© 2017 IJSRST | Volume 3 | Issue 3 | Print ISSN: 2395-6011 | Online ISSN: 2395-602X
Themed Section: Science and Technology

A Virtual Dressing Room Using Kinect

Jagtap Prajakta Bansidhar, Bhole Sheetal Hiraman, Mate Kanchan Tanaji, Prof. S. V. More, Prof. B. S. Shirole
Sanghavi College of Engineering, Nashik., Varvandi, Maharashtra, India

ABSTRACT

We Shows a novel virtual fitting room framework using a depth sensor, which provides a realistic fitting experience with customized motion filters, size adjustments and physical simulation. The proposed scaling method adjusts the avatar and calculate a standardized apparel size according to the user's measurements, prepares the collision mesh and the physics simulation, with a total of 1 s preprocessing time. The real-time motion filters prevent unnatural artifacts due to the bug from depth sensor or self barrier body parts. We apply bone splitting to realistically render the body parts near the joints. All components are combined efficiently to keep the frame rate higher than previous works while not sacrificing realism.

Keywords : virtual try-on, Kinect, HD camera, OpenNI, Kinect for Windows Augmented Reality, Human-Computer Interaction, Kinect.

I. INTRODUCTION

Trying clothes in clothing stores is usually a time consuming activity. Moreover, it might not even be possible to try on clothes in the store, such as when ordering clothes online. Here we propose a simple virtual dressing room application to make shopping for clothing faster, easier, and more accessible. The first problem we address in the design of our application is the correct position of the user and virtual cloth models. Detection and skeletal tracking of a user in a video stream can be implemented in several ways. For example, Kjaerside et al. [1] proposed a tag-based augmented reality dressing room, which needs sticking visual tags for motion capture. More recently, Shotton et al. [2] have developed a real-time human pose recognition system that predicts the 3D positions of body joints, using a single depth image without visual tags. In this project, we use Shotton et al.'s method and a Microsoft Kinect sensor to create a tag less, real-time augmented reality dressing room application. Developer tools, such as the ones included in the OpenNI framework and the Microsoft Kinect SDK, ease developing applications based on the Kinect sensor. We used the Kinect SDK as it includes a robust real-time skeletal body tracker based on [2].

Figure 1. The user interface of the application

or a better fitting experience. We present a novel virtual fitting room framework that provides all the basic features expected from such an application, along with enhancements in various aspects for higher realism. These enhancements include motion filtering, customized user scaling, and the use of a physics engine. The motion filtering process starts with temporal averaging of joint positions in order to overcome the high noise of the depth sensor. However, temporal averaging does not prove to be sufficient because unnatural movements take place due to limited recognition capabilities and self-occlusion. We implement customized joint angle filters, along with bone splitting, to let limbs twist in a more natural way. We also employ filtering on hip and knee joints to overcome the foot skating problem.

II. LITERATURE SURVEY

J. Shotton, T. Sharp[2], Markerless human motion tracking is a long-standing problem in computer vision.With the recent advances in depth cameras and sensors, especially the Kinect sensor research on human skeletal pose tracking has made great improvements . Our system builds on top of these techniques by utilizing publicly available SDKs that incorporate some of these state-of-the-art algorithms.

K. Kjærside[1], Hilliges, O., Kim,[4], Kinect has also enabled various interactive applications that are creative and fun, Most relevant to our Interactive Mirror is the ever-growing virtual fitting room systems available on the market, such as Fitnect and TriMirror [4]. However, we have not been able to find any technical details of these systems. From their demo videos alone, the major difference between our system and TriMirror, for example, is that we do not simulate clothes in our system.We simply render the deformed clothes on top of the user's video stream, and this requires a high-quality calibration between the Kinect

