Automatic Eyeglasses Replacement for a 3D Virtual Try-on System

Takumi Kobayashi
Keio University
Yokohama, Kanagawa, Japan
takumi_kobayashi@hvrl.ics.keio.ac.jp

Yuta Sugiura
Keio University
Yokohama, Kanagawa, Japan
sugiura@keio.jp

Hideo Saito
Keio University
Yokohama, Kanagawa, Japan
saito@hvrl.ics.keio.ac.jp

Yuji Uema
JINS Inc.
Tokyo, Japan
yuji-uema@jins.com

ABSTRACT

This paper presents a 3D virtual eyeglasses try-on system for practical use. For fitting eyeglasses in a shop, consumers wish to look at themselves in a mirror while trying on various eyeglass styles. However, for people who need to wear eyeglasses for correcting problems with eyesight, it is impossible for them to clearly observe their face in the mirror without wearing eyeglasses. This makes fitting them for new eyeglasses difficult. This research proposes a virtual try-on system that can be used while wearing eyeglasses. We replace the user's eyeglasses in the input video with new eyeglasses virtually. Moreover, a fast and accurate face tracking tool enables our system to automatically display 3D virtual glasses following a user's head motion. Experimental results demonstrate that the proposed method can render virtual glasses naturally while the user is wearing real eyeglasses.

CCS CONCEPTS

• **Human-centered computing** → **Mixed / augmented reality**; • **Computing methodologies** → *Graphics systems and interfaces*; Computer vision tasks.

KEYWORDS

mixed reality, augmented reality, virtual try-on, eyeglasses removal

ACM Reference Format:
Takumi Kobayashi, Yuta Sugiura, Hideo Saito, and Yuji Uema. 2019. Automatic Eyeglasses Replacement for a 3D Virtual Try-on System. In *Proceedings of Augmented Human International Conference 2019 (AH2019).* ACM, New York, NY, USA, Article 4, 4 pages. https://doi.org/10.1145/3311823.3311854

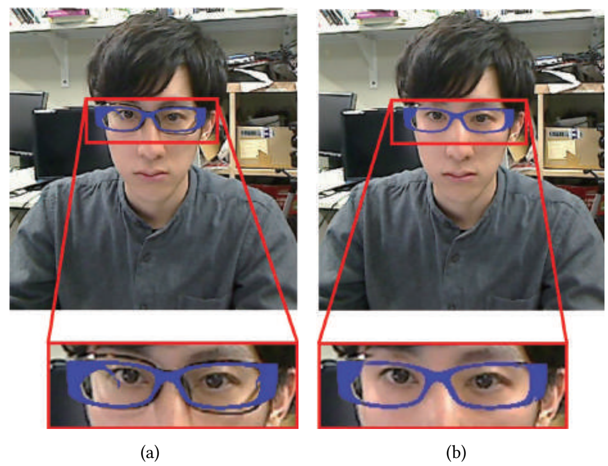


Figure 1: Contribution of our system. (a): An example of previous virtual reality try-on while wearing real eyeglasses, (b): eyeglasses replacement result by our method.

1 INTRODUCTION

Virtual try-on for eyeglasses enables consumers to select a desired item conveniently. In fact, consumers experience many difficulties when physically trying on glasses. One of the biggest complaints is that prescription lenses are not installed in glasses frames for try-on. Many of the consumers trying out glasses have poor eyesight, so they cannot see their appearances in a mirror clearly when removing their glasses and putting on new ones without lenses. They have no choice but to take pictures or to have their friends or family report on their own appearance each time. Besides, picking up and taking off new glasses repeatedly is troublesome and remembering appearances while trying out many different pairs of glasses is difficult. Furthermore, a user's appearance while wearing eyeglasses with and without prescription lenses is different because the eyes size and color seen through lenses change by refraction and reflection. The virtual try-on system can solve all these problems, and our system enables users to try on virtual glasses while wearing their own glasses. Moreover, virtual try-on services are very useful for online shopping because users can try on new eyeglasses anywhere and buy them.

