



(12) 发明专利申请

(10) 申请公布号 CN 113302694 A

(43) 申请公布日 2021.08.24

(21) 申请号 202080009459.1

(22) 申请日 2020.01.18

(30) 优先权数据

16/251,436 2019.01.18 US

16/251,472 2019.01.18 US

16/434,185 2019.06.07 US

16/551,756 2019.08.27 US

16/594,690 2019.10.07 US

16/594,771 2019.10.07 US

16/661,086 2019.10.23 US

16/661,122 2019.10.23 US

(85) PCT国际申请进入国家阶段日

2021.07.15

(86) PCT国际申请的申请数据

PCT/US2020/014225 2020.01.18

(87) PCT国际申请的公布数据

W02020/150692 EN 2020.07.23

(71) 申请人 斯纳普公司

地址 美国加利福尼亚州

(72) 发明人 维克托·沙布罗夫

帕维尔·萨夫琴科夫

亚历山大·马什拉博夫

德米特里·马托夫

索菲娅·萨维诺娃

阿列克谢·普切尼科夫

罗曼·戈洛布科维

(74) 专利代理机构 北京康信知识产权代理有限公司 11240

代理人 刘彬

(51) Int.Cl.

G11B 27/031 (2006.01)

G06T 11/60 (2006.01)

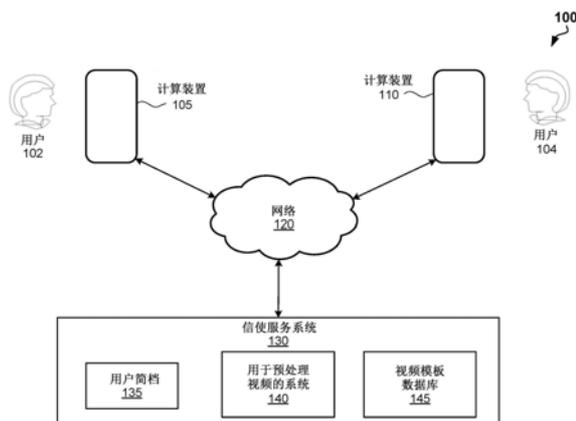
权利要求书3页 说明书12页 附图16页

(54) 发明名称

用于基于模板生成个性化视频的系统和方法

(57) 摘要

公开了用于基于模板生成个性化视频的系统和方法。示例方法可从接收视频配置数据开始,该视频配置数据包括:帧图像的序列、限定面部区域在帧图像中的位置的面部区域参数的序列、以及限定面部界标在帧图像中的位置的面部界标参数的序列。方法可继续至接收源面部的图像。方法还可包括生成输出视频。输出视频的生成可包括修改帧图像的序列的帧图像。具体地,可以修改源面部的图像以获得另外的图像,该另外的图像表征采用了与面部界标参数相对应的面部表情的源面部。另外的图像可在通过与帧图像相对应的面部区域参数确定的位置处插入到帧图像中。



1. 一种用于基于模板生成个性化视频的方法,所述方法包括:
  - 由计算装置接收视频配置数据,所述视频配置数据包括:
    - 帧图像的序列;
    - 面部区域参数的序列,所述面部区域参数限定面部区域在所述帧图像中的位置;以及
    - 面部界标参数的序列,所述面部界标参数限定面部界标在所述帧图像中的位置,其中,每个所述面部界标参数与面部表情相对应;
  - 由计算机装置接收源面部的图像;并且
  - 由所述计算装置生成输出视频,其中,生成所述输出视频包括通过以下步骤修改所述帧图像的所述序列中的一帧图像:
    - 基于与所述帧图像相对应的所述面部界标参数,修改所述源面部的所述图像以获得另外的图像,所述另外的图像表征采用了与所述面部界标参数相对应的面部表情的所述源面部;并且
    - 在由与所述帧图像相对应的所述面部区域参数确定的位置处,将所述另外的图像插入到所述帧图像中。
2. 根据权利要求1所述的方法,其中,所述帧图像的所述序列基于以下项中的一项生成:动画视频和实景真人视频。
3. 根据权利要求1所述的方法,其中,所述面部界标参数的所述序列基于实景真人视频来生成,所述实景真人视频表征面部同步演员的面部。
4. 根据权利要求1所述的方法,其中:
  - 所述视频配置数据包括皮肤掩模的序列,所述皮肤掩模限定所述帧图像中所表征的至少一个演员的身体的皮肤区域或另外的身体的2D动画/3D动画的皮肤区域;并且
  - 生成所述输出视频包括:
    - 确定与所述源面部相关联的颜色数据,并且
    - 基于所述颜色数据来对所述帧图像中的所述皮肤区域重新着色。
5. 根据权利要求1所述的方法,其中:
  - 所述视频配置数据还包括嘴部区域图像的序列,每个所述嘴部区域图像与至少一个所述帧图像相对应;并且
  - 生成所述输出视频包括将与所述帧图像相对应的嘴部区域插入到所述帧图像中。
6. 根据权利要求1所述的方法,其中:
  - 所述视频配置数据还包括眼部参数的序列,所述眼部参数限定所述帧图像中所表征的面部同步演员的巩膜中的虹膜的位置;并且
  - 生成所述输出视频包括:
    - 基于与帧相对应眼部参数来生成眼部区域的图像;并且
    - 将所述眼部区域的所述图像插入所述帧图像中。
7. 根据权利要求1所述的方法,其中:
  - 所述视频配置数据包括头部参数的序列,所述头部参数限定头部的旋转、转向、位置、以及比例中的一个或多个。
8. 根据权利要求1所述的方法,其中:
  - 生成所述输出视频包括:

基于所述源面部的所述图像确定头发掩模；

基于所述头发掩模生成头发图像；并且

将所述头发图像插入到所述帧图像中。

9. 根据权利要求1所述的方法，其中：

所述视频配置数据还包括动画对象图像的序列，其中，每个所述动画对象图像与至少一个所述帧图像相对应；并且

生成所述输出视频包括：将与所述帧图像相对应的动画对象图像插入到所述帧图像中。

10. 根据权利要求1所述的方法，其中：

所述视频配置数据包括配乐；并且

生成所述输出视频还包括将所述配乐添加到所述输出视频中。

11. 一种用于基于模板生成个性化视频的系统，所述系统包括至少一个处理器和存储处理器可执行代码的存储器，其中，所述至少一个处理器被配置为在执行所述处理器可执行代码时实现以下操作：

由计算装置接收视频配置数据，所述视频配置数据包括：

帧图像的序列；

面部区域参数的序列，所述面部区域参数限定面部区域在所述帧图像中的位置；以及

面部界标参数的序列，所述面部界标参数限定面部界标在所述帧图像中的位置，其中，每个面部界标参数与面部表情相对应；

由计算机装置接收源面部的图像；并且

由所述计算装置生成输出视频，其中，生成所述输出视频包括通过以下步骤修改所述帧图像的所述序列中的一帧图像：

基于与所述帧图像相对应的所述面部界标参数，修改所述源面部的所述图像以获得另外的图像，所述另外的图像表征采用了与所述面部界标参数相对应的面部表情的所述源面部；并且

在通过与所述帧图像相对应的所述面部区域参数确定的位置处，将所述另外的图像插入到所述帧图像中。

12. 根据权利要求11所述的系统，其中，所述帧图像的所述序列基于以下项中的一项生成：动画视频和实景真人视频。

13. 根据权利要求11所述的系统，其中，所述面部界标参数的所述序列基于实景真人视频来生成，所述实景真人视频表征面部同步演员的面部。

14. 根据权利要求11所述的系统，其中：

所述视频配置数据包括皮肤掩模的序列，所述皮肤掩模限定所述帧图像中所表征的至少一个演员的身体的皮肤区域或另外的身体的2D动画/3D动画的皮肤区域；并且

生成所述输出视频包括：

确定与所述源面部相关联的颜色数据；并且

基于所述颜色数据来对所述帧图像中的所述皮肤区域重新着色。

15. 根据权利要求11所述的系统，其中：

所述视频配置数据还包括嘴部区域图像的序列，每个所述嘴部区域图像与至少一个所

述帧图像相对应;并且

生成所述输出视频包括将与所述帧图像相对应的嘴部区域插入到所述帧图像中。

16. 根据权利要求11所述的系统,其中:

所述视频配置数据还包括眼部参数的序列,所述眼部参数限定所述帧图像中所表征的面部同步演员的巩膜中的虹膜的位置;并且

生成所述输出视频包括:

基于与帧相对应的眼部参数来生成眼部区域的图像;并且

将所述眼部区域的所述图像插入所述帧图像中。

17. 根据权利要求11所述的系统,其中:

所述视频配置数据包括头部参数的序列,所述头部参数限定头部的旋转、转向、位置、以及比例中的一个或多个。

18. 根据权利要求11所述的系统,其中:

生成所述输出视频包括:

基于源面部图像确定头发掩模;

基于所述头发掩模生成头发图像;并且

将所述头发图像插入到所述帧图像中。

19. 根据权利要求11所述的系统,其中:

所述视频配置数据还包括动画对象图像的序列,其中,每个所述动画对象图像与至少一个所述帧图像相对应;并且

生成所述输出视频包括:将与所述帧图像相对应的动画对象图像插入到所述帧图像中。

20. 一种非暂时性处理器可读介质,其上存储有指令,所述指令在被一个或多个处理器执行时,使所述一个或多个处理器实现一种用于基于模板生成个性化视频的方法,所述方法包括:

由计算装置接收视频配置数据,所述视频配置数据包括:

帧图像的序列;

面部区域参数的序列,所述面部区域参数限定面部区域在帧图像中的位置;以及

面部界标参数的序列,所述面部界标参数限定面部界标在所述帧图像中的位置,其中,每个面部界标参数与面部表情相对应;

由计算机装置接收源面部的图像;并且

由所述计算装置生成输出视频,其中,生成所述输出视频包括通过以下步骤修改所述帧图像的所述序列中的一帧图像:

基于与所述帧图像相对应的所述面部界标参数,修改所述源面部的所述图像以获得另外的图像,所述另外的图像表征采用了与所述面部界标参数相对应的面部表情的所述源面部;并且

在通过与所述帧图像相对应的所述面部区域参数确定的位置处,将所述另外的图像插入到所述帧图像中。

## 用于基于模板生成个性化视频的系统和方法

### 技术领域

[0001] 本公开总体涉及一种数字图像处理。更具体地,本公开涉及用于基于模板生成个性化视频的方法和系统。

### 背景技术

[0002] 共享诸如贴纸和表情符号的媒体已在消息传递应用中成为标准选项(在本文中也称为信使(messenger))。目前,一些信使向用户提供用于生成图像和短视频并经由通信聊天将该图像和短视频发送给其他用户的选项。某些现有的信使允许用户在传输之前修改短视频。然而,由现有信使提供的短视频的修改限于可视化效果、滤镜和文本。当前信使的用户不能执行复杂的编辑(例如,将一个面部替换成另一个面部)。这样的视频编辑无法由当前信使提供,并且需要复杂的第三方视频编辑软件。

### 发明内容

[0003] 此部分的目的是以简化的形式介绍选择的理念,该理念的具体内容如下文中的具体实施方式部分所述。本发明内容并非用于确定要求保护的的主题的关键特征或主要特征,也并非用于帮助确定要求保护的的主题的范围。

[0004] 根据本公开的一个实施例,公开了一种用于基于模板生成个性化视频的系统。系统可包括至少一个处理器和存储处理器可执行代码的存储器。至少一个处理器可被配置为由计算装置接收视频配置数据。视频配置数据可包括:帧图像的序列、限定面部区域在帧图像中的位置的面部区域参数的序列、以及限定面部界标在帧图像中的位置的面部界标参数的序列。每个面部界标参数可与面部表情相对应。至少一个处理器可被配置为由计算机装置接收源面部的图像。至少一个处理器可配置为通过计算装置生成输出视频。输出视频的生成可包括修改帧图像的序列的帧图像。具体地,可基于与帧图像相对应的面部界标参数来修改源面部的图像以获得另外的图像,该另外的图像表征采用了与面部界标参数相对应的面部表情的源面部。另外的图像可在通过与帧图像相对应的面部区域参数确定的位置处插入到帧图像中。

[0005] 根据本公开的一个示例性实施例,公开了一种用于基于模板生成个性化视频的方法。方法可从通过计算装置接收视频配置数据开始。视频配置数据可包括:帧图像的序列、限定面部区域在帧图像中的位置的面部区域参数的序列、以及限定面部界标在帧图像中的位置的面部界标参数的序列。每个面部界标参数可对应于面部表情。方法可继续至由计算机装置接收源面部的图像。方法还可包括由计算装置生成输出视频。输出视频的生成可包括修改帧图像的序列的帧图像。具体地,可以修改源面部的图像以获得另外的图像,该另外的图像表征采用了与面部界标参数相对应的面部表情的源面部。可基于与帧图像相对应的面部界标参数来执行图像的修改。另外的图像可在通过与帧图像先谷底应的面部区域参数确定的位置处插入到帧图像中。