III. RELATED WORK

A. Camera Calibration

Vision-based augmented reality systems need to trace the transformation relationship between the camera and the tracking target in order to augment the target with virtual objects. In our virtual try-on system, precise calibration between the Kinect sensor and the HD camera is crucial in order to register and overlay imaginary garments seamlessly onto the 2D HD video stream of the shoppers. Furthermore, we prefer a quick and semi-automatic calibration process because the layout between Kinect and HD camera with respect to the floor plan may be different for different stores, or even for the same store at different times. To this end, we use the CameraCalibrate and StereoCalibrate modules in OpenCV [3] for camera calibration. More specifically, we recommend to collect a minimum of 30 pairs of checkerboard picture seen at the same instant of time from Kinect and HD camera, and calculate each pair's correspondences, as shown in Fig. 7. In addition, the Kinect sensor is usually not perfectly horizontal to the ground plane, and its tilting angle is needed to calculate the height of user.We simply specify the floor area from the Kinect depth data manually, and the normal vector of the floor plane in Kinect's view can be estimated.

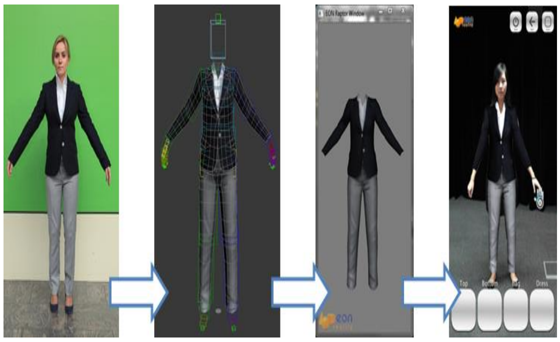


Figure 2. Major steps for content creation. Catalogue images are first manually modeled and textured offline in 3DS Max. We then augment the digital clothes with relevant size and skinning information. At runtime, 3D clothes are properly resized according to a user's height, skinned to the tracked skeleton, and then rendered with proper camera settings. Finally, the rendered clothes are com bind with the HD recording of the user in realtime.

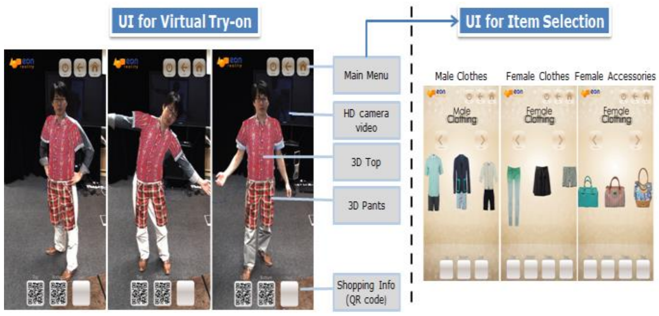


Figure 3. Left: the UI for virtual try-on. Right: the UI for clothing item selection. To summarize, the result of the camera calibration procedure adds: – extrinsic camera parameters (translation and rotation) of the HD camera with respect to the Kinect depth camera. – the tilting angles of the Kinect sensor and the HD camera with respect to the horizontal ground plane. – FoV of the HD camera.

B. Content creation

Our virtual 3D clothes are based on actual catalogue images, so that new fashion lines can be included to the system quickly. Fig. 4 shows the major steps of converting catalogue images to 3D digital clothes. In the preprocessing stage, our artists manually generatedone standard digital male mannequin and one female mannequin. Then they modeled the catalogue images into 3D clothes that fit the proportions of the default mannequins. Corresponding textures were also adapted and applied to the digital clothes. Then we augment the digital clothes with relevant size and skinning data Atruntime,3D

