Experiences in the Development of an Augmented Reality Dressing Room

Ugo Erra and Valerio Colonnese

Dipartimento di Matematica, Informatica ed Economia,
Università degli Studi della Basilicata,
Campus di Macchia Romana, Potenza, 85100, Italy
{ugo.erra,valerio.colonnese}@unibas.it

Abstract. Augmenting reality technology is changing the face of retail and altering consumer purchasing decisions wherever it is used, whether on home computers, mobile devices, or in kiosks in stores. This is because the technologies behind these applications are becoming increasingly affordable and sufficiently cheap that developers are exploiting their greatest potential to bring about a revolution in the way purchases are made. In this paper, we describe our experiences of designing an augmented reality dressing room where 3D models of the clothing are overlaid with a camera-captured color image to achieve the function of a virtual mirror. In this way, customers can move around to decide whether the clothing suits and fits them well. The project is implemented in Unity 4 Pro in combination with Microsoft Kinect 2 for the tracking process. In this paper, we discuss design issues and technical implementation and prospects for further development of the techniques.

Keywords: Augmented reality, human body tracking, dressing room

1 Introduction

Today, shopping has become a highly popular and time-consuming activity. For some people, clothes shopping is a leisure activity, whereas for others it is a stressful and tedious activity because of the guesswork involved in deciding whether a set of clothing suits and fits them well. Experiments with so-called augmented reality dressing room technology have been around for years. In these experiments, customers can try on clothing to check the fit or the style virtually rather than physically and make their purchases quickly and easily. These approaches implemented with augmented reality reflect the current state of the art in human body tracking technologies.

Recent developments in sensor technologies allow accurate and robust human body tracking, enabling a new set of possibilities. In particular, entertainment technologies such as Microsoft Kinect [1] and other related devices have lowered the barrier to entry to these possibilities in terms of both cost and development complexity. However, the technology behind Microsoft Kinect has gone much further than gaming and has been used in many contexts in very creative ways, ranging from advanced user interfaces to high-quality 3D scans, and even as a tool for stroke recovery.

Therefore, the goal of the present study was to develop an augmented virtual dressing room application that can run on a common desktop personal computer equipped with an off-the-shelf Microsoft Kinect 2 device. This application can enhance the way customers shop and help them to choose the correct type of item of clothing. The proposed approach is designed to be computationally efficient, and it can be used with cheap hardware. The major benefits from this type of application are expected to be: 1) an improved ability to make a correct buying decision, reducing the time required; 2) increasing opportunities for fashion designers to experiment creatively; and 3) parallel use for other goods and accessories such as jewelry, spectacles, handbags, and shoes.

This paper describes our experiences during the design, with a discussion of the technical solutions that we chose for our approach. A sample application with a user interface was developed to test practicality and performance. The user interface allows the user to choose an item of clothing by making a hand movement toward it. Preliminary results indicate that this is a promising approach, and further design experiments are needed.

2 Related Works

In the past, several approaches have been implemented for augmented reality dressing rooms. These reflect the current state of the art in human body tracking technologies. Three lines of research are briefly discussed in this section: image processing, fiducial marker, and hardware-based tracking.

Martin and Erdal [2] present an image-processing design flow for visualizing an augmented dressing room and that is designed to be compatible with a common webcam. The software is implemented by a three-stage algorithm: detection and sizing of the user's body, identification of reference points based on face detection and augmented reality markers, and superimposition of the clothing over the user's image. The limitations of this tracking method means that clothes can be superimposed only as 2D images. A similar approach is described in [3].

Fiducial marker-based tracking involves the automatic detection of patterns in digital images that are taken with a camera. Kjrside et al. [4] propose a tag-based approach that requires the manual labeling of body parts with one or more markers. The video frames received from the camera are analyzed in real time using image-processing techniques to determine the 3D position and orientation of the markers and then to create an augmented reality of the customer wearing the clothing. A similar approach is presented in Araki and Muraoka [5]. The capturing of a person is accomplished by positioning small colored markers on the user's joints. The markers have different colors according to the placement on the body. A negative point of this approach is that the user cannot be captured from the side. In addition, more generally, the printed marker pattern must be placed on the user's body, which may be time-consuming and cumbersome from the consumer's point of view. The manual labeling of body parts with tags may also be a source of error.