[0006] 根据本公开的又一方面,提供一种非暂时性处理器可读介质,其存储处理器可读

指令。当处理器可读指令由处理器执行时,它们使处理器实现上述用于基于模板生成个性化视频的方法。

[0007] 示例的另外目的、优点和新颖特征将部分地在以下描述中阐述,并且部分地在检查以下描述和附图后对于本领域技术人员将变得显而易见,或者可通过示例的生产或操作来获悉。理念的目的和优点可借助于在所附权利要求中特别指出的方法,手段和组合来实现和获得。

## 附图说明

[0008] 在附图中以示例而非限制的方式示出实施例,在该附图中,类似的附图标记表示相似的元件。

[0009] 图1是示出示例环境的框图,其中,可以实现用于基于模板生成个性化视频的系统和方法。

[0010] 图2是示出用于实现用于基于模板生成个性化视频的方法的计算装置的示例性实施例的框图。

[0011] 图3是示出根据本公开的一些示例性实施例的用于基于模板生成个性化视频的处理的流程图。

[0012] 图4是示出根据本公开的一些示例性实施例的用于基于模板生成个性化视频的系统的功能的流程图。

[0013] 图5是示出根据一些示例性实施例的生成用于生成视频模板的实景真人视频的处理的流程图。

[0014] 图6示出根据一些示例性实施例的用于生成视频模板的示例实景真人视频的帧。

[0015] 图7示出根据一个示例性实施例的面部的原始图像和具有归一化的光照的面部的图像。

[0016] 图8示出根据一个示例性实施例的分割的头部图像、具有面部界标的头部图像和面部掩模。

[0017] 图9示出根据一个示例性实施例的表征用户面部、皮肤掩模、以及对皮肤掩模重新着色的结果的帧。

[0018] 图10示出根据一个示例性实施例的面部同步演员的面部图像的图像、面部同步演员的面部界标的图像、用户的面部界标的图像、以及具有面部同步演员的面部表情的用户的面部的图像。

[0019] 图11示出根据一个示例性实施例的分割的面部图像、头发掩模、翘曲到目标图像的头发掩模、以及应用到目标图像的头发掩模。

[0020] 图12示出根据一个示例性实施例的眼部的原始图像、具有重建的眼部巩膜的图像、具有重建的虹膜的图像、以及具有移动的重建的虹膜的图像。

[0021] 图13和图14示出根据一些示例性实施例的基于视频模板生成的示例个性化视频的帧。

[0022] 图15是示出根据本公开的一个示例性实施例的用于基于模板生成个性化视频的方法的流程图。

[0023] 图16示出可用于实现用于基于模板生成个性化视频的方法的示例计算机系统。

## 具体实施方式

[0024] 以下对实施例的具体实施方式包括参照形成具体实施方式的一部分附图。此部分中描述的办法不是权利要求的现有技术,并且也不通过包括在此部分中而被承认为现有技术。附图示出根据示例性实施例的说明。足够详细地描述了这些在本文中也被称为“示例”的示例性实施例,以使本领域技术人员能够实践本主题。在不脱离所要求保护的的范围的情况下,可组合实施例,可利用其它实施例,或者可进行结构的,逻辑的和操作的改变。因此,以下具体实施方式不应被认为是限制性的,并且范围由所附权利要求及其等同物限定。

[0025] 出于本专利文件的目的,除非另外说明或在其使用的上下文中另外明确地意指,否则术语“或”和“和”应指“和/或”。除非另外说明或在“一个或多个”的使用明显不适当的情况下,否则术语“一个”应指“一个或多个”。术语“包括(comprise)”、“包括(comprising)”、“包含(include)”和“包含(including)”是可互换的并且不旨在限制。例如,术语“包含”应解释为指“包括但不限于”。

[0026] 本公开涉及用于基于模板生成个性化视频的方法和系统。本公开提供的实施例解决了现有技术的至少一些问题。本公开可设计成在诸如智能电话、平板电脑或手机的移动装置上实时工作,但实施例可扩展到涉及网络服务或基于云的资源。可通过在计算机系统上运行的软件和/或通过利用微处理器的组合或其它专门设计的专用集成电路(ASIC)、可编程逻辑装置或其任何组合的硬件来实现在本文中描述的方法。具体地,在本文中描述的方法可通过驻留在非暂时性存储介质(例如磁盘驱动器或计算机可读介质)上的一系列计算机可执行指令来实现。

[0027] 本公开的一些实施例可允许在诸如智能电话的用户计算装置上实时生成个性化视频。个性化视频可以以视听媒体(例如,视频、动画或任何其它类型的媒体)的形式生成,该视听媒体表征一个用户的面部或多个用户的面部。可以基于预生成的视频模板生成个性化视频。视频模板可包括视频配置数据。视频配置数据可以包括帧图像的序列、限定面部区域在帧图像中的位置的面部区域参数的序列、以及限定面部界标在帧图像中的位置的面部界标参数的序列。每个面部界标参数可与面部表情相对应。可基于动画视频或实景真人视频生成帧图像。可基于表征演员的面部的另一实景真人视频(也被称为如下面更详细描述的面部同步(facesync))、动画视频、音频文件、文本或手动来生成面部界标参数。

[0028] 视频配置文件还可包括皮肤掩模的序列。皮肤掩模可限定帧图像中所表征的演员的身体的皮肤区域或身体的2D动画/3D动画的皮肤区域。在一个示例性实施例中,可基于捕获不同演员(在本文中分别被称为演员和面部同步演员)的两个不同的实景真人视频来生成皮肤掩模和面部界标参数。视频配置数据还可包括嘴部区域图像的序列和眼部参数的序列。眼部参数可限定在帧图像中所表征的面部同步演员的巩膜中的虹膜的位置。视频配置数据可包括头部参数、以及头部的其他参数的序列,该头部参数限定头部的旋转和转向、位置、比例。当拍摄图像并直接看相机时,用户可保持其头部静止,因此,可手动调节头部的比例和旋转。头部参数可从不同的演员(在本文中也被称为面部同步演员)传递。如在本文中使用的,面部同步演员是其面部界标参数正被使用的人,并且演员是其身体正被用在视频模板中并且其皮肤可被重新着色的另一人,并且用户是拍摄他/她的面部的图像以生成个性化视频的人。因此,在一些实施例中,个性化视频包括被修改为具有面部同步演员的面部表情的用户面部,并且包括取自视频模板并重新着色以匹配用户面部颜色的演员的身体。

视频配置数据包括动画对象图像的序列。可选地，视频配置数据包括配乐和/或语音。

[0029] 预生成的视频模板可远程存储在基于云的计算资源中，并且可由计算装置（诸如智能电话）的用户下载。计算装置的用户可通过计算装置捕获面部的图像或从相机卷、从准备好的图像集合、或经由网络链接来选择面部的图像。在一些实施例中，图像可包括动物而不是人的面部，者可以是以画画的形式。基于面部的图像和预生成的视频模板中的一个，计算装置还可以生成个性化视频。用户可经由通信聊天将个性化视频发送给另一计算装置的另一用户、在社交媒体上共享、下载到计算装置的本地存储装置、或上传到云存储装置或视频共享服务。

[0030] 根据本公开的一个实施例，一种用于基于模板生成个性化视频的示例方法可包括通过计算装置接收视频配置数据。视频配置数据可包括帧图像的序列、限定面部区域在帧图像中的位置的面部区域参数的序列、以及限定面部界标在帧图像中的位置的面部界标参数的序列。每个面部界标参数可与面部同步演员的面部表情相对应。方法可继续通过计算装置接收源面部的图像并生成输出视频。输出视频的生成可包括修改帧图像的序列的帧图像。帧图像的修改可包括修改源面部的图像以获得另外的图像，该另外的图像表征采用了与面部界标参数相对应的面部表情的源面部，并且在通过与帧图像相对应的面部区域参数确定的位置处将该另外的图像插入到帧图像中。另外，例如可通过改变颜色，使眼部更大等来修改源面部。可基于与帧图像相对应的面部界标参数来修改源面部的图像。

[0031] 现在参照附图，描述示例性实施例。附图是理想化的示例性实施例的示意图。因此，不应将在本文中论述的示例性实施例理解为限于在本文中呈现的特定说明；准确地说，如对本领域技术人员将是明显的，这些示例性实施例可包括偏离并且不同于在本文中呈现的说明。

[0032] 图1示出示例环境100，其中，可以实现用于基于模板生成个性化视频的系统和方法。环境100可包括计算装置105、用户102、计算装置110、用户104、网络120和信使服务系统130。计算装置105和计算装置110可指诸如移动电话、智能电话或平板电脑的移动装置。在其它实施例中，计算装置110可指个人电脑、膝上型电脑、上网本、机顶盒、电视装置、多媒体装置、个人数字助理、游戏机、娱乐系统、信息娱乐系统、车载计算机或任何其它计算装置。

[0033] 计算装置105和计算装置110可以经由网络120通信地连接到信使服务系统130。信使服务系统130可实现为基于云的计算资源。信使服务系统130可包括在远程位置处可用并可通过网络（例如，互连网）访问的计算资源（硬件和软件）。基于云的计算资源可由多个用户共享，并且可基于需求动态地重新分配。基于云的计算资源可包括一个或多个服务器群/集群，该服务器群/集群包括可与网络交换机和/或路由器共处一地的计算机服务器的集合。

[0034] 网络120可包括任何有线网络、无线网络或光学网络（例如包括互连网、内联网、局域网（LAN）、个域网（PAN）、广域网（WAN）、虚拟专用网（VPN）、蜂窝电话网络（例如，全球移动通信系统（GSM））等）。

[0035] 在本公开的一些实施例中，计算装置105可被配置为启动用户102与计算装置110的用户104之间的通信聊天。在通信聊天期间，用户102与用户104可交换文本消息和视频。视频可包括个性化视频。可基于预生成的存储在计算装置105或计算装置110中的视频模板来生成个性化视频。在一些实施例中，可将预生成的视频模板存储在信使服务系统130中并

按需下载到计算装置105或计算装置110。

[0036] 信使服务系统130可包括用于预处理视频的系统140。系统140可基于动画视频或实景真人视频生成视频模板。信使服务系统130可包括用于存储视频模板的视频模板数据库145。视频模板可下载到计算装置105或计算装置110。

[0037] 信使服务系统130还可被配置以存储用户简档135。用户简档135可包括用户102的面部的图像、用户104的面部的图像、以及其他人的面部的图像。可按需并基于许可将面部的图像下载到计算装置105或计算装置110。另外,可使用计算装置105生成用户102的面部的图像并将该图像存储在计算装置105的本地存储器中。可基于存储在计算装置105中的其他图像来生成面部的图像。计算装置105还可使用面部的图像基于预生成的视频模板来生成个性化视频。类似地,计算装置110可用于生成用户104的面部的图像。用户104的面部的图像可用于在计算装置110上生成个性化视频。在其他实施例中,用户102的面部的图像和用户104的面部的图像可相互用于在计算装置105或计算装置110上生成个性化视频。

[0038] 图2是示出用于实现生成个性化视频的方法的计算装置105(计算装置110)的一个示例性实施例的框图。在图2所示的示例中,计算装置110包括硬件组件和软件组件两者。具体地,计算装置110包括用于获取数字图像的相机205或任何其他图像捕获装置或扫描仪。计算装置110还可包括处理器模块210和用于存储软件组件和处理器可读(机器可读)指令或代码的存储模块215,该指令或代码在被处理器模块210执行时使计算装置105执行如本文所述的用于基于模板生成个性化视频的方法的至少一些步骤。计算装置105可包括图形显示系统230和通信模块240。在其它实施例中,计算装置105可包括附加的或不同的组件。此外,计算装置105可包括执行与图2中描绘的功能类似或等效的功能的更少的组件。

[0039] 计算装置110还可包括用于启动与另一计算装置(诸如计算装置110)的通信聊天的信使220和用于基于模板生成个性化视频的系统250。系统250将下面参考图4被更详细地描述。信使220和系统250可实现为存储在存储器存储装置215中的软件组件和处理器可读(机器可读)指令或代码,该指令或代码在被处理器模块210执行时使计算装置105执行如本文所述的用于提供通信聊天和生成个性化视频的方法的至少一些步骤。

[0040] 在一些实施例中,用于基于模板生成个性化视频的系统250可集成在信使220中。信使220的用户界面和用于基于模板的个性化视频的系统250可经由图形显示系统230提供。可经由通信模块240和网络120启动通信聊天。通信模块240可包括GSM模块、WIFI模块、蓝牙TM模块等。

