Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 108C (2017) 355–364

Procedia Computer Science

International Conference on Computational Science, ICCS 2017, 12-14 June 2017, Zurich, Switzerland

3D High-quality Garment Reconstruction with Synthesized Texture

Pengpeng Hu[1], Taku Komura[2], Duan Li[1], Ge Wu[1] and Yueqi Zhong[1, 3, *]

[1]*Donghua University, Shanghai 201620, China*

[2]*Edinburgh University, Edinburgh EH8 9AB, UK*

[3]*Key Lab of Textile Science and Technology, Ministry of Education, China*

Abstract:

3D garment model plays an important role in the fashion industry. However, not much work focus on extracting high-quality clothes from scanned data. The texture is limited by photography methods in 3D scanning. This paper proposes a novel framework of reconstructing high-quality 3D garment models with synthesized texture. Firstly, a pipeline of 3D garment processing is proposed to obtain a better 3D model based on KinectFusion. Then, DeepTextures is used to synthesize a new texture. To our best knowledge, this is the first paper combining 3D garment reconstruction and texture synthesis. Experimental results show that our method can conveniently obtain 3D garment models and realistic textures.

Peer-review under responsibility of the scientific committee of the International Conference on Computational Science

Keywords: 3D garment reconstruction, Texture synthesis, KinectFusion

1 Introduction

3D scan technology has been widely used in the fashion industry to capture the 3D geometric shapes of clothes and the human body. Texture, as one of the most important features for clothes, is still a challenging task in today's fashion industry. One important parameter of clothes is the shape in three dimensions, which can give human more intuitive information than that from two-dimension images. With the rapid development of commercial depth cameras, e.g. Kinect, commercial systems for human body scanning based on depth cameras have been proposed [1,2]. However, previous work usually focus on capturing the geometry of clothes, and the texture usually is captured by the RGB camera, which is limited by the photography method. To visualize different textures for 3D garments is very interesting and useful in garment design.

In the fashion industry, the texture is mainly designed by the artist which is expensive and time-consuming. An increasing number of people, especially young people, like having their own pattern on the T-shirt which is called cultural T-shirt. Hence, there is a permanent interest in the development of rapid and automatic texture generation method for the garment. However, the human is extremely sensitive to the texture of clothes. It is not trivial that a realistic 3D textured garment model should be obtained.

*Corresponding author: zhyq@dhu.edu.cn

1877-0509 © 2017 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the International Conference on Computational Science
10.1016/j.procs.2017.05.004

This paper presents a novel framework for 3D garment reconstruction with synthesized texture. Our framework is shown in Fig.1. An efficient pipeline is designed for 3D garment reconstruction based on KinectFusion. DeepTextures is used to synthesize a new texture. To our best knowledge, it is the first paper to combine 3D garment reconstruction with texture synthesis.

2 Overview of the Framework

Our goal is to develop a framework of reconstructing 3D garment models with synthesized textures, as shown in Fig.1. We start by capturing the 3D garment model using KinectFusion with the help of a turntable (Section 3.1). Then a pipeline is designed for 3D garment mesh processing to increase the quality of the garment shape (Section 3.2). The DeepTextures is utilized for synthesizing a new texture from the original texture. (Section 3.3). The rest of this paper is structured as follows: Section 4 shows the results; Section 5 is the discussion and conclusion.

3 Methodology

3.1 3D Scan

KinectFusion is a popular single-Kinect 3D scanning system, which can be used for Kinect V1 and Kinect V2 [3]. Kinect V2 is chosen in our work as it has a higher depth resolution. We ensure the distance between Kinect v2 and the object to be no more than 1.0 m to obtain good depth resolution according to the work of Butkiewicz et al.[4]. As KinectFusion always collapses during scanning large objects, like the human body, and moving the Kinect at a constant slow speed will obtain a better reconstructed 3D model. A turntable is used as the platform, which the object is located on, to increase the accuracy and robustness of KinectFusion. It costs 30 seconds to scan the whole object while the turntable rotates a round.