Tracking hardware allows for a more accurate and robust solution that enables the investigation of various augmented reality dressing approaches. The solution proposed by Furkan and Gokcehan [6] is based on a 2D model that is scaled based on the distance between the user and the Microsoft Kinect sensor and then overlaid with the video image. However, this technique is designed specifically for t-shirts, and the researchers do not show the treatment in use for other items of clothing. A similar approach based on depth data from the Microsoft Kinect is presented in Presle [7]. Recently, several commercial applications have appeared based on the Microsoft Kinect. Such virtual dressing rooms are available from, for example, FaceCake [8] or Fitnect [9].

Our work also attempts to exploit body tracking technologies based on Microsoft Kinect. However, a key aspect of our solution that differs from other approaches is the use of 3D clothing with skeleton animation. This aspect allows the user complete freedom of the room and adds movement flexibility by utilizing smooth continuous tracking.

3 Fundamentals

This section provides a basic outline of the fundamentals of our approach, in particular for the two technologies used in our approach: Microsoft Kinect 2 (also known as Kinect for Xbox One) and Unity 3D.

Microsoft Kinect 2 [10] is a device for the Xbox One game console that allows users to control and interact with games through a natural user interface by using gestures and spoken commands. Microsoft Kinect 2 has a cone-shaped tracking area of 70°. The user comes into full view of the Microsoft Kinect 2 camera at approximately 1.4 m. At distances closer to the camera, only partial skeleton tracking is possible, and the end of the Microsoft Kinect 2 tracking range is 4.2 m from the camera. At its closest full body tracking range, the user can move up to 1 m to each side of the camera. At its furthest range, the user can move up to 2.9 m from each side of the camera. This results in a total tracking range of just over 10 m^2 [11]. The tracking process of the Microsoft Kinect 2 is based on retrieval of the particular body joint positions [12]. This algorithm allows real-time detection and tracking of the user's skeleton in a stable and efficient way. Moreover, the algorithm allows a full rotation of the body and a robust distinction between the left and right side of the user's body.

Unity 3D [13] is a feature-rich, fully integrated development engine that provides out-of-the-box functionality for the creation of interactive 3D content. With Unity, this content can be published on multiple platforms such as personal computer, Web, iOS, Android, and Xbox, and allows the augmented reality dressing room in combination with the Microsoft Kinect to be executed on several operating systems. Unity's complete toolset, intuitive workspace, and on-the-fly play testing and editing feature can save developers time and effort. Unity enables developers to extend its functionality by using platform-specific native code li-

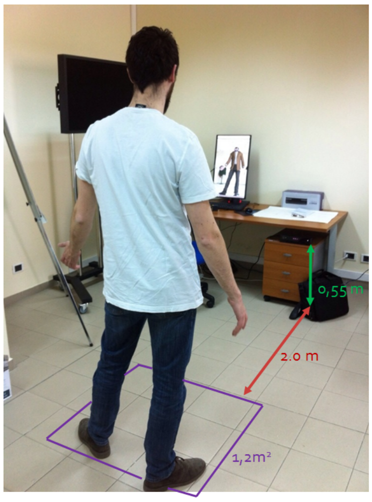

Fig. 1. A proposed basic setup of the dressing room. It consists of the Microsoft Kinect 2 device, a vertical display, and a computer. The person in front of the Microsoft Kinect is interacting at a certain distance and inside a limited area. The green line indicates the appropriate height placement of the Microsoft Kinect.

braries called native plugins. These allow access to features such as OS calls and third-party code libraries that would otherwise not be available in Unity. Through these plugins, the Microsoft Kinect API set can be made available in Unity Pro (the commercial full version), giving developers full access to the Microsoft Kinect core functionality. In particular, the plugins enable vision detection and tracking functionality within Unity and allow developers to create augmenting applications and games relatively easily [1].

4 Design

The idea behind our augmented reality dressing room is to allow users to try on clothes virtually in front of a vertical large screen, which acts as an augmented mirror, so that the users can quickly decide whether the clothing fits them physically and aesthetically. In this way, customers can try on more clothes in less time. This may improve the mood of customers and help to affect their decision to buy the clothing or try it on for real.