Many studies on virtual glasses try-on have been conducted, and some eyewear brands have already put the system into practical use. Several studies [1, 6] realized a 3D mixed reality system, but their

AH2019, March 11–12, 2019, Reims, France

ACM ISBN 978-1-4503-6547-5/19/03…$15.00
https://doi.org/10.1145/3311823.3311854

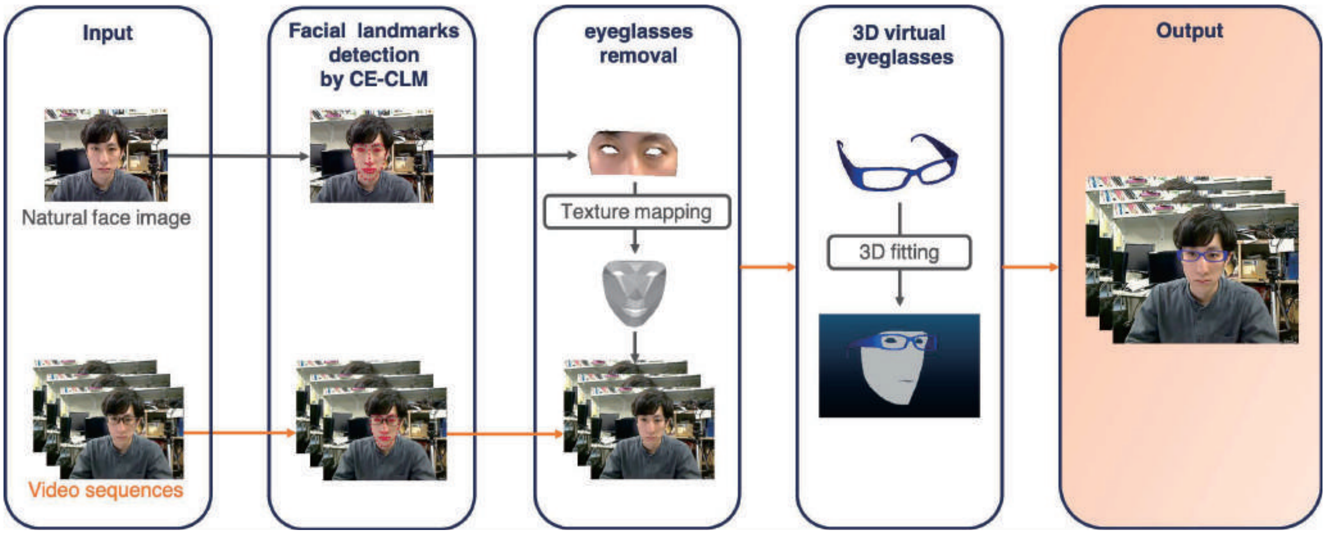


Figure 2: Overviw of the pipeline for the 3D virtual try-on system.

methods need a depth sensor for acquiring 3D maps of the face. In [8], large setup and manual operations are necessary, so it cannot be described as a simple system. [10] also simulates the reflection and refraction effects caused by prescription glasses lenses to to attain the maximum realism while virtually trying on. However, none of these mixed reality systems considered processing in case users wear their own glasses. Figure 1(a) shows an example of such a case. Needless to say, if a person use these virtual try-on systems while wearing glasses, they will find it looks strange because both of the glasses, virtual and physical, will be in the image.

On the other hand, some research has proposed methods to remove the eyeglasses from the face in images. The principal component analysis (PCA)-based method [4, 7] removes eyeglasses from the facial images but it only uses the frontal face images as input. There are also learning-based methods applying Deep Convolution Neural Networks (DCNN) [3] and Generative Adversarial Network (GAN) [5] to remove eyeglasses from the images of the face with eyeglasses. However, these methods are not applied to videos and the reconstructed image is actually not the original face, so they are not suitable for a virtual try-on system.

In this paper, we propose a new 3D virtual eyeglasses try-on system that can be used without taking off the eyeglasses. Our contributions are:

(1) We achieve the proposed system using just a single camera, so that the system can easily be used in various conditions.
(2) The system does not require users to take off their own eyeglasses, and takes only one shot of the frontal face image without glasses to obtain an image of the region around the eyes before the system works.
(3) We employ a Convolutional Experts Constrained Local Model (CE-CLM) algorithm [2, 9] as a facial landmark detection and face tracking method for fast and accurate face tracking, so that we can display 3D virtual glasses following user's head motion in realtime.
(4) We display the eyes area using the images where the user is actually wearing their eyeglasses, so the regions inside the virtual glasses frame looks natural to some extent.

