单幅图像真实感虚拟试戴技术①

杜 瑶, 王兆仲

(北京航空航天大学 图像处理中心, 北京 100191)

摘 要: 提出了一种针对眼镜、帽子等头部饰品的单幅图像真实感虚拟试戴技术, 其关键在于虚拟饰品的三维注册和虚实图像的合成. 首先提出了一种将人脸关键点检测与刚体姿态估计相结合, 求解单幅图像中人脸三维注册信息的算法. 然后阐述了借助像素颜色混合和深度缓冲检测技术解决虚实图像合成中遮挡关系和模型材质问题的方法. 在 AFLW(Annotated Facial Landmarks in the Wild)人脸数据库上对三维注册算法进行了量化测评, 结果表明该算法的精度满足虚拟试戴技术的要求. 在较大角度姿态变化以及部分遮挡条件下的实验结果表明, 提出的虚拟试戴技术快速准确, 试戴效果自然逼真.

关键词: 虚拟试戴; 三维注册; 人脸关键点检测; 头部姿态估计; 虚实图像合成

Real-Like Virtual Fitting for Single Image

DU Yao, WANG Zhao-Zhong

(Image Processing Center, Beihang University, Beijing 100191, China)

Abstract: An image based real-like virtual fitting system focusing on head accessories such as glasses and hat is proposed in this paper. The key techniques are 3D registration and virtual-real synthesis. Firstly, a facial landmark detection and pose estimation based algorithm for capturing registration data is presented. Then the approach using color blending and depth buffering to solve occlusion and model material problem is discussed. The registration algorithm is evaluated on the AFLW(Annotated Facial Landmarks in the Wild) dataset and testified to be precise enough for virtual fitting. Finally, a series of virtual fitting results are showed. The experimental results, which are obtained on the condition of pose variation and part occlusion, indicate that the proposed virtual fitting technique is fast, accurate and real-like.

Key words: virtual fitting; 3D registration; facial landmark detection; head pose estimation; virtual-real synthesis

虚拟试戴技术的目的是使得人们可以在电脑前体验到尽可能真实的试戴效果, 解决网购中无法亲身体验的问题. 目前已有一些眼镜试戴系统, 根据系统的输入方式可分为视频和图像两类, 视频类在增强现实技术(Augmented Reality)的基础上发展而来, 例如DITTO[1]和[2], 它们的问题在于试戴者必须配备摄像头, 且拍摄需从正脸开始作为初始化, 不够便捷. 图像类以图像编辑(Image Editing)相关技术为基础, 根据眼镜数据的来源又可分两类, 一类如[3]和[4], 它们的问题在于眼镜是从二维图像中通过抠图得来的, 角度单一, 限制了试戴中人脸的姿态, 且由于拍摄条件如光线的差异, 使得合成后的图像真实感较差. 另一类如欧诺虚拟配镜系统[5]、SmartBuyGlasses[6], 它们虽然是用三维饰品模型与图像进行融合的, 但只允许用户使用正面图像或需要手动调整角度和尺寸. HARMONY EYECARE[7]和[8]成功解决了以上两类问题, 但在镜片材质的处理方面效果较差, 同样导致缺乏真实感的问题.

从以上分析可知, 一套具有真实感的虚拟试戴系统需要解决两个难题: 第一, 求取虚拟模型的三维注册信息; 第二, 解决虚实图像合成的相关问题. 本文提出的虚拟试戴技术根据单张图像人脸关键点的检测结果, 应用刚体姿态估计算法计算出眼镜、帽子等头

① 收稿时间:2014-08-01;收到修改稿时间:2014-08-28

部饰品模型在三维注册中所需的旋转矩阵和平移向量，并且利用像素颜色混合和深度缓冲检测技术成功地解决了虚实图像合成中遮挡关系和镜片材质的问题，流程图如图 1 所示，各部分具体原理将在以下各节中详细阐述.