Most approaches based on body tracking involve mapping a 2D texture as cloth to the user's body. When the user moves around, the clothing does not accurately capture the user's position and movement, therefore causing several unaesthetic effects. To achieve a more realistic simulation of the process of dressing, we based our approach on the adoption of 3D clothing models. This approach has several advantages. First, it does not make any assumption of the user's dimensions (body shape, height, width, length of limbs, etc.) from the

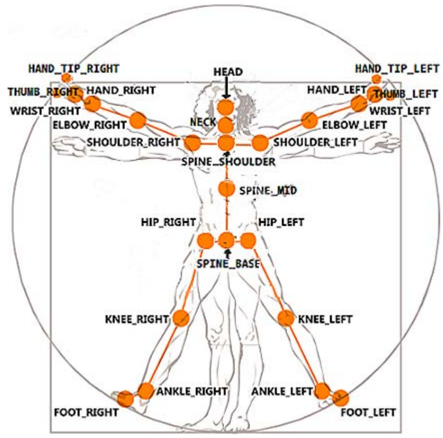

Fig. 2. The 20 joints that make up a Microsoft Kinect skeleton [14].

data captured by Microsoft Kinect and therefore does not require a prior 3D scan. Second, the full 3D clothing model will always follow the motion of the user captured by the Microsoft Kinect. As the user moves around, the Microsoft Kinect will capture the skeleton tracking information, which will be mapped onto the 3D clothing model. In this way, the clothes will match the movements of the user. Therefore, a realistic simulation of dressing is achieved by the interaction between the skeleton of the user and the virtual skeleton of the 3D clothing model.

To use this approach, we require the editing of a skeleton animation in the 3D clothing model. Skeleton animation is a well-known technique used in computer animation in which a character is represented in two parts: a surface representation called a mesh used to render the character, and a hierarchical set of interconnected bones called the skeleton. Each bone in the skeleton is associated with some portion of the mesh's visual representation. In this way, the movement of a portion of the skin is influenced by one or more associated bones [15]. The rigging is the process of constructing the series of bones used to animate the mesh. Usually, a 3D computer graphics program is used to provide a default skeleton to animate humans, for example, 3D Studio Max, which is the program we used for our implementation. The modeler must place the joints exactly where they would be in a real-world skeleton and associate the bones with the mesh (Figure 3). Although this technique is often used to animate human characters, it can be used to control the deformation of any object. In our case, the rigging is easy because we can use the default skeleton provided for human characters, but we only need to associate the bones that are bound to the mesh of the 3D clothing model (left and right foot bones are never used). For example, in the case of a long skirt, we require only the spine base and right and left knees.

The Microsoft Kinect SDK 2.0 provides information about the location of users standing in front of the Microsoft Kinect sensor array with detailed position and orientation data. These data are input into the application code as a set of 20 body joints (Figure 2), namely skeleton position. These joints are used to locate the parts of the 3D clothing model and therefore represent the user's current position and pose. The application can therefore use the skeleton data to measure various dimensions of users' parts and control. In particular, we can use the Euclidean distance from the head to one of the ankles to estimate the user height and the distance between the left and right shoulders to estimate user width. The skeleton data are retrieved with the abovementioned image retrieval method—calling a frame retrieval method and passing a buffer—while our application can then use an event model by hooking an event to an event handler to capture the frame when a new frame of the skeleton data is ready.

Our approach can be summarized as follows: 1) extraction of the user tracking information from the video stream; 2) positioning of the 3D clothing model by using the skeleton tracker of the Microsoft Kinect SDK; 3) scaling of the model by using the Euclidean distance between the body joints and the user's distance from the sensor; and 4) superimposition of the 3D clothing model on the user.

Because the clothing model is in 3D, the application allows horizontal rotation of the user. Thus, users can perform a full rotation in front of the augmented mirror to see their front and back. We tested this action and found that the body joints are adequately detected within the distance range of approximately 2–3.2 m (see Figure 1). A drawback of this approach is that the 3D model is superimposed on the top layer and the user always stays behind the model. This causes some inevitable artifacts for some types of clothing when users perform certain actions such as folding their arms.

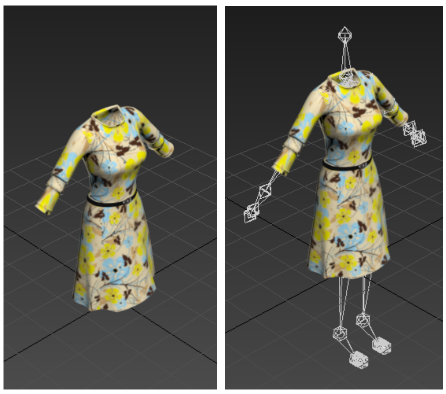

Fig. 3. A 3D item of clothing without (left) and with (right) the virtual skeleton.

Figure 4 shows the user interface elements in our application with some examples of clothing. The button on the upper right side can be used to force the application to re-acquire the user's dimensions. The buttons on right and left central sides are functions for the selection of previous and next item of clothing. A hand position indicator shows the current coordinates of the hand on the screen, similar to a mouse. This facilitates user interaction because it allows users to switch easily from one item of clothing to another by simply holding their hand above the elements for two seconds.