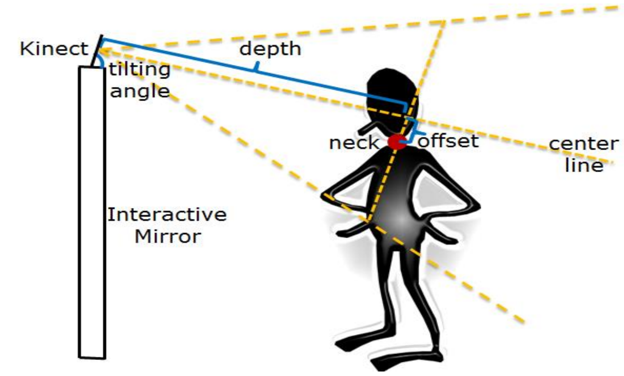
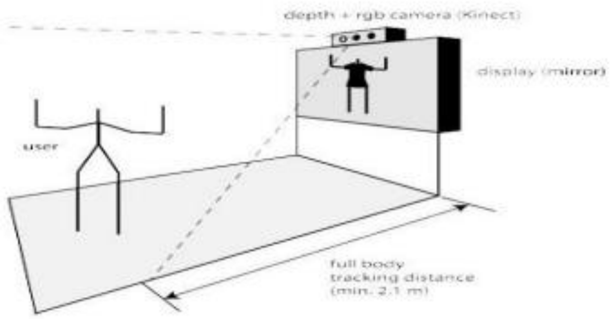


Figure 4. Shoulder height estimation when the user's feet are not in the field of view of Kinect.

The tilting angle of the Kinect sensor, the depth of the neck joint, and the offset of the neck joint with respect to the center point of the depth image can jointly determine the physical height of the neck joint in the world space. clothes are properly resized according to a user's height, skinned to the tracked skeleton, and then rendered with proper camera settings. Lastly, the rendered clothes are merged with the HD recording of the user in realtime. Our content development team modeled 115 clothing items in total, including male clothes, female clothes, and accessories. On average it took about two man days to create and test one item for its inclusion into the virtual try-on system.

C. User interface

Fig. 5 depicts the user interface of the Interactive Mirror. Because our clothes are 3D models rather than 2D images, users are able to turn their body within a reasonable range in front of the Interactive Mirror and still have the digital clothes properly fit to their body, just like what they can see in front of a real mirror. The user selects menu items and outfit items using hand gestures. Different tops, bottoms, and accessories can be added and matched on the fly.



D. Height estimation

Digital clothes need to be rescaled according to users' body size, for good fitting and try-on experiences. We propose two methods to estimate a user's shoulder height. The first one simply uses the neck to feet height difference, when both the neck and the feet joints are detected by Kinect skeletal tracking SDKs.

IV. PROPOSED SYSTEM

Our virtual try-on system consists of a vertical TV screen, a Microsoft Kinect sensor, an HD camera, and a desktop computer. Fig. 6 shows the front view of the Interactive Mirror together with the Kinect and HD camera. The Kinect sensor is an input device marketed by Microsoft, and intended as a gaming interface for Xbox 360 consoles and PCs. It consists of a depth camera, an RGB camera, and microphone arrays. Both the depth and the RGB camera have a horizontal viewing range of 57.5 degrees, and a vertical viewing range of 43.5 degrees. Kinect can also tilt up and down within -27 to +27 degrees. The range of the depth camera is [0.8_4]m in the normal mode and [0.4_3]m in the near mode. The HD camera supports a full resolution of 2080 _ 1552, from which

Figure 6. The front view of the Interactive Mirror with Kinect and HD camera placed on top.

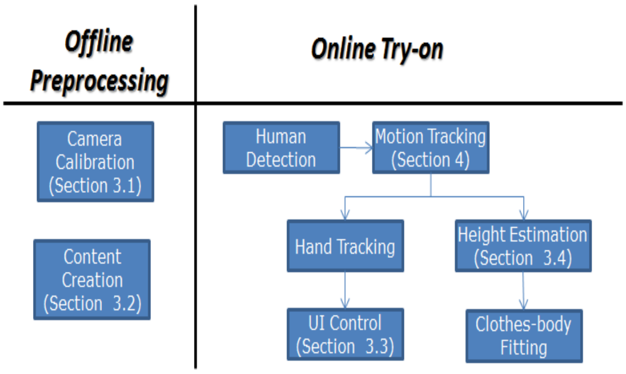


Figure 7. Major software components of the virtual try-on system.

