SAKARYA UNIVERSITY JOURNAL OF COMPUTER AND INFORMATION SCIENCES
VOL. 2, NO. 2, AUGUST 2019
DOI: 10.35377/saucis.02.02.596400

# Controlling of Virtual Mirror with Voice and Hand Motion

Gozde Yolcu Öztel[1], Serap Kazan[2]
[1]Corresponding Author; Department of Computer Engineering, Sakarya University, Sakarya, Turkey, https://orcid.org/0000-0002-6912-0057; gyolcu@sakarya.edu.tr
[2] Department of Computer Engineering, Sakarya University, Sakarya, Turkey, https://orcid.org/0000-0002-3682-0831; scakar@sakarya.edu.tr

Received 24 July 2019; Revised 2 August 2019; Accepted 28 August 2019; Published online 29 August 2019

## Abstract

Although online shopping has been very popular recently, customers have some disadvantages such as the impossibility to try the product before buying. Owing to the recent technological advances, people can try on products to be purchased on a virtual mirror. Inspired by this situation, it has been developed a system which users can try handbags on a monitor and see different poses of handbags as if they wear them. Also, users can see different colors of handbags by saying the color that they want to see to the system. In this application, it is taken advantage of Kinect Sensor to follow user motions and hand skeleton joints which is placed on the bag. The voice command recognition is used which is based on correlation method.

**Keywords:** Virtual mirror, correlation method, Microsoft Kinect Sensor, virtual reality

# El Hareketleri ve Ses ile Sanal Ayna Kontrolü

## Öz

Online alışveriş son zamanlarda popüler olmasına rağmen bu alışveriş türünde ürün deneyememe gibi bazı dezavantajlar mevcuttur. Son zamanlardaki teknolojik gelişmeler ile birlikte sanal aynalar aracılığı ile insanlar satın almak istedikleri ürünleri deneyebilmektedir. Bu yaklaşımdan esinlenilerek, kullanıcıların el çantalarını sanal olarak deneyebildikleri ve farklı pozlarını bir ekran üzerinde görebildikleri bir sistem geliştirilmiştir. Ayrıca, kullanıcılar ses komutları ile çantaların farklı renklerini de deneyebilmektedir. Bu çalışmada, kullanıcı hareketlerinin ve el iskelet bağantılarının takibi için Kinect Sensor kullanılmıştır. Ses komutlarını tanımak için korelasyon yönteminden faydalanılmıştır.

**Anahtar kelimeler:** Sanal ayna, korelasyon metodu, Microsoft Kinect Sensor, sanal gerçeklik

## 1. Introduction

Human computer interaction (HCI) projects have been very popular recently. Natural User Interface (NUI) which is specific type of HCI, provides interaction between machines and human without requiring an additional device on human's body [1]. NUI focuses on natural human abilities, such as vision, touch, speech, handwriting, motion to provide interaction between human and machine [2]. It provides direct and intuitive interaction. Usually NUIs are invisible or they are so unobtrusive that they quickly seem invisible to the user [3]. Touch-scrren interfaces, gesture recognition systems, speech recognition systems, gaze-tracking interfaces are some examples of natural user interfaces.

In the literature several NUI applications have been developed in many areas such as education, entertainment, sport, shopping etc. Echeverria et al. [4] developed a NUI system for teaching Mathematics to students. Their system allows gestural and touch interactions. Cassola et al. [5] developed online gymnastic NUI system. They also used Microsoft Kinect Sensor. Cho et al. [6] developed a system for trying on glasses. Eisert et al. [7] proposed a 3-D system for tracking of shoes. Blum et al. [8] proposed a virtual mirror for anatomy education. They presented a new interaction metaphor that makes use of the depth camera. Casas et al. [9] proposed a system for mental health treatment. Nikam et al. [10] proposed a system for shopping. Ishikiriyama et al. [11] proposed a system

to support makeup for visually impared people. In our previous study, we developed a system [12] which people can try virtual garments on a monitor. We used Kinect Sensor for this application. When user moves his/her arms or legs, different poses of the garments are showed on the virtual mirror. Thus it increased sense of virtual reality. In another previous study [13] we developed a NUI application for teaching painting and basic geometric shapes to children. We used Kinect Sensor Skeleton Map to provide NUI. In the application a child can draw and paint pictures with hand motions. Owing to the system children can learn by having fun.