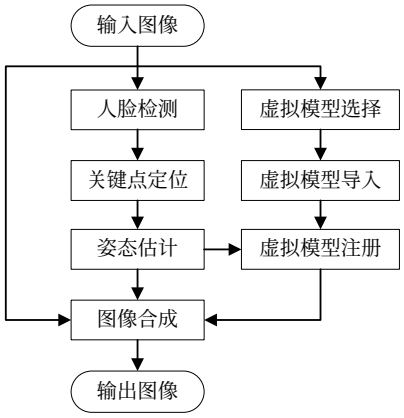

图 1　虚拟试戴系统流程图

1　虚拟模型三维注册

三维注册是指通过计算机图形学分析，计算出物体在三维空间中的准确坐标，据此将计算机生成的虚拟物体绑定拼接到真实的三维环境中去，以达到虚拟物体和真实环境的准确、具有真实感的融合. 三维注册技术是增强现实中最重要也是难度最大的一个环节，它的准确程度直接决定虚实图像融合的真实度[9]. 本文利用 Flandmark(Facial Landmark Detection)算法[10]的检测结果，采用 POSIT(Pose from Orthography and Scaling with Iterations)算法[11]求取虚拟模型在三维空间中的旋转矩阵和平移向量. 三维注册过程如图 2 所示，图中以左半边脸为例，头部模型为经过简化处理的统计标准模型[12]，模型上各关键点的三维坐标是先验知识，眼镜模型在准备阶段调整为与标准模型相吻合的坐标和尺度.

1.1　人脸关键点检测

本文中人脸关键点检测采用由 M. Uricar 等人在 2012 年提出以 DPM(Deformable Part Model)[13]为基础的 Flandmark 算法[10]. 该算法以 Viola 和 Jones 在 2001 年[14]提出的人脸检测算法输出的人脸区域作为输入，输出为人脸上的中心点、两眼内、外眼角、鼻尖和两嘴角共八个关键点位置，如图 3(a)所示，图中的方形框为人脸检测器检测出的人脸区域，用圆点标注的点分别为检测出的各关键点.

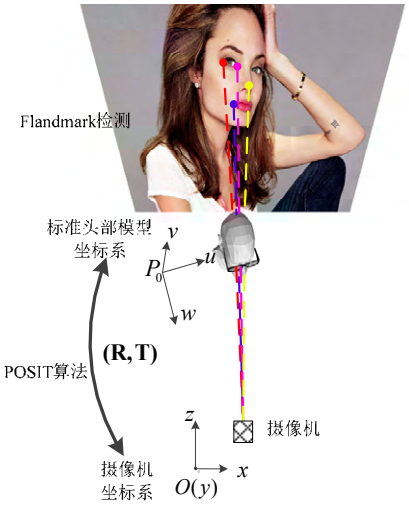

图 2　完成虚拟饰品模型的三维注册过程.

关键点检测模型由外观信息和几何形状两部分构成，用 $G=(V,E)$ 表示几何形状，其中 $V=\{0,\ldots,7\}$ 表示关键点的集合，$E\subset V^2$ 表示相邻两关键点之间位置关系的集合. 算法首先将输入的人脸图像大小归一化为 40×40 像素，设归一化后的图像为 I，如图 3(c)所示，图中矩形框为各关键点的搜索区域，任意一组待选关键点 $l^*=(l_0^*,\cdots,l_7^*)$ (如图 3(b))是最佳组合的可能性由一个打分函数表示:

$$S(I,l^*)=\sum_{i\in V}q_i(I,l_i^*)+\sum_{(i,j)\in E}g_{ij}(l_i^*,l_j^*)$$

其中，第一项表示外观信息，衡量每个待选关键点所在位置的局部外观特征与训练所得模型之间的相似度；第二项表示几何形状，衡量相邻关键点之间的几何位置关系与训练所得模型之间的相似度，

