











Augmented Reality Fashion Show Using Personalized 3D Human Models

Shihui Xu¹ , Jingyi Yuan¹ , Xitong Sun¹ , Yuhan Liu¹ , Yuzhao Liu¹ ,
Kelvin Cheng² , Soh Masuko² , and Jiro Tanaka¹ 

¹ Waseda University, Kitakyushu, Japan

{shxu, jingyyuan}@toki.waseda.jp, {sunxitong, liuyuhan-op,
liuyuzhao131}@akane.waseda.jp, jiro@aoni.waseda.jp

² Rakuten Institute of Technology, Rakuten, Inc., Tokyo, Japan

{kelvin.cheng, so.masuko}@rakuten.com

Abstract. When purchasing clothes online, consumers usually have to imagine what they might look like when wearing them in real life surroundings. This can hinder their ability to make appropriate purchasing decisions. In order to alleviate this uncertainty, we propose using augmenting fashion show, a new immersive experience to enable consumers to create personalized 3D models of themselves and have them fitted with a variety of purchasable clothing using Augmented Reality. Furthermore, their models are enhanced with animations (walking and a variety of poses) depending on the real environment. In this way, consumers can see a representation of themselves wearing a variety of clothes, and get a sense of what they would look like in real life surroundings. A prototype of our system was demonstrated, and a preliminary evaluation was conducted to verify our system. We have received positive feedback in terms of effectiveness, assistance and potential application.

Keywords: Augmented Reality · Virtual fashion show · Personalized human model

1 Introduction

A fashion show is an event to showcase the new and upcoming clothing and accessories designs, usually organized by fashion designers each season during a period known as Fashion Week [1]. Typically, there is a runway with light and special effects decorated, where models dressed in clothes designed by designers can walk the catwalk. With the widespread adoption of ready-to-wear collections [2], ready-to-wear fashion shows are becoming more and more diversified, with examples being held in street and supermarket scenes [3], and promote the retail of fashion industry to a certain extent [2].

Despite the popularity of ready-to-wear fashion shows, consumers may still have difficulty knowing whether the clothes are suitable for themselves or not. Clothes are perfectly tailored at fashion shows, and appear at their best on the models. However, they may not be suitable for consumers themselves, because of the differences in body

shapes. In addition, it may be difficult for customers to imagine wearing these clothes in their daily life. Some of the current fashion shows are not limited to the runway or stage, instead the models are static, standing or sitting in an artificially constructed environment. But the constructed environment is still far from customers' everyday lives.

Considering the above, we present an AR fashion show system in this paper, using augmented and mixed reality. Our system enables users to have their own 3D models which are customized for users and can reflect the appearance of the users (i.e. body shapes and facial appearance). Virtual fitting room is provided for users, where users can select the purchasable clothes virtually, and try them on their 3D human model. With dressed 3D models, users can join in fashion show conducted in real-world environment. Fashion show can be adaptively augmented to the daily life scenes where the users are. In other words, according to real environment's scene type (e.g. an office scene, a street scene and a home scene) and spatial layout (e.g. size and shape of floor plane, number and distribution of wall plane), the form of fashion show can vary, e.g. with different postures or walking route.

In this way, we intend to offer ordinary customers immersive experience of attending a ready-to-wear fashion show by themselves (Fig. 1). By using AR fashion show, we provide customers with a preview of themselves wearing those fashion items in their daily lives, which in turn can help them in their decision-making process when purchasing fashion items. In addition, we have conducted a preliminary evaluation to validate our system and received positive feedback in terms of effectiveness, assistance and potential application.

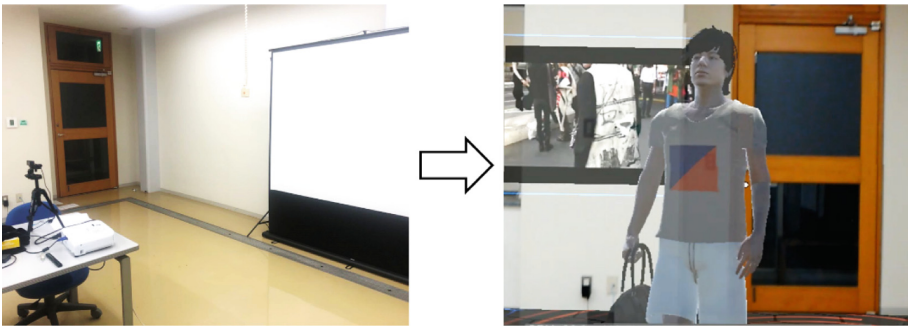


Fig. 1. Our system provides users a way of joining in fashion show using personalized 3D human models of themselves, making the real-world environment (left) an AR fashion show (right).

