The Performance of Self in the Context of Shopping in a Virtual Dressing Room System

Yi Gao, Eva Petersson Brooks, and Anthony Lewis Brooks

Centre for Design, Learning and Innovation,
Department of Architecture and Media Technology,
Aalborg University, Niels Bohrs Vej 8, 6700 Esbjerg, Denmark
gao@create.aau.dk

Abstract. This paper investigates the performance of self in a virtual dressing room based on a camera-based system reflecting a full body mirrored image of the self. The study was based on a qualitative research approach and a user-centered design methodology. 22 participants participated in design sessions, semi-structured interviews and a questionnaire investigation. The results showed that the system facilitated self-recognition, self-perception, and shared experience, which afforded an enriched experience of the performing self.

Keywords: Virtual dressing room, mirroring, self-perception, self-recognition, shared experience, hedonic shopping experience.

1 Introduction

In this paper, we investigate how a virtual dressing room facilitated the performance of self in terms of self-recognition and self-perception to identify how potential shoppers organize such interactions and experiences. In line with [1] online shopping is considered as a creative and social activity incorporating diverse meanings where both shopping and commodities invoke personal as well as collective interests and motivations. The study illustrates in different ways that the virtual dressing room system user interface invoked certain interactions that are afforded by the technique of mirroring in line with [2].

This on-going research is financed by the Danish National Advanced Technology Foundation to realize a turnkey solution of a Virtual Dressing Room (VDR), which should reduce customer purchase returns. In this regard, the practice of shopping clothes online is considered as framed by shoppers through the influence of affordances and personal agency. This is in line with [3], who states that:

> "When the behavior of the computer is coherent and the application is designed so that a human interactor knows what to do and receives clear and immediate feedback on the results of their actions, the interactor experiences the pleasure of agency, of making something happen in a dynamically responsive world." [3, p. 100].

F.F.-H. Nah (Ed.): HCIB/HCII 2014, LNCS 8527, pp. 307–315, 2014.

2 Background

A virtual dressing room, in which shoppers cannot only view apparel commodities but also view as an overlay on their bodies, addresses many of the concerns that shoppers have about purchasing apparel online. The Forrester report [4] states how customers prefer shopping in stores as this offers possibilities to touch and feel the items and, also, avoiding issues related to returns. 56% of shoppers considered lack of touch and feel as a primary concern when it comes to online shopping. Currently, online garment purchases are subject of approximately 30% returns [5]. Forrester research [6] and [7] claims that approximately 35% of online garment clothes are returned. This is costly for the e-commerce retail outlet and creates logistical problems in re-selling returned stock. We contribute to this field by introducing a turnkey solution that allows e-commerce clothing retailers to create digital clothing by scanning real clothes using a RGB-D sensor. Creating digital clothing using specialized programs such as Marvelous Designer 2 is a time consuming process, whereas the 3D scanning process is relatively fast [8]. With this kind of virtual dressing room solution, shoppers will view accurate models of the self, wearing clothing in different colors and sizes and receive information about how the garments fit [9]. [10] identified constraints and affordances related to a previous prototype of the virtual dressing room system involved in this paper, focusing on ludic activities, where access, movement based interaction, social activity, self-image recognition, and authority constituted core factors to be enhanced in order to provide a more hedonic shopping experience. The freedom to freely move was a required affordance, which enabled expressions of self and fostered ludic engagement.

Another body of work that intends to address the problem of returns on purchased-online apparel is Fits.me, a company that utilizes a robotic mannequin created by a consortium involving roboticists from the University of Tartu in Estonia, who enlisted the help of Human Solutions, a German firm specializing in body dimensions and ergonomic simulation. The mannequin has 50 actuators embedded, which push panels in and out to form different shapes. It is covered in Pedilin, a material used on prosthetic limbs, so fabrics will drape like they would on real skin. Fits.me collects user's data to compile a database of customers, which each individual can use to view clothed in chosen apparel. However, the company also plans to sell retailers garnered information on shoppers, which could help outfitters design clothing and target sales.