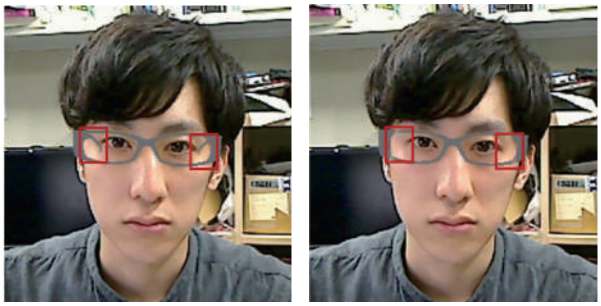
2 THE PROPOSED METHOD

Our virtual try-on system replaces the user's glasses with virtual eyeglasses automatically. Figure 2 presents an overview of the proposed system pipeline. There are two important stages in our proposed method: eyeglasses removal stage and mixed reality stage for displaying 3D virtual glasses. Before these processes, we need to perform facial landmark detection on input video and a natural face image. We can detect 68 facial landmarks in each frame by CE-CLM [9]. The face model includes a 3D representation of facial landmarks, and the method projects them onto the image in order to detect facial landmarks. Internally, the correspondence between 2D and 3D points is calculated, so we can estimate the head pose by solving the PnP problems. Based on the face tracking result, we replace the actual glasses with virtual glasses.

2.1 Eyeglasses Removal

We must remove the eyeglasses from the facial image to realize a virtual try-on system that is usable while wearing glasses. It is possible to detect the area of the eyeglasses in the image using an image segmentation technique and synthesize it with the neighboring pixels or a skin color. However, the boundary of the synthesized area would not be smooth and the resulting image would look as if a transparent glasses have been put on. In addition, since the eyeglasses in the video sequences are very thin and small, it is very difficult to accurately segment the area. In the proposed method, we perform image synthesis using the region around eyes in the facial image taken when not wearing glasses. With this approach, we can remove eyeglasses without having to account for the shape and size of glasses.

First, a user needs to take a natural face snapshot in advance, and this image is used as the texture image for image synthesis.



(a) (b)

Figure 3: Example of occlusion problem. (a): occlusion problem, (b): the result after handling occlusion problem.

Next, we detect facial landmark points in the face image. In the CE-CLM method [9], an index is assigned to each landmark, so we can recognize where a region of interest is in the facial image. Thus, when our virtual try-on system is used while wearing eyeglasses, the area covered by the eyeglasses can be automatically estimated by the facial landmark positions and synthesized with the patch clipped from the natural face image. We detect the 3D facial landmarks in the input video taken while wearing eyeglasses and then map the texture image to the region around eyes in 3D. Then, we project it to 2D using head pose estimation and generate a synthesized image. However, if the texture image is synthesized to the video sequences as it is, the boundary becomes novel and we cannot obtain the desirable result. Therefore, we blend two images using the alpha mask. First, we generate the binary image representing the target area to blend. Then, we apply a smoothing filter to the mask image and normalize each pixel value from zero to one. We then blend two images according to the alpha value α. Each pixel value of a result image are calculated through the following equations:

$$I = \alpha S + (1 - \alpha)D \tag{1}$$

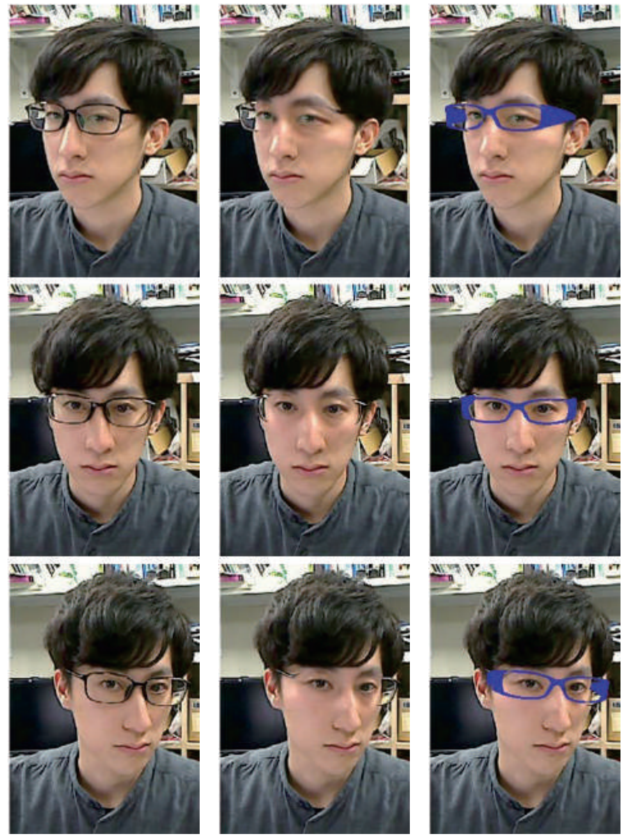
I represents each pixel value of the result image, S is the source image, and D is the destination image.

2.2 Mixed Reality

We display 3D virtual eyeglasses following the user's head motion and realize mixed reality system. Following the result of head pose estimation, we align the 3D eyeglasses model to the face position automatically. However, when projecting the 3D eyeglasses into image, we need to deal with the occlusion problem, because the temples of the virtual eyeglasses should actually be occluded by the user's face. An example of the occlusion problem is shown in Figure 3(a). We also reconstruct a 3D shape of the face in the input video, so we can judge which parts of virtual glasses should not be visible in order to display them correctly. Figure 3(b) shows the image after handling the occlusion problem.

3 EXPERIMENT AND RESULTS

We applied our virtual try-on system to the video of the user's face wearing eyeglasses. In the experiments, we overlaid the 3D virtual glasses model which comprise 8644 vertices and 15588 meshes. Our system runs on a PC with an Intel Core i7 @ 3.60 GHz CPU and 32GB RAM at 18 frames per second. The user moved his head to watch his appearance from various viewpoints and angles. Some images in the video sequences of the experimental results are shown in Figure 4. In comparison with Figure 4(a), Figure 4(b) shows that the area of the eyeglasses covering the face was removed and the images were naturally synthesized. The virtual glasses were properly rendered following the head motion in 3D, as shown in Figure 4(c). Furthermore, for some users who have very poor eyesight, our system would be convenient for them, as they can check their appearance with their prescription glasses.

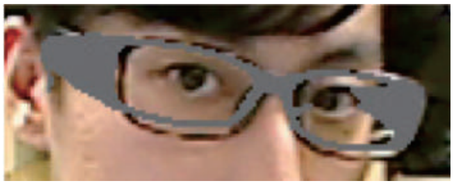


(a) (b) (c)

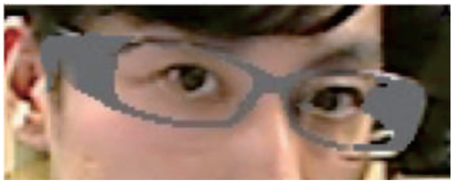
Figure 4: Experimental result. (a): input facial images in video sequences, (b): images after eyeglasses removal, (c): the results of 3D virtual eyeglasses try-on.

4 DISCUSSION AND LIMITATIONS

To demonstrate the effectiveness of the proposed method, we compare the appearance of try-on by our method and the image simply overlaid with virtual glasses. Figure 5(a) shows the appearance of virtual try-on without eyeglasses removal. The shapes of the real glasses and virtual glasses differed, so both of them were visible at the same time. Since the glasses the user actually wears can be seen, it is hard to check whether the shape and color of new glasses



(a)



(b)

Figure 5: Comparison of the virtual try-on result with and without eyeglasses removal. (a): image simply overlaid with virtual glasses, (b): the eyeglasses replacement result image our system produced.

look good on him or not. Figure 5(b) is the eyeglasses replacement result we produce. We confirm that the glasses are overlaid with skin textures and that it resembled actually putting on a different pair of eyeglasses.