2 Related Works

We considered related works from three aspects, *fashion show*, *personalized avatar* and *virtual clothes try-on*.

2.1 Fashion Show

With the development of AR technology, many designers have thought about taking AR into fashion show. For example, the autumn/winter 2014 fashion show by Goelia had used AR technology to decorate runway [4]; Xenium Digital and Magnum enabled virtual animals like tiger, panther and cheetah to walk together with models on runway using AR [5]; Three launched mixed reality show where models are companied with special effects during catwalk [6].

In this paper, we consider how to leverage advances in AR and MR to make it possible for ordinary customers to imagine themselves wearing the clothes that they are considering purchasing in the form of a fashion show. In this vein, previous systems in human-computer interaction and computer graphics, such as the motion capture system of Ryuzo Okada et al. [7] and the virtual fashion system of Stephen Gray [8], developed virtual fashion show system that renders a computer graphics clothes to a 3D virtual model. But these works lack of information of real world. In their works, fashion show is isolated to real world, which means users have no ideas about the fashion show in daily life scenes.

2.2 Personalized Avatar

In the field of computer vision, a personalized avatar is a graphical representation of a user. Human avatar needs to be as realistic as possible.

There are several previous works explored generating personalized avatars for users. Earlier work on personalized avatar were mostly based on parameter-driven modelling from images. Generally, 3D body modelling consists two parts: the representation of body joints or skeleton structures, the body shape of the model. Gavrilu [9] consider human body can be built using a 22-DOF model and the class of tapered super-quadratics [10]. Pascal Fua [11] proposed a framework which can form accurate body shape description by a small number of parameters and have an effective use of stereo and silhouette data.

In recent years, 3D scanning technique are used to acquire the shape and pose of a human body. Hasler [12] described an accurately model muscle deformations as a function of pose. Silvia Zuff [13] proposed stitched puppet model which has the realism of 3D body models and the graphical structure of the graphical body model. The body part can be described by low-dimensional state space and the whole parts will be connected by pairwise potentials. Skinned Multi-Person Linear model (SMPL) [14] is a learned model of accurate human shape and pose. Based on SMPL, Federica Bogo proposed SMPLify [15], a way of automatic estimation of 3D human pose and shape from an image.

For our system, we employ the method of Alldieck [16] called video avatar which is based on the parametric body model. Video avatar can infer highly accurate 3D model shapes requiring only an RGB camera. To generate a 3D face model from a single facial image, we use 68 feature key-points used the Dlib library.

2.3 Virtual Clothes Try-on

With the development of computer graphics, virtual clothes models are becoming more realistic. Earlier work on virtual try-on, mostly conducted in computer graphics [17–19]. Anna Hilsmann retextured garment overlay for real-time visualization of garments in a virtual mirror environment [20]. Yamada et al. proposed a method of reshaping the garment image based on human body shapes to make fitting more realistic [21]. However, like many other retexturing approaches they operate only in 2D without using 3D information in any way, which lacked the ability for users to view their virtual self from arbitrary viewpoints.

In recent years, some interactive 3D virtual try-on solutions have been reported. M Yuan et al. presented a mixed reality system for 3D virtual clothes try-on that enables a user to see herself wearing virtual clothes while looking at a mirror display [22]. Whereas, these kinds of system still do not allow users to see the textured garment from arbitrary viewing angles.

Other virtual try-on system based on virtual avatar have been proposed, which enables users to view the virtual garment from different angles. However, most virtual try-on systems provide virtual fitting experiences on a default virtual avatar, rather than one generated from user's own body [23–25]. And some virtual try-on systems provide virtual fitting experiences on a virtual avatar with user's face [22] without personalized user body. The absence of “true fit” may disappoint customers when shopping online.

For our system, we propose to create virtual personalized models for each user, which can reflect their body shape and facial appearance, making the 3D virtual try-on experience for users. Allowing user to view their virtual self with garment model from different viewpoints.

