

基于增强现实和人脸姿态估计的虚拟试戴技术^①



黄奕棋¹, 黄起豹², 杨民强¹

¹(兰州大学 信息科学与工程学院, 兰州 730000)

²(凯迪雷拉大学 经济管理学院, 碧瑶 2600)

通信作者: 杨民强, E-mail: yangmq@lzu.edu.cn

摘要: 得益于增强现实技术, 虚拟信息能够在真实世界中予以体现和融合, 使得脱离实体的虚拟信息在现实世界中的应用场景不断泛化. 基于此, 本文提出了一种高效实时的虚拟试戴技术, 可运用于多种实际场景中. 比如在电商场景下, 用户购买商品前可在先线上选择相应款式的模型文件进行虚拟试戴, 并根据效果辅助决策. 拟议方法依据人脸姿态参数将模型文件映射至可与实时视频流加合的图形状态, 在特定区域加合后反馈给视频帧中, 最终加合的模型文件能够自适应头部的位置变化. 实验结果表明, 拟议方法在人脸远近位置、图形渲染以及佩戴实时性等方面具有较好的效果.

关键词: 增强现实; 姿态估计; 图像掩码; 映射变换; 虚拟试戴技术

引用格式: 黄奕棋, 黄起豹, 杨民强. 基于增强现实和人脸姿态估计的虚拟试戴技术. 计算机系统应用, 2022, 31(2): 335-341. <http://www.c-s-a.org.cn/1003-3254/8277.html>

Virtual Try-on Technique Based on Augmented Reality and Face Pose Estimation

HUANG Yi-Qi¹, HUANG Qi-Bao², YANG Min-Qiang¹

¹(School of Information Science and Engineering, Lanzhou University, Lanzhou 730000, China)

²(Economics and Management School, University of the Cordilleras, Baguio 2600, Philippines)

Abstract: Thanks to the augmented reality technology, virtual information can be embodied and integrated in the real world, enabling the increasingly wide applications of virtual information detached from the physical world in real-world scenarios. On this basis, this study proposes an efficient real-time virtual try-on technique that can be applied to a variety of practical scenarios. For example, in the e-commerce scenario, users can select the corresponding model file of a certain style online for virtual try-on before purchasing the product and make their decisions upon the virtual try-on results. The proposed method maps the model file to a graphical state that can be added with the real-time video stream on the basis of the face pose parameters. After the addition is performed in a specific region, the result is fed back to the video frame. The final added model file is able to adapt itself to the position change of the head. The experimental results show that the proposed method delivers a good performance in face distance and position, graphics rendering, and real-time wearability.

Key words: augmented reality; pose estimation; image mask; projection transformation; virtual try-on method

增强现实技术的兴起让越来越多的非实体信息能够在真实世界中虚拟化体现, 随着日常生活场景不断信息化、虚拟化 (如网络购物的普及、各类云端体验的兴起), 增强现实技术的应用一直是计算机视觉领域

经久不衰的话题, 因为这些应用可以将虚拟化场景不断深化从而改变人们的一些固有的生活习惯. 例如, 在网络购物时, 用户无法体验商品的适合度, 只能凭借臆测来决定是否下单. 通过虚拟穿戴技术, 用户能够在缺

① 收稿时间: 2021-04-04; 修改时间: 2021-04-29; 采用时间: 2021-05-17; csa 在线出版时间: 2022-01-17

少试戴的情况下对商品有一个更好的感知,并辅助用户进行决策^[1]。

但是由于相关子技术过于分散并缺乏整合,且相关虚拟试戴技术一直未有一个明确的方案路线。此外,部分与拟议的虚拟试戴技术相关联的子技术尚未成熟,或其关注的重点与拟议技术相距甚远,还存在有如面部区域的识别易受姿态与表情的影响、面部姿态的角度表达无法对应于模型文件的观测视角等问题,这些问题都给虚拟试戴技术的研究与发展带来了重大的挑战。但是,放眼于未来,增强现实相关领域的发展、相关子技术的完善、以及技术侧重点的更改是十分必要的。