Fig. 7illustrates the major software components of the virtual try-on system. During the the offline preprocessing stage, we need to calibrate the Kinect and HD cameras, and create 3D clothes and accessories. During the online virtual try-on, we first detect the nearest person among the people in the area of interest. This person will then become the subject of interest to be tracked by the motion tracking component implemented on two publicly available Kinect SDKs, as will be discussed in Section 4. The user interacts with the Interactive Mirror with her right hand to control the User Interface (UI) and select clothing items. The UI layout will be discussed in more details

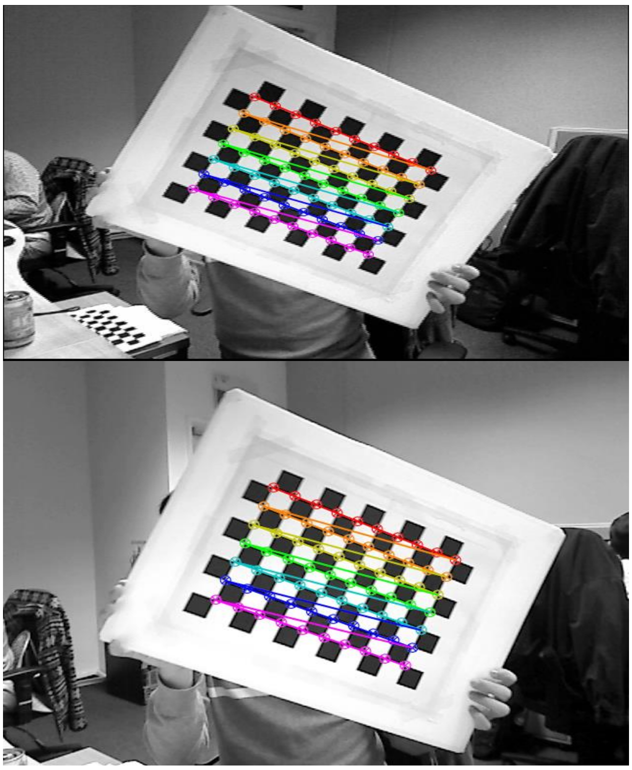


Figure 8. The camera calibration process. The checkerboard images seen by the Kinect RGB camera (left) and the HD camera (right) at the same instant of time.