The VDR system included in this study shows promising results of creating digital clothing of high quality by scanning real clothes. The paper discusses the major affordances and constraints originating from apparel online shopping in a virtual dressing room and addresses how shoppers experiences derives from the concept of mirroring [2]. However, before the discussion, it is necessary to detail the system and the design concepts.

3 Analysis of Design Concepts: Mirroring for the Virtual Dressing Room

The catalyst of the Virtual Dressing Room (VDR) is unencumbered interaction with a computer-based system. Unencumbered interaction relates to the field of Natural User Interfaces (NUIs) in that a participant interacts with a system via gesture without wearing, holding or otherwise operating a tangible device. The latest gesture-control peripherals for games are based on such interactions that are a direct result from the advent of camera-based software and related advances in computer-vision and hardware such as computer graphic cards.

In our VDR project, two systems were initially explored to establish the required tracking of gesture matched to affordability as the work targets a commercially viable product for the general public. The two systems that were explored spanned a wide cost range in order to determine if a more costly solution would be justified in the final product. These systems are (1) the Microsoft Kinect, and (2) the OpenStage® V2.0 Markerless Motion Capture System by Organic Motion.

3.1 The Microsoft Kinect

The Kinect is a well-known affordable video game camera-based peripheral operating at 30 frames per second that enables gesture control of software via predictive algorithm mapping in a scan area 1.2–3.5 m and with a latency of 250 milliseconds. The device uses a “time-of-flight camera” for its range imaging. Such technology enables the tracking of humans through providing distance images in real time. The cost of the stand-alone device is approximately 150 UK pounds.

Fig. 1. 1st generation Kinect device sold with Microsoft X-Box (*source: Microsoft*)

The second-generation of the Kinect can be plugged into a standard PC (so the need for the X-Box gaming platform module is redundant) where a SDK is used to program the interactions. The system is mobile for indoors and outdoors use (but a power supply is needed). It also operates through clear glass, thus, allowing shop-frontage use 24/7.

The most-recent Kinect is a self-contained platform that can also accept spoken commands via a microphone array. It uses a motion-sensing camera that tracks the whole body, thus, when one interacts it is about more than solely a hand or wrist gesticulation as arms, legs, knees, hips, and torso are all also involved.

When one interacts, a digital skeleton is created that is based upon the depth data received, which enables the player’s integration into the specific software, be it a game or a system such as the VDR. The Kinect ID has an in-built memory that has the capability to ‘remember’ each player from the physical information stored in the game console or VDR system. Thus, after an initial set-up and profile calibration, each session can begin straight away through an ID configuration recall. Voice recognition is via four in-built microphones that can determine room noise or player verbal commands. 24/7. The Kinect sold 18 million units in 2011. The built in biofeedback for the player advances this field.

In summary, find below an overview of the Kinect technology:

- Relies on interpretation and predictive algorithms to determine the most probably upcoming pose.
- Uses a library containing a large number of preloaded poses.
- Each detected area in the scan space is matched with a potential body part and assigned the probability that it actually matches such body part.

Based on the above probabilities, Kinect comes up with the most probably skeleton building on its experience and pre-programmed kinematics models. This skeleton is then outfitted as a 3D avatar. In context of the VDR, the avatar may be clothed with selected apparel, cf. [11].

3.2 The Organic Motion OpenStage®

The OpenStage® V2.0 Markerless Motion Capture System by Organic Motion is a multiple-camera (optimal 18 color cameras) full 3D real-time data tracking system (i.e. 360 degrees) that captures movement at an adjustable rate of 60 to 120 frames per second in a scalable scan area of up to 5m by 5m with a maximum latency of 50 to 100 milliseconds. Human bones (21 points) are tracked accurately with each having 5 to 6 degrees of freedom each. Objects can also be tracked in 360 degrees. The OpenStage® is only for indoors use, cf. [11].