在人脸检测标注中,Ghenescu等人提供的通用DNN目标检测方法提供了一种鲁棒的人脸检测方法^[2]。本文基于此方法调用集成的OpenCV接口对实时视频流中的人脸检测,但这种方法缺乏对于瞳孔中心的定位。结合Chen等人提供的眼部定位方法^[3],缩小待操作区域以减少误差。此外,文献[4]中提供的带有深度信息的姿态估计方法结合 n 点透视位姿求解法,为本方法得到人脸姿态参数提供了一种良好的思路,将参数运用文献[5]中所述的矩阵关系进行运算,将矩阵转换成相关角度。

对于某些系统^[6],其虚拟试戴技术主要关注于2D图像的3D重建过程,该操作需要大量算力与运行时间,本文简化了重建过程,采用映射方式在图像上进行操作,为实时视频流操作提供了可能。而系统^[7]对于操作区域与模型关键特征的定位操作存在偶然性,文献[3]和文献[8]中提供了一种眼部定位与眼睛检测的方法,本文结合两种方法并通过实时检测的面部坐标差值缩小操作范围,文献[9]中采用HSV颜色空间的方法为本文提供了一个很好的图像分割思路。

结合上述子技术并为解决现有系统的不足,本文尝试提供一种方法路线,结合相关子技术提出一种基于增强现实与面部姿态估计的虚拟试戴技术。此方法主要是一种用于加合实时面部场景与3D模型文件的虚拟试戴技术,利用深度神经网络人脸标志检测器,估计实时场景中的人脸位态信息,返回对应的方位角与旋转角^[5]。在此方法中,相关姿态参数会被用于调整3D模型文件相对于摄像机坐标系的位置,从而将模型文件映射至屏幕坐标系中。同时采取图像分割技术,以68个标志点为基准,定向分割出待加合区域,模型文件还将根据成像坐标系与摄像机坐标系的相对位置进行

自适应。

1 拟议的虚拟试戴方法

图1显示了拟议方法的示意图,该方法主要针对于人脸部分的虚拟佩戴。对于输入的原始实时视频流(real-time video streaming, RVS),使用预先训练的深度神经网络进行人脸检测。检测人脸存在后,通过初始化相关关键点标识面部轮廓,采用Perspective-N-Point方法估计面部的姿态信息,并返回相关角度参数。由相应的数学关系实现角度向elev与azim的转化,将转化后的参数传入view_init方法对3D模型文件进行视角转换。在不同的视角下,利用相关数学原理对矩阵进行归一与消元操作实现3D文件的映射降维。基于面部关键点标注框选眼部区域,基于深度颜色空间设定阈值并构建掩码位图,提取出的细化区域与框选区域进行加合,得到部分加合结果。并依此法对RVS逐帧操作,实现面部虚拟试戴。

1.1 DNN人脸检测与标注

首先对于输入的实时视频流中的每一帧进行检测,使用传统视觉方法中的由Papageorgiou等人提出的Haar特征^[10],以及Viola和Jones提出的Haar-like小波特征和积分图方法进行人脸检测^[11,12],本方法选择直接使用改进后的OpenCV的Haar分类器。基于对静态图的检测方法,在视频中的每一帧均重复相同操作,实现对视频流的人脸检测,并对相应面部区域进行框选,返回置信度。

紧接着对框选区域进行关键点检测,绘制出68个面部关键点并以此为基准对人脸进行对齐,即对于某些具有初始姿态的面部图像,进行相应的仿射变换以返回正对于屏幕坐标轴的对齐图像,仿射关系可表示为:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (1)$$

图像进行仿射变换后,在面部检测器中初始化一个DNN,将对齐后的面部图像输入到该检测器中,并提取layer6中名为final_dense的张量,将其用于面部预测,得到尺寸为136×1的矩阵,将其转换为68×2的矩阵,分别对应于人脸68个关键点的像素坐标,并由这些坐标定向至像素坐标系下对应的人脸位置,进行标注,图2展示了在某一个姿态下的向量与矩阵以及对应的面部标注情况。