[0041] 图3是示出根据本公开的一些示例性实施例的用于基于模板生成个性化视频的处理300的步骤的流程图。处理300可包括制作305、后期制作310、资源准备315、皮肤重新着色320、嘴唇同步和面部重演325、头发动画化330、眼部动画化335、以及部署340。资源准备315可通过用于预处理信使服务系统130(在图1中示出)中的视频的系统140执行。资源准备315的结果是生成可包括视频配置数据的视频模板。

[0042] 皮肤重新着色320、嘴唇同步和面部重演325、头发动画化330、眼部动画化335、和部署340可通过用于在计算装置105中的(在图2中示出的)基于模板生成个性化视频的系统250执行。系统250可接收用户面部的图像和视频配置数据,并且生成表征用户面部的个性化视频。

[0043] 皮肤重新着色320、嘴唇同步和面部重演325、头发动画化330、眼部动画化335、和

部署340可通过信使服务系统130(在图1中示出)中的用于预处理视频的系统140执行。系统140可接收用户面部的测试图像和视频配置文件。系统140可生成表征用户面部的测试个性化视频。操作者可检查测试个性化视频。基于检查的结果,视频配置文件可存储在视频模板数据库145中,然后可下载到计算装置105或计算装置110。

[0044] 制作305可包括理念和场景创建、预制作(在预制作期间识别位置、道具、演员、服饰和效果)、以及可要求一个或多个记录会话的制作本身。在一些示例性实施例中,可通过在色度键背景(在本文中也被称为绿色屏幕或色度键屏)上记录场景/演员来执行记录。为了允许随后的头部跟踪和资源清理,演员可佩戴具有跟踪标记的色度键面部掩模(例如,巴拉克拉瓦盔式帽(balaclavas)),该跟踪标记覆盖演员的面部,但漏出颈部和下巴的底部。在图5中详细示出理念和场景创建。

[0045] 在一个示例性实施例中,预制作和随后的制作步骤305是可选的。取代记录演员,可创建二维或三维动画或者可使用第三方镜头/图像。此外,可使用用户的图像的原始背景。

[0046] 图5是示出生成实景真人视频的处理500的框图。实景真人视频还可用于生成视频模板以生成个性化视频。处理500可包括在步骤505生成理念并在步骤510创建场景。处理500可在步骤515继续进行预制作,随后进行制作305。制作305可包括使用色度键屏525或在现实生活位置530进行记录。

[0047] 图6示出用于生成视频模板的示例实景真人视频的帧。在现实生活位置530处记录视频605和视频615的帧。使用色度键屏525来记录视频610、视频620和视频625的帧。演员可佩戴色度键面部掩模630,该色度键面部掩模630具有覆盖演员面部的跟踪标记。

[0048] 后期制作310可包括视频编辑或动画化、视觉效果、清理、声音设计和语音记录。

[0049] 在资源准备315期间,准备进一步部署的资源可包括以下组件:没有演员的头部的背景镜头(即,准备移除演员的头部的清理后背景);演员在黑色背景上的镜头(只用于所记录的个性化视频);帧的前景序列;具有通用头部和配乐的示例镜头;头部位置、旋转和比例的坐标;附着于头部的动画元素(可选的);有旁白和无旁白的配乐;单独文件中的旁白(可选的)等。所有这些组件都是可选的,并且可以以不同的格式呈现。组件的数量和配置取决于个性化视频的格式。例如,对于定制的个性化视频,不需要旁白,如果使用来自用户的图片的原始背景等,则不需要背景镜头和头部坐标。在一个示例性实施例中,可指示(例如,手动地)面部需要位于的区域,而不是准备具有坐标的文件。

[0050] 皮肤重新着色320允许将个性化视频中的演员的皮肤的颜色与用户的图像上的面部的颜色匹配。为了实施本步骤,可准备具体指示背景的哪个部分必须重新着色的皮肤掩模。优选地,对于演员的每个身体部分(颈部、左手和右手等)具有单独的掩模。

[0051] 皮肤重新着色320可包括面部图像光照归一化。图7示出根据一个示例性实施例的面部的原始图像705和具有归一化光照的面部的图像710。由不均匀光照引起的阴影或亮点影响颜色分布,并可能导致重新着色后肤色太暗或太亮。为了避免这种情况,可检测并去除用户的面部中的阴影和亮点。面部图像光照归一化处理包括以下步骤。可使用深度卷积神经网络来变换用户的面部的图像。网络可接收在任意光照下拍摄的在肖像图像形式下的原始图像705,并在保持原始图像705中的主题相同的同时,改变原始图像705的光照以使原始图像705光照均匀。因此,面部图像光照归一化处理的输入包括用户的面部的图像形式的原

始图像705和面部界标。面部图像光照归一化处理的输出包括具有归一化光照的面部的图像710。

[0052] 皮肤重新着色320可包括掩模创建和身体统计。可以仅存在用于整个皮肤的掩模或用于身体部分的单独掩模。此外,可针对视频中的不同场景(例如,由于显著的光照变化)创建不同的掩模。掩模例如可通过诸如键控的技术在一些人类指导下半自动地创建。可将准备好的掩模合并到视频资源中,然后在重新着色中使用。并且,为了实时避免不必要的计算,可预先针对每个掩模计算颜色统计。统计可包括每个颜色通道的平均值、中间值、标准偏差和一些百分位数。可在红色、绿色和蓝色(RGB)颜色空间以及其它颜色空间(色调、饱和度、值(HSV)颜色空间、CIELAB颜色空间(也称为CIEL\*a\*b\*或缩写为“LAB”颜色空间)等)中计算统计。掩模创建处理的输入可包括以视频或图像序列形式的针对具有未覆盖的皮肤的演员的身体部分的灰度掩模。掩模创建处理的输出可包括针对每个掩模被压缩并合并成视频和颜色统计的掩模。

[0053] 皮肤重新着色320可包括面部统计计算。图8示出根据一个示例性实施例的分割的头部图像805,分割的头部图像805具有面部界标810、以及面部掩模815。基于对用户的头部图像和面部界标的分割,可创建用户的面部掩模815。面部掩模815中可不包括诸如眼部、嘴、头发或配饰(如眼镜)的区域。用户的分割的头部图像805和面部掩模可用于计算用户的面部皮肤的统计。因此,面部统计计算的输入可包括用户的分割的头部图像805、面部界标810和面部分割,并且面部统计计算的输出可包括用户的面部皮肤的颜色统计。

[0054] 皮肤重新着色320还可包括肤色匹配和重新着色。图9示出根据一个示例性实施例的表征用户面部的帧905、皮肤掩模910和对皮肤掩模910重新着色的结果915。可使用描述演员的皮肤和用户的皮肤中的颜色分布的统计来执行肤色匹配和重新着色,并且可在计算装置上实时地执行背景帧的重新着色。对于每个颜色通道,可执行分布匹配并且可修改背景像素的值,以便使变换值的分布接近面部值的分布。分布匹配可在假设颜色分布是正常的情况下执行,或者通过应用诸如多维概率密度函数转移的技术来执行。因此,肤色匹配和重新着色处理的输入可包括背景帧、帧的演员皮肤掩模、每个掩模的演员身体皮肤颜色统计、以及用户面部皮肤颜色统计,并且输出可包括所有身体部分未覆盖的皮肤被重新着色的背景帧。

[0055] 在一些实施例中,为了应用皮肤重新着色320,可记录具有不同肤色的几个演员,然后可使用具有与用户的图像的肤色最接近的肤色的个性化视频的版本。

[0056] 在一个示例性实施例中,取代皮肤再着色320,可使用预定查找表(LUT)来针对场景的光照调整面部的颜色。LUT还可用于改变面部的颜色,例如,使面部为绿色。

[0057] 嘴唇同步和面部重演325可产生真实感的面部动画。图10示出嘴唇同步和面部重演325的示例处理。图10示出根据一个示例性实施例的面部同步演员面部的图像1005、面部同步演员面部界标的图像1010、用户的面部界标的图像1015、以及具有面部同步演员的面部表情的用户面部的图像1020。嘴唇同步和面部重演325的步骤可包括记录面部同步演员和预处理源视频/图像,以获得面部同步演员面部的图像1005。然后,如面部同步演员面部界标的图像1010所示,可提取面部界标。步骤还可包括注视跟踪面部同步演员。在一些实施例中,取代记录面部同步演员,可使用预先准备的动画2D或3D面部和嘴部区域模型。可通过机器学习技术来生成动画2D或3D面部和嘴部区域模型。

[0058] 可选地,可执行面部界标的微调。在一些示例性实施例中,手动地执行面部界标的微调。这些步骤可在准备视频配置文件时在云端执行。在一些示例性实施例中,这些步骤可在资源准备315期间执行。然后,如用户的面部界标的图像1015所示,可提取用户的面部界标。同步和面部重演325的下一步骤可包括用提取的界标来动画化目标图像以获得具有面部同步演员的面部表情的用户面部的图像1020。步骤可基于用户面部的图像在计算装置上执行。在美国专利申请号16/251,472中详细描述了动画化方法,其公开内容通过引用整体合并与此。嘴唇同步和面部重演325还可用人工智能制造的头部转向来丰富。

[0059] 在一些示例性实施例中,在用户拍摄图像之后,可创建用户头部的三维模型。在本实施例中,可省略嘴唇同步和面部重演325的步骤。

[0060] 可执行头发动画化330以动画化用户的头发。例如,如果用户具有头发,则当用户移动或旋转他的头部时,头发可以是动画。头发动画化330在图11中示出。图11示出根据一个示例性实施例的分割的面部图像1105、头发掩模1110、移动到面部图像的头发掩模1115以及应用到面部图像的头发掩模1120。头发动画化330可包括以下步骤中的一项或多项:对头发类型进行分类、修改头发的外观、修改发型、使头发更长、改变头发的颜色、切割头发并动画化头发等。如图11所示,可获得以分割的面部图像1105形式的面部图像。然后,可将头发掩模1110应用到分割的面部图像1105。图像1115示出移动到面部图像的头发掩模1110。图像1120示出应用到面部图像的头发掩模1110。在美国专利申请号16/551,756中详细描述了头发动画化330,其公开内容通过引用整体合并与此。

[0061] 眼部动画化335可使用户的面部表情更真实。在图12中详细示出眼部动画化335。眼部动画化335的处理可由以下步骤组成:用户面部的眼部区域的重建、注视移动步骤以及眨眼步骤。在眼部区域的重建处理中,将眼部区域分割成以下部分:眼球、虹膜、瞳孔、睫毛、眼睑。如果眼部区域的一些部分(例如,虹膜或眼睑)不是完全可见的,则可合成该部分的完整纹理。在一些实施例中,可拟合眼部的3D可变形模型,并且可获得眼部的3D形状以及眼部的纹理。图12示出眼部的原始图像1205,具有眼部的重建巩膜的图像1210、以及具有重建虹膜的图像1215。

[0062] 注视移动步骤包括在面部同步演员的视频中跟踪注视方向和瞳孔位置。如果面部同步演员的眼部移动不够丰富,则可手动编辑数据。然后,可通过合成具有变换的眼部形状和与面部同步演员的虹膜位置相同的虹膜位置的新眼部图像,来将注视移动转移到用户的眼部区域。图12示出具有重建的移动的虹膜的图像1220。

[0063] 在眨眼步骤期间,可通过跟踪面部同步演员的眼部来确定用户眼部的可见部分。可基于眼部区域的重建来生成眼睑和睫毛的改变的外观。

[0064] 如果使用生成对抗网络(GAN)进行面部重演,则可明确地(如前所述)或隐含地进行眼部动画化335的步骤。在后一种情况下,神经网络可隐含地捕获来自用户面部和源视频的图像的所有必要信息。

[0065] 在部署340期间,用户面部可被逼真地动画化并且自动地插入镜头模板中。来自先前步骤(资源准备315、皮肤重新着色320、嘴唇同步和面部重演325、头发动画化330和眼部动画化335)的文件可用作配置文件的数据。可生成具有预定的一组用户面部的个性化视频的示例以用于初始检查。在消除在检查期间识别的问题之后,可部署个性化视频。

[0066] 配置文件还可包括允许指示定制的个性化视频的文本参数的组件。定制的个性化

视频是一种允许用户在最终视频顶层添加用户想要的任何文本的个性化视频。在2019年10月23日提交的题为“用于生成具有定制文本消息的个性化视频的系统和方法 (SYSTEMS AND METHOD FOR GENERATING PERSONALIZED VIDEOS WITH CUSTOMIZED TEXT MESSAGES)”的美国专利申请号16/661,122中更详细地描述了具有定制文本消息的个性化视频的生成,其公开内容通过引用整体合并与此。

[0067] 在一个示例性实施例中,个性化视频的生成还可包括生成用户的头部的明显头部转向的步骤;身体动画,改变服装;面部增强(诸如发型改变、美化、添加附件等);改变场景光照;合成可读/唱用户输入的文本的声音或将语音转换为与用户的语音相匹配的语音;性别切换;根据用户输入构造背景和前景;等。