Fig. 2. OpenStage® V2.0 Markerless Motion Capture System (*source: Organic Motion*)

OpenStage harnesses Organic Motion's core computer vision technology, enabling computers to cognitively "see" people's complex movements and generate accurate 3D tracking data in real-time: A video sub-system acquires lens and space calibrated video from 8 - 18 cameras and delivers these synchronized streams to the 3D reconstruction processor. The 3D reconstruction system turns the 2D video streams into 3D point and surface clouds by triangulating the various 2D viewpoints. In this way the 3D reconstruction system acts much like a 3D scanner. The final step involves "recognizing" the human figure in this 3D data cloud. Here Organic Motion uses a complex rules based approach that maps a 3D humanoid skeleton into the data.

The output data OpenStage delivers is the X, Y and Z positions and orientation of 21 segments of this skeleton. This information is then ready to be loaded directly via plug-ins or SDK into any form of animation software, game engine, biomechanical or other processing software, all in real-time. Tracking customized objects, non-typical humans, or non-human shapes requires the modelling of new character fitting systems, which OpenStage offers as part of its new software architecture.

Organic Motion sold systems in over 20 countries worldwide, and is used in both commercial and academic settings for multiple applications in various markets including:

- Digital media & arts (animation, game development, VFX).
- Life sciences (bioengineering, physical therapy and rehabilitation, neuroscience, sport).
- Training and simulation (military and defence).
- Public Installations (theme parks, museums).

OpenStage interfaces with various 3D animation systems, 3D game engines and Virtual World systems and 3D immersive visuals, biofeedback and other applications. For therapeutic applications, movement 'rules' may be incorporated to encourage people to be more actively engaged in the recovery regime. This improves outcome and reduces recovery times.

The high accuracy of OpenStage allows clinicians and researchers to identify multi-level movement disorders, develop predictive models of pathology and gather statistical relevant data for long-term improvements, see also [11].

Summary. The Kinect was selected for the VDR system included in this study. This meant a more affordable system that enabled mirroring of the participant, however, problems were apparent with different sized participants – e.g. adults vs. children; and with gesture interactions with the interface (resolution/efficiency). This was apparent through the evaluation with various participants, see [12] and [10]. The next section presents the methodology and methods used in this study.

4 Method

The study applied a qualitative research approach and was based on a user-centered design methodology, including video observations, questionnaires and interviews. In line with a user-centered approach, participants also took part in a design session followed by an interview. The task for the participants was to provide suggestions on improvements for the VDR system prototype, particularly addressing reflections on self-performance in terms of the perception and recognition of the projected self-image.

A total of 22 people were involved in the study. The data was analyzed using an interpretative approach [13], in which the researchers draw on the understanding, and shared perspective of the users, as well as the domain of their actions, to determine the reality of the prototype (VDR system).

The procedure of the study started with that the participants were introduced to the VDR system and the evaluation procedure. They were instructed to freely interact with the system and to try on different pieces of clothes until they experienced they were done. After this, the participants were requested to fill out a questionnaire, to take part in a semi-structured interview, and, finally, to participate in a design session directed towards possible design improvements.

The questionnaires focused on general participant background information as well as preferences when it comes to the 'projected self'. The interviews and the design session were directed towards different aspects of the user interface in terms of usability and user experience; however, the focus of this paper is on the affordances and constraints when it comes to the performance of self. The video observations focused on the user's interaction with the VDR system (the prototype); facial expressions and actions indicating self-recognition and perceptions of self.

The next section discusses the results of the study and concludes with an outlook indicating potential directions of development for a sufficient virtual dressing room system.

5 Analysis and Discussion

It was deemed optimal to use such a sensing device to achieve the desired state-of-the-art capturing of a customer's torso image for further system processing. Using a camera-based device means that the system is non-invasive to enable data-generation from unencumbered motion, i.e. no need to wear, hold or touch any input device. In this way unencumbered gesture controls responsive feedback that is pleasing, direct and immediate so as to digitally mirror input to stimulate the user to further react intuitively in such a way as to become immersed in the interaction. The mirroring technique reinforces a participant's awareness of movement and proprioception and is developed from observations of traditional silver mirroring use in research.