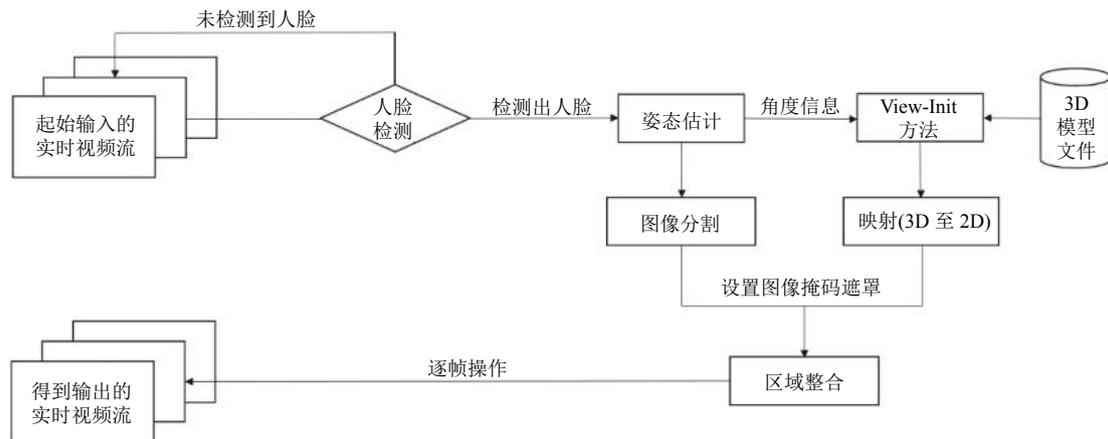


图1 拟议方法的示意图

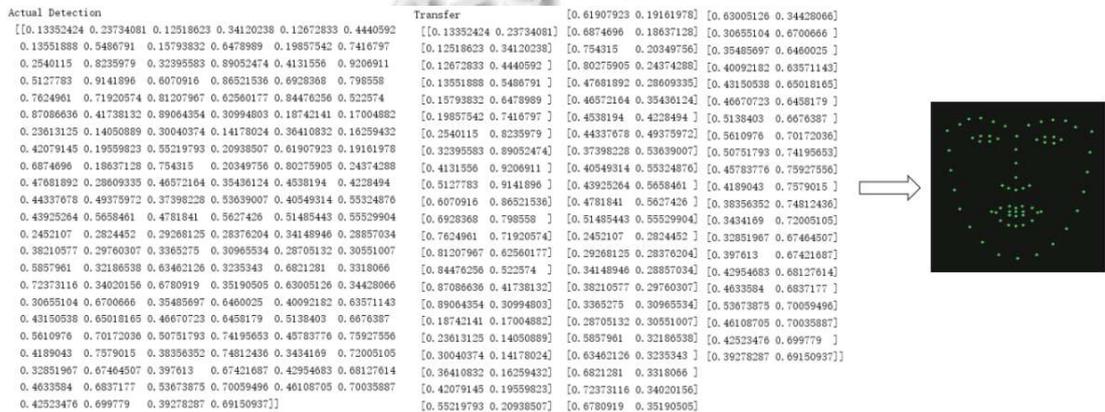


图2 实际检测、转换后的矩阵与相应的面部标记

1.2 基于面部标注的姿态估计

本方法主要是针对于面部的虚拟试戴,而面部朝向的角度信息则是虚拟佩戴适应度与真实性的基础,因此我们将在DNN网络人脸检测器标识的基础上,求解所有面部标志的姿态信息^[2]。面部的姿态估计主要分析3个维度的角度信息,分别是俯仰角Pitch,偏航角Roll和滚转角Yaw三个欧拉角,如图3所示,分别对应于实际场景中头部的上下抬低、左右摇摆和立轴转动,面部的所有姿态均可分解为这3个方面的变化。

基于上述描述,本文选择采用5点透视位姿求解的方法,利用点云配准方法求解摄像机坐标系相对于世界坐标系的平移矩阵(T)以及旋转矩阵(R)^[5]。假设世界坐标系用X、Y、Z表示,像素坐标系用U、V、W表示,由摄像机成像原理可知,绕各自轴旋转一定角度和平移一段距离,能够获得世界物体呈现在以光心为原点的相机坐标下。而针对像平面,其间存在着投影

