Customized 3D Clothes Modeling for Virtual Try-on System based on Multiple Kinects

Shiyi Huang and Won-Sook Lee
School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, Canada

Keywords: 3D Clothes Modeling, Virtual Try-on, Kinect.

Abstract: Most existing 3D virtual try-on systems put clothes designed in one environment on a human captured in another environment, which cause the mismatching brightness problem. And also typical 3D clothes modeling starts with manually designed 2D patterns, deforms them to fit on a human, and applies stitching to sew those patterns together. Such work usually relies on labour work. In this paper, we describe an approach to reconstruct clothes and human that both are from the same space. With multiple Kinects, it models the 3D clothes directly out of a dressed human without the need of the predefined 2D clothes patterns, and fits them to a human user. Our approach has several advantages: (1) a simple hardware setting consisted of multiple Kinects to capture a human model; (2) 3D clothes modeling directly out of captured human model. To the best of our knowledge, our work is the first one which separates clothes out of captured human figures; (3) resizing of clothes adapting to any sized human user; (4) a novel idea of virtual try-on where clothes and human are captured in the same location.

1 INTRODUCTION

Virtual try-on for clothes is the technique that can visualize the human dressed in selected clothes virtually without need to undress and dress themselves physically. Basically, it consists of two objects: human and clothes. And the main objective is to fit the clothes to a human body.

Based on the data type of human and clothes, we define that virtual try-on systems can be divided into 2D and 3D systems. And we define 2D systems to be 2D clothes and 2D/3D human while 3D systems to be 3D clothes and 2D/3D human. Substantial progress has been achieved in this area during last decades.

(Hauswiesner et al., 2013) proposed a 2D virtual try-on system through image-based rendering approach. The system has one obvious limitation, which is that a large garment database is required to solve the mismatching problem. The EON Interactive Mirror (Giovanni et al., 2012) was considered as the 3D virtual try-on system consisted of two stages. In the offline stage, 3D clothes are modelled based on actual catalogue images. While online, it used one Kinect to detect the nearest person in the scenario and track his/her motion to merge 3D clothes with the user's HD video frames from one high-definition camera. One of its limitations is that 3D clothes are resized uniformly based on the customer's shoulder height. (Zhang et al., 2014) also adopted one Kinect to construct a 3D virtual try-on system. In their system, one customized 3D static human model was generated from a template model based on the anthropometry method. And animation of the 3D static human model was performed based on the recorded customer's motion data captured by Kinect. While several predefined 2D garment patterns were mapped onto the 3D human model to generate the 3D clothes. However, special data format of 2D clothes patterns was required to avoid manually pre-positioning them onto the human model for this system. Other systems such Fitnect, and the demo video about Kinect for Windows Retail Clothes described the similar idea as the EON Interactive Mirror.

In this paper, we introduce a new idea of virtual try-on system. A mannequin is dressed nicely and a customer wants to check the clothes on him/her body but without actually putting on the clothes. We just capture the customer's body data using a scanner (or the customer's previously stored data can be used in the shop) and we transfer the

Huang, S. and Lee, W-s.
Customized 3D Clothes Modeling for Virtual Try-on System based on Multiple Kinects.
DOI: 10.5220/0005782001160121
In *Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2016) - Volume 1: GRAPP*, pages 118-123
ISBN: 978-989-758-175-5

mannequin's clothes to the customer virtually. We emphasize that the clothes and human are from the same space to avoid the mismatching brightness problem: bright/dark human skin color and dark/bright clothes. We use multiple Kinects to capture different views of one mannequin dressed in one shirt and one trouser. Clothes are extracted from the dressed mannequin and separated into two pieces of clothes. Multiple point clouds of the clothes with respect to the different views are aligned. And then surface reconstruction technique was applied to the aligned point clouds to generate the clothes mesh. Next, we applied mesh post-processing and texture mapping the clothes mesh to obtain the textured clothes. Besides, we used KinectFusion to scan the naked mannequin and a user dressed in the tight clothes to generate two 3D human models. Finally we resize the clothes to fit to the user based on the same transformation that deforms the naked mannequin to have the user's body size and shape. A detailed overview of our method is provided in Section 3.