Fig. 4. The user interface of the application and some examples of clothing.

5 Conclusions

In this paper, we have described our experiences during the design of an augmented reality dressing room. We introduced our approach, which combines the visualization capabilities of the game development tool Unity Pro with the position and body tracking capabilities of Microsoft Kinect 2. The overall system software does not require calibration, is inexpensive, and is easy to use. The low cost and ease of use makes it accessible to a wider group of vendors who do not normally have access to a professional augmenting reality facility.

Compared with other approaches to augmenting reality dressing rooms, our approach offers good visualization results in terms of appearance. Moreover, our approach allows customers to move around relatively freely with an item of clothing because the user's position and movement are captured with skeleton mapping. At the same time, the initial setup in terms of the design of the 3D clothing model is relatively inexpensive. Although it would mean that every piece of clothing in a vendor's inventory would have to be modeled and textured and added to the 3D repository in advance, several ad-hoc applications are available for designing clothing, for example, Marvelous Designer [16].

In future work, our augmented reality dressing room may be enhanced in some areas. First, we will conduct an empirical evaluation of the proposed system to investigate customer satisfaction in relation to service quality. The illumination of clothes by the lighting conditions of the real-world captured camera images could be used to enhance the realism of the rendering. Another possible improvement is the implementation of a complete 3D scanning procedure of the clothing to enable a new item to be added into a 3D repository more quickly and easily.

Acknowledgement

Work in this paper has been funded by Basiliata Innovazione under the agreement "Studio di fattibilità finalizzato allo Sviluppo di una Vetrina 3D basata sulla Realtà Aumentata" in 2014.

References

1. Kinect SDK 2.0. `https://www.microsoft.com/en-us/kinectforwindows/develop/`.
2. Cecilia Garcia Martin and Erdal Oruklu. Human Friendly Interface Design for Virtual Fitting Room Applications on Android Based Mobile Devices. *Journal of Signal and Information Processing*, 3(4):481–490, November 2012.
3. A. A. Shaikh, P. S. Shinde, S. R. Singh, S. Chandra, R. A. Khan. A Review on Virtual Dressing Room for e-shopping using Augmented Reality. *International Journal of Soft Computing and Engineering (IJSCE)*, 4(5):98–102, November 2014.
4. Krista Kjrside, Karen Johanne Kortbek, Henrik Hedegaard, and Kaj Grnbk. ARDressCode: Augmented Dressing Room with Tag-based Motion Tracking and Real-Time Clothes Simulation. In *Central European Multimedia and Virtual Reality Conference*, 2005.
5. N. Araki and Y. Muraoka. Follow-the-trial-fitter: Real-time dressing without undressing. In *Digital Information Management, 2008. ICDIM 2008. Third International Conference on*, pages 33–38, Nov 2008.
6. Furkan Isikdogan and Gokcehan Kara. A Real Time Virtual Dressing Room Application using Kinect, 2012. Computer Vision Course Project.
7. Philipp Presle. A Virtual Dressing Room based on Depth Data. Master's thesis, Institut für Softwaretechnik und Interaktive Systeme, 2012.
8. Facecake marketing technologies, inc. `http://www.facecake.com/`.
9. Fitnect. `http://www.fitnect.hu/`.
10. Kinect for Windows features. `https://www.microsoft.com/en-us/kinectforwindows/meetkinect/features.aspx`.
11. Stefan Greuter and David J. Roberts. Spacewalk: Movement and interaction in virtual space with commodity hardware. In *Proceedings of the 2014 Conference on Interactive Entertainment*, IE2014, pages 1:1–1:7, New York, NY, USA, 2014. ACM.
12. Jamie Shotton, Andrew Fitzgibbon, Mat Cook, Toby Sharp, Mark Finocchio, Richard Moore, Alex Kipman, and Andrew Blake. Real-Time Human Pose Recognition in Parts from a Single Depth Image. In *CVPR*. IEEE, June 2011.
13. Unity Game Engine. `https://unity3d.com/`.
14. Tracking Users with Kinect Skeletal Tracking. `https://msdn.microsoft.com/en-us/library/jj131025.aspx`.
15. Rick Parent. *Computer Animation: Algorithms and Techniques*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 3 edition, 2012.
16. Marvelous designer. `http://www.marvelousdesigner.com/`.