This study presents a NUI system for shopping area. While online shopping has advantages such as saving time, it has some disadvantages such as impossibility of trying the product. So there is a need to develop a system for trying on virtual products before shopping. In the proposed system a user can try handbag. User can interact with the system by hand motions and voices He / she can select a bag using his/ her right hand, see different colors of the bag by saying the related color. Owing to the proposed system customers can try on products during online shopping. It provides natural, easy and funny way to try on virtual handbags.

Remaining of the paper is structured as follows. Section 2 describes the methology, Section 3 presents experimental results. Section 3 concludes the paper.

## 2. Methodology

In this study, Kinect Sensor is used to detect human skeleton joints. Owing to this skeleton joints users can iteract with the system by hand movements and the handbags are localized on human body. Also users can interact with the system with voice commands. Thus he/she can see different colors of selected handbags. In order to recognize voice commands correlation method are used. The design steps are illustrated in Figure 1.

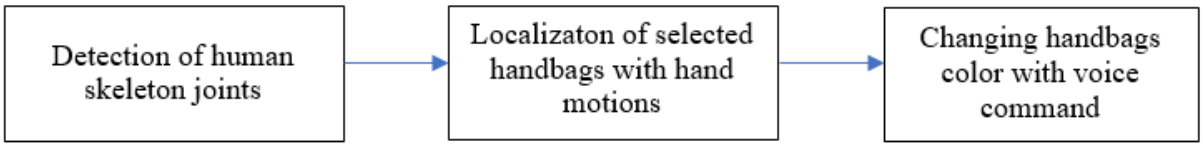



Figure 1 Design Steps of the Study

### 2.1 Voice Command Recognition with Correlation Method

In order to recognize user's voice command, correlation method is used. Firstly reference voice signals are collected in a database. When user interact with system by voice command, this voice command are correlated with all reference voice signals. According to correlation results; highest peak level obtains for similar signals.

Correlation method is used to measure the similarity of two signals. For *n* dimensional signals *x* and *y* the correlation equation is shown in Equation 1 [14].

$$r_{xy}(n) = \sum_{k=0}^{2(N-1)} x(k).y(k-n) \tag{1}$$

Equation 1 shows that by making the shift operation elements of the data, the comparison process is implemented. Similar signals at the highest rate will give the highest peak (Figure 5). If the $r_{xy}(n)$ >threshold then "Recognition procedure" or else "Recognition failure".

### 2.2 System Hardware: Microsoft Kinect Sensor

Various devices have been developed to detect human motions. Previously the implanted sensors were used. Recently new sensors which don't need to implant have been developed, such as Kinect.

Kinect Sensor has been developed for Xbox game console by Microsoft (Figure 2). Then it has been adapted to computer software world. Kinect can connect to computer with Universal Serial Bus (USB) and also requires 5V power from an external source [15]. The vision system of the Kinect sensor is composed of two cameras; RGB (Red-Green-Blue) camera and IR camera. And it contains an IR projector which is responsible of shooting infrared rays toward the environment. The distortion degree of each ray projected against the scene is used to estimate a depth map of the environments [16].

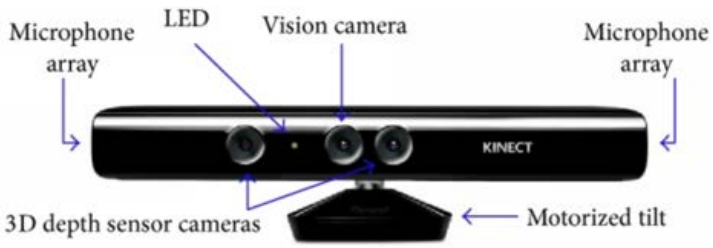



Figure 2 Microsoft Kinect Sensor [17]

Kinect sensor detects 6 people owing to the camera emitting infrared rays. It can follow 2 of them active movements. As shown in Figure 3 Kinect can find 20 skeleton joints of human body. Various movements of human are defined in Kinect software. With giving motion sensing command to Kinect, IR camera emits infrared rays to the possible human's hand, arm and foot points. If human is detected at these points, CMOS sensors open and movement begins to be defined. If this movement is defined in Kinect Sensor, signal is sent to the computer else Kinect is put on hold [15].