[0068] 图4是示出根据一些示例性实施例的用于基于模板生成个性化视频的系统250的功能400的示意图。系统250可接收作为用户面部图像405示出的源面部的图像和包括视频配置数据410的视频模板。视频配置数据410可包括数据序列420。例如,视频配置数据410可包括:帧图像的序列、限定面部区域在帧图像中的位置的面部区域参数的序列、以及限定面部界标在帧图像中的位置的面部界标参数的序列。每个面部界标参数可与面部表情相对应。可基于动画视频或基于实景真人视频生成帧图像的序列。可基于表征面部同步演员的面部的实景真人视频来生成面部界标参数的序列。视频配置数据410还可包括皮肤掩模、眼部参数、嘴部区域图像、头部参数、动画对象图像、预设文本参数等。视频配置数据可包括限定帧图像中所表征的至少一个演员的身体的皮肤区域的皮肤掩模的序列。在一个示例性实施例中,视频配置数据410还可包括嘴部区域图像的序列。每个嘴部区域图像可与至少一个帧图像相对应。在另一示例性实施例中,视频配置数据410可包括限定在帧图像中所表征的面部同步演员的巩膜中的虹膜的位置的眼部参数的序列和/或限定头部的旋转、转向、比例和其他参数的头部参数的序列。在另一示例性实施例中,视频配置数据410还可包括动画对象图像的序列。每个动画对象图像可与至少一个帧图像相对应。视频配置数据410还可包括配乐450。

[0069] 系统250可基于用户面部图像405确定用户数据435。用户数据可包括用户的面部界标、用户面部掩模、用户颜色数据、用户头发掩模等。

[0070] 系统250可基于用户数据435和数据序列420生成作为个性化视频440显示的输出视频的帧445。系统250还可将配乐添加到个性化视频440中。个性化视频440可通过修改帧图像的序列的帧图像来生成。帧图像的修改可包括:修改用户面部图像405以获得另外的图像,该另外的图像表征采用了与面部界标参数相对应的面部表情的源面部。可基于与帧图像相对应的面部界标参数来执行修改。另外的图像可在由与帧图像相对应的面部区域参数确定的位置处插入到帧图像中。在一个示例性实施例中,输出视频的生成还可包括确定与源面部相关联的颜色数据,并且基于颜色数据来对帧图像中的皮肤区域重新着色。此外,输出视频的生成包括将与帧图像相对应的嘴部区域插入到帧图像中。生成输出视频的其他步骤可包括基于与帧相对应眼部参数来生成眼部区域的图像,并且将该眼部区域的图像插入到帧图像中。在一个示例性实施例中,输出视频的生成还可包括基于源面部图像来确定头发掩模,基于该头发掩模和与帧图像相对应的头部参数来生成头发图像,并且将该头发图像插入到帧图像中。此外,输出视频的生成包括将与帧图像相对应的动画对象图像插入到帧图像中。

[0071] 图13和图14示出根据一些示例性实施例的基于视频模板生成的示例个性化视频的帧。图13示出拍摄的具有演员的个性化视频1305,在该个性化视频1305中,已执行重新着色。图13进一步示出基于从第三方获得的库存视频创建的个性化视频1310。在个性化视频1310中,用户面部1320插入到库存视频中。图13进一步示出个性化视频1315,该个性化视频1315是具有添加在二维动画顶层的用户头部1325的2D动画。

[0072] 图14示出个性化视频1405,该个性化视频1405是具有插入在3D动画中的用户面部1415的3D动画。图14进一步示出具有效果、动画元素1420以及可选地添加在用户面部的图像顶层的文本的个性化视频1410。

[0073] 图15是示出根据本公开的一些示例性实施例的用于基于模板生成个性化视频的方法1500的流程图。方法1500可通过计算装置105执行。方法1500可从在步骤1505接收视频配置数据开始。视频配置数据可包括帧图像的序列,限定面部区域在帧图像中的位置的面部区域参数的序列,以及限定面部界标在帧图像中的位置的面部界标参数的序列。每个面部界标参数可与面部表情相对应。在一个示例性实施例中,可基于动画视频或实景真人视频生成帧图像的序列。可基于表征面部同步演员的面部的实景真人视频来生成面部界标参数的序列。视频配置数据可包括以下的一个或多个:限定帧图像中所表征的至少一个演员的身体的皮肤区域的皮肤掩模的序列;嘴部区域图像的序列,其中,每个嘴部区域图像可与至少一个帧图像相对应;限定帧图像中所表征的面部同步演员的巩膜中的虹膜的位置的眼部参数的序列;限定头部的旋转、比例、转向和其他参数的头部参数的序列;动画对象图像的序列,其中,每个动画对象图像与至少一个帧图像相对应;等。

[0074] 方法1500可在步骤1510继续接收源面部的图像。方法还可包括在步骤1515生成输出视频。具体地,输出视频的生成可包括:修改帧图像的序列的帧图像。可通过修改源面部的图像来修改帧图像以得到另外的图像,该另外的图像表征采用了与面部界标参数相对应的面部表情的源面部。可基于与帧图像相对应的面部界标参数来修改源面部的图像。另外的图像可在通过与帧图像相对应的面部区域参数确定的位置处插入到帧图像中。在一个示例性实施例中,输出视频的生成还可包括以下步骤中的一个或多个:确定与源面部相关联的颜色数据,并且基于该颜色数据对帧图像中的皮肤区域重新着色;将与帧图像相对应的嘴部区域插入到帧图像中;基于与帧相对应眼部参数来生成眼部区域的图像,将该眼部区域的图像插入到帧图像中;基于源面部图像确定头发掩模并基于该头发掩模和与帧图像相对应的头部参数来生成头发图像,将该头发图像插入到帧图像中;以及将与帧图像相对应的动画对象图像插入到帧图像中。

[0075] 图16示出可用于实现在本文中描述的方法的示例计算系统1600。计算系统1600可在与计算装置105和110、信使服务系统130、信使220和用于基于模板生成个性化视频的系统250的类似的环境中实现。

[0076] 如图16所示,计算系统1600的硬件组件可包括一个或多个处理器1610和存储器1620。存储器1620部分地存储用于由处理器1610执行的指令和数据。存储器1620可在系统1600运行时存储可执行代码。系统1600还可包括可选的大容量存储装置1630、可选的便携式存储介质驱动器1640、一个或多个可选的输出装置1650、一个或多个可选的输入装置1660、可选的网络接口1670、以及一个或多个可选的外围装置1680。计算系统1600还可包括一个或多个软件组件1695(例如,可实现如本文描述的用于基于模板生成个性化视频的方法)

法的软件组件)。

[0077] 图16所示的组件描绘为经由单个总线1690连接。组件可通过一个或多个数据传送设备或数据网络连接。处理器1610和存储器1620可经由本地微处理器总线连接,并且大容量存储装置1630、外围装置1680、便携式存储装置1640和网络接口1670可经由一个或多个输入/输出(I/O)总线连接。

[0078] 可用磁盘驱动器、固态硬盘驱动器或光盘驱动器来实现的大容量存储装置1630是用于存储供处理器1610使用的数据和指令的非易失性存储装置。大容量存储装置1630可存储用于实现在本文中描述的实施例的系统软件(例如,软件组件1695)。

[0079] 便携式存储媒体驱动器1640结合便携式非易失性存储介质(诸如,压缩盘(CD)或数字视频盘(DVD))操作以将数据和代码输入到计算系统1600和从计算系统1600输出数据和代码。用于实现在本文中描述的实施例的系统软件(例如,软件组件1695)可存储在这样的便携式介质上并经由便携式存储介质驱动器1640输入到计算系统1600。

[0080] 可选输入装置1660提供用户界面的一部分。输入装置1660可包括用于输入字母数字和其他信息的字母数字键盘(诸如键盘)或定点装置(诸如鼠标、跟踪球、指示笔或光标方向键)。输入装置1660还可包括相机或扫描仪。此外,图16所示的系统1600包括可选的输出装置1650。合适的输出装置包括扬声器、打印机、网络接口和监视器。

[0081] 网络接口1670可用于经由一个或多个通信网络与外部装置、外部计算装置、服务器和联网系统通信,该通信网络诸如一个或多个有线网络、无线网络或光学网络,包括例如互连网、内联网、局域网(LAN)、广域网(WAN)、蜂窝电话网络、蓝牙无线电和基于IEEE802.11的射频网络等。网络接口1670可以是网络接口卡(诸如以太网卡、光收发器、射频收发器)或能够发送和接收信息的任何其他类型的装置。可选外围装置1680可包括任何类型的计算机支持装置,以向计算机系统添加附加功能。

[0082] 包含在计算系统1600中的组件旨在表示一大类计算机组件。因此,计算系统1600可以是服务器、个人电脑、手持式计算装置、电话、移动计算装置、工作站、小型计算机、大型计算机、网络节点或任何其它计算装置。计算系统1600还可包括不同的总线配置、联网平台、多处理器平台等。可使用各种操作系统(OS),包括UNIX、Linux、Windows、Macintosh OS、Palm OS和其他合适的操作系统。

[0083] 上述功能中的一些可由存储在存储介质(例如,计算机可读介质或处理器可读介质)上的指令组成。指令可由处理器检索并执行。存储介质的一些示例是存储装置、磁带、磁盘等。指令在被处理器执行时是可操作的,以指导处理器根据本发明进行操作。本领域技术人员熟悉指令、处理器和存储介质。

[0084] 值得注意的是,适用于执行在本文中描述的处理的任何硬件平台都适用于本发明。在本文中使用的术语“计算机可读存储介质”和“计算机可读存储介质”是指参与向处理器提供指令以供执行的任何介质。这样的介质可采取许多形式,包括但不限于非易失性介质、易失性介质和传输介质。非易失性介质例如包括光盘或磁盘(诸如固定盘)。易失性介质包括动态存储器(诸如系统随机存取存储器(RAM))。

[0085] 传输介质包括同轴电缆、铜线和光纤等,该传输介质包括包含总线的一个实施例的导线。传输介质还可采用声波或光波的形式(诸如在射频(RF)和红外(IR)数据通信期间生成的那些声波或光波)。计算机可读介质的常见形式包括例如软盘、柔性盘、硬盘、磁带、

任何其他磁性介质、CD只读存储器 (ROM) 盘、DVD、任何其它光学介质、具有标记或孔的图案的任何其它物理介质、RAM、可编程序只读存储器 (PROM)、可擦除可编程只读存储器 (EPROM)、带电可擦可编程只读存储器 (EEPROM)、任何其它存储芯片或盒式磁带、载波或计算机可从其读取的任何其它介质。

[0086] 各种形式的计算机可读介质可涉及将一个或多个指令的一个或多个序列携带到处理器以供执行。总线将数据携带到系统RAM,处理器从该系统RAM检索并执行指令。由系统处理器接收的指令可以可选地在由处理器执行之前或之后存储在固定磁盘上。

[0087] 因此,已经描述了用于基于模板生成个性化视频的方法和系统。尽管已经参照特定示例性实施例描述了实施例,但是很明显,在不脱离本申请的更广泛的精神和范围的情况下,可对这些示例性实施例进行各种修改和改变。因此,说明书和附图应被认为是说明性的而不是限制性的。