$$q_i(I,l_i^*)=\left\langle w_i^q,\Psi_i^q(I,l_i^*)\right\rangle$$

$$g_{ij}(l_i^*,l_j^*)=\left\langle w_{ij}^g,\Psi_{ij}^g(l_i^*,l_j^*)\right\rangle$$

w_i^q，w_{ij}^g 为由 SO-SVM(Structured Output Support Vector Machine)算法[15]根据大量训练数据得出的参数向量；$\Psi_i^q(I,l_i^*)$ 表示局部外观特征描述子，文中选用 LBP(Local Binary Patterns)[16] 金字塔特征，它是由多尺度上对局部像素点进行 LBP 编码构成的；$\Psi_{ij}^g(l_i^*,l_j^*)$ 表示形变损耗，采用平方距离函数[13]，其定义如下:

$$\Psi_{ij}^g(l_i^*,l_j^*)=(dx,dy,dx^2,dy^2)$$

$$(dx,dy)=(x_j,y_j)-(x_i,y_i)$$

遍历搜索框内所有像素点(如图 3(f))得到的关键点集合为 $L=L_0\times\cdots\times L_7$，用图 3(e)表示，图中箭头表示

相邻关键点之间的依附关系. 最终, 算法通过最大化打分函数 $S(I,l)$, 从 L 中求得打分最高的关键点组合作为最终关键点定位结果(如图 3(d)):

$$\tilde{l} = \arg\max_{l\in L} S(I,l)$$

算法假设 $G=(V,E)$ 是无环的, 上述最大化问题用动态规划的方法得以快速解决.

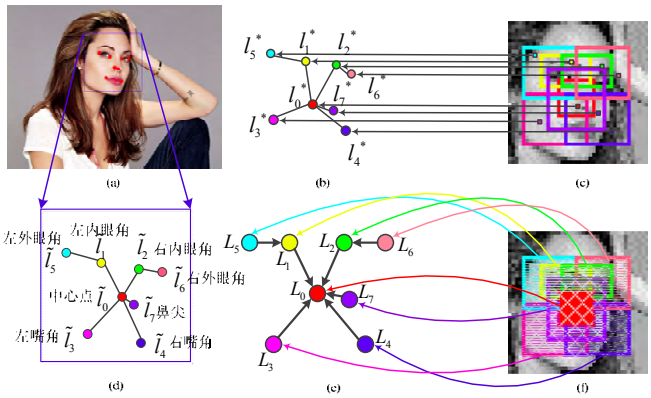

图 3

Flandmark 检测算法利用局部外观和几何形状信息(b), 由打分函数 $S(I,l^*)$ 来衡量任意一组搜索框内的待选关键点 l^* (c)的质量, 遍历所有组合(f)得到集合 L (e), 从中求取使得打分函数 $S(I,l)$ 取得最大值的关键点组合 $\tilde{l}$ (d), 即为最终检测结果(a).

1.2 姿态估计

在 1.1 节中, 已求得人脸上各个关键点 $\tilde{l}=(\tilde{l}_0,\cdots,\tilde{l}_7)$, 在本节中将详细阐述图 2 中利用 POSIT(Pose from Orthography and Scaling with Iterations)算法[11], 根据关键点检测结果 $\tilde{l}_1,\cdots,\tilde{l}_7$ 以及各关键点在三维标准头部模型上的相对位置关系, 求取旋转矩阵 **R** 和平移向量 **t** 的姿态估计方法.