Figure 3 Skeleton Joints Found by Kinect Sensor

## 3. Experiments

### 3.1 System Interaction by Gestures

In proposed virtual mirror system, users can choose a handbag on a monitor with hand motions (Figure 4). In this way he /she can see on the monitor as if he/she wears this handbag. The Microsoft Kinect Sensor follows his/her left hand and when the user brings his/her left hand to the relevant bag picture,

this handbag picture is inserted on the right hand joints by the system (Figure 5). To placement the handbag on the human skeleton body it was also benefited Kinect Sensor Skeleton Map.



Figure 4 Proposed Virtual Mirror System

### 3.2 System Interaction by Voices

If the user wants to see different colors of the chosen bag, it will be sufficient to tell the color that he / she wants to see. In this instance, voice command recognition methods are activated in the system.

Owing to the reference voices, the system can compare between the voices. Before starting the application, the voices were collected in a database. Voices were obtained two second long, .wav format in a noiseless environment. Then the voices were divided into the audio component in Matlab platform.



Figure 5 Selecting a Handbag using Hand Motions

In order to compare the input voice, with the reference voices; both of them must be same or similar qualification. Thus input voices are processed in reference audio format. Firstly, the length of the voice which will be recorded is defined as two second using a timer. Then the frequency of voice is chosen as 22050 Hz which is reference voices frequency. The obtained record is stored in database in .wav format.

Input voice is compared with the all reference voices in the database and the nearest vioice is selected. Then the color of handbag is changed with target color which is requested by the user. Figure 6 shows

that the result of correlation method. In the figure, the correlation method is applied for user's 'blue' voice command. Thus user's 'blue' command is correlated with all reference voices. The reference voices are blue, red and green for this sample. As shown in graphs the correlation between blue reference voice and user's blue voices give the highest peak. X axis of the graphics shows that length of the correlation vector, y axis of the graphics show that results of correlation.

For changing color of the virtual handbags RGB color space is converted to HSV color space, because it is required not to lose shadows and saturation for more reality. So, only Hue (H) value is changed. After that color space is converted to RGB space again.

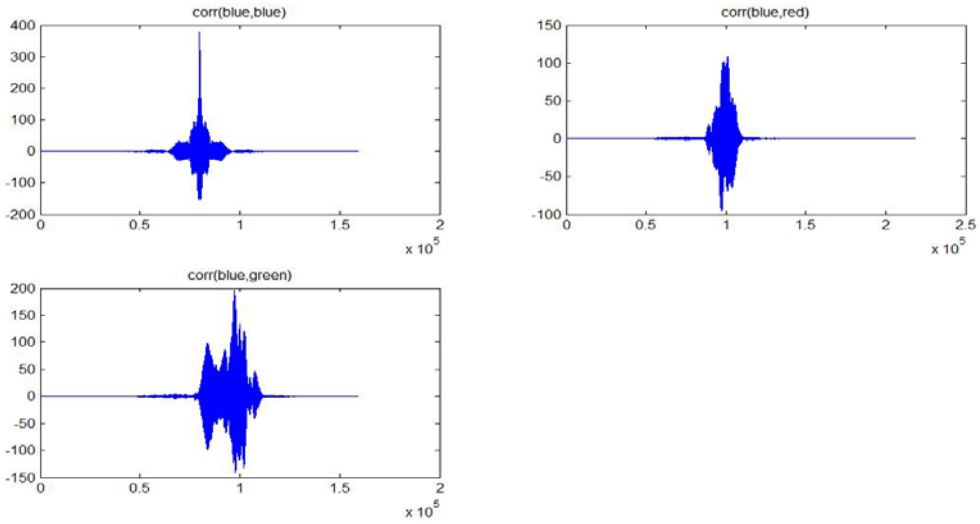



Figure 6 Results of Correlation Method

## 4. Conclusion

This study presents a NUI application which is developed for shopping. Kinect Sensor is utilized to ensure that the NUI. Users can interact with the system using voices and hand motions. To follow hand motions it has been benefitted from Kinect Skeleton Map; and to interact with the system by speech we used correlation method. Also user can see different poses of the bags by turning his / her arm. In this case system calculates subtraction of x axis coordinate of elbow joints and x axis coordinate of wrist joints. If the result is positive, the front side of the bag is displayed; if the result is negative, the back side of the bag is shown; if the result is zero, side view of the bag is shown.