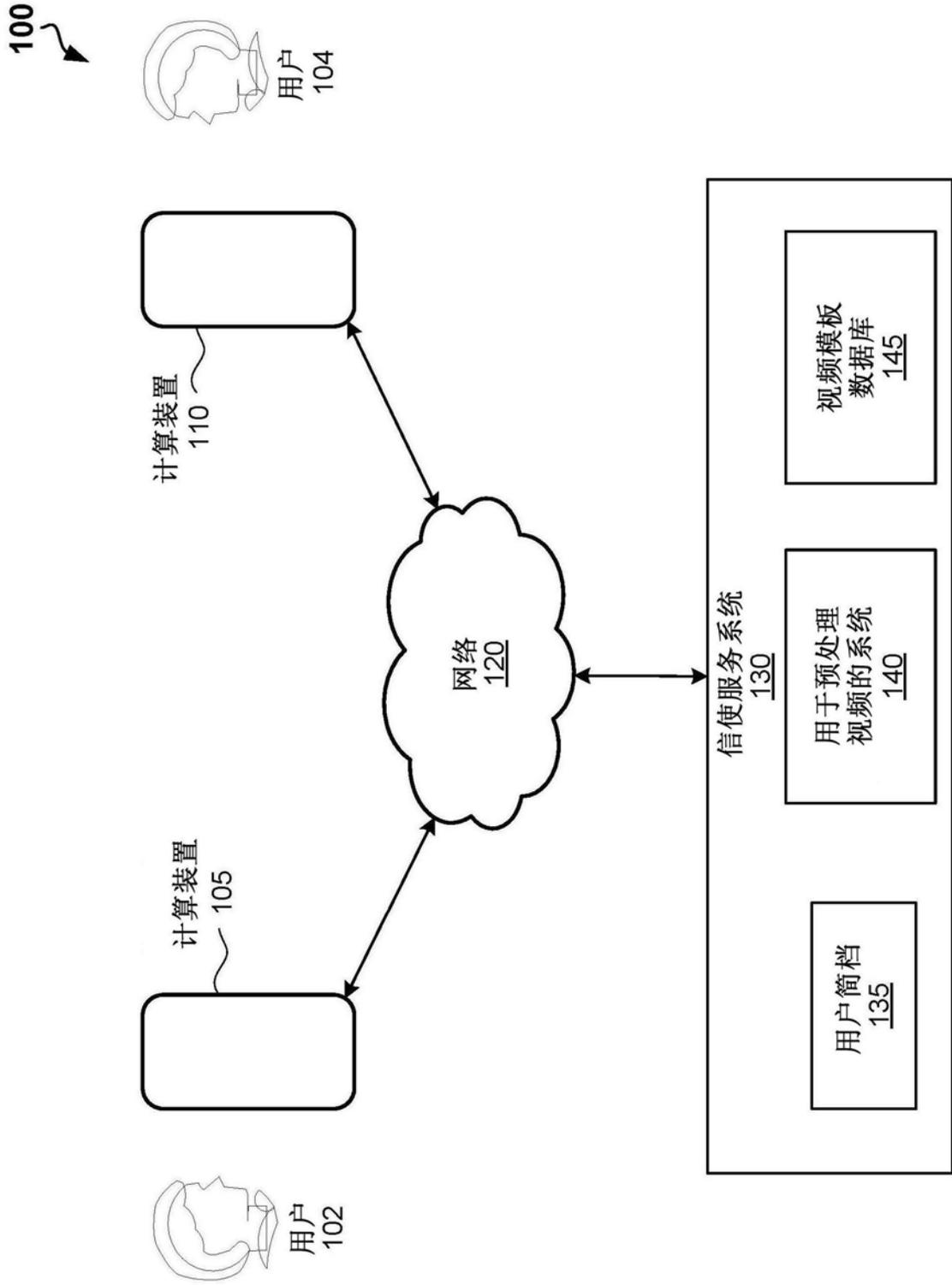


图1

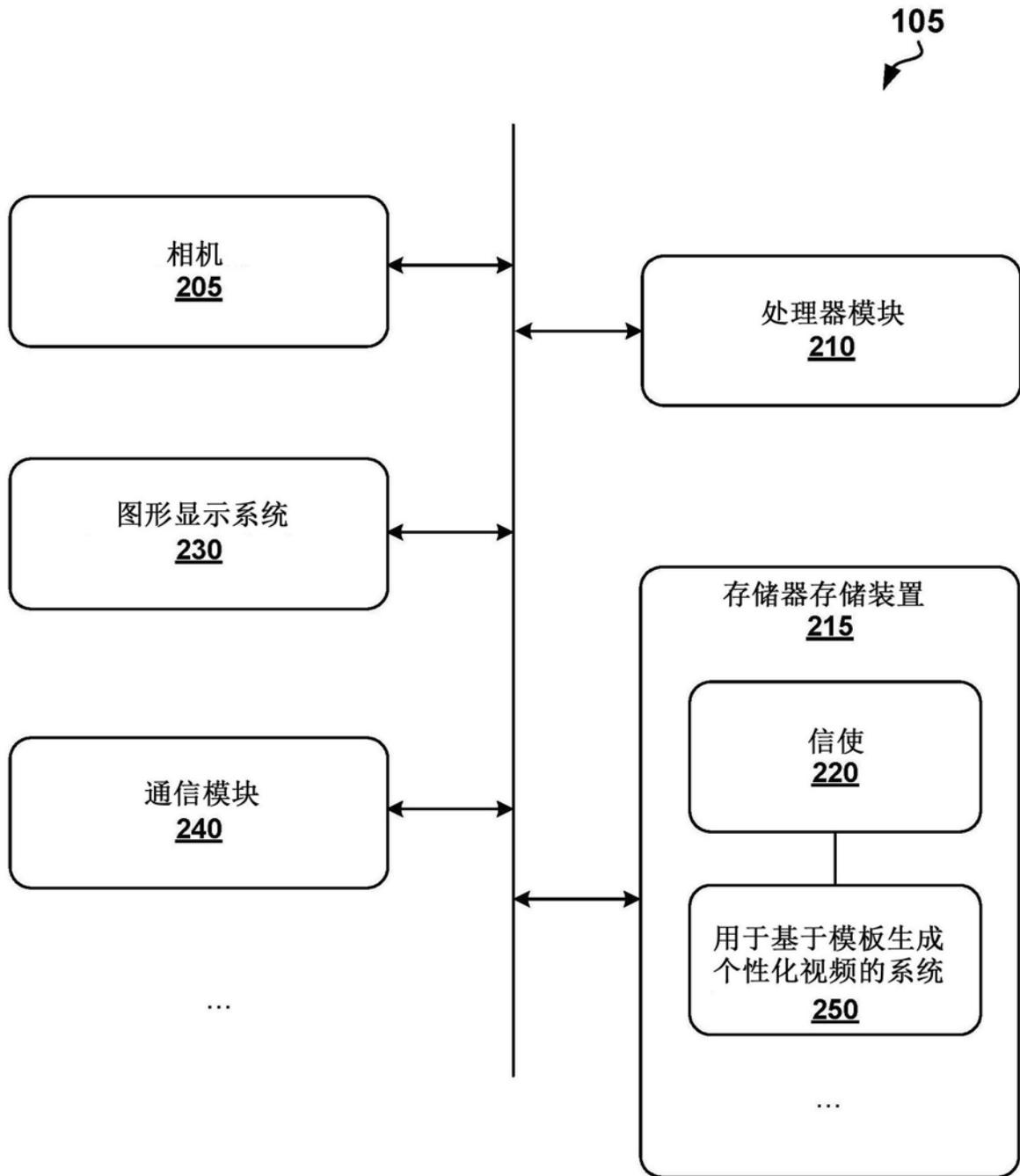


图2

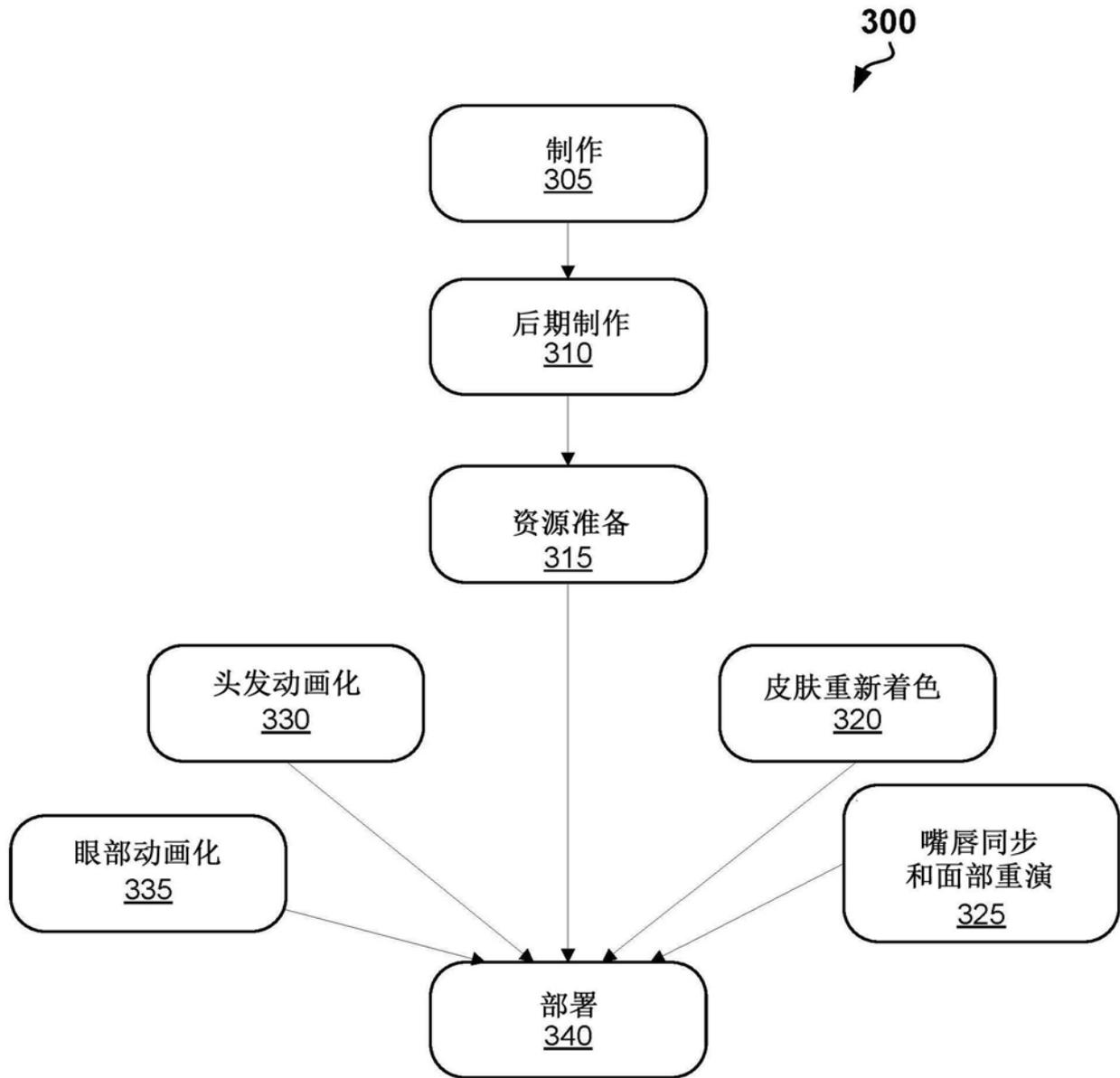


图3

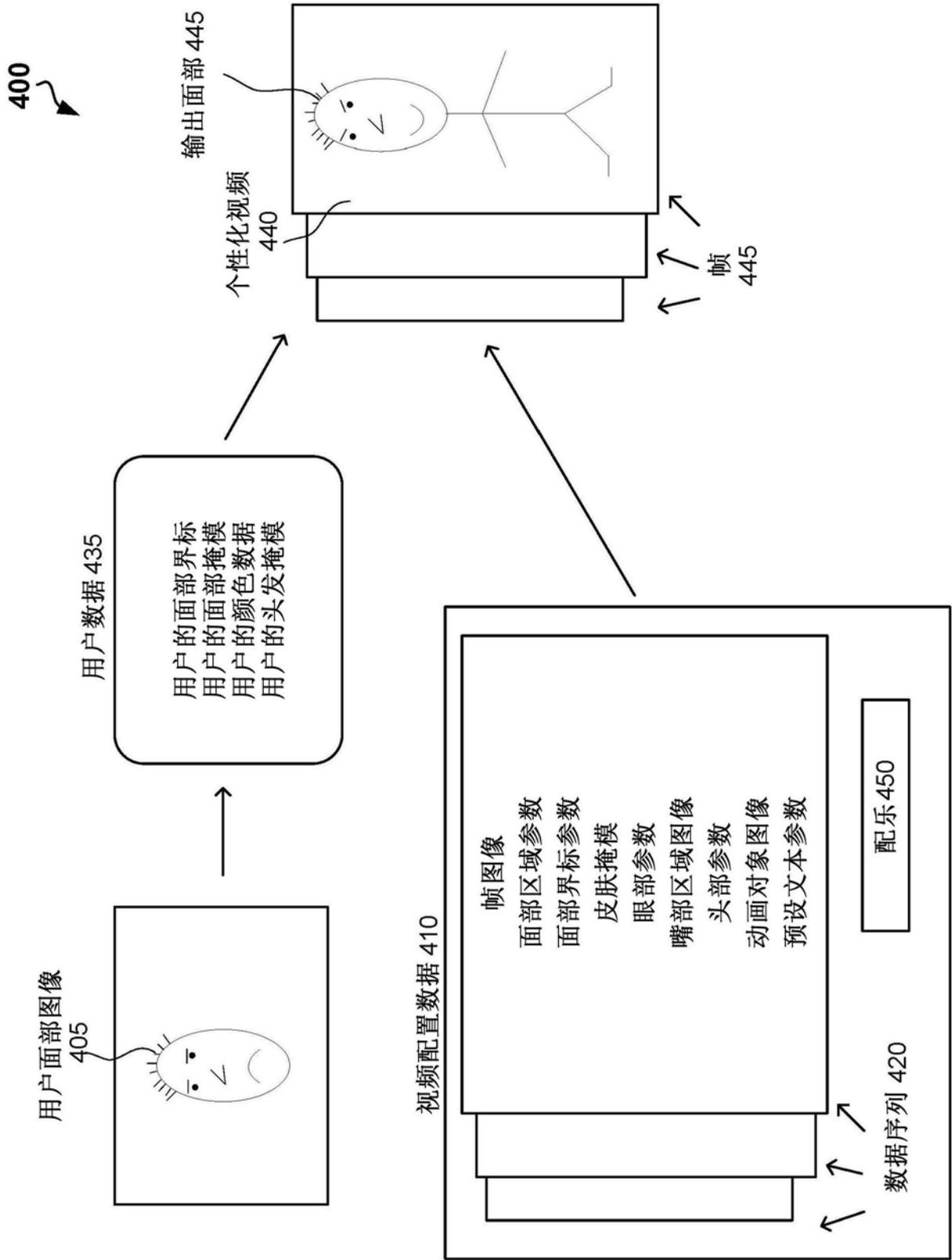


图4