2 RELATED WORK

This section briefly describes related work about virtual try-on systems, clothes modeling, clothes customization, and provides a short overview of data processing techniques related to our work.

2.1 Virtual Try-on

Over the last several years, various methods for virtual try-on have been designed. Due to the lack of the available high-performance and high-accuracy capturing device, some early work usually generated relatively poor result. In the following, we mainly discuss the latest development.

Virtual Try-on in 2D: (Araki and Muraoka, 2008) designed a 2D system that resized the triangulated clothes images based on several markers on the user's joints to fit his/her motion captured by a web camera. (Hauswiesner et al., 2013) constructed a system that transferred the appearance of the recorded clothes onto the user's body by matching his/her pose from the input video frames with that from the recorded video frames. (Isıkdogan and Kara, 2013) mapped and resized 2D clothes images to fit to the user based on the tracked skeleton from one Kinect.

One obvious disadvantage of the 2D virtual try-on system is that the original clothes in the user's body usually can be seen even after putting on selected clothes.

Virtual Try-on in 3D: With the advance of several 3D sensing devices, especially Kinect, the development of the 3D virtual try-on system is unprecedented. (Giovanni et al., 2012) fitted the manually modelled 3D clothes to the user' high-definition video frames from one HD camera by using one Kinect to track his/her skeleton and align clothes with the skeleton. (Yuan et al., 2013) uses one Kinect to capture user's data, such as skeleton data and body skin colour for customizing one avatar, and then transferred 3D clothes onto the customized avatar. (Zhang et al., 2014) also utilized one Kinect to capture user's data for customizing and refining one template human model. And several 2D clothes patterns were fitted to the human model and then sewed together to achieve the virtual try-on effect.

Those 3D virtual try-on systems have a common drawback, which is that clothes and human are from different space. And those 3D clothes modeling parts are usually 2D patterns based.

2.2 Clothes Modeling for Virtual Try-on

3D clothes modeling is one part of the virtual try-on system, which basically can be divided into three types: manually modeling, automatically 2D patterns to 3D clothes modeling, and scanner-based 3D modeling. Other methods are also proposed for 3D clothes modeling, for example, (Furukawa et al., 2000) created 3D clothes based on 2D photo input where the body was created from three photo views and then clothes were extracted in 3D by colour segmentation. As it was based on feature-based human cloning technique, accuracy was relatively low.

Manually clothes modeling usually requires the designer to use some 3D computer graphic software, such as Blender, 3ds Max or Maya to model 3D clothes. Several 2D clothes patterns were designed virtually, and the linkages were specified by manually selecting corresponding points between nearby patterns. Then those patterns were deformed to fit a digital human model and textured to generate the 3D clothes. This approach usually requires a lot of manual work in order to generate a high quality result.

The automatically 2D patterns to 3D clothes modeling was introduced by (Zhang et al., 2014), which required special data format of 2D clothes patterns to prevent manually preposition 2D patterns onto customized human model. However, it still required designers to position those 2D patterns to the template model manually.

Scanner-based 3D clothes modeling is the idea that reconstruct 3D clothes based on the 3D scanner device. The scanner-based method is popular for human modeling and hair modeling. However, to the best of our knowledge, we haven't seen any research about scanner-based 3D clothes modeling yet, which can be an area that we can have a contribution to.

2.3 Clothes Customization for Virtual Try-on

Fitting the 3D clothes to the 3D human is a crucial part in the 3D virtual try-on system. The fitting step directly affects the visualization result. Most existed 3D virtual try-on systems achieve the clothes customization by deforming several 2D patterns to fit the human body. However, the approach that directly resizing 3D clothes to fit to the 3D human is more intuitive. (Li et al., 2010) deformed the clothes mesh to match the shapes of the human model through several steps, which made the approach relatively complicated.

2.4 Related Data Processing Techniques

For our 3D scanner-based virtual try-on system, several data processing techniques have been used: color segmentation, 3D registration, surface reconstruction, texture mapping, and resizing.

(Zhang, 2000)'s calibration method is widely used to calibrate camera for 3D reconstruction, which only requires one simple pattern that has several feature points to be detected so that the intrinsic parameters can be calculated. Stereo camera calibration is used to calculate the extrinsic parameters between two cameras. In order to extract concerned objects from the image, the color segmentation techniques are the ideal choice. JSEG system developed by (Deng and Manjunath, 2001) was used in our system for clothes segmentation. From the captured images and depth data, several point clouds can be generated. Since each point clouds is with respect to different coordinate system, we used ICP algorithm (Besl and Neil, 1992) to align multiple point clouds. Poisson Surface Reconstruction method proposed by (Kazhdan et al., 2006) was used to create mesh. A texture mapping technique similar to (Zhou et al., 2005) was utilized to apply texture information to the mesh. And for resizing the clothes to fit the human body, we utilized the Radial Basis Function.

3 OVERVIEW

Our method of 3D clothed modeling and customization based on multiple Kinects consists of three key components: data acquisition, 3D clothes modeling and their customization. In Sections 4 to 6 we elaborate on the components of our method. We then show the results in Section 7, and conclude with a discussion of the method.

4 DATA ACQUISITION

This section describes our acquisition setup, including the hardware and calibration.

Our capture environment is displayed in Figure 1.

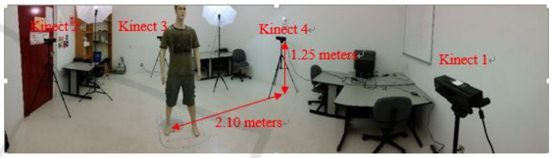


Figure 1: Capture environment setting.

We utilized Kinect V2 as the main capturing device, which is equipped with one higher resolution RGB camera and one higher depth fidelity depth camera compared with the Kinect V1. Three kinds of images can be generated from the two cameras: RGB image, depth image and infrared image. Besides, depth image and infrared image are generated from the same camera. We utilized (Zhang, 2000)'s method to do camera calibration for every Kinect built-in cameras. A checkerboard consists of 12 x 20 squares was used as the pattern to implement the calibration. 25 synchronized RGB images and infrared images for each Kinect built-in RGB camera and depth camera were captured. The captured checkerboard images are displayed in Figure 2.

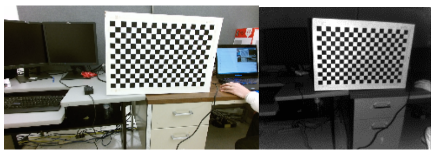


Figure 2: Captured checkerboard images. Left: RGB image; Right: Infrared image.

Through stereo camera calibration, the transformation between RGB camera space and depth camera space can be derived. (Macknojia et al., 2013) introduced the idea of camera calibration of two built-in cameras and stereo camera calibration for Kinect V1 in detail. Different from Kinect V1, there is no offset between depth image and infrared image for Kinect V2.

In the data acquisition stage, we used KinectFusion (Newcombe et al., 2011) to scan the naked mannequin and a user dressed in tight clothes to obtain their corresponding point clouds, which would be processed based on surface reconstruction algorithm to model two 3D mesh out of them. And then we utilize four Kinects to capture four views of the naked mannequin so that we can align each view with the naked mannequin model from KinectFusion to calculate the transformation between every two Kinects, which can be used to align multiple views of the clothes. Next, we captured four views of the mannequin with one shirt and one trouser.

5 3D CLOTHES MODELING

5.1 Clothes Extraction

Four views' data of clothes are acquired during the data acquisition step. Clothes extraction was implemented through two steps: (1) extracting the dressed mannequin out of the common background; (2) extracting clothes out of the dressed mannequin and separating them into one shirt and one trouser.

We utilized Kinect Body Tracking Function to extract the dressed mannequin from the common background in front and back view. While for left and right view, since the Function usually doesn't work, depth thresholding was used to roughly extract the dressed mannequin.

To extract clothes out of the dressed mannequin, we used the JSEG algorithm proposed by (Deng and Manjunath, 2001). And then several manual work was performed to remove some unconcerned regions. And due to the block of mannequin's hands, data missing area existed in the left and right view, which require the step of region filling to repair those areas. The final result is displayed in Figure 3.

5.2 3D Registration of Kinect Data

We have five groups of 3D data, four from four views of Kinect (four views of clothes data and four views of the naked mannequin) and one from KinectFusion (the naked mannequin) and we need to register them in one space.

Figure 3: Four views shirt and trouser images after performing clothes extraction, manually removing unconcerned regions and region filling.

We used ICP proposed by (Besl and Neil, 1992) to align multiple point clouds to create the 360° point clouds. Our proposed approach to do 3D registration for clothes lied in four steps: (1) using KinectFusion to scan the naked mannequin to get the well-aligned point clouds; (2) using four Kinects to capture four views of the naked mannequin; (3) aligning each view of the naked mannequin with the well-aligned point clouds to obtain the relation between every two Kinects; (4) aligning four views point clouds of clothes using the same relation calculated in (3).

The aligned point clouds of clothes usually contain several noise and outliers, and then we manually removed many outliers. The result is displayed in Figure 4.

Figure 4: Left to right: aligned shirt and trouser point clouds after manually removing several outliers.

5.3 Surface Reconstruction of Clothes

Based on the aligned point clouds of clothes, we utilized Poisson surface reconstruction algorithm to create the mesh. The result is displayed in Figure 5.

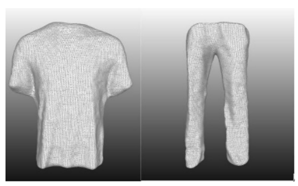


Figure 5: Left to right: reconstructed shirt and trouser mesh after using Poisson surface reconstruction.

The model reconstructed using Poisson surface reconstruction is watertight, which makes the reconstructed shirt and trouser quite different from the real clothes. Several steps are implemented to process the watertight 3D model of clothes to match the clothes regions in the image. Our proposed approach was divided into six steps: (1) using Kd-Tree to do roughly texture mapping for the clothes; (2) selecting several points on the mesh manually; (3) deleting triangular faces along the selected points; (4) removing other unconcerned components; (5) finding the boundary edges of the mesh left; (6) creating several triangular faces to link boundary edges and the selected points. The result is displayed in Figure 6.

Figure 6: Left to right: reconstructed shirt and trouser mesh after mesh post-processing.

5.4 Texture Mapping of Clothes

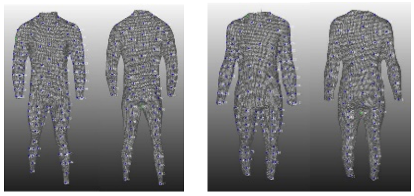
After the surface reconstruction and mesh post-processing, we obtained the mesh without texture information. Then we segmented the mesh into four parts with respect to each view for texture mapping. Several points were manually selected on the mesh to define the intersection between every two views. Then we applied the algorithm used for deleting triangular faces along the selected points to the clothes mesh. The textured shirt and trouser are displayed in Figure 7.

Figure 7: Top to Down: the well-reconstructed and textured shirt and trouser mesh after texture mapping.

6 3D CLOTHES CUSTOMIZATION

We used Radial Basis Function for clothes customization. The approach to resize the 3D clothes to fit to the user's body can be divided into two steps: (1) deforming the naked mannequin to have the user's body size and shape using Radial Basis Function (2) resizing the 3D clothes to fit to the user's body using the same Radial Basis Function.

Using Radial Basis Function for resizing requires two data groups: control points and target points. Multiple points were selected on source and target models manually. The point selections are displayed in Figure 8.



(a) (b)

Figure 8: (a): control points selection on the naked mannequin's body; (b): target points selection on the user's body.

The raw naked mannequin model, user's 3D model, and the deformed naked mannequin model are displayed in Figure 9.

3D clothes were resized and fitted to the user. But the fitting result has the surface penetrating problems: several faces of the 3D clothes penetrating into the user's body. We applied the surface correction algorithm based on the point-to-plane project theory to pull those penetrated faces out of the body. And the result is displayed in Figure 10.

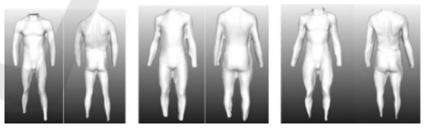


Figure 9: Left to Right: raw naked mannequin model; user's 3D model; deformed naked mannequin model.

Figure 10: Left to Right: user with reconstructed clothes before and after surface correction.

7 RESULT

The evaluation is based on the visualization comparison between the captured clothes image and the reconstructed clothes. As shown in Figure 11, the clothes shape and texture in the captured images are passed to the 3D clothes model. The patterns on shirt are much similar to the raw patterns, and the virtual mannequin dressing the reconstructed clothes also shares a variety of similarities with the actual one. Clothes with different styles (such as pants, skirt, trouser, and short/long sleeves shirt) and various colour and texture (such as blue, green, and black) are all successfully reconstructed using the proposed approach.

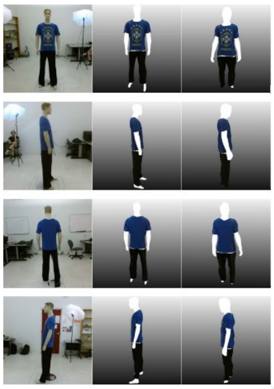


Figure 11: Left to Right: raw images of the mannequin dressed in selected clothes; reconstructed mannequin with the reconstructed clothes; user dressed in deformed clothes.

8 CONCLUSIONS

In this paper, a novel approach for 3D clothes modeling and customization is proposed. Different from previous systems, we captured clothes and human data that both are from the same space to avoid the mismatching brightness problem. Besides, we modelled 3D clothes directly out of the captured human figures without need to use the predefined 2D patterns to avoid so much labour work. Those characteristic makes our system more practical.

ACKNOWLEDGEMENTS

We wish to thank Iman Eshraghi for his help during data acquisition.

REFERENCES

Araki, N., and Muraoka, Y. 2008. Follow-the-Trial-Fitter Real-time addressing without undressing. In *Third International Conference on Digital Information Management*, 33-38.

Besl, P. J., and McKay, N. D. 1992. Method for registration of 3-D shapes. In International Society for Optics and Photonics, *Robotics-DL tentative*, 586-606.

Deng, Y. and Manjunath, B. S. 2001. Unsupervised segmentation of colour-texture regions in images and video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 800-810.

Furukawa, T. Gu, J., Lee, W., and Magnenat-Thalmann, N. 2000. 3D clothes modeling from photo cloned human body. In *Springer Berlin Heidelberg, Virtual Worlds*, 159-170.

Giovanni, s., Choi, Y. C., Huang, J., Khoo, E. T., and Yin, K. 2012. Virtual try-on using Kinect and HD camera. In *Springer Berlin Heidelberg, Motion in Games*, 55-65.

Hauswiesner, S., Straka, M., and Reitmayr, G. 2013. Virtual try-on through image-based rendering. *IEEE Transactions on Visualization and Computer Graphics*, 1552-1565.

Isikdogan, F., and Kara, G. 2013. A real time virtual dressing room using Kinect.

Kazhdan, M., Bolitho, M., and Hoppe, H. 2006. Poisson surface reconstruction. In *Proceedings of the fourth Eurographics symposium on Geometry processing*.

Li, J., Ye, J., Wang, Y., Bai, L., and Lu, G. 2010. Fitting 3D garment models onto individual human models. In *Computers and Graphics*, 742-755.

Macknojia, R., Chavez-Aragon, A., Payeur, P., and Laganiere, R. 2013. Calibration of a network Kinect sensors for robotic inspection over a large workspace. In *IEEE Workshop on Robot Vision*, 184-190.

Newcombe, R. A., Izadi, S., Hilliges, O., Molyneaux, D., Kim, D., Davison, A. J., and Fitzgibbon, A. 2011. KinectFusion: Real-time dense surface mapping and tracking. In *10th IEEE international symposium on Mixed and augmented reality*, 127-136.

Yuan, M., Khan, I. R., Farbiz, F., Yao, S., Niswar, A., and Foo, M. H. 2013. A mixed reality virtual clothes try-on system. *IEEE Transactions on Multimedia*, 1958-1968.

Zhang, Z. 2000. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1330-1334.

Zhou, K., Wang, X., Tong, Y., Desbrun, M., Guo, B., and Shum, H. Y. 2005. TextureMontage. In *ACM Transactions on Graphics*, 1148-1155.

Zhang, Y., Zheng, J., and Magnenat-Thalmann, N. 2014. Clothes Simulation and Virtual Try-on with Kinect Based on Human Body Adaptation. In *Spring Singapore Simulation, Serous Games and Their Applications*, 31-50.