Proposed System Screenshots

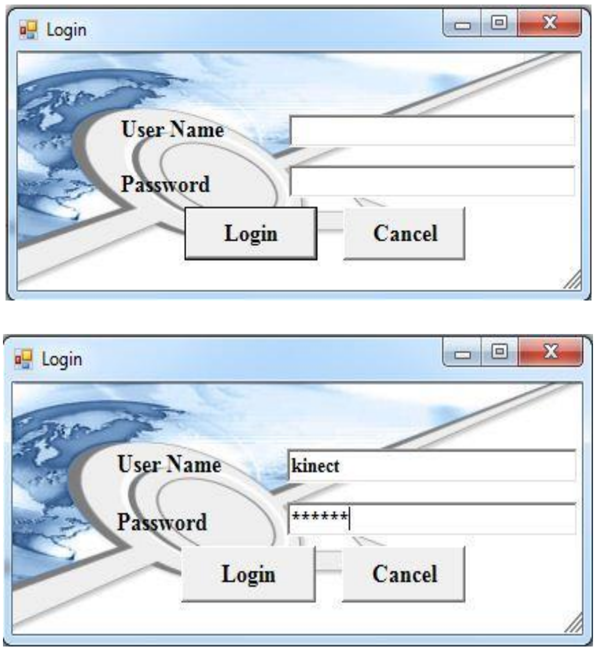


Figure 9. Login Window

Figure 10. Splash Window

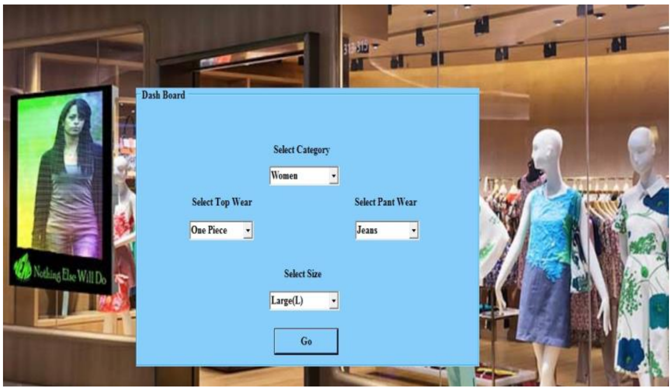


Figure 11. DashBoard

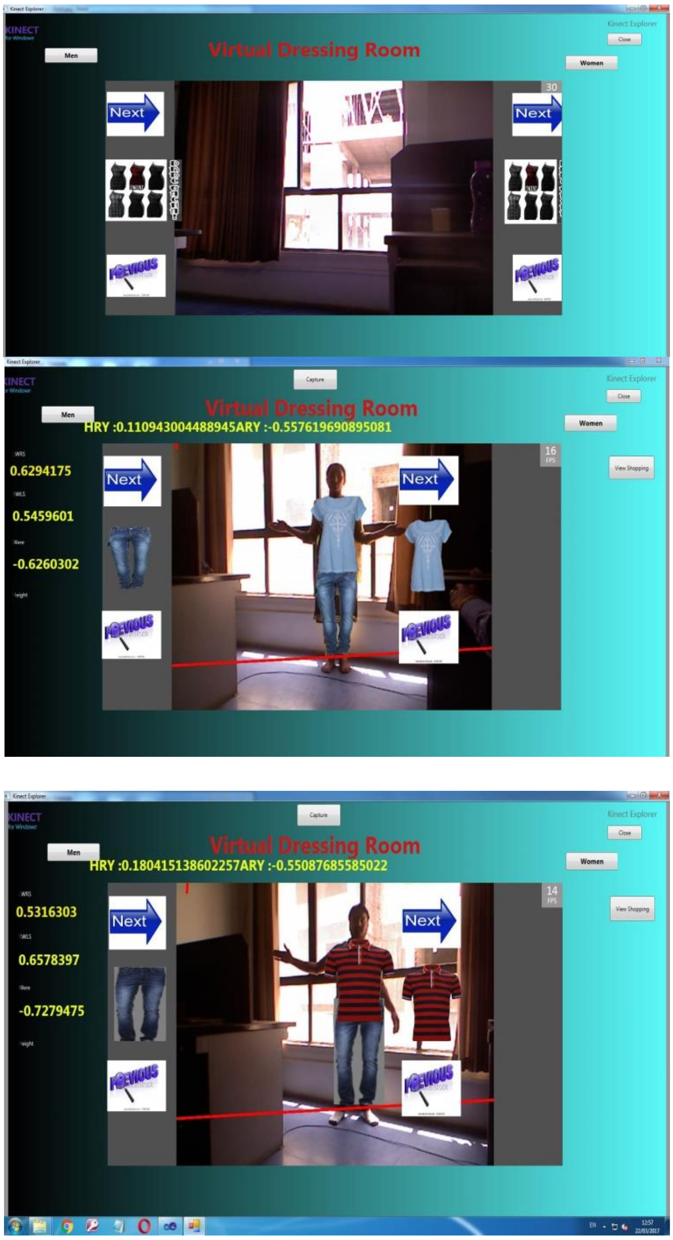


Figure 12. Dressing Room

V. EXPRIMENTAL RESULTS

We evaluated the performance of the application on a set of 12 poses with different angles of rotation and distance from the sensor (Fig. 13). We measured the performance as the amount of overlap between the constructed cloth model and manually labeled ground truth data as $P = \frac{A_c \cap A_g}{A_c \cup A_g}$ (6) where Ac is the area of the constructed model and Ag is the area of the ground truth model in terms of the number of pixels. We observed an average overlap of 83.97 or higher between the ground truth and the computed models, within a rotation range of 0−45◦ . Rotations along the vertical axis dropped the performance as the fitting is performed in only 2-dimensions. The best results were obtained when the distance from the sensor was 2 meters. The results are summarized in Table 1.

