图6,作为对上述图2所示方法的实现,本申请提供了一种身份认证装置的一个实施例,该装置实施例与图2所示的方法实施例相对应,该装置具体可以应用于各种电子设备中。

[0103] 如图6所示,本实施例上述的身份认证装置600包括:获取单元601,配置用于获取显示有用户的唇部口型变化过程的视频数据;解析单元602,配置用于对上述视频数据进行解析,生成构成上述视频数据的连续图像中每一个图像所显示的人脸区域的标注信息和上述连续图像中每一个图像对应的人脸特征信息;生成单元603,配置用于基于所生成的标注信息和上述连续图像,生成唇部特征信息序列和人脸局部特征信息序列;发送单元604,配置用于将所生成的人脸特征信息、上述唇部特征信息序列和上述人脸局部特征信息序列发送至服务器,并接收上述服务器返回的身份验证结果。

[0104] 本申请的上述实施例提供的装置,通过解析单元602对获取单元601所获取的显示有用户唇部口型变化过程的视频数据进行解析,以便生成标注信息和人脸特征信息,而后生成单元603基于所生成的标注信息和连续图像,生成唇部特征信息序列和人脸局部特征信息序列,最后发送单元604将所生成的人脸特征信息、唇部特征信息序列和人脸局部特征信息序列发送至服务器,并接收服务器返回的身份验证结果,从而可以结合人脸的面部识别、用户使用唇语时的嘴型动作以及用户使用唇语时的脸部局部行为的识别进行用户身份认证,由于用户使用唇语时的嘴型动作以及脸部行为是用户自身独有的生物特征,攻击者不易效仿,因此,提高了身份认证的安全性。

[0105] 进一步参考图7,作为对上述图5所示方法的实现,本申请提供了一种身份认证装置的一个实施例,该装置实施例与图5所示的方法实施例相对应,该装置具体可以应用于各种电子设备中。

[0106] 如图7所示,本实施例上述的身份认证装置700包括:接收单元701,配置用于接收客户端发送的身份认证请求,其中,上述身份认证请求包含用户的人脸特征信息、唇部特征信息序列、人脸局部特征信息序列,上述人脸特征信息、上述唇部特征信息序列和上述人脸局部特征信息序列为上述客户端基于显示有上述用户唇部口型变化过程的视频数据的解析而生成的;提取单元702,配置用于提取预置的、上述用户的目标人脸特征信息、目标唇部特征信息序列和目标人脸局部特征信息序列;匹配单元703,配置用于基于上述人脸特征信息、上述唇部特征信息序列、上述人脸局部特征信息序列与上述目标人脸特征信息、上述目标唇部特征信息序列、上述目标人脸局部特征信息序列的匹配结果,生成用于表征上述用户是否通过身份认证的身份认证结果;发送单元704,配置用于向上述客户端发送上述身份

认证结果。

[0107] 本申请的上述实施例提供的装置,通过匹配单元703将接收单元701所接收的人脸特征信息、唇部特征信息序列、人脸局部特征信息序列与提取单元702所提取的目标人脸特征信息、目标唇部特征信息序列、目标人脸局部特征信息序列进行分别匹配,以便基于匹配结果生成用于表征上述用户是否通过身份认证的身份认证结果,而后发送单元704向上述客户端发送上述身份认证结果,从而可以结合人脸的面部识别、用户使用唇语时的嘴型动作以及用户使用唇语时的脸部局部行为的识别进行用户身份认证,由于用户使用唇语时的嘴型以及脸部行为是用户自身独有的生物特征,攻击者不易效仿,因此,提高了身份认证的安全性。

[0108] 下面参考图8,其示出了适于用来实现本申请实施例的终端设备/服务器的计算机系统800的结构示意图。图8示出的终端设备/服务器仅仅是一个示例,不应对本申请实施例的功能和使用范围带来任何限制。

[0109] 如图8所示,计算机系统800包括中央处理单元(CPU)801,其可以根据存储在只读存储器(ROM)802中的程序或者从存储部分808加载到随机访问存储器(RAM)803中的程序而执行各种适当的动作和处理。在RAM 803中,还存储有系统800操作所需的各种程序和数据。CPU 801、ROM 802以及RAM 803通过总线804彼此相连。输入/输出(I/O)接口805也连接至总线804。

[0110] 以下部件连接至I/O接口805:包括摄像头等的输入部分806;包括诸如液晶显示器(LCD)等以及扬声器等的输出部分807;包括硬盘等的存储部分808;以及包括诸如LAN卡、调制解调器等的网络接口卡的通信部分809。通信部分809经由诸如因特网的网络执行通信处理。驱动器810也根据需要连接至I/O接口805。可拆卸介质811,诸如磁盘、光盘、磁光盘、半导体存储器等等,根据需要安装在驱动器810上,以便于从其上读出的计算机程序根据需要被安装入存储部分808。

[0111] 特别地,根据本公开的实施例,上文参考流程图描述的过程可以被实现为计算机软件程序。例如,本公开的实施例包括一种计算机程序产品,其包括承载在计算机可读介质上的计算机程序,该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中,该计算机程序可以通过通信部分809从网络上被下载和安装,和/或从可拆卸介质811被安装。在该计算机程序被中央处理单元(CPU)801执行时,执行本申请的方法中限定的上述功能。需要说明的是,本申请所述的计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质或者是上述两者的任意组合。计算机可读存储介质例如可以是——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件,或者任意以上的组合。计算机可读存储介质的更具体的例子可以包括但不限于:具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、光纤、便携式紧凑磁盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本申请中,计算机可读存储介质可以是任何包含或存储程序的有形介质,该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。而在本申请中,计算机可读的信号介质可以包括在基带中或者作为载波一部分传播的数据信号,其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式,包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读的信号介质还可以是计算机

可读存储介质以外的任何计算机可读介质,该计算机可读介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输,包括但不限于:无线、电线、光缆、RF等等,或者上述的任意合适的组合。

[0112] 附图中的流程图和框图,图示了按照本申请各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分,该模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意,在有些作为替换的实现中,方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如,两个接连地表示的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这依所涉及的功能而定。也要注意,框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合,可以用执行规定的功能或操作的专用的基于硬件的系统来实现,或者可以用专用硬件与计算机指令的组合来实现。

[0113] 描述于本申请实施例中所涉及到的单元可以通过软件的方式实现,也可以通过硬件的方式来实现。所描述的单元也可以设置在处理器中,例如,可以描述为:一种处理器包括获取单元、解析单元、生成单元和发送单元。其中,这些单元的名称在某种情况下并不构成对该单元本身的限定,例如,获取单元还可以被描述为“获取显示有用户的唇部口型变化过程的视频数据的单元”。

[0114] 作为另一方面,本申请还提供了一种计算机可读介质,该计算机可读介质可以是上述实施例中描述的装置中所包含的;也可以是单独存在,而未装配入该装置中。上述计算机可读介质承载有一个或者多个程序,当上述一个或者多个程序被该装置执行时,使得该装置:获取显示有用户的唇部口型变化过程的视频数据;对该视频数据进行解析,生成构成该视频数据的连续图像中每一个图像所显示的人脸区域的标注信息和该连续图像中每一个图像对应的人脸特征信息;基于所生成的标注信息和该连续图像,生成唇部特征信息序列和人脸局部特征信息序列;将所生成的人脸特征信息、该唇部特征信息序列和该人脸局部特征信息序列发送至服务器,并接收该服务器返回的身份验证结果。

[0115] 以上描述仅为本申请的较佳实施例以及对所运用技术原理的说明。本领域技术人员应当理解,本申请中所涉及的发明范围,并不限于上述技术特征的特定组合而成的技术方案,同时也应涵盖在不脱离上述发明构思的情况下,由上述技术特征或其等同特征进行任意组合而形成的其它技术方案。例如上述特征与本申请中公开的(但不限于)具有类似功能的技术特征进行互相替换而形成的技术方案。

100

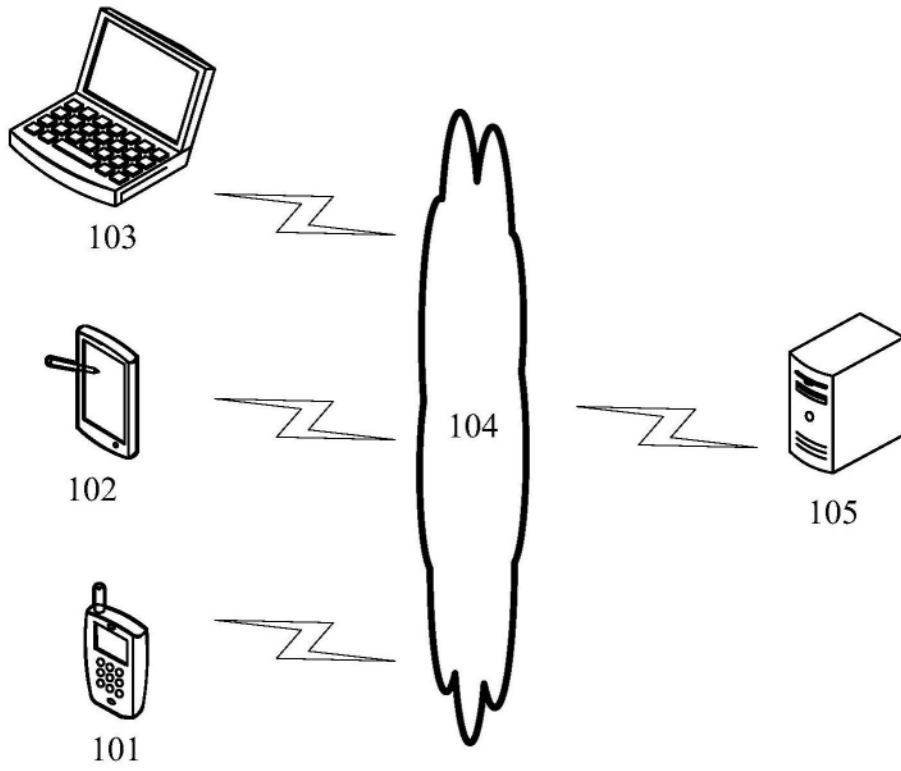


图1

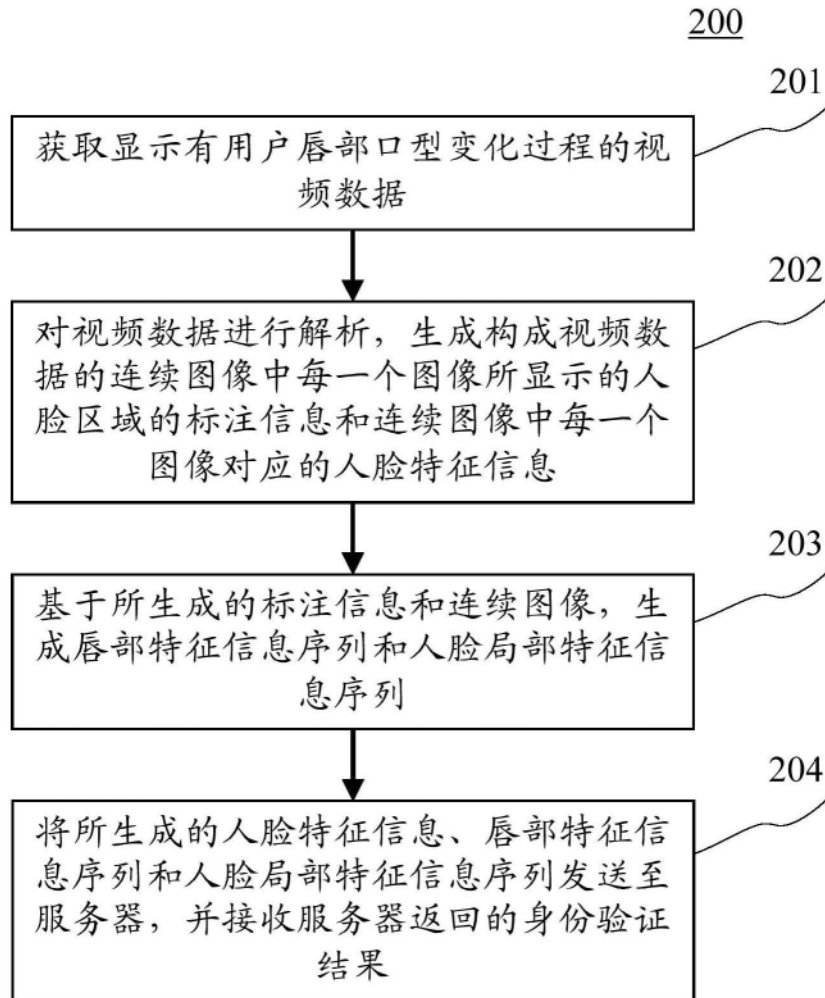


图2

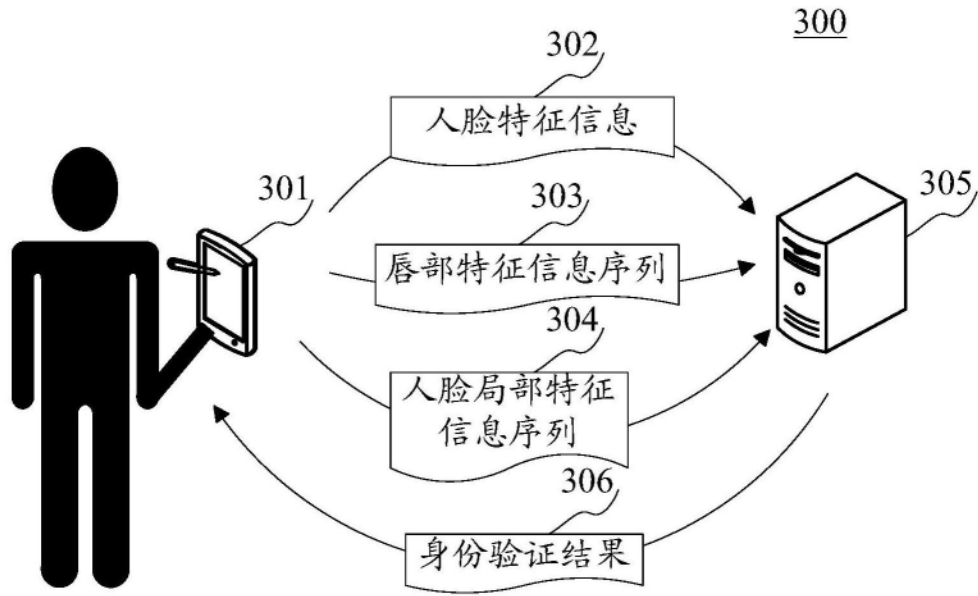


图3

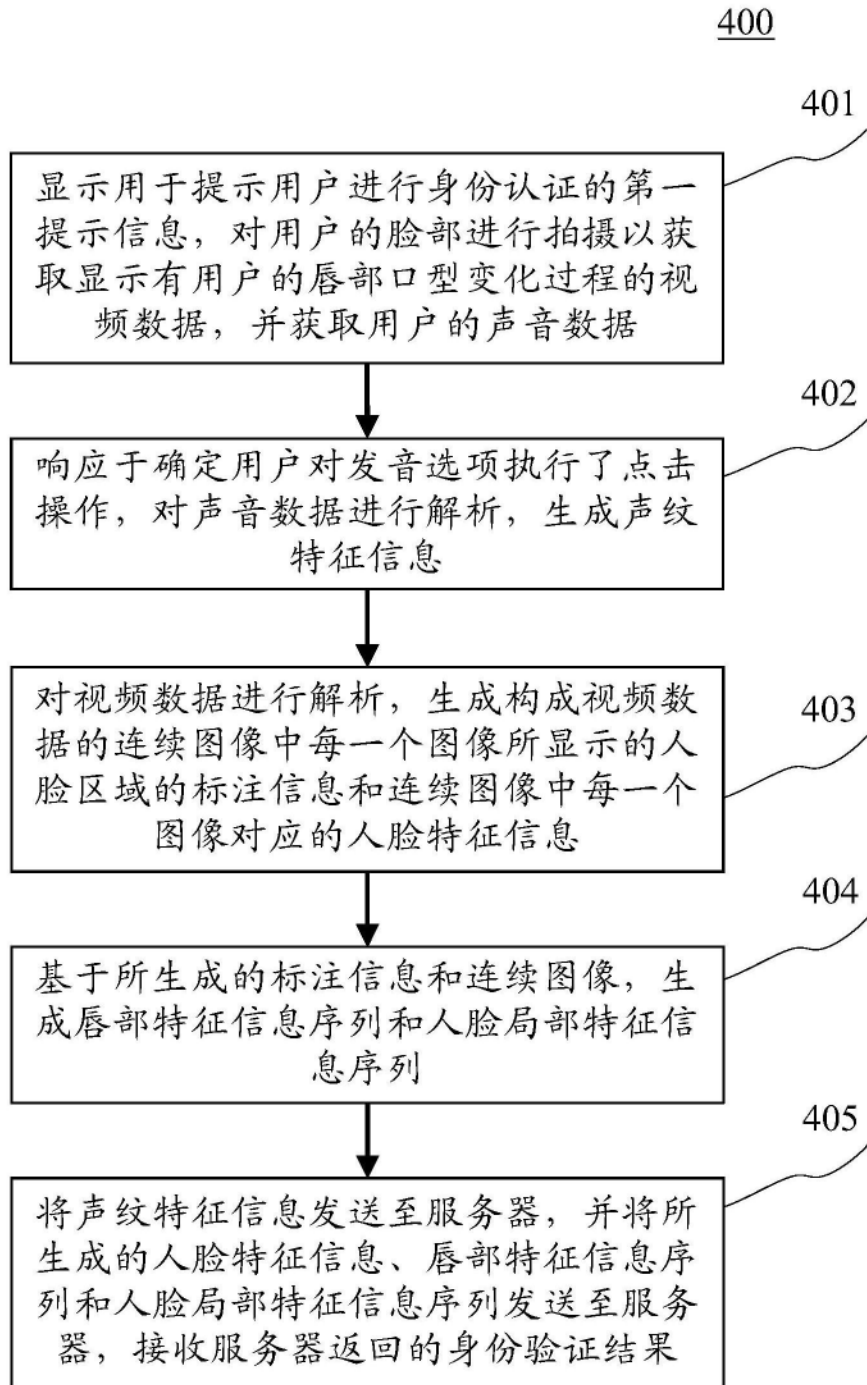


图4

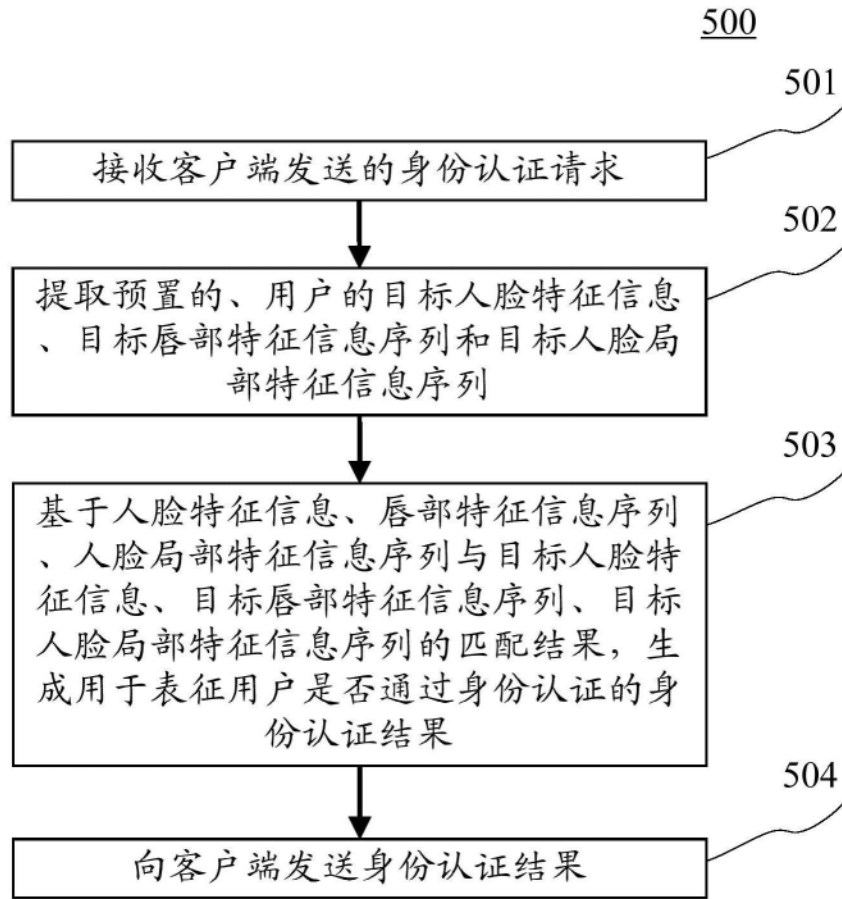


图5

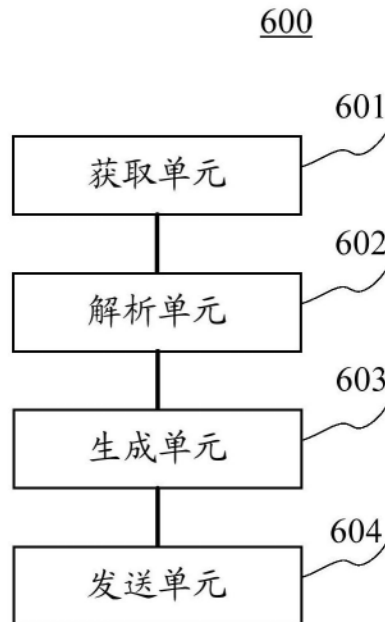


图6

700