透射关系,是由三维坐标向像素点的转换,针对相应的数学运算,我们可建立屏幕坐标系下映射的坐标与世界坐标系中实际点的关系,可用于求解矩阵R和矩阵T:

$$s \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} [R|T] \begin{bmatrix} U \\ V \\ W \\ 1 \end{bmatrix} \quad (2)$$

其中,s表示尺度参数(scale parameter), f_x 和 f_y 表示各自方向上的焦距, c_x 和 c_y 表示光学中心,x和y为图像上各自点的横纵坐标。所述的5点透视位姿求解的方法,只能得到T矩阵和R矩阵,而本文设定得到的姿态信息应由图2所示的3个欧拉角表示,因此对于得到R矩阵还需进行相应计算得到最终的3个欧拉角。T矩阵作为人脸平移量的表示,直接加合至3D模型文件矩阵中,表示平移量。旋转矩阵到欧拉角的关系如式(3):

$$R = \begin{bmatrix} r_{00} & r_{01} & r_{02} \\ r_{10} & r_{11} & r_{12} \\ r_{20} & r_{21} & r_{22} \end{bmatrix} = \begin{bmatrix} \cos\phi\sin\gamma & -\cos\phi\sin\gamma & \sin\phi \\ \cos\phi\sin\gamma + \sin\phi\sin\phi\cos\gamma & \cos\phi\sin\gamma + \sin\phi\sin\phi\cos\gamma & -\sin\phi\sin\phi \\ \sin\phi\sin\phi - \cos\phi\sin\phi\cos\gamma & \sin\phi\cos\gamma + \cos\phi\sin\phi\sin\gamma & \cos\phi\sin\phi \end{bmatrix}$$

$$\Rightarrow \begin{cases} \varphi = \text{atan}(-r_{12}, r_{22}) \\ \phi = \text{atan}(r_{02}, \sqrt{r_{12}^2 + r_{22}^2}) \\ \gamma = \text{atan}(-r_{01}, r_{00}) \end{cases} \quad (3)$$

其中, φ, ϕ, γ 分别对应于 Pitch, Roll 和 Yaw. 依据上述求得的欧拉角, 可将面部姿态的关系数值化, 并依此将最终经姿态变换后的人脸标注映射至屏幕坐标系, 得到对应的 2D 成像, 如图 4 所示, 其中相机坐标系的 z 轴指向相面的中心, 也称为光心.

依据上述姿态估计方法, 任意世界坐标系下的人脸姿态都将被准确估计, 即每个人脸可检测范围内的朝向信息都将被转换为相应的欧拉角信息, 所有人脸的姿态信息将被用于标注该朝向的人脸关键点, 并分别在 3D 空间中与 2D 平面上呈现.

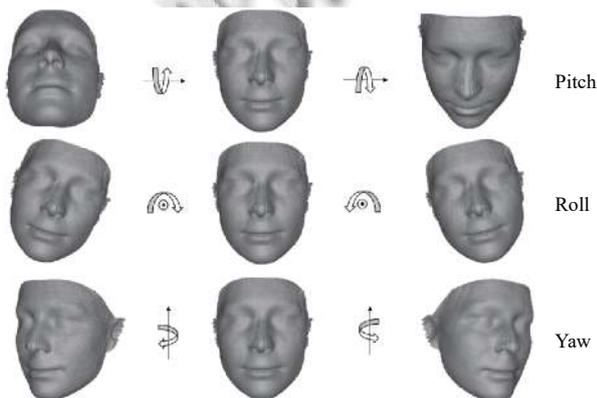


图 3 3 种基础的面部姿态

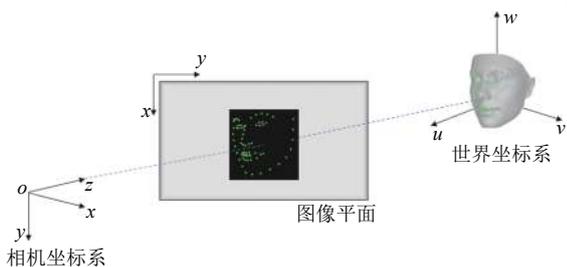


图 4 从世界坐标到相机坐标的转换

1.3 转换参数以设定视角

计算人脸姿态信息的目的是得出人脸在该姿态下所对应的 3D 模型物体的视角信息, 因此, 本步骤主要聚焦于将欧拉角向视角信息转化. 由上述步骤可知, 人脸姿态信息最终体现为俯仰角 Pitch, 偏航角 Roll 和滚转角 Yaw 三个欧拉角. 我们假设以人脸鼻尖作为原点,