The participants experienced the VDR system as a goal-oriented solution [14] where the navigation was defined and the selection and purchasing of clothes were efficient and convenient, for example by following certain steps and browsing the collection of clothes. However, the selection of clothes was limited, which constrained the performance and, furthermore, the sensorial adventure of exploring different kinds of clothes in terms of textiles, shape, size, and colors. This was in many cases described as closely related to an optimal clothes-shopping event. This addresses an issue concerned with a mismatch of the correlation between the participants' expectations and the computational expectations of the VDR system.

The system provided the participants with a quick access time due to that it simply required them to stand in front of the Kinect (camera) and the clothes were then auto-scaled and applied to the image of the user. Again, this experience was constrained due to the limited access to a variety of clothes, but on the other hand the freedom to move back and forth to judge the fitting of the clothes afforded a sense of exploration value. These results also show that the participants considered that their body could be used to manipulate the system in an interactive mode.

Due to the camera-based system, the participants clearly could recognize themselves via the mirrored image presented on the screen. The fact that it was a full body image seemed to facilitate the performance of self, for example in terms of more movements by the participants in order to express themselves with different clothes and different postures, gestures, and facial expressions. This kind of performance of self refers to a direct self-recognition [cf. 15], which is related to performing actions such as pointing at oneself, or rearranging clothes as a reaction to the captured self-image.

The system user interface provides tracking of multiple participants, which opened up for the participants to try out clothes together with others. For example, a mother brought her son (3 years of age) to try out the system. While doing so, she tried to pick up the virtual clothes placed on her body and pass them on to her son. This action was, then, repeated by the son, who moved the clothes placed on his body back to the mother's body. Two male participants acted in a similar way when they tried on different clothes. If we relate this to the concept of mirroring, which can be defined as a personal experience, but in this paper it relates more to the instances of seeing a projection of the self in a virtual world designed for the motion sensing input device; the Microsoft Kinect. Mirroring in this sense, influenced the self-perception where the self was capable of taking the role of another or, like in the mother-son example, being influenced by someone 'like me' [15, p. 235]. Furthermore, the interaction between the mother and the son as well as between the two males mentioned in the above text, shows that a system providing opportunities for multiple participants to interact when trying out clothes afford shared experiences. This includes a joint performance fostering a hedonic [16] shopping experience.

5.1 Conclusions

This study investigated how a virtual dressing room can facilitate the performance of self in terms of self-recognition and self-perception. The performance of self was fostered by a camera-based system, which efficiently reflected a full body mirrored image of the self. This, in turn, facilitated a range of movements and expressions (postures, gestures, facial expressions) where the participants, for example, rearranged their projected clothes in a direct self-recognition manner. Furthermore, the performance of self was afforded by the mirroring experience, which influenced the self-perception where the self was capable of acting together with another person, who was considered as 'someone like me'. Finally, social aspects such as joint performance, which was not only fostering a hedonic shopping experience, but also afforded an enriched experience of the performing self.

The constraints related to the performance of self in a virtual dressing room were merely related to the technical accuracy of the system, for example the fit between the mirrored body and the tried on clothes. Also, exploring different kinds of clothes in terms of textiles, shapes, sizes and colors constrained the performance of self. These constraints address technical challenges related to the further development of the system, which will be elaborated upon in the following section.

Outlook. Because of the lacking of feel of apparel it is important to make an effort to enhance another sense of the clothing that a person is considering to purchase. Advancing the VDR display could be a glasses-free 3D-Multi-View Portrait format auto stereoscopic display (Fig. 3). In line with the argument presented in the previous section related to the lack of feel of cloth, is that the SKOPOS Institute for market and communication research conducted a study on advertising in 3D in October 2010. In the study, 312 people were divided into two equal groups, each of whom was presented a commercial either in 3D or in 2D. The differences between the test groups were clear: the 3D viewers found the commercial to be modern, original and unique. After the test, 82% of the 3D viewers were convinced of the product. In the 2D group, only 64% were convinced. After the 3D broadcast, viewers also felt more of a desire to try the advertised product; in other words, the purchase probability was significantly greater.

