Game Engine Compatible 3D Clothes Modeling from a Single Image

Soyoung Yoon[1], Sojin Yun[1], and In Kyu Park[1][0000-0003-4774-7841]

Department of Information and Communication Engineering, Inha University
Incheon 22212, Korea
thdud679@gmail.com, sj0524sj@gmail.com, pik@inha.ac.kr

Abstract. In this paper, we propose a clothes modeling technique for 3D human synthesis, as a key ingredient of virtual human modeling. Given a single image, 3D clothes model is reconstructed automatically while being compatible with commercial game engines and graphics tools like Unreal and Maya. Among the pre-selected clothes categories proposed by DeepFashion2 dataset, we classify the clothes types to provide the generic 3D model of clothes. UV map is automatically captured to represent the texture of clothes. Experimental results show that the clothes model is generated with reasonable accuracy and compatibility to Unreal game engine.

Keywords: 3D clothes modeling · Single image · Game engine.

1 Introduction

Metaverse is a combination of 'Meta' and 'Universe', which means a virtual world linked to reality. In the existing metaverse applications, characters are created by selecting from a few predefined characters and modifying the shape and appearance using the provided assets. With this method, it is difficult to create a 3D virtual human that reflected the variety of human appearances. Therefore, this work proposes a 3D virtual human clothes reconstruction from one's own single image.

Previously, the work of reconstructing a 3D body model has progressed a lot. Examples of this are [5, 4], which produce 3D surface and texture from images. However, the results of these works are volumetrically modeled with an overall 3D mesh. This integrated mesh is not compatible with the game engine that is the basis of platforms such as Metaverse, so it cannot be used as a character in Metaverse. Therefore, the work of modeling the clothes separately is carried out in [3].

Our previous work [3] estimates the body shape from a single image and models the clothes. However, there are several limitations to this method. For example, very few clothes types are reconstructed, and the texture is often misaligned. Therefore, in this paper, we attempt to solve this problem. This paper focuses on the clothes part and proposes a method for expanding the clothes

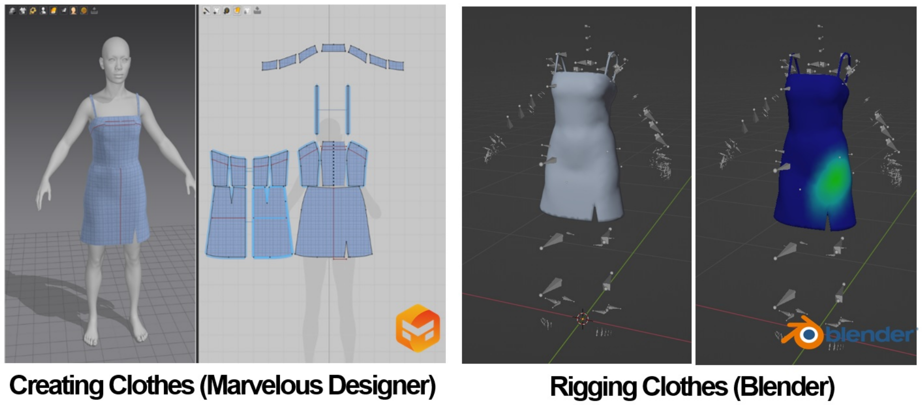

Fig. 1. 3D virtual human custom clothing creation process. Marvelous designer creates clothes that fit the 3D virtual human body type. And then, the clothes are rigged for animation in the Blender.

category and improving the accuracy of texture mapping. To this end, we extract the keypoints for each category of clothing proposed by DeepFashion2 [2] and proceed with a classification of clothes patterns. Thereafter, UV mapping using keypoints and clothes types is performed to represent the 3D clothes from the input image.

2 Proposed Method

2.1 Clothes Geometry Modeling

Essential clothing and body models use files provided by the MetaHuman Creator. However, the basic model of clothes supplied by MetaHuman Creator is limited. 3D tools such as Marvelous Designer and Blender are used to solve these limitations and create various clothes suitable for the 3D virtual human body.

We use Marvelous designer, which is mainly used as a tool for producing 3D clothes. Marvelous designer imports basic 3D virtual human body models as avatars to dress up. Afterward, the clothes are fitted to the body of the 3D virtual human set as an avatar. And the fitted clothes are created as 3D virtual human custom clothes and are exported as an FBX file for compatibility with other software such as blender and unreal engine.

The blender, which is used as a 3D Object editing tool, performs a rigging operation that allows clothes to be animated the same as the body. A 3D virtual human skeleton is combined with the clothes FBX file generated by Marvelous designers, and weight paint is performed on each bone to enable clothes animation. In addition, mesh deformation and UV Map modification are performed to generate clothes optimized for the 3D virtual human body.

Fig. 2. 3D virtual human custom clothing types. Through the process of creating custom clothes, clothes corresponding to 13 categories classified in the dataset DeepFashion2 were composed.

The process of creating custom clothes using two 3D tools can be seen in Fig. 1. And by this process, the results of constructing clothes corresponding to 13 clothes categories can be seen in Fig. 2.

2.2 Clothes Texture Modeling

UV Map is a plane of a 3D model used to wrap a texture. And it can express the texture of clothes. For UV Map generation, the pipeline as shown in Fig. 3 is proposed. In the previous study, clothes parsing was performed from the input image to generate UV Map, and in the case of the top, Pix2Pix was used to blur the boundary. However, inaccuracy was a problem because warping from incorrect points was performed in the UV Map production of the clothes. Therefore, to solve these problems, keypoint detection and clothes pattern classification is added to the existing pipeline.

DeepFashion2 [2] is used as the dataset required for learning clothes keypoint detection. DeepFashion2 has 13 clothes categories, and keypoints proposed for each category are presented according to each clothes characteristic. For learning, clothes classes and bounding boxes, which are results from clothes segmentation, are used together as inputs. And as a model, we proceed with learning using HRNet [6], which has features of maintaining high resolution throughout the entire process.

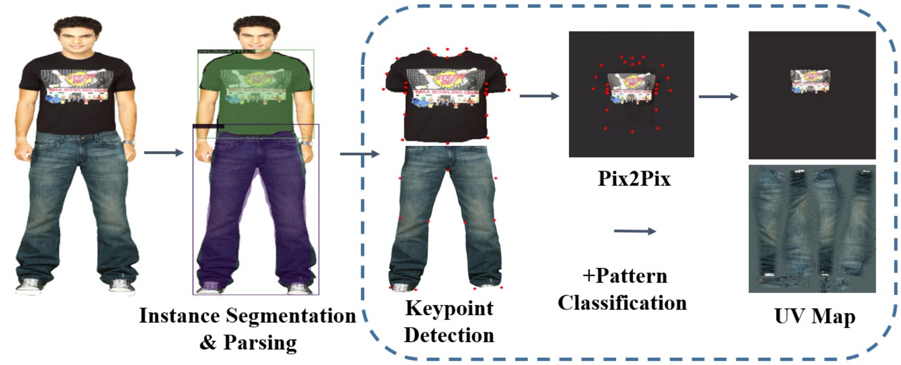

Fig. 3. Clothes UV Mapping Pipeline. When the input image comes in, keypoint detection is performed simultaneously with instance segmentation and parsing. After that, UV Map based Warping is performed with the Pix2Pix and pattern classification results.

The texture of the clothes on the invisible side during UV mapping is inferred. Therefore, according to the characteristics of the clothes, the clothes pattern classification is conducted on whether the pattern is repeated and continued, or not. (Such as the graphic or logo is centered) For the dataset, part of [1] provided by Kaggle was used, and data were organized and learned by image crawling to reflect the diversity of pattern clothes.

A method of generating a clothes UV Map using the above-described information is as follows. With the keypoint coordinate information of the Pix2Pix generation result and the clothes segmentation result, it is mapped by applying a function related to the perspective transformation of OpenCV to the coordinates to be mapped in the UV Map format. In addition, when mapping, it determines whether to reflect the back side according to the clothes pattern classification results and uses a boundary processing function to add naturalness.

3 Experimental Results

Experimental results are shown in Fig. 4. Note that we present the qualitative result only because there is no ground truth in this field. A clothes FBX combination is provided by classifying clothing types among 13 clothes categories from the input image. In addition, according to the UV Map format of the clothes FBX, a UV Map that reflects the texture of the clothes is provided to express the texture of the clothes model.

Looking at the results of Fig. 4, it can be seen that the clothes corresponding to the input image are provided according to the clothes category. However, the generated clothes did not completely match those in the input image. This is because the clothes categories are categorized into 13 categories and generalized, but there are more diverse lengths and fits in reality. Previous work [3] has

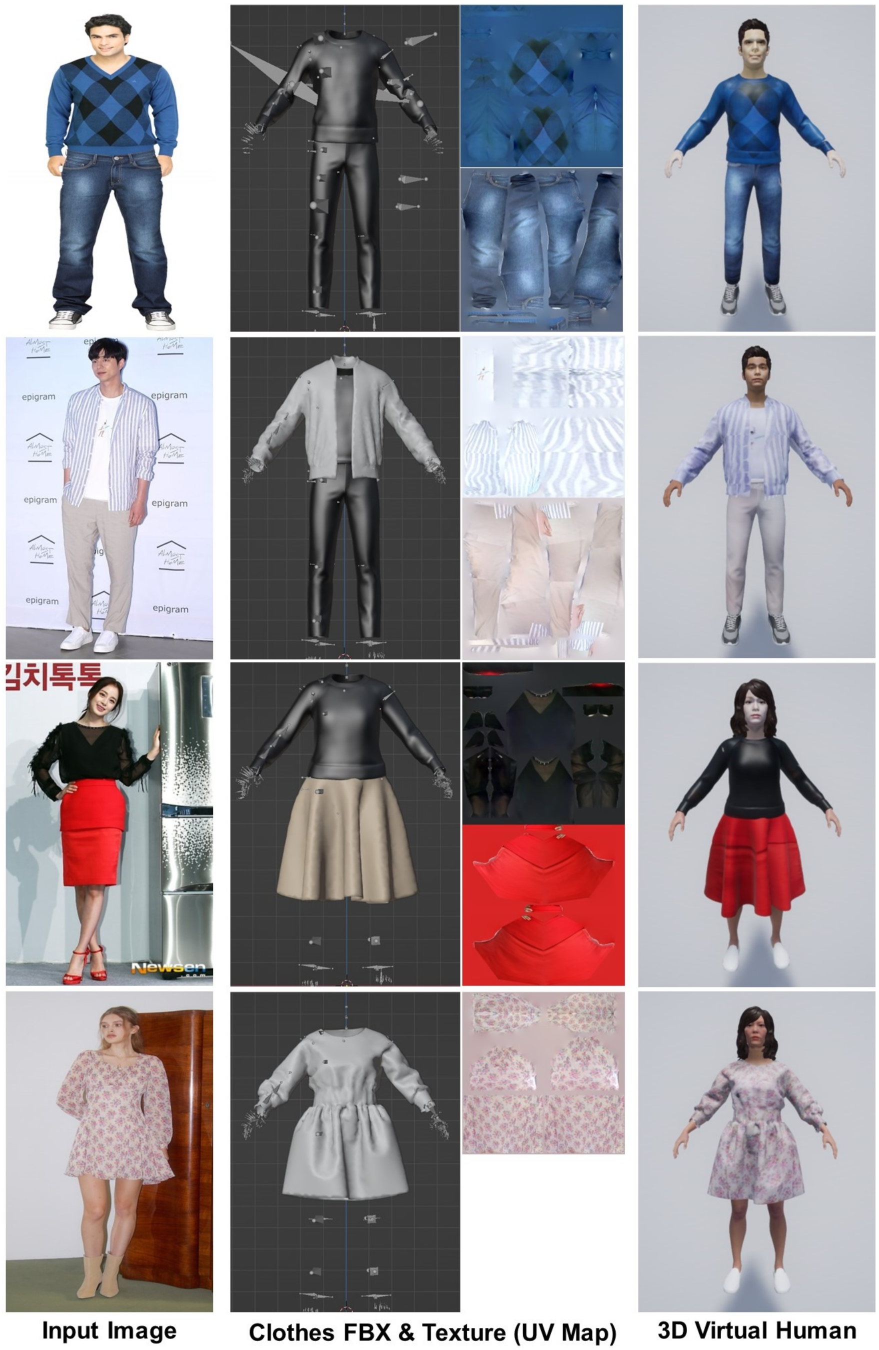

Fig. 4. Results of creating a 3D virtual human. The clothes of the corresponding category are classified and provided from the input image. It also provides UV Map reflecting the texture of the clothes. And the 3D virtual human created with the compatible clothes can be checked on the Unreal Engine.

mentioned errors in UV Map generation due to the inability to find feature points for images that are not frontal or static, or the problem of poor detail because only the center of the clothes is reflected. By checking the UV Map, it can be seen that these problems are solved and the overall details are reflected. However, obstacles such as hands appear as there is no dealing with the occlusion. To overcome this limitation, it seems necessary to handle the hidden side using the information on the occlusion provided by the deepfashion2 Dataset.

Looking at the final 3D virtual human in Fig. 4, the body and face materials are used with costumes to create a complete virtual human being. Although not covered in this paper, the clothes are transformed along with the body according to the body type estimate and the face is also reflected. The 3D virtual human finally created through the automation process can be identified in the Unreal Engine. By performing the clothes FBX rigging work, it is possible to animate the clothes to move the same as 3D virtual human body animation. And these animations can be checked along with the basic animations of the Unreal Engine.

4 Conclusion

This paper presented a clothes modeling method during 3D virtual human restoration from a single image. It was observed that 3D characters that captured the appearance of a real person could be created within the 3D virtualization platform by reflecting the increased types of clothes and detailed textures. We believe it would increase the satisfaction of metaverse users.

Acknowledgements This work was partly supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2022R1A4A1033549). This work was partly supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2022-0-00981, Foreground and background matching 3D object streaming technology development and No.RS-2022-00155915, Artificial Intelligence Convergence Innovation Human Resources Development (Inha University)).

References

1. Fashion product images dataset https://www.kaggle.com/datasets/paramaggarwal/fashion-product-images-dataset
2. Ge, Y., Zhang, R., Wang, X., Tang, X., Luo, P.: DeepFashion2: a versatile benchmark for detection, pose estimation, segmentation and re-identification of clothing images. In: Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 5337–5345 (2019)
3. Kim, H.W., Kim, D.E., Kim, Y., Park, I.K.: 3D clothes modeling of virtual human for metaverse. Journal of Broadcast Engineering **27**(5), 638–653 (2022)
4. Remelli, E., Bagautdinov, T., Saito, S., Wu, C., Simon, T., Wei, S., Guo, K., Cao, Z., Prada, F., Saragih, J., Sheikh, Y.: Drivable volumetric avatars using texel-aligned features. In: Proc. ACM SIGGRAPH 2022 Conference. pp. 1–9 (2022)

5. Saito, S., Huang, Z., Natsume, R., Morishima, S., Kanazawa, A., Li, H.: Pifu: Pixel-aligned implicit function for high-resolution clothed human digitization. In: Proc. IEEE/CVF International Conference on Computer Vision. pp. 2304–2314 (2019)
6. Sun, K., Xiao, B., Liu, D., Wang, J.: Deep high-resolution representation learning for human pose estimation. In: Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognitionn. pp. 5693–5703 (2019)