On the other hand, our system has several limitations. Since we perform image synthesis based on facial landmark detection and head pose estimation, our method depends on the accuracy of CE-CLM [9]. Therefore, if we fail to track the user's face, we cannot display the virtual glasses in an appropriate position. As shown in Figure 6, the virtual glasses and skin texture patch may be misaligned when the user's head largely rotates or quickly moves. In addition, since the proposed method overlays the warped skin surface image onto the eyeglasses area for removing the eyeglasses and does not detect the area of the eyeglasses, we cannot completely remove the parts of glasses protruding from the face.

5 CONCLUSIONS

We proposed a 3D virtual try-on system, which replace the eyeglasses in videos with virtual eyeglasses. The proposed system is very simple and practical, as this method does not require any special sensors or multiple cameras. For users who have poor eyesight, our system is very useful, since they can utilize it while wearing their own eyeglasses. Experimental results proved that the proposed system displays virtual glasses naturally and realistically, even when the user is wearing glasses.In future work, We need to work on removing the edge of glasses not fully covered with skin textures. Furthermore, the smoothness of the image synthesis result can be improved by adapting to changes in skin color due to lighting and shadows.

ACKNOWLEDGMENTS

This work was supported by JST AIP-PRISM Grant Number JPMJCR18Y2, JAPAN.

Figure 6: Limitations of our system. In the case that the head pose estimation fails, the virtual eyeglasses and skin texture patch will be misaligned.

REFERENCES

[1] Pedro Azevedo, Thiago Oliveira Dos Santos, and Edilson De Aguiar. 2016. An Augmented Reality Virtual Glasses Try-On System. In *2016 XVIII Symposium on Virtual and Augmented Reality (SVR)*. IEEE, Gramado, Brazil, 1–9. https://doi.org/10.1109/SVR.2016.12

[2] Tadas Baltrusaitis, Peter Robinson, and Louis-Philippe Morency. 2013. Constrained Local Neural Fields for Robust Facial Landmark Detection in the Wild. In *2013 IEEE International Conference on Computer Vision Workshops*. IEEE, Sydney, NSW. https://doi.org/10.1109/ICCVW.2013.54

[3] Mao Liang, Yueju Xue, Kunnan Xue, and Aqing Yang. 2017. Deep Convolution Neural Networks for Automatic Eyeglasses Removal. (2017). https://doi.org/10.12783/dtcse/aiea2017/14988

[4] Guo Pei and Su Fei. 2014. Enhanced PCA reconstruction method for eyeglass frame auto-removal. In *2014 4th IEEE International Conference on Network Infrastructure and Digital Content*. IEEE, 322–336. https://doi.org/10.1109/ICNIDC.2014.7000325

[5] Wei Shen and Rujie Liu. 2017. Learning Residual Images for Face Attribute Manipulation. In *Computer Vision and Pattern Recognition*. IEEE.

[6] Difei Tang, Juyong Zhang, Ketan Tang, Lingfeng Xu, and Lu Fang. 2014. Making 3D Eyeglasses Try-on practical. In *2014 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*. IEEE, Chengdu, China, 1–9. https://doi.org/10.1109/ICMEW.2014.6890545

[7] Chenyu Wu, Ce Liu, Heung-Yueng Shum, Ying-Qing Xy, and Zhengyou Zhang. 2004. Automatic eyeglasses removal from face images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26, 3 (March 2004), 359–363. https://doi.org/10.1109/TPAMI.2004.1262319

[8] Miaolong Yuan, Ishtiaq Rasool Khan, Farzam Farbiz, Arthur Niswar, and Zhiyong Huang. 2011. A mixed reality system for virtual glasses try-on. In *the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry*. 363–366. https://doi.org/10.1145/2087756.2087816

[9] Amir Zadeh, Tadas BaltruÅąaitis, and Louis-Philippe Morency. 2017. Convolutional Experts Constrained Local Model for Facial Landmark Detection. In *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. IEEE, Honolulu, HI. https://doi.org/10.1109/CVPRW.2017.256

[10] Qian Zhang, Yu Guo, Pierre-Yves Laffont, Tobias Martin, and Markus Gross. 2017. A Virtual Try-On System for Prescription Eyeglasses. *IEEE Computer Graphics and Applications* 37, 4 (August 2017), 84–93. https://doi.org/10.1109/MCG.2017.3271458