本节中将 $\tilde{l}_7,\cdots,\tilde{l}_1$ 标记为点 $p_0,\cdots,p_6$, 将其在三维头部模型上的对应点标记为 $P_0,\cdots,P_6$, 以鼻尖 P_0 作为参考点建立头部模型坐标系 $(P_0\mu,P_0v,P_0w)$, 则图 2 中关键点在模型和图像上的对应关系可抽象为图 4 所示的透视成像模型, 其中, 摄像机坐标系为 (Ox,Oy,Oz), i,j,k 为坐标轴上的单位向量, 成像平面位于 Oz 轴上焦距 f 处. 由图 2 可知, 我们要求取的是头部模型坐标系 $(P_0\mu,P_0v,P_0w)$ 与摄像机坐标系 (Ox,Oy,Oz) 之间的转换关系 **(R,T)**, 其中, 旋转矩阵 **R** 的每行分别为 **i,j,k** 在坐标系 $(P_0\mu,P_0v,P_0w)$ 下的坐标, 平移向量 **T** 为向量 $\mathbf{OP_0}$. 设 P_i 在 (Ox,Oy,Oz) 坐标系下的坐标为 (X_i,Y_i,Z_i), p_i 在图像平面上的坐标为 (x_i,y_i), 其中 $i=0,\cdots,6$. 考虑透视投影模型以及 $P_iQ_iN_i$和Omp_i 的相似三角形关系, 有

$$\begin{aligned}\mathbf{P_0P_i} &= \mathbf{P_0N_i}+\mathbf{N_iQ_i}+\mathbf{Q_iP_i}\\ &= \frac{Z_0}{f}\mathbf{p_0p_i}+\frac{\mathbf{P_0P_i}\cdot\mathbf{k}}{f}\mathbf{mp_i}+\mathbf{Q_iP_i}\end{aligned}$$

定义 $$\mathbf{I}=\frac{f}{Z_0}\mathbf{i},\mathbf{J}=\frac{f}{Z_0}\mathbf{j},\varepsilon_i=\frac{1}{Z_0}\mathbf{P_0P_i}\bullet\mathbf{k},$$

则有

$$\begin{aligned}\mathbf{P_0P_i}\bullet\mathbf{I} &= x_i(1+\varepsilon_i)-x_0\\ \mathbf{P_0P_i}\bullet\mathbf{J} &= y_i(1+\varepsilon_i)-y_0\end{aligned} \tag{1}$$

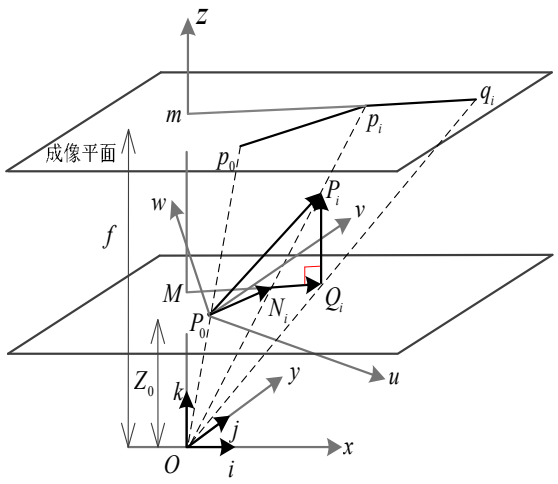

图 4 透视成像模型

其中, $p_0,\cdots,p_6$ 为检测到的人脸关键点, $P_0,\cdots,P_6$ 为各关键点在三维标准头部模型上的对应点.

$\mathbf{P_0P_i}$ 是三维标准头部模型上各关键点 P_i 与鼻尖 P_0 之间的相对位置关系, 可事先从模型上取得; x_i,y_i 是各关键点在图像上的位置信息, 已在 1.1 节中求得. 给定 ε_i, 可从式(1)中求得 **I**和**J**, 单位化得 **i**和**j**, 并利用模长求得 Z_0, **k**可由**i**×**j**求得, 进而由 **k** 求得更加准确的 ε_i, 如此循环直至满足终止条件: 迭代次数 $n\geq 1000$ 或收敛误差 $|\varepsilon_i(n)-\varepsilon_i(n-1)|\leq 1.0\times10^{-5}$. 在结束迭代后可得所求 $\mathbf{R}=(\mathbf{i},\mathbf{j},\mathbf{k})^T$ 和 $\mathbf{T}=(X_0,Y_0,Z_0)^T$, 算法描述如下:

算法 1. POSIT

输入: 由模型上关键点之间的位置关系 $\mathbf{P_0P_i}$ 求得的广义逆矩阵 $\mathbf{B}=((\mathbf{P_0P_1},\cdots,\mathbf{P_0P_6})^T)^+$, 图像上关键点 p_i 位置信息 (x_i,y_i), $i=0,\cdots,6$; 焦距 f.