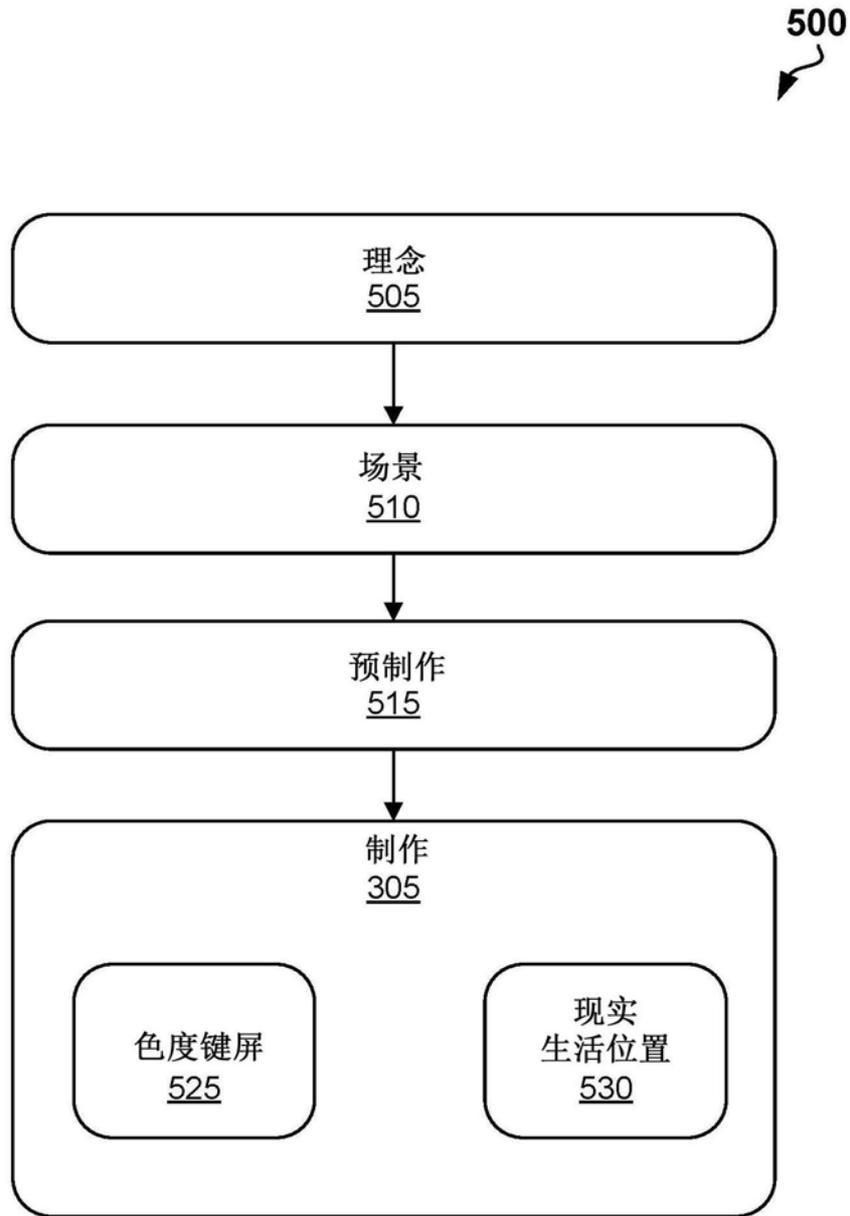


图5

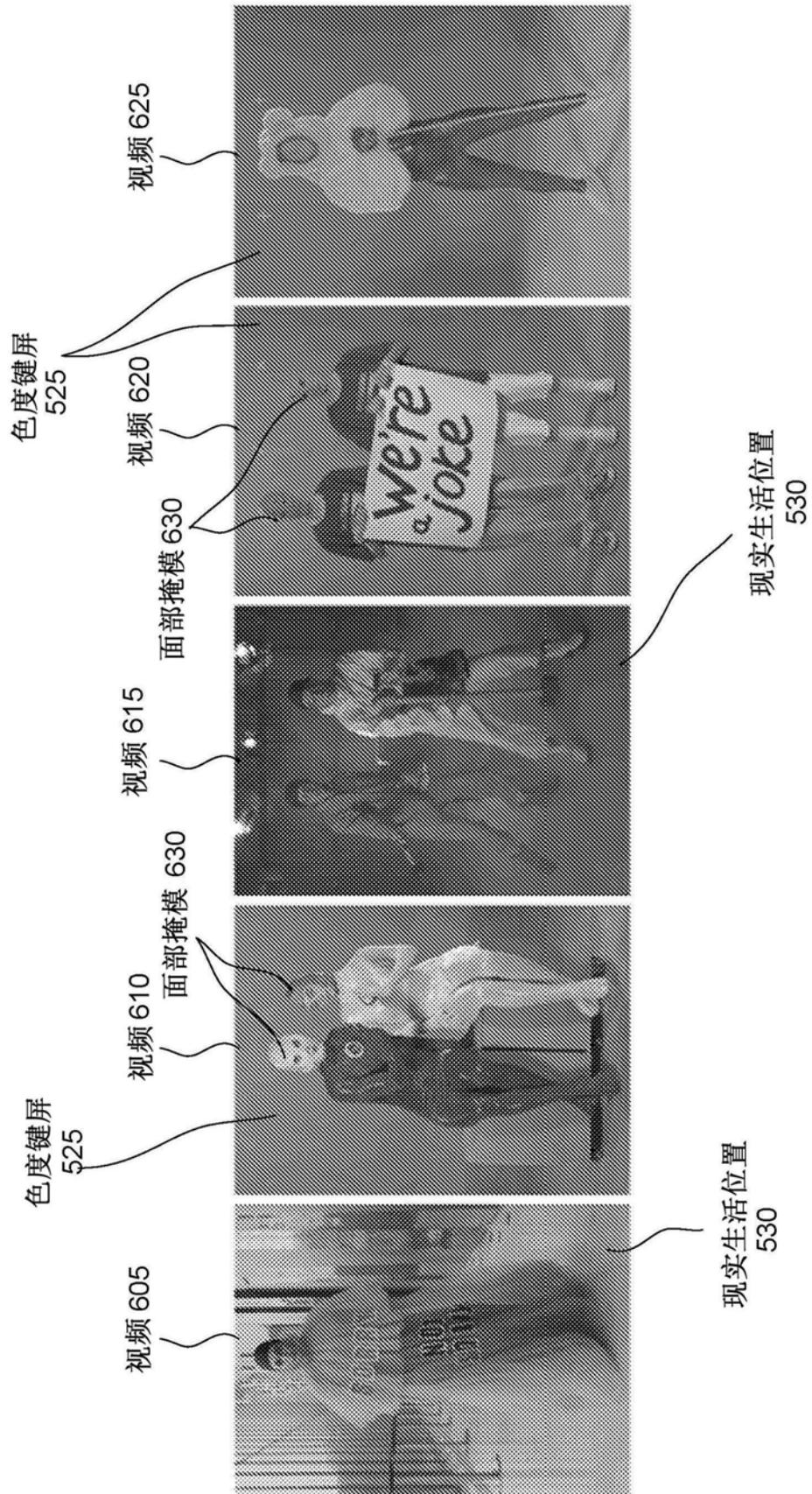


图6

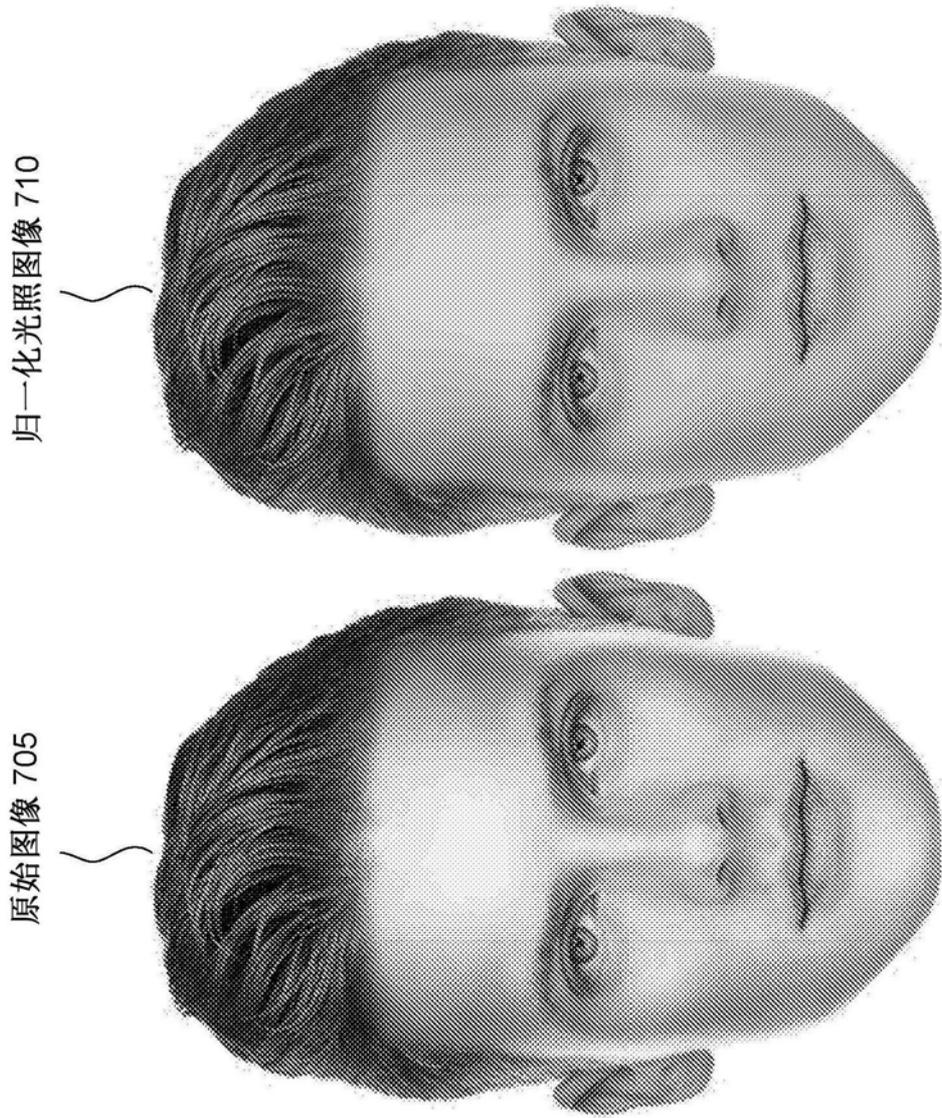


图7

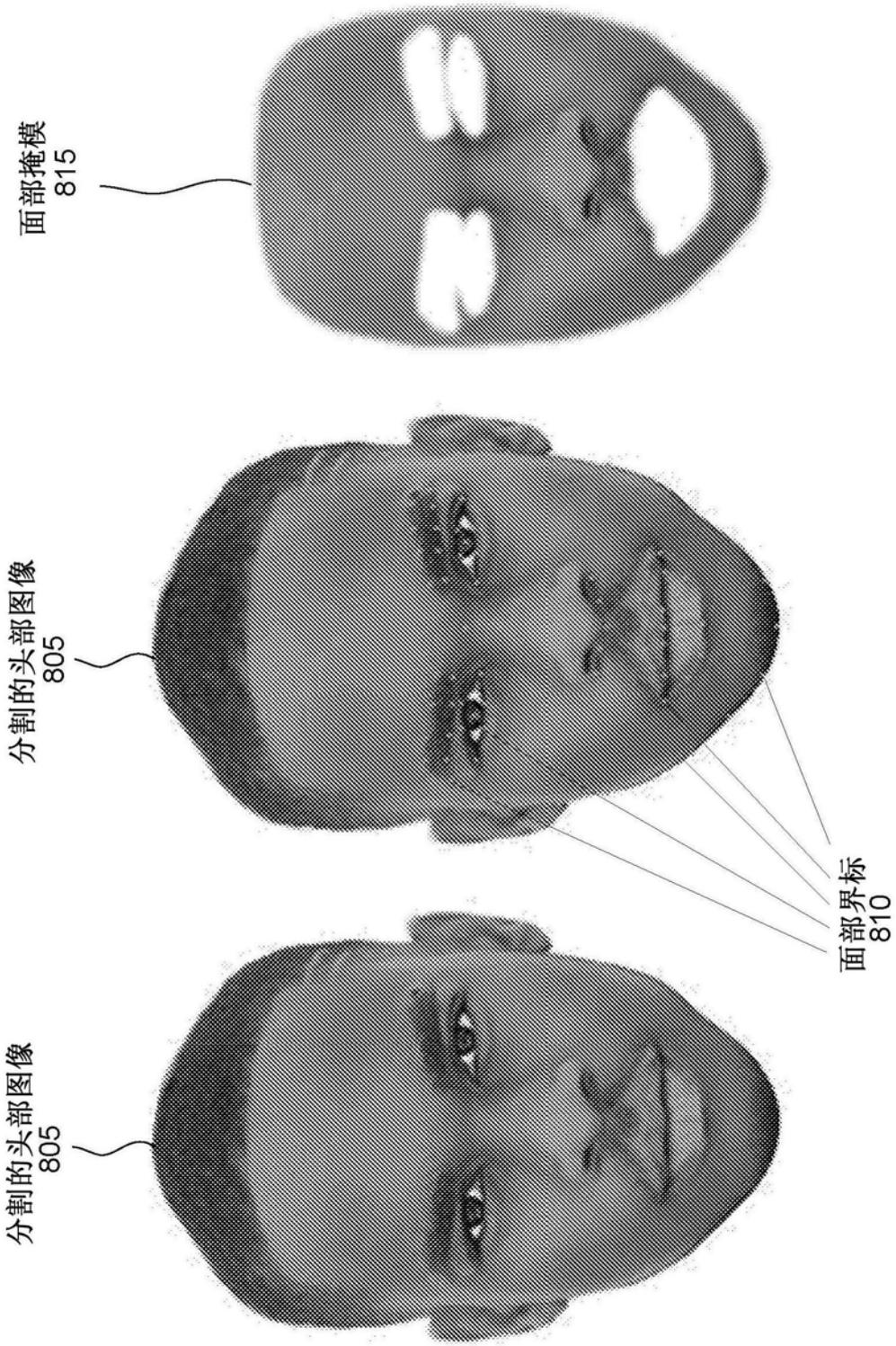
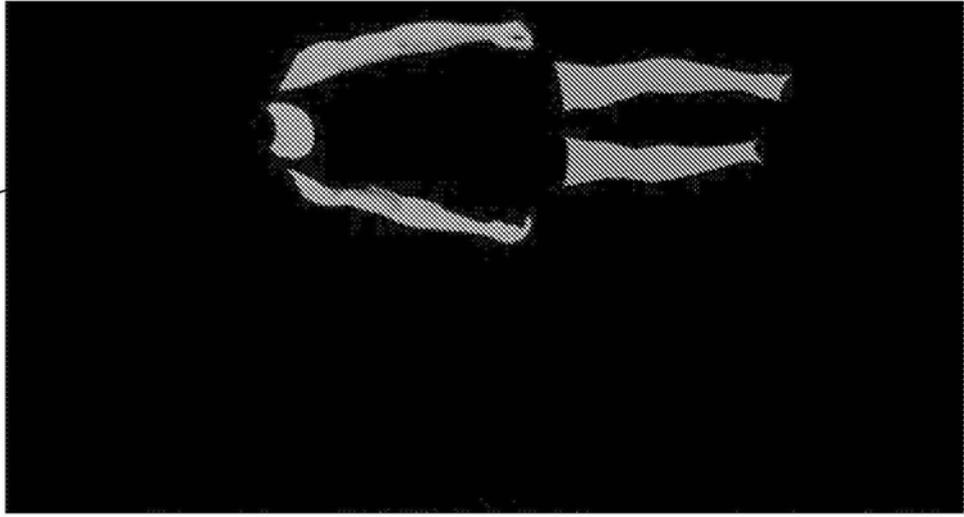