3 AR Fashion Show

Different from typical fashion shows, our purpose is to make a daily life fashion show using users' life-size 3D human models, which means it could adapt to the scenes where users are physically located. As shown in Fig. 1, our system aims to make the real-world environment where the user virtual self become part of the fashion show using AR. Our AR fashion show consists of three parts: scene preprocessing, immersive show and user interaction, which will be introduced in detail below.

3.1 Scene Preprocessing

Scene means the real-world environment where the user is standing. To make a daily life fashion show, we need to do some preprocessing to the daily life scene, including recognition of scene and awareness of spatial layout (Fig. 2).

Recognition of Scene. Currently, scene information is seldom used as input in clothes try-on system. While previous works had shown that the scene type has had a crucial influence on outfits of customers [26, 27]. We propose to recognize the scene where users are and make it an input to AR fashion show system. For example, the scene type could be office, street, supermarket and so on. With different scene as input, AR fashion show is given different theme and form.

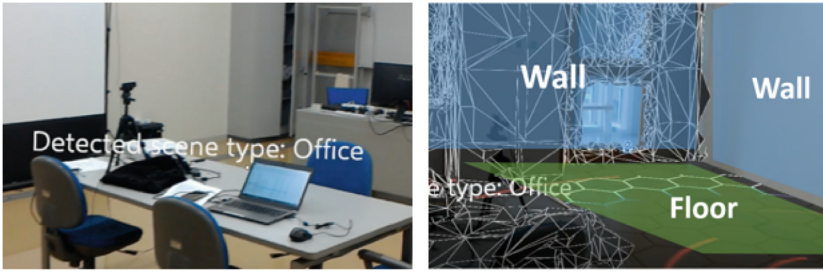


Fig. 2. Scene preprocessing: recognition of scene type (left) and awareness of spatial layout (right).

Awareness of Spatial Layouts. To align virtual items to the real-world, spatial awareness system is needed. Our system scans the real environment through depth cameras and understands the planes in real-world, and then decides the position of users' models in real-world according to the spatial layouts of real-world.

With scene preprocessing, AR fashion show can adapt to multiple scenes in real life, which can facilitate users to have previews of wearing clothes in different daily scenes.

3.2 Immersive Show

Immersive show is provided for users in which users can see a 3D model of themselves wearing the fashion items in AR and get some sense about taking a show. Generally, most people watch fashion shows online or on TV, but few have the chance to watch fashion show in person or take part in a fashion show. We therefore propose immersive show, making the life scene a fashion show. With the personalized 3D human models of users, users can get a sense of taking a show. At the same time, they can also watch the immersive show from various points of view.

The styles of immersive show contain dynamic walking and animated pose.

Dynamic Walking. During the immersive show, 3D human models can dynamically walk in AR environment, giving users a dynamic view of wearing fashion items on their own body models while moving around.

The 3D models will find ways and walk around the real-world space. The walking route is as the dotted line in Fig. 3. Wayfinding of model is based on the result of scene preprocessing, so the ways may vary according to the scene change. From the result of scene preprocessing, the max floor plane (green plane in Fig. 3) is found as the plane where fashion show is superimposed. Next, to get the route, our system will assign four points on the floor plane as the vertexes of rectangle, based on the position and size of floor plane. The points are the quarter points of max floor's length or width. The four sides of defined rectangle will be the route of dynamic walking.

The postures of models' dynamic walking are diversified and associated with the type of scene. Usually, people walk differently in different occasions. To make the virtual 3D human models walk naturally in various real-world scenes and give users a realistic feeling, we construct a library of walking animations for every user model. Every

animation has its suitable application scenes. The walking animations contain catwalk, walking with bag, happy walking, texting while walking etc. According to scene type we get from scene recognition, 3D models will be triggered to walk in different forms. Figure 4 shows some examples of walking postures in an office scene. Figure 5 shows the examples of walking postures in a street scene. In this case, users can better understand what they will look like walking in those daily life scenes.