垂直纸面向外为 u 轴, 竖直向上为 w 轴, 垂直 $u-o-w$ 平面向右为 v 轴, 如图 4 中的世界坐标系, 我们易知俯仰角为人脸姿态与 v 轴的夹角, 偏航角为人脸姿态绕 u 轴转动, 滚转角则为姿态信息与 w 轴分量的夹角.

然而, 对于 3D 模型文件, 其方向信息主要由观测者的视角进行体现, 分别对应于 view_init 方法中的 elev 参数 (方位视角) 和 azim 参数 (海拔视角). 假设初始视角为 3D 模型文件的正视图, 图 5(a) 以眼镜 3D 模型为例展现了视角变化过程. 在不同的视角下, 均有相对应的面部姿态, 图 5(b) 中的人脸分别对应于图 5(a) 中的 5 个视角.



图 5 多姿态眼镜及与眼镜相对应的面部姿态

通过大量试验, 我们得出了在不同姿态下, 3D 模型文件的视角与人脸姿态欧拉角的关联关系, 式 (5)–式 (9) 展现出了部分视角与欧拉角的等价关系, 其中, Yaw 的值均为 0:

$$\begin{cases} azim = -90 \\ elev = 90 \end{cases} \Leftrightarrow \begin{cases} roll = 0 \\ pitch = 0 \end{cases} \quad (4)$$

$$\begin{cases} azim = -45 \\ elev = 90 \end{cases} \Leftrightarrow \begin{cases} roll = 45 \\ pitch = 0 \end{cases} \quad (5)$$

$$\begin{cases} azim = -135 \\ elev = 90 \end{cases} \Leftrightarrow \begin{cases} roll = -45 \\ pitch = 0 \end{cases} \quad (6)$$

$$\begin{cases} azim = -90 \\ elev = 45 \end{cases} \Leftrightarrow \begin{cases} roll = 0 \\ pitch = 45 \end{cases} \quad (7)$$

$$\begin{cases} azim = -90 \\ elev = 135 \end{cases} \Leftrightarrow \begin{cases} roll = 0 \\ pitch = -45 \end{cases} \quad (8)$$

由此类等价关系, 我们可推得 azim 与 elev 对应

于 Roll 与 Pitch 的等价关系. 对于 Yaw, 它是相对于 w 轴的变化, 即为视角上的远近关系, 这类关系并非由视角与滚转角直接对应.

1.4 映射变换

本步骤是为了将 3D 模型文件通过映射变换, 转换为 2D 图像渲染的形式. 本文采取投影方法进行映射, 此处成像面为一个二维平面, 此处不考虑 3D 模型向规范视域体 (canonical view volume) 的投影. 相较于 2D 平面图形, 3D 模型包含有深度信息, 在投影的过程中, 原本的 3D 物体会失去一些角度和距离信息, 只会保留原物体中的一些属性.

在常见的相机模型中, 光线将经由三维物体的各个点通过小孔模型投影在像平面上, 假设世界坐标系中存在有一个三维物体, 其某个面上的顶点坐标为 (x, y, z) , 若要将其映射至二维平面, 则该点在投影时将对应着一个投影矩阵, 此投影矩阵通常为这个顶点向量提供了一个点乘运算, 从而实现 3D 齐次坐标向 2D 齐次坐标转换^[5].

1.5 眼部区域框选

3D 模型文件经过映射变换后, 在像素平面上得到了对应文件的二维图像形式. 为了确保虚拟试戴的准确性, 我们需要根据所操作模型的特征对操作区域进行缩小, 即只对输入帧的局部进行加合操作.