输出: 旋转矩阵 **R**, 平移向量 **T**.

1: $\varepsilon_i(0)=0(i=0,\cdots,6),n=1;$

2 $\mathbf{x}(n)=(x_1(1+\varepsilon_1(n-1))-x_0,\cdots,x_6(1+\varepsilon_6(n-1))-x_0)$,

$\mathbf{y}(n)=(y_1(1+\varepsilon_1(n-1))-y_0,\cdots,y_6(1+\varepsilon_6(n-1))-y_0)$;

3: $\mathbf{I}=\mathbf{B}\cdot\mathbf{x},\mathbf{J}=\mathbf{B}\cdot\mathbf{y}$;

4: $s_1=(\mathbf{I}\bullet\mathbf{I})^{1/2},s_2=(\mathbf{J}\bullet\mathbf{J})^{1/2},s=(s_1+s_2)/2$;

5: $\mathbf{i}=\mathbf{I}/s_1,\mathbf{j}=\mathbf{J}/s_2$;

6: $\mathbf{k}=\mathbf{i}\times\mathbf{j}$, $\mathbf{k}=\mathbf{k}/|\mathbf{k}|$;

7: $Z_0=f/s$;

8: $\varepsilon_i=\mathbf{P_0P_i}\bullet\mathbf{k}/Z_0$;

9: if $n\le 1000$ and $|\varepsilon_i(n)-\varepsilon_i(n-1)|\ge 1.0\times10^{-5}$,
n=n+1;go to 2;

10: else
$X_0=Z_0\cdot x_0/f;Y_0=Z_0\cdot y_0/f;$
$\mathbf{R}=(\mathbf{i},\mathbf{j},\mathbf{k})^T$; $\mathbf{T}=(X_0,Y_0,Z_0)^T$.

在求得 **R**和**T** 后，将它们应用于标准头部模型的视角转换与平移变换，完成虚拟模型的三维注册.

2 虚实图像合成

虚实图像的合成技术决定着用户的真实感，是整个系统最终效果的关键所在. 在本系统中，虚实图像的合成借助 OpenGL 图形程序接口实现，共分为遮挡关系和模型材质两个主要环节，由于两个环节均借助于像素颜色混合和深度缓冲检测检测技术，以下将首先对此技术作简要阐述.

2.1 像素颜色混合与深度缓冲检测

像素颜色混合是指把图像某一像素位置原来的颜色(目标颜色)和将要渲染的颜色(源颜色)通过某种方式混在一起，从而实现透明、半透明等特殊效果的渲染技术[17]. 基本原理是将目标颜色和源颜色的 RGBA 分量分离后，分别乘以各自的混合系数并对应加和，重新合并，其中, A 是指 alpha 分量表示透明度. 具体在[17]中有详细论述.

深度缓冲检测是在三维渲染中为实现被遮挡物的消隐、增强物体立体感、消除三维物体二义性而开发的三维渲染技术[17]. 基本原理为划分出深度缓冲区存储渲染图像每个像素位置的深度值，在渲染过程中，若当前渲染的图元通过了深度检测，则此图元被渲染，并用其深度值替换深度缓冲区中原有值；否则，图元被忽略. 在[17]中亦有详细论述.

2.2 遮挡关系

在真实的三维空间中，眼睛获取的图像是沿视线方向对所有三维物体进行投影的结果，这些三维物体之间存在着相互的遮挡关系，例如，从正面观察时，佩戴眼镜者耳后的镜架部分被耳朵遮挡，是不可见的. 然而，在虚拟试戴的情况下，人脸图像是无深度信息的二维数据，这种自然的遮挡效果不复存在，使得试戴效果不真实，如图 5(a)所示.