3.2 3D Mesh Processing

With the help of the turntable, 3D garments can be easily obtained. Without loss of generality, we take reconstructing trousers model as an example to further explain our method. It is useless to scan a flat garment laid on the ground. Therefore, a human body dressing trousers are scanned and the trousers model is extracted next. The 3D model from KinectFusion is shown in Fig.2 (a). It can be seen that the 3D model is not perfect. It needs to be further processed, which is called 3D mesh processing. In our pipeline, as shown in Fig.1 (a), the noises and outliers should be removed first. Bilateral Mesh Denoising [5] is used in this step. As KinectFusion merges too many frames of depth images, redundant data cannot be avoided. Too large data is also not conveniently used in practice. Secondly, 3D model is subsampled. Poisson-disk sample [6] is used in this step. The occlusion is a generic issue in almost all 3D scanning work, especially when we scan garments and bags in our work. Thirdly, a hole-filling algorithm [7] is applied to get complete 3D garments. Next, the surface of the 3D garment is smoothed [8]. Universally, a pretty good 3D model can be obtained through these four steps, as shown in Fig.2 (b). However, due to features of garments, these processes are not sufficient. Here, we just list part of garment specificities:

- **Diversity**: Containing various kinds of objects, e.g. T-shirt, pants, shorts and so on. So garments have many different topologies;

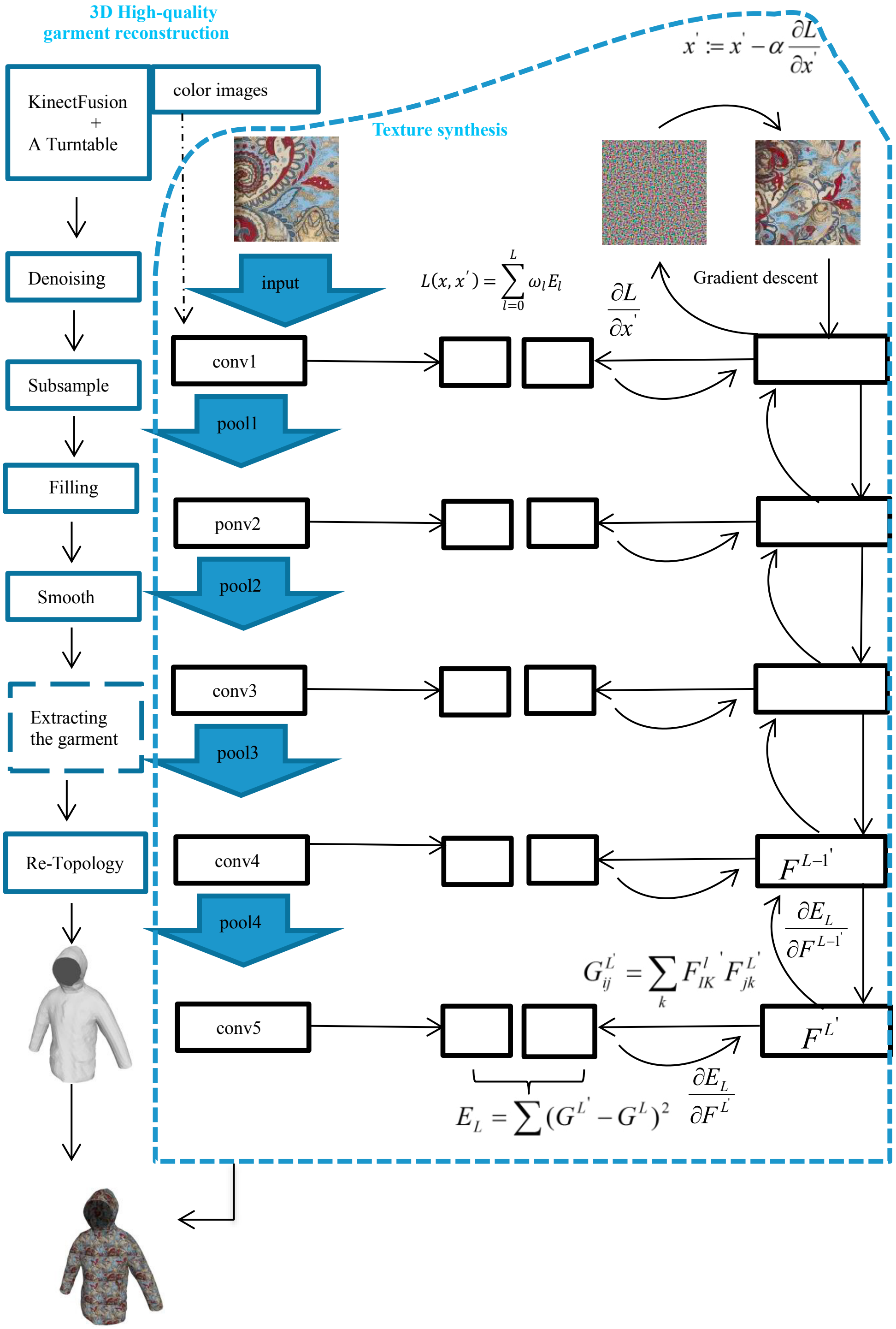


Figure 1. The framework of 3D high-quality garment reconstruction with synthesized texture. A better 3D garment model is obtained based KinectFusion using our method (left). The texture is generated using DeepTextures (right). An initial noise image x is passed through the CNN and a loss function E_l is computed

on each layer of the texture model. L is a weighted sum of the contributions E_l from each layer. A new texture is found by producing the same Gram matrices $G_l^{'}$ as the original texture.