Fig. 3. A 65” glasses-free 3D-Multi-View Portrait format autostereoscopic display[1]

In addition, in both test groups, 43% said they would also like to watch 3D at home. These test results show that there is generally a strong willingness to view 3D content, which can thus be exploited very advantageously for the advertising industry.

In the recent past, 3D technology has experienced a boom, especially in the consumer market. Early in 2010, the Korean display market research company DisplayBank added glasses-free 3D solutions to their list of the most important display innovations for the next ten years. And it’s only a matter of time until

[1] http://www.tridelity.com/fileadmin/user_upload/Produkte/MP5700_3.jpg

companies in all industries adapt to the market changes. Latest advances are 3D glasses-free display solutions in portrait mode – the question is whether this can be a real-time solution for a VDR in boutique or online, to enhance the shopping experience.

Acknowledgements. The Danish National Advanced Technology Foundation who funds the Virtual Dressing Room project. The authors would also like to thank Virtual Lab ApS and Commentor ApS for the VDR prototype.

References

1. Miller, D.: A Theory of Shopping. Cornell University Press, Ithaca (1998)
2. Brooks, A., Petersson, E.: Play Therapy Utilizing the Sony EyeToy®. In: The Proceedings of Presence 2005, London, UK, pp. 303–314 (2005)
3. Murray, J.H.: Inventing the Medium. Principles of Interaction Design as a Cultural Practice. The MIT Press, Cambridge (2012)
4. Mulpuru, S.: U.S. E-commerce Forecast: 2008 to 2012. Forrester Research Report, January 18 (2008)
5. Beck, B.: Key Strategic Issues in Online Apparel Retailing. The Need for an Online Fitting Solution (2000), `http://www.yourfit.com` (retrieved October 31, 2013)
6. Evans, P.F.: Retail Executive Survey, 2008: Using Maslow To Set Site Investment Priorities. Forrester Report, April 1 (2009)
7. Grimaldi, P.: Day of Rejects: Store Clerks Gear Up for Gift Exchanges. Tribune Business News, Washington, December 26 (2008)
8. Holte, M.: 3D Scanning of Clothing using a RGB-D Sensor with Application in a Virtual Dressing Room (under review, 2014)
9. Valente, A., Marchetti, E.: The Elusive Phenomenon of Returns in Online Apparel Shopping. The International Journal of Technology, Knowledge, and Society 10 (2014)
10. Gao, Y., Petersson Brooks, E.: Designing Ludic Engagement in an Interactive Virtual Dressing Room System – A Comparative Study. In: Marcus, A. (ed.) DUXU 2013, Part III. LNCS, vol. 8014, pp. 504–512. Springer, Heidelberg (2013)
11. Brooks, A.L., Czarowicz, A.: Markerless Motion Tracking: MS Kinect and Organic Motion OpenStage. In: Sharkey, P., Klinger, E. (eds.) Proceedings of the 9th International Conference on Disability, Virtual Reality and Associated Technologies, pp. 435–437 (2012)
12. Kristensen, K., Borum, N., Christensen, L.G., Jepsen, H.W., Lam, J., Brooks, A.L., Brooks, E.P.: Towards a Next Generation Universally Accessible 'Online Shopping-for-Apparel' System. In: Kurosu, M. (ed.) HCII/HCI 2013, Part III. LNCS, vol. 8006, pp. 418–427. Springer, Heidelberg (2013)
13. Walsham, G.: Interpreting Information Systems in Organizations. Wiley, Chichester (1993)
14. Parsons, A.G.: Non-Functional Motives for Online Shoppers: Why we Click. The Journal of Consumer Marketing 19(5), 380–392 (2002)
15. Brooks-Gunn, J., Lewis, M.: The development of early visual self-recognition. Developmental Review 4(3), 215–239 (1984)
16. To, P.-L., C. L.-H.: Shopping Motivations on Internet: A Study Based on Utilitarian and Hedonic Value. Technovation, 774–787 (2007)