以眼镜模型为例, 本方法会将操作区域缩小至眼睛区域^[3,8], 上至眉毛, 下至眼袋, 并以左右眼角作为两侧边界, 如图 6 所示. 基于第 1.1 节中得到的面部关键点标注, 依据标注点与面部区域的对应关系选择出所要规划边界的关键点. 对于眼部区域而言, 将起始标注记为 1 (对应面部左侧太阳穴), 两眉最高点 (对应于标注 19 和 25) 为上边界, 左右眼角外侧 (对应于标注 37 和 46) 与左右眉毛边缘点内侧 (对应于标注 18 和 27) 竖线作为左右边界, 过鼻梁位置 (对应于标注 29) 与脸颊外侧关键点 (对应于标注 2) 连线作为下边界, 各边界构成一个类矩形区域. 对于框选出的局部区域, 计算该区域的长度和宽度作为模型文件的尺寸, 使用 OpenCV 中的 `resize` 方法, 调整模型文件的相对大小, 以自适应人脸在远近移动时带来的眼间距宽度的绝对数值变化.



图 6 矩形眼部区域与对应配戴眼镜效果

在选取局部图像后, 降维后的模型图片定向至目标区域, 这种做法不仅降低了加合过程中的复杂程度, 还消除了无关背景对于虚拟试戴区域定位的影响, 可在一定程度上提高佩戴的准确性.

1.6 基于深度颜色空间的掩码生成

经由前述阶段对于人脸视频流与 3D 模型文件的操作后, 本阶段将在此基础上对截取后的局部图像与输入的人脸视频流进行加合处理. 然而, 对于映射降维后的 2D 图像, 其大体呈现为以一定视角成像的图形嵌入白色背景中, 直接加合会导致 2D 图像所在图层将覆盖人脸对应区域. 因此, 本文采用生成图像遮罩层的方法, 基于 HSV 颜色空间 (色调 H , 饱和度 S 以及明度 V) 计算白色背景外的模型图形各像素点的 HSV 值, 并绘制如图 7 所示的阈值曲线.

此外, 根据图 7 所示的阈值曲线, 我们发现样例模型文件各点的 HSV 颜色阈值主要集中在 $[100\ 205\ 70]$ – $[105\ 215\ 170]$ 之间, 由此界定上下限并生成掩码矩阵. 由于我们可将图像视为一系列数组, 这些数组的值即为像素值, 并对应着不同的颜色. 对于重叠部分, 我们期望映射后的模型图形能够覆盖对应部位而其余面部区域保持不变, 因此我们利用掩码界定区域将重叠部分的面部像素值设置为 1, 其余保持不变. 在对相应像素值进行修改后, 将视频帧中人脸区域矩阵与模型 2D 映射图像进行点积运算, 最终得到局部区域的模型佩戴结果, 由此结果反馈至全局帧区域, 以得到增强现实下的虚拟试戴效果.

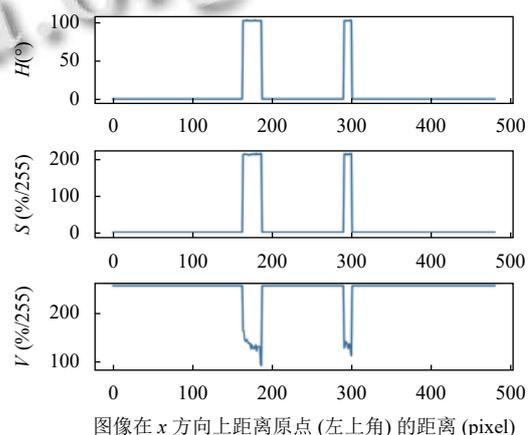


图 7 HSV 阈值曲线

2 现存方法比较

文献 [13] 中提出的人脸检测方法具有较大的局

限性,其基于人脸肤色的检测方法在背景色与肤色十分接近的情况下,会产生较大的偏差,因此该检测方法对于背景色有较为严格的要求.此外,该方法中设定的眼间距是基于经验确定,这种确定方法并不具有鲁棒性,对于某些特殊人群该间距可能并不适合.本文提出的人脸检测方法,是基于面部关键点进行区分,并不受背景色与不同人群的不同特征影响,能够根据返回的眼部关键点坐标进行相应的计算,并自适应于不同特征的人脸.对于文献[13]中的结果而言,其最终佩戴在人脸面部的眼镜,其镜片颜色将遮挡眼部区域,导致佩戴的不真实性,针对此类问题,本文提出了图像掩码设定,利用HSV阈值设定模型文件的框架范围,将眼部区域取代镜片内的白色区域,以此增加虚拟佩戴的真实性.