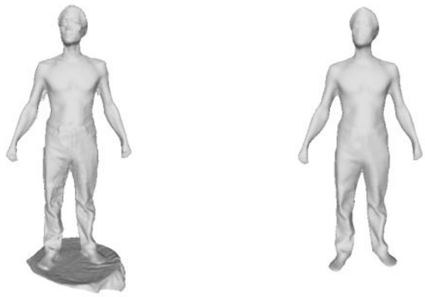
- **Non-rigidity**: Garments are non-rigid, and once the geometry of the garment is changed by external force it is impossible to recover its original shape;
- **Utility:** Garments are designed for meeting kinds of needs, like keeping warm, protecting and so on. So, the 3D model of the garment in usage is needed.

Majorities of garments are scanned together with other accessories due to above features, e.g. the garment is dressed on a body for 3D scanning. Therefore, differing from generic 3D mesh processing, the trousers extraction from the human body model is very important. As the data structure of the 3D model is triangle soup, zigzag edges along the cutting line will happen. As shown in Fig.3 (a), the waist of trousers and leg openings have bad shape. Our solution to this issue is to do re-topology for the trousers model.

Meshing algorithms can be classified into local and global methods. The former ones are usually simple, robust and scalable. Global algorithms solve optimization problems whose size depends on the entire dataset. In our work, a re-topology method called Instant Field-aligned Meshing [9] is chosen. It computes a mesh that is globally aligned with a direction field using local orientation and position-field smooth operators. The mesh is then extracted from the fields and optionally post-processed. Comparing to other re-topology methods, Instant Field-aligned Meshing has following advantages:

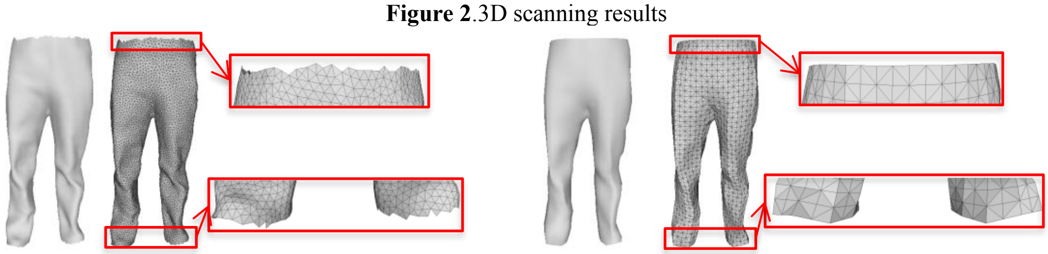
- This method is simple to implement and parallelize, and it can process a variety of input surface representations, such as point clouds, range scans and triangle meshes;
- This method can process extremely large meshes and point clouds with sizes exceeding several hundred million elements, as it avoids any global optimization;
- This method is interactive.

Due to these features, Instant Field-aligned Meshing algorithm is used to implement re-topology in our work. The extracted trousers model using our method is shown in Fig.3 (b). It can be seen that the waist of trousers and leg openings are perfect.



(a) 3D model from KinectFusion (b) 3D model after generic processing

Figure 2.3D scanning results



(a). Trousers using generic 3D mesh processing (b). Trousers using our method

Figure 3. 3D trousers model

3.3 Texture Generation

The convolutional neural network has been proven to be an excellent approach to texture synthesis [10,11,12] and image style transfer [13,14] for nature images. In our work, the DeepTextures method proposed by Gatys et al. [11] is chosen to synthesize the new texture. The basic idea of their approach is to consider each layer in the convolutional network as a non-linear filter bank, and its activations in response to an image from a set of feature maps. The overall procedure of this method follows:

Step 1. Extract features of different sizes homogeneously from the original image.

Step 2. Compute a spatial summary statistic on the feature responses to get a stationary representation of the original image, shown in Fig.1 (b).

Step 3. Synthesize a new image with the same stationary representation by performing gradient descent on a random image which has been initialized with white noise, as shown in Fig.1 (b).

More detailed explanation on this method can be found in [11].

4 Results

To evaluate the proposed method, we first scan some physical garments to obtain their 3D models. The original textures are from Texture website [15] and captured by the RGBD camera. The obtained 3D textured trousers can be seen in Fig. 4.