针对这一问题，本文在像素颜色混合和深度缓冲检测的基础上，通过渲染一个标准头部模型而处理出真实遮挡关系. 基本流程为：首先渲染距离观察者最远的图像，然后渲染一个三维注册头部模型，并将其 alpha 分量设置为全零，最后渲染饰品模型. 其关键在于：第一，运用像素颜色混合(如 2.1 节所述)技术渲染全透明的头部模型，在不影响原图像显示的条件下建立一个真实的三维场景；第二，通过深度缓冲检测(如 2.1 节所述)消隐被饰品遮挡的图像区域和被头部遮挡的部分饰品. 实验证明，上述遮挡处理算法准确可靠，结果自然真实，如图 5(b)所示.

(a) (b)

图 5 (a)遮挡关系处理前镜架全部显示和(b)处理后被遮挡部分消隐对比图

2.3 模型材质

除遮挡关系外，饰品模型材质的处理方法也决定着最终试戴的真实感，其中以镜片材质最为典型. 真实的眼镜镜片是透明或半透明的反光材质，在光线照射下有反射、折射等效果，但虚拟的镜片模型并不具备这种特性，渲染效果如图 6(a)所示. 针对镜片材质的特点，本文首先采取 2.1 节所述像素颜色混合的方法，结合纹理贴图，根据不同镜片调整 alpha 分量，解决半透明问题. 其次，在虚拟的三维场景中设立光源，模仿图像中的真实场景. 最后，通过合理设置镜片模型的漫反射光和镜面反射光并加入高光效果，最终完成镜片材质的仿真，如图 6(b)所示. 表 1 中列出了本文中使用的四种眼镜模型镜片的参数设置，其渲染效果在下文图 7 中可见. 表 1 的第一列表示各眼镜模型效果图在图 7 中的位置，第二、三列表示漫反射光和镜面光的 RGBA 设置，第四列表示高光参数. 实验结果表明，采用上述方法处理的镜片模型材质与真实镜片效果相一致.

(a) (b)

图 6 镜片材质处理前不透明、无光感(a)和处理后半透明、有光照效果(b)对比图

表 1 眼镜模型镜片参数设置

眼镜效果图在图 7 中位置	漫反射光 RGBA	镜面反射光 RGBA	高光系数
第 1 行第 1 组	(255,255,255,60)	(255,255,255,255)	100
第 2 行第 1 组	(40,40,30,50)	(255,255,255,255)	50
第 2 行第 2 组	(50,50,50,80)	(255,255,255,255)	100
第 3 行第 2 组	(100,100,100,120)	(128,128,128,255)	80

3 实验结果

3.1 算法精度测试

为测试上述姿态估计算法的精度, 本文选用带有关键点位置和姿态 ground truth 数据的 AFLW[18] (Annotated Facial Landmarks in the Wild)数据库进行测试. AFLW 数据库共有 21997 张现实生活中采集的图像, 包含分辨率、背景、光线、被采集者年龄、种族、姿态、表情等的丰富变化, 其中头部姿态变化涵盖转动(沿竖直轴)、俯仰(沿水平轴)、摆动(沿垂直于图像平面的轴线), 每个方向角度变化范围均大于-90°到 90°. 本文随机抽取了其中 1500 张图像作为测试集, 由于算法需要检测到面部各关键点, 因此剔除掉了关键点被遮挡的图像, 最终约为 1200 张. 测试中将 POSIT 算法得出的旋转矩阵转化为与 ground truth 相一致的欧拉角以方便比对结果, 转化方法如下:

(1) 将旋转矩阵转化为四元数

$$q_w=\frac{1}{2}\sqrt{1+r_{11}+r_{22}+r_{33}}$$

$$q_x=\frac{r_{32}-r_{23}}{4q_w},q_y=\frac{r_{13}-r_{31}}{4q_w},q_z=\frac{r_{21}-r_{12}}{4q_w}$$

其中, q_w,q_x,q_y,q_z表示四元数的四个分量, r_{ij}表示旋转矩阵**R**的第i行, 第j列元素;

(2) 将四元数转化为欧拉角

$$\alpha=\text{-}\arctan(\frac{2(q_wq_x+q_yq_z)}{1-2q_x^2-2q_y^2})$$

$$\beta=\arcsin(2(q_wq_y-q_zq_x))$$

$$\gamma=\arctan(\frac{2(q_wq_z+q_xq_y)}{1-2q_y^2-2q_z^2})$$

其中, a, β, γ,分别表示头部俯仰、转动和摆动的角度, q_w, q_x, q_y,q_z 表示四无数的四个分量.

测试数据如表 2 和表 3 所示. 表 2 中统计了头部不同姿态范围内的平均误差, 考虑到本文虚拟试戴的背景, 角度范围最大取; 表 3 中统计了各误差范围内涵盖的图像数量百分比.

表 2 姿态估计平均误差

角度范围	俯仰	转动	摆动
±15°	8.6°	5.3°	8.8°
±30°	8.7°	7.5°	11.2°
±60°	9.3°	8.2°	11.6°

表 3 姿态估计误差统计结果

误差范围	俯仰(%)	转动(%)	摆动(%)
±10°	67.8	71.7	56.1
±15°	83.9	85.9	70.9
±20°	91.8	91.3	81.6

由以上结果可知, 本文提出的 Flandmark 检测与 POSIT 算法相结合的姿态估计算法在虚拟试戴用到的角度范围 –60°到60° 内平均误差约为 10°, 对于俯仰和转动, 90%以上的图像误差范围在 ±20° 之内, 对于误差较大的摆动, 仍有 80%以上的图像误差范围在 ±20° 之内, 可以满足虚拟试戴技术的精度要求.

3.2 虚拟试戴结果

本文实验语言环境为 C++, 人脸检测和 POSIT 算法基于计算机视觉库 OpenCV 实现, 虚拟模型渲染基于 OpenGL 图形程序接口实现. 在此环境下选取了部分较为典型的图像对虚拟试戴的效果进行测试, 实验结果如图 7 所示, 左右两张图片为一组, 左图为从网络上获取的原图像, 右图为虚拟试戴后的效果图. 由图可见, 本文提出的虚拟试戴技术在试戴者头部俯仰(如第一行第一组), 摆动(如第二行第一组), 转动(如第三行第一组), 背景复杂有干扰物(如第一行第二组), 有少量头发遮挡(如第三行第二组)的情况下均能较为准确的实现三维注册, 对于遮挡关系及镜片材质的处理自然逼真. 在图 8 中我们提供了与本文相类似的虚拟试戴技术效果的对比图, 图中第一列为原图像, 第二列为 HARMONY EYECARE 系统[7]的试戴效果, 第三列为本文系统的效果, 分析可知, 本文提出的虚拟试戴技术三维注册精度更高, 试戴真实感更强. 处理速度方面, 在双核 Intel i7-2600 3.4GHz 处理器, 4G 内存的 PC 机上, 各典型尺寸图像的平均消耗时间如表 4 所示, 表中数据以大小 84KB 的 3DS 格式饰品模型为

例，模型载入时间约为 0.04s. 最后一列为从读入图像到合成图像渲染结束总消耗时间，满足图像自动编辑速度的需求，亦可作为实时性增强现实系统中替代人机交互的自动初始化技术方法.

图 7　虚拟试戴前后效果对比图，左右两张图像为一组，分别为原图像和试戴图像，允许试戴者头部俯仰、摆动、转动，背景有干扰物或部分头发遮挡

图 8　与本文相类似的虚拟试戴技术效果对比图

其中，第一列为原图像，第二列为 HARMONY EYECARE 系统[7]的试戴效果图，第三列为本文算法的试戴效果图.