与文献[13]类似,文献[7]中所提出的试戴方案,仅测量左、右眼角与鼻梁3个点的位置,这种定位方法基于人眼对称的先验知识,具有一定的偶然性.对于某些具有特殊眼部特征的人群,可能无法达到与文献[7]中所述的定位效果.而本文中采用多关键点的人脸检测,能够准确识别眼眶、眼角等关键位置,进而为眼部区域的框选提供了可能.不可否认的是,文献[7]中使用的等腰直角三角形拟合的方式对于眼镜之于对称眼部的佩戴有较为良好的效果.

就定位准确性与模型加载的真实性而言,文献[9]很好地完成此类要求,其基于大量无眼镜的人脸数据集进行训练,并使用ROI提取方法,使得其定位效果良好.而且利用2D图像重建令其眼镜模型具有较强的真实性并能较好地解决因为成像角度而带来的遮挡问题,但是人脸重建过程对于具体实现而言具有较大的难度^[14],且算法通常会存在较大的时延,而此数值已远超出人眼所能适应延迟的阈值,因此在实时场景下并不适用.本文所提出的技术方案简化了重建过程,采用图像映射的方法实现低维度的图像加合,一定程度上减轻了实时计算时可能因为重建过程而带来的计算负担.

3 试验结果

3.1 头部姿态估计结果

本文拟议方法的操作对象为现实场景下包含人脸的实时视频流,该视频流中必然包含有多姿态、多表情的面部帧.因此,为了贴近实际应用场景,本次实验选取了20名受试者,受试者任意转动头部角度并改变

面部表情,直接调用本地摄像头运行本算法.我们通过姿态估计矩形与面部检测点的稳定程度以及姿态估计准确度来判断头部姿态估计效果的好坏,若估计矩形具有明显抖动,则认为本次测试效果不佳;若姿态估计有明显偏差,亦认为本次测试效果不佳.图8表示其中一位受试者的头部姿态估计结果.经过多次试验,本方法中所使用到的姿态估计算法具有较强的鲁棒性.



图8 头部姿态估计结果

3.2 试戴效果

本方法能实现在实时场景下利用增强现实技术将选择的模型文件定向添加至面部区域,该方法具有鲁棒性和可扩展性.模型文件可选择眼镜、口罩、贴纸等面部修饰品,修改对应佩戴区域的参数即可得到不同类型的模型文件向面部区域的加合,因此本方法具有较强的可扩展性.以眼镜模型文件为例,图9展示了受试者在正脸姿态下的佩戴情况.从图中我们可以看到,受试者佩戴口罩,本方法仍能检测出人脸区域并向眼部区域添加模型文件,对于不同受试者,本方法均可以针对受试者的面部区域虚拟佩戴模型文件.此外,对于被试者的远近移动,本方法能够自适应眼镜模型大小,以贴合受试者的眼部区域,因此本方法具有较强的鲁棒性,图9展示了本方法的作用下最终的虚拟试戴结果.



图9 最终的佩戴效果

4 结论与未来工作

本文提出了一种增强现实技术的虚拟试戴方法,主要针对实时场景下的人脸对应区域,在摆脱实体的情况下确保用户的整体感知,通过对虚拟信息的加载实现相关区域的模型文件试戴.

为了确保准确性,我们基于先前的姿态估计技术进行相应地修改以适应本方法的目标,依据相关参数变换以及模型文件的映射成像实现虚拟佩戴技术,本

方法整合并改进现有子技术,较为系统地呈现了一个虚拟试戴方案.未来本方法还可将进行进一步改良,通过载入更完善的纹理与材质信息使得本技术具有更强的真实性,并运用到电子商务、云端体验等新兴领域.