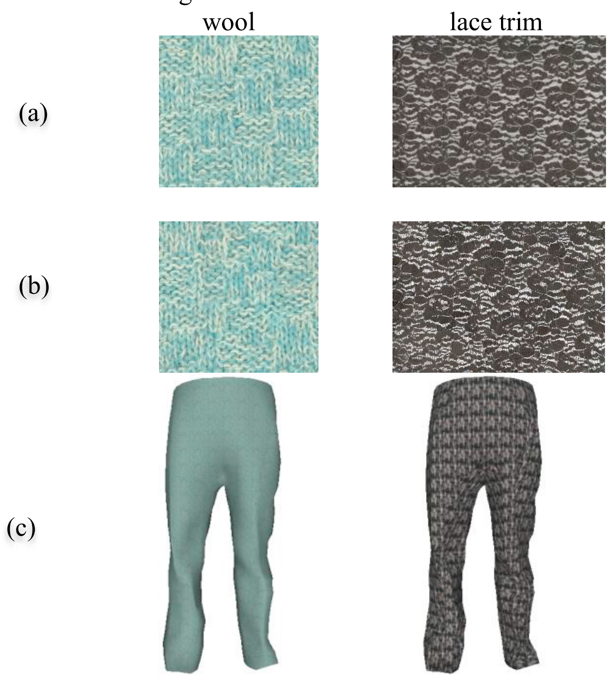


Figure 4. Images of the first row (a) are the original textures; Images of the second row (b) are the synthesized textures; (c) shows the 3D textured trousers model.

We further show synthesized textures using a different number of layers that are used to constrain the gradient descent, as shown in Fig.5. It means that the images in the first row were generated only from the texture representation of the first layer ('conv1_1') of the VGG-19 network. The images in the second row were generated by jointly representations of the layers of 'conv1_1', 'conv1_2' and 'pool1'. From Fig.5, we can see that the texture generated by constraining all layers to layer 'pool4' are nearly indistinguishable from the original texture (Fig.5, the last row). More results are shown in Fig.6.

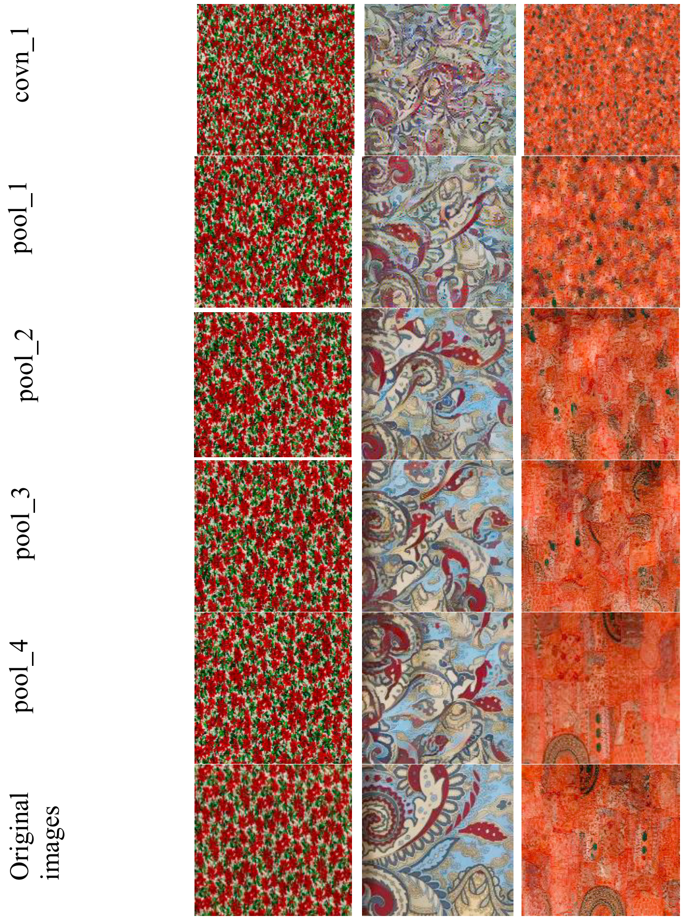
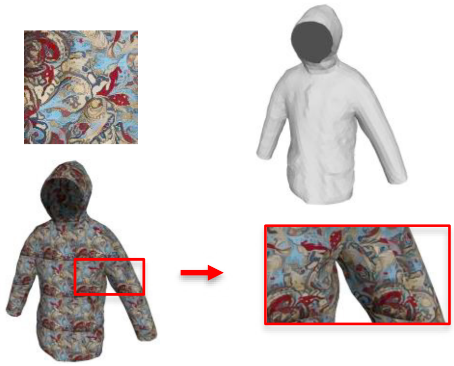


Figure 5. Each row corresponds to a different processing stage in the network

(a) 3D textured T-shirt (b) 3D textured suit



(c) 3D textured hoodie

Figure 6. More 3D garment models with synthesized textures using our method

(a)

(b)

Figure 7. Wrong synthesized textures: (a) Original texture; (b) Synthesized textures

5 Discussion and Conclusion

We introduced a novel framework of reconstructing high-quality 3D garment models with synthesized textures. Our main contributions are as follows:

- A novel pipeline is designed to obtain 3D high-quality garment models based on KinectFusion;
- Texture synthesis is first introduced to fashion industry;
- To our best knowledge, this is the first paper combining 3D scanning with texture synthesis. It overcomes the restriction of photography texture in the area of 3D scanning to obtain realistic 3D synthesized textured garment models, which is the first attempt to explore Artificial Intelligence of fashion industry.

A 3D garment model with better geometry can be obtained using our method. And the application of texture synthesis can generate new texture for 3D garment model, which is overcoming the limitation of texture mapping from photography methods in 3D scanning. The final textured 3D garment model from our method is high-quality and realistic, as shown in Fig. 6.

There are still some limitations on our method. The 3D garment model cannot be extracted automatically due to its complicated geometry, so the quality of 3D garment models highly depends on manual precision. As to the texture synthesis, an original texture should be given in our work, the synthesized texture has a similar style to the original texture. As shown in Fig. 7, DeepTextures cannot preserve the regular structure of the texture. It will be improved in our future work.

In the future, we hope to generate the arbitrary texture without any input images. Also, it is interesting to synthesize the 3D garment geometry rather than using 3D scanning. A similar concept has been proven in the field of human body modeling called Body Talk [16]. Only a brief word description, e.g. “medium thin” and “tall”, can generate a good human body shape.

Acknowledgment

This work is supported by Natural Science Foundation of China (Grant No.61572124).

Reference

1 Izadi S, Kim D, Hilliges O, et al. KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera[C]//Proceedings of the 24th annual ACM symposium on User interface software and technology. ACM, 2011: 559-568.
2 Smisek J, Jancosek M, Pajdla T. 3D with Kinect[M]//Consumer Depth Cameras for Computer Vision. Springer London, 2013: 3-25.
3 Amon C, Fuhrmann F, Graf F. Evaluation of the spatial resolution accuracy of the face tracking system for kinect for windows v1 and v2[C]//Proceedings of the 6th Congress of the Alps Adria Acoustics Association. 2014.
4 Butkiewicz T. Low-cost coastal mapping using Kinect v2 time-of-flight cameras[C]//2014 Oceans-St. John's. IEEE, 2014: 1-9.
5 Fleishman S, Drori I, Cohen-Or D. Bilateral mesh denoising[C]//ACM transactions on graphics (TOG). ACM, 2003, 22(3): 950-953.
6 Dunbar D, Humphreys G. A spatial data structure for fast Poisson-disk sample generation[J]. ACM Transactions on Graphics (TOG), 2006, 25(3): 503-508.
7 Zhao W, Gao S, Lin H. A robust hole-filling algorithm for triangular mesh[J]. The Visual Computer, 2007, 23(12): 987-997.
8 Jones T R, Durand F, Desbrun M. Non-iterative, feature-preserving mesh smoothing[C]//ACM Transactions on Graphics (TOG). ACM, 2003, 22(3): 943-949.
9 Jakob W, Tarini M, Panozzo D, et al. Instant field-aligned meshes[J]. ACM Transactions on Graphics, 2015, 34(6): 15.
10 Ulyanov D, Lebedev V, Vedaldi A, et al. Texture Networks: Feed-forward Synthesis of Textures and Stylized Images[C]. In ICML 2016 (pp. 1349-1357).
11 Gatys L, Ecker A S, Bethge M. Texture synthesis using convolutional neural networks[C]//Advances in Neural Information Processing Systems. 2015: 262-270.
12 Berger G, Memisevic R. Incorporating long-range consistency in CNN-based texture generation[J]. arXiv preprint arXiv:1606.01286 (2016).
13 Gatys L, Ecker AS, Bethge M. A neural algorithm of artistic style.[J]. arXiv preprint arXiv:1508.06576 (2015).
14 Gatys L, Ecker AS, Bethge M. Image style transfer using convolutional neural networks. [C] Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.
15 www.textures.com
16 Streuber S, Quiros-Ramirez M A, Hill M Q, et al. Body talk: crowdshaping realistic 3D avatars with words[J]. ACM Transactions on Graphics (TOG), 2016, 35(4): 54.