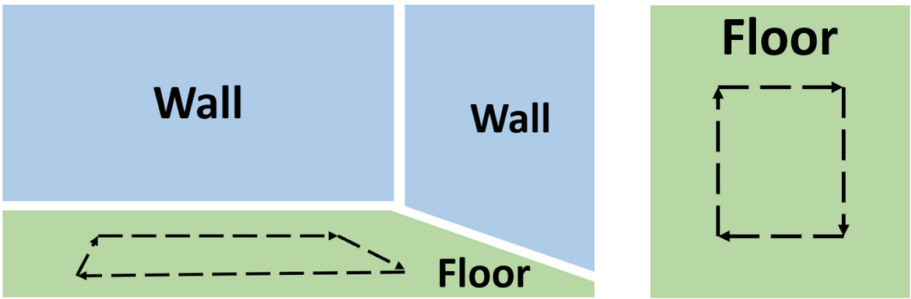


Fig. 3. The route of dynamic walking is illustrated using dotted line with arrows. Route of dynamic walking is defined by four points on the floor plane. (Color figure online)



Fig. 4. Dynamic walking: female catwalk walking, walking with bag and male happy walking.

Animated Pose. To make the fashion show more diverse and varied, we enable 3D human model to make pose during the show. Animated pose is performed in the middle of dynamic walking, i.e. the second vertex of route rectangle, which is called the pose point. When 3D human model walks to the pose point, it will stop to make an animated pose. As the dynamic walking, a library of pose animations is also prepared for each 3D human model. Models poses according to the current scene. For example, as the Fig. 6

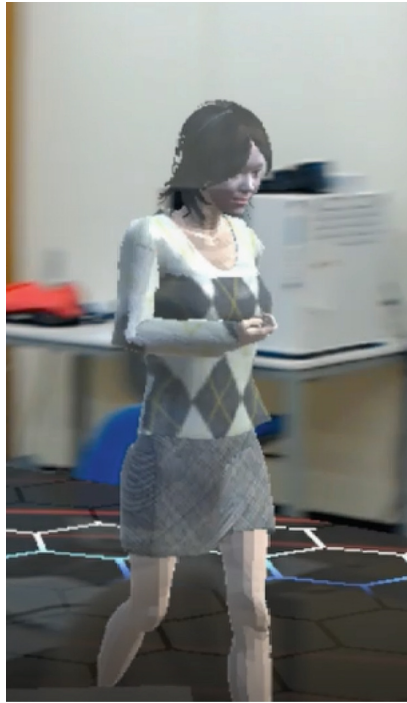


Fig. 5. Text while walking

shows, the animated pose ranges from looking around pose, female stand pose, waving pose, saying hello pose and so on, which is relevant to our daily activities. The scene shown in Fig. 6 is an office scene, so models will do the animated pose related to the office environment. This is designed for achieving a simulation of users' daily actions so that the AR fashion show is more relevant to people's daily life.

3.3 User Interaction

Interactions are available to enhance the engagement of users, including the interactions among multiple users and the interactions with environments.

Show Partner. Show partner is a multi-user interaction. AR fashion show system enables users to invite another user as the partner of show. Show partner is designed to give users some outfit inspirations about how to look good as a couple. Matching styles for couples have been popular in recent fashion industry. Users could be families, friends or lovers, and want to select clothes for matching with each other. But currently, users could not know how they and their partners will look like wearing matching styles clothes. In light of this, show partner is provided for users to let them have a broad understanding of the looks as a couple. Once a user is invited as a partner of another user, the pair of users' models will be assigned as show partners. Show partners will walk together and make paired poses, as shown in Fig. 7.



Fig. 6. Animated pose



Fig. 7. A pair of show partners.

Interactions with Models. Interaction with models directly is also made possible in our system. Users can give a “like” to any model by clicking the “like” button beside the

models. The number of likes received will be shown on the left of every model's like button, as shown in Fig. 8. Users can use hand gestures to interact with their own model as well, including moving or rotating it.



Fig. 8. Like a model

4 Personalized 3D Human Model

Personalized 3D human model is generated for every user so that they can virtually try on clothes and take part in AR fashion show using their own model. The generation of personalized 3D human model consists of the generation of 3D body shape model and 3D face model. As the Fig. 9 shows, we generated the 3D body shape model of user using Alldieck et al's method [16], and generated the 3D face model of users using Deng et al's method [28]. The two parts are then combined to get the complete user's personalized 3D human model.

4.1 3D Body Shape Model

3D body shape model of user is generated from a 2D video in which user rotates 360 degrees in front of a camera. To realize this, we use video from user as the input of the whole system. To get the body model, we followed the method of Thiemo Alldieck [16].

We see human body modeling as part of preprocessing. After getting the video from user, the system will split the video into a series of frames. These frames can be used as

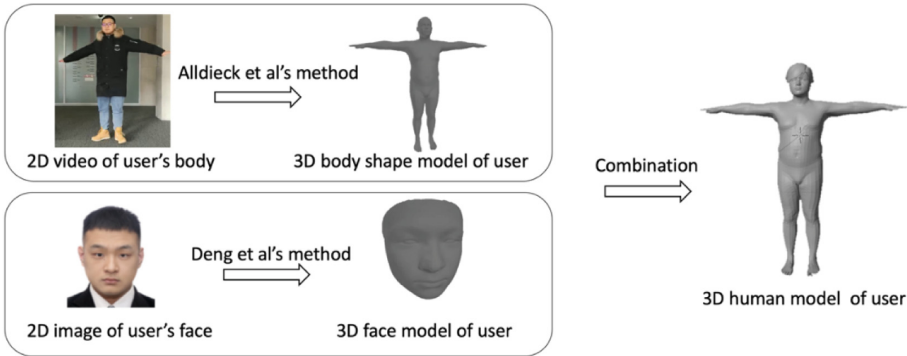


Fig. 9. Process of personalized 3D human model generation

input of preprocessing. Each frame will generate a corresponding mask using a method called OSVOS [29]. This method can tackle the problem of semi-supervised video object segmentation using a fully-convolutional neural network architecture.

Then we need to detect body joints data (such as human body, hand, foot key points) using OpenPose [30]. Pose can be represented by many formats. We choose COCO as pose output format which uses 18 key points to represent body joints. This information will be stored into the json file. And each frame will generate a corresponding json file.

After getting the whole masks and key points files, we can use Alldieck's work to generate 3D human body model by SMPL model. Every SMPL model has two kinds of parameters: pose and shape. We only care about the shape parameter because the system will set every model to T-pose to rig the model. Alldieck's work will calculate shape parameter. Afterwards, the system will generate user's body model as the output.

4.2 3D Human Face Modeling

In this part, we need to generate a 3D face model from a single facial image as the face of the 3D avatar of the user. To realize this, we first need a face image of the user. After that, the facial feature points of the image are required for modeling. To extract the feature points, here we used the Dlib library to get the 68 feature points of the user's face. From these points, we select the 5 points, the left eye, the right eye, the nose, the left mouth corner and the right mouth corner and got their pixel coordinates for the latter use. After getting these 5 points' information, we followed Yu Deng's work to generate the 3D face model of the user. In the process, we used the points information and the face image as the input, and we could get the 3D face model.obj file as the output. In this way, we can generate the 3D human face model of the user.

We also need a hair model which is suitable for the user's hair in order to make our 3D avatar more realistic. We collected many hair models, and collected them as a library. From this library, the most similar hair model compared to the user's haircut will be chosen to generate the 3D avatar.

After having both the 3D body shape model and the 3D face model, we used Blender to combine them together. Finally, the 3D user model will be served as the user's avatar.

5 Virtual Fitting

For virtual fitting, users can select clothes in the form of 2D images from shopping website such as H&M. Then corresponding virtual clothes will be fitted on user’s personalized 3D human model. They can select the garments fitting in a real-life scene and view the virtual garment from various angles. The overall architecture of our system for virtual try-on are shown in Fig. 10.

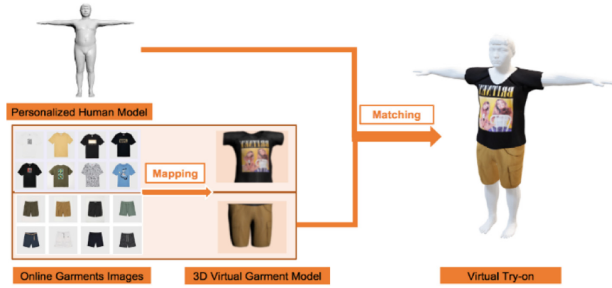


Fig. 10. The overall architecture of the virtual try-on

We gather various clothing design and information from websites, and map 2D Garment images to generated 3D virtual garment templates, then match the clothes to personalized body model of users.

5.1 Garment Model Generation

Our approach collects garment image information from existing shopping websites (e.g. H&M [31], Zara [32]) to create a virtual garment library. Textures are extracted from the online garment image and mapped onto the 3D garment model in 3Ds Max.

Cloth weaver [33] is a blender garment template library which is used as the basis for creating various garment models. We used Cloth Weaver to build several 3D garment model templates for personalized human models. Some 3D clothing templates provided to users are shown in Fig. 11. Such as short T-shirts, pants, skirts and long sleeves.

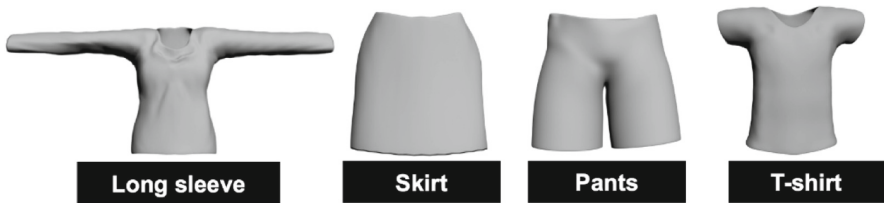


Fig. 11. Some 3D garment templates provided for users

We collect garment images on existing shopping websites such as H&M [31], ZARA [32] ...) and mapped these garment images onto the generated 3D garment model templates in 3Ds Max. Some textured garment models are shown in Fig. 12 below.



Fig. 12. Textured garment models

5.2 Virtual Fitting Room

In virtual fitting, users can select clothes in the form of 2D images from shopping website such as H&M. The corresponding virtual clothes will be fitted on user's personalized model. The users' personalized models wearing the virtual clothing will appear standing on the floor in front of them.

The animation of human body enables users to view the dressed human model in a dynamic and interactive way, from various perspectives. Therefore, users can have a better understanding of whether the garment is suitable or not, while they move. Most of the previous work focused on fitting on a static body model [19, 25]. So far, there is a lack of research exploring virtual try-on with motion. Therefore, we provide dynamic interaction with the virtual human model in virtual fitting room.

6 Implementation

AR fashion show is developed using Unity 3D. We implemented our system on a see-through type head-mounted display, Microsoft HoloLens, with Intel 32-bit CPU, 2 GB RAM, and a $120^\circ \times 120^\circ$ depth camera.

6.1 Scene Type and Spatial Layout

We used the Mixed Reality Capture API [35] provided by Microsoft to capture the photo of real environment. Then the captured photo is processed by using Google cloud vision API [36] to get the labels of photo. The labels usually contain objects such as table, road, room and so on. The depth camera embedded in HoloLens is used to calculate the spatial layout of the real environment.

6.2 Libraries of Animations

To build the libraries of dynamic walking animations and animated pose animations, we have used Maximo [37] from Adobe, which enables to empower various animations

with 3D characters. We classify the animations from Maximo into several sets according to scene usage. In another word, we added the information of the scene to the animations we get from Maximo to ensure the animations will be triggered correctly in different scenes.

7 Preliminary Evaluation

We have conducted a user study to evaluate our system. The evaluation is conducted from the aspects of effectiveness, assistance and potential application.

7.1 Participant

Ten participants were recruited for this evaluation, including 7 males and 3 females aged from 21 to 25. All of the participants have a background in human-computer interaction with a basic knowledge of AR technology.

7.2 Task and Procedure

Each participant was asked to virtually try on clothes by the personalized 3D human model which is generated by our system. Then using the dressed 3D models of themselves, participants can take part in an AR fashion show. During the show, they can watch their models and interact with them.

The study was conducted in an office which was about 10 by 10 m.

After the show, participants were asked to complete a questionnaire to evaluate our system using 5-point Likert scale (1: strongly disagree to 5: strongly agree).

7.3 Results

We have asked 5 questions which covers the effectiveness, assistance and potential application of our system. The results from our questionnaire reveals positive feedback from our users. Table 1 shows the 5 questions and their average scores.

Effectiveness. Questions 1, 2 and 3 are designed to evaluate the effectiveness of our system. The effectiveness testing includes the effectiveness of personalized 3D human model, the effectiveness of augmenting fashion show to real-world scenes and the effectiveness of user interaction.

For question 1, we intend to know whether users can get a better understanding of how the clothes would look on their body by using a virtual avatar. We see an average score of 4, indicating that the personalized 3D human model of our system can reflects the personality of the user. We have gotten feedbacks about the photorealism of personalized 3D human model from participants. Most participants feel that body model part is more similar to the real person than the face model. We thought it might be because people are more familiar with their face than their body shape, so that they could recognize quite small differences between face models and real faces. But participants also referred that similarity of body shape model is more important than face model during the selecting of clothes, so there is no problem if the face model has a few differences with real face.

Table 1. Questions and corresponding scores

	Question	Average score
1	I could get feelings about wearing clothes on my own virtual body by using this system	4
2	I could get feelings about wearing clothes in real life scene by using this system	3.5
3	I feel the user interactions are interesting and make me engage with the show	4
4	I could make better purchasing decisions when buying clothes using this system	4.5
5	I would like to use the system when purchasing clothes in the future	4.5

From question 2, we intend to know whether users can get a better understanding of how the clothes would look like in a real-life scene. The average score for question 2 is higher than neutral value, suggesting that augmenting a fashion show to the real environment is effective for users to imagine the look in their daily life.

Question 3's result shows that the user interaction can increase the amusement and engagement of AR fashion show experience.

Assistance. Question 4 is to evaluate whether our system can benefit customers, leading to better decisions on the purchase of fashion items like clothes. The result is between agree and strongly agree, indicating that our system do indeed help users in general.

Potential Application. Question 5 explored the application prospect of our system. Result shows that our system has a lot of potential.

In general, the average scores of all questions were higher than 3 (neutral), which implies that our system was generally viewed as positive in all aspects. This might verify that our system designs are reasonable and practical.

8 Conclusion and Future Work

In this paper, we present an AR fashion show system which uses personalized 3D human models of users. Our system enables ordinary customers an immersive fashion show experience within their real environment. A preliminary evaluation was conducted to validate our system. It demonstrated that our system is effective and can help customers make better decisions on the purchase of clothes, having potential applications in future.

Since the fashion industry has the features of diversity and is rapidly changing, we need to supplement or change the clothes models in time. In future, we hope to improve the quality and quantity of clothes model in order to increase the universality of our system.

References

1. Fashion show. https://en.wikipedia.org/wiki/Fashion_show. Accessed 30 Jan 2020
2. Ready-to-wear. <https://en.wikipedia.org/wiki/Ready-to-wear>. Accessed 30 Jan 2020
3. Chanel ready-to-wear collection. https://en.wikipedia.org/wiki/Chanel_ready-to-wear_collection. Accessed 30 Jan 2020
4. <https://www.youtube.com/watch?v=YMGNLDcx-mA>. Accessed 30 Jan 2020
5. <https://www.youtube.com/watch?v=4r73e2MjaB4>. Accessed 30 Jan 2020
6. <http://www.threemediacentre.co.uk/news/2019/three-launch-worlds-first-5g-mixed-reality-catwalk-at-london-fashion-week-featuring-lennon-gallagher.aspx>. Accessed 30 Jan 2020
7. Okada, R., Stenger, B., Ike, T., Kondoh, N.: Virtual fashion show using real-time markerless motion capture. In: Narayanan, P.J., Nayar, Shree K., Shum, H.-Y. (eds.) ACCV 2006. LNCS, vol. 3852, pp. 801–810. Springer, Heidelberg (2006). https://doi.org/10.1007/11612704_80
8. Gray, S.: In virtual fashion. *IEEE Spectr.* **35**(2), 18–25 (1998)
9. Gavrilu, D.M., Davis, L.S.: 3-D model-based tracking of humans in action: a multi-view approach. In: Proceedings of CVPR IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 73–80. IEEE (1996)
10. Metaxas, D., Terzopoulos, D.: Shape and nonrigid motion estimation through physics-based synthesis. *IEEE Trans. Pattern Anal. Mach. Intell.* **15**(6), 580–591 (1993)
11. Plankers, R., Fua, P.: Articulated soft objects for video-based body modeling. In: Proceedings Eighth IEEE International Conference on Computer Vision, ICCV 2001, vol. 1, pp. 394–401. IEEE (2001)
12. Hasler, N., Stoll, C., Sunkel, M., Rosenhahn, B., Seidel, H.P.: A statistical model of human pose and body shape. In: Computer Graphics Forum, vol. 28, no. 2, pp. 337–346. Blackwell Publishing Ltd., Oxford (2009)
13. Zuffi, S., Black, M.J.: The stitched puppet: a graphical model of 3D human shape and pose. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3537–3546. IEEE (2015)
14. Loper, M., Mahmood, N., Romero, J., Pons-Moll, G., Black, M.J.: SMPL: a skinned multi-person linear model. *ACM Trans. Graph. (TOG)* **34**(6), 1–16 (2015)
15. Bogu, F., Kanazawa, A., Lassner, C., Gehler, P., Romero, J., Black, M.J.: Keep it SMPL: automatic estimation of 3D human pose and shape from a single image. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9909, pp. 561–578. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46454-1_34
16. Alldieck, T., Magnor, M., Xu, W., Theobalt, C., Pons-Moll, G.: Video based reconstruction of 3D people models. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 8387–8397. IEEE (2018)
17. Han, X., Wu, Z., Wu, Z., Yu, R., Davis, L.S.: VITON: an image-based virtual try-on network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7543–7552. IEEE (2018)
18. Sekine, M., Sugita, K., Perbet, F., Stenger, B., Nishiyama, M.: Virtual fitting by single-shot body shape estimation. In: Proceedings of International Conference on 3D Body Scanning Technologies, pp. 406–413. Citeseer (2014)
19. Decaudin, P., Julius, D., Wither, J., Boissieux, L., Sheffer, A., Cani, M.P.: Virtual garments: a fully geometric approach for clothing design. In: Proceedings of Computer Graphics Forum, vol. 25, no. 3, pp. 625–634. Blackwell Publishing, Inc., Oxford (2006)
20. Hilsmann, A., Eisert, P.: Tracking and retexturing cloth for real-time virtual clothing applications. In: Gagalowicz, A., Philips, W. (eds.) MIRAGE 2009. LNCS, vol. 5496, pp. 94–105. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-01811-4_9

21. Yamada, H., Hirose, M., Kanamori, Y., Mitani, J., Fukui, Y.: Image-based virtual fitting system with garment image reshaping. In: Proceedings of 2014 International Conference on Cyberworlds, pp. 47–54. IEEE (2014)
22. Yuan, M., Khan, I.R., Farbiz, F., Yao, S., Niswar, A., Foo, M.H.: A mixed reality virtual clothes try-on system. *IEEE Trans. Multimed.* **15**(8), 1958–1968 (2013)
23. Meng, Y., Mok, P.Y., Jin, X.: Interactive virtual try-on clothing design systems. *Comput. Aided Des.* **42**(4), 310–321 (2010)
24. Li, R., Zou, K., Xu, X., Li, Y., Li, Z.: Research of interactive 3D virtual fitting room on web environment. In: Proceedings of 2011 Fourth International Symposium on Computational Intelligence and Design, vol. 1, pp. 32–35. IEEE (2011)
25. Warehouse. <https://www.warehouselondon.com/row/homepage>. Accessed 30 Jan 2020
26. Yu, L.F., Yeung, S.K., Terzopoulos, D., Chan, T.F.: DressUp!: outfit synthesis through automatic optimization. *ACM Trans. Graph.* **31**(6), 1–134 (2012)
27. How to dress to make the right first impression in 7 scenarios. <https://www.mentalfloss.com/article/79760/how-dress-make-right-first-impression-7-scenarios>. Accessed 30 Jan 2020
28. Deng, Y., Yang, J., Xu, S., Chen, D., Jia, Y., Tong, X.: Accurate 3D face reconstruction with weakly-supervised learning: from single image to image set. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. IEEE (2019)
29. Caelles, S., Maninis, K.K., Pont-Tuset, J., Leal-Taixé, L., Cremers, D., Van Gool, L.: One-shot video object segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 221–230. IEEE (2017)
30. Cao, Z., Simon, T., Wei, S.E., Sheikh, Y.: Realtime multi-person 2D pose estimation using part affinity fields. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7291–7299. IEEE (2017)
31. Wang, L., Villamil, R., Samarasekera, S., Kumar, R.: Magic mirror: a virtual handbag shopping system. In: Proceedings of 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp. 19–24. IEEE (2012)
32. H&M. <https://www.hm.com>. Accessed 30 Jan 2020
33. ZARA. <https://www.zara.com>. Accessed 30 Jan 2020
34. Cloth-weaver. <https://clothweaver.com>. Accessed 30 Jan 2020
35. Mixed reality capture for developers. <https://docs.microsoft.com/en-us/windows/mixed-reality/mixed-reality-capture-for-developers>. Accessed 30 Jan 2020
36. Cloud Vision documentation. <https://cloud.google.com/vision/docs/?hl=en>. Accessed 30 Jan 2020
37. Maximo. <https://www.mixamo.com>. Accessed 30 Jan 2020