表 4　系统运行平均消耗时间表

图像尺寸	三维注册运算(s)	图形图像渲染(s)	运行总时间(s)
1024 × 768	0.13	0.06	0.27
800 × 600	0.08	0.05	0.21
640 × 480	0.06	0.04	0.17

4　结语

本文提出了一种具备真实感的单幅图像虚拟试戴技术，其中基于 Flandmark 和 POSIT 算法的三维注册信息求取算法为虚拟模型注册提供了较为准确的旋转和平移数据，借助像素颜色混合和深度缓冲检测实现的虚实图像合成技术真实感强. 实验结果表明，该虚拟试戴技术能够满足对眼镜、帽子等头部饰品的试戴要求，实用性强，速度快，效果逼真. 目前，主要误差来源于 POSIT 算法中使用的标准人脸模型与输入图像中的目标人脸之间的差异. 在今后的科研工作中，我们将尝试通过网格迭代形变的方法修正模型，使其尽量贴近目标人脸，以减小误差. 另外，我们还将针对光照不均匀的图像研究光照估计算法并增加阴影效果，提高虚实合成图像的真实度.

参考文献

1 DITTO. http://www.ditto.com/.

2 Deniz O, Castrillon M, Lorenzo J, et al. Computer vision based eyewear selector. Journal of Zhejiang University -SCIENCE C (Computers & Electronics), 2010, 11(2): 79–91

3 李娟.基于特征点定位的虚拟试戴的研究[学位论文].上海:上海交通大学,2011.

4 刘丽余.基于虚拟试戴技术的眼镜销售系统的研究与实现[学位论文].成都:电子科技大学,2011.

5 欧诺虚拟配镜, http://www.ono.com.cn/.

6 SmartBuyGlasses, http://www.smartbuyglasses.cn/3D-Try-On.

7 Eyecare H. http://harmonyeyecare.opticalestore. com.

8 Zhu JE, Hoi SCH, Lyu MR. Real-time non-rigid shape recovery via Active Appearance Models for Augmented Reality. Proc. of 9th European Conference on Computer Vision. Graz, Austria. 2006. 186–197.

9 Azuma RT. A survey of Augmented Reality. Presence: Teleoperators and Virtual Environments, 1997, 6(4): 355–385.

10 Uricar M, Franc V, Hlavac V. Detector of facial landmarks learned by the Structured Output SVM. Proc. of the 7th International Conference on Computer Vision Theory and Applications. Rome, Italy. 2012. 547–556.

11 DeMenthon DF, Davis LS. Model-based object pose in 25 lines of code. International Journal of Computer Vision, 1995, 15(1-2): 123–141.

12 Storer M, Urschler M, Bischof H. R3D-MAM: 3D morphable appearance model for efficient fine head pose estimation from still images. Proc. of the International Conference on Computer Vision Workshops. Kyoto, Japan. 2009. 192–199.

13 Felzenszwalb PF, Girshick RB, McAllester D, Ramanan D. Object detection with discriminatively trained part based models. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2009, 99(1).

14 Viola P, Jones M. Robust real-time object detection. International Journal of Computer Vision, 2004, 57(2): 137–154.

15 sochantaridis I, Joachims T, Hofmann T, Altun Y, Singer Y. Large margin methods for structured and interdependent output variables. Journal of Machine Learning Research, 2005, 6:1453–1484.

16 Heikkila M, Pietikainen M, Schmid C. Description of interest regions with local binary patterns. Pattern Recognition, 2009, 42(3): 425–436.

17 Shreiner D, Sellers G, Kessenich JM, Licea-Kane BM. OpenGL Programming Guide: The Official Guide to Learning OpenGL: 8th Edition. United States: Addison-Wesley Professional, 2013.

18 Koestinger M, Wohlhart P, Roth PM, Bischof H. Annotated facial landmarks in the wild: A large-scale, real-world database for facial landmark localization. First IEEE International Workshop on Benchmarking Facial Image Analysis Technologies, 2011.