图8

重新着色的结果  
915



皮肤掩模  
910



具有用户面部的帧  
905



图9

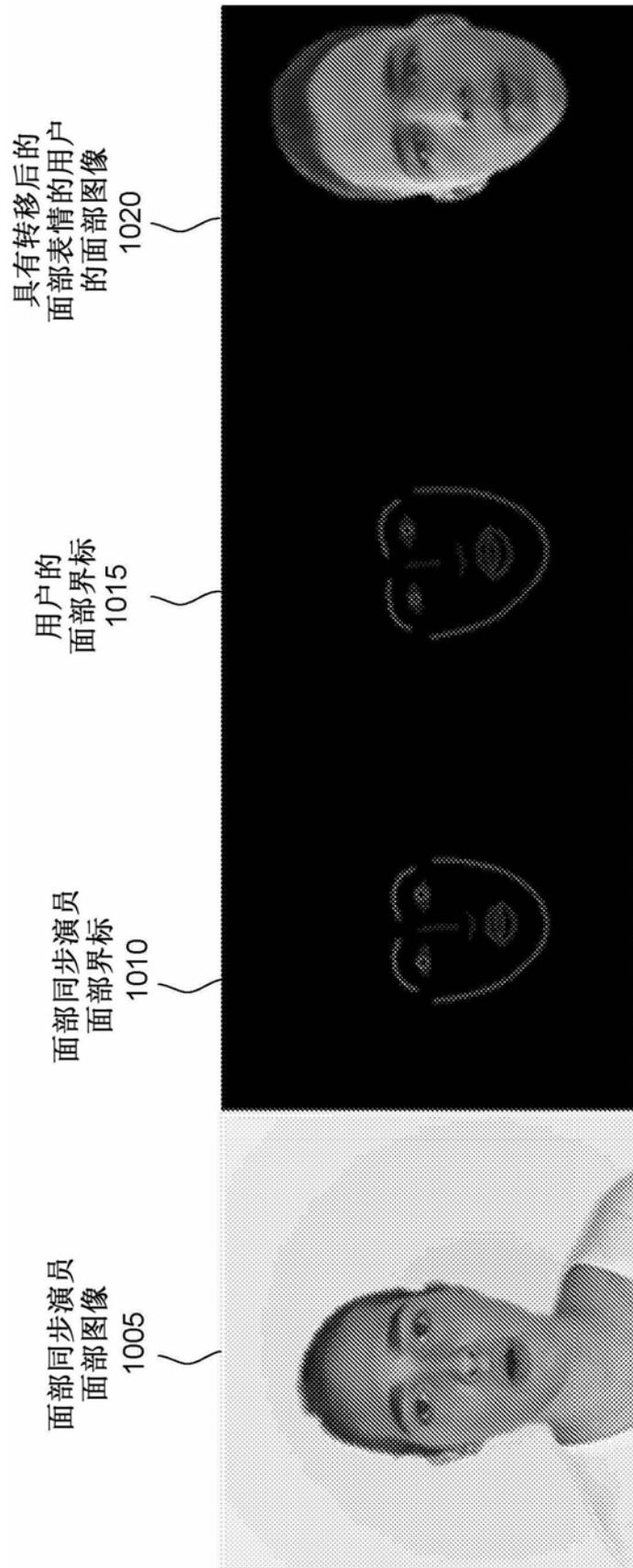


图10

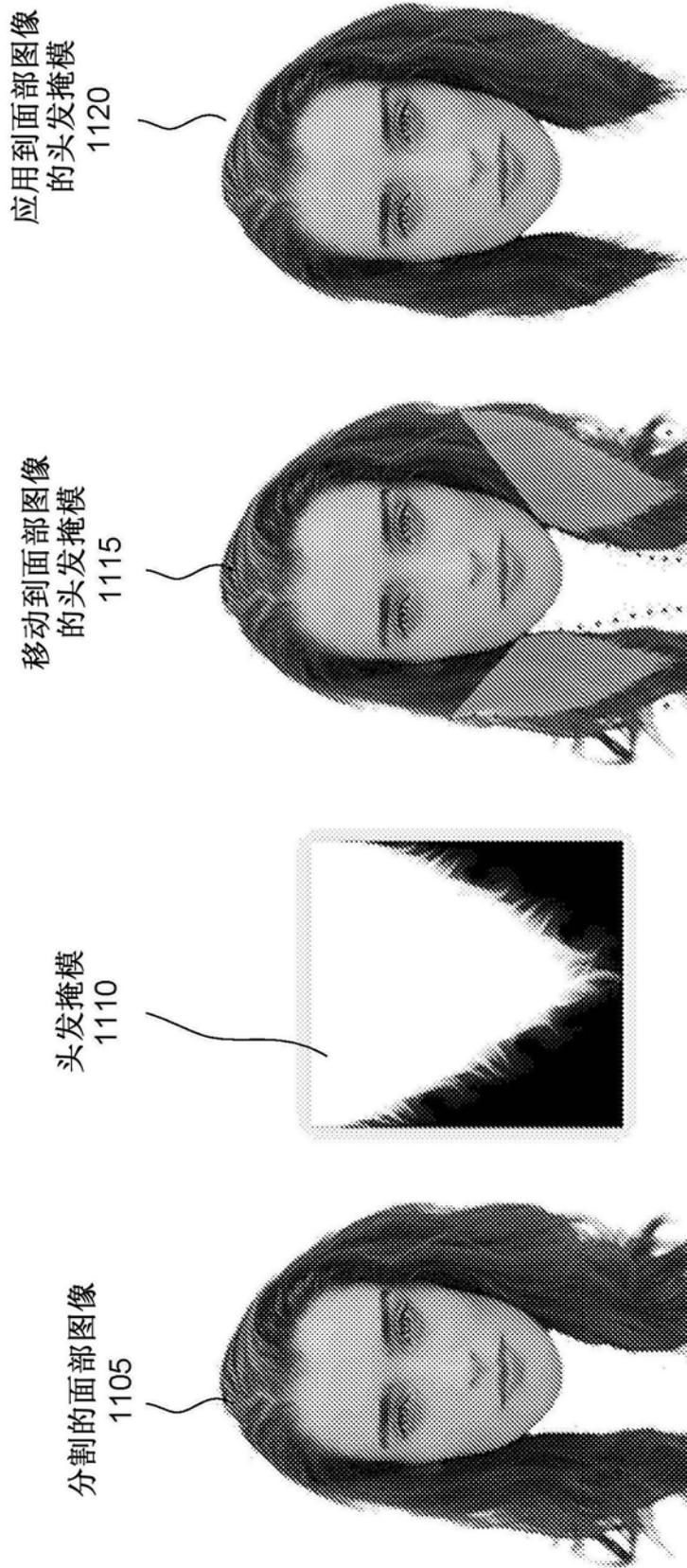


图11

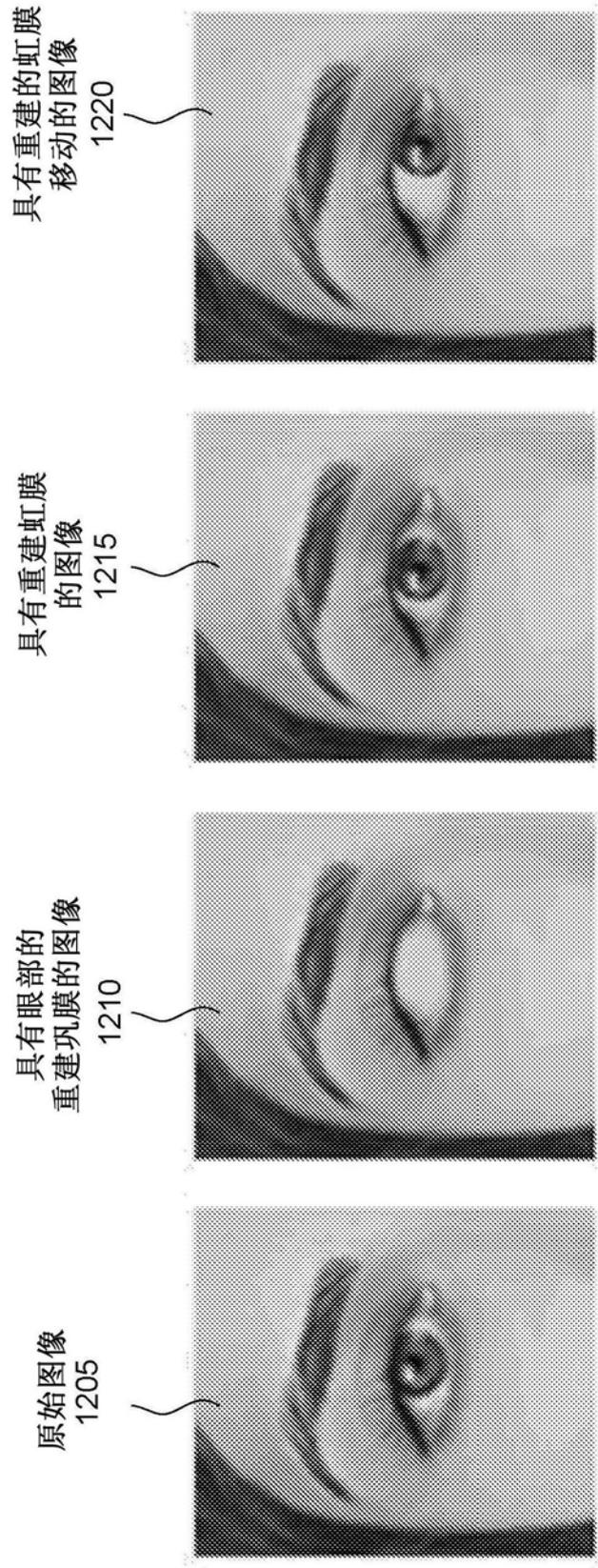


图12



图13

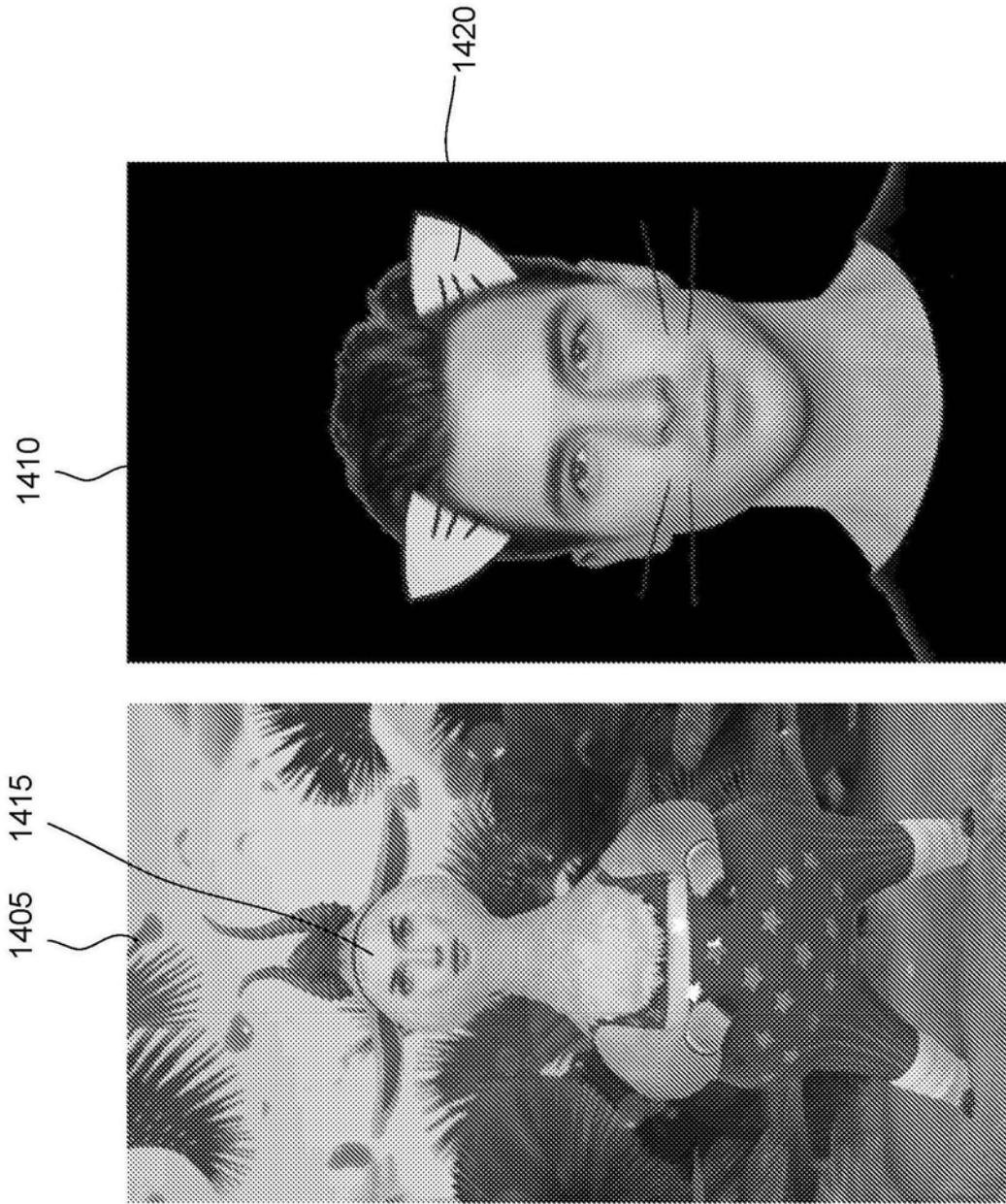


图14

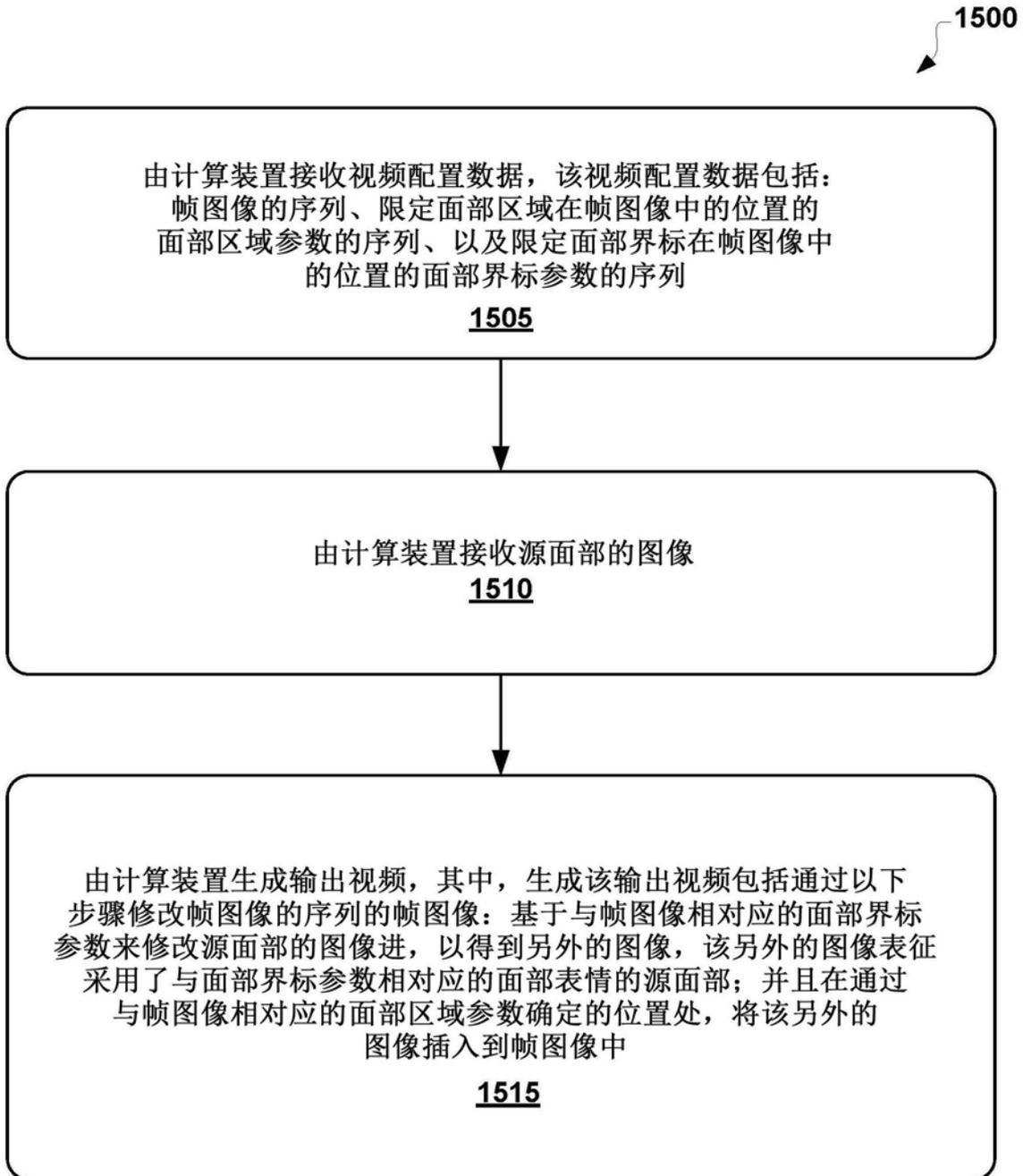


图15

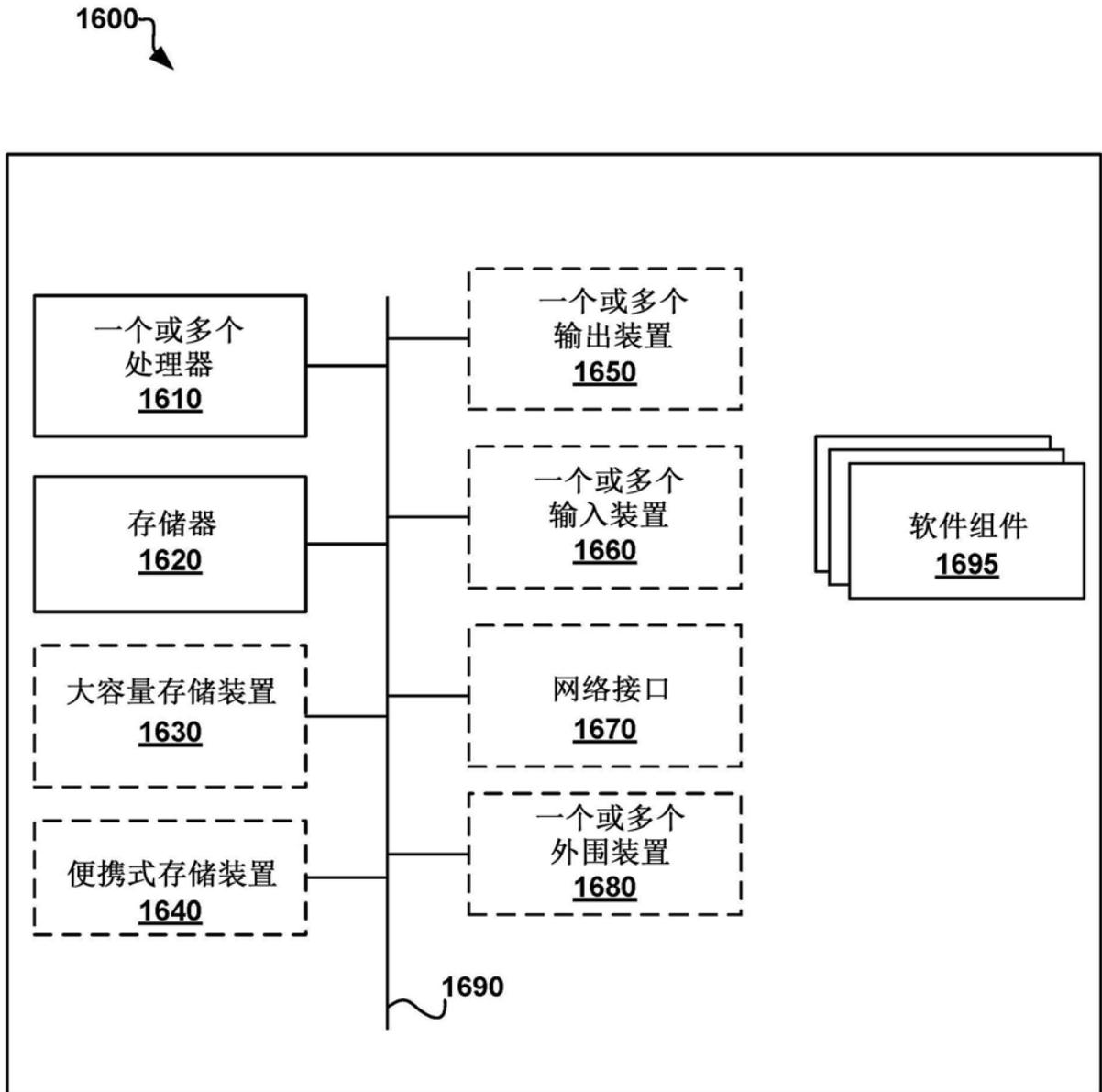


图16