参考文献

- 1 Yuan ML, Khan IR, Farbiz F, *et al.* A mixed reality virtual clothes try-on system. *IEEE Transactions on Multimedia*, 2013, 15(8): 1958–1968. [doi: [10.1109/TMM.2013.2280560](https://doi.org/10.1109/TMM.2013.2280560)]
- 2 Ghenescu V, Mihaescu RE, Carata SV, *et al.* Face detection and recognition based on general purpose DNN object detector. *Proceedings of 2018 International Symposium on Electronics and Telecommunications (ISETC)*. Timisoara: IEEE, 2018. 1–4. [doi: [10.1109/ISETC.2018.8583861](https://doi.org/10.1109/ISETC.2018.8583861)]
- 3 Chen XF, Luo X, Liu XY, *et al.* Eyes localization algorithm based on prior MTCNN face detection. *Proceedings of the 2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*. Chongqing: IEEE, 2019. 1763–1767. [doi: [10.1109/ITAIC.2019.8785430](https://doi.org/10.1109/ITAIC.2019.8785430)]
- 4 Kong SG, Mbouna RO. Head pose estimation from a 2D face image using 3D face morphing with depth parameters. *IEEE Transactions on Image Processing*, 2015, 24(6): 1801–1808. [doi: [10.1109/TIP.2015.2405483](https://doi.org/10.1109/TIP.2015.2405483)]
- 5 蔡自兴. 机器人学基础. 2版. 北京: 机械工业出版社, 2015.
- 6 Yuan XY, Tang DF, Liu YB, *et al.* Magic glasses: From 2D to 3D. *IEEE Transactions on Circuits and Systems for Video Technology*, 2017, 27(4): 843–854. [doi: [10.1109/TCSVT.2016.2556439](https://doi.org/10.1109/TCSVT.2016.2556439)]
- 7 Huang WY, Hsieh CH, Yeh JS. Vision-based virtual eyeglasses fitting system. *Proceedings of 2013 IEEE International Symposium on Consumer Electronics (ISCE)*. Hsinchu: IEEE, 2013. 45–46. [doi: [10.1109/ISCE.2013.6570244](https://doi.org/10.1109/ISCE.2013.6570244)]
- 8 Lin KH, Huang JY, Chen JW, *et al.* Real-time eye detection in video streams. *Proceedings of the 2008 Fourth International Conference on Natural Computation*. Jinan: IEEE, 2008. 193–197. [doi: [10.1109/ICNC.2008.278](https://doi.org/10.1109/ICNC.2008.278)]
- 9 Sari Y, Alkaff M, Maulida M. Classification of rice leaf using fuzzy logic and Hue Saturation Value (HSV) to determine fertilizer dosage. *Proceedings of the 2020 5th International Conference on Informatics and Computing (ICIC)*. Gorontalo: IEEE, 2020. 1–6. [doi: [10.1109/ICIC50835.2020.9288585](https://doi.org/10.1109/ICIC50835.2020.9288585)]
- 10 Papageorgiou CP, Oren M, Poggio T. A general framework for object detection. *Proceedings of the 6th International Conference on Computer Vision (IEEE Cat. No. 98CH36271)*. Bombay: IEEE, 1998. 555–562. [doi: [10.1109/ICCV.1998.710772](https://doi.org/10.1109/ICCV.1998.710772)]
- 11 Viola P, Jones M. Rapid object detection using a boosted cascade of simple features. *Proceedings of 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Kauai: IEEE, 2001. I. [doi: [10.1109/CVPR.2001.990517](https://doi.org/10.1109/CVPR.2001.990517)]
- 12 Viola P, Jones MJ. Robust real-time face detection. *International Journal of Computer Vision*, 2004, 57(2): 137–154.
- 13 卢洋, 王世刚, 赵文婷, 等. 基于人脸姿态估计的虚拟眼镜试戴技术. *中国光学*, 2015, 8(4): 582–588.
- 14 Tang DF, Zhang JY, Tang KT, *et al.* Making 3D eyeglasses try-on practical. *Proceedings of 2014 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*. Chengdu: IEEE, 2014. 1–6. [doi: [10.1109/ICMEW.2014.6890545](https://doi.org/10.1109/ICMEW.2014.6890545)]