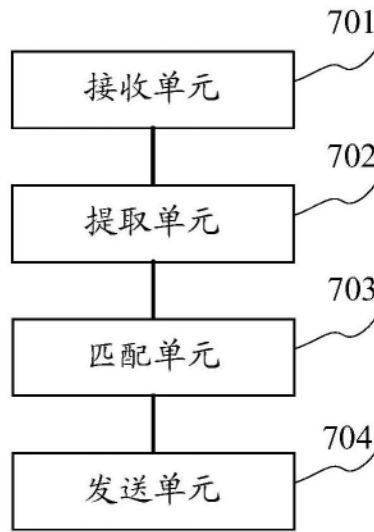


图7

800

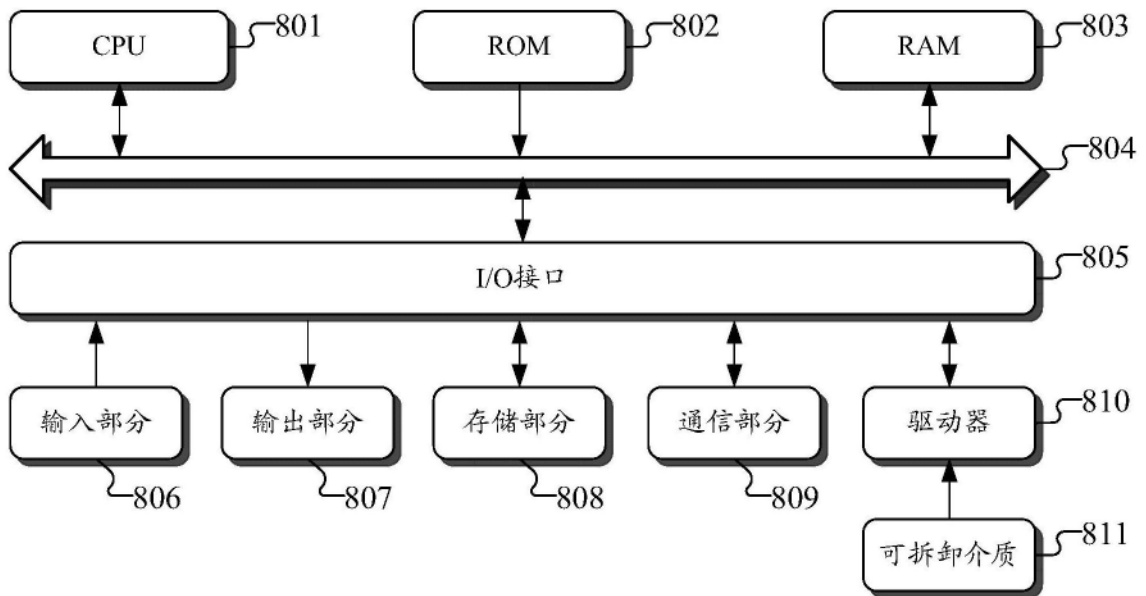


图8