While the system usually runs correctly, sometimes some inaccuracies such as wrong determination of vioce and synchronization disorder between the bag and the body can be (Figure 7). To avoid this, more posed-handbag images and reference voices can be uploaded to the database. In the future 3D system can be developed for handbag trying.



Figure 7 Example of Synchronization Failure

**References**

[1] Weiyuan Liu, “Natural user interface- next mainstream product user interface,” in 2010 IEEE 11th International Conference on Computer-Aided Industrial Design & Conceptual Design 1, 2010, pp. 203–205.

[2] K.-L. Kramer, User Experience in the Age of Sustainability. Science Direct, 2012.

[3] Y. El Miedany, Rheumatology Teaching. Springer.

[4] M. A. M. Echeverría, P. C. Santana-Mancilla, H. F. Q. Carrillo, and E. A. F. Enciso, “Natural User Interfaces to Teach Math on Higher Education,” Procedia - Soc. Behav. Sci., vol. 106, pp. 1883–1889, Dec. 2013.

[5] F. Cassola, L. Morgado, F. de Carvalho, H. Paredes, B. Fonseca, and P. Martins, “Online-Gym: A 3D Virtual Gymnasium Using Kinect Interaction,” Procedia Technol., vol. 13, pp. 130–138, 2014.

[6] H. Cho and N. Schwarz, “I like those glasses on you, but not in the mirror: Fluency, preference, and virtual mirrors,” J. Consum. Psychol., vol. 20, no. 4, pp. 471–475, Oct. 2010.

[7] P. Eisert, P. Fechteler, and J. Rurainsky, “3-D Tracking of shoes for Virtual Mirror applications,” in 2008 IEEE Conference on Computer Vision and Pattern Recognition, 2008, pp. 1–6.

[8] T. Blum, V. Kleeberger, C. Bichlmeier, and N. Navab, “mirracle: An augmented reality magic mirror system for anatomy education,” in 2012 IEEE Virtual Reality (VR), 2012, pp. 115–116.

[9] S. Casas, C. Portalés, L. Vera, and J. V. Riera, “Virtual and Augmented Reality Mirrors for Mental Health Treatment,” pp. 95–117.

[10] R. Nikam, C. Thete, S. Sonawane, S. Gadhave, and D. S. Thosar, “Virtual Shopping Using Kinect,” in National Level Conference On "Advanced Computing and Data Processing.

[11] J. Ishikiriyama and K. Suzuki, “An interactive virtual mirror to support makeup for visually impaired persons,” in 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2017, pp. 1393–1398.

[12] G. Yolcu, S. Kazan, and C. Oz, “Real Time Virtual Mirror Using Kinect,” , Balkan Journal vol. 2, no. 2, 2014.

[13] G. Yolcu, C. Oz, and T. Tasci, “Developing and Establishing a Painting Program Controlled by Hand Motions Using Kinect,” 2Nd Int. Symp. Innov. Technol. Eng. Sci., pp. 1–7, 2014.

[14] M. R. Alam, "Isolated Speech Recognition System Based on Cross-Correlation Technique", ResearchGate, 09 September 2016.

[15] A. Suzen, “Home automation for disabilities using kinect technology,” Suleyman Demirel, 2012.

[16] P. Cottone, G. Maida, and M. Morana, “User Activity Recognition via Kinect in an Ambient Intelligence Scenario,” IERI Procedia, vol. 7, pp. 49–54, 2014.

[17] D. Tuvshinjargal, B. Dorj, and D. J. Lee, “Hybrid Motion Planning Method for Autonomous Robots Using Kinect Based Sensor Fusion and Virtual Plane Approach in Dynamic Environments,” J. Sensors, vol. 2015, pp. 1–13, 2015.