Figure 13. A set of poses that are used in the evaluation of the performance

Arm Rotation	**0°**	**45°**	**90°**
Performance	89.80%	90.54%	**76.64%**
Body Rotation	**45°**	**0°**	**45°**
Performance	83.68%	90.19%	88.84%
Horizontal Rotation	**-45°**	**0°**	**45°**
Performance	74.88%	87.74%	77.87%
Distance From Sensor	**1.5m°**	**2m°**	**3m°**
Performance	86.64%	89.87%	70.98%
Average Overlap	**0°** 83.97%		

TABLE 1 Experimental Results

VI. CONCLUSION AND FUTURE WORK

EON Interactive Mirror offers several advantages over traditional retailing. It attracts more customers through providing a new and exciting retail concept, and creates interest in the brand and store by viral marketing campaigns through customers sharing their experiences in Social Media such as Facebook. Furthermore, it reduces the need for floor space and fitting rooms, thereby reducing rental costs and shortening the time for trying on different combinations and making purchase decisions. We encourage interested readers to search our demo videos with keywords EON Interactive Mirror at http://www.youtube.com.

We developed a real-time virtual dressing room application that requires no visual tags. We tested our application under different conditions. Our experiments showed that the application performs well for regular postures. The application can be further improved towards creating more realistic models by using 3D cloth models and a physics engine.

VII. REFERENCES

[1]. K. J. Kortbek, H. Hedegaard, and K. Grønbæk, "Ardresscode: augmented dressing room with tag-based motion tracking and real-time clothes simulation," in Proceedings of the Central European Multimedia and Virtual Reality Conference, 2005.

[2]. J. Shotton, T. Sharp, A. Kipman, A. Fitzgibbon, M. Finocchio, A. Blake, M. Cook, and R. Moore, "Real-time human pose recognition in parts from single depth images," Communications of the ACM, vol. 56, no. 1, pp. 116–124, 2013.

[3]. Benko, H., Jota, R., Wilson, A.D.: Miragetable: Freehand interaction on a projected augmented reality tabletop. In: CHI'12 (2012)

[4]. Hilliges, O., Kim, D., Izadi, S., Weiss, M., Wilson, A.D.: Holodesk: Direct 3d interactions with a situated see-through display. In: Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems. CHI '12 (2012)

[5]. Izadi, S., Newcombe, R.A., Kim, D., Hilliges, O., Molyneaux, D., Hodges, S., Kohli, P., Shotton, J., Davison, A.J., Fitzgibbon, A.: Kinectfusion: real-time dynamic 3d surface reconstruction and interaction. In: SIGGRAPH 2011 Talks. p. Article 23 (2011)

[6]. Kim, K., Bolton, J., Girouard, A., Cooperstock, J., Vertegaal, R.: Telehuman: effects of 3d perspective on gaze and pose estimation with a life-size cylindrical telepresence pod. In: CHI'12. pp. 2531–2540 (2012)

[7]. Schwarz, L.A., Mkhitaryan, A., Mateus, D., Navab, N.: Estimating human 3d pose from time-of-flight images based on geodesic distances and optical flow. In: IEEE Conference on Automatic Face and Gesture Recognition (FG). pp. 700–706 (2011)

[8]. Shotton, J., Fitzgibbon, A.W., Cook, M., Sharp, T., Finocchio, M., Moore, R., Kipman, A., Blake, A.: Real-time human pose recognition in parts from single depth images. In: CVPR. pp. 1297–1304 (2011)

[9]. Weise, T., Bouaziz, S., Li, H., Pauly, M.: Realtime performance-based facial animation. ACM Trans. Graph. 30(4), Article 77 (2011)

[10]. Ye, M., Wang, X., Yang, R., Ren, L., Pollefeys, M.: Accurate 3d pose estimation from a single depth image. In: ICCV. pp. 731–738 (2011)

[11]. Zhu, Y., Dariush, B., Fujimura, K.: In: CVPRW'08, IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshop